Tempo Extraction From Electroencephalography Using a Single EEG Channel

Andreas Papagiannis

January 2019

1 Introduction

It has been shown that oscillatory neural activity is sensitive to accented tones in a rhythmic sequence as neural oscillations synchronize to rhythmic sequences [2]. In other words, when subjects hear rhythmic sequences, the magnitude of the oscillations changes for frequencies related to the metrical structure of the rhythm.

One could argue that as the brain processes the perceived music, it generates a transformed representation which is captured by the EEG electrodes. Hence, the recorded EEG signal could in principle be seen as a mid-level representation of the original music piece that has been heavily distorted by two consecutive black-box filters: the brain and the EEG equipment. So far, experiments suggest that using 64 EEG channels, the tempo of a song can be accuratelly estimated [1].

The goal of this paper is to discover if it is possible to correctly estimate the tempo of a rhythm by analyzing a single EEG channel. More specifically, the computation of tempo information is done by applying **beat tracking techniques** to extract the tempo from electroencephalography (EEG) recordings obtained from people listening to music stimuli.

Throughout this paper the term BPM (beats per minute) will be used to describe tempo in an accurate way.

2 Dataset

The selected dataset [3] is a result of ongoing joint work between the Owen Lab and the Music and Neuroscience Lab at the Brain and Mind Institute of the University of Western Ontario and is kindly released under the PDDL licence, which means that you can freely use it without any restrictions.

Figure 1: OpenMIIR logo

It contains EEG recordings of a number of participants who were listening to a small selection of music stimuli. The total number of participants was nine and the listened to 12 music stimuli samples. Also five trials were conducted for every participant and every music stimuli sample. This means that in total the dataset contains 540 EEG recordings. Each of these recordings were recorded with the BioSemi Active-Two system at a sample rate of 512kH using 64 EEG channels.

In this paper we analyze a subset of the whole dataset, namely all EEG recordings of one trial done for music stimuli labeled 's14' in OpenMIIR. That specific music stimuli sample is a part of famous children's song 'Mary Had a Little Lamb' and it's true BPM value is 160.

3 Computation of Tempo Information

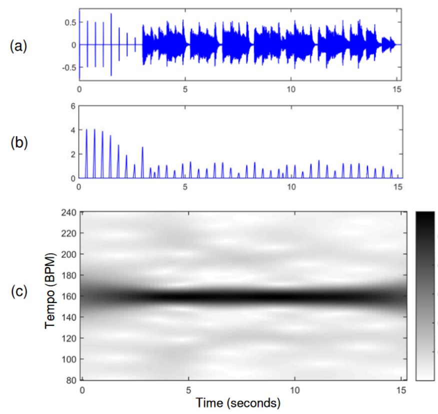

Figure 2: Tempogram computation for 'Mary Had a Little Lamb'. **(a)** Waveform signal. **(b)** Novelty curve. **(c)** Tempogram representation

The previously mentioned beat tracking technique used to extract tempo is as follows. From the inputted audio signal, it's waveform signal is transformed to a novelty curve. As the waveform signal typically peaks where the beats occur, it's novelty curve will spike at every peak of the audio signal representing where every beat is located in the audios time frame. Then the **tempogram** of the signal is computed by applying DFT (Discret Fourrier Transform) to the novelty curve of the audio. A tempogram should be thought as the eqivalent of a spectogram but instead of measuring pitch frequency on the y axis, the overall BPM of the audio signal is computed. Figure 2 depicts the whole procedure so far. After computing the tempogram of the audio signal, it's histogram is also computed. Lastly, the highest peak of the histogram is that of the estimated BPM.

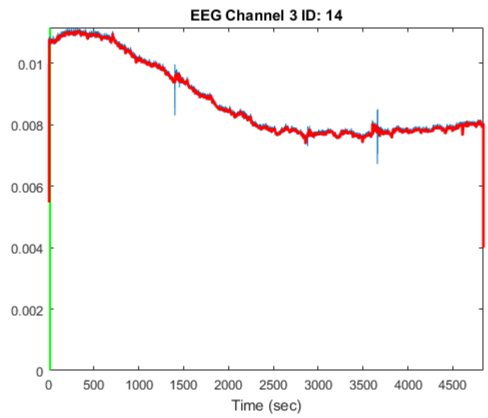

Figure 3: Local average curve of the 3rd EEG channel of participant 'P03'. The blue line represents the original EEG signal while the red line is it's local average curve.

In order to estimate BPM analyzing the EEG data instead of audio signal, the aforementioned procedure is repeated. The difference is that instead of computing the novelty curve of the audio waveform, we compute the local average curve of the recorded EEG signal (Figure 3) in order to normalize the EEG signal as shown in Figure 4. The signal is normalized between $(-1, 1)$ and should show the spikes that appear in the raw EEG recording. Besides the curvature of the raw EEG signal, it's normalised version is expected to be identical.

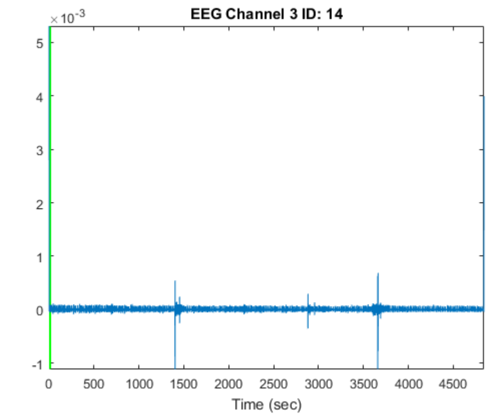

Figure 4: Normalized EEG signal of the 3rd channel of participant 'P03'

After having the EEG signal normalized, applying the DF transform to it results to the tempogram of the same EEG signal (Figure 5). The tempogram itself gives almost no information about the BPM of the music stimuli sample, which why the BPM histogram (Figure 6) must be computed.

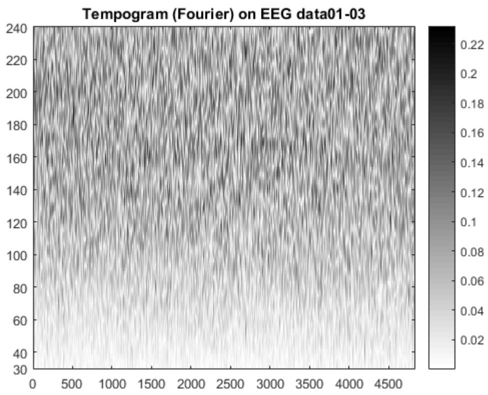

Figure 5: Tempogram of 3rd channel EEG for participant 'P03'

The last step is finding the maximum value of the y axis in the histogram. Since it is derived from the previously computed tempogram, the x axis projection of the maximum y axis value will represent the most "perceived" BPM value of the selected EEG channel and participant. Thus this value is sought as the estimation of the BPM of the audio stimuli the participant listened.

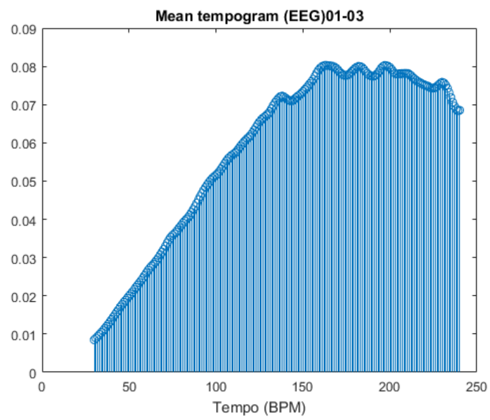

Figure 6: BPM Histogram of 3rd channel EEG for participant 'P03'

4 Results & Evaluation

In order to extrapolate any results, the procedure described in the previous section in repeated for all participants and all channels. This means that the computation of tempo information happens 576 times since the dataset in use consists of nine participants for each one of which there are 64 available EEG channels. After computing the BPM estimation for each channel and participant in that way, all estimations are agregated in a $9x64$ matrix.

In order to evaluate the results an the per EEG channel BPM error rate is defined as:

$$\left| \left(\sum_{1}^{\#\text{ participants}} \text{channel BPM estimation} \right) - \text{true BPM} \right|,$$

for each channel.

Since the true BPM of the audio stimuli sample is known, the per channel BPM error is simply the difference of the mean of all BPM estimations of that channel by all participants and of the true value of the stimulus's BPM. So if for example a channel has a BPM error of value 10 and the true BPM value is 120, it means that the respective channel BPM estimation was between $(110, 130)$.

Table 1: BPM errors of the five best performing EEG channels

Channel Index	BPM Error
3	0.22
2	1.33
35	2.00
1	2.56
36	3.33

Figure 7 shows the distribution of the BPM error for all channels sorted in an ascending way by the BPM error value. At least 10 of the EEG channels produce an accurate estimate of the BPM but for sure not all of them yield estimations close to the real value.

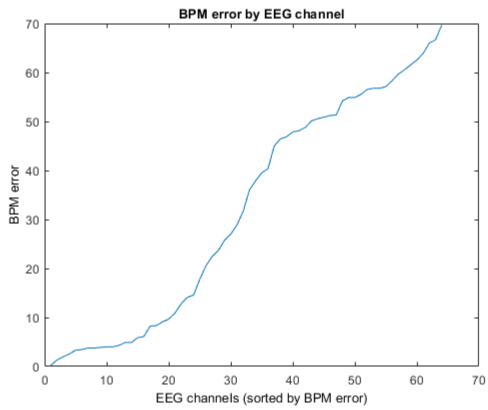

Figure 7: BPM error distribution of EEG channels for stimuli 's14'

Figure 8 shows the location of all EEG channels. The channels colored red are the five channels that produced the most accurate estimations. Next five best performing channels are the ones colored dark orange. The light orange and yellow channels obtained an BPM error < 10. The blue ones a BPM error between 10 and 20. The rest had BPM error values greater than 20 BPM.

The most important channel are located in the central area of the head. This comes in contrast with what is indicated in [1] where the aggregation filter used was heavily depending on the channels located close to the ears.

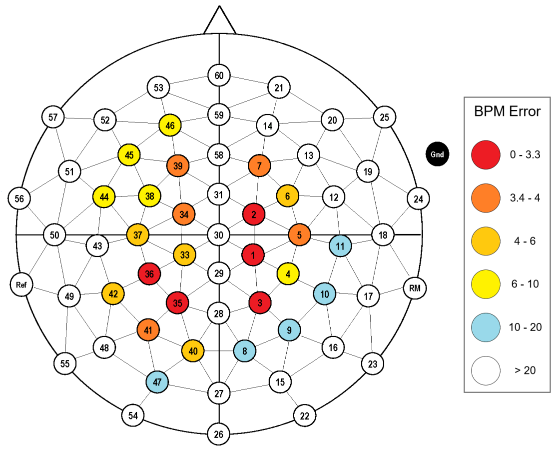

Figure 8: Topographic depiction of the most important EEG channels

5 Summary, Discussion & Future work

All the above work shows that BPM information **can** be accurately estimated by only one EEG channel under the condition that the recordings are replicated. This means suppose there existed only one participant, the accuracy of the BPM estimation would be worse. This method of estimating BPM is largely dependent on the EEG equipment used and also the participant himself. The estimations derived from the EEG channels of some participants yielded almost always accurate BPM estimations whereas for other participants no EEG channels could produce accurate results. Something also very important in the accuracy of the BPM estimations is the audio stimuli itself as repetitive audio waveforms with strong beat tend to produce more accurate results.

In order to find out which exactly channel is best of all for processing EEG data this way, this experiment should be replicated with more audio stimulus and participants. However, as mentioned above this paper shows that BPM estimation using only one EEG channel is possible.

References - Web Resources

[1] S. Stober, T. Thomas Prätzlich, M. Müller; *Brain Beats: Tempo Extraction From EEG Data*; Conference: Proceedings of the International Conference on Music Information Retrieval (ISMIR)January 2016. http://bib.sebastianstober.de/ismir2016.pdf

[2] M.S. Treder, H. Purwins, D. Miklody, I. Sturm, and B. Blankertz; *Decoding auditory attention to instruments in polyphonic music using single-trial EEG classification*; Journal of Neural Engineering, 11(2):026009, April 2014. https://www.ncbi.nlm.nih.gov/pubmed/24608228

[3] S. Stober; *OpenMIIR: a public domain dataset of EEG recordings for music imagery information retrieval* ; https://github.com/sstober/openmiir

[4] Swartz Center for Computational Neuroscience; *EEGLAB Toolbox*; https://sccn.ucsd.edu/eeglab/index.php