Proceedings of Institute of Mathematics of NAS of Ukraine 2004, Vol. 50, Part 3, 1410–1414

Symmetry Analysis in Spikes (Bursts) Recognition and Classifications

Alexander MAKARENKO †, *Anatoliy POLIARUSH* † *and Igor TETKO* ‡

† *Institute of Applied System Analysis, National Technical University of Ukraine (KPI), 37 Peremogy Ave., Kyiv, Ukraine*
E-mail: *makalex@mmsa.ntu-kpi.kiev.ua, apolixus@yahoo.com*

‡ *Institute of Bioorganic and Petroleum Chemistry of NAS of Ukraine, Kyiv, Ukraine; Institute for Bioinformatics, Ingolstraedter Lanstrasse 1, D-85764, Neuherberg, Germany*
E-mail: *itetko@vvclab.org*

New approach to spikes (burst) recognition is described. The symmetry properties of recognized objects are the foundation of the methodology. The objects are described by ordinary differential equations in special phase space. Then the symmetry is connected with phase flow of ODE systems. Neuronal spikes recognition is the example of approach application.

1 Introduction

In recent research of data processing, pattern recognition, modelling the problem of recognition of abrupt changes and intrinsic structures is very important. Especially interesting are also applications of recognition of the structural elements in time series data. The examples are the cycles in economics, spikes in neurophysiology, textures in pattern recognition. Analysis follows to the conclusion that structural properties are geometrical properties of some objects. This requires the application of geometrical approach to geometrical properties recognition. Proposed study covers the problem of recognition of objects that have the form of mappings of a segment onto a manifold. The method presented is based on symmetry analysis.

2 Abstract statement of the problem

Let us consider the following function of class k:

$$\varphi : [a; b] \to M, \tag{1}$$

where $[a; b] \subset \mathbb{R}$, and M is some manifold of class k. We shall designate the set of such functions as Φ. So, functions (1) describe geometrical objects in M.

As is known, manifolds are metrizable. Let us designate one of possible distances

$$d_M : M \times M \to [0; +\infty)\,. \tag{2}$$

Similarly we may introduce a distance function between geometrical objects in M

$$d : \Phi \times \Phi \to [0; +\infty)\,. \tag{3}$$

Consider an arbitrary transformation T that acts on Φ:

$$T : \Phi \to \Phi. \tag{4}$$

Then we can define classes $K_T(\varepsilon)$ of geometrical objects in M:

$$K_T : (0; +\infty) \to 2^{\Phi}, \tag{5}$$

where 2^Φ is a set of subsets in Φ, and

$$K_T(\varepsilon) = \{\varphi : d(T\varphi, \varphi) < \varepsilon\}. \tag{6}$$

Suppose that $A \subset 2^\Phi$ consists of several classes $K_{T_i}(\varepsilon_i)$, $i = 1, \dots, n$. Then the problem of classification consists in finding T_i if A is given.

3 Geometrical objects described by an ordinary differential equation

Suppose we have n ordinary differential equations of order m:

$$\begin{aligned} &\frac{dy^{(0)}}{dt} = y^{(1)}, \qquad \frac{dy^{(2)}}{dt} = y^{(2)}, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &\frac{dy^{(m-1)}}{dt} = f_i\left(y^{(0)}, y^{(1)}, \dots, y^{(m-1)}\right), \end{aligned} \tag{7}$$

where $i = 1, \dots, n$, and f is bounded function defined on some area $D \subset \mathbb{R}^m$. Then all functions $y(t)$ that are solutions of (7) induce a set of geometrical objects on $\mathbb{R}^m$. In this particular case (1) takes the form

$$\varphi : A_\varphi \subset \mathbb{R} \to \mathbb{R}^m, \tag{8}$$

where

$$\varphi : t \mapsto \left(y^{(0)}(t), y^{(1)}(t), \dots, y^{(m-1)}(t)\right)^T \tag{9}$$

and φ is a function of class 1. Then for an arbitrary point $p \in D$ there exists the only function $y_p(t)$, that $\left(y_p^{(0)}(0), y_p^{(1)}(0), \dots, y_p^{(m-1)}(0)\right)^T = p$. Let us define such mapping $X$of D into $\mathbb{R}^m$, that

$$X : p \mapsto \left(y_p^{(1)}(0), y_p^{(2)}(0), \dots, y_p^{(m)}(0)\right)^T. \tag{10}$$

It is a vector field on D, induced by equation (7). Similarly, we can define a local one-parametric group G, that acts on D:

$$g_\tau : p \mapsto \left(y_p^{(0)}(\tau), y_p^{(1)}(\tau), \dots, y_p^{(m-1)}(\tau)\right)^T. \tag{11}$$

Consider the action of G on set Φ of geometrical objects represented by functions

$$\phi : A_\phi \subset \mathbb{R} \to D. \tag{12}$$

by formula

$$g_\tau : \Phi \to \Phi, \tag{13}$$

$$g_\tau(\varphi)(t) = g_\tau(\varphi(t - \tau)). \tag{14}$$

Let us introduce a set of functions

$$Q_\tau : \Phi \to [0; +\infty), \tag{15}$$

that act as

$$Q_\tau : \varphi \mapsto d(g_\tau(\varphi), \varphi). \tag{16}$$

It is obvious that $Q_\tau(\varphi) = 0$ for any admissible τ if and only if φ is a solution of (7). Then we have a simple criterion of evaluation of adjustment of the geometrical object to given differential equations (7).

4 Neuronal spikes recognition as example of symmetry analysis application

Analysis of populations of neurons represents an important step for better understanding of brain functioning. The present theories of brain functioning put accent on the neuronal interactions in terms of syncronized firing across many neurons, or spatio-temporal interactions and presence of specific patterns. Many such neuronal interactions cannot be observed with recordings from single neuron. Thus, analysis of neuronal populations does not simply provide an additive scheme for increase experimental data but makes possible to detect some features of brain functioning that could never be observed with recording of only single neuron. We represent at Fig. 1 typical example of spike in neurophysiological data.

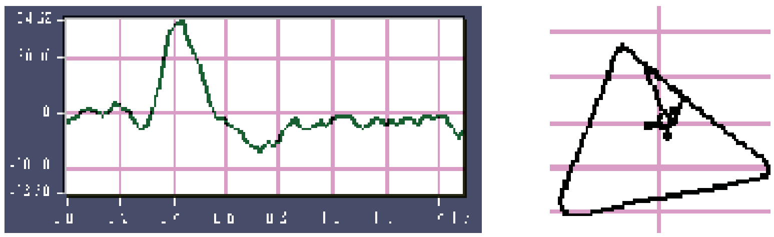

Figure 1. Spike example and its portrait in phase space. The centers of the spikes are at 0.41 ms. The axis x corresponds to time and axis y corresponds to voltage in electrode.

The basic hypothesis used to detect and separate action potential of neurons assumes that spikes generated by the same neurons have similar shapes and these shapes are unique and conservative for each recorded neuron. The shape of neuron spikes detected by the electrode depends on the distance, relative position, properties of the media, e.g., presence and distribution of glia cells, between electrode and neurons. The shape also depends on resistance of electrode and neuron. Thus, even if two neurons are identical, their action potential detected at the microelectrode can be different and, thus, spikes from such neurons could be separated.

The number of different methods used in the neurophysiology for spike sorting dramatically increases. One of the most popular methods is template matching. These techniques use templates that represent some typical waveform shapes of neurons in time domain. A classification of a candidate spike is done by its comparison to all templates and selecting the best matching one.

There are many basic problems that should be solved for successful spike sorting. The first basic question concerns the number of different types of neurons that should be separated in the experimental data. The usual practice is to use a “supervisor”, i.e. the experimentator, who can provide a preliminary classification of data, e.g. selection of template spikes, based on his experience and knowledge. However, even if the number of classes is identified, the separation of spikes remains very difficult problem due to extracellular and intracellular noise that can disturb the form of the action potential.

The extracelular noise is usually taken into account by the most of models as an additive noise. The intracellular noise that can produce variation in the spike waveform is more difficult to account for. Recently we have proposed a new method for spike sorting that consider the problem of spike sorting in phase space and describe the spike waveform as an ordinary differential equation with perturbation. This approach made possible to account for both extracellular and intracellular noise. The differential equation describing the activity of a neuron was supposed to have a limit trajectory in phase space, and noise was treated as deviation of the signal from that trajectory. Current study provides further development of this idea. In contrast to that we proposed a numerical method that takes into account that the variety of spike waveforms generated by the same neuron cannot be explained only by influence of both kinds of noise on some "typical" for a given neuron impulse, but it should be considered as a whole.

Let us suppose that every observed spike $x_i(t)$ of the i-th neuron is a solution of an ordinary differential equation with perturbation

$$\frac{d^3x}{dt^3} = f_i\left(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}\right) + F\left(t, x, \frac{dx}{dt}, \frac{d^2x}{dt^2}\right), \tag{17}$$

where $F(\cdot)$ is a perturbation function. The perturbation function $F(x, \dots, t)$, bounded by a small value, is a random process with zero mean and small correlation time $\tau^* \ll T$. The solution of the equation

$$\frac{d^3x}{dt^3} = f_i\left(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}\right) \tag{18}$$

describes a self-oscillating system. The duration of spikes is limited, so we can suppose that $x(t)$ is defined on some interval $(a; b)$. In practice parameters a and b are chosen manually by an expert. Set of differential equations

$$\frac{dx^{(0)}}{dt} = x^{(1)}, \qquad \frac{dx^{(1)}}{dt} = x^{(2)}, \qquad \frac{dx^{(2)}}{dt} = f_i\left(x^{(0)}, x^{(1)}, x^{(2)}\right), \tag{19}$$

where $x^{(k)}$ is the k-th derivative of $x(t)$ is equivalent to (18).

If function f_i is defined in open domain $D \subset \mathbb{R}^3$ then any solution $x(t)$ of (19) induces phase trajectory $\vec{X}(t) = \left(x^{(0)}(t), x^{(1)}(t), x^{(2)}(t)\right)^T$. In this case the trajectory is also an integral curve for (19). The theorem about existence of a solution for an ordinary differential equation gives that for an arbitrary point $p \in D$ there exists such integral curve $\vec{X}_p(t)$ that $\vec{X}_p(t_0) = p$ for some t_0 and $\vec{X}_p(t)$ is a solution of (19).

The equation (18) induces a local one-parameter group G on D. Let us define some action $g_\tau \in G$ on D. Consider integral curve $\vec{X}_p(t)$ that passes p when $t = t_0$. Then

$$g_\tau : p \mapsto \vec{X}_p(\tau + t_0), \tag{20}$$

i.e., $g_\tau(p)$ describes the state of system (19) at time moment $\tau + t_0$ under initial conditions p.

If the action of G on D is known we can introduce a criterion that allows to determine whether an arbitrary function $x(t)$, defined on interval $(a; b)$, is a solution of (19). That is, $\vec{X}(t) = \left(x^{(0)}(t), x^{(1)}(t), x^{(2)}(t)\right)^T$ is an integral curve for (19) if and only if

$$g_\tau\left(\vec{X}(t - \tau)\right) = \vec{X}(t) \tag{21}$$

for any $t \in (a; b)$ and any permissible $\tau \in (-\varepsilon; +\varepsilon)$.

Therefore, the new theoretical background uses a wider class of differential equations, not necessarily describing self-oscillating systems. Another essential difference is that the activity of a neuron is classified with a symmetry transformation in phase space. Criterion of classification is the steadiness of the portrait in phase space of a spike against the transformation that corresponds to the given neuron (under the action of local one-parameter group). A computational method for modelling transformations is introduced and tested.

The algorithm to separate neuronal spikes several intermediate steps:

1. Spike detection from the noisy signal;
2. Calculation of distances between the phase trajectories of the detected spikes;
3. Detection of spikes that hypothetically belong to the same neuron;

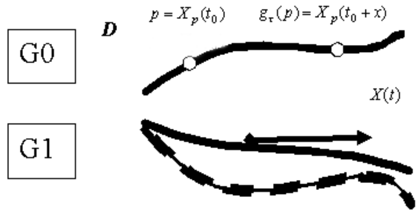

Figure 2. A vector field and its integral curves that correspond to solutions of equation (3). The action of g_t on point p is shown. The curve G0 corresponds to trajectory of equations (19) invariant under the action of one-parameter group. Any other curve G1 (not the solution of the system (19)) is not invariant under the action of group which corresponds to system (19). The difference between G1 and transformed curve G1 (dashed line) is the measure of dissymmetry and is the tool for classification.

4. Building a numerical model of transformation group that corresponds to every neuron observed;
5. Classification of spikes.

The three first steps were performed similarly to our previous analysis and are described in details in elsewhere [1]. In this article we only briefly summarize the main implementation details and parameters of calculations of these steps that are important to reproduce our work. The vector field approach (steps 4–5) is described and main differences between old and new algorithm are emphasized.

The first 4 steps corresponded to the training phase on which the number of spike classes was estimated and the vector field for each class was constructed. In order to perform this analysis a few dozens of spike occurrences, usually corresponding to several minutes of recording time were required. Like in our previous approach a human expert could participate at step 3 were the number of spike classes was detected.

The method described in the article is development of the method based on template matching in phase space. The stages of spike detection and forming spike classes are generally similar to the latter. The essential difference of the method presented from template matching is implementation of spike classification after spike classes are formed. The template matching method characterizes the whole class of spikes generated by the same neuron by the phase portrait of its typical spike.

In contrast to other methods, the method based on symmetry analysis uses computational modelling of vector field that conforms the differential equation describing the activity of the chosen neuron and thus it is more stable against alterations of spike form. It showed better results than the former and can be useful in experiments, where spikes tend to change their form. More details are adduced in [1–4].

[1] Polyarush A.I., Makarenko A.S. and Tetko I.V., Spike separation based on symmetries analysis in phase space, *System Research and Informational Technologies*, 2003, N 2, 118–126.

[2] Polyarush A.T., Tetko I.V. and Makarenko A.S., Geometrical-invariant approach to recognition of the structures in the time series and abstract maps, *Int. J. of Computing Anticipation Systems*, submitted.

[3] Koch C., Biophysics of computation, Oxford University Press, 1998.

[4] Letelier J.C. and Weber P.P., Spike sorting based on discrete wavelet transform coefficients, *Journal of Neuroscience Methods*, 2000, V.101, 93–106.