

PADRÕES DE ATIVAÇÃO CORTICAIS GERADOS DURANTE A AUDIÇÃO DE MÚSICAS BRASILEIRAS

Leonardo Antunes Ferreira¹, Estela Ribeiro, Carlos Eduardo Thomaz
Departamento de Engenharia Elétrica, Centro Universitário FEI
leonardoantunes123@gmail.com, cet@fei.edu.br

Resumo: A música nos permite investigar a formação estrutural e funcional do cérebro humano. Nesse trabalho, a combinação de um método de extração de características acústicas com a análise de sinais de eletroencefalograma de voluntários permitiu revelar diferenças nos padrões de ativação corticais entre músicos e não-músicos. Como estímulos, foram selecionadas duas músicas brasileiras reconhecidas internacionalmente.

1. Introdução

Nas últimas décadas houve um aumento significativo no número de estudos científicos que identificaram diferenças cognitivas entre músicos e não-músicos [1, 5, 6]. Trabalhos mais recentes vêm utilizando músicas completas para indicar padrões cognitivos relacionados ao processamento musical humano.

Todos esses trabalhos são baseados na extração de características acústicas diretamente do sinal de áudio para identificação dos *triggers*, que são momentos correspondentes a instantes no tempo da música em que há um alto contraste em uma característica acústica que pode estar relacionado com respostas cognitivas [5], ou seja, Potenciais Relacionados ao Evento (ERPs). ERPs são sinais elétricos gerados nas estruturas cerebrais em resposta a eventos específicos ou estímulos [4]. Os sinais obtidos de um ERP consistem em uma sequência de deflexões de tensão positiva e negativa ao longo do tempo, que são chamadas de componentes. A maioria dos componentes é referenciada por uma letra (N/P) indicando polaridade (negativa/positiva), seguida por um número indicando a latência em milissegundos, por exemplo, N100 para um pico negativo em 100 ms [4].

No entanto, não há na literatura afim estudos sobre músicas brasileiras. Portanto, um dos objetivos deste trabalho foi utilizar a metodologia científica que segue as diretrizes propostas por [5], nas quais há o uso de músicas completas como estímulos, extração de características acústicas, seleção de *triggers* e por último, análise dos sinais de EEG, com músicas brasileiras. Adicionalmente, foram realizados experimentos com um grupo de voluntários, por meio da análise de sinais de EEG, para investigar uma possível diferença cortical entre músicos e não-músicos.

2. Metodologia

A metodologia adotada segue 5 etapas: (I) Extração de características acústicas; (II) Seleção de características acústicas; (III) Seleção de *triggers*; (IV) Processamento dos sinais de EEG; (V) Análise dos sinais de EEG de músicos e não-músicos.

Na primeira etapa, são extraídas do sinal de áudio 12 características acústicas, já bem consolidadas na

literatura, com a MIRtoolbox (Versão 1.71) [3] e Matlab 2015a, decompondo o sinal de áudio em janelas de 50 ms sobrepostas em 50%, sendo essas: (1) Root Mean Square Energy (RMS), (2) Zero Crossing Rate (ZCR), (3) Spectral Rolloff, (4) Spectral Roughness, (5) Brightness, (6) Spectral Entropy, (7) Spectral Flatness, (8) Spectral Skewness, (9) Spectral Kurtosis, (10) Spectral Centroid, (11) Spectral Spread e (12) Spectral Flux.

Na segunda etapa, é determinado se todas as 12 características acústicas são estatisticamente relevantes por meio de uma análise de agrupamentos utilizando *Factor Analysis*, ou Análise Fatorial (FA), com rotação *varimax* [2] e o algoritmo *K-means*. O número de fatores usados pelo FA é igual ao número de componentes principais extraídas dos dados que possuem autovalor maior que 1. As características acústicas são selecionadas pela sua representatividade em cada agrupamento (*cluster*), representada como aquela que possui maior carga fatorial em cada fator. Vale ressaltar que a matriz de dados é normalizada com média 0 e variância 1 já que cada característica possui uma ordem de grandeza distinta.

Durante a terceira etapa, são selecionados e marcados nos instantes de tempo da música os *triggers* das características mais relevantes do conjunto seguindo a metodologia de referência científica descrita em [5].

A quarta etapa envolve o processamento dos dados adquiridos de EEG com as *toolboxes* EEGLAB (Versão 14.1.2) e ERPLAB (Versão 7.0.0). Todos os sinais adquiridos são filtrados por meio de um filtro digital *Butterworth* de ordem 10 Passa Banda de 1 Hz a 30 Hz elaborado no Matlab, para remoção do nível DC e do sinal da rede elétrica. Adicionam-se então aos sinais os *triggers* selecionados que correspondem aos eventos a serem analisados e então delimita-se o sinal em épocas de 1000 ms, sendo 200 ms pré-estímulo e 800 ms pós-estímulo. O sinal de EEG foi filtrado utilizando-se o método *Moving Window peak-to-peak Threshold*, presente no ERPLAB, para remoção de artefatos. Como o piscar de olhos e movimentos do voluntário.

Por último, fez-se o uso de duas abordagens para a análise dos dados de EEG obtidos classificando músicos e não-músicos. A primeira consiste em verificar graficamente através dos sinais de ERPs a ocorrência dos componentes N100-P200. A segunda abordagem tem como objetivo classificar os sinais utilizando o algoritmo MLDA (*Maximum Uncertainty Linear Discriminant Analysis*) [7] do qual foi obtido os sinais médios para músicos e não-músicos.

3. Resultados

Para este trabalho, foram selecionadas 2 músicas para a análise: Garota de Ipanema em 2 versões, uma

executada por Tom Jobim (GI1) e a outra uma versão instrumental por Zimbo Trio (GI2); e Bachianas Brasileiras N°2 - O Trenzinho do Caipira de Heitor Villa Lobos (BB). Todas as músicas foram obtidas por meio do Spotify, um serviço de *streaming* de música *online*, e gravadas com o software de áudio Audacity (Versão 2.2.2) com taxa de amostragem de 44,1 kHz, áudio mono e em formato WAV (*Waveform Audio Format*).

Para todas as músicas usando somente 3 componentes principais é possível explicar mais de 80% da variância dos dados. Então o número de fatores utilizado no FA é igual a 3. As cargas fatoriais obtidas pelo FA revelam a correlação entre as características acústicas e os *clusters* apresentados na Figura 1 foram obtidos com o algoritmo *K-means*.

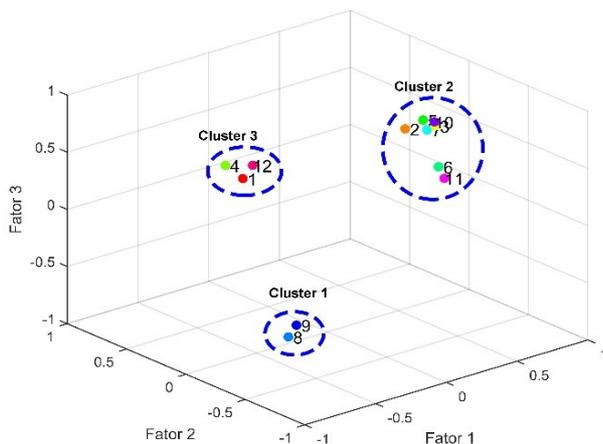


Figura 1 – Exemplo dos agrupamentos das características acústicas para a música Garota de Ipanema de Zimbo Trio (GI2).

Os resultados obtidos pelo FA revelaram-se praticamente iguais para as 3 músicas analisadas, com o agrupamento das mesmas características em 3 *clusters*. É interessante notar, no entanto, que para cada música, em cada *cluster*, houve uma característica mais relevante. Para cada uma das músicas (GI1, GI2, BB) há representatividades diferentes, como pode ser visto a seguir [cluster1, cluster2, cluster3]: GI1 [9, 10, 1]; GI2 [8, 5, 1]; BB [8, 10, 1].

A análise dos sinais de ERPs foi realizada na região central do escalpo. Para os 2 grupos (3 músicos e 2 não-músicos) obteve-se o sinal médio para cada característica acústica selecionada. Os resultados mostraram que existe um comportamento diferente entre os sinais médios de ERPs para os dois grupos de estudo. Notou-se também que as componentes detectadas mostraram-se predominantes no grupo de músicos e nem todas características apresentaram ERPs.

Os resultados obtidos pelo MLDA (Figura 2) foram analisados por meio de mapas topográficos, que oferecem uma maneira visual e discriminante para identificar diferenças ao contrário dos sinais de ERPs. Analogamente à análise de ERPs, nem todas as características causaram as ativações esperadas, mas mesmo em baixo número tais ativações se mostraram mais presentes no grupo de músicos.

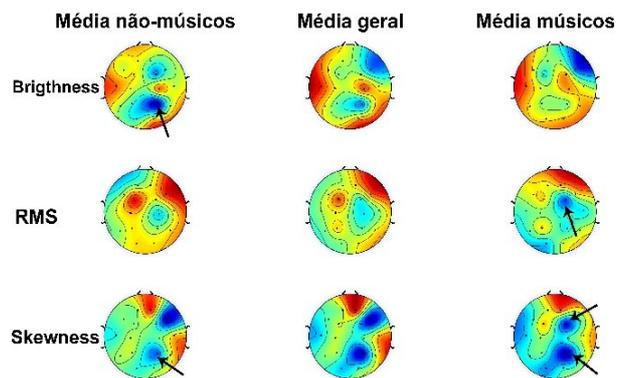


Figura 2 – Exemplo dos mapas topográficos para a música Garota de Ipanema de Zimbo Trio (GI2). As setas indicam o componente N100.

4. Conclusão

Nesse trabalho foi mostrado que é possível representar estatisticamente 12 características acústicas em 3 *clusters* por meio dos algoritmos FA e *K-means*. As análises dos sinais de EEG pelo método de ERPs e o algoritmo MLDA apresentaram resultados que comprovaram diferenças, entre músicos e não-músicos principalmente pela presença do componente N100 em grande maioria nos sinais cerebrais dos músicos.

5. Referências

- [1] ALLURI, V. et al. Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *NeuroImage*, v. 59, n. 4, p. 3677–3689, fev. 2012.
- [2] JOHNSON, R. A.; WICHERN, D. W. **Applied Multivariate Statistical Analysis**. 6th. ed. New Jersey: Pearson, 2007.
- [3] LARTILLOT, O. MIRtoolbox 1.6 users manual. **Aalborg: Department of architecture, design e media technology**, nov. 2014.
- [4] LUCK, S. J. **An introduction to the event-related potential technique**. 2nd ed. Massachusetts: MIT press, 2014.
- [5] POIKONEN, H. et al. Event-related brain responses while listening to entire pieces of music. *Neuroscience*, v. 312, p. 58–73, jan. 2016.
- [6] SAARI, P. et al. Decoding Musical Training from Dynamic Processing of Musical Features in the Brain. *Scientific Reports*, v. 8, n. 1, dez. 2018.
- [7] THOMAZ, C. E.; KITANI, E. C.; GILLIES, D. F. A maximum uncertainty LDA- based approach for limited sample size problems – with application to face recognition. *Journal of the Brazilian Computer Society*, Springer, v. 12, n. 2, p. 7–18, 2006.

Agradecimentos

À FEI pela concessão da bolsa, e ao meu orientador e coorientadora pelo conhecimento que adquiri sob a orientação dos mesmos.

¹ Aluno de IC da FEI. Projeto com vigência de 06/18 a 07/19.