

An efficient method applied to spike pattern detection

Thang Nguyen Duc *, Tae-Seong Kim**, Young-Koo Lee*, and Sungyoung Lee*

*Dept. of Computer Engineering, KyungHee University

**Dept. of Biomedical Engineering, KyungHee University

Abstract

The detection of neural spike activity is a technical challenge that is very important for studying many types of brain function. On temporal recordings of firing events or interspike interval series of neural signal, spike pattern correspond to action will be repeated in the presence of background noise and they need to be detected to develop higher applications. We will introduce new method to find these patterns in raw multitrial data and is tested on surrogate data sets with the main target to get meaningful analysis of electrophysiological data from microelectrode arrays (MEA).

1. Introduction

In the neuroinformatic fields, microelectrode array recordings reveal complex structures in the firing of spatially distributed neurons and by analysis neural signal we can discovery decoding neural information and predict human action[4]. It is possible to measure population activity by using multiple electrodes that are spaced far enough apart so that each can function as a single independent electrode. An automatic classification process called spike sorting can be applied to recognize signal of single neural and then sample to get interspike interval series (ISI). This process is described by step by step. First, the spike trains are analyzed in the timing, the interval, or the bitstring representation. This means, that for a certain bin τ_i one checks for each train $j = 1 \dots S$ if a spike appears in this bin at position j , one writes '1', otherwise '0'. The results of this procedure are sequences of Bitstrings of length S representing the joint-spike activity. With multi-channel recordings, joint single channel together, we will have an ISI series correlative is presented as form $\{x_1, x_2, \dots, x_L\}$ where $x_i > 0$ is the number of spikes t_k, \dots, t_l of spike train $\{t_1, \dots, t_k \dots t_l \dots t_L\}$ for which $(i-1)\Delta\tau \leq t_k \leq \dots \leq t_l < i\Delta\tau$ represent the value of the i -th bin. Spike pattern are defined as those part of ISI series that repeat significantly more often than they would in a randomized series based upon the identical distribution.

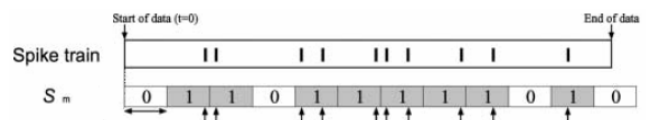


Figure 1. Spike pattern as Bitstring representation

Previous researches have been proposed based on time series analysis theory to build fundamental to solve this problem. All methods developed technique to estimate approximate pattern length because if we can find the pattern length, the next step to determinate spike pattern is performed easily by clustering algorithm on time series domain. Therefore, the method to estimate pattern length is very important and must satisfy efficient and exactly requirement. The histogram method is easily to implement, but by analysis on next session, we can see that it does not define clearly criteria to estimate the number of cluster. The template-based method [5][6] depends on pre-chosen templates and is not efficient algorithm with $O(n^k)$ complexity computation (k is pattern length). The correlation integral methods[1] although is computationally inexpensive, but still have $O(n^2)$ complexity computation. On this article, we will propose an $O(n)$ optimized algorithm relating to Kurtosis definition, not only fast but also exactly, and easily for widely implementing.

2. Spike pattern length estimation

Consider an arbitrary scalar time series of measurements $\{x_i\}, i = 1, \dots, L$ where L denotes the length of time series [2]. An embedded space is constructed based on L-m $\zeta_k^{(m)}$ vectors, with

* This research was supported by the MIC (Ministry of Information and Communication), Korea, Under the ITFSIP (IT Foreign Specialist Inviting Program) supervised by the IITA (Institute of Information Technology Advancement).

$\zeta_k^{(m)} = \{x_k, x_{k+1}, \dots, x_{k+m-1}\}$ and m is embedding dimension or also the pattern length we are finding.

The histogram based method will estimate the histogram of $\left\{ \sum_{k=1}^n x_k, \dots, \sum_{k=L-n}^L x_k \right\}$ or n -th order histogram and $\left\{ \zeta_1^{(m)}, \dots, \zeta_{L-m+1}^{(m)} \right\}$ mD histogram. By comparing the n or m step with $n+1$ and $m+1$ step correlating with n -th order and mD histogram, while the histogram diagram displays almost one cluster, we can show that the pattern length is n or m depend on the kind of histogram we have chosen. But we can see that how to compute the number of cluster is actually a hard problem and this method requires necessary manual aids from users.

The correlation integral method calculates based on the definition of correlation integral $C_N^{(m)}(r) = \frac{1}{N(N-1)} \sum_{i \neq j} \theta(r - \|\zeta_i^{(m)} - \zeta_j^{(m)}\|)$ with $\begin{cases} \theta(x) = 0, x \leq 0 \\ \theta(x) = 1, x > 0 \end{cases}$ and N is the number of embedded points ($N \leq L - m + 1$). By mathematic proving, we found that correlation integral is another kind of histogram estimation but its algorithm is clearly to get results. Increasing step by step the value of m until the arcs presenting $C_N^{(m)}, C_N^{(m+1)}, \dots$ are merged, at this point we have calculated the length m of pattern.

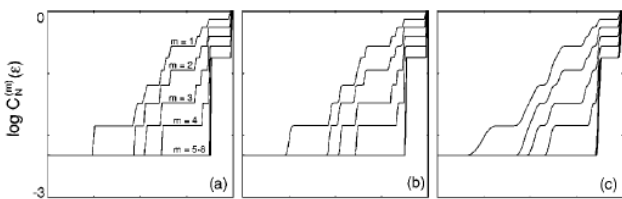


Figure 2. Estimate spike pattern length by correlation integral function

3. Improved spike pattern length estimation by comparing Kurtosis value

From observations of the histogram diagram, we can generally assume that the clusters corresponding to pattern has Gaussian distribution form on embedding space. Another thing is that comparing mD -histogram is actually projecting embedding vectors on $\vec{e} = \left(\frac{1}{\sqrt{m}}, \dots, \frac{1}{\sqrt{m}} \right)$ direction and then finding the condition one cluster on $1D$ embedding space. We have developed our algorithm based on these bases. First, rather than chose \vec{e} vector as t

his form, we applied CCIPCA algorithm[3] to find the best optimized project direction with minimum Mean Square Error (MSE). A set of vectors $\{\zeta_1^{(m)}, \dots, \zeta_{L-m+1}^{(m)}\}$ will be projected to $\Omega = \{v_1, v_2, \dots, v_n\}$ with v_i is $1D$ vectors. Increasing value of m , histogram distribution present set Ω will become more Gaussian distribution by the condition existing one pattern on dataset. When m exceeds pattern length, the histogram distribution will be more flat and convergence. With the special properties of Kurtosis value of data set $\kappa = \frac{\kappa_4}{\kappa_2^2} = \frac{\mu_4}{\sigma^4} - 3$, going to negative when data set distribution has flat shape, zero with Gaussian distribution and nonnegative on the other cases, we can easily estimate with high degree of accuracy the pattern length.

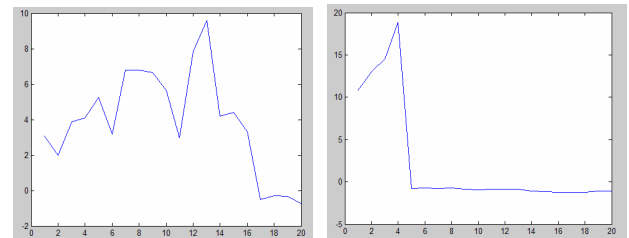


Figure 3. Estimate spike pattern length by comparing Kurtosis value

4. Improved spike pattern length estimation by comparing Kurtosis value

On this article, based on previous fundamental research, we have proposed an efficient, simple and fast algorithm used for spike pattern recognition. This algorithm is not only can be applied to neuroinformatics fields but also extended to various fields relating to time series analysis, data mining and pattern recognition.

Reference

[1] M.Christen, A.Kern, A.Nikitchenko, W.-H Steed and R.Stoop, "Fast spike pattern detection using the correlation integral" *Physical Review E* 70 (2004)
 [2] Markus Christen, "The Role of Spike Patterns in Neuronal Information Processing", *Doctoral Thesis* 07.10.1969
 [3] Juyang Weng, Yilu Zhang, Wey-Shiuan, "Candid Covariance-Free Incremental Principal Component Analysis", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol 25. August 2003
 [4] Jeab Marc Fellous, Paul H.E.Tiesinga, Peter J.Thomas and Terrence J.Sejnowski, "Discovering spike patterns in neuronal responses", *The journal of Neuroscience* March 24th 2004
 [5] J.E. Dayhoff and G.L. Gertein, *J.Neurophysiol* **49**, 1334 (1983)
 [6] I.V.Tetko and A.E.P.Villa, *J.Neurosci. Methods* **105**, 1 (2001).