Neural correlates of affective context in facial expression analysis: A simultaneous EEG–fNIRS study

3 authors:

Yanjia Sun
Automatic Data Processing
9 PUBLICATIONS   **69** CITATIONS

SEE PROFILE

Hasan Ayaz
Drexel University
258 PUBLICATIONS   **3,951** CITATIONS

SEE PROFILE

Ali Akansu
New Jersey Institute of Technology
252 PUBLICATIONS   **4,342** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   En Route Decision Support Automation Research View project

Project   NeuroErgonomics Methods and Applications View project

Neural Correlates of Affective Context in Facial Expression Analysis: A Simultaneous EEG-fNIRS Study

Yanjia Sun, Hasan Ayaz*, and Ali N. Akansu

Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, New Jersey USA
School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, Pennsylvania USA*
E-mail: yanjia.sun@njit.edu, hasan.ayaz@drexel.edu, akansu@njit.edu

Abstract— **We present a framework to measure the correlation between spontaneous human facial affective expressions and relevant brain activity. The affective states were registered from the video capture of facial expression and related neural activity was measured using wearable and portable neuroimaging systems: functional near infrared spectroscopy (fNIRS), electroencephalography (EEG) to asses both hemodynamic and electrophysiological responses. The methodology involves the simultaneous detection and comparison of various affective expressions by multi-modalities and classification of spatio-temporal data with neural signature traits. The experimental results show strong correlation between the spontaneous facial affective expressions and the affective states related brain activity. We propose a multimodal approach to jointly evaluate fNIRS signals and EEG signals for affective state detection. Results indicate that proposed method with fNIRS+EEG improves performance over fNIRS or EEG only approaches. These findings encourage further studies of the joint utilization of video and brain signals for face perception and brain-computer interface (BCI) applications.**

Index Terms— functional near infrared spectroscopy, electroencephalography, facial emotion recognition

I. INTRODUCTION

Facial expression has been regarded as one of the most efficient and instinctive ways to reflect human feelings and intentions in the form of nonverbal communication. Moreover, the psychology literature [1] empirically demonstrates that people can infer others' preferences through spontaneous facial expressions. The accurate interpretation of facial expressions has applications spanning from patient monitoring to marketing analytics.

In this study, we explored the verification of the facial expressions by simultaneous analysis of affective states through relevant brain activity. Six basic emotions defined by Ekman [2] are the building blocks although they cover a rather small part of our daily emotional displays [3]. Affect [4] refers to the emotion elicited by stimuli and objectively reflects a person's mental state. It has been generally classified in the field of psychology by using positive and negative dimensions. This is an important finding since affect is commonly used in natural communication among people.

High recognition rates of facial emotions have been achieved in existing systems based on the benchmarked databases containing posed facial emotions. However, they were generated by asking subjects to perform a series of exaggerated expressions that are quite different from spontaneous ones [5]. An automatic facial emotion recognition system recently proposed in [6] is capable of classifying the spontaneous facial emotions. We use that system in this study. The block diagram of the system is shown in Fig. 1.

Parts of the human brain have been demonstrated to be related with the expression of emotions [7]. The development of brain-computer interface (BCI) gives a new way to enhance the interactions between computers and humans. Functional near-infrared spectroscopy (fNIRS) [8] and electroencephalography (EEG) [9] successfully detect human brain activity although they measure different physiological responses to the changes in the mental state. Thus, a brain imaging technique using both of them simultaneously is utilized to improve the analysis of inner affective states.

The contributions of the paper are as follows. 1) To the best of our knowledge, this is the first attempt in detecting human spontaneous affective states simultaneously using fNIRS and EEG signals for neural measures, and also facial expressions captured by camera. 2) The spontaneous facial affective expressions are demonstrated to be aligned with the affective states coded by brain activities. Consistent with mobile brain/body imaging (MoBI) approaches [10], we have shown that the proposed technique outperforms methods using only a subset of these signal types for the task.

II. RELATED WORK

The face is one of the information sources used daily to infer people's mental states. Lucey et al. [11] show the automatic pain detection from the captured expressions on the face of the participants who were suffering from shoulder pain. The relationship between the facial responses and ad effectiveness is studied in [12] where authors demonstrated that ad liking can be predicted accurately from webcam facial responses.
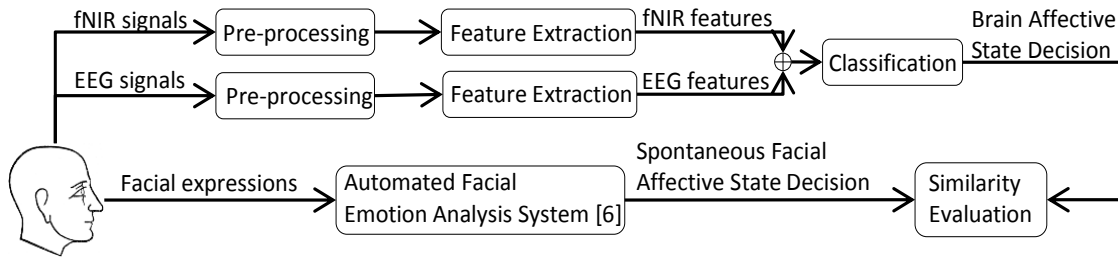
Figure 1. Block diagram of a system to evaluate spontaneous facial expression using the brain activity.

Physiological mechanisms under the emotional state have been investigated. fNIRS and EEG techniques are widely utilized due to their non-invasive and mobile natures [13], [14]. [15] built the database of brain signals and facial expressions using fNIRS, EEG, and video. Therefore, they have been a primary option for emotive state recognition. However, whether the inner affective states translated by human brain activity are consistent with those expressed on the face has not been explored as of yet. In this study, we investigated the spontaneous facial affective expressions through detecting the brain activity using both fNIRS and EEG signals.

III. EXPERIMNET DESIGN

A. Experimental Protocol

Twelve male participants (age: mean=27.58, SD=4.81) are selected as perceivers to participate in this experiment. The perceiver was assigned to watch twenty emotionally stimulating images. An image was shown for six seconds such that the perceiver can recognize the type of affect. After observing an image, a perceiver was asked to answer two simple questions (e.g. Were you able to observe the photo carefully? What were you seeing?) in order to verify that the image content was perceived. The perceiver was required to wear both fNIRS and EEG when watching the images during the experiment. Moreover, his facial reaction to the stimuli was recorded as video for further emotion analysis. The experimental environment is displayed in Fig. 2 (a). The study was approved by the Institutional Review Board of the New Jersey Institute of Technology.

B. Brain Data Acquisition

The neuroimaging systems used in this study consists of two commercial products: a wireless fNIRS system and a portable EEG system (see Fig. 2). A compact and battery-operated wireless fNIRS system (Model 1100W) developed by Drexel University (Philadelphia, PA) and manufactured by fNIR Devices, LLC (www.fnirdevices.com) was used to monitor the prefrontal cortex of the perceiver. The system is composed of three modules: a sensor pad that holds light sources and detectors to enable a fast placement of four optodes, a control box for hardware, and a computer that runs COBI Studio [16] receives the data wirelessly. Additional information about the device was reported in [13]. Emotiv EPOC headset was used to acquire EEG signals with 128 Hz sampling rate. It transmits wirelessly to a Windows PC

(www.emotiv.com). It has 14 electrodes located over 10-20 international system positions AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 using 2 reference electrodes.

(a)
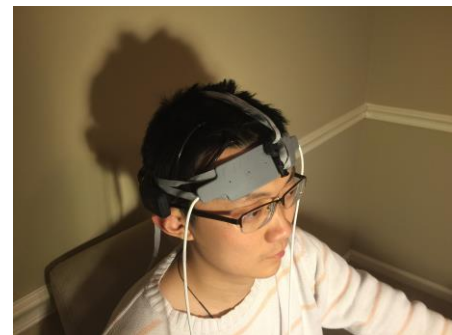
(b)

Figure 2. (a) Experimental environment and (b) Neuroimaging apparatus.

C. Automatic Facial Emotion Recognition System

The system processes video clips and utilizes Regional Hidden Markov Model (RHMM) as classifier to train the states of three face regions: the eyebrows, eyes and mouth. 41 facial feature points are identified on a human face, as displayed in Fig. 3. They are comprised of 10 salient points on the eyebrows region, 12 points on the eyelids, 8 points on the mouth, 10 points on the corners of the lips, and one anchor feature point on the nose. The 2D coordinates of facial feature points in various face regions are extracted to form corresponding observation sequences for classification. As reported in [6], the system not only outperforms the state-of-

the-art for the posed expressions, but it also provides satisfactory performance on the spontaneous facial emotions. The system has been used to read CEO's facial expressions to forecast firm performance in the stock market [17].

Figure 3. Facial feature points on a face image.

The system brings benefits to this study for three main reasons. First, the objective and measurable response to person's emotions by the system will be perceived as more natural, persuasive, and trustworthy. This allows us to plausibly analyze the measured data. Second, the system is able to recognize the various affective states. The last but most important advantage is that the system automatically analyzes and measures the live facial video data in the manner that is completely intuitive and prompt to the different kinds of applications. It benefits the user to take actions based on this type of analysis.

IV.   MODEL AND ANALYSIS

A.   Data Pre-processing

Noise reduction in fNIRS signals was achieved by utilizing a low-pass filter with 0.1 Hz cutoff frequency [8]. Motion artifacts were eliminated prior to extracting the features from fNIRS signals by applying a fast Independent Component Analysis (ICA) [18]. We selected Independent Components through modeling the hemodynamic response using gamma function [19]. The perceiver was also instructed to minimize head movements during the viewing of images in order to avoid the artifacts from the muscular activity at maximum level.

The EEG signal was passed through a low-pass filter with 30 Hz cutoff frequency. The artifacts in the raw EEG signals were detected and removed using ICA in EEGLAB which is an open-source toolbox for analysis of single-trial EEG dynamics [20]. The average of the 5-second baseline brain signal before each trail was subtracted from the brain response data for the baseline adjustment.

Inference of perceiver's facial affective expression was performed by a facial emotion recognition system. The system recognized emotion type of a facial video clip v by calculating the largest probability among Anger, Disgust, Fear, Happiness, Neutral, Sadness, and Surprise expressed as

$$P_v^{Emotion} = \max(P_i), \quad v = 1,...,N_{video}, \ i = 1,..,7 \qquad (1)$$

where N_{video} is the total number of videos, each of which contains the perceiver's response to test images. P_i is the probability of the ith emotion type, that is obtained by summing the overall probabilities p_{ij} of this emotion type in the jth frame of a video clip. p_{ij} is normalized to one.

We coded Happiness as the positive affect and Anger, Disgust, Fear, and Sadness were combined as the negative affect. Since Surprise may be revealed as a result of positive or negative affect, it was not used in this experiment. The ratings of both the positive affect P_{pos} and the negative affect P_{neg} for a video clip were separately calculated by the system as

$$P_v^{Affect} = \max(P_{pos}, P_{neg}), \ v = 1,...,N_{video},$$

$$\text{subject to} \begin{cases} P_{pos} = \dfrac{P_{Happiness}}{P_{Happiness} + \max(P_{Anger}, P_{Disgust}, P_{Fear}, P_{Sadness})} \\[4mm] P_{neg} = \dfrac{\max(P_{Anger}, P_{Disgust}, P_{Fear}, P_{Sadness})}{P_{Happiness} + \max(P_{Anger}, P_{Disgust}, P_{Fear}, P_{Sadness})} \end{cases} \qquad (2)$$

where the larger affect rating P_v^{Affect} of the two values is selected as the recognized affective state of the corresponding video clip.

B.   Feature Extraction

The recorded data of raw fNIRS light intensity was converted to the relative changes in hemodynamic responses in terms of oxy-hemoglobin (Hbo) and deoxy-hemoglobin (Hbr) using a modified Beer-Lambert Law [16]. Total blood volume (Hbt), and the difference of Hbo and Hbr (Oxy) were also calculated for each optode. As features, we calculated the mean, median, standard deviation, maximum, minimum, and the range of maximum and minimum of four hemodynamic response signals, giving 4×4×6=96 fNIRS features for each trial.

The spectral power of EEG signals in different bands has been used for emotion analysis [21]. The logarithms of the power spectral density (PSD) from theta (4Hz < f < 8Hz), slow alpha (8Hz < f < 10Hz), alpha (8Hz < f < 12Hz), and beta (12Hz < f < 30Hz) bands were extracted from all 14 electrodes as features. In addition, the difference between the spectral power of all the possible symmetrical pairs on the right and left hemisphere were extracted to measure the possible asymmetry in the brain activity due to the valance of emotional stimuli [22]. The asymmetry features were extracted from four symmetric pairs over four bands, AF3-AF4, F7-F8, F3-F4, and FC5-FC6. The total number of EEG features of a trial for 14 electrodes were 14×4+4×4=72.

C.   Affective State Detection

We conducted twenty trials for each perceiver. The stimuli were selected from the Nencki Affective Picture System (NAPS) [23]. It is a standardized, wide-range, high-quality picture database that is used for brain imaging realistic studies. The pictures are labeled according to the affective states. Polynomial Support Vector Machine (SVM) was

employed for classification. fNIRS features and EEG features were concatenated to form a larger features vector before feeding them into the model. For comparison, we also applied the univariate modality of either fNIRS features or EEG features for the affective state detection. We applied leave-one-out approach for training and testing. Namely, the features extracted in nineteen trials over all perceivers were used for training, and the remaining one was used for testing purpose. The experimental outcome was validated through twenty-time iterations.

The performance results are shown in Fig. 4 and Fig. 5. The proposed multi-modal method performs about 20% better than the others for the same stimuli. Although both EEG and fNIRS are non-invasive, safe and portable neuroimaging techniques for assessing the brain function, they measure different and actually complimentary aspects of the brain activity. The performance of recognizing affective states using the combination of fNIRS features and EEG features (in red bar) almost dominates over all trial. The overall performance comparison can be found in Fig. 5.
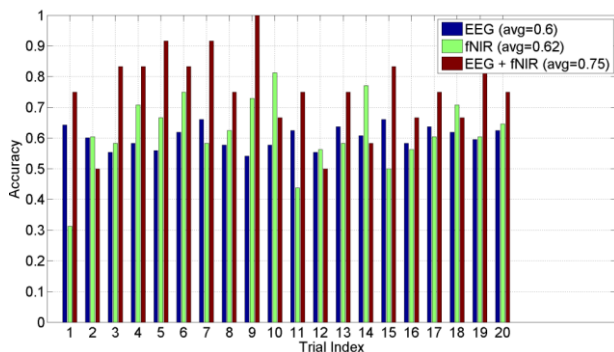

Figure 4. Performance comparisons of the proposed multimodal approach (EEG+fNIRS) and others overall image-content trials.
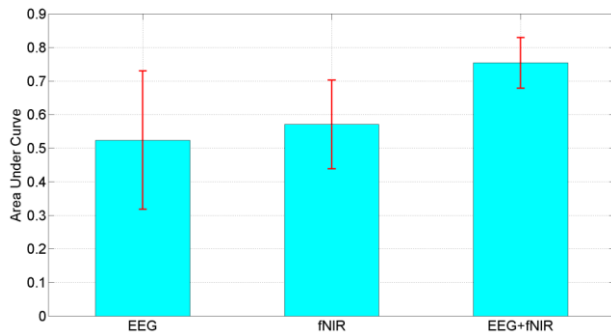

Figure 5. Average performances of EEG only, fNIRS only, and the propopsed fNIRS+EEG methods are measured in classifications area under a ROC curve displayed in error bars.

D. Similarity of the Spontaneous Facial Affective States and the Affective States Coded by the Brain Activity

Affect recognition accuracy of the system was measured by comparing its results with the actual labels of test images.

The accuracy reached 0.74. The averages of perceivers' facial inferences per trial are given in Table I. T_i, $i = 1,...,20$ represents the ith trial.

TABLE I. AVERAGE OF PERCEIVERS' FACIAL AFFECTIVE EXPRESSION PER TRIAL

T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
0.92	0.75	0.67	0.75	0.83	0.75	0.67	0.83	0.83	0.83
T11	T12	T13	T14	T15	T16	T17	T18	T19	T20
0.67	0.75	0.67	0.5	0.75	0.75	0.92	0.75	0.67	0.58

Similarity scores were calculated by the percentages of spontaneous facial affect recognized by the system same as the inner affective states translated by perceiver's brain signals. For this purpose, we performed Wilcoxon Signed Rank Test where affective states were classified as positive and negative. Wilcoxon Signed Rank Test (W=65, p=0.04) confidently shows that the spontaneous facial affective states are consistent with those translated by human brain activity. That is, the spontaneous facial affective states can reflect true human affective response to stimuli.

To further check the reliability of our experimental results, we examine whether some perceivers were expressive which may bias the average of all perceivers' facial expressions in a trial. To do that, we explored the perceiver's neutral affective state over trials (Table II). One-sample t-test was used to assess whether the average of measured neutral states is close to the population mean (μ=0.3) (there are three types of affective states). The results of one-sample t-test (Mean=0.23, SD=0.11, t(11)=1.75, p>0.1) explains that all perceivers were not expressive during the experiment.

TABLE II. AVERAGE NEUTRAL AFFECTIVE STATE OF A PERCEIVER OVER ALL TRIALS

P1	P2	P3	P4	P5	P6
0.46	0.11	0.13	0.25	0.24	0.33
P7	P8	P9	P10	P11	P12
0.21	0.18	0.29	0.10	0.14	0.37

V. CONCLUSION

This paper demonstrates that human spontaneous facial expressions convey the affective states translated by the brain activity. The proposed method combining fNIRS signals and EEG signals is shown to outperform the methods using either fNIRS or EEG as they are known to contain complimentary information [24]. We also examined the perceiver's expressiveness. The results indicate that there is no perceiver expressing exaggerative expressions over all trials performed.

To the best of our knowledge, this is the first attempt in detecting the affective states using fNIRS, EEG, and video capture of facial expressions and demonstrates the utility of the multimodal approach. The experimental results validate the proposed method. In addition, the results explain the reliability of spontaneous facial expressions used in various applications.

REFERENCES

[1] M.S. North, A. Todorov, and D.N. Osherson, "Inferring the preferences of others from spontaneous, low-emotional facial expressions," *Journal of Experimental Social Psychology*, vol. 46, no.6, pp. 1109-1113, 2010.

[2] P. Ekman, "Universals and Cultural Differences in Facial Expressions of Emotion," Proc. Nebraska Symp. Motivation, 1971, pp. 207-283.

[3] Z. Zeng, M. Pantic , G. I. Roisman and T. S. Huang, "A survey of affect recognition methods: audio, visual, and spontaneous expressions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, no. 1, pp.39 -58, 2009.

[4] P.T. Trzepacz and R.W. Baker, *The psychiatric mental status examination*, Oxford University Press, 1993.

[5] M.F. Valstar, H. Gunes, and M. Pantic, "How to distinguish posed from spontaneous smiles using geometric Features," *Ninth ACM Int'l Conf. Multimodal Interfaces*, 2007, pp. 38-45.

[6] Y. Sun and A. N. Akansu, "Automatic Inference of Mental States from Spontaneous Facial Expressions," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2014, pp. 719-723.

[7] V. Zotev, R. Phillips, K. D. Young, W. C. Drevets, and J. Bodurka, "Prefrontal control of the amygdala during real-time fMRI neurofeedback training of emotion regulation," *PloS one*, vol. 8, no. 11, 2013.

[8] H. Ayaz, P.A. Shewokis, S. Bunce, K. Izzetoglu, B. Willems, and B. Onaral, "Optical brain monitoring for operator training and mental workload assessment," *NeuroImage*, vol. 59, no. 1, pp. 36–47, 2012.

[9] M. De Vos and S. Debener, "Mobile EEG: Towards brain activity monitoring during natural action and cognition," *International Journal of Psychophysiology*, vol. 91, no. 1, pp. 1-2, 2014.

[10] K. Gramann, J. T. Gwin, D. P. Ferris, K. Oie, T.-P. Jung, C.-T. Lin, L.-D. Liao, and S. Makeig, "Cognition in action: imaging brain/body dynamics in mobile humans," *Reviews in the Neurosciences*, vol. 22, no. 6, pp. 593-608, 2011.

[11] P. Lucey, J.F. Cohn, K.M. Prkachin, P.E. Solomon, and I. Matthews, "Painful data: the UNBC-McMaster shoulder pain expression archive database," *In: Proc. of IEEE Int. Conf. on Automatic Face and Gesture Recognition*, 2011, pp. 57–64.

[12] D. McDuff, R. Kaliouby, J. Cohn, and R. Picard, "Predicting Ad Liking and Purchase Intent: Large-scale Analysis of Facial Responses to Ads," *IEEE Transactions on Affective Computing*, vol.PP, no.99, 2014.

[13] H. Ayaz, B. Onaral, K. Izzetoglu, P. A. Shewokis, R. McKendrick, and R. Parasuraman, "Continuous monitoring of brain dynamics with functional near infrared spectroscopy as a tool for neuroergonomic research: Empirical examples and a technological development," *Frontiers in Human Neuroscience*, no. 7, pp. 1-13, 2013.

[14] Y. Liu, H. Ayaz, A. Curtin, B. Onaral, and P. A. Shewokis, "Towards a Hybrid P300-Based BCI Using Simultaneous fNIR and EEG," *In Foundations of Augmented Cognition, Springer Berlin Heidelberg*, vol. 8027, pp. 335-344, 2013.

[15] A. Savran, K. Ciftci, G. Chanel, J. C. Mota, L. H. Viet, B. Sankur, L. Akarun, A. Caplier, and M. Rombaut, "Emotion Detection in the Loop from Brain Signals and Facial Images," *Proc. of the eNTERFACE Workshop*, July, 2006.

[16] H. Ayaz, P. A. Shewokis, A. Curtin, M. Izzetoglu, K. Izzetoglu, and B. Onaral, "Using MazeSuite and Functional Near Infrared Spectroscopy to Study Learning in Spatial Navigation," *Journal of visualized experiments*, no. 56, 2011.

[17] Y. Sun, A. N. Akansu, J. E. Cicon, "The Power of Fear: Facial Emotion Analysis of CEOs to Forecast Firm Performance," *IEEE International Conference on Information Reuse and Integration*, 2014.

[18] A. Hyvärinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis," *IEEE Trans on Neural Networks*, vol. 10, no. 3, pp. 626-634, 1999.

[19] M. A. Lindquist, J. M. Loh, L. Y. Atlas, and T. D. Wager, "Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling," *Neuroimage*, vol. 45, no. 1, 2009.

[20] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9-21, 2004.

[21] S. Koelstra and I. Patras, "Fusion of facial expressions and EEG for implicit affective tagging," *Image and Vision Computing*, vol. 31, no. 2, pp. 164-174, 2013.

[22] S. K. Sutton, and R. J. Davidson, "Prefrontal brain asymmetry: A biological substrate of the behavioral approach and inhibition systems," *Psychological Science*, vol. 8, no. 3, pp. 204-210, 1997.

[23] A. Marchewka, L. Żurawski, K. Jednoróg, and A. Grabowska, "The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database," *Behavior research methods*, vol. 46, no. 2, pp. 596-610, 2014.

[24] S. Fazli, J. Mehnert, J. Steinbrink, G. Curio, A. Villringer, K. R. Müller, and B. Blankertz, "Enhanced performance by a hybrid NIRS-EEG brain computer interface," *Neuroimage*, vol. 59, no. 1, pp. 519-529, 2012.