

AGRUPAMENTO ESTATÍSTICO DE CARACTERÍSTICA ACÚSTICA E TRIGGERS

Vinícius Santos Barbosa, Estela Ribeiro, Carlos Eduardo Thomaz
 Engenharia Elétrica, Centro Universitário FEI
unievbarbosa@fei.edu.br, cet@fei.edu.br

Resumo: Existem instantes nas músicas capazes de gerar respostas neurais significativas (*triggers* ou gatilhos), importantes para se analisar o processamento musical. Neste trabalho, analisa-se a influência da quantidade de *triggers* ao comparar sinais neurais decorrentes do uso de n° gatilhos identificados em um conjunto de características acústicas agrupadas por meio de análise de cluster com o uso apenas das características acústicas dominantes. Resultados demonstram que utilizar todos os *triggers* que compõem um agrupamento é uma oportunidade de melhorar a estimativa do sinal médio de EEG para cada um dos voluntários já que as características agrupadas são estatisticamente semelhantes

1. Introdução

Tocar um instrumento musical é uma tarefa que altera o cérebro de tal forma que é possível dizer se uma pessoa é um músico ou não por meio dos sinais neurais registrados durante a escuta musical (RIBEIRO; THOMAZ, 2019). É possível fazer a discriminação desses grupos ao se utilizar ferramentas de estatística e aprendizado de máquina. Para isso, é necessária uma melhor representação dos dados adquiridos para realizar a classificação.

Pesquisas recentes demonstram que é possível identificar, por meio da análise de características acústicas de baixo nível extraídas de sinais de áudio, instantes na música capazes de gerar respostas neurais significativas, denominados *triggers* (POIKONEN, 2016). Essas pesquisas descrevem a associação entre as mudanças dinâmicas nas características dos áudios e nas ativações neurais registradas em voluntários utilizando eletroencefalografia (EEG) (RIBEIRO; THOMAZ, 2019). Além disso, tem se demonstrado estatisticamente que é possível reduzir o número de características acústicas necessárias para investigar os padrões de ativações corticais gerados durante a escuta musical (FERREIRA; RIBEIRO; THOMAZ, 2019).

Este trabalho busca comparar os sinais neurais resultantes de dois cenários: (I) Considerando apenas os *triggers* encontrados nas características acústicas dominantes de cada agrupamento encontrado; (II) Considerando os *triggers* encontrados em todas as características acústicas de cada agrupamento. Desta forma, espera-se que haja uma melhor representação dos dados no segundo cenário para uma análise posterior de discriminação de voluntários músicos e não-músicos.

2. Metodologia

Neste trabalho foram utilizados os dados de EEG adquiridos no trabalho de Ribeiro e Thomaz (2019). Neste, foi apresentada a música Dança Húngara n.5 de Johannes Brahms a 26 voluntários (13 músicos e 13 não-músicos) que deveriam ouvir a música enquanto seus

sinais de EEG eram registrados. A metodologia utilizada (RIBEIRO; THOMAZ, 2019; FERREIRA; RIBEIRO; THOMAZ, 2019) é dividida em quatro etapas, conforme apresentado na Figura 1, sendo as três primeiras relacionadas ao processamento do sinal de áudio e a última relacionada ao processamento do sinal de EEG registrados dos voluntários, sendo essas: (I) Extração de características acústicas; (II) Seleção das características acústicas; (III) Seleção de Triggers; (IV) Processamento dos sinais de EEG. Na primeira etapa, são extraídas as características acústicas já empregadas em outros trabalhos (ALLURI et al., 2012, RIBEIRO, THOMAZ, 2019), sendo estas: *Root Mean Square (RMS) Energy*, *Zero Crossing Rate*, *Brightness*, *Spectral (S.) Centroid*, *S. Spread*, *S. Flatness*, *S. Skewness*, *S. Kurtosis*, *S. Rolloff*, *S. Flux*, *S. Roughness* e *S. Entropy*. A definição de cada característica pode ser consultada em Lartillot (2014). O valor de janelamento para extração das características acústicas foi fixado em 50 ms e sobreposição de 25 ms, ou 50% do tempo de janelamento, já o comprimento do sinal de áudio é de 192 s.

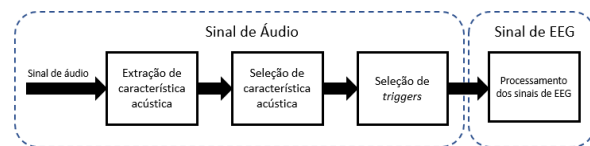


Figura 1 - Metodologia empregada

A seguir, realiza-se a seleção das características acústicas agrupando-as em uma matriz de dados M_{JxC} , na qual cada coluna representa uma característica acústica e as linhas os valores encontrados para cada janela de tempo. Normaliza-se M_{JxC} com média 0 e variância 1, e a seguir aplica-se a a Análise Fatorial (FA), como uma ferramenta de análise de agrupamentos, com o objetivo de definir quais das doze características acústicas mencionadas são estatisticamente mais relevantes para o estudo e agrupá-las de acordo com a similaridade estatística da informação contida nelas.

Na terceira etapa são selecionados os *triggers*. Dois parâmetros que caracterizam (Figura 2) um *trigger* são: Magnitude de rápido crescimento (MoRI) que define os limiares superior e inferior de amplitude do sinal, e Preceding Low-Feature Phase (PLFP) que é o tempo de permanência da característica abaixo do limiar inferior de amplitude e antecede a MoRI. O processamento dos sinais de EEG é feito utilizando os instantes determinados pelos *triggers* como estímulos no EEG. É feito um janelamento no sinal de EEG determinado por um intervalo de 200 ms que antecede o estímulo e outro de 800 ms que o sucede, formando sucessivas épocas de 1000 ms. Desta forma, são analisados dois cenários,

sendo que no primeiro analisa-se os sinais de EEG considerando apenas os *triggers* encontrados nas características acústicas dominantes de cada agrupamento, e no segundo considerando os *triggers* encontrados em todas as características acústicas de cada agrupamento.

3. Resultados

Demonstrou-se que a variação do intervalo de tempo de permanência da característica abaixo do limiar inferior (PLFP) influencia na quantidade de *Triggers* presentes nos fatores resultantes da Análise Fatorial. Observou-se um comportamento exponencialmente decrescente na quantidade de *triggers* em função do intervalo de PLFP. A Figura 2 demonstra o comportamento utilizando a característica acústica *S. Centroid* como exemplo.

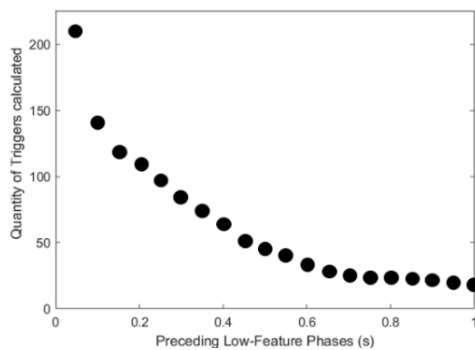


Figura 2 - Quantidade de triggers em função de PLFP para *S. Centroid*

De acordo com os resultados obtidos da análise de agrupamentos, observa-se que as características acústicas para o sinal de áudio estudado podem ser categorizadas em três agrupamentos de maior similaridade estatística. Cada um destes agrupamentos possui uma característica acústica com maior carga fatorial capaz de explicar a variação dos dados ao longo do sinal. A quantidade de *triggers* para cada característica acústica pode ser levada em consideração ao se definir a característica dominante de um agrupamento. A Tabela 1 apresenta a quantidade de *triggers* encontrados utilizando apenas as características dominantes e a quantidade de *triggers* para todas as características nos agrupamentos.

Além disso, foi feita a verificação de como o sinal médio de EEG se comporta para cada um dos voluntários se considerarmos dois cenários. No primeiro cenário, utilizou-se todos os *triggers* não repetidos que compõem cada um dos clusters. No segundo cenário, utilizou-se os *triggers* da característica acústica dominante selecionada para o cluster, ou seja, os *triggers* de RMS, *S. Kurtosis* e *S. Rollof*.

Tabela 1 - Quantidade de triggers encontrados nos diferentes cenários propostos.

Cluster	Carac. dominante	Quant. triggers dominante	Quant. triggers agrupamento	Quant. carac. agrupadas
1	RMS	6	32	3
2	<i>S. Kurtosis</i>	8	14	2
3	<i>S. Rollof</i>	7	30	7

Esses resultados demonstram que utilizar todos os *triggers* que compõem um agrupamento é uma oportunidade de melhorar a estimativa do sinal médio de EEG para cada um dos voluntários já que as características agrupadas são estatisticamente semelhantes. Com isso, o sinal médio de EEG torna-se mais uniforme do que quando comparado ao sinal médio gerado utilizando apenas os *triggers* da característica dominante. A Figura 3 demonstra o sinal médio composto pelo EEG de voluntários músicos.

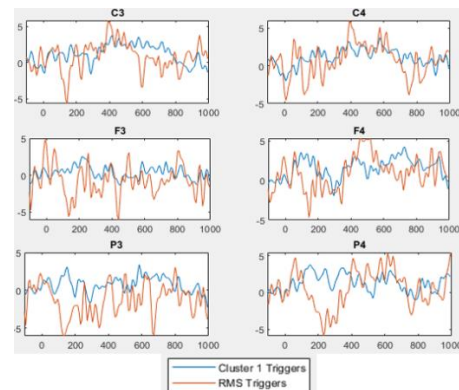


Figura 3- Sinal médio de EEG para Voluntários Músicos

4. Conclusões

Triggers de um mesmo agrupamento estatístico podem ser utilizados para compor o sinal médio de EEG durante a escuta musical, podendo melhorar a classificação dos sinais de EEG de músicos e não-músicos. Além disso, a curva exponencial inversa entre quantidades de *triggers* em função dos parâmetros para seleção de *triggers* demonstra que valores baixos de PLFP geram uma quantidade mais elevada de *triggers*. Porém, podem não ter o potencial necessário para gerar um alto-contraste perceptível nos sinais de EEG. Posteriormente pretende-se utilizar essas informações para realizar a discriminação de voluntários músicos e não-músicos, comparando os resultados obtidos com e sem o agrupamento das características acústicas, com a expectativa de que haja uma melhoria na classificação dos voluntários com o agrupamento.

5. Referências

- [1] ALLURI, V. et al. Large-scale brain networks emerge from dynamics processing of musical timbre, key and rhythm. *NeuroImage*, p. 3677 – 3689, 2012.
- [2] FERREIRA, L. A.; RIBEIRO, E.; THOMAZ, C. E. A Cluster Analysis of Benchmark Acoustic Features on Brazilian Music. In: *SBCM*. 2019.
- [3] LARTILLOT, O. *MIRtoolbox 1.6.1 User’s Manual*. 1. ed. [S.l.], 2014.
- [4] RIBEIRO, E.; THOMAZ, C. E. *Whole Brain EEG Analysis of Musicianship*. *Music Perception*, 2019.
- [5] POIKONEN, H. et al. *Event-Related Brain Responses While Listening to Entire pieces of Music*. *Neuroscience*, 2016.

¹ Aluno de IC CNPq. Projeto PBIC001/20 com vigência de 08/20 a 10/21. R.A 12.119.336-1