Towards a novel biometric facial input for emotion recognition and assistive technology for virtual reality

J T McGhee[1], M Hamedi[2], M Fatoorechi[3], D Roggen[4], A Cleal[5], R Prance[6], C Nduka[7]

[1]Department of Medicine, Imperial College London,
South Kensington Campus, London, UNITED KINGDOM

[2,5,7]Emteq Ltd,
Sussex Innovation Centre, Science Park Square, Brighton, UNITED KINGDOM

[3,4,6]Department of Engineering and Design, Sussex University,
Falmer, Brighton, UNITED KINGDOM

james.thomas.mcghee@gmail.com, *Hamedi.mahyar@gmail.com*

ABSTRACT

Preliminary work using facial EMG to identify facial expressions is reported in this paper. Ten subjects performed 14 different facial expressions following an agreed protocol. Facial EMG signals, measured from surface electrodes, were processed and analysed using a machine learning algorithm. Our system is able to differentiate facial expressions for assistive input to a high degree of accuracy (99.25%) and posed emotional responses with 100% accuracy. We conclude facial EMG technology has the potential for both assistive input and emotion detection and could replace conventional assistive input devices or video based techniques for use with VR technologies.

1. INTRODUCTION

Virtual reality (VR) has been called the 4th digital communications platform after personal computers, the internet and mobile phones. However, humans use facial expression to convey large amounts of information which infer their underlying affective states. These expressions have developed far more universally than language and six universal expressions; anger, disgust, fear, sadness, surprise and happiness have been described (Ekman et al, 1990). While affective states are suggested by physiological parameters (Picard, 2000) there is no substitute for spontaneous facial expression for research, product development and usability studies (Vandal et al, 2016).

Virtual reality headsets cover the area of the face traditional facial recognition software uses to detect expressions. Facial gestures can be used as an input or measured to assess real time emotional responses to stimuli. Facial gestures could be of particular benefit to those with limb disabilities who typically rely on current devices which are often based on hand tracking technology. Real time emotional response measurement has wide applications across media, research and commercial applications. The Facial Action Coding System requires video footage of the area obscured by a VR headset. By measuring facial EMG universal facial expressions can be identified and therefore affective states extrapolated.

Muscle computer interaction (MuCI) is a method of increasing engagement opportunities for people with disabilities through the use of alternate inputs or outputs. Traditionally based on muscle activation in the upper limb MuCI has been shown to be viable even for complex applications such as wheelchair navigation (Firoozabadi et al, 2008). Current research has shown facial EMG to be a viable method of MuCI (Hamedi et al, 2011, 2013, 2014, 2015). Even high spinal cord lesions will leave the facial nerve, the nerve responsible for facial expressions, intact. Therefore, demonstrating the potential uses for this technology as a method of human computer interaction.

Facial EMG analysis have been explored and refined using techniques including neural networks and machine learning. The high computational demands processing the data has the result that higher accuracy is achievable at the expense of performance speed (Hamedi et al, 2014). For these systems increasing the number of expressions which can be recognised increases the complexity of the task. Previous groups have considered up to 10 facial expressions, increasing the total to 14 including differentiating between left and right sided expressions could increase the applications of EMG as an input device (Hamedi et al, 2015). This development could encourage further research into other applications including medical rehabilitation of facial palsy.

Virtual reality provides an emotionally engaging and fully immersive visual environment for users. The nature of VR headsets obscuring the face disadvantages communication and social applications for VR. In this respect EMG is promising as sensors can be placed within the foam padding of existing VR devices potentially broadening the scope of VR interaction. Duel use of these sensors for both measuring emotional response and expression recognition for gesture control could widen the audience for VR and make it more accessible to those with disabilities.

2. METHODS AND MATERIALS

A range of facial expressions were considered, taking into account the action units (AU) of the facial action coding system (FACS) and how this could related to emotions and gesture input (Ekman & Frisen, 1978). For emotion recognition a subset of 6 expressions representing the emotional states of neutral, surprise (AU1, AU2), disgust (AU9, AU10), anger (AU4, AU5) and happiness were chosen. Happiness was divided into two states, a closed mouth (AU12) and a wider smile (AU6, AU12) representing a stronger emotional response. For assistive input a more comprehensive range of 14 expressions representing those that the majority of people could create voluntarily were selected. Figure 1 shows the final expressions selected for analysis.

2.1 *Data Collection Protocol*

The custom goggles with BIOMETRICS ltd. silver chloride (AgCl) electrodes (SEN3001) were positioned on subjects with freshly washed and dried faces to ensure good contact and standardised positioning. Five bipolar channels were used to gather EMG data. Maximum input of ±3V and resolution of 0.732mV. Power supply and current per channel was <4.6V and <20mA respectively. The wireless EMG system (BIOMETRICS ltd. DataLOG MWX8) can be worn on a belt and streams data via Bluetooth to a PC.

Expressions were recorded and performed in the same order using a standardised presentation on Microsoft PowerPoint. Subjects were shown a captioned photograph of the expression before and during each expression capture period. Expression capture was standardised with an audio recording signalling 10 seconds of rest a warning and 2 seconds of expression repeated 10 times for each expression. Subjects were instructed to hold a "neutral expression" between expression captures and asked not to blink in the capture period. Recordings took place in a single sitting without adjustment of the custom platform or positioning of the surface electrodes.

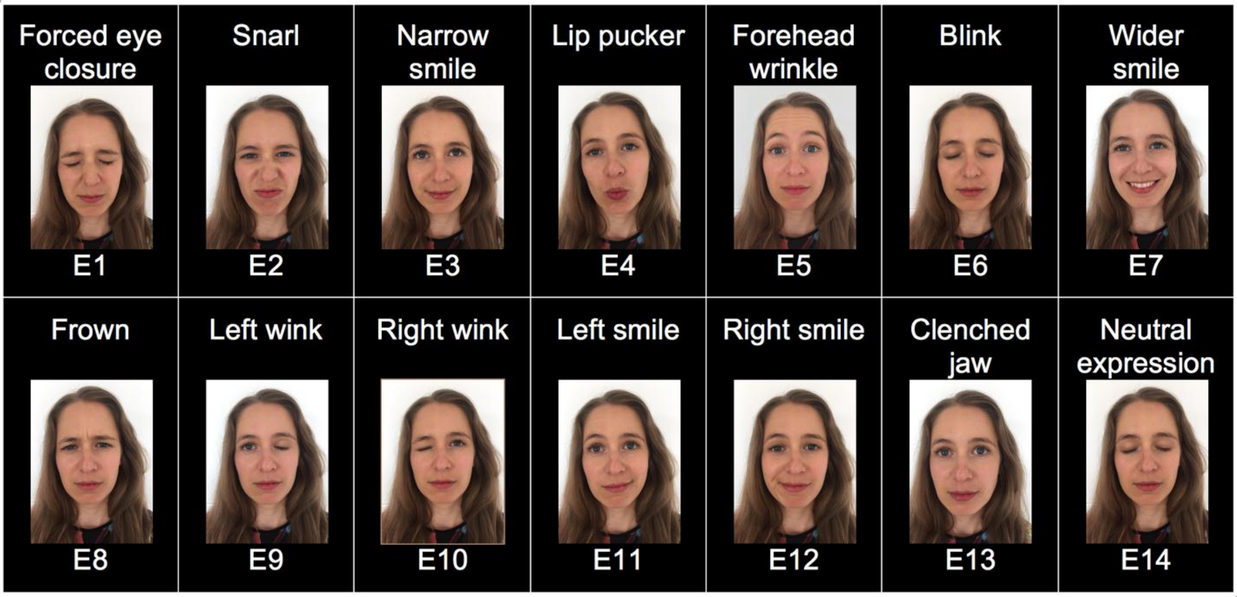

Figure 1. *Final expressions selected for analysis and coding. Snarl, Forehead wrinkle, Frown, Narrow smile and Wide smile have been selected to represent Disgust, Surprise, Anger and happiness respectively.*

2.2 *EMG Pre-processing*

Raw EMG data in all channels was passed through a 6th order Butterworth band-pass filter in the range of 30-450Hz to envelop the most significant spectrum of signals. Facial EMG signals vary if the set up, including positioning, of the electrodes has been altered. The amplitude of the signals is therefore best measured in relative rather than absolute values. The filtered EMGs were therefore separately normalised using the maximum and minimum values of the EMG in that channel for each expression.

2.2 *EMG Segmentation and feature extraction*

To recognise the EMG patterns of different facial expressions, the most significant, discriminative features of the EMGs should be estimated. In pattern recognition the accuracy and success of the final performance is highly dependent on the quality of the signal features. Normalised EMG signals were segmented into non-overlapped windows of 256 ms length. Mean absolute values (*MAV*) were computed using equation 1 and extracted from each signal segment, where *N* is the length of the segment, *n* is the current segment and x_k is the current point.

$$MAV_n = \frac{1}{N}\sum_{k=1}^{N}|x_k| \tag{1}$$

2.3 *Feature Classification*

In order to recognise the facial expressions the extracted features need to be classified into distinct classes through a formal technique which provides a high level of accuracy at a low computational cost. This is desirable, particularly for assistive input technology. Previous studies have shown that least square support vector machine (LS-SVM) is a robust algorithm for classification of facial EMG patterns with a very short training time (Hamedi et al, 2015, 2016). This study used LS-SVM constructed by radial basis function kernel where the regularisation and smoothing parameters were set to 10 and 0.6 respectively. Multi-class LS-SVM was trained by considering One-Vs-One encoding method. The 10-fold cross validation strategy is used for classification evaluation.

3. RESULTS AND DISCUSSION

Of the 10 subjects 6 recordings were appropriate for full assessment of 14 expressions due to inability of some of the subjects to perform additional expressions. All 10 subjects were appropriate for assessment of posed emotional responses.

3.1 *Assistive Input*

Recognition accuracy varied between 98.03-100% for our system. More expressions were classified in this paper compared to previous publications, system performance shows our custom platform provides a stable framework to perfectly detect different facial expressions. Table 1 provides the confusion matrix of the system performance averaged over all the subjects. Forehead wrinkle, blink and left and right smiles were the most recognisable facial expressions within all subjects. Lip pucker was the most misclassified expression, mainly for left wink. On investigation this was an issue in one subject and was most likely due to electrode misplacement rather than similar signalling source.

Table 1. *Confusion matrix of the system performance averaged over all subjects.*

Actual Class Label

Predicted Class Label

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14
E1	99.93	0	0	0	0	0	0.04	0	0	0	0	0	0	0
E2	0	98.81	0.11	0	0	0	0	0	0	0	0	0	1.17	0.08
E3	0	0.18	98.8	0	0	0	0	1.6	0	0	0	0	0.04	0
E4	0	0	0	96.88	0	0	0	0	1.47	0	0	0	0	0
E5	0	0	0	0	100	0	0	0	0	0	0	0	0	0
E6	0	0	0	0	0	100	0	0	0	0.58	0	0	0	0
E7	0.07	0	0	0	0	0	99.96	0	0	0	0	0	0	0
E8	0	0	1.09	0	0	0	0	98.4	0	0	0	0	0	0
E9	0	0	0	3.13	0	0	0	0	98.53	0	0	0	0	0
E10	0	0	0	0	0	0	0	0	0	99.42	0	0	0	0
E11	0	0	0	0	0	0	0	0	0	0	100	0	0	0
E12	0	0	0	0	0	0	0	0	0	0	0	100	0	0
E13	0	0.91	0	0	0	0	0	0	0	0	0	0	98.8	0
E14	0	0.11	0	0	0	0	0	0	0	0	0	0	0	99.93

3.2 Emotion Recognition

This is the first study investigating emotion detection for VR using Facial EMG. The placement of the electrodes round the custom platform mirrors current VR headsets while maximising muscle coverage. The system detects activity of the corrugator, frontalis, orbicularis oculi, nasalis, levator labii and zygomaticus muscles at different amplitudes for different expressions. As this section of our results focuses on emotion recognition and universal expressions all 10 subjects are included. Our system achieved 100% recognition accuracy for all 6 of our posed emotional expressions.

4. CONCLUSIONS

This paper is the first to report preliminary work using facial EMG signals to detect and identify facial expressions as a measure of an instant emotional response. Previous research has used video to measure facial expression and infer affective state. However, current methods are ineffective due to placement of VR headsets. We have demonstrated recognition of positive and negative expressions with an extremely high degree of accuracy with technology we feel could be integrated into current VR headsets.

Hamedi et al. (2011, 2013, 2014, 2015) have assessed different expressions but this is the first paper to recognise 14 different expressions, the largest number of facial expressions to be recognised by EMG as far as we are aware. Our system was able to differentiate between the expressions with a very high degree of accuracy, encouraging further research into this area. We plan to create an open hardware and software platform to enable other researchers and device makers to investigate opportunities for facial EMG in VR.

5. REFERENCES

Ekman, P., & Frisen, W. (1978). *The facial action coding system: a technique for the measurement of facial movement.* San Fransisco: Consulting Psycologists Press.

Ekman, P., Rolls, E., Perrett, D., & Ellis, H. (1990). Facial Expressions of EmotionL An Old Controversy and New Findings. . *JSTOR Health & General Sciences , 335* (1273), 63-69.

Firoozabadi, S., Oskoei, M., & Hu, H. (2008). A human-computer interface based on forehead multi-channel bio-signals to control a virtual wheelchair. *ICBME*, (pp. 272-277).

Hamedi, M., Salleh, S., & Noor, A. (2015). Facial Neuromuscular Signal Classification by means of Least Square Support Vector Machine for MuCI. *Applied Soft Computing , 30,* 83-93.

Hamedi, M., Salleh, S., Noor, A., Harris, A., & Majid, N. (2014). Multiclass Least-Square Support Vector Machine for Myoelectric-Based Facial Gesture Recognition. *The 15th International Conference on Biomedical Engineering* (pp. 180-183). Springer.

Hamedi, M., Salleh, Sh-H, Astaraki, M. and Noor, A.M. (2013), EMG-based facial gesture recognition through versatile elliptic basis function neural network, *Biomedical engineering online*, *12, 1*, 73-94.

Hamedi, M., Salleh, S.H., Ting, C.M., Astaraki, M. and Noor, A., Robust Facial Expression Recognition for MuCI: A Comprehensive Neuromuscular Signal Analysis. *IEEE Transactions on affective computing*, 99 (2016).

Hamedi, M., Salleh, S. H., Tan, T. S., Ismail, K., Ali, J., Dee-Uam, C., and Yupapin, P. P. (2011), Human facial neural activities and gesture recognition for machine-interfacing applications. *International journal of nanomedicine*, *6*, 3461-3472.

Picard, R. (2000). *Affective computing.* MIT Press.

Vandal, T., McDuff, D., & El Kaliouby, R. (n.d.). *Affectiva.* http://www.affectiva.com/wp-content/uploads/2016/02/Event-Detection-Ultra-Large-scale-Clustering-of-Facial-Expressions.pdf