Available online at www.sciencedirect.com

SCIENCE DIRECT®

NEUROCOMPUTING

Neurocomputing 69 (2006) 1380–1384

www.elsevier.com/locate/neucom

A spike sorting framework using nonparametric detection and incremental clustering

Mingzhou (Joe) Song[a,*], Hongbin Wang[b]

[a]Department of Computer Science, New Mexico State University, Las Cruces, NM 88003, USA

[b]Doctoral Program in Computer Science, Graduate Center, City University of New York, NY 10016, USA

Available online 3 March 2006

Abstract

We introduce a statistical computing framework to address two important issues in spike sorting: flexible spike shape modeling and realtime spike clustering. In this framework, spikes are detected based on a nonparametric shape distribution; detected spikes are further grouped by an incremental clustering algorithm involving the second-order statistics–covariance matrix. We performed experiments on both simulated and real signals to study spike detection accuracy and cluster separation.

Keywords: Spike sorting; Nonparametric modeling; Incremental clustering

1. Introduction

Spike sorting is the process of detecting action potentials from extracellular signals and assigning them to individual neurons. Early development aims at assisting researchers in studying brain functions off-line. Recent applications include brain–computer interfaces and neural prostheses [11] for people suffering from nervous system traumas. These efforts all build on the capability of accurate automatic decoding of neuronal signals, which imposes a statistical computing task replete with open problems [2].

Most recent work in spike detection [5,8–10] assumes a parametric form or a template for spike shapes. Some methods require exact knowledge of spike shapes such as matched filters; others utilize a combination of shape bases. For example, nine wavelet bases are manually constructed in [5] after analyzing real spikes. Some quantitative performance studies have been reported. The morphological filter [8] achieves a correct detection rate of 80 ± 4% on data from a simulated cortex containing 90 neurons. The wavelet transform approach [5] obtains a correct detection rate of 93% and a false alarm rate of 10%. However, the method was not evaluated on independent test data. An opportunity is to find a flexible and realistic representation for spike shapes beyond templates.

Realtime spike clustering is another hard challenge [2]. In order for neural prosthetic devices and brain–computer interfaces to be practical, it is essential to decode spike signals in realtime. In spite of substantial progress in algorithms [1,3] for clustering points arrived in a data stream, no work has been established for realtime spike clustering.

We introduce a statistical computing framework that supports a two-phase statistical spike sorting strategy. During the first phase, signal segments are detected for spikes using a grid representation of probability density functions (p.d.f.s). Uniform quantization has been applied on a spectral representation of spike features and on time [7], but no probability is associated. In the second phase, an incremental clustering algorithm is employed to group detected spikes from the first phase. Experiments on simulated and real spike signals show encouraging results for spike sorting under the framework to be described.

2. Spike detection with nonparametric shape modeling

Spike detection is the process of selecting neuronal action potentials from background noises. A real signal in Fig. 1 illustrates the ambiguousness of spikes.

*Corresponding author.
E-mail addresses: joemsong@cs.nmsu.edu (M.(J.) Song), hwang2@gc.cuny.edu (H. Wang).

0925-2312/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2005.12.111

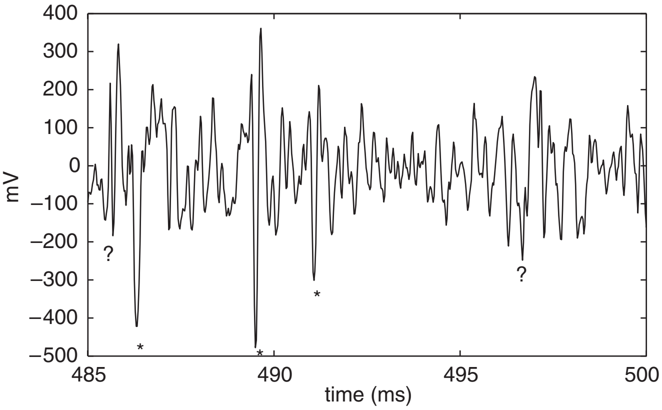

Fig. 1. A real signal, where "*" indicates spikes and "?" for possible ones.

To overcome strong constraints imposed by parametric or template models on spike shapes, we describe a nonparametric spike detection approach under a Bayesian framework. In this framework, the posterior probability of a signal segment X being a spike, given the shape of X and the repetitiveness of X, is used as an indicator for a spike. A segment X is obtained by finding the maximum value within a window and then taking two chunks of discrete signal before and after the maximum value.

A random vector Z and a random variable R, both defined in the probability-theoretical sense, represent the shape and the repetitiveness of X, respectively. Although they are statistically dependent, we consider Z and R conditionally independent for a given neuron. This is a strong statistical assumption, but it is justifiable for spike signals. For a given neuron, its firing rate is not statistically related to its spike shape. Accordingly, the posterior probability can be written as

$$\begin{aligned}&P(X \text{ is a spike}|Z,R)\\&\quad\propto P(Z,R|X \text{ is a spike})P(X \text{ is a spike})\\&\quad= P(Z|X \text{ is a spike})P(R|X \text{ is a spike})P(X \text{ is a spike}).\end{aligned} \tag{1}$$

The prior probability $P(X$ is a spike) is provided by a domain expert. The two conditional p.d.f.s are computed, respectively, by a nonparametric quantization algorithm using a multi-dimensional nonuniform grid. The grid is estimated by maximizing a performance measure, which is defined by the data log likelihood plus an entropy penalty [12].

$P(Z|X$ is a spike) suggests how probable it is to observe the vector Z when X is indeed a spike. It is learned by off-line training. We gathered groundtruth spikes by setting a very high threshold on many real signals. We consider those signal segments that cross the threshold to be true spikes. Each spike unit lasts about 1 ms. At the sampling rate of 40 kHz, a spike is quantified to a vector of 400 dimensions. Then, we reduce the dimensions to 4 by principal component analysis, which preserves 90% of the variance. The marginal entropies—computed automatically using marginal histograms—determine the relative quantization levels of each dimension. An estimated $P(Z|X$ is a spike), by grid quantization, is shown in Fig. 2(a).

The repetitiveness of a segment is a critical factor in spike sorting because a neuron consistently fires spikes in similar shape. We perform grid quantization again after the signal segment space is reduced to a 4-D space. We measure the repetitiveness R by the density value of a segment X. Fig. 2(b) shows the marginals of an estimated p.d.f. of signal segments. Eventually, we compute the conditional p.d.f. of repetitiveness by $P(R|X$ is a spike$) \propto R$.

Spike detection is achieved via a Bayesian hypothesis test on the ratio

$$P(X \text{ is a spike}|Z,R)/P(X \text{ is background}|Z,R), \tag{2}$$

where we apply a uniform distribution for $P(Z,R|X$ is background).

3. Realtime spike clustering without historical signals

A static clustering method uses the entire recorded signal to group detected spikes, which is infeasible for realtime spike sorting. We use a density-based data stream clustering method [13] relying on the second-order statistics, i.e., the covariance, to cluster spikes arriving in realtime. Sliding a time window, it incrementally keeps track of all clusters using only newly arrived data in a current time window. Within each sliding window, clusters are modeled as a Gaussian mixture distribution. The number of components in a mixture model is selected by the Bayesian information criterion. Each mixture model is identified through the expectation maximization algorithm. Newly discovered clusters are compared to and may be merged with historical clusters by a strategy via multivariate statistical tests for equality of covariance and mean.

3.1. Equality-of-covariance test

Let $x_1, x_2, \ldots, x_n \in \mathbf{R}^d$ be a d-dimensional sample of size n in a cluster of the current window. Let Σ_x be the covariance matrix of this cluster. We use Σ_0 to represent the covariance matrix of a historical cluster. In this test, we determine if Σ_x is statistically equal to Σ_0. The null hypothesis H_0 is $\Sigma_x = \Sigma_0$. The data are first transformed by $Y = L_0^{-1}X$, where L_0 is a lower triangular matrix obtained by Cholesky decomposition of Σ_0. An equivalent null hypothesis H_0' becomes $\Sigma_y = I$, where Σ_y is the covariance matrix of Y and I is d-dimensional identity matrix. The W statistic [6] induces a method to achieve the test without inverting the sample covariance matrix, defined by

$$W = \frac{1}{d}\mathbf{tr}[(S_y - I)^2] - \frac{d}{n}\left[\frac{1}{d}\mathbf{tr}(S_y)\right]^2 + \frac{d}{n}, \tag{3}$$

where S_y is the sample covariance matrix of Y and $\mathbf{tr}(\cdot)$ is the trace of a matrix. Under the null hypothesis of covariance equality, the statistic $(nW - d)d/2 + d$ is

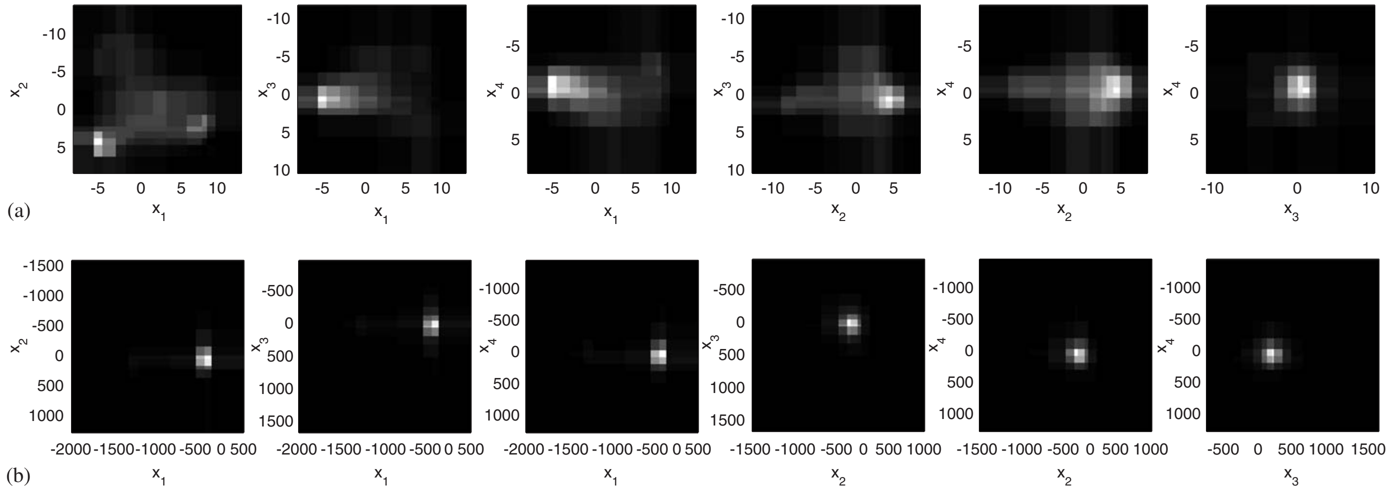

Fig. 2. Two-dimensional marginals of the two 4-D grid p.d.f.s for spike shapes and signal segments, respectively. Each dimension corresponds to one of the first four principal components for spike shapes and signal segments, respectively. (a) Two-dimensional marginals of $P(Z|X$ is a spike) for spike shapes. (b) Two-dimensional marginals of the p.d.f. for all segments in a signal as a measure of repetitiveness.

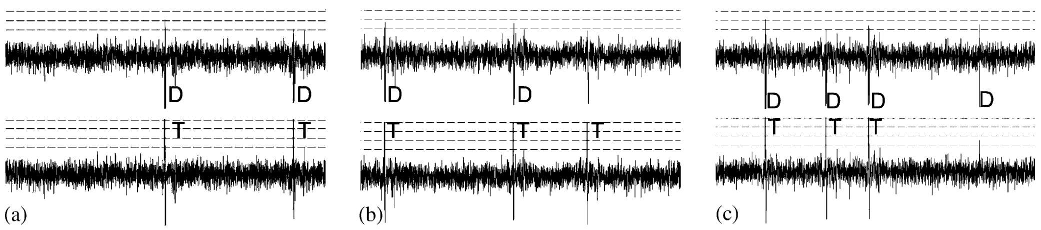

Fig. 3. Spike detection on a simulated signal. The letter D indicates the peak location of a detected spike; the letter T marks the peak location of a groundtruth spike. (a) Correct detections. (b) A miss-detection. (c) A false-alarm.

asymptotically χ^2 distributed with $d(d+1)/2$ degrees of freedom.

3.2. Equality-of-mean test

If the two covariance matrices are statistically equivalent, equality of mean is further checked using Hotelling's T^2 statistic [4] $n(\bar{x}-\mu_0)^\top S_x^{-1}(\bar{x}-\mu_0)$, where S_x and $\bar{x}$ are the sample covariance and mean of a cluster, and μ_0 is the mean of another cluster with equal covariance. Under the null hypothesis of mean equality, $((n-d)/d(n-1))T^2$ has F distribution with d numerator and $n-d$ denominator degrees of freedom. In case of singularity of S_x due to small sample size, we use Σ_0 to replace S_x.

If they pass both tests, the two components are merged to create a new one in the updated model for all clusters. Each remaining component in the current time window is added to the updated model. New components in the updated model are merged further if they are equivalent. The mean and covariance of a merged cluster can be estimated directly [13] from the covariance and means of the two clusters as if done with all historical data. Thus, past data are not needed for the processing, which enables clustering in realtime without historical signals.

4. Experimental results on simulated and real spike signals

We simulated a spike signal lasting 7 s at a 40 kHz sampling rate, using real spikes as templates. The spike detection result is shown in Fig. 3. In this example, there were 15 mis-detected and 5 false-alarm spikes among a total of 165 spike events, corresponding to a recall of 91% and a precision of 97%. The average time difference between the matched spikes is 0.068 ms with a standard error of 0.41 ms.

We performed spike detection on a signal recorded at 40 kHz from a monkey (by courtesy of Plexon Inc.). Fig. 4 presents the result at different time scales, which is consistent with what an expert would expect.

The incremental clustering result, displayed in Fig. 5, accomplished good separation on a real spike signal.

However, we found that some spatially close clusters likely from the same neuron were not merged, e.g. the diamond and triangle clusters. Although they have different densities, it appears that a Gaussian or an elliptical distribution would emerge if they were combined. We are resolving this issue by working beyond the second-order statistics.

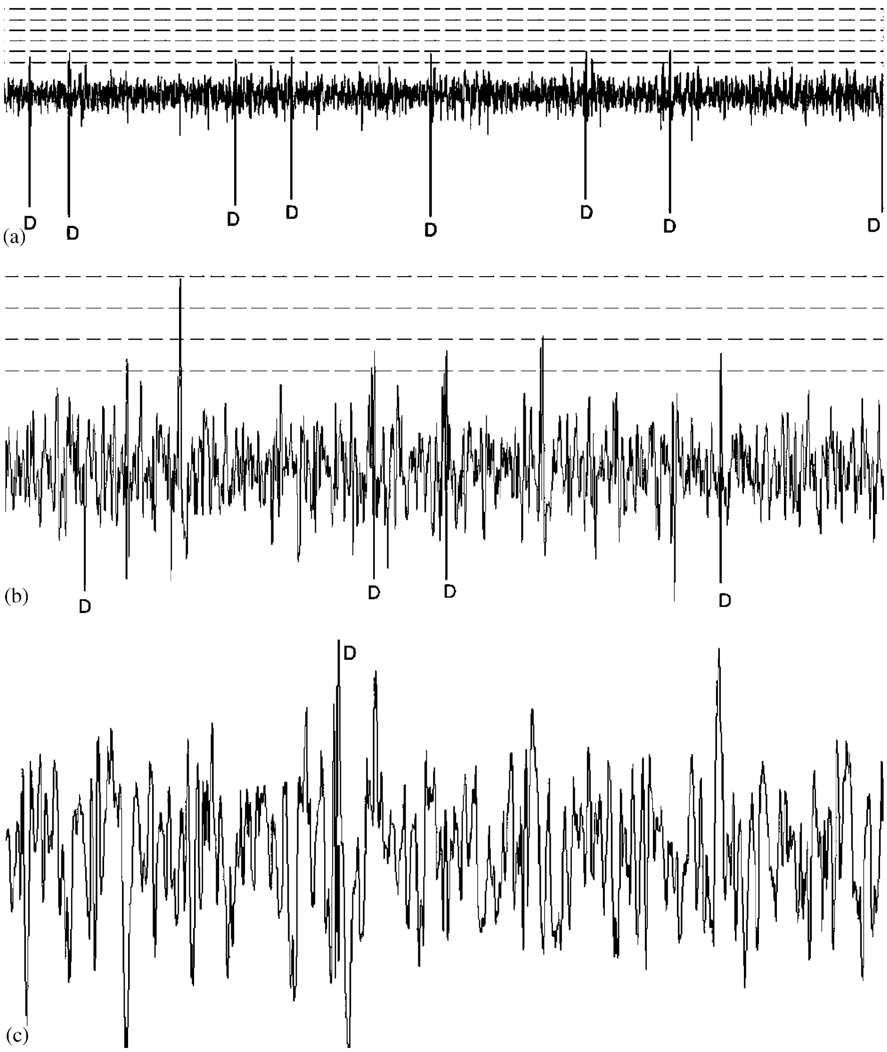

Fig. 4. Detected spike on a real signal. The letter D indicates spike peak locations. (a) Detected spikes. (b) A closer look. (c) An even closer look.

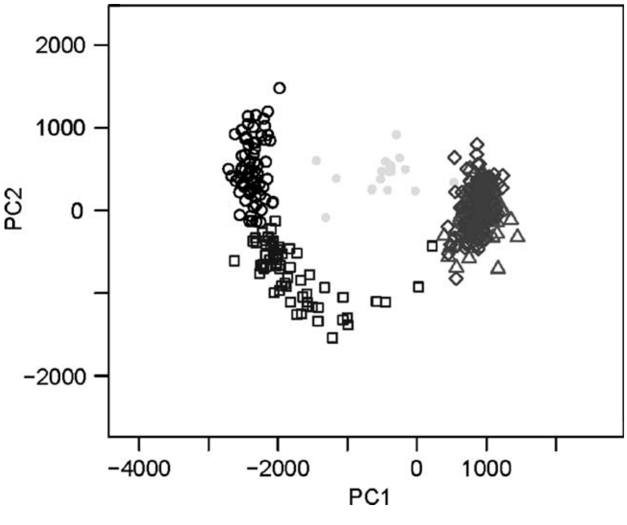

Fig. 5. Five spike clusters from a real signal, shown in two of the four dimensions.

5. Conclusion and future work

The nonparametric approach to spike detection dispenses with the requirement of templates or parametric forms for a spike in question. This gain is garnished with a nonparametric p.d.f. of spike shapes. The incremental strategy to realtime spike clustering is a major response to the demand from neural prosthetic applications. It facilitates reliable decoding using the complete second-order statistics.

There are plenty of opportunities to extend the reported work further. The incremental clustering can be enhanced by introducing higher-order statistics into cluster representation, for which we are establishing a theoretical framework. In a neural prosthetic device, the recordings are multi-channel with potential overlapping among neuronal spikes. It apparently poses great challenges to both nonparametric spike detection and incremental clustering.

Acknowledgment

The reported work was partially supported by a grant from CUNY Institute for Software Design and Development. The authors thank Plexon Inc. for providing neuronal spike data sets to conduct this research.

References

[1] C.C. Aggarwal, J. Han, J. Wang, P.S. Yu, A framework for clustering evolving data streams, in: Proceedings of the 29th International Conference on Very Large Data Bases, 2003.

[2] E.N. Brown, R.E. Kass, P.P. Mitra, Multiple neural spike train data analysis: state-of-the-art and future challenges, Nat. Neurosci. 7 (5) (2004) 456–461.

[3] S. Guha, A. Meyerson, N. Mishra, R. Motwani, L. O'Callaghan, Clustering data streams: theory and practice, IEEE Trans. Knowl. Data Eng. 15 (3) (2003) 515–528.

[4] H. Hotelling, The generalization of student's ratio, Ann. Math. Stat. 2 (1931) 360–378.

[5] E. Hulata, R. Segev, E. Ben-Jacob, A method for spike sorting and detection based on wavelet packets and Shannon's mutual information, J. Neurosci. Methods 117 (2002) 1–12.

[6] O. Ledoit, M. Wolf, Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size, Ann. Stat. 30 (4) (2002) 1081–1102.

[7] A. Luczak, N.S. Narayanan, Spectral representation—analyzing single-unit activity in extracellularly recorded neuronal data without spike sorting, J. Neurosci. Methods 144 (1) (2005) 53–61.

[8] K.M.L. Menne, A. Folkers, T. Malian, R. Maex, U.G. Hofmann, Test of spike sorting algorithms on the basis of simulated network data, Neurocomputing 44–46 (C) (2002) 1119–1126.

[9] K.G. Oweiss, D.J. Anderson, Spike sorting: a novel shift and amplitude invariant technique, Neurocomputing 44–46 (C) (2002) 1133–1139.

[10] M. Sahani, Latent variable models for neural data analysis, Ph.D. Thesis, California Institute of Technology, Pasadena, California, May 1999.

[11] M.D. Serruya, J.P. Donoghue, Section 7.9: design principles of a neuromotor prosthetic device, in: K.W. Horch, G.S. Dhillon (Eds.), Neuroprosthetics: Theory and Practice, World Scientific, New Jersey, 2004, pp. 1158–1196.

[12] M. Song, R.M. Haralick, Optimal multidimensional quantization for pattern recognition, in: Proceedings of SPIE, vol. 4875; Second International Conference on Image and Graphics, Part 1, Hefei, China, 2002, pp. 15–30.

[13] M. Song, H. Wang, Highly efficient incremental estimation of Gaussian mixture models for online data stream clustering, in: K.L. Priddy (Ed.), Proceedings of SPIE: Intelligent Computing—Theory and Applications III, vol. 5803, 2005, pp. 174–183.

Mingzhou (Joe) Song received his B.S. degree in Electrical Engineering from Beijing University of Posts and Telecommunications in 1992. He obtained his M.S. and Ph.D. degrees in Electrical Engineering from the University of Washington at Seattle in 1999 and 2002, respectively. He was assistant professor of computer science at Queens College, City University of New York from 2002 to 2005. Since August 2005, he is with the Department of Computer Science at New Mexico State University as an assistant professor. His research interests include statistical computing, computational neuroscience, computational molecular biology, and computer vision.

Hongbin Wang received her B.S. degree in Engineering from Qingdao University in China in 1994. She obtained her M.A. degree in Computer Science from Queens College, City University of New York in 2002. She is currently a doctoral student in the Computer Science Department at the City University of New York Graduate Center. Her research focuses on pattern discovery in data streams.