

Biometria em dados de EEG por conectividade funcional utilizando sincronização por motifs

Palavras-Chave: Biometria, eletroencefalografia, motifs, conectividade funcional

Autoras:

Manuela Von Ah Davanço [IFGW - Unicamp]

Marina Cabral de Paulo [IFGW - Unicamp]

Profa. Dra. Gabriela Castellano (orientadora) [IFGW - Unicamp]

INTRODUÇÃO:

A eletroencefalografia (EEG) é uma das técnicas mais antigas para a investigação da atividade cerebral [1]. Recentemente, a possibilidade de usar a EEG para produzir uma assinatura do indivíduo, isto é, para fins biométricos, tem sido estudada em vários trabalhos [2]–[4]. A maioria desses trabalhos utilizou medidas extraídas diretamente das séries temporais de eletrodos específicos, como a densidade espectral de potência [2], [5], [6].

No entanto, cada vez mais vem sendo explorado o fato de que regiões cerebrais atuam em conjunto para exercer tarefas cognitivas, sensoriais ou motoras. Isso levou ao estudo de redes cerebrais funcionais, ou da conectividade funcional [7]. A conectividade funcional é calculada comparando a atividade de pares de regiões cerebrais, utilizando alguma medida de similaridade [8]. Várias medidas de similaridade em sido exploradas no contexto de uso da EEG para biometria, como a coerência espectral [4], a correlação de

Spearman aplicada à transformada de Hilbert das séries temporais [9], o valor de bloqueio de fase ou *phase locking value* [10], a parte imaginária do *phase locking value* [11], o índice de atraso de fase ou *phase lag index* [12], e informação mútua aplicada a padrões ordinais [13], entre outras.

O objetivo do presente trabalho foi utilizar a medida de sincronização por motifs [14], também baseada em padrões ordinais [15], para o cálculo de redes cerebrais funcionais a partir de dados de EEG, e explorar a viabilidade de usar estas redes para fins biométricos. Em 2021, apresentamos um trabalho piloto [16] com dados de 20 indivíduos, obtidos de uma base de dados disponível online. No presente trabalho, a análise foi estendida para 50 indivíduos, e foram alterados os métodos de pré-processamento e classificação dos sinais.

METODOLOGIA:

Foi usada uma base de dados online chamada “EEG Motor Movement/ Imagery Dataset”

(<https://physionet.org/content/eegmmidb/1.0.0/>) para obter-se os dados de EEG usados [17], [18]. Nesta base, encontram-se dados de 109 indivíduos, obtidos por 64 eletrodos, cada um com 14 aquisições (diferentes tarefas sendo realizadas), sendo usadas, para este trabalho, apenas as no estado de repouso (uma com os olhos abertos, R1, e outra com os olhos fechados, R2, ambas com duração de um minuto) de 50 indivíduos.

Para o pré-processamento dos dados foi usado o software EEGLAB [19] implementado junto do MatLab. Inicialmente, foi feita uma limpeza visual dos sinais através da plotagem do gráfico da resposta temporal de cada eletrodo, e rejeitados os trechos com amplitudes muito discrepantes do padrão do sinal (trechos com amplitude 5x maior que o restante, por exemplo). Posteriormente, foi usada a ferramenta de *Independent Component Analysis* (ICA) para decompor cada sinal em suas componentes, podendo identificar as relacionadas à piscadas de olhos, por exemplo, ou movimento de músculos, as quais eram por fim rejeitadas. Em seguida, a banda alfa foi removida de todos os sinais por um filtro “*stop band*” no intervalo de 7 – 13 Hz, já que buscamos comparar sinais de olhos abertos com de olhos fechados, e esta banda é muito mais presente quando se está com os olhos fechados, o que poderia trazer diferenças significativas nos resultados. Por fim, um filtro “*pass band*” foi aplicado no intervalo de 4 – 50 Hz a fim de desconsiderar artefatos de baixas e altas frequências. Como última etapa do pré-processamento, foi aplicado o método de *Common Average Referencing* (CAR) [20], o

qual calcula a média entre os sinais dos eletrodos e subtrai esse valor de cada sinal, a fim de remover artefatos comuns a todos eletrodos.

Os 50 indivíduos analisados foram selecionados com a condição de que, após o pré-processamento, estivessem com um sinal de pelo menos 45 segundos de duração.

Nas aquisições de olhos abertos, R1, foram selecionados quatro intervalos, começando em 10 s, 20 s, 30 s e 40 s, enquanto nas de olhos fechados, R2, foi selecionado um intervalo começando em 30 s. A duração de cada intervalo foi primeiro de 1 segundo, e posteriormente de 5 segundos. Para cada intervalo foi calculada a respectiva matriz de conectividade funcional, sendo que uma matriz final de referência para R1 foi obtida pela média das quatro matrizes dessa aquisição.

Como medida de similaridade para o cálculo das matrizes de conectividade funcional, foi usado o método de sincronização por motivos, o qual consiste em “decompor” a série temporal do sinal de um eletrodo em uma série de motivos. Um motivo é um padrão de comportamento que pode ser retirado de um sinal de EEG, como mostra a Figura 1, onde têm-se os possíveis motivos de três pontos, usados no presente trabalho.

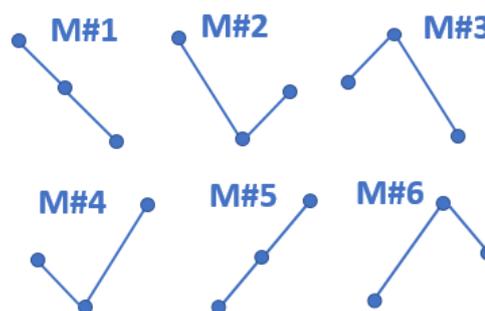


Figura 1: Motifs de três pontos utilizados no trabalho.

Comparando as séries de motifs de dois eletrodos, A e B, pode-se calcular a similaridade entre os sinais dos mesmos, calculando primeiro o coeficiente C_{AB} :

$$C_{AB} = \sum_{i=1}^{L_M} J_i$$

em que $J_i = 0$ se os motifs da posição i em A e B forem diferentes, e $J_i = 1$ se os motifs forem iguais, e L_M é o comprimento total da série de motifs.

Por fim, calcula-se o grau de sincronização entre os dois sinais:

$$Q_{AB} = \frac{C_{AB}}{L_M}$$

Ao final do processo, obtém-se uma matriz de conectividade 64x64 (pois os sinais foram obtidos de 64 eletrodos) onde cada elemento é o grau de sincronização entre o eletrodo linha e o eletrodo coluna.

Como método de comparação entre as matrizes de conectividade de R1 e R2, foi usada a correlação de Pearson, calculada pela função implementada em MatLab “*corrcoef()*”. Se o maior valor para a correlação fosse entre os sinais R1 e R2 do mesmo indivíduo, ele foi corretamente identificado dentre os outros.

RESULTADOS E DISCUSSÃO:

Os resultados obtidos podem ser encontrados na Tabela 1.

Tabela 1: Taxas de acerto para a identificação de indivíduos obtidos para intervalos com duração de 1 s e 5 s.

1 segundo	5 segundos
48%	48%

Dentre os 50 sujeitos analisados, 24 foram corretamente identificados, para as duas durações de intervalos, resultando então em uma acurácia de 48% para ambas.

Os resultados obtidos se encontram abaixo do esperado e do que poderia ser considerada uma acurácia alta para ser usada como um sinal biométrico (esperava-se um mínimo de 80%).

Por outro lado, é importante notar que foram comparados sinais de diferentes aquisições, R1 e R2, as quais possuem certas diferenças nas frequências dos sinais que, apesar da filtragem realizada, ainda podem estar presentes, já que em uma os olhos estão abertos e em outra estão fechados (maior atividade na banda alfa).

Também nota-se que a filtragem realizada, pelos passos de filtragem visual e ICA, por exemplo, são muito subjetivas, podendo ser retirados trechos ou componentes que, na verdade, deveriam ser consideradas.

CONCLUSÕES:

Os resultados encontram-se abaixo do esperado para concluir que há realmente um sinal biométrico nos dados, porém deve-se considerar o fato de se comparar duas aquisições em situações diferentes, as quais geram diferenças nas frequências de cada sinal. O número de 50 sujeitos analisados é um número pequeno para aplicação em biometria, portanto considera-se aumentar a base de dados utilizada.

A correlação de Pearson pode não ter sido o método de comparação mais adequado para os dados utilizados, sendo considerado

a utilização de outros métodos em estudos futuros, como calcular a distância geodésica entre as matrizes de conectividade.

BIBLIOGRAFIA

- [1] L. F. Haas, “Hans Berger (1873-1941), Richard Caton (1842-1926), and electroencephalography,” *J. Neurol. Neurosurg. Psychiatry*, vol. 74, no. 1, pp. 9–9, Jan. 2003.
- [2] S. Marcel and J. R. Millan, “Person Authentication Using Brainwaves (EEG) and Maximum A Posteriori Model Adaptation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 743–752, Apr. 2007.
- [3] P. Campisi, D. La Rocca, and G. Scarano, “EEG for Automatic Person Recognition,” *Computer (Long. Beach. Calif.)*, vol. 45, no. 7, pp. 87–89, Jul. 2012.
- [4] D. La Rocca *et al.*, “Human Brain Distinctiveness Based on EEG Spectral Coherence Connectivity,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 9, pp. 2406–2412, Sep. 2014.
- [5] J. Berkhout and D. O. Walter, “Temporal Stability and Individual Differences in the Human EEG: An Analysis of Variance of Spectral Values,” *IEEE Trans. Biomed. Eng.*, vol. BME-15, no. 3, pp. 165–168, Jul. 1968.
- [6] F. Su, H. Zhou, Z. Feng, and J. Ma, “A biometric-based covert warning system using EEG,” in *2012 5th IAPR International Conference on Biometrics (ICB)*, 2012, pp. 342–347.
- [7] K. J. Friston, “Functional and Effective Connectivity: A Review,” *Brain Connect.*, vol. 1, no. 1, pp. 13–36, Jan. 2011.
- [8] A. M. Bastos and J.-M. Schoffelen, “A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls,” *Front. Syst. Neurosci.*, vol. 9, Jan. 2016.
- [9] R. Cox, A. C. Schapiro, and R. Stickgold, “Variability and stability of large-scale cortical oscillation patterns,” *Netw. Neurosci.*, vol. 2, no. 4, pp. 481–512, Oct. 2018.
- [10] E. Pereda, M. García-Torres, B. Melián-Batista, S. Mañas, L. Méndez, and J. J. González, “The blessing of Dimensionality: Feature Selection outperforms functional connectivity-based feature transformation to classify ADHD subjects from EEG patterns of phase synchronisation,” *PLoS One*, vol. 13, no. 8, p. e0201660, Aug. 2018.
- [11] S. I. Dimitriadis, C. Salis, I. Tarnanas, and D. E. Linden, “Topological Filtering of Dynamic Functional Brain Networks Unfolds Informative Chronnectomics: A Novel Data-Driven Thresholding Scheme Based on Orthogonal Minimal Spanning Trees (OMSTs),” *Front. Neuroinform.*, vol. 11, Apr. 2017.
- [12] M. Garau, M. Fraschini, L. Didaci, and G. L. Marcialis, “Experimental results on multi-modal fusion of EEG-based personal verification algorithms,” in *2016 International Conference on Biometrics (ICB)*, 2016, pp. 1–6.
- [13] I. Kottlarz *et al.*, “Extracting Robust Biomarkers From Multichannel EEG Time Series Using Nonlinear Dimensionality Reduction Applied to Ordinal Pattern Statistics and Spectral Quantities,” *Front. Physiol.*, vol. 11, Feb. 2021.
- [14] R. S. Rosário, P. T. Cardoso, M. A. Muñoz, P. Montoya, and J. G. V. Miranda, “Motif-Synchronization: A new method for analysis of dynamic brain networks with EEG,” *Phys. A Stat. Mech. its Appl.*, vol. 439, pp. 7–19, Dec. 2015.
- [15] E. Olofsen, J. W. Sleight, and A. Dahan, “Permutation entropy of the electroencephalogram: a measure of anaesthetic drug effect,” *Br. J. Anaesth.*, vol. 101, no. 6, pp. 810–821, Dec. 2008.
- [16] M. V. A. Davanço, M. C. De Paulo, M. S. J. Preiss, G. Castellano. Estudo piloto sobre identificação de indivíduos a partir de dados de EEG por meio de conectividade funcional utilizando sincronização por motivos. In: XXIX Congresso de Iniciação Científica da UNICAMP, 2021, Anais eletrônicos. Campinas, Galoá, 2021.

Disponível em: <<https://proceedings.science/pibic-2021/papers/estudo-piloto-sobre-identificacao-de-individuos-a-partir-de-dados-de-eeg-por-meio-de-conectividade-funcional-utilizando--2>>. Acesso em: 20 jul. 2022.

[17] A. L. Goldberger *et al.*, “PhysioBank, PhysioToolkit, and PhysioNet,” *Circulation*, vol. 101, no. 23, Jun. 2000.

[18] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, “BCI2000: A general-purpose brain-computer interface (BCI) system,” *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, 2004.

[19] A. Delorme and S. Makeig, “EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis,” *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004.

[20] K. A. Ludwig, R. M. Miriani, N. B. Langhals, M. D. Joseph, D. J. Anderson, and D. R. Kipke, “Using a Common Average Reference to Improve Cortical Neuron Recordings From Microelectrode Arrays,” *J. Neurophysiol.*, vol. 101, no. 3, pp. 1679–1689, Mar. 2009.