MEDICON 2019 Scientific Challenge
Dataset Description

Setup description

The setup has two main components: i. Data acquisition module and ii. Stimuli and Feedback presentation.

Data acquisition module

To record EEG data we used the g.Nautilus (gTEC, Austria). The equipment has 16 active electrodes that do not require abrasive skin treatment and with completely wireless signal transmission. The EEG data were acquired from 8 electrodes positions (C3, Cz, C4, CPz, P3, Pz, P4, POz), the reference was placed at the right ear and the ground electrode at AFz. Sampling rate was set at 250Hz. Data were acquired notch filtered at 50Hz and passband filtered between 2 and 30Hz.

Stimuli and Feedback presentation

To train and execute the task proposed, we use the Vizard toolkit to visualize the virtual environment consisting of a bedroom with common type of furniture (shelves, a bed, a table, a chair, and a dresser) and objects (frames, books, lights, a printer, a radio, a ball, a door, a window, and a laptop).

The objects used throughout the experiment, and their respective labels, were:

1. books on a shelf, 2. a radio on top of a dresser, 3. a printer on a shelf, 4. a laptop on a table, 5. a ball on the ground, 6. a corkboard on the wall, 7. a wooden plane hanging from the ceiling, and 8. a picture on the wall.

The virtual environment was presented via the Oculus Rift Development Kit 2 headset (from Oculus VR).

Experimental design overview

The experiment was divided in 2 parts: calibration and online phase procedures.

Calibration

Consist of 20 blocks. Each block contains 10 runs of the experimental condition: Each one of the 8 objects in the scene flashed (green flashes) in a randomized order (Figure 1). The
highlight (flash) of each object occurred with a Interstimulus Interval of 200 ms. Each flash had the duration of 100 ms. The total number of ‘events’ per block is 80 (8 per run).

**Instruction and P300**

The target object in each block was selected based on pseudorandom number generator Mersenne Twister algorithm (i.e. 1 to 8). The participants were informed by the experimenter of the target object and instructed to count the number of times the target object would flash.

This protocol creates a rare event (target event probability of 1/8), thus generating a P300 brain response. At the end of each run, the participants were asked to confirm the object (behavioural control).

![Figure 1 - BCI training phase](image)

Calibration phase (with a total of 1600 events (200 target, 1400 non-target)) is used to optimize the number of runs used to classify the target object during the online phase (i.e. the number of runs is specific for each session, ranging from 3 to 10). The participant had an interval in the middle of the calibration phase.

**Online phase**

In the online phase of BCI (figure 2), we have a total of 50 blocks. Each block is composed by a number of runs previously selected based on the calibration phase. The structure of each run is similar to the calibration phase.
The total number of events is session-specific and is determined by the number of blocks\((50)\)\times number of runs\((\text{calibration-dependent})\)\times number of events\((8)\).

![Figure 2 - Online phase of BCI](image)

**Dataset description**

The competition dataset is divided into two parts: train and test sets. The train set is available with labels (the target for each block) for the contest participants to train their models. The test set is available without labels. The participants should submit the predicted labels.

![Figure 1 - Structure of the paradigm with its subdivisions by block, run and event.](image)
**Train set**

The train set consists of the calibration phase of each session. The true labels are available to train the models.

**Test set**

The test set consists of the online phase of each session, without the true labels.

**Procedure:**


2. Download the dataset from the link provided in the registration page.

4. Submit the predicted labels in the correct format to challenge@medicon2019.org

The data is provided in a folder structure:

```
SBJXX\n  \SYY\n    Train\n       trainData.mat
       trainEvents.txt
       trainTargets.txt
       trainLabels.txt
    Test\n       testData.mat
       testEvents.txt
       runs_per_block.txt
```

**Files Description:**

**Train Folder:**

- **trainEvents.txt** - one label [1..8] per line, corresponding to the order of objects flashed.
- **trainData.mat** - Data from the calibration phase, in the structure [channels x epoch x event], epoch the data samples from -200ms to 1200ms relative to the event stimulus.
- **trainTargets.txt** - 1 or 0 per line, indicating if the object flashed is target or not, respectively.
- **trainLabels.txt** - label of the target object [1..8] per line, one for each block.
Test Folder:

- **testData.mat** - Data from the online phase, in the same structure as the train data.
- **testEvents.txt** - one label [1..8] per line, corresponding to the order of objects flashed.
- **runs_per_block.txt** - The number of runs per block used in the online phase (3 to 10)

Additional Info:

- **data**: [8×350×1600 double] - each *.mat: channels per epoch per event
  - Number of epochs correspond to Number of Events * Number of Runs * Number of Blocks.

Training data:

- Number of Events: 8
- Number of Runs: 10
- Number of Blocks: 20

Test Data:

- Number of Events: 8
- Number of Runs: VARIABLE (check runs_per_block.txt)
- Number of Blocks: 50
  - Channels order: (C3, Cz, C4, CPz, P3, Pz, P4, POz)

- **xmin**: -0.2000 ms - start of epoch relative to the event trigger
- **xmax**: 1.2000 ms - end of epoch relative to the event trigger
- **srate**: 250 Hz - sample rate of the EEG acquisition

References
