MODELLING PERCEPTION OF RHYTHMIC COMPLEXITY: COMPUTATIONAL AND NEURAL MEASURES

Jay Appaji[1] **Jacek P. Dmochowski**[2] **Blair Kaneshiro**[1]
[1]Stanford University, USA [2]City College of New York, USA
{jay.appaji, blairbo}@gmail.com

ABSTRACT

Beat processing is a critical component of how humans experience music and a core topic in Music Information Retrieval (MIR). In this study we investigate relationships between computational features and human perception using real-world music spanning a variety of rhythmic categories. We observe a significant relationship between pulse clarity computed from the audio and inter-subject correlation of the neural responses. This finding suggests promising future lines of investigation integrating MIR audio features, neural correlation, and subjective behavioral reports in understanding beat processing and representation.

1. INTRODUCTION

Beat tracking is a central open topic in Music Information Retrieval (MIR) research. A better understanding of how humans process and track beat and rhythm has the potential to inform evaluation of rhythmic complexity and guide extraction of relevant audio features. One way this goal can be approached is by investigating relationships between computational representations of audio and cognitive correlates of rhythm.

A core MIR feature that can be extracted computationally from audio is *pulse clarity*, which aims to quantify strength of the underlying pulse in a musical excerpt [1]. On the cognitive side, neural responses—such as those measured with electroencephalography (EEG)—can provide objective measures of perception. While the study of EEG responses to real-world stimuli is an open challenge, recent approaches have made ecologically valid listening settings more feasible. One such approach involves the *inter-subject correlation* (ISC), which is thought to reflect brain states of attention and engagement [2]. Results from a recent study suggest that heightened EEG-ISC during music listening might be attributed to attention and expectation formulation, irrespective of enjoyment [3].

Leveraging these two measures, we follow up on a pilot study and dataset containing listeners' EEG responses to real-world musical excerpts of varying rhythmic complexity [4, 5]. We compute pulse clarity and EEG-ISC for each stimulus, and then correlate the two measures in order to determine whether a relationship exists between them.

2. METHODS

2.1 Dataset

For the present analysis, we used the Naturalistic Music EEG Dataset–Rhythm Pilot (NMED-RP) [4]. Stimuli for this study were twelve 30-sec audio excerpts derived from Bollywood music, the primary form of popular music in the Indian subcontinent. The excerpts span four evenly distributed rhythmic categories of different complexities: Even, Polyrhythm, Swing, and Syncopated. Each rhythmic category included three real-world excerpts plus one characteristic 'plain' rhythm devoid of any vocals or melodic instrumentation. 128-channel EEG was recorded from five participants, for a total of 27–28 trials recorded across participants for each stimulus. Participants listened attentively to the stimuli during each trial. After each trial, the participant rated the enjoyment, perceived complexity, and ease of finding the underlying pulse of the stimulus.

2.2 EEG and Audio Analysis

We focus on the audio and EEG responses for the present analysis. For each audio excerpt, we calculated pulse clarity using the Matlab MIR Toolbox [6]. This resulted in one pulse clarity measure per stimulus. For the EEG data, we first filtered and cleaned the data, downsampling the data from 1 kHz to 125 Hz and using Independent Components Analysis to remove ocular and EKG artifacts. After preprocessing, one electrode-by-time-by-trial matrix was produced for each stimulus. To compute ISC, we followed an established procedure whereby the multichannel EEG data were first spatially filtered to maximize ISC (i.e., reliability) while also reducing data dimensionality [2]. We applied a precomputed spatial filter that had been trained over a larger collection of EEG responses to Bollywood music [3] and retained the temporal activations from the first, maximally reliable component. ISC was computed using a one-against-all procedure: Trial-level ISC was the mean correlation of that trial with every other trial. Finally, we averaged ISC across trials to obtain a single value for each stimulus. These ISC values were correlated with the corresponding pulse clarity measures.

3. RESULTS AND DISCUSSION

In this study, we sought to determine whether a relationship exists between a computational measure of musical pulse (pulse clarity) and an objective measure of perception derived from EEG responses (ISC). The correlation between pulse clarity and ISC is visualized in Figure 1. Interestingly, in addition to being statistically significant ($r = -0.8$, $p < 0.001$), the correlation is also negative. This means that stimuli with lower pulse clarity evoked higher ISC in the EEG responses. Three of the four 'plain' stimuli, denoting rhythm exemplars not embedded in natural music, are among the stimuli with the highest pulse clarity and lowest ISC.

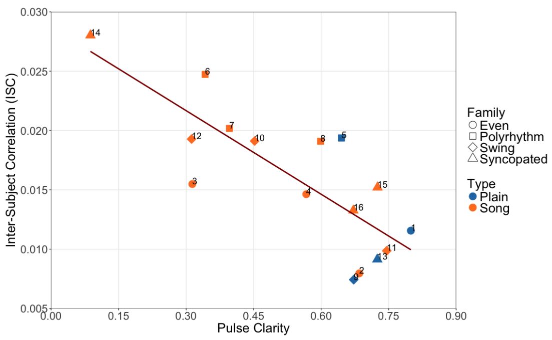

Figure 1. Correlation between pulse clarity and EEG-ISC ($r = -0.8$, $p < 0.001$). Shape denotes rhythm category; color denotes stimulus type.

In interpreting this inverse relationship between pulse clarity and ISC, we surmise that higher ISC might be driven by heightened attention when listening to stimuli with more complex beat structures. Previously reported behavioral results from this dataset [5] indicated that enjoyment and perceived complexity was low—while perceived ease of finding the beat was high—for the 'plain' stimuli relative to the real-world excerpts. Thus, a significant relationship may also exist between subjective reports and ISC/pulse clarity measures. Investigating this further will be an interesting avenue for future research. Our future work will also involve additional neural correlation metrics that have proven insightful in recent EEG studies of natural music listening, such as the extent to which EEG tracks time-varying stimulus features [3, 7].

4. CONCLUSION

MIR research will benefit from multimodal approaches that increase our understanding of the relationship between human and computational measures of perception. New methodologies for analyzing EEG permit the study of music processing under more naturalistic listening scenarios and offer new objective measures of perception. Our current results using one such approach show that significant relationships can be drawn between neural correlation and a computational measure of metrical pulse, enriching our understanding of beat processing.

5. REFERENCES

[1] O. Lartillot, T. Eerola, P. Toiviainen, and J. Fornari, "Multi-feature modeling of pulse clarity: Design, validation and optimization." in *ISMIR*, 2008, pp. 521–526.

[2] J. P. Dmochowski, P. Sajda, J. Dias, and L. C. Parra, "Correlated components of ongoing EEG point to emotionally laden attention—a possible marker of engagement?" *Frontiers in Human Neuroscience*, vol. 6, p. 112, 2012.

[3] B. Kaneshiro, D. T. Nguyen, A. M. Norcia, J. P. Dmochowski, and J. Berger, "Natural music evokes correlated EEG responses reflecting temporal structure and beat," *NeuroImage*, vol. 214, p. 116559, 2020.

[4] J. Appaji and B. Kaneshiro, "Naturalistic music EEG dataset—Rhythm Pilot (NMED-RP)," in *Stanford Digital Repository*, 2016. [Online]. Available: https://purl.stanford.edu/rz763kn3821

[5] ——, "Neural tracking of simple and complex rhythms: Pilot study and dataset," *Extended Abstracts for the Late-Breaking Demo Session of ISMIR*, 2018.

[6] O. Lartillot and P. Toiviainen, "A Matlab toolbox for musical feature extraction from audio," in *DAFx*, 2007, pp. 237–244.

[7] N. Gang, B. Kaneshiro, J. Berger, and J. P. Dmochowski, "Decoding neurally relevant musical features using Canonical Correlation Analysis," in *ISMIR*, 2017, pp. 131–138.