Journal of Neuroscience Methods 238 (2014) 43–53

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Computational Neuroscience

Spike sorting using locality preserving projection with gap statistics and landmark-based spectral clustering

Thanh Nguyen*, Abbas Khosravi, Douglas Creighton, Saeid Nahavandi

Centre for Intelligent Systems Research (CISR), Deakin University, Waurn Ponds Campus, Victoria 3216, Australia

HIGHLIGHTS

- An automatic unsupervised spike sorting method is proposed.
- The method uses locality preserving projection (LPP) algorithm for feature extraction.
- LPP features serve as inputs for the landmark-based spectral clustering (LSC) method.
- LPP–LSC is highly accurate and computationally inexpensive spike sorting.
- LPP–LSC can be applied into real-time spike analysis.

ARTICLE INFO

Article history:
Received 18 March 2014
Received in revised form 23 July 2014
Accepted 12 September 2014
Available online 26 September 2014

Keywords:
Spike sorting
Feature extraction
Locality preserving projection
Landmark-based spectral clustering
Gap statistics
Wavelet transformation
Superparamagnetic clustering

ABSTRACT

Background: Understanding neural functions requires knowledge from analysing electrophysiological data. The process of assigning spikes of a multichannel signal into clusters, called spike sorting, is one of the important problems in such analysis. There have been various automated spike sorting techniques with both advantages and disadvantages regarding accuracy and computational costs. Therefore, developing spike sorting methods that are highly accurate and computationally inexpensive is always a challenge in the biomedical engineering practice.
New method: An automatic unsupervised spike sorting method is proposed in this paper. The method uses features extracted by the locality preserving projection (LPP) algorithm. These features afterwards serve as inputs for the landmark-based spectral clustering (LSC) method. Gap statistics (GS) is employed to evaluate the number of clusters before the LSC can be performed.
Results: The proposed LPP–LSC is highly accurate and computationally inexpensive spike sorting approach. LPP spike features are very discriminative; thereby boost the performance of clustering methods. Furthermore, the LSC method exhibits its efficiency when integrated with the cluster evaluator GS.
Comparison with existing methods: The proposed method's accuracy is approximately 13% superior to that of the benchmark combination between wavelet transformation and superparamagnetic clustering (WT–SPC). Additionally, LPP–LSC computing time is six times less than that of the WT–SPC.
Conclusions: LPP–LSC obviously demonstrates a win–win spike sorting solution meeting both accuracy and computational cost criteria. LPP and LSC are linear algorithms that help reduce computational burden and thus their combination can be applied into real-time spike analysis.

1. Introduction

Neuroscience practice extracellularly records the activity of single neurons using thin electrodes implanted in the brain. Neurons in the vicinity of the electrode tip are picked up by the extracellular recordings and thus there is a demand to determine which spike corresponds to which neuron (Fig. 1).

Neurons, which are picked up by the same electrode, can fire in response to different activities. Even when nearby neurons have similar responses, it is important to distinguish them and observe their individual tuning properties, firing characteristics, and relationship with other neurons. Spike sorting refers to the process that assigns the detected spikes of a multichannel signal into clusters based on the similarity of their shapes.

* Corresponding author at: Centre for Intelligent Systems Research (CISR), Deakin University, Waurn Ponds Campus, 75 Pigdons Road, Waurn Ponds, Victoria 3216, Australia. Tel.: +61 3 52278281; fax: +61 3 52271046.
E-mail address: thanh.nguyen@deakin.edu.au (T. Nguyen).

http://dx.doi.org/10.1016/j.jneumeth.2014.09.011

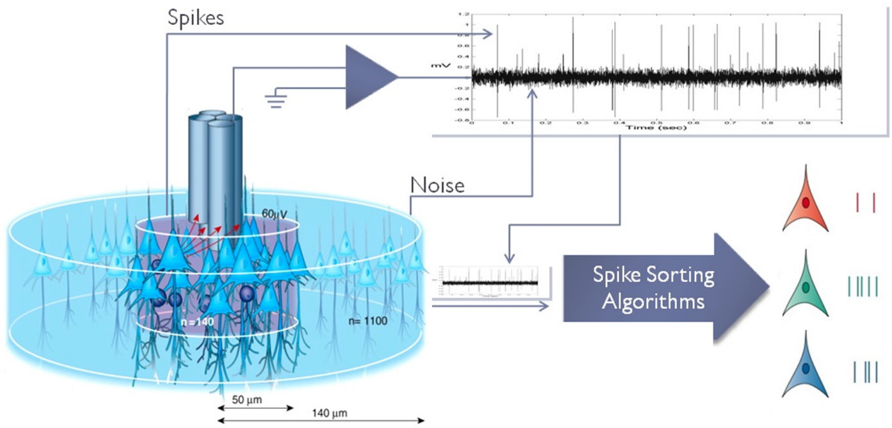

Fig. 1. Extracellular recordings and spike sorting (Buzsáki, 2004; Adamos et al., 2012).

In the literature, there exist a number of methods from machine learning or statistical mechanics dealing with neural spike analysis in general or spike sorting in particular. Lewicki (1998) and Brown et al. (2004) reviewed state-of-the-art techniques and challenges in analysing neural spike training data.

Alternatively, a method that combines wavelet transformation (WT) with superparamagnetic clustering without assumptions such as low variance or Gaussian distributions was proposed in Quiroga et al. (2004). Vollgraf et al. (2005) presented a spike sorting application that uses an optimal linear filter to reduce the distortions of the peak amplitudes of action potentials in extracellular multitrode recordings.

Hill et al. (2011) on the other hand recommended that four quality metrics of false-positive and false-negative errors should accompany spike sorting regardless of the algorithm used to sort. These metrics would facilitate the assessment regarding the performance of the sorter relative to the level of contamination of the data. Oliynyk et al. (2012) constructed a new tool for fast and robust online classification of single neuron activity based on the fuzzy c-mean clustering. The method is particularly useful for the analysis of large parallel recordings, which are practically impossible or inconvenient for human supervision, and thus is helpful in the decoding of neural ensembles or other clinical applications.

An algorithm for automatic unsupervised detection of action potentials in extracellular recordings was introduced in Shalchyan et al. (2012). A new manifestation variable for detection is defined based on the combination of denoised wavelet coefficients over selected scales. Tiganj and Mboup (2012) used an iterative application of independent component analysis and a deflation technique in two nested loops for spike sorting with multi-channel recordings. Each loop of the algorithm improves the final sorting results and thus significantly increases the overall spike sorting performance.

More recently, Pillow et al. (2013) investigated the geometry of failures of traditional spike sorting algorithms and developed a sorting model, which explicitly accounts for the superposition of spike waveforms. Otherwise, a divide and conquer approach for spike sorting, which uses a modified gradient ascent clustering algorithm, was examined in Swindale and Spacek (2014). Ekanadham et al. (2014) in another approach investigated a unified sparse estimation methodology for spike sorting that iteratively optimizes both the waveform shapes and their respective spikes.

Though various methods have been suggested, obtaining high accuracy in spike sorting is always a big challenge in neuroscience and biomedical engineering. Furthermore, the computational burden in spike sorting is massive. This paper presents an integrated approach, which combines locality preserving projection (LPP) (He and Niyogi, 2004), gap statistics (GS) (Tibshirani et al., 2001), and landmark-based spectral clustering (LSC) (Chen and Cai, 2011), for a computationally inexpensive unsupervised spike sorting method. According to our best knowledge, this is the first proposal on application of the LPP for spike feature extraction, GS in determining the number of clusters, and LSC for clustering spike sorting data. The accuracy of the proposed approach is compared to the renowned benchmark spike sorting method that is a combination between the wavelet transformation (WT) and the superparamagnetic clustering (SPC) in Quiroga et al. (2004).

The arguments are organized as follows. The next section describes details of steps in the proposed methodology. Section 3 is devoted for experiments and results whilst discussions and concluding remarks are presented in Sections 4 and 5 respectively.

2. Spike sorting methodology

The proposed methodology is graphically illustrated in Fig. 2 where LPP method is employed for spike feature extraction. The automatic unsupervised clustering is deployed by a combination of GS and LSC.

The first step in the methodology is spike detection, which aims to identify data points that form an action potential. The voltage threshold detection is utilized where the automatic threshold (*Thr*) is set to:

$$Thr = 4\,median\left\{\frac{|x|}{0.6745}\right\} \tag{1}$$

where x is the bandpass-filtered signal. For each detected spike, 64 samples are assembled for further process. Details of other steps are described in the following subsections.

2.1. Feature extractions

Feature extraction is one of the most important steps in which the silent features of the spikes are derived based on spike wave shapes. The features should be able to well differentiate spikes of different neurons and preferably low-dimensional. Simple features like peak-to-peak amplitude, maximum spike amplitude and spike width can be used (Gibson et al., 2012). These approaches however are sensitive to noise and intrinsic variations in spike shapes. Alternatively, principal component analysis (PCA) is one of the popular methods used for feature extraction (Jung et al., 2006; Tiganj and Mboup, 2011; Wild et al., 2012). WT also has emerged as a competitive feature extraction method for spike sorting (Chan et al., 2010; Yang et al., 2011; Lai et al., 2011; Yuan et al., 2012). For ease of comparison, we briefly present both WT and the suggested method LPP in the following subsections.