Tetrode Spike Detection Method Based on Quaternion Principle Component Feature Extraction

Yong Zhao
Collegel of Electrical Engineering
Yanshan University
Qinhuangdao, P.R. China
e-mail: zhaoyong@ysu.edu.cn

Yibo Wang
College of Electrical Engineering
Yanshan University
Qinhuangdao, P.R. China
e-mail: 18033500490@163.com

Ailing Tan
College of Information Science and Engineering
Yanshan University
Qinhuangdao, P. R. China
The Key Laboratory for Special Fiber and Fiber Sensor of Herbei Province
Yanshan University
Qinhuangdao, P. R. China
e-mail:tanailing@ysu.edu.cn

Abstract—Multi-unit recording of tetrode has been used in spike detections for many years, but traditional feature extraction methods of spikes sorting based on analyzing correlations of single channel data don't consider relations of each channel data. A new feature extraction method based on quaternion principle component analysis (QPCA) algorithm used for sorting the tetrode spikes trains is presented. A quaternion vector formed of four-dimensional numbers can replace a set of tetrode spikes data. Firstly, the four channels of spikes were achieved using dual threshold detection method. We used the modulus values of the vectors grouped by 4 channels corresponding data to constitute the new features vectors of spikes through QPCA algorithm. The features vectors contain correlations of 4 channels. Our method fully fuses the information of tetrode data. Thus it has higher clustering accuracy than traditional feature extraction methods.

Keywords-tetrode; spike; spike detection; QPCA; K-means clustering

I. INTRODUCTION

In the last decades, multichannel electrode recordings have been proved to be a key technique in studying the population activities of the brain [1] [2]. Especially tetrode technology made it easier to record simultaneously from large numbers of neurons [3] [4]. But this also generated some new challenges in analysis of multichannel spike datasets.

Many spike separation and classification methods have been proposed. Yuan Yuan, and Chenhui Yang presented a robust and automatic spike detection and sorting system named the M-Sorter which was equipped with the multiple correlation of wavelet coefficients (MCWC) and a template matching classification [5]. Chaitanya Ekanadham presented a unified frame for automatic identification of neural spike, which used the Continuous Basis Pursuit (CBP) algorithm to classify spikes [6]. Quiroga presented a unsupervised classification method on the wavelet and super-paramagnetic clustering [7]. Hulata proposed a spike detection and sorting algorithm based on wavelet packets and Shannon's mutual information [8]. The Independent component analysis and SOM algorithm was applied to the multi electrode record spike classification of the central nervous system by Hermle [9]. Wood and Black applied nonparametric Bayesian alternative to spike of classification [10]. Chah E proposed Laplacian eige-nmaps and k-means clustering algorithm for automated spike classification [11].

All the proposed algorithms handle multichannel spikes with serial processing methods, which is a common drawback. They deal with individual single-channel spike data one by one, and then connect the features of each channel to constitute a long feature vector. These methods separate the link between multi-channels and losts the coupling information between the multichannels.

The quaternion is a member of noncommutative division algebra which was invented by William Rowan Hamilton. Recently quaternion-based signal processing methods have been proposed such as quaternion principle component analysis [12], quaternion independent component analysis [13], quaternion Fourier transform [14], quaternion Gabor filter etc [15] [16].These quaternion-based algorithms have offered many useful applications in pattern recognition, vector-sensor signal processing, colorful image processing, blind extraction, adaptive filtering [17-22].

This paper proposes a novel approach for tetrode data model and spike sorting based on the algorithm of quaternion signal processing. The paper is organized as follows. We first introduce the concept of quaternion and extend the principle component analysis (PCA) algorithm to the field of quaternion—quaternion principle component analysis (QPCA). Secondly, we propose our algorithm combined quaternion principle component analysis feature extraction and K-mean clustering method. Finally, we conduct simulation for tetrode spike databases to verify the effectiveness of our proposed algorithm.

II. Quaternion Principle Component Analysis

A brief introduction about quaternion and quaternion principle component analysis is presented in this section. The concept of the quaternion was introduced by Hamilton in 1843, which is a generalization of a complex number. The set of quaternion is denoted by H, which is a single example of a more general class of hypercomplex numbers. The quaternion is not commutative, it is associative, and it forms a group known as the quaternion group. By analogy with the complex numbers being indicated as a sum of real and imaginary parts, a quaternion can also be written as a linear combination

$$H=a+bi+cj+dk \tag{1}$$

The quaternion satisfies the following identities which known as Hamilton's rules,

$$i^2=j^2=k^2=-1 \tag{2}$$

$$ij=-ji=k \tag{3}$$

$$jk=-kj=i \tag{4}$$

$$ki=-ik=j \tag{5}$$

The quaternion can be represented in complex 2×2 matrices

$$H=\begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix}=\begin{bmatrix} a+ib & c+id \\ -c+id & a-ib \end{bmatrix} \tag{6}$$

where z and w are complex numbers, a, b, c, and d are real, and $\overline{z}$ is the complex conjugate of z.

Next, we introduce some quaternion matrix algebra tools. The decompositions of quaternion matrices as a PCA techniques are presented such as quaternion singular value decomposition, quaternion eigenvalue decomposition and quaternion Karhunen-L`oeve transform. N training samples build a matrix $S_{mn}=\left[Q_1 Q_2 \cdots Q_n\right]$. Each column is a quaternion vector of sample data and its length is $m(64\times 64)$. Given the matrix $S_{m\times n}$, its covariance is $C_{m\times n}$. However, calculating the eigenvectors and the eigenvalues of the matrix $C_{m\times n}$ is not easy for such a high dimension. In a small number of samples n, it is still easy to compute eigenvectors and eigenvalues of matrix $C_{m\times n}$.

$$C=\frac{1}{n-1}E^{T^*}E \tag{7}$$

$$E=S-\overline{S} \tag{8}$$

$$\overline{S(x,y)}=\frac{1}{n}\sum_{l=1}^{n}S(x,l), x=1\cdots m, y=1\cdots n \tag{9}$$

Where T^* is the conjugate-transposition operator for a quaternion matrix. $C_{n\times n}$ is a Hermitian matrix. so through Householder transformations, it could be tridiagonalized to obtain B, which is a real tridiagonal symmetric matrix, and P, which is used to compute matrix:

$$C_{n\times n}=P^{T^*}BP \tag{10}$$

Matrix B is a real matrix, so eigenvectors of B are easily calculated. V_B represents the eigenvectors matrix of B, each column of V_B is an eigenvector of matrix B. Thus eigenvectors of $C_{n\times n}$ can be calculated as:

$$V_C=P^{T^*}V_B \tag{11}$$

Eigenvalues of $C_{n\times n}$ are eigenvalues of matrix B as well. D_C is used to represent eigenvalues of $C_{n\times n}$ .Now we have the eigenvectors and the eigenvalues of matrix $C_{n\times n}$. Finally eigenvalues and eigenvectors of matrix $C_{m\times m}$ is computed by:

$$V=EV_C \tag{12}$$

$$D=\frac{n-1}{m-1}D_C \tag{13}$$

Where V consists of eigenvectors of $C_{m\times m}$, and D consists of eigenvalues of $C_{m\times m}$. Eigenvalues D is sorted in descending order, and an energy ratio is defined as:

$$\text{ratio}=\sum_{x=1}^{P}D_x \Big/ \sum_{x=1}^{n}D_x \times 100\% \tag{14}$$

Where D_X is the xth eigenvalue. Given an energy ratio, the number of the eigenvalues p can be calculated by using (14). We can get the projection matrix $\hat{P}$ by the first p eigenvectors in the matrix V. In general, the energy ratio is set to 90% for a balance between accuracy and feature dimension.

If there is an input quaternion sample s, the QPCA feature f_{QPCA} is calculated as:

$$f_{QPCA}=\hat{P}^{T^*}s \tag{15}$$

For more details on this subject, interested readers can consult other papers for complete overview and state in the field of quaternion in [23] [24].

III. Data Simulation

Firstly, we constructed simulated single channel spikes using a dataset of 126 different spike shapes. Three kinds of distinct spikes were selected randomly in the dataset and were superimposed on the background noises at random times. The amplitude of the selected spike was normalized to have a peak value of 1. The standard deviation of noise signals is equal to 0.05, 0.10 and 0.15 relative to the amplitude of the spikes. The three kinds of spikes on the basis of Poisson distribution of inter spike intervals with a mean firing rate is equal to 25Hz.

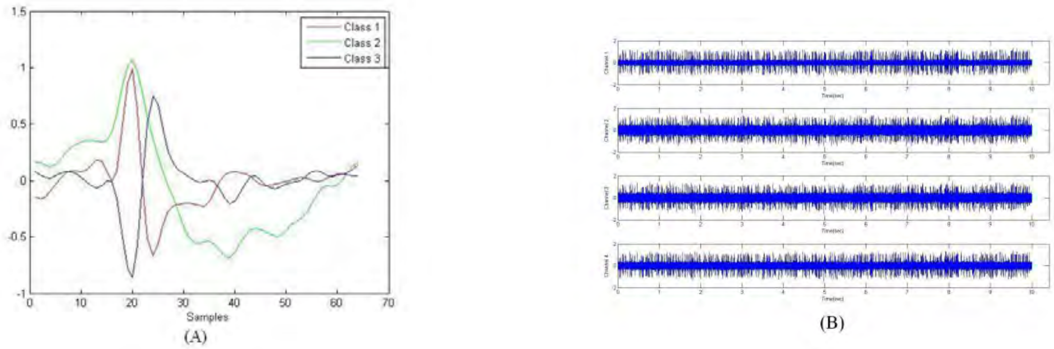

Figure 1. Waveforms of the Simulated Data

Fig .1 shows the simulated spike data with noise level 0.1. Fig .1 A shows the three kinds of distinct spike; Fig .1 B shows the simulated four channel spikes

IV. DESCRIPTION OF THE METHOD

(1) Reprocessing data; (2) quaternion principle component analysis algorithm; (3) K-mean cluster. Fig .2 gives the three fundamental steps of the algorithm. Each step will be explained in subsequent sections.

A. *Spike detection*

Spike detection is achieved with dual-threshold detection method. In this way, the false positive and missed spikes of threshold detection method in peak potential have a certain degree of reduction. According to spike waveforms with positive and negative characteristics of two-way voltage thresholds, we choose positive or negative thresholds. Assuming that x_i is the amplitude of the electrical signal peaks in each sampling point, n is the peak electrical sampling points. Commonly automatic threshold based on the standard deviation is used below:

$$Thr_i = a \times s\,,\ s=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}\,,\quad a=3\sim5 \qquad (16)$$

When a negative peak voltage signal is below of the negative threshold or positive peak voltage is higher than the positive threshold, the signals are considered to be a spike. Get all greater points than the threshold by setting a threshold to find where the points of description of the maximum amplitude method. Depending on the length of the spike and the sampling frequency we get each spike points. Due to the duration of a spike in 1-2ms, the length of the signal conforms spike duration, before judging it as a spike. We repeated the process until finding all the spikes.

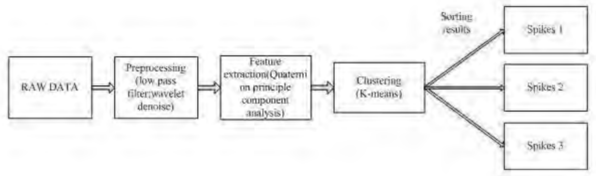

Figure 2. Description of the Method

The detected spikes are then up sampled based on the sampling theorem, and the spike waveforms were realigned by peak position.

B. *Feature vector extraction*

For the multichannel data, traditional method stacks each waveform matrix to yield a new vector, and concatenate stacked waveforms to obtain the multichannel data matrix. And then QPCA is used to get the low-dimensional representations.

Feature vectors extraction process has four steps:

Firstly, for each channel, the resampling waveforms with 64 point are aligning and corresponding points in four channels construct a quaternion. Thus, we construct quaternion spike matrix from the 4 channel spikes.

Secondly, we employ quaternion principal component analysis (QPCA) to get the first five quaternion principal components of the quaternion waveform set.

Thirdly, we use geometrical odd algorithm to create new feature vectors.

Finally, A 5-dimensional quaternion feature vector is then created for four channels spikes by projecting the multichannel spike waveforms onto the three principal components.

C. *Clustering*

Since quaternion principal components analysis feature extraction algorithm widely make use of the data of four channels to reflect the information. Subsequently selection of clustering algorithm is relatively simple. We choose K-means clustering algorithm to classify different spikes. This method defines the clustering locations as the mean of data. A spike is classified to whichever cluster has the closest mean based on Euclidean distance. Given a set of observations $(x_1, x_2, \ldots, x_n)$, where each observation is a D-dimensional real vector, K-means clustering aims to partition the n observations into k sets ($k \le n$) $S = \{S_1, S_2, \ldots, S_k\}$ so as to minimize the within-cluster sum of squares (WCSS):

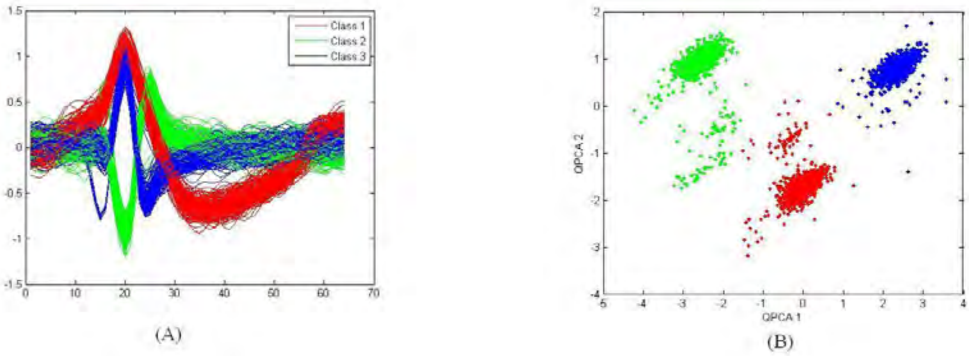

Figure 3. The Result of Spike Clustering

$$J=\sum_{j=1}^{k}\sum_{i=1}^{n}\left\|x_i^{(j)}-c_j\right\|^2 \tag{17}$$

where $\left\|x_i^{(j)}-c_j\right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j.

Fig .3 shows the result of spikes clustering, Fig .3 A shows waveforms of different kinds of spikes in different color.Fig .3 B is the scatter diagram of clustering result.

Though the number of clusters should correspond to the number of neurons ideally, but there are some factors that affect such a simple interpretation. So choosing the number of clusters is a most difficult problem. In this paper we choose a method which based on inter spike interval histogram to guide decisions about whether a class represents a single neuron

V. CONCLUSION

Tetrode recording technology has been widely applied in assessing simultaneous neural activity in neural scientific studies. Spike detection and classification is worth discussing in the field of neural signal analysis. In this paper, we propose a novel feature extraction method for multichannel spike data based on quaternion principle component algorithm. By using the data simulated with four channels spike, this method presents a better classification performance than traditional methods.

ACKNOWLEDGMENT

The authors are grateful to the support of Hebei Provincial Natural Science Foundation(No.F2014203245) and the support of Hebei Educational Committee Nature Science Youth Fund (No.Q2012129,No.QN2014034)

REFERENCE

[1] Brozoski T J, Caspary D M, Bauer C A. Marking multi-channel silicon-substrate electrode recording sites using radiofrequency lesions[J]. Journal of neuroscience methods, 2006, 150(2): 185-191..

[2] Branner A, Stein R B, Normann R A. Selective stimulation of cat sciatic nerve using an array of varying-length microelectrodes[J]. Journal of neurophysiology, 2001, 85(4): 1585-1594.

[3] Harris K D, Henze D A, Csicsvari J. Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements[J]. Journal of Neurophysiology, 2000, 84(1): 401-414.

[4] Takahashi S, Anzai Y, Sakurai Y. A new approach to spike sorting for multi-neuronal activities recorded with a tetrode—how ICA can be practical[J]. Neuroscience research, 2003, 46(3): 265-272.

[5] Yuan Y, Yang C, Si J. The M-Sorter: An automatic and robust spike detection and classification system[J]. Journal of neuroscience methods, 2012, 210(2): 281-290.

[6] Ekanadham C, Tranchina D, Simoncelli E P. A unified framework and method for automatic neural spike identification[J]. Journal of neuroscience methods, 2014, 222: 47-55.

[7] Quiroga R Q, Nadasdy Z, Ben-Shaul Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering[J]. Neural computation, 2004, 16(8): 1661-1687.

[8] Hulata E, Segev R, Ben-Jacob E. A method for spike sorting and detection based on wavelet packets and Shannon's mutual information[J]. Journal of Neuroscience methods, 2002, 117(1): 1-12.

[9] Hermle T, Schwarz C, Bogdan M. Employing ICA and SOM for spike sorting of multielectrode recordings from CNS[J]. Journal of Physiology-Paris, 2004, 98(4): 349-356.

[10] Wood F, Black M J. A nonparametric Bayesian alternative to spike sorting[J]. Journal of neuroscience methods, 2008, 173(1): 1-12.

[11] Chah E, Hok V, Della-Chiesa A. Automated spike sorting algorithm based on Laplacian eigenmaps and k-means clustering[J]. Journal of neural engineering, 2011, 8(1): 016006.

[12] Le Bihan N, Sangwine S J. Quaternion principal component analysis of color images[C]//Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on. IEEE, 2003, 1: I-809-12 vol. 1.

[13] Le Bihan N, Buchholz S, ENSIEG B P. Quaternionic independent component analysis using hypercomplex nonlinearities[C]//Proc. 7th IMA Conf. Math. Signal Process. 2006: 1-3.

[14] Guo C, Ma Q, Zhang L. Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform[C]//Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008: 1-8.

[15] Bulow T, Sommer G. Quaternionic Gabor filters for local structure classification[C]//Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on. IEEE, 1998, 1: 808-810.

[16] ones III C F. Color face recognition using quaternionic Gabor filters[D]. Virginia Polytechnic Institute and State University, 2004.

[17] Guo C, Ma Q, Zhang L. Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform[C]//Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008: 1-8.

[18] Le Bihan N, Mars J. Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing[J]. Signal processing, 2004, 84(7): 1177-1199.

[19] Pei S C, Cheng C M. Color image processing by using binary quaternion-moment-preserving thresholding technique[J]. Image Processing, IEEE Transactions on, 1999, 8(5): 614-628.

[20] Bas P, Le Bihan N, Chassery J M. Color image watermarking using quaternion Fourier transform[C]//Acoustics, Speech, and

Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on. IEEE, 2003, 3: III-521-4 vol. 3.

[21] Javidi S, Took C C, Jahanchahi C. Blind extraction of improper quaternion sources[C]//Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on. IEEE, 2011: 3708-3711.

[22] Took C C, Mandic D P. A quaternion widely linear adaptive filter[J]. Signal Processing, IEEE Transactions on, 2010, 58(8): 4427-4431.

[23] Le Bihan N, Sangwine S J. Quaternion principal component analysis of color images[C]//Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on. IEEE, 2003, 1: I-809-12 vol. 1.

[24] Xu X, Guo Z, Song C. Multispectral palmprint recognition using a quaternion matrix[J]. Sensors, 2012, 12(4): 4633-4647.