Published in IEEE 12th International Symposium on Biomedical Imaging (ISBI)
Year: 2015
Pages: 351 - 354, DOI: 10.1109/ISBI.2015.7163885

TOPOLOGICAL SEIZURE ORIGIN DETECTION IN ELECTROENCEPHALOGRAPHIC SIGNALS

Yuan Wang[⋆] *Hernando Ombao*[†] *Moo K. Chung*[⋆]

[⋆] Department of Biostatistics and Medical Informatics, UW Madison, U.S.A.
[†] Department of Statistics, UC Irvine, U.S.A.

ABSTRACT

We propose a method for detecting seizure origin in epileptic electroencephalographic (EEG) data based on a novel multi-scale topological technique called persistent homology (PH). Among several PH descriptors, persistence landscape (PL) possesses many desirable properties for rigorous statistical inference. By building PLs on EEG epilepsy signals smoothed by a weighted Fourier series (WFS) expansion, we compared the before and during phases of a seizure attack in a patient diagnosed with left temporal epilepsy and successfully identified site T3 as the origin of the seizure attack.

Index Terms— EEG, epilepsy, persistence landscape, persistent homology

1. INTRODUCTION

Epilepsy is a neurological problem that can negatively affect the brain's visual, audial and motor functions [1]. Researchers are pursuing all possible avenues to gain a better understanding of the disease. EEG, an important imaging modality for understanding the function and dysfunction of the human brain, is a popular tool for studying epileptic seizures because of its non-invasive procedure and high temporal resolution. It indirectly records neuronal activity by capturing signals recorded by electrodes on the scalp. Many statistical methods have been developed over the past decades to study the patterns of these nonlinear electrical signals [2]. This paper aims to push the boundaries by exploring the topological detection of epileptic seizure based on multichannel EEG signals.

The PH technique reveals the topological structures of space X through a dynamic ranking of multi-scale features in the observed sample [3]. When PH is applied to functional data, the basic building blocks are the sub-level sets of functions. By changing the threshold values, we create a dynamic sequence of nested sub-level sets called a *Morse filtration*. As the threshold (called the *filtration value*) increases, the sub-level sets change from multiple isolated components into connected clusters. Since the ground truth of X is unknown in scientific experiments, we seek a compact PH descriptor that ranks topological features by their lifetime in the filtration. A collection of bars known as *barcode* serves exactly to such purpose. As a new PH descriptor, PL builds landscape-like structures on barcode that possess desirable statistical properties [4].

Our main contribution is the first ever application of the topological descriptor PL to the seizure detection problem in EEG while building a related signal analysis pipeline.

2. METHOD

Weighted Fourier series. For noisy functional measurements, it is necessary to smooth out high frequency noise in the function. We will use the WFS expansion, which is equivalent to the diffusion wavelet transform [5], to filter out noise in an EEG signal. Consider isotropic diffusion in an interval $[-T, T] \in \mathbb{R}$:

$$\begin{cases} \frac{\partial g(x,\sigma)}{\partial \sigma} & = -\frac{\partial^2 g}{\partial x^2} g(x,\sigma), \sigma \geq 0, x \in [-T,T], \\ g(x,0) & = f(x), \end{cases} \quad (1)$$

where the observed functional measurement f is treated as the initial condition. By imposing the additional periodic boundary conditions

$$g(-T,\sigma) = g(T,\sigma),\ \frac{\partial g}{\partial x}(-T,\sigma) = \frac{\partial g}{\partial x}(T,\sigma),$$

we obtain WFS

$$a_0 + \sum_{j=1}^{k} e^{-\frac{j^2\pi^2}{T^2}\sigma}\Big[a_j \cos\Big(\frac{j\pi x}{T}\Big) + b_j \sin\Big(\frac{j\pi x}{T}\Big)\Big], \quad (2)$$

which gives the unique solution of (1) when $k = \infty$. The Fourier coefficients are given by

$$\begin{aligned} a_j &= \frac{1}{T}\int_{-T}^{T} f(x)\cos\Big(\frac{j\pi x}{T}\Big)dx, \\ b_j &= \frac{1}{T}\int_{-T}^{T} f(x)\sin\Big(\frac{j\pi x}{T}\Big)dx. \end{aligned}$$

The diffusion time or scale σ modulates the smoothness of the estimation.

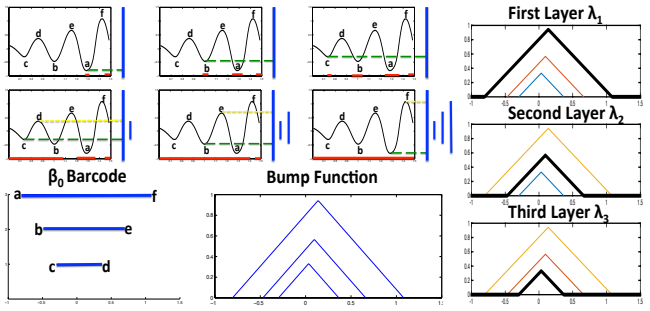


Fig. 1. Top left: barcode construction on a one-dimensional Morse function g as the filtration value increases. Bottom left: barcode and its bump function (3). Right: the PL traces all layers of outlines of the bump function.

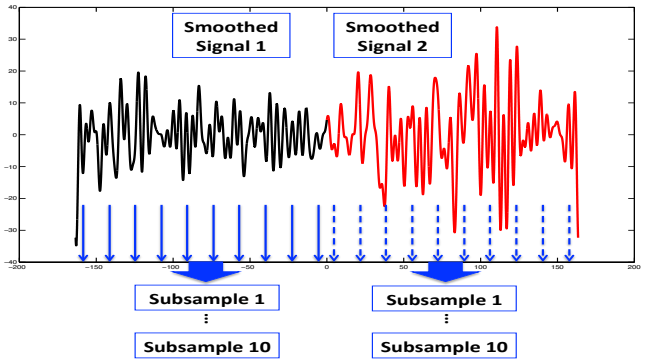


Fig. 2. Subsampling scheme: for each channel, smoothed functional values at every tenth time points starting at $t = s$ make up the sth subsample, $s = 1, \ldots, 10$.

Morse filtration. We apply PH to a smoothed 1D function g. Define the sub-level set of g thresholded at λ as

$$X_\lambda = \{x \in \mathbb{R} : g(x) \leq \lambda\}.$$

By varying λ, we create a dynamic sequence of nested subsets called a *Morse filtration*:

$$X_{\lambda_1} \subset X_{\lambda_2} \subset X_{\lambda_3} \subset \cdots \text{ for } \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots .$$

As the filtration value λ increases, the sub-level set change from multiple isolated components into connected clusters. These topological changes can be tabulated by the barcode [3]. A β_i-barcode is a set of bars with their endpoints representing birth and death times of i-th-dimensional holes in the filtration. A long bar in a barcode indicates a persistent feature in X in the face of growth, whereas short intervals in a barcode correspond to transient signals, noise or inadequate sampling. Figure 1 shows an example of β_0-barcode construction on a one-dimensional Morse function with unique local minima $a < b < c$ and maxima $d < e < f$. As the filtration value γ increases from $-\infty$ to ∞, new connected components are born consecutively in the sublevel sets $X_{g(a)}$, $X_{g(b)}$ and $X_{g(c)}$, and are merged in the sublevel sets $X_{g(d)}$, $X_{g(e)}$ and $X_{g(f)}$ according to the *Elder Rule*: *older components live on at a merging junction* [3]. The birth and death times of components are paired in as bars in a β_0 barcode.

Persistence landscape. It is difficult to perform statistical inference on barcodes. PL is a new PH descriptor that possesses many desirable statistical properties such as a Central Limit Theorem, on which it is easier to build test statistics [4]. Given an interval (a, b) on the barcode with $a \leq b$, we can define the piecewise linear *bump function* $h_{(a,b)} : \mathbb{R} \to \mathbb{R}$ by

$$h_{(a,b)}(t) = \max(\min(t - a, b - t), 0). \qquad (3)$$

The geometric representation of the bump function (3) is a right-angled isosceles triangle with height equal to half of the base of the corresponding interval in the barcode. The PL of $\{(a_i, b_i)\}_{i=1}^N$ is the set of functions defined by

$$\nu_k(t) = k\text{-th largest value of } \{f_{(a_i,b_i)}(t)\}_{i=1}^N, \qquad (4)$$

with $\nu_k(t) = 0$ for $k > N$. Geometrically, the PL $\{\nu_k\}_{k=1,2,3}$ traces the k-th outermost outline of these crossover triangles. Figure 1 illustrates PL for the underlying Morse function.

For each subsampled EEG signal, we construct PL. Then, for n PLs, $\{\nu_k^1, \ldots, \nu_k^n\}$, we compute the *average persistence landscape*, $\bar{\nu}(k, t) = \frac{1}{n}\sum_{i=1}^n \nu_k^i(t)$. Then we compare two average PLs $\bar{\nu}_1$ and $\bar{\nu}_2$ by the l_2-distance

$$d_2 = \left(\int \sum_{k=1}^{N} ||\bar{\nu}_1(k, t) - \bar{\nu}_2(k, t)||^2 dt \right)^{1/2},$$

where the interval is then discretely approximated by a trapezoidal Riemann sum with the partition intervals $\Delta t_j = t_{j+1} - t_j, j = 1, \ldots, m - 1$, being identical across the two sets of landscapes $\bar{\nu}_1$ and $\bar{\nu}_2$.

Subsampling and Permutation test. For statistical inference in a single-trial context, we subsample a smoothed EEG signal independently before and during seizure. Given a smoothed univariate signal, ten subsamples are created by pooling functional values at every sth time points starting at s for the sth subsample ($s = 1, \ldots, 10$). The subsampling scheme is shown in Figure 2. This gives two groups of 10 samples each. Then we compute d_2 distance between the average PLs before and during seizure. Since the distribution for d_2 is unknown, we utilize a permutation test and obtain its empirical distribution for statistical inference.

3. SIMULATIONS

We performed two simulation studies. Two blocks of signals were generated in each simulation of a study. In Study 1, two blocks were almost identical, whereas in Study 2, the second block showed significant frequency variation. In each simulation, we rejected the null hypothesis that the population mean PLs for the two phases were identical if the p-value was small.

Study 1. We simulated 100 epochs of two blocks of EEGs in the form Epoch = Baseline + Noise. In the first block, the baseline signal was generated by smoothing the autoregressive AR(1) model $y(t) = -0.9y(t-1) + \varepsilon(t)$, where $t = 1, \ldots, 100$, $\mu_1(0) = 0$, and the innovation process was Gaussian $\varepsilon \sim N(0, 0.1^2)$, with a robust LOESS procedure [6]. In the second block, the baseline signal was 0.8 times the baseline signal in the first block. In each simulation, independent Gaussian noise $N(0, 0.5^2)$ was added at every time point of first block of the baseline signal, whereas the second block was assigned independent noise $N(0, 1^2)$. Time was scaled down by a factor of 10.

Study 2. In each of the 100 simulations, we simulated two consecutive blocks of signals μ_1 and μ_2 by AR models inducing different frequencies. The first half of the signal μ_1 was generated by the AR(1) model: $\mu_1(t) = 0.9\mu_1(t-1) + \varepsilon(t)$; the second block of the signal μ_2 was generated by the AR(1) model: $\mu_2(t) = -0.9\mu_2(t-1) + \varepsilon(t)$, where $t = 1, \ldots, 100$, $\mu_1(0) = 0$, and the innovation process was Gaussian $\varepsilon \sim N(0, 0.1^2)$. Time was scaled down by a factor of 10.

Results. Figure 3 shows one representative simulation result and p-values for 100 simulations in each study. In Study 1 (left), there were three cases where the p-value was smaller than the 0.01 threshold, and so 3% false positives were detected at the threshold. In Study 2 (right), there were 97 cases where the p-values were smaller than 0.01, so the proposed method had 97% power in detecting frequency variation in the signal.

4. APPLICATION

In the current study, EEG data was retrieved by the Department of Neurology at the University of Michigan from a female subject already diagnosed with epilepsy on the left temporal lobe. Figure 4 shows a montage of the eight channels at which the EEG signals were sampled at a rate of 100 Hz with a total number of 32,680 time points. For this particular episode, the seizure initiates at the left temporal site (T3 channel) approximately halfway through the recording. We applied the topological detection method to the EEG data to identify the seizure origin.

Denoising, barcode and PL. The effect of denoising by WFS with $k = 85$ and $\sigma = 0.01$ is shown in Figure 4. Barcodes and their corresponding PLs calculated for signals smoothed at $k = 85$ and $\sigma = 0.01$ are shown in Figures 5(a) (before seizure) and 5(b) (during seizure) for the channels C4 and T3. We can see that T3 has more dominant landscape features than C4.

Seizure origin detection. Table 1 shows the p-values for the observed d_2 distance between average PLs to be greater than the shuffled distance based on 2000 permutations. The method consistently identified site T3 as the most significant at a 1% significance level with Bonferroni correction. Since there are two free parameters k and σ in the WFS representation, we performed additional analyses on the different combinations of parameters to see how sensitive the changing parameters are on the final outcome. Due to space limitation, only the combinations of degrees $k = 75, 85, 95$ and bandwidth $\sigma = 0.01$ are shown but other parameter combinations produced similar results. Neurologists may utilize the proposed topological analysis pipeline to guide them in examining regions that show the strongest significant changes during a seizure episode.

Table 1. Summary of p-values for combinations of $k = 75, 85, 95$ and $\sigma = 0.01$ at 2000 permutations.

		$k = 75$	$k = 85$	$k = 95$
$\sigma = 0.01$	C3	0.132	0.136	0.088
	C4	0.098	0.108	0.118
	Cz	0.346	0.324	0.122
	P3	0.140	0.338	0.106
	P4	0.108	0.018	0.104
	T3	0.001	0.001	0.001
	T4	0.118	0.022	0.024
	T5	0.102	0.112	0.121

5. DISCUSSION

Previous analysis on the same epilepsy data only showed the evolution of the spectral power and coherence during the seizure episode but was not abel to identify seizure location [2]. Here we demonstrated the potential utility of our proposed procedure for seizure localization. As part of our future study, we shall consider the option of a parametric test for distances between average PLs, and explore the use of PH in comparing network difference in phases of a seizure attack.

Acknowledgements. The research was funded by NIH grant UL1-TR000427 and the Vilas Associate Award from University of Wisconsin-Madison.

6. REFERENCES

[1] I Fried, "Auras and experiential responses arising in the temporal lobe.," *The Journal of Neuropsychiatry and Clinical Neurosciences*, vol. 9, no. 3, pp. 420–8, Jan. 1997.

[2] H. Ombao, R. von Sachs, and W. Guo, "SLEX analysis of multivariate nonstationary time series," *Journal of the American Statistical Association*, vol. 100, no. 470, pp. 519–531, June 2005.

[3] H. Edelsbrunner and J. Harer, *Computational Topology*, American Mathematical Society, Jan. 2010.

[4] P. Bubenik, "Statistical topology using persistence landscapes," *arXiv 1207.6437*, July 2013.

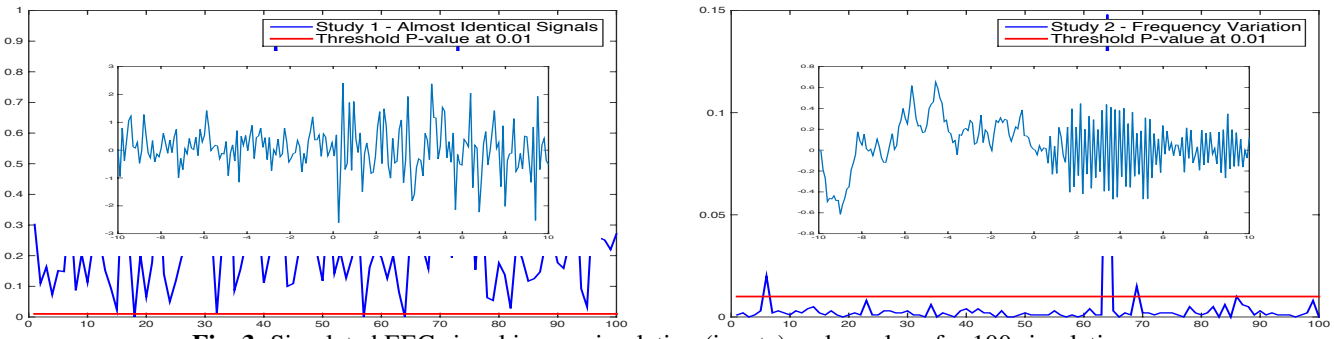


Fig. 3. Simulated EEG signal in one simulation (innate) and p-values for 100 simulations.

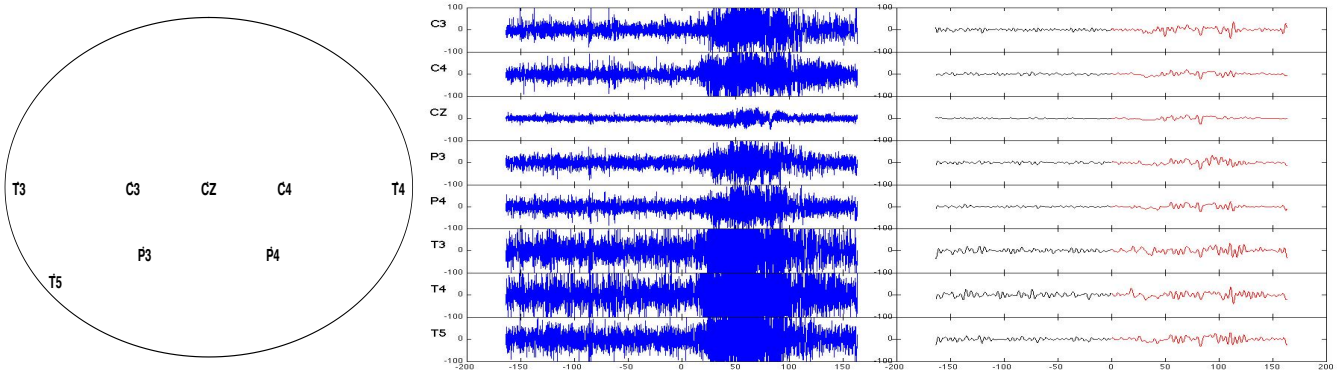
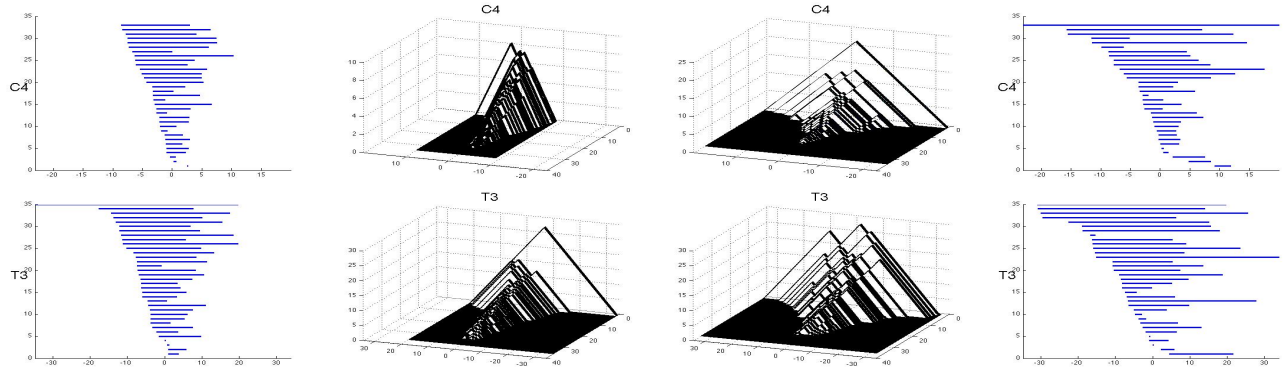


Fig. 4. Left: EEG channels in accordance with the international 10-20 system; T=temporal, P=parietal, C=central; odd (even) numbers indicate the left (right) hemisphere. Middle: EEG recording with a sampling rate of 100Hz over 163.4s; time 0 indicates the start of the seizure attack. Right: WFS denoising with degree $k = 85$ and bandwidth $\sigma = 0.01$. All amplitude units are in μV.



(a) Pre-seizure. (b) Seizure.

Fig. 5. Examples of observed barcodes (side) for one subsample and average persistence landscapes (middle) over all subsamples with smoothing parameters $k = 85$ and $\sigma = 0.01$ before and during seizure; each subsample produces one barcode, an average persistence landscape is the average bump function corresponding to barcodes across all subsamples.

[5] M.K. Chung, S.M. Schaefer, C.M. van Reekum, L. Peschke-Schmitz, M.J. Sutterer, and R.J. Davidson, "A unified kernel regression for diffusion wavelets on manifolds detects aging-related changes in the amygdala and hippocampus," *17th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2014.

[6] William S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, vol. 74, no. 368, pp. 829–836, 1979.