



Estudo piloto sobre identificação de indivíduos a partir de dados de EEG por meio de conectividade funcional utilizando recorrências espaço-temporais

Palavras-Chave: biometria, eletroencefalografia, recorrências espaço-temporais

Marina Cabral de Paulo [IFGW - UNICAMP]

Manuela Von Ah Davanço [IFGW - UNICAMP]

Marcela Stracieri Janchevis Preiss [INSTITUTO PRINCIPIA]

Prof.^a Dr.^a Gabriela Castellano (orientadora) [IFGW - UNICAMP]

INTRODUÇÃO:

A eletroencefalografia (EEG) é uma das técnicas mais antigas utilizadas para o estudo da dinâmica cerebral [1]. Para fazer o exame de EEG, eletrodos são colocados sobre o escalpo do indivíduo, e estes medem sinais elétricos provenientes dos disparos de grandes populações de neurônios. A EEG tem sido utilizada para várias aplicações clínicas, particularmente para diagnóstico de algumas doenças neurológicas como epilepsia [2] e distúrbios do sono [3]. Estudos mais antigos também mostraram que esta técnica fornece informação sobre diferenças entre indivíduos, relacionadas a traços cerebrais anatômicos e funcionais [4], [5]. Mais recentemente, a ideia de usar sinais de EEG para distinguir indivíduos e implementar um possível sistema biométrico tem sido explorada de forma mais aprofundada [6]–[8]. Para isso, vários trabalhos têm usado sinais de EEG obtidos no estado de repouso [9], [10]. A vantagem de utilizar este tipo de paradigma (“estado de repouso”) é que pode ser aplicado a qualquer tipo de indivíduo (mesmo pacientes com deficiências, que teriam dificuldade para realizar algum experimento que requeresse uma resposta motora ou cognitiva) e, além disso, diminui o problema dos artefatos de movimento, já que a pessoa deve ficar imóvel durante a aquisição.

Dentre os trabalhos que visam obter informação biométrica a partir de sinais de EEG, estão os que têm usado características extraídas de eletrodos específicos, e aqueles que têm explorado as relações entre os sinais de EEG obtidos por diferentes eletrodos, método conhecido como conectividade cerebral. A ideia por trás das abordagens baseadas em conectividade é que muitas funções cerebrais são executadas por um conjunto de regiões (ou rede cerebral) e, portanto, entender como estas interagem entre si pode fornecer informação adicional sobre o indivíduo.

Os métodos para avaliar conectividade cerebral podem ser agrupados em conectividade anatômica (ou estrutural), conectividade funcional e conectividade efetiva [11]. O objetivo deste trabalho foi utilizar medidas baseadas em conectividade funcional, obtidas a partir de dados de EEG

no estado de repouso, para tentar identificar diferentes indivíduos. Este é um trabalho piloto, que utilizou dados de 20 indivíduos, obtidos de uma base de dados disponível online.

METODOLOGIA:

Os dados de EEG utilizados foram obtidos do banco de dados online “*EEG Motor Movement/Imagery Dataset*” (<https://physionet.org/content/eegmmidb/1.0.0/>) [12], [13]. Este banco contém dados de 109 indivíduos, mas para este trabalho piloto, foram utilizados dados de 20 indivíduos. Cada indivíduo do banco possui 14 aquisições de EEG, sendo que cada aquisição foi feita com diferentes atividades. Em particular, as duas primeiras aquisições foram feitas no estado de repouso (uma com olhos abertos e outra com olhos fechados), e foram as utilizadas neste trabalho.

Primeiramente, foi realizado o pré-processamento (ou “limpeza”) dos sinais, para eliminar ou reduzir ruído e artefatos. Para isso, foi utilizado o software EEGLAB, que roda na plataforma MATLAB, que permitiu a remoção de artefatos por meio visual através da plotagem gráfica do sinal. Depois de removidos os trechos com artefatos detectados visualmente, foi usada a técnica de *independent component analysis* (ICA), também disponível no EEGLAB, que viabiliza a remoção de componentes do sinal que caracterizam ruídos do momento de aquisição da imagem, como movimentos musculares do indivíduo, por exemplo. A técnica de ICA também foi utilizada para eliminar a componente da banda α de frequência do EEG (8-12 Hz), para as aquisições de olhos fechados. Isto pois esta faixa de frequências apresenta um pico característico quando o indivíduo fecha os olhos [14], e como a ideia era comparar sinais de olhos abertos e fechados, tentamos eliminar esta diferença.

Ainda com a finalidade de filtrar os sinais, foi utilizada a técnica de *common average referencing* (CAR), que consiste em calcular a média dos sinais de EEG de todos os eletrodos e subtrair este resultado do sinal de cada eletrodo. Esse método é usado para remover artefatos comuns a todos os eletrodos.

Após aplicadas as técnicas de pré-processamento, foi implementado o método de recorrências espaço-temporais, também na linguagem de programação MATLAB. Este método permite investigar como os dados em questão progridem ao longo do tempo através de mapas de recorrência [15].

Os mapas de recorrência são uma representação computacional que avalia se os dados do espaço de fase m -dimensional de um sistema dinâmico se aproximam de uma localização na qual eles já estiveram anteriormente [16]. Isso levanta a hipótese de que o método de recorrências espaço-temporais possa ser usado para avaliar a proximidade entre duas séries temporais de EEG estimando esta proximidade ou semelhança entre os dados. Portanto, é possível buscar conexões que permitam distinguir o sinal biométrico, já que os sinais pertencentes ao mesmo indivíduo devem ser mais próximos do que de indivíduos diferentes.

Assim, foi calculada uma matriz de recorrência, que foi transformada em uma matriz de adjacência (ou conectividade), que congrega mapas de recorrência de um dado intervalo de tempo.

O elemento i, j da matriz de recorrência espaço-temporal (STR) entre os eletrodos i e j para a n -ésima amostra é obtida da seguinte forma:

$$STR_{i,j}(\varepsilon, n) = \theta[\varepsilon - |x_i(n) - x_j(n)|]$$

onde tem-se uma estrutura de dados tridimensional $M \times M \times N$, sendo M a quantidade de eletrodos, N o total de instantes de tempo (amostras da série temporal) e $\theta(\cdot)$, a função de Heaviside, dada por

$$\theta(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

ε é um limiar de distância escolhido de forma arbitrária; neste caso, foi escolhido $\varepsilon = 50\%$ da maior distância entre o maior e menor valor dos dados sendo analisados.

Após estabelecida a matriz de recorrência STR , foi possível estabelecer a matriz de adjacência:

$$A_{i,j} = \theta \left[\left(\frac{1}{N} \sum_{n=1}^N STR_{i,j}(\varepsilon, n) \right) \right]$$

Para verificar se essas matrizes de adjacência podem ser utilizadas como um sinal biométrico, foram calculadas matrizes para trechos de 1 s ($N = 160$ amostras) dos sinais, extraídos das aquisições da seguinte forma: dois trechos da 1ª aquisição (10º segundo e penúltimo segundo), e um trecho da 2ª aquisição (penúltimo segundo). Foram feitas duas comparações: entre os dois trechos da 1ª aquisição, e entre o 2º trecho da 1ª aquisição e o trecho da 2ª aquisição.

As respectivas matrizes de adjacência foram comparadas por meio da distância euclidiana, dada por:

$$d_{XY} = \sqrt{\sum_{i=1}^M \sum_{j=1}^M (X_{i,j} - Y_{i,j})^2}$$

em que $X_{i,j}$ e $Y_{i,j}$ representam respectivamente o elemento i, j das matrizes de conectividade (adjacência) dos indivíduos X e Y .

RESULTADOS E DISCUSSÃO:

A Tabela 1 consiste na comparação entre dados da 1ª aquisição. Foi feita uma tabela similar para a comparação entre as duas aquisições (mas não está mostrada aqui por questões de espaço). Na Tabela 2 estão explicitadas as porcentagens de acerto no reconhecimento da biometria dos indivíduos das duas comparações apresentadas previamente.

Ao comparar os resultados para os trechos de mesmo sinal e de sinais diferentes na Tabela 2, percebemos que a porcentagem de acerto para trechos de um mesmo sinal é significativamente maior, ou seja, estes trechos carregam uma semelhança maior entre si, o que já era esperado, uma vez que os trechos foram adquiridos em uma única medida.

Tabela 1: Comparação entre os dois trechos da 1ª aquisição de cada indivíduo usando a distância euclidiana. SN designa cada indivíduo, onde N é o número atribuído no banco de dados. Estão marcados na cor verde os testes que tiveram resultados corretos (menor distância) ao tentar identificar o indivíduo (dados mais parecidos estavam dentro da mesma aquisição), em vermelho estão os valores errados (o menor valor da coluna não correspondia ao mesmo sujeito).

	S01	S02	S03	S05	S06	S07	S08	S09	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S22
S0	8,9	16,6	9,2	7,8	9,6	7,9	11,1	12,3	17,7	46,7	7,9	7,6	8,9	8,0	8,0	8,0	10,9	10,3	16,3	13,0
S0	3,7	8,6	3,1	1,6	4,3	2,2	6,1	7,1	13,4	40,0	2,4	2,1	3,9	2,8	3,2	2,8	7,9	4,2	11,5	7,4
S0	4,0	11,7	1,1	0,7	6,9	5,9	8,3	10,5	16,3	42,4	4,9	4,7	6,2	5,2	5,6	5,2	8,4	8,2	14,3	8,9
S0	4,6	12,0	2,4	0,3	3,8	2,8	5,8	7,5	13,2	40,8	1,9	1,6	3,3	2,8	2,8	2,1	8,0	5,9	11,4	7,0
S0	17,6	28,4	26,7	25,4	20,5	21,5	29,1	28,7	32,9	63,2	24,9	24,6	26,9	23,5	25,7	25,7	30,0	25,0	33,6	27,4
S0	4,8	11,7	5,7	4,5	6,0	1,8	7,6	9,0	14,7	41,9	4,5	4,2	5,8	4,0	5,0	4,8	10,0	5,4	12,7	8,8
S0	7,6	14,2	6,5	5,2	6,9	5,7	8,3	9,6	16,2	43,0	5,1	4,8	6,3	5,6	5,5	5,3	10,6	8,9	13,4	9,9
S0	4,0	11,5	2,4	1,0	2,9	2,0	4,9	3,2	13,3	40,8	0,9	0,6	2,5	1,6	1,7	1,1	7,1	5,2	10,6	6,2
S1	4,5	11,8	3,0	1,6	13,4	12,0	14,4	16,6	19,1	48,5	11,4	11,1	12,4	11,6	12,0	11,6	16,4	12,5	20,7	16,1
S1	5,0	15,1	6,2	5,4	21,2	19,8	25,6	27,7	31,1	57,1	22,1	21,7	23,6	21,1	22,6	22,4	27,3	24,0	31,4	23,3
S1	6,3	13,5	4,7	3,3	6,0	4,9	7,8	9,7	16,0	43,9	3,5	3,8	5,5	5,0	5,0	4,2	9,8	8,0	13,7	9,5
S1	5,4	11,5	5,2	4,0	21,5	18,6	23,4	23,2	29,2	55,6	20,0	20,1	22,3	20,4	20,9	20,9	23,1	22,2	26,0	24,9
S1	7,8	14,1	6,1	5,0	6,9	6,1	8,7	10,6	13,7	44,5	4,9	4,7	4,4	5,8	5,8	5,2	11,0	8,6	14,6	10,2
S1	6,9	16,0	8,0	6,9	8,5	7,0	10,5	10,6	16,1	45,6	6,8	6,6	7,7	5,0	7,2	7,0	11,2	9,8	15,3	11,1
S1	3,7	11,0	2,0	0,6	3,5	2,3	5,0	6,6	13,5	37,1	1,6	1,3	2,9	2,0	0,7	1,6	7,6	4,9	10,6	6,9
S1	4,7	12,2	2,7	1,3	2,6	2,0	4,5	6,3	13,0	40,1	0,7	0,4	2,2	1,6	1,2	0,2	7,0	5,0	10,3	6,2
S1	7,0	11,5	8,3	7,0	8,7	6,4	10,7	12,0	17,5	42,5	6,7	6,9	8,6	7,0	7,9	7,5	8,9	8,2	14,6	11,0
S1	3,5	9,7	2,8	1,4	3,0	1,5	5,0	5,5	12,9	40,4	1,1	0,8	2,6	1,3	1,6	1,3	6,8	3,8	10,3	6,1
S2	11,1	18,3	10,4	9,7	2,8	2,0	4,7	6,3	13,1	40,9	0,9	0,6	2,4	1,8	1,8	1,1	7,1	5,2	9,9	6,1
S2	5,9	13,2	4,2	3,0	2,1	1,7	4,4	6,0	13,0	40,4	0,5	0,2	1,6	1,5	1,4	0,7	6,8	4,7	10,3	5,7

Tabela 2: Porcentagem de acertos da comparação entre trechos da 1ª aquisição e entre os trechos das duas aquisições.

Trechos diferentes do mesmo sinal	Trechos de sinais diferentes
45%	15%

Entretanto, o resultado geral obtido não era esperado, pois tratam-se, ainda, de porcentagens muito baixas para conseguir identificar um indivíduo apenas através da metodologia apresentada.

Assim, é possível levantar certas hipóteses acerca de modificações plausíveis na forma como estes dados foram tratados, como por exemplo, se a banda α (frequência de sinal que caracteriza relaxamento de olhos fechados) não deveria ter sido filtrada em ambas as aquisições, uma vez que ela foi removida na 2ª aquisição, mas não na 1ª.

Outro questionamento seria se o método das recorrências espaço-temporais é realmente adequado para salientar diferenças individuais de conectividade. Ou ainda, se o número de indivíduos estudados nesta amostra (20 indivíduos) é pequeno demais para avaliar a acurácia do método e se o intervalo de tempo usado para o cálculo das matrizes de adjacência (1 segundo) é apropriado para a análise proposta.

CONCLUSÕES:

Tendo realizado a filtragem dos dados e a análise pelo método de recorrências espaço-temporais com a finalidade de encontrar um traço biométrico nos dados de EEG, obtivemos resultados que, apesar de ainda não estarem condizentes com o que era esperado, podem ser revistos e comparados através de outros critérios.

Futuramente, esperamos testar esta análise para uma amostra maior de indivíduos, após filtrar os sinais das duas aquisições aqui apresentadas da mesma forma, e também testar a aplicação do método para diferentes durações de sinal (ao invés de 1 s, como usado aqui).

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] M. Koutroumanidis *et al.*, "The role of EEG in the diagnosis and classification of the epilepsy syndromes: a tool for clinical practice by the ILAE Neurophysiology Task Force (Part 1)," *Epileptic Disord.*, vol. 19, no. 3, pp. 233–298, Sep. 2017.
- [3] I. G. Campbell, "EEG Recording and Analysis for Sleep Research," *Curr. Protoc. Neurosci.*, vol. 49, no. 1, Oct. 2009.
- [4] 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.
- [5] H. Van Dis, M. Corner, R. Dapper, G. Hanewald, and H. Kok, "Individual differences in the human electroencephalogram during quiet wakefulness," *Electroencephalogr. Clin. Neurophysiol.*, vol. 47, no. 1, pp. 87–94, Jul. 1979.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] D. Mantini, M. G. Perrucci, C. Del Gratta, G. L. Romani, and M. Corbetta, "Electrophysiological signatures of resting state networks in the human brain," *Proc. Natl. Acad. Sci.*, vol. 104, no. 32, pp. 13170–13175, Aug. 2007.
- [10] P. Campisi and D. La Rocca, "Brain waves for automatic biometric-based user recognition," *IEEE Trans. Inf. Forensics Secur.*, vol. 9, no. 5, pp. 782–800, May 2014.
- [11] K. J. Friston, "Functional and Effective Connectivity: A Review," *Brain Connect.*, vol. 1, no. 1, pp. 13–36, Jan. 2011.
- [12] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, Jun. 2000.
- [13] 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.
- [14] H. Berger, "Über das elektroencephalogramm des menschen," *Arch. Für Psychiatr.*, vol. 87, no. 1, pp. 527–570, 1929.
- [15] P. G. Rodrigues, C. A. S. Filho, R. Attux, G. Castellano, and D. C. Soriano, "Space-time recurrences for functional connectivity evaluation and feature extraction in motor imagery brain-computer interfaces," *Med. Biol. Eng. Comput.*, vol. 57, no. 8, pp. 1709–1725, Aug. 2019.
- [16] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence Plots of Dynamical Systems," *Europhys. Lett.*, vol. 4, no. 9, pp. 973–977, Nov. 1987.