

2008 年 BCI 竞赛数据提取方法

2012.05.30 LWC

1. 创建此文档的目的有二：

- 1) 概括 2008 年 BCI 竞赛数据的实验设计、数据采集参数；
- 2) 介绍数据提取脚本的思路，以帮助后继者快速熟悉数据，查找可能存在的错误；

2. 实验介绍：

根据《BCI Competition 2008 - Graz data set A》^[1]，该数据集包含 9 名受试者，受试者执行 4 类运动想象任务：想象左手运动，想象右手运动，想象双脚运动，想象舌头运动。每位受试者采集 2 个 session 数据，每个 session 包含 6 个 run，run 之间有短暂休息。每个 run 有 48 个 trial，每类各 12 个，每个 session 共 288 个 trial。数据采集频率 250Hz，有 22 个 EEG 电极，3 个 EOG 电极。详细实验设计和 EOG 采集见前述文档[1]。

3. 数据提取依据：

根据文档[1]，我们获知以下信息：

- 1) 实验时间安排如图 1 所示

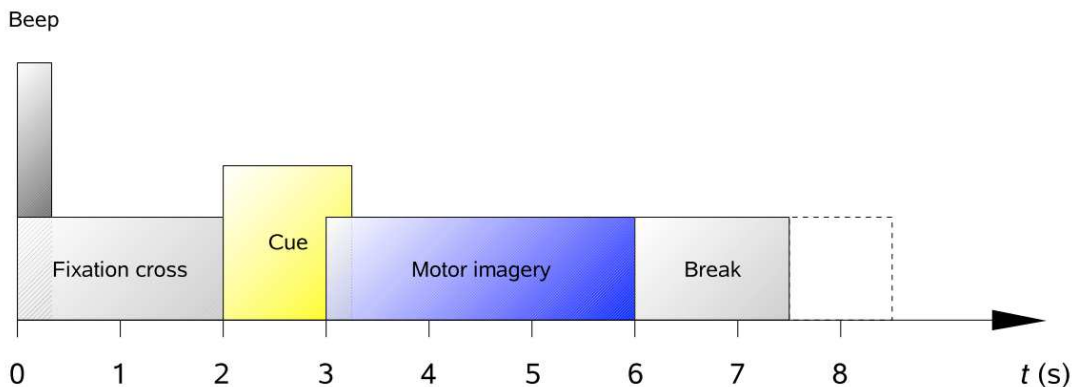


图 1 实验时间安排

- 2) 通过 BioSig 工具箱里的 sload 函数加载 GDF 格式数据，加载示例：

```
[s, h] = sload('A01T.gdf');
```

S 变量保存实际数据，h 保存 GDF 数据头，主要根据以下 3 个变量提取数据：

h.EVENT.TYP – 数据类型标识（详细类型后文）；

h.EVENT.POS – 数据起始位置；

h.EVENT.DUR – 数据持续行数；

这 3 个变量的各行对应。另外可以根据 h.Classlabel 确定运动想象类型，h.ArtifactSelection 确定 trial 是否有伪迹，

- 3) 标识含义见图 2

此处注意 768 标识图 1 中 Fixation cross 的开始位置，769-772 和 783 标识 Cue 起始位置，数据中有舍弃的 trial。

Event type		Description
276	0x0114	Idling EEG (eyes open)
277	0x0115	Idling EEG (eyes closed)
768	0x0300	Start of a trial
769	0x0301	Cue onset left (class 1)
770	0x0302	Cue onset right (class 2)
771	0x0303	Cue onset foot (class 3)
772	0x0304	Cue onset tongue (class 4)
783	0x030F	Cue unknown
1023	0x03FF	Rejected trial
1072	0x0430	Eye movements
32766	0x7FFE	Start of a new run

图 2 标识含义

4. 数据提取方法：

因为废弃 trial（标识为 1023）的存在，768 标识位置向后推移 750 行（第 1-3 秒）并不一定是适合的 trial 数据，若从 768 标识位置开始计算数据位置，有多种情形要考虑，而从 769-772 和 783 标识位置处可以很容易计算数据起始位置，计算思路：

搜寻 h.EVENT.TYP 中值为 769-772 或 783 的位置 -> 在 h.EVENT.POS 查找对应位置处的值 -> 找到的值 + 250（第 2-3 秒的提示）即为运动想象数据起始行 -> 由起始行向后取 750 行数据为 1 个 trial（共 3 秒，第 3-6 秒）。

该脚本采取的是这种方法提取。提取结果有 2 种版本：

1) subjectX_V1.mat 和 subjectX_EV1.mat

X 是 1-9 的数字，对应 9 名受试者。公共后缀 V1 的数据中各个 trial 数据放在一起，都在 data 变量中。没有 E 表示竞赛数据的训练数据，有分类标识；有 E 表示竞赛数据的测试数据，没有分类标记。

2) subjectX_V2.mat 和 subjectX_EV2.mat

V2 版本的数据中各 trial 数据分别存放，保存形式为 dX(X 为数字)。其他同上。