

# Um estudo sobre predição de musicalidade por meio da análise de sinais de EEG

Estela Ribeiro, Carlos Eduardo Thomaz

Departamento de Engenharia Elétrica – Centro Universitário FEI  
São Bernardo do Campo, São Paulo, Brasil  
e-mail: estela.eng@hotmail.com, cet@fei.edu.br

**Abstract** - *This work describes and implements the use of a computational model for the classification of the Electroencephalogram (EEG) signals recorded from a group of volunteers categorized as musicians and non-musicians. This model is based on the extraction and selection of audio features from the songs Hungarian Dance No.5 (Johannes Brahms) and The Barber of Seville - Overture (Gioachino Rossini), used as stimuli in such analysis. Experiments were carried out with 26 volunteers, 13 non-musicians and 13 musicians (all amateurs), who listened the songs while their EEG signals were recorded using dry electrodes. The non-parametric k-NN classifier was used to discriminate the samples and the model proposed was able to deliver an accuracy rate higher than 80% for both songs.*

**Palavras-chave:** EEG, k-NN, Music, Computational Model.

## Introdução

A música está em todos os lugares e saber como o cérebro processa essa informação acústica é uma questão que vem sendo estudada cada vez mais nos últimos tempos, pois a música “oferece uma oportunidade única de entender a organização do cérebro humano” [1]. Sabe-se que, devido a plasticidade cerebral, a forma como um músico processa a música é diferente de um não-músico, havendo ainda diferenças entre os músicos dependendo do instrumento ou estilo musical tocado [2].

Diversas pesquisas têm sido feitas sobre este tema, analisando as diferenças estruturais existentes no encéfalo de músicos e não-músicos [1, 2], e para a compreensão de como o cérebro processa a música, utilizando o fMRI [3, 4] ou EEG [5, 6].

Grande parte dos estudos tradicionais de Potencial Evocado Auditivo são realizados a partir de sons artificiais, que possuem intervalos de silêncio bem definidos, porém, recentemente as músicas têm ganhado um papel importante nestas análises, sendo possível encontrar os mesmos potenciais evocados das pesquisas tradicionais, mesmo que as músicas sejam caracterizadas por um fluxo con-

tínuo de informação, sem intervalos de silêncio bem definidos, havendo interação e sobreposição de notas [4, 6].

Neste trabalho, musicalidade será definida, não de maneira quantitativa, indicando o grau de aptidão, mas sim de uma maneira qualitativa, onde não-músicos são considerados como pessoas que não possuem conhecimento formal em música e não tocam nenhum instrumento musical e músicos são pessoas que possuem prática musical e conhecimento formal em música.

O objetivo deste trabalho é classificar os sinais de EEG obtido de um grupo de voluntários em músicos e não-músicos, a partir de características acústicas extraídas do áudio, tendo como referência a amplitude do sinal medido de EEG de todos os eletrodos, analisando, desta forma, as respostas neurais induzidas ao longo de todo o escalpo (limitado ao número de eletrodos).

## Metodologia

A metodologia está dividida em 6 etapas, conforme a Figura 1, que apresenta o fluxograma das etapas propostas e implementadas através da plataforma Matlab R2015a.

Na Etapa 1, são descritas e extraídas  $c$  características acústicas que descrevem o sinal de áudio selecionado, utilizando a MIRtoolbox (versão 1.6.1) [7]. Portanto, em vez de analisar o sinal de áudio como um todo, serão analisadas as características acústicas que o compõem e o descrevem quantitativamente. Estas características não necessariamente apresentam propriedades capazes de serem diretamente interpretadas a nível de percepção auditiva [8], porém as alterações que ocorrem nelas são capazes de gerar respostas neurais significativas [6].

Para reduzir a redundância destes dados, são selecionadas, na Etapa 2, por meio da Análise Fatorial (FA) [9], as  $p$  características acústicas mais significativas (maior carga fatorial), onde  $p \leq c$ .

Na Etapa 3 determinam-se os *triggers* [6], que são instantes no tempo da música onde há um alto-contraste nas  $p$  características acústicas sele-

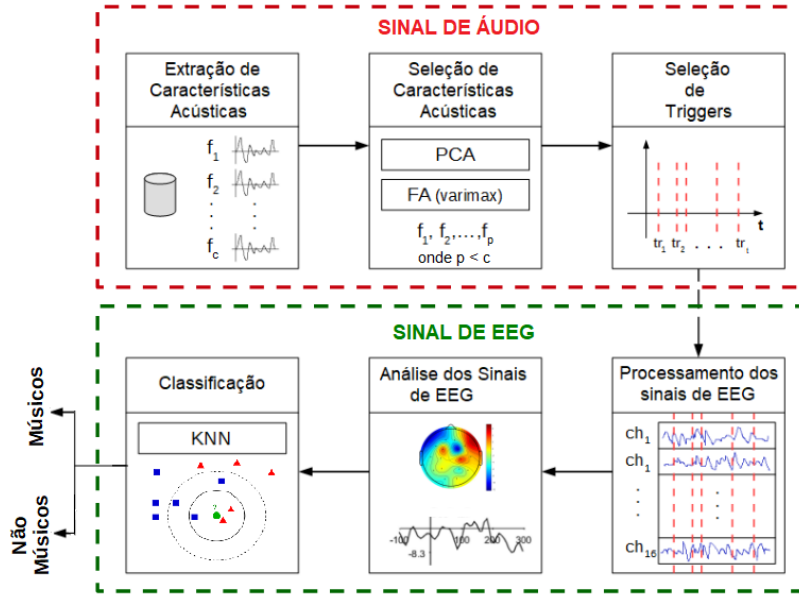


Figura 1 – Fluxograma da metodologia de classificação de sinais de EEG.

cionadas. Estes instantes são considerados como estímulos para o processamento do sinal de EEG na Etapa 4, onde os sinais são filtrados (1-30 Hz), demarcados por uma janela de tempo de 200 ms pré-estímulo, ou seja, 200 ms antes do instante em que os *triggers* ocorreram, até 800 ms pós-estímulo e, por fim, é realizada uma média dentro destas janelas, resultando em um único sinal de 1000 ms de duração, para cada característica acústica selecionada.

Deste sinal médio são gerados os mapas topográficos corticais dos grupos de voluntários, na Etapa 5, utilizando a toolbox ERPLab [10], onde é feita uma análise dos sinais apresentados pelos dois grupos e, por fim, na Etapa 6 é realizada a classificação dos sinais de EEG de cada voluntário, utilizando o classificador  $k$ -NN, onde são utilizadas as amplitudes medidas em todos os eletrodos nos instantes 100 ms e 200 ms pós-estímulo para cada característica acústica selecionada.

Cada voluntário será representado por um vetor de entrada composto pela concatenação dos valores registrados para todos os eletrodos  $Ch$ , para cada característica acústica selecionada  $f_p$ , para os dois instantes analisados. Os dados obtidos para cada voluntário são agrupados em uma matriz  $\mathbf{S}_{d \times 26}$  onde as 26 colunas se referem ao número de voluntários e  $d = f_p \times Ch \times 2$ .

## Resultados

Os experimentos foram realizados utilizando a plataforma OpenBCI, onde o sinal de EEG foi

gravado utilizando 16 eletrodos posicionados no escalpo de acordo com o Sistema 10-20, mais 2 eletrodos posicionados nos lóbulos de cada orelha como referência. Para o registro dos sinais foi utilizado o Ultracortex IV [11], com eletrodos a seco, a uma frequência de amostragem de 125 Hz.

Para a realização dos experimentos foram utilizados voluntários que não possuíam conhecimento musical, ou seja, que não tiveram nenhum treinamento musical formal e nunca aprenderam a tocar nenhum instrumento, que foram categorizados como *Não-Músicos* e pessoas com conhecimento musical, que possuíam no mínimo 2 anos de treinamento musical formal, que foram categorizadas como *Músicos*. Todos os participantes assinaram um Termo de Consentimento Livre Esclarecido para a realização dos experimentos.

Portanto, para o experimento foram considerados 26 voluntários (17 homens e 9 mulheres), 13 não-músicos ( $31,2 \pm 7,4$  anos) e 13 músicos ( $26,8 \pm 6,3$  anos) que praticam seus instrumentos há, em média, 14 anos.

Foram extraídas 12 características acústicas ( $c = 12$ ) do sinal de áudio, de acordo com a Tabela 1, utilizando um janelamento de 50 ms ( $l = 0,05s$ ) com fator de sobreposição  $h = 50\%$ . Tais parâmetros resultaram em uma matriz  $\mathbf{M}$  de dimensão  $24880 \times 12$ , a qual foi aplicado o FA utilizando 3 fatores, que representam mais de 85% da variância dos dados.

Para cada fator foram selecionadas para a análise as características acústicas que possuíam maior carga fatorial, sendo elas:

Tabela 1 – Características acústicas extraídas

Root Mean Square (RMS)
Zero Crossing Rate (ZCR)
Spectral Rolloff
Spectral Roughness
Brightness
Spectral Entropy
Spectral Flatness
Spectral Kurtosis
Spectral Centroid
Spectral Spread
Spectral Flux

1. *S. Rolloff*: Representa 89% de carga fatorial no F1, definido como a frequência abaixo da qual a porcentagem  $\beta$  (ex.:  $\beta = 85\%$ ) da magnitude da distribuição do espectro está concentrada;
2. *RMS*: Representa 96% de carga fatorial no F2, definido como uma medida da energia do sinal de áudio;
3. *S. Kurtosis*: Representa 95% de carga fatorial no F3, que mede se a forma da distribuição do valor da magnitude espectral é semelhante a distribuição Gaussiana ou não, podendo indicar transientes no sinal de áudio.

Utilizando os parâmetros propostos para a determinação dos *triggers* [6], foram encontrados para a música de Johannes Brahms 13 *triggers* para a característica RMS, 11 para o Spectral Rolloff e 7 para o Spectral Kurtosis, e para a música de Gioachino Rossini 14 *triggers* para a característica RMS, 16 para o Spectral Rolloff e 15 para o Spectral Kurtosis.

Para cada voluntário, portanto, é realizada a média dos sinais de EEG para cada *trigger* resultando em um único sinal médio para cada característica acústica selecionada, sendo obtidos 3 sinais de 1000 ms de duração, a partir do qual são gerados os mapas topográficos e são formados os vetores de entrada do classificador.

A Figura 2 apresenta os mapas topográficos corticais médios dos voluntários músicos e não-músicos, que representam, de forma visual, uma visão global da atividade elétrica, onde é possível notar a distinção entre os grupos de voluntários que são resultantes dos estímulos que estão sendo apresentados e processados pelo cérebro.

Sendo assim, cada voluntário é representado por um vetor composto pela concatenação dos valores

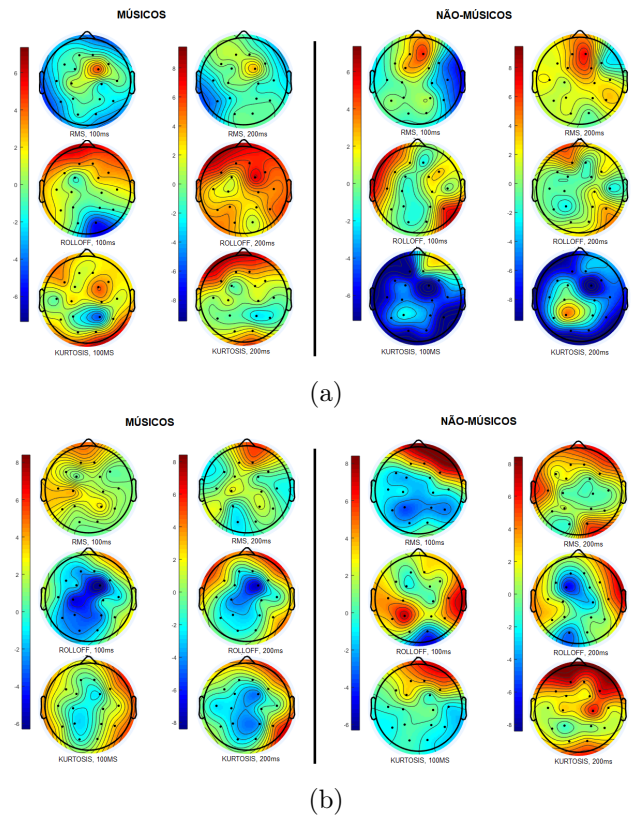


Figura 2 – Mapa topográfico cortical do sinal médio de músicos e não-músicos para as músicas (a) Dança Húngara No. 5, de Johannes Brahms, e (b) O Barbeiro de Sevilha - Abertura, de Gioachino Rossini, nos instantes 100 ms e 200 ms pós-estímulo.

registrados para todos os eletrodos ( $Ch = 16$ ), para cada característica acústica selecionada ( $f_p = 3$ ), para os dois instantes analisados, resultando em um vetor de dimensionalidade  $d = 96$ . Portanto, os dados obtidos de cada voluntário foram agrupados em uma matriz  $S_{96 \times 26}$ , onde as 26 colunas se referem ao número de voluntários.

Para a classificação foi utilizado o  $X$ -fold *cross-validation*, onde o conjunto de dados é separado aleatoriamente em  $X$  partes, garantindo que haja a mesma proporção de músicos e não-músicos em cada parte e, a cada iteração, uma das  $X$  partes é selecionada como teste e o restante é selecionado como treinamento. Neste trabalho foi utilizado o 5-fold (20% dos dados para teste e 80% para treinamento) e o 10-fold (10% dos dados para teste e 90% para treinamento).

Foi realizado o PCA sobre a matriz  $S_{96 \times 26}$ , projetando os dados nos autovetores, variando o número de componentes principais entre 5 e 25, e variando o número de  $k$  vizinhos mais próximos entre 1 e 3, selecionando como resultado aquele que apresentava melhor taxa de acerto na classi-

ficação onde, para a música de Johannes Brahms as 10 primeiras componentes principais representam 78,3% da variância dos dados enquanto que na música de Gioachino Rossini elas representam 81,8% da variância dos dados.

Na Tabela 2 são apresentados os melhores resultados da classificação para cada música encontrados utilizando  $k = 3$ ,  $CP = 15$  e  $CP = 10$  para as músicas de Rossini e Brahms, respectivamente, apresentando o valor da taxa de acerto e o desvio padrão, além da taxa de acerto por classe para os Músicos (M) e Não-músicos (NM).

Tabela 2 – Resultado da classificação dos voluntários para o classificador  $k$ -NN.

Xfold	Música	Média(std)	M	NM
5fold	Rossini	81,6%(12,7)	94,1%	68,2%
	Brahms	90,6%(8,8)	90,7%	90,4%
10fold	Rossini	82,9%(10,2)	98,5%	66,2%
	Brahms	88,9%(11,3)	86,9%	90,8%

## Conclusões

Neste trabalho foi proposto e implementado um método para a classificação de um grupo de voluntários músicos e não-músicos, por meio dos sinais de EEG registrados a partir da audição de duas músicas instrumentais, obtendo como resultado uma taxa de acerto na classificação superior a 80% para ambas as músicas, onde é possível verificar que apesar de os voluntários músicos serem amadores, são observadas diferenças entre os grupos nos mapas topográficos corticais, encontrando uma boa classificação.

A partir deste trabalho, algumas possibilidades podem ser executadas como trabalhos futuros, tais como: Diversificar os estilos das músicas apresentadas como estímulo, aumentar o número de eletrodos para a análise dos sinais cerebrais e aumentar o número de voluntários na análise. Por fim, espera-se que utilizando músicos profissionais a classificação se torne mais assertiva, havendo maior distinção entre os grupos.

## Agradecimentos

Os autores agradecem o suporte financeiro provido pela FEI e CAPES.

## Referências

- [1] I. Peretz and R. J. Zatorre, “Brain organization for music processing,” *Annual Reviews Psychology*, vol. 56, pp. 89–114, 2004.
- [2] P. Vusst, E. Brattico, M. Sppänen, R. Näätänen, and M. Tervaniemi, “The sound of music: Differentiating musicians using a fast, musical multi-feature mismatch negativity paradigm,” *Neuropsychologia*, vol. 61, pp. 1432–1443, 2014.
- [3] D. A. Abrams, S. Ryali, T. Chen, P. Chordia, A. Khouzam, D. J. Levitin, and V. Menon, “Inter-subject synchronization of brain response during natural music listening,” *European Journal of Neuroscience*, vol. 37, pp. 1458–1469, 2013.
- [4] V. Alluri, P. Toivainen, I. P. Jääskeläinen, E. Glerean, M. Sams, and E. Brattico, “Large-scale brain networks emerge from dynamics processing of musical timbre, key and rhythm,” *NeuroImage*, vol. 59, pp. 3677–3689, 2012.
- [5] P. Virtala, M. Huotilainen, E. Partanen, and M. Tervaniemi, “Musicianship facilitates the processing of western music chords - an erp and behavioral study,” *Neuropsychologia*, vol. 61, pp. 247–258, 2014.
- [6] H. Poikonen, V. Alluri, E. Brattico, O. Lartillot, M. Tervaniemi, and M. Huotilainen, “Event-related brain responses while listening to entire pieces of music,” *Neuroscience*, pp. 58–73, 2016.
- [7] O. Lartillot, *MIRtoolbox 1.6.1 Users Manual*. Department of Architecture, Design and Media Technology, 1st ed., 2014.
- [8] A. Lerch, *An Introduction to audio content analysis. Applications in signal processing and music informatics*. IEEE Press, 1st ed., 2012.
- [9] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*. Pearson, 6th ed., 2007.
- [10] S. J. Luck, *An introduction to the Event-Related Potential Technique*. MIT Press, 2nd ed., 2014.
- [11] OpenBCI, “Openbci - open source biosensing tools.” Distribuído digitalmente, nov 2017.