EEG-based emotion classification using graph signal processing

Seyed Saman Saboksayr*, Gonzalo Mateos*, and Mujdat Cetin*†

*Department of Electrical and Computer Engineering, University of Rochester, Rochester, NY, USA

†Goergen Institute for Data Science, University of Rochester, Rochester, NY, USA

Abstract

The question of whether there exists a brain pattern associated with a specific emotion is the theme of many affective neuroscience studies. In this work, we bring to bear graph signal processing (GSP) techniques to tackle the problem of automatic emotion recognition using brain signals. With the help of GSP, we propose a new framework for learning class-specific discriminative graphs. To that end, firstly we assume for each class of observations there exists a latent underlying graph representation. Secondly, we consider the observations are smooth on their corresponding class-specific sough graph while they are non-smooth on other classes' graphs. The learned class-specific graph-based representations can act as sub-dictionaries and be utilized for the task of emotion classification.

Graph signal processing - 101

- Undirected $\mathcal{G}$ with adjacency matrix $\mathbf{W}$
- Define a signal $\mathbf{x}$ on top of the graph $\Longrightarrow x_i$ = Signal value at node i
- Associated with $\mathcal{G}$ is the combinatorial graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{W}$
  - $\Longrightarrow$ $\mathbf{D}$ is diagonal matrix where its elements are vertices degree
  - $\Longrightarrow$ Graph Laplacian is positive semidefinite so $\mathbf{L} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\top$
- GSP $\longrightarrow$ Exploit structure encoded in $\mathbf{W}$ or $\mathbf{L}$ to process $\mathbf{x}$
  - $\Longrightarrow$ Use GSP to learn the underlying $\mathcal{G}$ or a meaningful network model
- Graph Fourier Transform (GFT): $\tilde{\mathbf{x}} := \mathbf{V}^\top\mathbf{x}$ for undirected graphs
  - $\Longrightarrow$ Decompose $\mathbf{x}$ into different modes of variation
  - $\Longrightarrow$ Inverse (i)GFT $\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$, eigenvectors as frequency atoms
- Total variation (or *Dirichlet energy*) of signal $\mathbf{x}$ with respect to $\mathbf{L}$

$$\mathrm{TV}(\mathbf{x}) := \mathbf{x}^\top\mathbf{L}\mathbf{x} = \frac{1}{2}\sum_{i\neq j} W_{ij}\left(x_i - x_j\right)^2 = \sum_{k=1}^{N}\lambda_k\tilde{x}_k^2$$

  - $\Longrightarrow$ Smoothness measure on the graph $\mathcal{G}$
- For Laplacian eigenvectors $\mathbf{V} := [\mathbf{v}_1, \ldots, \mathbf{v}_N] \Longrightarrow \mathrm{TV}(\mathbf{v}_k) = \lambda_k$
  - $\Longrightarrow 0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N$ can be viewed as frequencies

Classification of network data

- labeled graph signals $\mathcal{X}_c := \{\mathbf{x}_p^{(c)}\}_{p=1}^{P_c}$ from C different classes.
  - $\Longrightarrow$ Signals in each class possess a very distinctive structure
- Assumption: Class c signals are smooth w.r.t. unknown $\mathcal{G}_c(\mathcal{V}, \mathcal{E}_c)$
- Multiple linear subspace model
  - $\Longrightarrow$ Signals spanned by few Laplacian modes (GFT components)
  - $\Longrightarrow$ Like susbpace clustering, but with supervision

Problem statement

Given training signals $\mathcal{X} = \bigcup_{c=1}^{C}\mathcal{X}_c$, learn discriminative graphs $\mathbf{W}_c$ under smoothness priors to classify test signals via GFT projections.

Learning graphs from observations of smooth signals

- Given a set of signal observations from P trials $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_P] \in \mathbb{R}^{N\times P}$
  - $\Longrightarrow$ A link between smoothness and sparsity

$$\sum_{p=1}^{P}\mathrm{TV}(\mathbf{x}_p) = \mathrm{trace}(\mathbf{X}^\top\mathbf{L}\mathbf{X}) = \frac{1}{2}\|\mathbf{W}\circ\mathbf{Z}\|_1$$

  - $\Longrightarrow$ Considering $\mathbf{Z} \in \mathbb{R}_+^{N\times N}$ as Euclidean-distance matrix
- Framework for learning graphs under a smoothness prior [1]

$$\min_{\mathbf{W}} \|\mathbf{W}\circ\mathbf{Z}\|_1 + f(\mathbf{W}) \quad \text{s. t.} \quad \mathrm{diag}(\mathbf{W}) = \mathbf{0},\ W_{ij} = W_{ji} \geq 0,\ i \neq j$$

  - $\Longrightarrow$ Convex function $f(\mathbf{W})$ encodes assumptions about the network $\mathcal{G}$
  - $\Longrightarrow$ Primal-dual solver amenable to parallelization, $\mathcal{O}(N^2)$ cost [1]

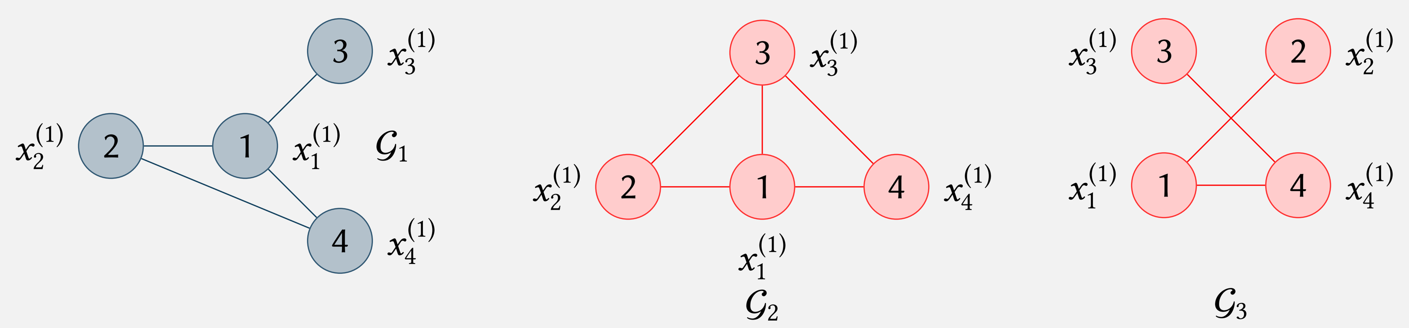
Discriminative graph learning

- Discriminative graph learning per class c

$$\min_{\mathbf{W}_c\in\mathcal{W}_m} \|\mathbf{W}_c\circ\mathbf{Z}_c\|_1 - \alpha\mathbf{1}^\top\log(\mathbf{W}_c\mathbf{1}) + \beta\|\mathbf{W}_c\|_F^2 - \gamma\sum_{k\neq c}^{C}\|\mathbf{W}_c\circ\mathbf{Z}_k\|_1$$

$$\mathcal{W}_m = \left\{\mathbf{W} \in \mathbb{R}_+^{N\times N} : \mathbf{W} = \mathbf{W}^\top, \mathrm{diag}(\mathbf{W}) = 0\right\}$$

  - $\Longrightarrow$ Capture the underlying graph topology (class c structure)
  - $\Longrightarrow$ Discriminability to boost classification performance
  - $\Longrightarrow$ Logarithmic barrier forces positive degrees
  - $\Longrightarrow$ Penalize large edge-weights to control sparsity



- Q: Given graphs $\{\hat{\mathbf{W}}_c\}_{c=1}^{C}$, how do we classify a test signal $\mathbf{x}$?
  - $\Longrightarrow$ Pass $\mathbf{x}$ through a filter-bank with C low-pass filters

$$\tilde{\mathbf{x}}_{F,c} = \mathrm{diag}(\tilde{\mathbf{h}})\hat{\mathbf{V}}_c^\top\mathbf{x} \quad\Longrightarrow\quad \hat{c} = \operatorname{argmax}_c\left\{\|\tilde{\mathbf{x}}_{F,c}\|_2^2\right\}$$

  - $\Longrightarrow$ $\tilde{\mathbf{h}} = [\tilde{h}_1, \ldots, \tilde{h}_N]^\top$ is the frequency response of an ideal low-pass filter
  - $\Longrightarrow$ Learned class-c GFT basis $\hat{\mathbf{V}}_c$

EEG emotion recognition

Emotion is an important aspect of human life

- Brain pattern associated with a specific emotion

Why EEG? 1. Non-invasive 2. Affordable 3. High temporal resolution

DEAP data-set: 32 participants rated 40 one-minute-long music video [2]

- Brain signals (EEG) recorded via 32 channels
- Rating levels of valence and arousal from 1 to 9

Results

- Outperform state-of-the-art in classifying low and high valence/arousal

Study	Valence	Arousal
Our method	**92.73**	**93.44**
Kalofolias [1]	86.56	88.91
Chao and Liu [3]	77.02	76.13
Chen et al. [4]	76.17	73.59
Tripathi et al. [5]	81.40	73.36

- Any meaningful pattern in the underlying learned graph?
- Learn two graphs corresponding to low and high emotions per person
  - $\Longrightarrow$ We do classification via low-frequency components
    - $\longrightarrow$ Mapping by low-frequency eigenvectors
- The asymmetrical pattern of the frontal EEG activity [6]
- Different patterns in: 1. Left frontal 2. Right temporal 3. Parietal [7]

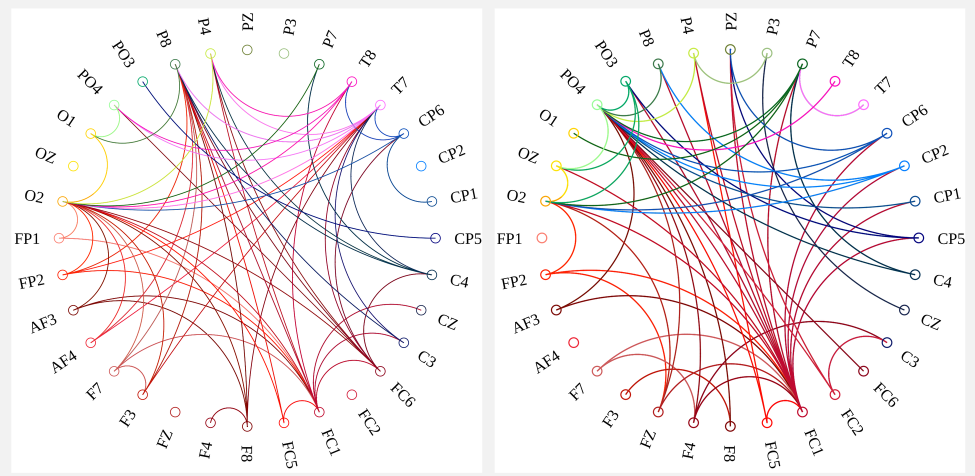


Figure 1:Significantly different connections between low and high for valence (left) and arousal (right).

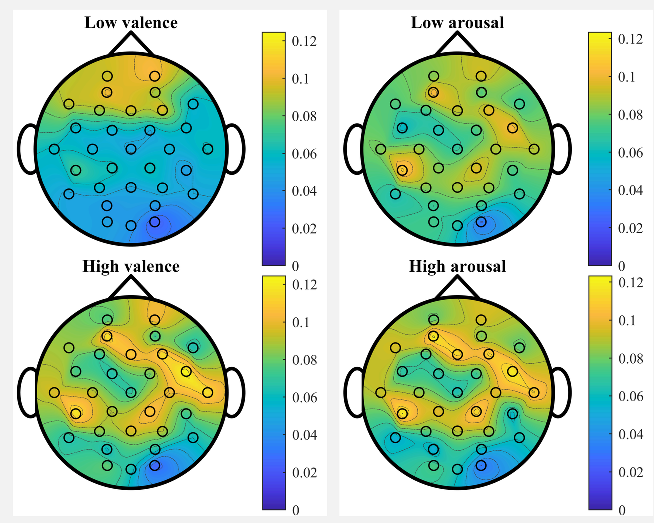


Figure 2:Mean of eigenvectors magnitude corresponding to low frequencies.

References

[1] V. Kalofolias, "How to learn a graph from smooth signals," in *Artif. Intel. and Stat. (AISTATS)*, 2016, pp. 920–929.

[2] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, 2011.

[3] H. Chao and Y. Liu, "Emotion recognition from multi-channel EEG signals by exploiting the deep belief-conditional random field framework," *IEEE Access*, vol. 8, pp. 33002–33012, 2020.

[4] M. Chen, J. Han, L. Guo, J. Wang, and I. Patras, "Identifying valence and arousal levels via connectivity between EEG channels," in *Intl. Conf. Affect. Comput. and Intel. Interaction (ACII)*, 2015, pp. 63–69.

[5] S. Tripathi, S. Acharya, R. D. Sharma, S. Mittal, and S. Bhattacharya, "Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset," in *Proc. of the Thirty-First AAAI Conference on Artificial Intelligence*, 2017, pp. 4746–4752.

[6] L. A. Schmidt and L. J. Trainor, "Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions," *Cognition & Emotion*, vol. 15, no. 4, pp. 487–500, 2001.

[7] D. Nie, X. Wang, L. Shi, and B. Lu, "EEG-based emotion recognition during watching movies," in *Intl. IEEE/EMBS Conf. Neural Eng.*, 2011, pp. 667–670.