EEG Hyperscanning for Eight or more Persons - Feasibility Study for Emotion Recognition using Deep Learning Technique

Sunghan Lee, Sangjun Han and Sung Chan Jun*
*Gwangju Institute of Science and Technology, Gwangju, Korea
E-mail: scjun@gist.ac.kr Tel: +82-62-7152216

Abstract— **Multi-user electroencephalogram (EEG) system is necessary to study concurrent activity among many persons. It is difficult to find a system that measures multiple EEG signals from more than even three people simultaneously. Therefore, we suggested a framework that is able to acquire EEG signals of more than eight persons at the same time and investigated the feasibility of this system. Acquisition was performed by using OpenViBE software developed by INRIA. Wireless EEG devices for our proposed framework were manufactured by BioBrain, Corp. in Korea. A device consists of eight channels measuring frontal EEG at a speed of 1 KHz sampling rate. While participants wore this system and did emotional video watching task as a group audience, their brain signals were acquired. To show its feasibility and efficacy, our preliminary result is analyzed using deep learning technique.**

I. INTRODUCTION

Recently, many attempts have been made to understand the interaction among humans. Investigating brain activities during interaction of two or more persons, using techniques simultaneously acquiring brain signals is called hyperscanning [1]–[4]. Human interaction activities appear in the form of one-to-one (i.e. conversation), one-to-many (i.e. class), and many-to-many (i.e. debate). Type of interaction is very diverse, but most hyperscanning studies have been done in a one-to-one interaction environment. For example, according to the recent review paper of F. Babiloni and L. Astolfi [3], 25 among a total of 28 hyperscanning studies were one-to-one interaction studies, except for three EEG hyperscanning studies. There may be some reasons why it is so rare for multilateral interactions. For example, it is not easy to analyze the correlations of signals in the multilateral system, and there are difficulties in recruiting a large number of participants. Also, another important reason is difficult to construct a system to study interactions for multiple persons. It does not only need to acquire multiple brain signals at the same time, but also requires various functions such as presenting stimulus, synchronization between multiple devices, and recording trigger information for event. However, it is hard to find a well-designed system that could measure signals more than even three persons. Therefore, in this study, we designed and implemented a framework for acquiring multiple signals of eight or more persons. In order to verify its applicability, we designed and conducted an experiment that can utilize the functionality of the developed framework.

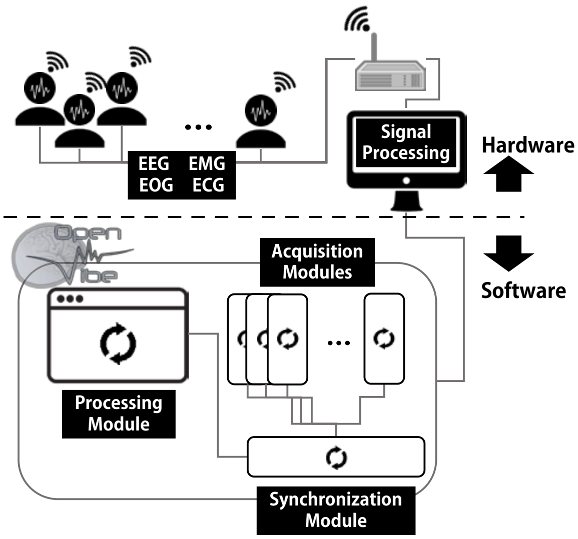

Fig. 1 Hardware and software configuration diagram of the proposed framework

Using proposed system, signals from multiple persons were successfully acquired and stored simultaneously. In addition, emotion classification was performed on acquired data using deep neural networks (DNN) technique.

II. FRAMEWORK DEVELOPMENT

The design and specifications of the hardware and software that make up the proposed framework are introduced. The system aims at acquiring electroencephalogram (EEG) signals at high speed in a wireless environment from eight or more persons. A schematic diagram of the developed framework is shown in Fig. 1.

A. Hardware Configurations

The EEG measurement hardware module was developed by BioBrain, Corp. in Korea. The detailed appearance is illustrated in Fig. 2. The equipment has 8 channels consisting of 7 monopolar channels and 1 bipolar channel. The hardware acquires signal at a speed of 1 kHz and has a 24 bit quantization resolution. The ADS 1299 Chipset from Texas Instrument was used as a signal measurement chip. A 9 volt

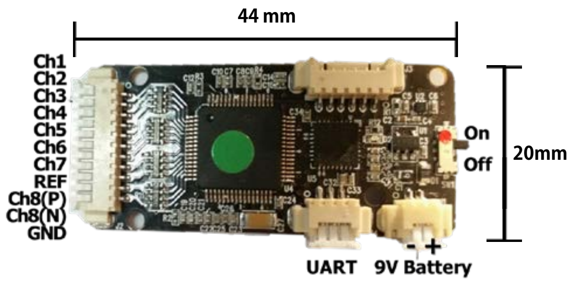

Fig. 2 EEG measurement device appearance

battery was used for the independent power source (for wireless communication).

Packet structure consists of header and payloads. Header consists of synchronization part (2 bytes), device information (1 byte), and 1 byte of packet number as a timestamp. The payloads are a total of 24 bytes with 8 channels of 3 bytes (24 bit resolution per channel). Therefore, the transmission rate of one device is 28 kB per second.

It supports both wired and wireless communications. Particularly, universal asynchronous receiver/transmitter method was used for wired communication. Universal serial bus interface was used to connect the server with devices. 2.4 GHz radio frequency was used as carrier for a wireless communication. Custom transmitter and receiver hardware were manufactured.

B. *Software Configurations*

Signal acquisition software was developed in two directions. The first was signal acquisition through custom-developed software and the second was interfacing with the acquisition module of OpenViBE [5].

Firstly, we developed signal acquisition module controlled by multithread in one process. Each device was assigned to each thread and multiple devices were controlled at the same time. Using self-developed software is advantageous in that it has a high degree of freedom in implementation and is easy to add functions as needed.

However, in order to use the software for conducting the experiment, it is necessary not only to acquire the signal of each device, but also to present stimulus, insert an event trigger, and so on. Implementing all these functions takes a long time and tremendous cost. Fortunately, there are well-established open-source frameworks for experiment. OpenViBE, one of well-organized open source software, was utilized for our own purpose. Thus, custom-acquisition modules for OpenViBE that receive signals from all devices has been developed.

OpenViBE is open source software for real-time neuroscience developed in INRIA. It provides a variety of functions such as MATLAB and python scripting, file input/output, real-time filtering, and real-time signal acquisitions. The structure for simultaneous acquisition of multiple signals and synchronization of the acquisition modules used in this study is illustrated in Fig.1.

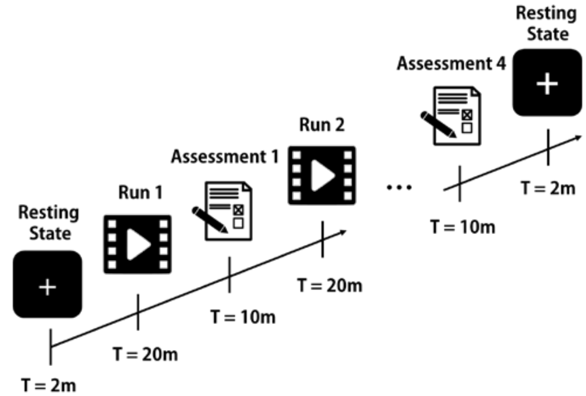

Fig.3 Outline of experiment progress. The lowercase m represents minutes

III. Test of Proposed Framework

It is necessary to verify that the developed framework may work properly in real environment. Therefore, our research team conducted emotion-evoking video viewing experiment with 8 people at the same time using the developed framework. The DNN technique was applied to detect emotion.

A. *Experiments*

A total of 40 undergraduate students (17 female, age 21.25±1.04) participated in the experiment. A group of eight persons participated in one experiment at the same time. Five experiments were conducted.

Eight participants were asked to sit in seats about 5 meters away from the screen and watched the video clips together. The video clips had been reasonably selected to do induce four types of emotions: fear, pleasure, boredom, and sadness.

The experiment consisted of 4 runs and 2 resting phases. The resting state was measured for two minutes before the first run and after the last run. The duration of each run is 20 minutes, and the order of the presented video clip was set randomly for each experiment. About 10 minutes of rest and evaluation time were given between runs. During this evaluation time, participants surveyed about the emotional

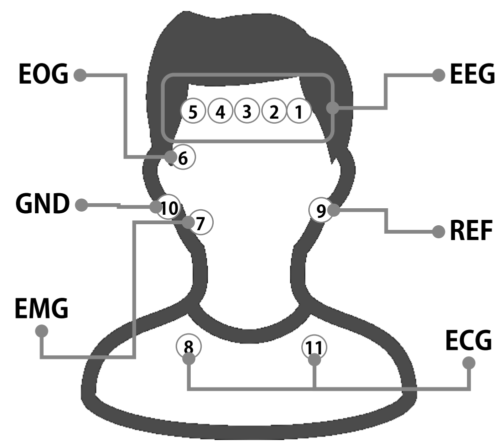

Fig. 4 Electrode location used in the experiment.

changes caused by the presented videos. Also, participants were asked to indicate their emotional states according to the quadrant of the Russell [6] model.

The location of the electrodes used is depicted in Fig. 4. Detachable form electrode was used. Five channels of the eight channels were used to measure the EEG of the frontal area (AF8, Fp1, Fpz, Fp2, and AF8) and the remaining channels were measured for electrocardiogram (ECG), horizontal electrooculogram (EOG), and electromyogram (EMG) near the chin. Left and right mastoids were used as reference channels.

B. *Data Classification*

Raw signals acquired from the experiments were contaminated due to movement, environmental noise, and detachment of electrodes. Thus, signals were inspected manually and ones with high contamination were removed. After rejection, due to severe contamination, signals from 13 participants only (out of 40) were used for analysis.

It is not easy to remove eye movement noise by applying independent component analysis (ICA) because recorded data had a few channels. Signals were bandpass filtered to 8-25 Hz to eliminate the frequency bands that might be caused by blinking and muscle movement. Also, a 60 Hz notch filter was applied to eliminate power noise. 10 minutes signals only were chosen by excluding the first 5 minutes and the last 5 minutes. After that, signals were split into 8000 samples (eight seconds long for each) without overlapping and used as an epoch. Each epoch was used as an input to a one dimensional convolutional neural network (CNN) and one classifier was created for all subjects. The classification results were verified using 5-fold cross validation.

EEG data recorded during each run were labeled with corresponding emotions, and four classes were classified through the CNN technique. All layers used rectifier linear unit (ReLU) as activation function and included drop out layer (with 0.2 ratio) except last dense layers for the classifier. The initial weight values of neurons were assigned through

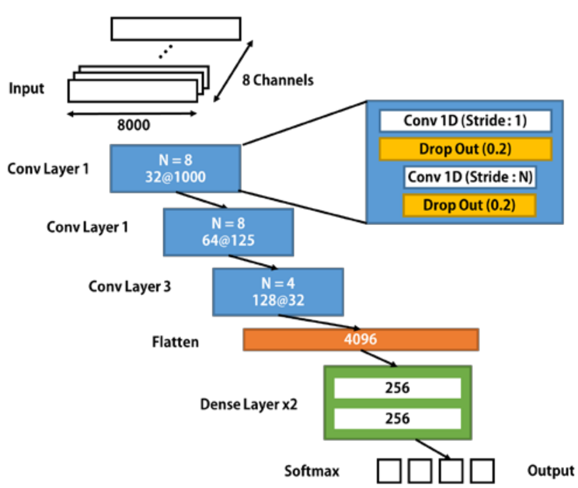

Fig. 5 Structure of proposed 1-D CNN. Each channel behaves like an RGB channels of image in a 2-D convolution.

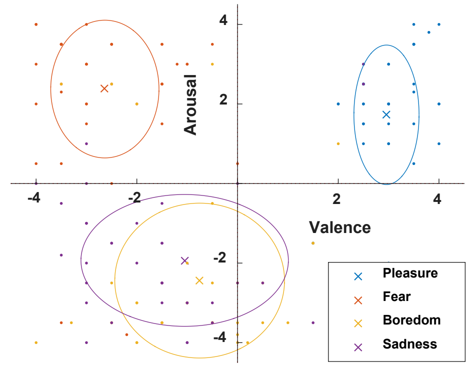

Fig. 6 Evaluation result of each video clips by participants

Tab. 1 averaged classification accuracy (%)

Classified Input	Pleasure	Fear	Boredom	Sadness
Pleasure	88.65	6.73	2.24	2.37
Fear	27.47	48.97	10.46	13.10
Boredom	25.16	12.01	59.09	3.73
Sadness	27.05	13.83	11.82	47.29

He initializer[7] and Adam [8] was used as the optimizer. Finally, the total number of parameters in this network was 1,276,100. The detailed structure of CNN used for emotion classification is illustrated in Fig. 5.

IV. RESULTS

Emotion evaluation results according to the Russell [6] model were depicted in Fig. 6. Colored dot is the raw data marked by the participants, 'x' mark indicates the average of the raw data, and colored ellipses are the standard deviations. It is observed that these evaluation results were quite similar to those of the previous literatures [6], [9], [10]. As a result, we could assume that the participants are likely to feel emotions induced by the video-clips.

Tab.1 shows the classification accuracy from 5-fold cross validation. The average classification accuracy using 1-D CNN was 61%. Among the four classes, accuracy of the emotion pleasure was notably higher than others. Interestingly, the probability of properly matching the pleasure class was very high, but other emotions were highly misclassified as the class of pleasure.

V. DISCUSSIONS

In terms of synchronization, the ideal speed control is to use synchronization hardware in a wired environment. Hardware-based synchronization has the advantage of minimizing delays in high-speed environments, but it is not easy to manufacture and takes a lot of cost. Also, the number of wired connections increases if the number of participants increases. In reality, wired connection may be very inconvenient if hyperscanning experiments are conducted in a situation considering a number of users. Therefore, we choose software synchronization through wireless connection considering the convenience of connection, scalability and cost.

OpenViBE's acquisition module is basically designed to acquire signals from one equipment. In this study, we developed software modules to communicate with multiple EEG devices based on the basic form of acquisition module in OpenViBE. Each device requires the same number of processes in such a structure. A process which is active for signal acquisition continues to receive incoming signals from the device in real time until the user stops operating. We observed that these processes did not generate any problems and seemed working properly when the number was small. However, increased processes consumed a lot of hardware resources individually, which puts a heavy load on the system. Central processing unit (CPU) utilization becomes close to 100 % and it becomes difficult to function when more than 12 devices are activated at the same time (environment under i7-6700k CPU, 32 GB memory and Windows 10 operating system).

Therefore, it is necessary to change the current structure for stable connection with more than 12 devices. We will pursue to modify the structure in which only one process is executed when connecting to the device. Multi-threads will be in communication with the acquisition devices instead of many processes. It is expected that the load will be reduced compared to the previous structure because threads shares many resources such as code area and data area in the same process.

In terms of time-synchronizations, useful software tool called the lab streaming layer (LSL) [11] has been developed in the Swartz Center for Computational Neuroscience. LSL is open source software that has various powerful features for handling time series; it includes time-synchronization modules for multi-modality, thus it has the potential to be used for multi-user hyperscanning environment. It is quite interesting to test the LSL operation in the multi-user environment for more than 10 devices. This implementation of signal acquisition modules through LSL will be done in the future study.

In terms of classification accuracy, obtained results seem not high. We believe that the data used for the learning was limited and it may result in low performance. As described previously, data of 13 subjects were used due to unavoidable noise issue. Two main causes of noise were found. First, it was too difficult to control many participants at the same time during the experiment. Second, the watching time per video was too long to maintain concentration of participants. The longer video was more likely to make movements due to yawning and stretching. Therefore, a well-controlled experimental paradigm should be designed. Effort to reduce noise may lead to improvement of classification performance by making the sufficient data available to the DNN. Although the classification rate was not high enough, we believe that it achieved reasonable results with only a small number of channels and little pre-processing, and even a simple DNN structure. It was found that DNN might be effectively used in end-to-end learning even under the worst conditions with very few channels. It is expected that we may be able to achieve more enhanced performance through in-depth analysis of the data, change of the DNN structure, transfer learning using the pre-trained networks such as VGG, ResNet, etc., and parameter adjustment.

We implemented simultaneous EEG acquisition framework for multi-user (up to 8 persons) and successfully acquired the signals using this framework. It would greatly benefit the study of multilateral interactions because such a system is currently hard to develop. Although the framework has been developed, well-organized experimental paradigm and advanced analysis methods for multi-users should be proposed. Further, it is necessary to intensive analysis of multilateral interactions. Future research should focus on the stabilization of our proposed framework in multi-user environment with more than at least 12 persons. Also, hyperscanning experiments to study interaction between multi-users using our framework are currently under investigation.

VI. Conclusions

We have designed and implemented a system that has a capability of simultaneous acquisition from multi devices for the purpose of multi-user hyperscanning research. A proposed framework could be used for designing and conducting the hyperscanning experiments. We introduced components and detailed specifications of our developed system. To verify the feasibility and usability of the system, we designed a simultaneous emotion-evoking video watching experiment. We successfully acquired eight person's biological signals simultaneously in wireless environment using the developed system. We used the DNN technique to classify emotion states from the acquired biological signals. Using the proposed system, it is expected that we can conduct one-to-many or many-to-many interaction hyperscanning studies. Thus, the framework introduced here is expected to contribute to the future research of various interactions and multi-mind brain-computer interfaces.

Acknowledgment

This work was supported by Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (No. 2017-0-00451).

References

[1] E. Delaherche, G. Dumas, J. Nadel, and M. Chetouani, "Automatic measure of imitation during social interaction: A behavioral and hyperscanning-EEG benchmark," *Pattern Recognit. Lett.*, vol. 66, pp. 118–126, 2015.

[2] T. Koike, H. C. Tanabe, and N. Sadato, "Hyperscanning neuroimaging technique to reveal the 'two-in-one' system in social interactions," *Neuroscience Research*, vol. 90. pp. 25–32, 2015.

[3] F. Babiloni and L. Astolfi, "Social neuroscience and hyperscanning techniques: Past, present and future," *Neurosci. Biobehav. Rev.*, vol. 44, pp. 76–93, 2014.
[4] P. R. Montague *et al.*, "Hyperscanning: Simultaneous fMRI during linked social interactions," *NeuroImage*. 2002.
[5] Y. Renard *et al.*, "OpenViBE: An Open-Source Software Platform to Design, Test, and Use Brain–Computer Interfaces in Real and Virtual Environments," *Presence Teleoperators Virtual Environ.*, 2010.
[6] J. A. Russell, "A circumplex model of affect," *J. Pers. Soc. Psychol.*, 1980.
[7] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015.
[8] D. P. Kingma and J. L. Ba, "Adam: a Method for Stochastic Optimization," *Int. Conf. Learn. Represent. 2015*, 2015.
[9] J. Abdur-Rahim *et al.*, "Multi-sensor based state prediction for personal mobility vehicles," *PLoS One*, 2016.
[10] G. Paltoglou and M. Thelwall, "Seeing stars of valence and arousal in blog posts," *IEEE Trans. Affect. Comput.*, 2013.
[11] A. Delorme *et al.*, "EEGLAB, SIFT, NFT, BCILAB, and ERICA: New tools for advanced EEG processing," *Comput. Intell. Neurosci.*, 2011.