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

Yanjia Sun[1], Hasan Ayaz[2], and Ali N. Akansu[1]

Department of Electrical and Computer Engineering,

New Jersey Institute of Technology, Newark, New Jersey USA[1]

School of Biomedical Engineering, Science and Health Systems, Drexel University, Philadelphia, Pennsylvania USA[2]

E-mail: yanjia.sun@njit.edu, hasan.ayaz@drexel.edu, akansu@njit.edu

NJIT New Jersey's Science & Technology University

Drexel UNIVERSITY

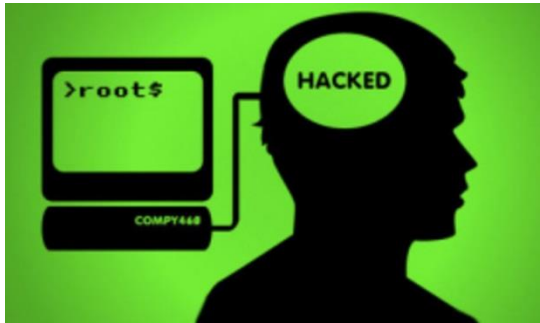
Entertainment

High Tech

Brain Power

Disability

BCI



Overview of Related Work

- **Univariate analysis of emotional response**
  - Peripheral physiological system
    - Automatic facial emotion recognition
  - Central nervous system
    - Emotion-based brain-activity detection

- **Multivariate analysis of emotional response**
  - Fusion of Electroencephalography (EEG), Functional Near Infrared Spectroscopy (fNIRS), skin conductance, and blood pressure
  - Fusion of EEG and facial emotion recognition system

Functional Near Infrared Spectroscopy (fNIRS)

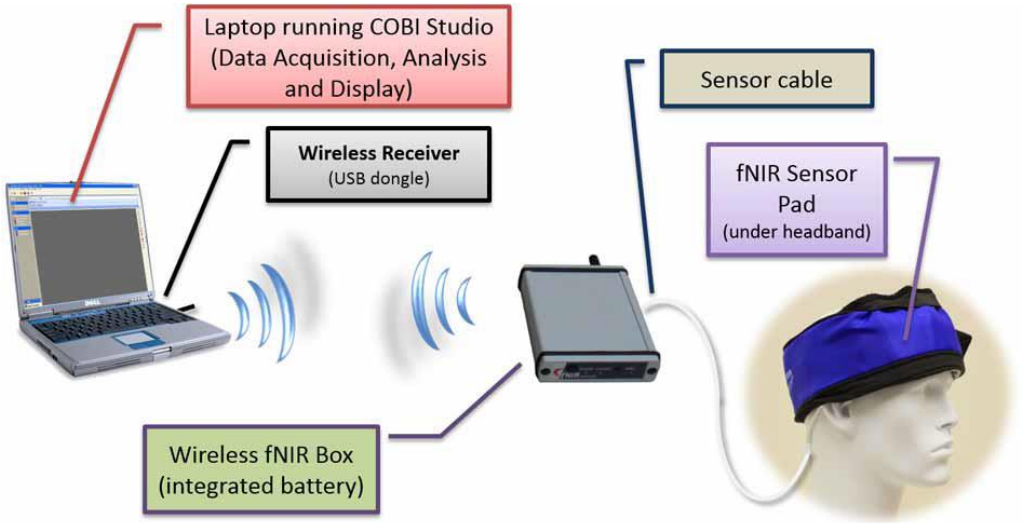
- fNIRS uses optical fibers placed on forehead to send light in the wavelengths of 730nm and 850nm into head.

- Photons that enter the forehead are
  - Scattering in the different layers of head
  - Absorbed by tissues (oxy-Hemoglobin and deoxy-Hemoglobin) in the near-infrared range.
  - Changes of concentration of these oxy-Hemoglobin and deoxy-Hemoglobin are reflected into optical density (OD) at two wavelengths (730nm and 850nm).

- Photodetectors capture the photons
  - Those are not absorbed
  - Those travel through a specific volume between the source and detectors

Wireless fNIRS Apparatus

- **Functional Near Infrared Spectroscopy (fNIRS)**
  - A sensor pad: 4 light sources and 10 detectors (4 optodes; 12 channels)
  - A control box
  - A computer that run data acquisition



[1] 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.

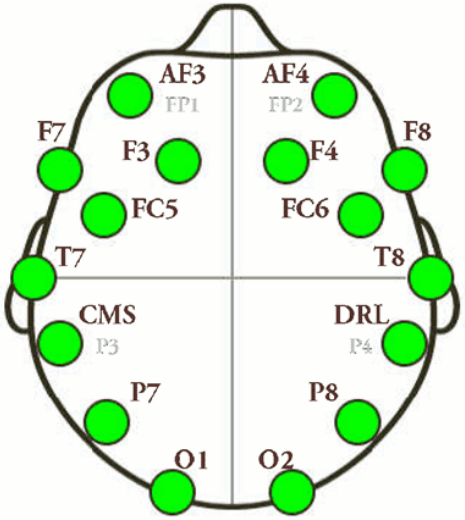
fNIRS Signals

- **Light intensity data**
  - 4 optodes in two wavelengths (730nm and 850nm)

- **Oxy data**
  - Deoxygenated hemoglobin (deoxy-Hb)
  - Oxygenated hemoglobin (oxy-Hb)
  - Hbt: sum of deoxy-Hb and oxy-Hb
  - Oxy: difference of deoxy-Hb and oxy-Hb

Electroencephalography (EEG)

- **Emotiv EPOC headset [2]**
  - EEG records electrical activities along the scalp.
  - EEG measures voltage fluctuations resulting from current flows within the neurons of the brain.
  - 10-20 system: 14 electrodes + 2 references



[2] https://emotiv.com/epoc.php

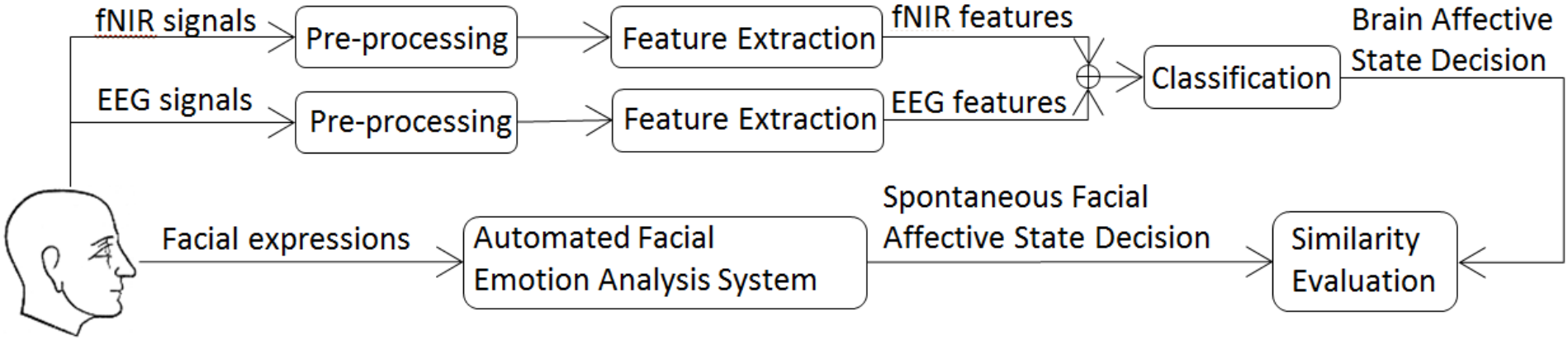
Problem Statement

- Neurophysiological correlates of emotional state have been investigated.
- fNIRS and EEG techniques are widely utilized due to their non-invasive and mobile natures.

Question: Whether are the inner affective states translated by human brain activity consistent with those expressed on the face?

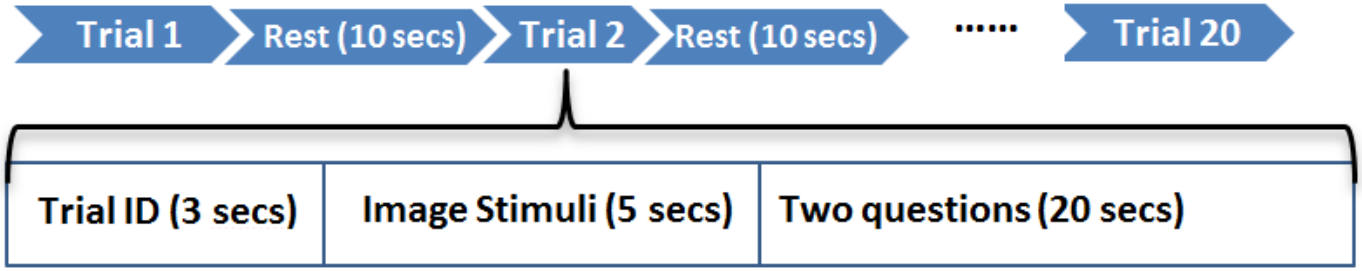
- **Proposed idea**
  - Explore the relationship of human facial spontaneous affective states and relevant brain activity by simultaneous use of fNIRS, EEG and facial expressions registered in video.

Framework of Methodology



Experiment Protocol

- 12 male participants (age: mean=27.58, SD=4.81) volunteered for the study.
- Task: watch 20 image-based stimuli
- Simple questions to verify he understood the image content.
- All perceivers have no knowledge of the video/image content in advance.
- During the entire experiment, the participant is asked to wear fNIRS head pad and EEG headset.
- In order to collect the useful face response, the participant's frontal face is required and recorded when he/she watches videos.



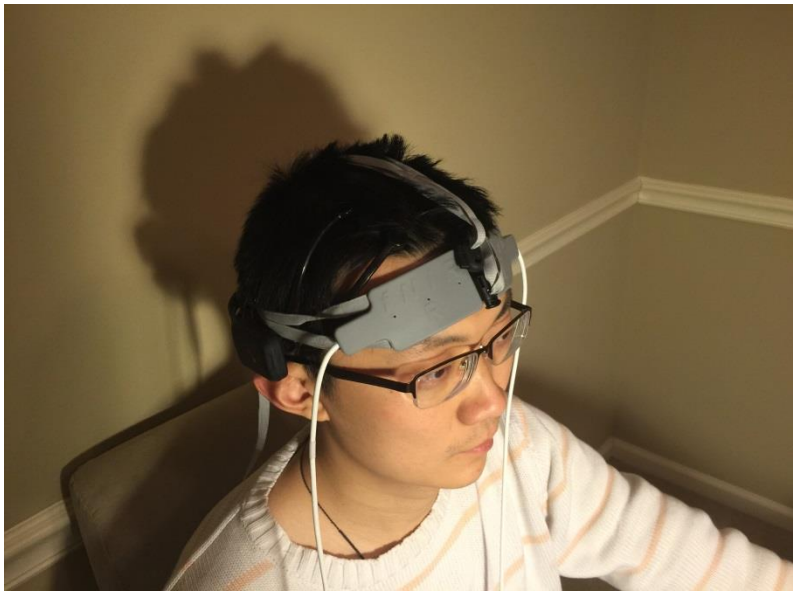
Note: The study has been approved by the Institutional Review Broad of New Jersey Institute of Technology. Before the experiment, each perceiver was asked to sign an agreement to participate in the study.

NJIT New Jersey's Science & Technology University

Drexel UNIVERSITY

Experimental Apparatus

- **fNIRS**: Functional Near Infrared Spectroscopy
- **EEG**: Electroencephalography
- **FE**: Automatic Facial Emotion Recognition System



Signal Processing

- **fNIRS signal**
  - Artifact correction
    - Clean the raw oxy data from heart pulsation, respiration, etc using Independent component analysis (ICA).
  - Signal filtering
    - Filter the noise in the high frequency.
    - A finite impulse response (FIR) low-pass filter with a cut-off frequency at 0.1 Hz and order of 20 is used, with blackman-harris windows.
  - Baseline adjustment
    - Subtract the mean of the signal at the 3 seconds prior to the beginning of the video.
  - Feature extraction
    - mean, median, standard deviation, max, min, and difference of max and min of hbo, hbr, hbt, and oxy, respectively for each optode in a trial.

Artifact Correction

- The modeled hemodynamic response represents the expected hemodynamic response to the given stimulus calculated by convolving the stimulus function and a canonical hemodynamic response function (HRF). The HRF consists of a linear combination of two Gamma functions [3] as

$$h(t) = A\left(\frac{t^{\alpha_1 - 1}\beta_1^{\alpha_1} e^{-\beta_1 t}}{\Gamma(\alpha_1)} - c\frac{t^{\alpha_2 - 1}\beta_2^{\alpha_2} e^{-\beta_2 t}}{\Gamma(\alpha_2)}\right)$$

where A controls the amplitude, α and β control the shape and scale, respectively, and c determines the ratio of the response to undershoot.

- The t-test was used to select the independent component associated with the hemodynamic response.

[3] 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, Mar. 2009.

Signal Processing

- **EEG signal**
  - Artifact correction
    - Clean the raw EEG data from facial muscle movement and noise
  - Feature extraction
    - Power spectral densities (PSD) of signals from 14 electrodes
      - AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4
    - 4 bands in the range of 4-30 Hz
      - theta (4-8Hz), slow alpha (8-10Hz), alpha (8-12Hz),beta (12-30Hz)
      - The asymmetry features
        - » The dfference between PSD of 4 systemrical pairs: AF3-AF4, F7-F8, F3-F4, FC5-FC6

Automatic Facial Emotion Recognition System

- **Feature extraction**

41 facial feature points

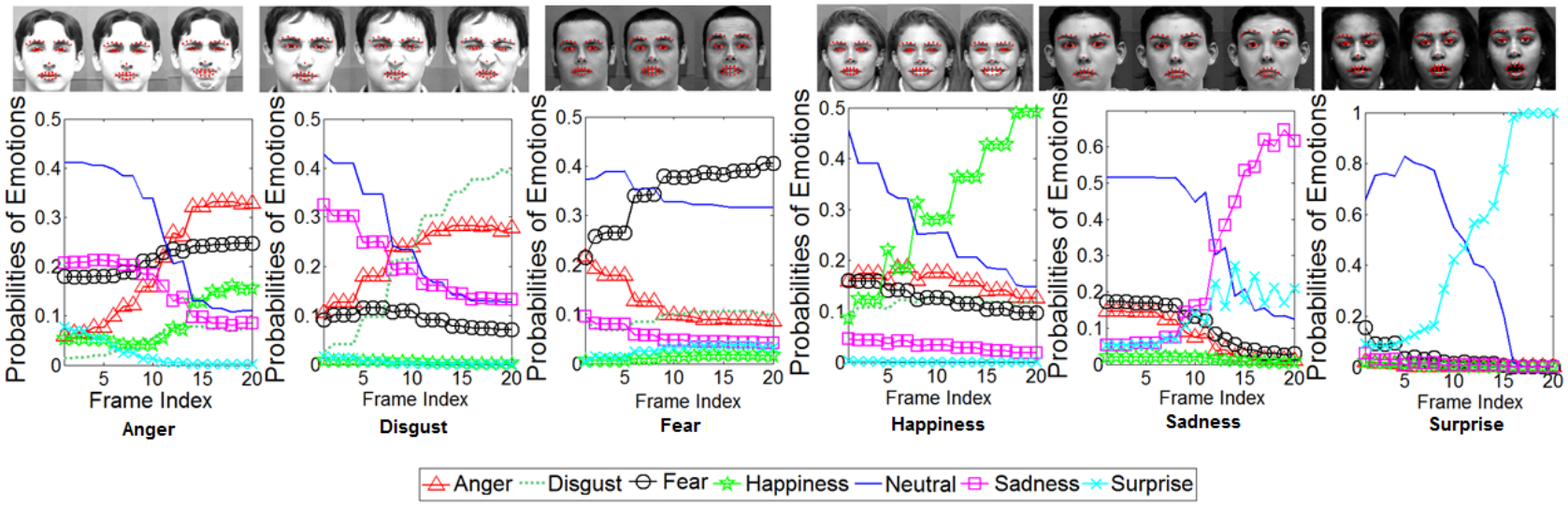
Eyebrows: 10 points	Mouth: 8 points
Eyelids: 12 points	Lips Corners: 10 points
Nose: 1 point (the anchor)	

- **Classifier**
  - Regional Hidden Markov Model
  - Y. Sun and A.N. Akansu, "Automatic Inference of Mental States from Spontaneous Facial Expressions", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2014.
  - Online demo system: http://emotions.njit.edu/

Automatic Facial Emotion Recognition System

- **Recognize seven types of basic facial emotions**
  - Anger, Disgust, Fear, Happiness, Neutral, Sadness, and Surprise.

- **Infer three types of Affects**
  - Positive, Negative, and Neutral



Automatic Facial Emotion Recognition System

- Emotions are recognized by the system by calculating the largest probability $P_v^{Emotion}$ of the emotion type

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

where N_{video} is the total number of videos. P_i is the probability of one emotion, which is average of the overall probabilities P_{ij} of this emotion in each frame of a video clip. P_{ij} is normalized to one. P_i is calculated as

$$P_i = \frac{\sum_{j=1}^{N_{frame}} p_{ij}}{N_{frame}}, \quad subject\ to \sum_{i=1}^{7} P_{ij} = 1$$

where N_{frame} is the number of fames of a video.

Affective States

The Affect ratings of both the positive affect P_{pos} and the negative affect P_{neg} for a video are separately generated by the system by normalizing the probability of Happiness and the largest probability among Anger, Disgust, Fear and Sadness.

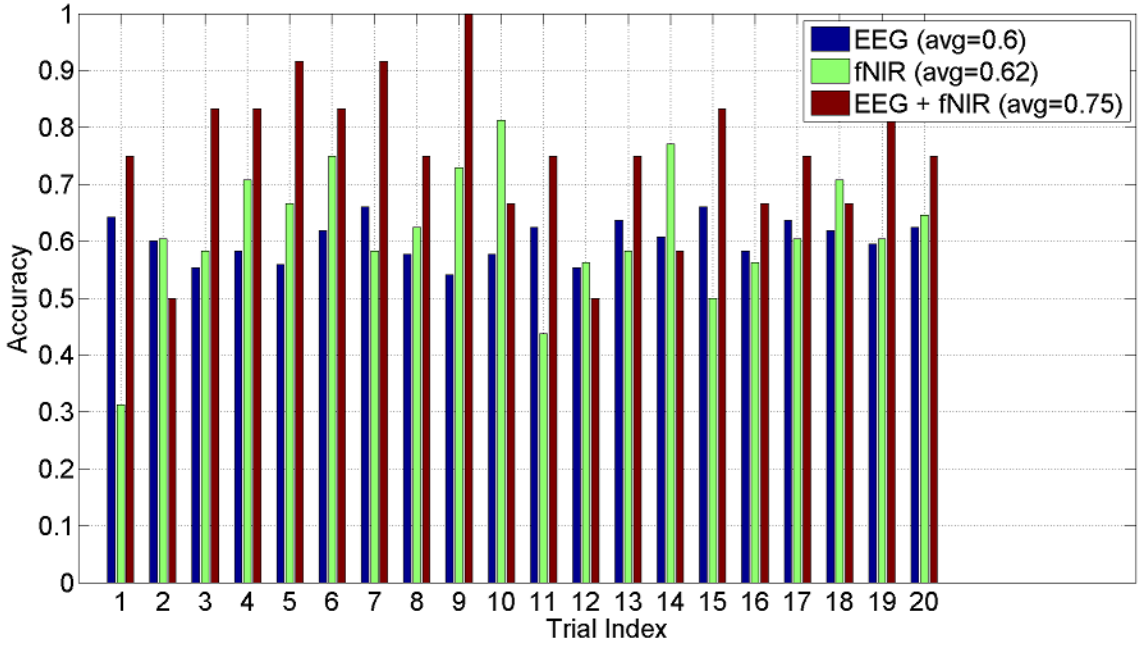
- The larger affect rating P_v^{Affect} of the two affect values is selected as the affect recognized in the corresponding video clip.

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

$$\text{subject to} \begin{cases} P_{pos} = \dfrac{P_{Happiness}}{P_{Happiness} + \max(P_{Anger}, P_{Disgust}, P_{Fear}, P_{Sadness})} \\ P_{neg} = \dfrac{\max(P_{Anger}, P_{Disgust}, P_{Fear}, P_{Sadness})}{P_{Happiness} + \max(P_{Anger}, P_{Disgust}, P_{Fear}, P_{Sadness})} \end{cases}.$$

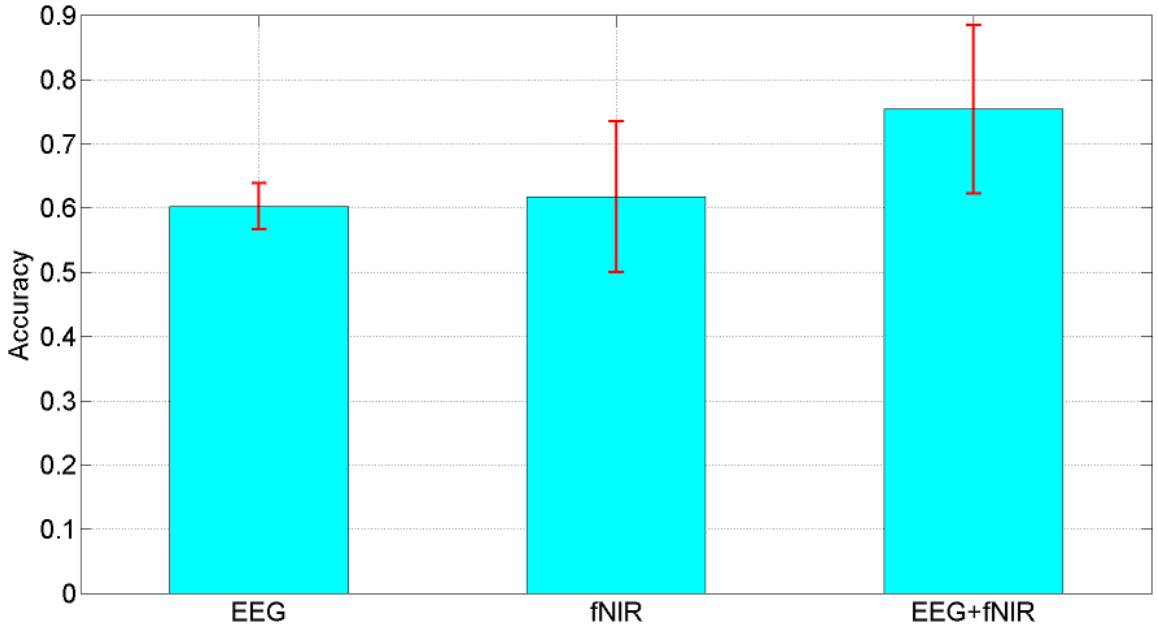
Experimental Performance

- Polynomial Support Vector Machine (SVM) was used for classification.
- For comparison, the univariate modality of either fNIRS or EEG features for recognition of the affective state is also applied.
- The leave-one-out approach for training and testing was applied.



Comparison

- **Image-content stimuli**
  - Standard Deviation:
    - 0.0416 (EEG only)
    - 0.1171 (fNIRS only)
    - 0.1310 (EEG+fNIRS)



Similarity

Similarity scores are calculated by the percentages of spontaneous facial affect recognized by the system same as the inner affective states translated by perceiver's brain signals.

Average Similarity of Perceivers' Spontaneous Facial Affective Response to the **Image-content** Stimulus Each Trial (avg_all_trial = 0.75)

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

- **Wilcoxon Signed Rank Test** is used to test the significance.
- Image-content trials
  - W=65, p=0.04>0.001

The result indicates that the spontaneous facial affective states are consistent with those translated by human brain activity.

Evaluation of Perceivers' Expressiveness

- The one-sample t-test is used to assess whether the average of measured neutral states is close to the population mean (μ=0.3) as there are three types of affective states.

- Image-content trials
  - Mean=0.26, SD=0.09, $t(11)=1.01$, $p>0.5$

The result explains that all perceivers are not expressive during the experiment.

Conclusion

- People's spontaneous facial expressions are demonstrated to be able to 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 alone as they are known to contain complementary information.
- 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 altogether.
- The experimental results validate the proposed method. In addition, the results explain the reliability of spontaneous facial expressions used in various applications.

NJIT New Jersey's Science & Technology University

Drexel UNIVERSITY

Thank you !

Q&A