

# A Multi-Day and Multi-Band EEG Dataset for Steady-State Visual Evoked Potential

Contact: Han-Jeong Hwang ([h2j@kumoh.ac.kr](mailto:h2j@kumoh.ac.kr))

Citation: Ga-Young Choi, Chang-Hee Han, Young-Jin Jung, **Han-Jeong Hwang\***, "A multi-day and multi-band dataset for steady-state visual evoked potential-based brain-computer interface", Gigascience, potentially accepted.

## Experimental Paradigm

Thirty subjects (9 females and 21 males;  $23.8 \pm 1.3$  years) were recruited for this study. Figure 1(a) shows the schematic diagram of the SSVEP stimulation. In order to control the stimulator, we used a LAUNCHXL-F28027 Board powered by C2000 MCU (Texas Instrument). A duty cycle was set at 50%, meaning that an LED has 50% on-time and 50% off-time. The three different frequency bands (low: 1-12 Hz, middle: 12-30 Hz, and high: 30-60 Hz) were independently used for SSVEP stimulation to get multi-band SSVEP datasets in this study. Three sets of four stimulation frequencies selected for each frequency band were as follows: 5, 5.5, 6, and 6.5 Hz for the low frequency band; 21, 21.5, 22, 22.5 Hz for the middle frequency band; 40, 40.5, 41, and 41.5 Hz for the high frequency band. We assigned four stimulation frequencies to four LEDs, depending on the stimulation frequency band, as shown in figure 1(b). Note that all instructions were presented on the center of the monitor and the subjects could perceive them through the center whole of the stimulator. For each trial, a blank screen was presented for 5 s, and then an arrow indicating one of the four LEDs was presented for 6 s, during which the subject was asked to gaze at a target LED according to the direction of an arrow. After that, a white cross mark ('+') was presented for 6 s as a short break for the next trial. A short beep sound was also presented for every transition of visual stimuli in order to lead to more explicit attention of the subjects. The direction of an arrow was randomly presented 20 times (20 trials) for each direction, resulting in a total of 80 trials, which was repeated for each frequency band (low, middle, and high), respectively. Each subject performed six sessions of an SSVEP experiment (2 sessions x 3 frequency bands), which was conducted twice for different days with an interval of at least one day. The three stimulation frequency bands were alternatively used for counterbalancing between subjects. In particular, all possible order combinations of the three frequency bands were as follows: (low-middle-high), (low-high-middle), (middle-low-high), (middle-high-low), (high-low-middle), (high-middle-low). Each order was randomly assigned to five subjects, respectively (6 orders x 5 subjects = 30 subjects), and a same order was used for two days once an order was assigned to each subject in the first day experiment.

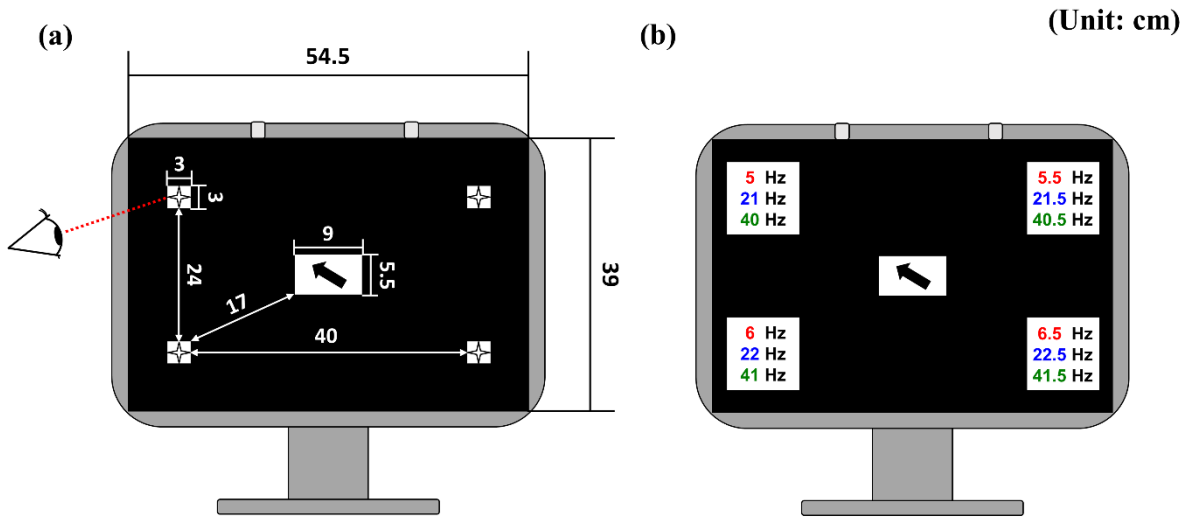
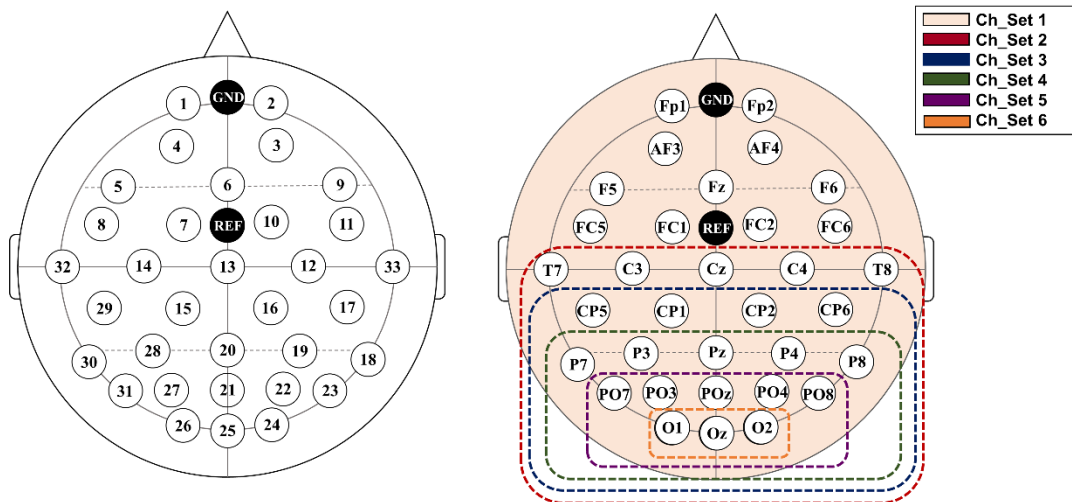


Figure 1. (a) Schematic diagram of the SSVEP stimulation, (b) Placement of four stimulation frequencies for each of three stimulation bands (5, 5.5, 6, and 6.5 Hz for the low frequency band; 21, 21.5, 22, 22.5 Hz for the middle frequency band; 40, 40.5, 41, and 41.5 Hz for the high frequency band).

### Data recording

EEGs were measured using a BrainAmp EEG amplifier (Brain products, GmbH Ltd, Germany) with a sampling rate of 1,000 Hz, where the ground and reference electrodes were attached on Fpz and FCz, respectively (figure 2). We used thirty-three active electrodes mounted based on the international 10-10 system to measure EEGs (FP1, FP2, AF4, AF3, F5, Fz, FC1, FC5, F6, FC2, FC6, C4, Cz, C3, CP1, CP2, CP6, P8, P4, Pz, POz, PO4, PO8, O2, Oz, O1, PO3, P3, CP5, P7, PO7, T7, and T8), where electrodes were more densely mounted around occipital areas than the others because SSVEP is mainly originated from the occipital lobe.



**Figure 2.** Electrode position used in the experiment with respect to (a) number and (b) position name. Note that six different channel sets are used for data analysis to see the impact of the number of electrodes on classification performance.

## Data file description

Because data analysis was performed using Matlab R2013b (MathWorks, Natick, MA, USA), we provide our dataset in the form of Matlab files (.mat). Each data folder named as subject initial (e.g., S1) has two sub-folders for two sub-datasets measured from two different days (i.e., day1 and day2). Each sub-folder has cnt and mrk files, which contain continuous time series data for all physiological measurement (cnt) and the trigger information of the corresponding data (mrk), respectively. The cnt and mrk files have suffixes corresponding to three frequency bands and session numbers. For example, cnt\_Low(1) means time series data measured using the low frequency band for SSVEP stimulation in the first session. Thus, the sub-folder of each subject contains the following six pairs of cnt and mrk files: cnt\_Low(1), mrk\_Low(1), cnt\_Low(2), mrk\_Low(2), cnt\_Middle(1), mrk\_Middle(1), cnt\_Middle(2), mrk\_Middle(2), cnt\_High(1), mrk\_High(1), cnt\_High(2), and mrk\_High(2). All data were down-sampled to 200 Hz when converting the raw data into Matlab compatible files. Table 1 shows all data files provided for each sub-folder.

**Table 1.** Data format. Each data folder of all subjects has two sub-folders for two sub-datasets measured from two different days, and each sub-folder has six pairs of *cnt* and *mrk* files (shown below) according to stimulation frequency band and session number.

Frequency Band	Stimulation Frequencies	Data format (*.mat)
Low	5 Hz	
	5.5 Hz	<i>cnt_Low(1)</i> , <i>cnt_Low(2)</i>
	6 Hz	<i>mrk_Low(1)</i> , <i>mrk_Low(2)</i>
	6.5 Hz	
Middle	21 Hz	
	21.5 Hz	<i>cnt_Middle(1)</i> , <i>cnt_Middle(2)</i>
	22 Hz	<i>mrk_Middle(1)</i> , <i>mrk_Middle(2)</i>
	22.5 Hz	
High	40 Hz	
	40.5 Hz	<i>cnt_High(1)</i> , <i>cnt_High(2)</i>
	41 Hz	<i>mrk_High(1)</i> , <i>mrk_High(2)</i>
	41.5 Hz	