ARTICLE IN PRESS

JID: NEUCOM [m5G;December 22, 2017;12:37]

Neurocomputing 000 (2017) 1–8

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Low-dimensional feature fusion strategy for overlapping neuron spike sorting

Hong Ge Li*, Rui Qi Song, Jian Wei Liu

School of Electronic and Information Engineering, Beihang University, Beijing 100191, China

ARTICLE INFO

Article history:
Received 2 September 2016
Revised 13 September 2017
Accepted 1 December 2017
Available online xxx

Communicated by Deng Cai

Keywords:
Overlapping spike
Feature fusion
Wavelet coefficients
Principal component analysis (PCA)

ABSTRACT

Complex action potentials of the brain neurons from extracellular recordings are usually represented as overlapping superimposed spikes resulting from the neuron spike bursts. In this paper, a novel feature fusion strategy is proposed for the high-precision classification of the neuron spikes. Based on the wavelet coefficients and the principal component analysis (PCA) features of the spikes, the eigenvectors of the adjacency matrices are constructed using the Locality Preserving Projections (LPP) algorithm. The fusion adjacency matrix of different features is constructed using different weighted adjacency matrices method. The feature fusion data based on the wavelet coefficients and PCA features are projected from high-dimensional space to low-dimensional space. The experiment results show that the noise level is 0.4, and the accuracy of this method is increased by 9% on average compared with non-fusion feature classification. According to the experimental results, a high-precision classification is obtained based on the feature fusion method with a low-dimensional feature space. The proposed feature fusion method effectively reduces the dimension number of the features and remedies the missing information.

1. Introduction

Neuroscientists have extensively studied neuron spikes recorded extracellularly [1]. Spike classification poses a great technical challenge for research in neural information science [2]. Neural spike sorting is already used in brain science research. Neural spikes are a comprehensive reflection of the electrophysiological activity of brain cells. The recording, detection, and classification of neural spikes are a hotspot in the field of brain science engineering [3,4]. The extracellular electrode technology is a common method of neural spike recording. However, the simultaneous discharges of multiple adjacent neurons near the microelectrodes often result in overlapping spikes. Therefore, it is essential to separate the spikes according to each neuron.

Related studies have been conducted on the classification of overlapping spikes. The template-matching method has been used to solve the problem of overlapped spikes [1] by using PCA to extract the spike features, which are compared with the template to identify the types of spikes. However, the applicability of this method is limited because of the need for a relevant template. Another approach achieves overlapping spike sorting using neural networks to learn more about the general decision boundaries [2]. This approach is that the neural network must be trained using the spikes data recorded.

The noise-assisted strategy has been used to identify and classify overlapping spikes, and the further classification is performed by a trained learning machine [3]. The independent component analysis (ICA) method can also separate overlapped spikes, but it must meet the requirement that there are more channels of overlapping spikes than the number of spike components [4,7]. The relaxation algorithm has been used to solve the problem of overlapping spikes by using the Fast Fourier Transformation (FFT) to extract the spike features and then matching and identifying the types of spikes [5]. However, this method is sensitive to noise. A single feature method is usually used to identify the spikes. The type of overlapping spike also increases with its number. Therefore, a single feature of spike is often insufficient to characterize the differences between the different types of spikes.

In this paper, a feature fusion strategy is proposed to classify overlapping spikes. This strategy aims at utilizing the advantages of different features to work out different spike extraction methods. It can make up for the deficiency of using a single feature and provide a new measurement mode of the different spikes. Both the wavelet coefficients and the PCA features extraction are widely used in the classification of the spikes. The wavelet coefficients' reflect local difference of the spikes. On the other hand, the PCA features extraction reflects the spike amplitudes. The purpose of the feature fusion is to fully utilize different features of the

* Corresponding author.
E-mail addresses: honggeli@buaa.edu.cn (H.G. Li), songruiqi@buaa.edu.cn (R.Q. Song), liujianwei@buaa.edu.cn (J.W. Liu).

https://doi.org/10.1016/j.neucom.2017.12.004
0925-2312/

Please cite this article as: H.G. Li et al., Low-dimensional feature fusion strategy for overlapping neuron spike sorting, Neurocomputing (2017), https://doi.org/10.1016/j.neucom.2017.12.004

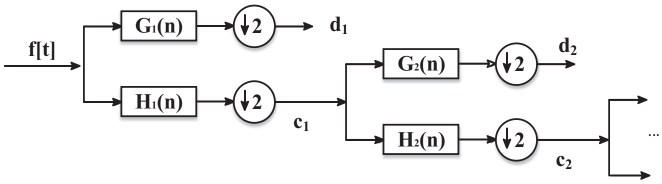

Fig. 1. Schematic diagram of discrete wavelet decomposition, G is the high pass filter coefficient, H is the low pass filter coefficient, $\downarrow 2$ denotes subsampling.

same objects. Therefore, the proposed methods with two kinds of features simultaneously are integrated to construct a new feature vectors using the feature fusion algorithm based on a locality preserving projections (LPP). Finally, the support vector machines (SVM) [6] and K-means method be used to classify the spikes.

2. Representation of spikes with multi-features

2.1. Feature extraction of wavelet coefficients

Wavelet transforms can be applied to any finite-energy function to map the function from the time domain into a compound time–frequency space. The neural spike is decomposed into an orthogonal function space with the wavelet transform, which consist of a scaling function and a wavelet function. The discrete wavelet transform (DWT) is a linear transformation that operates on a data vector whose length is an integer power of 2, transforming it into a numerically different vector of the same length, as shown in Fig. 1. The discrete wavelet decomposition of the low-frequency part can be continued to decompose into a lower-frequency part and a relatively high-frequency part according to the needs of the analysis [9,10].

For any spike, the wavelet decomposition is

$$f(t) = \sum_k c_{j,k}\phi_{j,k} + \sum_k d_{j,k}\psi_{j,k} \tag{1}$$

where $c_{j,k}$ is the kth lower frequency portion of scale j. $d_{j,k}$ is the kth higher frequency portion of scale j. $\phi_{j,\,k}$ is the scaling function and $\psi_{j,\,k}$ is the wavelet function, which can be expressed as

$$\begin{cases} \psi_{j,k}(t) = 2^{-\frac{j}{2}}\psi\left(2^{-j}t - k\right) \\ \phi_{j,k}(t) = 2^{-\frac{j}{2}}\phi\left(2^{-j}t - k\right) \end{cases} \tag{2}$$

where t is the sampling point of the spike, j is the scale parameter and k is the shift parameter. The coefficients of the scaling function and the wavelet function have the relationships, where, c is the lower frequency portion and d is higher frequency portion.

$$\begin{cases} c_{j-1,k} = \sum\limits_{l\in Z} h(l-2k)c_{j,k} \\ d_{j-1,k} = \sum\limits_{l\in Z} g(l-2k)d_{j,k} \end{cases} \tag{3}$$

This relationship indicates that both the wavelet function coefficients and scaling function coefficients under scale $j-1$ can be obtained through function coefficients under scale j. $h(l-2k)$ represents a low-pass filter, and $g(l-2k)$ represents a high-pass filter, l is the number of sampling points of spike [6].

Each spike waveform consists of 64 sampling points. As mentioned above, each spike waveform is decomposed into 6 scales using Symlet8 wavelets [8]. Symlet8 wavelets are chosen due to their compact support and orthogonality. Then, the distinguishing features of the spikes are expressed using a few wavelet coefficients. The wavelet coefficients represented with $[c_6, d_6, \ldots, d_2, d_1]$ can be obtained with six-level decomposition. c_6 is the approximation level, and $d_6, \ldots, d_2, d_1$ are the detail levels.

2.2. Feature extraction of PCA

Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables. The PCA method transforms a set of high dimensional vectors into a lower dimension space. A feature extraction method is used to determine an appropriate subspace mapping in the original feature space.

PCA features extraction is to characterize the maximum change direction of data sets using the ordered orthogonal basic vector set. Then, the original high-dimensional relativity vector can be mapped into a low-dimensional vector space by an eigenvector matrix. The new vector not only reflects the original information but also realize the irrelevance between new indicators [11].

The spikes are defined as $[X_1, X_2, \ldots, X_m]$, and each spike consists of an n-dimensional vector that can be expressed as $[x_{i1}, x_{i2}, \ldots, x_{in}]$. Then, the whole data set of spikes can be expressed as

$$\mathbf{X} = [X_1\ X_2 \ldots X_m]^T = \begin{bmatrix} x_{11}, & x_{12}, \ldots, x_{1n} \\ x_{21}, & x_{22}, \ldots, x_{2n} \\ \vdots & \vdots \quad \vdots \\ x_{m1}, & x_{m2}, \ldots, x_{mn} \end{bmatrix} \tag{4}$$

where m, n is number of spikes, the features number for each spike, respectively. The extracted PCA features $\mathbf{Y_p}$ can be expressed with the whole dataset of spikes as:

$$Y_p = XG \tag{5}$$

where matrix G is the linear mapped matrix, which is composed of the eigenvectors with the maximum eigenvalues of the $\mathbf{X}$' covariance matrix. The mapped vector of PCA is $Y_p = [Y_{p1}, Y_{p2}, \ldots, Y_{pn}]$, where Y_{p1} represents the indicator composed of the first linear combination of the original variables, that is $Y_{p1} = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1m}X_m$. The information quantity extracted from each main component can be indicated by the variance. Y_{p1} is the first principal component, and the mapped vector $\mathbf{Y}_\mathrm{p}$ can be defined as:

$$\mathbf{Y}_p = \begin{bmatrix} Y_{p1} \\ Y_{p2} \\ \vdots \\ Y_{pm} \end{bmatrix} = \begin{bmatrix} a_{11}x_{11} + a_{12}x_{12} + \cdots + a_{1n}x_{1n} \\ a_{21}x_{21} + a_{22}x_{22} + \cdots + a_{2n}x_{2n} \\ \vdots \quad \vdots \quad \vdots \\ a_{m1}x_{m1} + a_{m2}x_{m2} + \cdots + a_{mn}x_{mn} \end{bmatrix} \tag{6}$$

According to the analysis above, we can derive that Y_{pi} and Y_{pj} are uncorrelated, that is, $Cov(Y_{pi}, Y_{pj}) = 0$. Y_{pi} has the greatest variance among all the linear combinations of $X_1, X_2 \ldots X_m$ and is called the first principal component. $Y_{p2}, Y_{p3}, \ldots, Y_{pm}$ are the 2th, 3th, …, mth principal components, respectively [12].

The PCA features of the spikes are extracted, and a 32-dimensional PCA eigenvector is constructed by the first 32 dimensions of the main component. In the third step of the feature extraction, we obtain the eigenvalues of the sample data covariance matrix. In this paper, the first 32 eigenvalues represent more than 96% of the sum of all eigenvalues. Thus, that PCA eigenvector can characterize more than 96% of the amplitude information of the spikes [13].

3. Feature fusion

3.1. Multiple features

The high-frequency coefficients of the wavelet decomposition indicate the sharp information of the spikes. For the feature extraction with the PCA method, the amplitude of the spike is mapped into a new group of coordinates, resulting in an eigenvector that reflects the amplitude information.