Naturalistic Music EEG Dataset - Tempo (NMED-T)

Citation

Steven Losorelli, Duc T. Nguyen, Jacek P. Dmochowski, and Blair Kaneshiro (2017). Naturalistic Music EEG Dataset - Tempo (NMED-T). Stanford Digital Repository.

If using the dataset, please also cite the following paper:
Steven Losorelli, Duc T. Nguyen, Jacek P. Dmochowski, and Blair Kaneshiro (to appear). NMED-T: A Tempo-Focused Dataset of Cortical and Behavioral Responses to Naturalistic Music. In Proceedings of the 18th International Society for Music Information Retrieval Conference, Suzhou, China.

Links

Dataset (Stanford Digital Repository):
https://purl.stanford.edu/jn859kj8079

Losorelli et al. (to appear) ISMIR preprint:
http://bit.ly/NMEDT_preprint

Contact

Blair Kaneshiro <blairbo@ccrma.stanford.edu>

Abstract

This dataset contains cortical (EEG) and behavioral data collected during natural music listening. Dense-array EEG was recorded from 20 adult participants who each heard a set of 10 full-length songs with electronically produced beats at various tempos. In a separate subsequent listen, each participant tapped to the beat of a 35-second excerpt from each song. Participants also delivered ratings of familiarity and enjoyment for each full-length song during the EEG recording. Finally, the dataset includes basic demographic information about the participants, as well as Matlab scripts to perform the illustrated analyses presented in the paper introducing the dataset (Losorelli et al., to appear). Cleaned and aggregated data are published in Matlab format; raw EEG is published in Matlab format, while raw tapping data are published in .txt format. Stimulus audio is not published, but metadata links are provided. Cleaned EEG files have been aggregated across participants on a per-song basis and range in size from 656-721MB. Raw EEG recordings (two recordings per participant) range in size from 753-861MB. All other files are less than 1MB in size. In total the dataset comprises 55 downloadable files, with a total size of 39GB.

License

Dataset files

- **NMED-T_README.pdf** (this document)
  Informational document describing the dataset (1 file).
- **songSS_Imputed.mat**
  Cleaned and aggregated EEG recordings for song SS (10 files).
- **PP_R_raw.mat**
  Single raw EEG recording for participant PP, recording R (40 files).
- **behavioralRatings.mat**
  Aggregated behavioral ratings (1 file).
- **TapIt.zip**
  Archive of aggregated and raw tapping responses (1 file).
- **participantInfo.mat**
  Basic demographics of participants (1 file).
- **Code.zip**
  Collection of Matlab scripts for performing illustrative analyses in the Losorelli et al. (2017) ISMIR paper, as well as helper functions and files (1 file).

The contents of these files are described in detail in the Losorelli et al. (2017) ISMIR paper.

Stimulus and trigger information

Stimulus triggers 21-30 correspond to the 10 songs used as stimuli, as noted in the table below. More information about the stimuli can be found in Table 1 of Losorelli et al. (2017).

Trigger	Song #	Title	Artist	min:sec
21	1	First Fires	Bonobo	4:38
22	2	Oino	LA Priest	4:31
23	3	Tiptoes	Daedelus	4:36
24	4	Careless Love	Croquet Club	4:54
25	5	Lebanese Blonde	Thievery Corporation	4:49
26	6	Canopée	Polo & Pan	4:36
27	7	Doing Yoga	Kazy Lambist	4:52
28	8	Until the Sun Needs to Rise	Rüfüs du Sol	4:52
29	9	Silent Shout	The Knife	4:54
30	10	The Last Thing You Should Do	David Bowie	4:58

In addition to the triggers listed above, the Raw EEG files contain triggers resulting from audio timing clicks and key presses. These triggers are all stored in the *DIN_1* variable. Trigger numbers and descriptions are noted in the table below.

Trigger	Description
128	Click sent from 2nd audio channel. Clicks are delivered 1 second into the audio, and in 60-second intervals thereafter (e.g., 1 second, 61 seconds, 121 seconds, etc.).
1-9	Ratings given in response to the questions after each song (familiarity first, then enjoyment).
10	Enter (initiate the very start of the experiment, and conclude it at the end).
11	Space bar (exit break screen and begin next trial).

Note: The *DIN_1* variable is a cell array which stores a combination of trigger-related numeric characters and the letters “DIN”. The Code.zip archive contains a helpful function `parseDIN.m` in the Misc folder, which takes in the *DIN_1* variable and outputs 2 numeric arrays: One with the trigger numbers, and the other with the corresponding timings of those triggers (in samples, a.k.a. msec for fs = 1000).

Mapping participant numbers from raw to cleaned files

The participants whose data are presented in the cleaned and aggregated response files do not reflect continuous numbering of participants in the raw files (i.e., participant numbering in the raw files goes from 1-23, with participants 1, 18, and 22 omitted). The table below maps the participant numbers of the raw files to their respective position in the preprocessed data files.

Clean	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Raw	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	19	20	21	23