Stronger correlation of music features with brain signals predicts increased levels of enjoyment

Pandey, Pankaj; Bedmutha, Poorva Satish; Miyapuram, Krishna Prasad; Lomas, Derek

DOI
10.1109/APSCON56343.2023.10101229

Publication date
2023

Document Version
Final published version

Published in
APSCON 2023 - IEEE Applied Sensing Conference, Symposium Proceedings

Stronger correlation of music features with brain signals predicts increased levels of enjoyment

Pankaj Pandey *
IIT Gandhinagar
India

Poorva Satish Bedmutha*
UCSD
USA

Krishna Prasad Miyapuram
IIT Gandhinagar
India

Derek Lomas
TU Delft
Netherland

* These authors contributed equally

2023 IEEE Applied Sensing Conference (APSCON) | 978-1-6654-6163-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/APSCON56343.2023.10101229

***Abstract*—Music recommendation systems struggle with predicting the aesthetic responses of listeners based solely on acoustic characteristics, which are dependent on the listener's perception. This research correlates acoustic music features with brain responses to report the neural aesthetic hypothesis that the intensity of an aesthetic experience can be decoded based on the degree of correlation to brain responses. We employ hybrid encoding-decoding model (Canonical Correlation Analysis) to identify music features that maximally covary with brain responses. EEG signals of 20 participants are analyzed while they listen to 12 songs and mark their enjoyment on a scale of 1 to 5. Firstly, 18 acoustic features are extracted from music signals and transformed into the first principal component (PC1). In addition, two other features used for analysis are root mean square (RMS) and Spectral Flux (Flux). The first principal canonical component (CC1) with PC1 determines significant ($p<0.05$) evidence of correlating with brain responses that increasing correlation reflects increased enjoyment. We consider each participant's average CC1 values and enjoyment rating over all 12 songs, followed by plotting a correlation graph to decode the relationship. We observe a significant ($p<0.05$) positive linear correlation with increasing CC1 scores of PC1 features against increased enjoyment rating. PC1 shows the maximum Pearson correlation ($r = 0.48$, $p = 0.03$). In addition, we segregate the brain responses based on low (1,2) and high (3,4) enjoyment ratings and find that higher CC1 values correspond to brain responses of high enjoyment and low values to low enjoyment in all three features. Our experiments reveal that Canonical correlation reflects music-induced pleasure and can be employed in EEG-enabled headphones to decode the user experience, leading to better recommendations.**

***Index Terms*—EEG, Aesthetic, Music, CCA**

I. Introduction

Contemporary music recommendation systems are prone to several biases and solely present suggestions based on the acoustic characteristics of rhythm, melody, genre, and many more. However, recent advances in neurosensory technology and wearable Electroencephalography (EEG) sensors have enabled the recording of real-time brain signals while listening to music. This possibility embarks on identifying the computational neural marker for the assessment of the aesthetic response of an individual.

Recent years have shown a gradual increase in studies decoding musical appraisal from brain activity. Brain activity is primarily decomposed into five brain waves, and slower theta rhythms evoked in the frontal cortical brain region were found to discriminate between pleasant and unpleasant music [1]. Lin et al. reported the classification of emotional states in music listening with an accuracy of 82.29% [2]. Recently few studies incorporating techniques of time-frequency representation and dynamic functional connectivity measures support evidence of classification of liked and disliked music [3], [4]. The recent work by Gulshan et al. demonstrated enjoyment score incorporated while song recognition enhanced the classification performance [5].

Previous studies were solely based on brain responses, and our presented article involves acoustic features extracted from music signals and correlating with brain responses. This research is based on the neural aesthetic hypothesis that individuals with a stronger positive aesthetic response to music will exhibit a stronger relationship with the music features, leading to resonance (synchronization) in brain areas [6].

We employ the hybrid encoding-decoding technique, i.e., Canonical Correlation Analysis (CCA), to determine the relationship between audio stimuli and brain responses. CCA has been used previously to correlate music rhythms to brain responses [7].

II. Dataset Description

EEG activity to music was collected from 20 participants listening to 12 naturalistic music stimuli. Data were recorded using Geodesic 128 EEG sensors with 1000 Hz sampling rate. After each song, a double beep sound was played to indicate that the participants should open their eyes and rank their enjoyment level on a scale of 1 (low) to 5 (high). Each participant recording had 12 segments corresponding to 12 songs. Raw data were preprocessed using EEGLAB software, and high pass filtering of 0.2 Hz was applied, followed by a clean line method to remove the 50 Hz line noise. EEG data were downsampled to 256 Hz and removed bad channels based on spectrum criteria of 3 deviations. Machine learning-based Multiple Artifact Rejection Algorithm (MARA) was applied to remove the artifact components. Finally, we interpolated the channels using the spherical function of EEGLAB and followed by an average reference. Dataset and preprocessing script are publicly available at openneuro platform [8].

We thank to SERB and PlayPowerLabs for PMRF to PP. We thank FICCI for facilitating the PMRF.

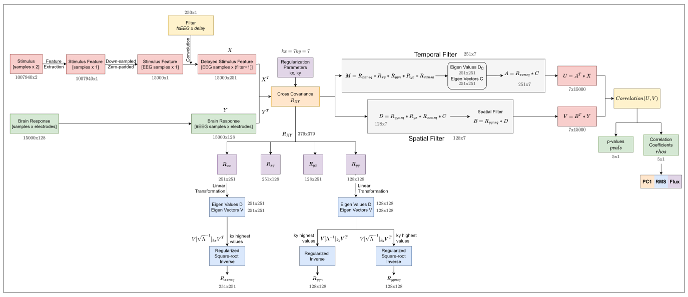

Fig. 1. The complete flow of Canonical Correlation Analysis with the mathematical equation is displayed. On the left, audio features (PC1, RMS, Flux) and brain responses are passed to compute the correlation between the two and report the first canonical correlation (CC1). (High resolution image available at [9])

III. Methodology

A. Stimulus Feature Extraction

Using the Music Information Toolbox (MIR), we extracted several acoustical features from the stimuli including zero crossing rate, spectral entropy, spectral centroid, spectral spread, high to low energy ratio, spectral roll-off, spectral flatness, roughness, root mean square energy (RMSE), spectral flux for eight octave-wide sub-bands, and broadband spectral flux. These features are similar to the set of 18 short-term features used by Alluri et. al. [10]. All the above features were retrieved with 50% overlap in 25-msec analysis windows between frames, resulting in an 80 Hz feature sampling frequency. We orthogonalized the features using principal component analysis, providing a lower dimensional stimulus representation, i.e., the first principal component (PC1) that encoded all features in direction of maximal variance. PC1 was used in all subsequent analyses, with two individual features comprising RMS indicating amplitude envelope and spectral flux (Flux) representing timbre. RMS and Flux of music stimulus features are widely used in literature to correlate brain activity with music stimulus features.

B. Canonical Correlation Analysis

Stimulus-Response correlation is the emerging endeavor of neuroscience to understand the fundamental link between sensory stimulus and their neural responses, and one rapidly growing technique is Canonical Correlation Analysis [11]. CCA is a hybrid encoding-decoding approach that decomposes neural activity into several components and reflects a relationship with the stimulus. CC1 is the strongest component observed in the previous studies [7]. Previous studies have utilized this to correlate auditory and visual features to brain responses. We performed Canonical Correlation by passing PC1, RMS, and Flux features with brain responses. In Fig. 1, we have provided a flow of computation, where X represents stimulus features with a 2-sec delay to brain responses Y. The code implementation used publicly available at [12].

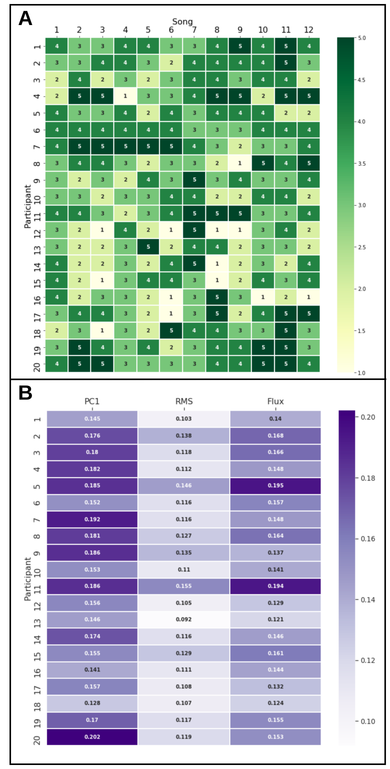

Fig. 2. [A] Enjoyment Ratings of every participant. [B] Average CC1 of PC1, RMS and Flux features with respective enjoyment ratings of 12 songs for each participant

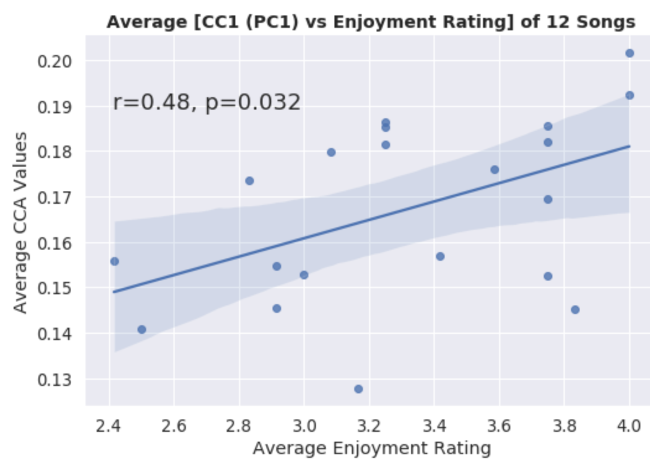

Fig. 3. CC1 (PC1) vs Enjoyment Ratings on average of 12 songs, and each data point represents a participant.

IV. Experimental Result

Each participant's brain responses were recorded while listening to 12 songs. Brain responses were extracted against each song and computed the Canonical Correlation with features extracted from the music signal. The three features correlated to brain responses were the first principal component of 18 audio features (PC1), root mean square (RMS), and spectral Flux (Flux). We reported the values of the first

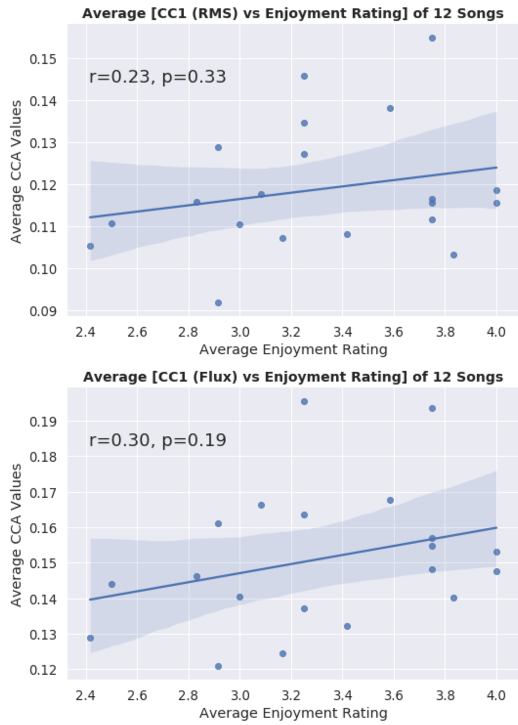

Fig. 4. The top and bottom plots represent RMS & Flux for CC1 (RMS) vs Enjoyment Ratings on an average of 12 songs. Each data point represents a participant.

Canonical Component (CC1) ($p < 0.05$). In Fig. 2 (A), we presented the behavioral ratings provided by participants.

A. Increasing correlation of music features to brain responses reflects the increasing level of enjoyment

We observed the cumulative effect of 12 songs listened by each participant. Therefore we averaged the CC1 values of all songs respective to every participant and presented the Table 2 (B). We presented a scatter plot for each feature, as shown in Fig. 3,4. The plot indicates a relationship of averaged enjoyment with CC1 averaged across all 12 songs for every participant. We computed the Pearson correlation coefficient between average CC1 values and enjoyment ratings for all participants and observed a maximum correlation (r) of 0.48 with PC1 with a significance (p-value) of 0.03. In contrast, we did not obtain significance in the other two features. PC1 was high and showed an increasing CC1 value with an increasing level of enjoyment rated by participants.

B. Low vs High

We divided the enjoyment ratings based on low (1,2) and high (4,5), including all songs, participants and corresponding CC1 values segregated and plotted for PC1, Flux, and RMS in Fig. 5. The high enjoyment indicates higher CC1 values in all three features than low enjoyment. We performed a two-sample t-test with equal variance and observed the one-tailed significance ($p<0.05$) across three features.

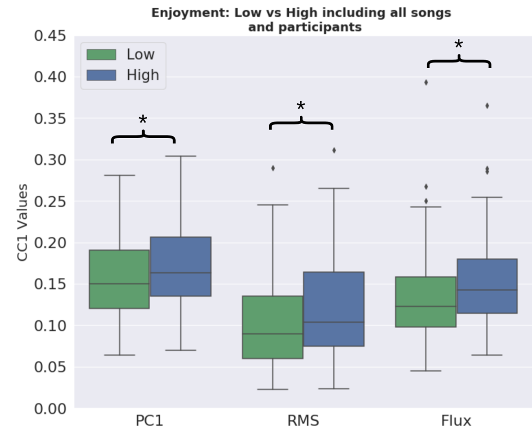

Fig. 5. Total responses are segregated on low and high enjoyment and (*) indicates significant difference between two states.

V. Conclusion

Our findings suggest that Canonical Correlation between acoustic features and brain responses could serve as real-time metric for decoding the aesthetic experience.

References

[1] D. Sammler, M. Grigutsch, T. Fritz, and S. Koelsch, "Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music," *Psychophysiology*, vol. 44, no. 2, pp. 293–304, 2007.

[2] Y.-P. Lin, C.-H. Wang, T.-P. Jung, T.-L. Wu, S.-K. Jeng, J.-R. Duann, and J.-H. Chen, "Eeg-based emotion recognition in music listening," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 7, pp. 1798–1806, 2010.

[3] S. K. Hadjidimitriou and L. J. Hadjileontiadis, "Toward an eeg-based recognition of music liking using time-frequency analysis," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 12, pp. 3498–3510, 2012.

[4] S. Bakas, D. A. Adamos, and N. Laskaris, "On the estimate of music appraisal from surface eeg: a dynamic-network approach based on cross-sensor pac measurements," *Journal of Neural Engineering*, vol. 18, no. 4, p. 046073, 2021.

[5] G. Sharma, P. Pandey, R. Subramanian, K. Miyapuram, A. Dhall *et al.*, "Neural encoding of songs is modulated by their enjoyment," *arXiv preprint arXiv:2208.06679*, 2022.

[6] P. Lawhatre, B. R. Shiraguppi, E. Sharma, K. P. Miyapuram, and D. Lomas, "Classifying songs with eeg," *arXiv preprint arXiv:2010.04087*, 2020.

[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.

[8] K. P. Miyapuram, P. Pandey, N. Ahmad, B. R. Shiraguppi, and e. a. Esha Sharma, ""music eeg dataset"," 2022. [Online]. Available: https://openneuro.org/datasets/ds003774/versions/1.0.2

[9] P. Pandey, "Cca_image." [Online]. Available: https://github.com/Pandey-Pankaj/Resonance-Toolbox/blob/main/stimulus-response-correlation/CCA.png

[10] V. Alluri, P. Toiviainen, I. P. Jääskeläinen, E. Glerean, M. Sams, and E. Brattico, "Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm," *Neuroimage*, vol. 59, no. 4, pp. 3677–3689, 2012.

[11] J. P. Dmochowski, J. J. Ki, P. DeGuzman, P. Sajda, and L. C. Parra, "Extracting multidimensional stimulus-response correlations using hybrid encoding-decoding of neural activity," *NeuroImage*, vol. 180, pp. 134–146, 2018.

[12] Dmochow, "Matlab code to compute stimulus-response correlation." [Online]. Available: https://github.com/dmochow/SRC