

Automatic spike sorting for extracellular electrophysiological recording using unsupervised single linkage clustering based on grey relational analysis

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2011 J. Neural Eng. 8 036003

(<http://iopscience.iop.org/1741-2552/8/3/036003>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 140.113.38.11

This content was downloaded on 24/04/2014 at 23:53

Please note that [terms and conditions apply](#).

Automatic spike sorting for extracellular electrophysiological recording using unsupervised single linkage clustering based on grey relational analysis

Hsin-Yi Lai¹, You-Yin Chen^{1,2,11}, Sheng-Huang Lin^{3,4}, Yu-Chun Lo⁴, Siny Tsang⁵, Shin-Yuan Chen⁶, Wan-Ting Zhao⁷, Wen-Hung Chao⁸, Yao-Chuan Chang¹, Robby Wu⁹, Yen-Yu I Shih¹⁰, Sheng-Tsung Tsai⁵ and Fu-Shan Jaw³

¹ Department of Electrical Engineering, National Chiao Tung University, No 1001, Ta-Hsueh Rd, Hsinchu, Taiwan 300, Republic of China

² Department of Biomedical Engineering, National Yang Ming University, No 155, Sec. 2, Linong St, Taipei, Taiwan 112, Republic of China

³ Department of Neurology, Tzu Chi General Hospital, Tzu Chi University, No 707, Sec. 3, Chung Yang Rd, Hualien, Taiwan 970, Republic of China

⁴ Institute of Biomedical Engineering, College of Medicine, National Taiwan University, No 1, Sec. 1, Jen-Ai Rd, Taipei, Taiwan 100, Republic of China

⁵ Department of Psychology, University of Virginia, No 102 Gilmer Hall, PO Box 400400, Charlottesville, VA 22904-4400, USA

⁶ Department of Neurosurgery, Tzu Chi General Hospital, Tzu Chi University, No 707, Sec. 3, Chung Yang Rd, Hualien, Taiwan 970, Republic of China

⁷ Department of Applied Science, National Hsin Chu University of Education, No 521, NanDa Rd, Hsinchu City, Taiwan 300, Republic of China

⁸ Department of Biomedical Engineering, Yuanpei University, No 306, Yuanpei St, Hsinchu, Taiwan 300, Republic of China

⁹ Philadelphia College of Osteopathic Medicine, 4170 City Avenue, Philadelphia, PA 19131, USA

¹⁰ Research Imaging Institute, University of Texas Health Science Center at San Antonio, 8403 Floyd Curl Drive, San Antonio, TX 78229-3900, USA

E-mail: irradiance@so-net.net.tw

Received 28 October 2010

Accepted for publication 14 January 2011

Published 4 April 2011

Online at stacks.iop.org/JNE/8/036003

Abstract

Automatic spike sorting is a prerequisite for neuroscience research on multichannel extracellular recordings of neuronal activity. A novel spike sorting framework, combining efficient feature extraction and an unsupervised clustering method, is described here. Wavelet transform (WT) is adopted to extract features from each detected spike, and the Kolmogorov–Smirnov test (KS test) is utilized to select discriminative wavelet coefficients from the extracted features. Next, an unsupervised single linkage clustering method based on grey relational analysis (GSLC) is applied for spike clustering. The GSLC uses the grey relational grade as the similarity measure, instead of the Euclidean distance for distance calculation; the number of clusters is automatically determined by the elbow criterion in the threshold-cumulative distribution. Four simulated data sets with four noise levels and electrophysiological data recorded from the subthalamic nucleus of eight patients with Parkinson's disease during deep brain stimulation surgery are used to evaluate the performance

¹¹ Author to whom any correspondence should be addressed.

of GSLC. Feature extraction results from the use of WT with the KS test indicate a reduced number of feature coefficients, as well as good noise rejection, despite similar spike waveforms. Accordingly, the use of GSLC for spike sorting achieves high classification accuracy in all simulated data sets. Moreover, J -measure results in the electrophysiological data indicating that the quality of spike sorting is adequate with the use of GSLC.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Extracellular recording of brain activity is an important method for examining the nervous system. The recorded signals typically consist of action potentials (i.e. spikes) from several neurons adjacent to the electrode site; however, the signals are often contaminated by noise from distant neuron activities, as well as from the signal recording system (Lewicki 1998). In order to understand the electrophysiological process in neuronal networks, it is necessary to isolate spikes from the background noise, extract crucial features from the detected spike waveforms, and correctly assign each spike to its originating neuron using a clustering algorithm (Wood *et al* 2004, Chapin 2004, Rutishauser *et al* 2006, Takekawa *et al* 2010). This process is universally known as spike sorting, which is essential for neurophysiological studies.

In conventional detection methods, spikes are usually separated from noise by amplitude threshold discrimination, where the threshold can be designated by several standard variations in background noise (Lewicki 1998). Once a spike is detected, extracting crucial feature coefficients from the spike is an important process of spike sorting since clustering quality may be affected (Quiroga *et al* 2004). Principal component analysis (PCA) is the most commonly adopted dimension reduction method for spike feature extraction (Lewicki 1998, Thakur *et al* 2007, Adamos *et al* 2008). Recently, the use of wavelet transform (WT) has become popular for spike discrimination (Hulata *et al* 2002, Quiroga *et al* 2004, Brychta *et al* 2007, Pavlov *et al* 2007, Geng *et al* 2010, Takekawa *et al* 2010). Some studies have indicated that WT may outperform PCA in feature extraction when spike waveforms manifest a small-scale structure or under heavy recording noise (Aharon *et al* 2006, Pavlov *et al* 2007, Takekawa *et al* 2010). Other feature extraction methods are also available, such as the reduced features of spike morphology (Stewart *et al* 2004), the inter-spike interval-based algorithms (Delescluse and Pouzat 2006), the 3-Gaussian model (Roh *et al* 2008), and waveform derivatives (Yang *et al* 2009). Although several viable methods for spike feature selection are available, the establishment of a feature extraction method with accurate classification results and dimension reduction remains a challenging issue in the field.

Spikes from different neurons are distinguished by the unique and reproducible spike waveforms produced by each neuron (Lewicki 1998). Manual spike sorting may yield significant variability due to individual differences in experience and subjective judgment (Wood *et al* 2004, Takekawa *et al* 2010). For the past several decades, many

clustering algorithms have been used for spike sorting. The utility of a fully automated algorithm for spike sorting is indispensable and important, especially for multichannel recordings (Kim 2006). In general, clustering algorithms can be divided into supervised and unsupervised clustering methods. Supervised clustering methods, such as artificial neural network (Chandra and Optican 1997, Kim and Kim 2000), K-means clustering (Hulata *et al* 2000), fuzzy C-means methods (Zouridakis and Tam 2000, Anil *et al* 2000), template matching (Bankman *et al* 1993, Vollgraf and Obermayer 2006, Carlos and Donoghue 2007), and multiwavelet transform and hierarchical clustering (Geng *et al* 2010), have been widely used. As supervised clustering methods are restricted by the need to identify a 'true' label in each spike and a 'true' number of classes in advance, they are not feasible for real-time analysis. Alternatively, the unsupervised clustering method assigns spikes to the classes by appropriate pattern recognition. Previous studies have employed unsupervised clustering, such as a Gaussian mixture model (Lewicki 1998), a t -distribution mixture model (Shoham *et al* 2003), superparamagnetic clustering (Quiroga *et al* 2004), Bayesian clustering (Aharon *et al* 2006), evolving mean shift clustering (Yang *et al* 2009), and robust variational Bayes clustering (Takekawa *et al* 2010) for spike sorting. In addition, hidden Markov models (HMMs) are an alternative approach that incorporates spike detection and classification into a single computational procedure (Herbst *et al* 2008). Although a variety of clustering algorithms are available, the development of an optimal unsupervised clustering method remains a worthy topic for study.

The current study presents an automatic spike sorting method that uses unsupervised single linkage clustering based on grey relational analysis (GRA). The current approach extracts features from spike waveforms by applying a wavelet transform; the crucial feature coefficients are then selected with the Kolmogorov–Smirnov test (KS test). Instead of a standard single linkage clustering, GRA (Deng 1982) is used to estimate the similarities between a pair of spikes. Furthermore, the new approach determines the number of clusters according to the threshold-cumulative distribution, which serves as an objective guideline for automatic spike sorting.

2. Methods

2.1. Overall procedure

Spike sorting is a classifying technique used for the identification of individual cell activity in electrophysiological

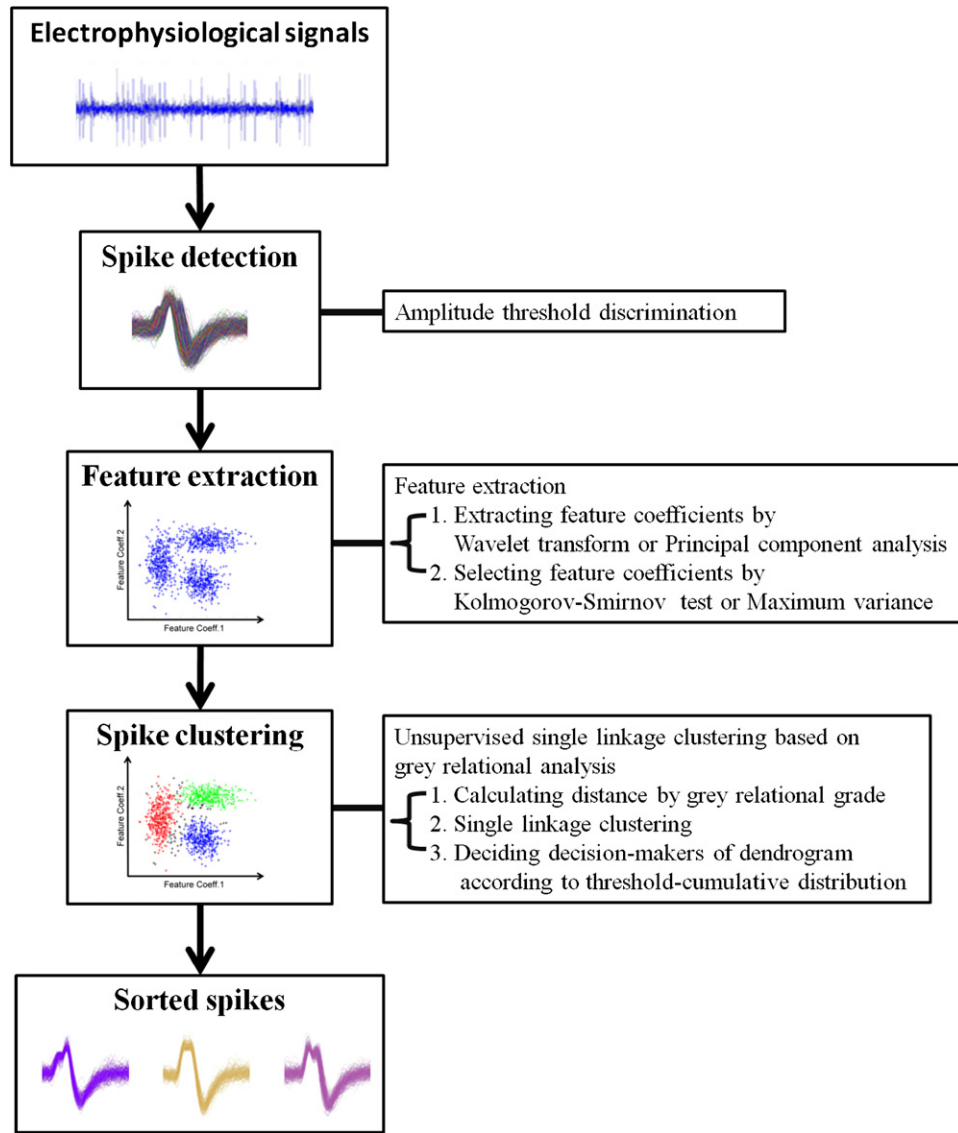


Figure 1. Overview of the automatic spike sorting procedure. Spikes are detected according to an amplitude threshold; a set of the feature coefficients is then selected from the spikes' wavelet coefficients with the KS test. Finally, the GSLC algorithm is used for automatic spike clustering.

data. The flowchart of the spike sorting procedure is shown in figure 1. Spike sorting consists of three major steps: (1) spike detection by amplitude threshold discrimination, (2) feature extraction by wavelet transform (WT) with the KS test; and (3) spike clustering by single linkage clustering based on GRA (GSLC). GSLC is an unsupervised clustering method that determines the optimum number of clusters automatically. The use of simulated data serves as an objective evaluation of the spike sorting algorithm hereby proposed. Finally, GSLC is implemented to classify electrophysiological data collected from eight patients with Parkinson's disease. The procedures of GSLC are described in detail in the following sections.

2.2. Neural spike detection

Spike detection is used to segment spike waveforms from the background noise based on an amplitude threshold (Donoho

and Johnstone 1994). The threshold (Th) was set as

$$Th = 4 \cdot \text{median} \left\{ \frac{|x|}{0.6745} \right\}, \quad (1)$$

where x was the neural signal (including spikes and background noises). The threshold Th was estimated based on a fourfold median of the neural signal. The conventional estimation of the amplitude threshold using the standard deviation (S.D.) of neural signals (Lewicki 1998) may result in very large threshold values due to high firing rates and large spike amplitudes. Therefore, using the median as an estimation of the background noise may diminish spike interferences (Quiroga *et al* 2004). Each detected spike consisted of 48 samples (i.e. 2 ms), which were aligned at the point of threshold crossing. The significant feature coefficients were extracted from each spike, as discussed below.

2.3. Extracting and selecting the crucial feature coefficients

The Haar wavelet transform (Daubechies 1992) was implemented to extract the four-level decomposition wavelet coefficients from the spike waveforms. The Haar wavelet is a compact support orthogonal wavelet, representing spikes with few wavelet coefficients, with no prior assumption to the spike shapes (Hariharan and Kannan 2010, Venkatesh *et al* 2010, Quiroga *et al* 2004). Sixty-four wavelet coefficients characterizing spike waveforms at the temporal and frequency domain for each spike were obtained (Hulata *et al* 2002). After extracting the spike features, the KS test (Press *et al* 2002) was used to select crucial wavelet coefficients as the feature coefficients for the determination of the best spike class.

The KS test was used to compute the differences between the distribution of wavelet coefficients and the Gaussian distribution. $F(x)$ was the cumulative distribution function of the data set x . $G(x)$ was a Gaussian distribution with the same mean and variance as $F(x)$. D was the deviation of the data set from a Gaussian distribution, quantified by

$$D = \max_{-\infty < x < \infty} |F(x) - G(x)|. \quad (2)$$

Six wavelet coefficients with the largest deviation from normality were used as the feature coefficients; these wavelet coefficients accounted for approximately 95% of the energy of all wavelet coefficients in the simulation data sets. For cases in which there were more than two spike classes, such feature coefficients would present as a multimodal distribution.

2.4. Single linkage clustering based on GRA (GSLC)

The selected feature coefficients were applied as feature vectors of each spike to be clustered by GSLC. Instead of the Euclidean distance, the grey relational grade (GRG) (Deng 1982, Yi and Pin 1992) was used in the single linkage clustering (Gower and Ross 1969) to assess the overall degree of similarities between two spikes (Azzeh *et al* 2010). The major advantage of GRG is its ability to handle incomplete information precisely, despite a reduced number of feature vectors and an unknown data distribution (Morán *et al* 2006). In order to compact elements in an inner cluster, the feature coefficients were computed again with a relational threshold set as the average of the top 3% of GRGs between any elements. These revised spike feature coefficients presented compacted elements in an inner cluster, whereas separated elements were presented in an outer cluster. The number of clusters was subsequently estimated by choosing the segment parameters that affected the linkage threshold. The optimum segment parameter was decided according to the segment parameters with the maximum number of clusters and spikes. Further details of the GSLC are described below.

2.4.1. Step 1: calculating the GRG between spikes. $x_i(k)$ and $x_j(k)$ were the feature coefficients of spike i and spike j , respectively, where $i = 1, 2, \dots, n$, $j = 1, 2, \dots, n$, $i \neq j$ and $k = 1, 2, \dots, m$. n was the number of detected spikes and m was the number of feature coefficients for each spike. The correlation between two spikes was computed by the GRG.

First, the grey relational coefficient (GRC), $\gamma(x_i(k), x_j(k))$, was expressed as

$$\gamma(x_i(k), x_j(k)) = \frac{\min_{\forall i, \forall j} \min_{\forall k} \|x_i(k) - x_j(k)\| + \zeta \max_{\forall i, \forall j} \max_{\forall k} \|x_i(k) - x_j(k)\|}{\|x_i(k) - x_j(k)\| + \zeta \max_{\forall i, \forall j} \max_{\forall k} \|x_i(k) - x_j(k)\|}. \quad (3)$$

The distinguishing coefficient, $\zeta \in (0, 1]$ (i.e. $0 < \zeta \leq 1$), controlled the resolution scale, which was set as 1 in the current implementation. $\|x_i(k) - x_j(k)\|$ was the absolute value (norm) of the difference between $x_i(k)$ and $x_j(k)$; $\min_{\forall i, \forall j} \min_{\forall k} \|x_i(k) - x_j(k)\|$ was the secondary minimum difference in all i and j ; $\max_{\forall i, \forall j} \max_{\forall k} \|x_i(k) - x_j(k)\|$ was the secondary maximum difference in all i and j . After obtaining the GRC, the GRG was obtained by averaging the GRCs with

$$\gamma_{ij} = \gamma(x_i, x_j) = \frac{1}{m} \sum_{k=1}^m \gamma(x_i(k), x_j(k)). \quad (4)$$

The GRG series for x_i was defined as $R_i = \{\gamma_{i1}, \gamma_{i2}, \dots, \gamma_{in}\}$.

2.4.2. Step 2: compacting elements by setting the relational threshold. Revised spike feature coefficients that compacted the neighbouring elements and separated the distant elements were computed. The relational threshold, ω , was set to compact the elements as γ_{ij} equaled or exceeded ω . The relational threshold was defined as

$$\omega = \frac{1}{n} \sum_{i=1}^n R'_i, \quad (5)$$

where R'_i was the average of the top 3% of GRGs, $R_i = \{\gamma_{i1}, \gamma_{i2}, \dots, \gamma_{in}\}$.

Next, the spike feature coefficients, $X = \{x_1, x_2, \dots, x_n\}$, were recalculated as $Y = \{y_1, y_2, \dots, y_n\}$, according to the threshold relational,

$$y_i(k) = \frac{\sum_{j=1}^n n_{ij} x_j(k)}{\sum_{j=1}^n n_{ij}} \quad \text{where} \quad n_{ij} = \begin{cases} 1 & \text{if } \gamma_{ij} \geq \omega \\ 0 & \text{if } \gamma_{ij} < \omega. \end{cases} \quad (6)$$

After the new feature coefficients of the spikes were obtained, the GRG of Y was computed as

$$\gamma_{y_i y_j} = \gamma(y_i, y_j) = \frac{1}{m} \sum_{k=1}^m \gamma(y_i(k), y_j(k)). \quad (7)$$

2.4.3. Step 3: setting the linkage threshold to cluster the spikes by single linkage clustering. The linkage threshold was set to determine whether the elements were to be merged into the same cluster or separated as different clusters. The linkage threshold, v_q , depending on the segment parameter q , was defined as

$$v_q = \omega + (q - 1) \times \frac{(1 - \omega)}{Q}, \quad (8)$$

$q, Q \in N, q \leq Q$

where $q = 1, 2, \dots, Q$; the maximum segment parameter Q was set as 20, which was sufficient for adequate performance (please refer to the appendix for the quantitative determinations of the maximum segment parameter Q). Setting the initial linkage threshold as the relational threshold, ω , the linkage threshold was increased by incrementing the segment parameter. Initially, $Y = \{y_1, y_2, \dots, y_n\}$ were classified into distinct clusters, C_1, C_2, \dots, C_n . As the GRG of Y , $\gamma_{y_i y_j}$ was conformed to $\gamma_{y_i y_j} \geq v_q$, y_i and y_j were classified into the same cluster so that the elements were clustered. As the linkage threshold was increased with an added segment parameter, the merging of elements became more difficult. In other words, when the GRG, $\gamma_{y_i y_j}$, of some clusters was smaller than the linkage threshold v_q , these clusters remained separated. In general, a sample size of 30 was adopted as a reasonable sample size for valid statistical results. Thus, cluster(s) in which the number of elements was fewer than 30 were regarded as a singularity, and were subsequently removed. This procedure, also known as an outline removal process, has been applied in other studies (Quiroga *et al* 2004, Geng *et al* 2010).

2.4.4. Step 4: selecting the optimum segment parameters.

An optimum segment parameter was selected for the best spike clustering results. In order to develop an automatic and unsupervised algorithm, the linkage thresholds were divided into 20 segment parameters to merge adjacent clusters. The segment parameter, q , thus ranged from 1 to 20. In the implementation, the number of clusters was estimated based on specific criteria. Spikes were clustered by using a linkage threshold, with incrementing segment parameters to cut the dendrogram. The segment parameter with the maximum number of clusters and spikes was considered to be the optimum segment parameter. Then, the number of clusters was determined by the linkage threshold with the optimum segment parameter. For example, the results of GSLC spike clustering are shown in figure 2. The threshold-cumulative distribution of the number of clusters and spikes in each cluster for linked thresholds with different segment parameters is shown in figure 2(A). The maximum number of clusters was 3; the segment parameter, q , ranged from 10 to 19; the number of spikes was optimum at segment parameter $q = 10$. The dendrogram for the grey relation grade of spikes is shown in figure 2(B). The solid line indicated the decision maker of the linkage threshold, with the optimum segment parameter, $q = 10$, as the GRG was 0.89, which produced the maximum number of clusters with the maximum number of spikes. The results of spike clustering with the optimum segment parameter, $q = 10$, are shown in figure 2(C).

2.5. Data source

In the current study, the performance of spike sorting using GSLC was examined with simulated data sets and electrophysiological data recorded from patients with Parkinson's disease. Published simulated data were obtained from a database of 594 different average spike waveforms, compiled from recordings in the

neocortex and basal ganglia (Quiroga *et al* 2004) (<http://www2.le.ac.uk/departments/engineering/research/bioengineering/neuroengineering-lab/spike-sorting>). Background noise was generated from randomly selected spikes in the database, superimposed with random times and amplitudes (Fee *et al* 1996, Pouzat *et al* 2002). A spike train consisting of three distinct spike waveforms with background noise at random points was utilized; the sampling rate was set as 24 kHz. The peak value of three distinct spike waveforms was normalized to 1 in each data set. The noise level was defined according to Quiroga *et al* (2004): four noise levels, determined from the S.D., were set as 0.05, 0.1, 0.15, and 0.2, relative to the amplitude of the spikes. For all spike trains, three distinct spikes had a Poisson distribution of inter-spike intervals, with a mean firing rate of 20 Hz and a refractory period of 2 ms. Accordingly, four simulated data sets with four noise levels were generated. The three distinct spike waveforms of each simulated data set are shown in table 1. The difference in spike waveforms in the simulated data set is defined by the average correlation coefficient (ACC) between class 1, class 2 and class 3. A low ACC indicates spike waveforms that are easy to differentiate, as shown in simulated data set 1 with an ACC of 0.2102. In contrast, the spikes in simulated data sets 2, 3 and 4 have similar waveforms, with ACCs of 0.9085, 0.8579 and 0.7936, respectively. Nonetheless, there are differences in the shape and time durations of the spikes.

Eight patients with Parkinson's disease at Buddhist Tzu Chi General Hospital, Taiwan, volunteered to participate; informed consent was obtained from all subjects in accordance with Buddhist Tzu Chi General Hospital's Institutional Review Board (IRB 097-08) Committee on research involving human subjects. The electrophysiological data were recorded from the subthalamic nucleus (STN) during a surgical operation for deep brain stimulation. The patients with Parkinson's disease (seven male, one female; age 57.8 ± 7.4 years, disease duration of 9.9 ± 3.3 years) were placed under local anesthesia with lidocaine (3%, s.c.). A microelectrode (Lot# 108874, FHC Inc., USA) with a 10 mm bare tungsten tip and 10–40 μm in diameter and an impedance in the range of 0.4–0.7 M Ω at 1 kHz was mounted on a micro-driver to send the tip into the STN target (target location, X: 11 ± 2 mm, Y: -3 to -4 mm and Z: -6 to -7 mm with respect to the midpoint of the median line in the third ventricle). Microelectrode recording (MER) has been widely used for locating target neural areas during DBS surgical operations. Electrophysiological data obtained from the microelectrode were band-pass filtered (300 Hz–5 kHz), amplified (gain: 20000), and sampled (24 kHz) with the Leadpoint recording system (Lead Point 4, Medtronic Inc., USA). Because spike information is used to measure neuronal activity of a targeted location, spike sorting of electrophysiological data is typically conducted prior to further analyses.

2.6. Evaluating the quality of spike sorting

The percentage of the correct number of neurons (CNN (%)) was used to evaluate the capability of cluster determination.

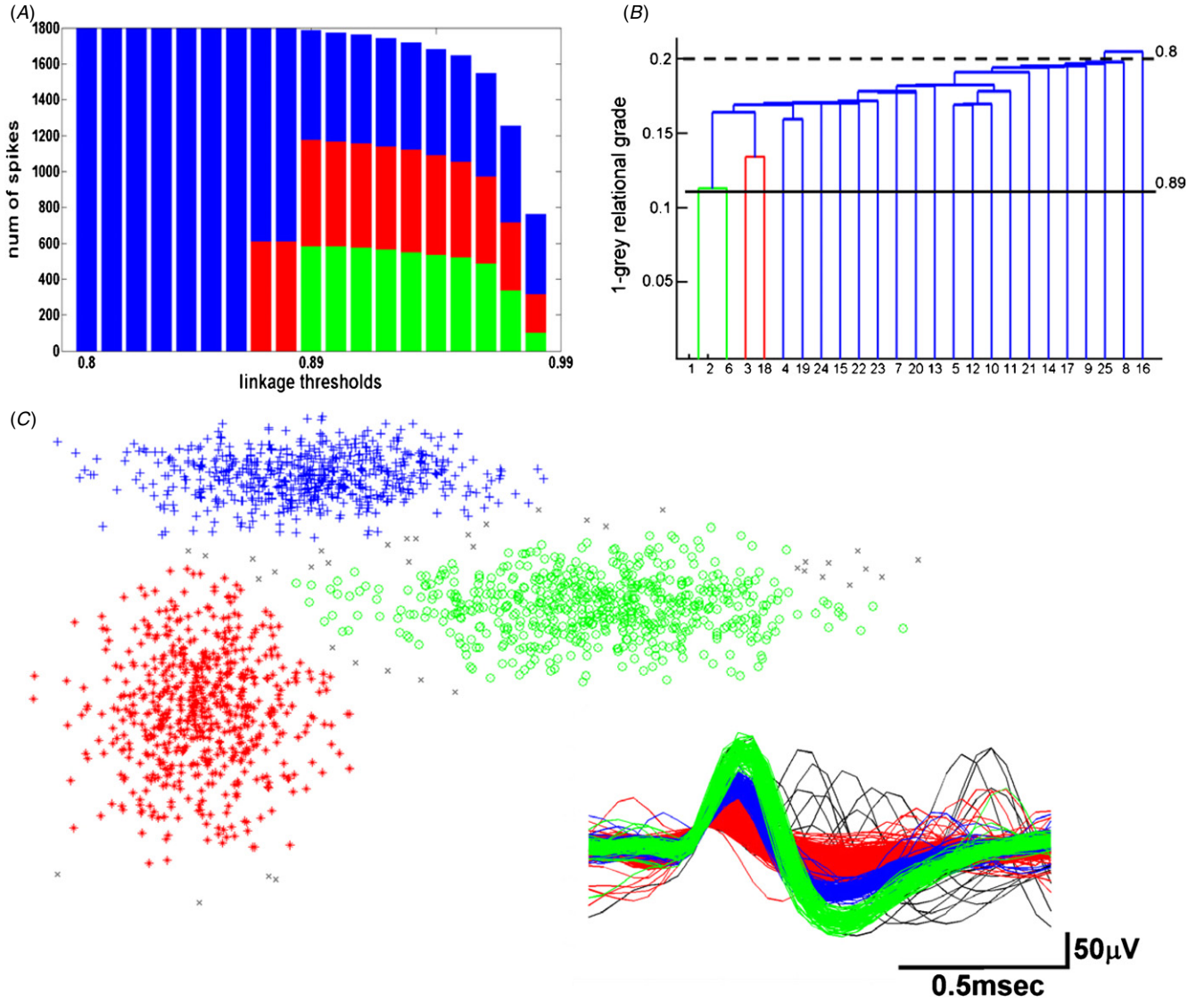


Figure 2. The linkage threshold-cumulative distribution and the dendrogram of the GSLC algorithm. The number of spikes in each cluster, with linkage threshold-cumulative at the segment parameter $q = 1, 2, \dots, Q$, where $Q = 20$, is shown in (A). The dendrogram for the grey relation grade between the spikes is shown in (B). The relational threshold, ω , where the initial linkage threshold was 0.8, is shown as a dotted line; the optimum segment parameter 10, where the linkage threshold was 0.89, is shown as a solid line. The two-dimensional (2D) projections of the first two feature coefficients, where the wavelet coefficients are selected by the KS test, are shown in (C); the three waveforms of the sorted spikes are shown at the bottom right of (C).

The CNN (%) indicated the percentage of detected neurons over the total number of neurons in the data set, defined as

$$\text{CNN (\%)} = \frac{\text{number of detected neurons}}{\text{total number of neurons}} \times 100\%. \quad (9)$$

A CNN (%) of 100% indicated that an accurate number of neurons in the simulated data set was detected. In contrast, if CNN (%) was less than 100%, some neurons were not identified. Furthermore, the rate of classification accuracy (CA (%)) was defined as the percentage of the number of correctly classified spikes over the total number of spikes,

$$\text{CA (\%)} = \frac{\text{number of correctly classified spikes}}{\text{total number of spikes}} \times 100\%. \quad (10)$$

In general, a decrease in CA (%) may result from two types of errors: false positive errors (false alarms) occur when spikes

are assigned to the wrong neuron, and false negative errors (misses) occur when spikes are omissive.

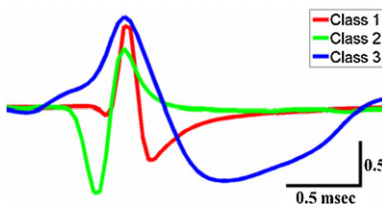
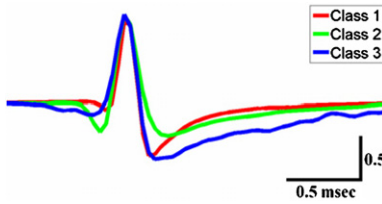
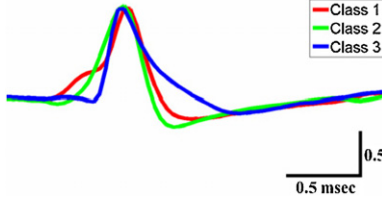
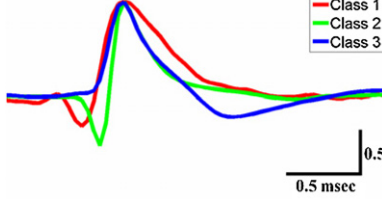
The quality of spike sorting for electrophysiological data could not be evaluated with CNN (%) and CA (%) due to the unknown number of originating neurons. Therefore, a non-parametric measure, J -measure, was used to evaluate the quality of spike sorting (Wheeler 1999). $E(x)$ indicated the squared Euclidean distance; the trace of a matrix was equal to the sum of its eigenvalues. Three equations for J -measure were defined,

$$J_1 = \sum_u \sum_i E(f(u, i) - m(u)) \quad (11)$$

$$J_2 = \sum_u N(u) E(m(u) - m) \quad (12)$$

$$J\text{-measure} = J_2 / J_1, \quad (13)$$

Table 1. The number of spikes in each class in four simulated data sets at four noise levels.

Simulated data		Noise level	Amount of spikes	Class 1	Class 2	Class 3
Set 1 ACC: 0.2102		0.05	3514	1165	1157	1192
		0.1	3522	1151	1134	1237
		0.15	3477	1132	1188	1157
		0.2	3474	1198	1128	1148
Set 2 ACC: 0.9085		0.05	3383	1115	1113	1155
		0.1	3448	1164	1155	1129
		0.15	3472	1159	1172	1141
		0.2	3414	1136	1099	1179
Set 3 ACC: 0.8579		0.05	3364	1120	1109	1135
		0.1	3462	1187	1136	1139
		0.15	3440	1142	1113	1185
		0.2	3493	1151	1195	1147
Set 4 ACC: 0.7936		0.015	3410	1130	1113	1167
		0.1	3520	1160	1146	1214
		0.15	3411	1181	1098	1132
		0.2	3526	1186	1188	1152

where J_1 was the measure of the average distance of points in a cluster $f(u, i)$ from their centre $m(u)$ in feature space, which was minimized for the compact class. The summations were over neurons u , and over feature vectors (points) in each neuron i . J_2 was a measure of the average distance between the clusters; it was maximized for well-separated class means. The summation was over neurons u , and $N(u)$ was the number of points in the neuron u , $m(u)$ was the cluster centre for neuron u , and m was the grand centre of all points in all neurons. Thus, J -measure referred to the maximum value for compact, well-separated clusters in the evaluation of spike sorting.

3. Results

Due to the significance of feature coefficients for spike sorting, the simulated data sets were used to examine the applicability of the proposed algorithm for feature extraction.

The simulated data sets with four different noise levels (0.05, 0.1, 0.15, and 0.2) were sorted using GSLC, with four compositions of the feature coefficients, WT and PCA accompanied by the KS test and MV. The simulated data sets were also used to evaluate the quality of the clustering method based on the CNN (%) and CA (%) using GSLC, as compared with the use of single linkage clustering (SLC) (Murtagh 1983), superparamagnetic clustering (SPC) (Quiroga *et al* 2004), K-means (Jain *et al* 1999), multiwavelets transform and hierarchical clustering (MWHC) (Geng *et al* 2010), efficient technology of spike sorting (EToS) (Takekawa *et al* 2010) and hidden Markov models (HMMs) (Herbst *et al* 2008). Furthermore, STN electrophysiological data recorded in the patients with Parkinson's disease were clustered by GSLC; results were then compared with SLC, SPC and T-distribution expectation-maximization (T-Dist E-M) (Shoham *et al* 2003).

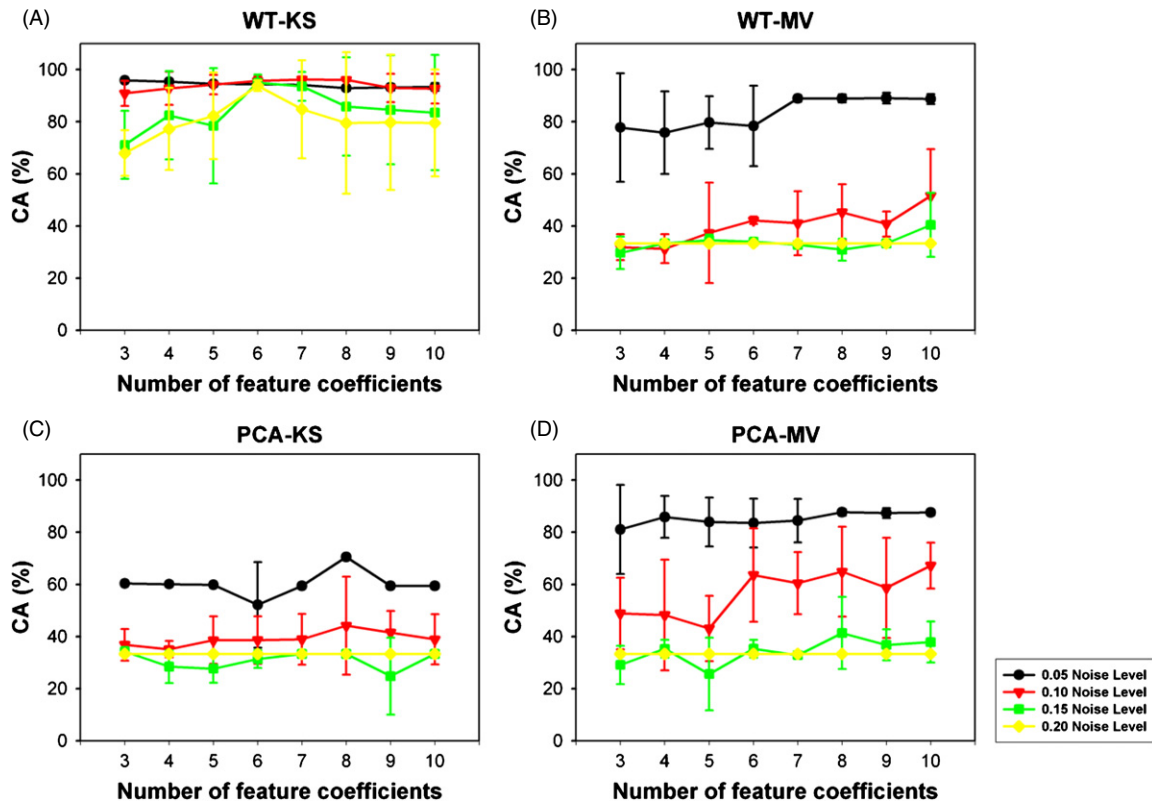


Figure 3. The mean CA (%) of three simulated data sets (simulated data sets 2, 3 and 4) at four noise levels is plotted with an increasing number of feature coefficients. The feature coefficients are selected by (A) WT-KS, (B) WT-MV, (C) PCA-KS and (D) PCA-MV, respectively. Classification accuracy is significantly affected by noise levels; the CA (%) is not increased despite the increase in the number of feature coefficients. As shown, the best results are attained with the WT-KS method. Error bar: \pm standard derivation (S.D.) for this study.

3.1. Determination of the crucial feature coefficients

Feature extraction is an important process of spike sorting, as classification accuracy may be affected. The current experiment aims to determine the optimum number of feature coefficients and feature spaces. The feature coefficients were determined by four different combinations: the wavelet coefficients selected by the KS test (WT-KS) or MV (WT-MV), and the principal components selected by the KS test (PCA-KS) or MV (PCA-MV).

3.1.1. Determination of the number of feature coefficients.

The relationship between the number of feature coefficients and noise levels was examined with three simulated data sets (simulated data sets 2, 3 and 4), with three similar spike waveforms. The mean CA (%) at four noise levels, with an increasing number of feature coefficients, is shown in figure 3. Six wavelet coefficients selected by the KS test are used to obtain the maximum CA at all noise levels, as shown in figure 3(A). As the number of wavelet coefficients increases, there is no increase in CA (%); CA (%) is limited by the noise levels when the feature coefficients are selected by WT-MV, PCA-KS and PCA-MV (figures 3(B)–(D)). Nonetheless, there was a slight increase in CA (%) proportional to increasing numbers of feature coefficients. Such results indicated that the WT-KS method significantly outperforms other methods for feature extraction.

In addition, the effect of similar spike waveforms on classification accuracy was examined. CA (%)s at a 0.15 noise level, with increasing numbers of feature coefficients, are illustrated in figure 4. In simulated data set 1, which displays three dissimilar spike waveforms, results indicate CA (%) of over 75% across different numbers of feature coefficients obtained from the four feature extraction methods across different numbers of feature coefficients, as shown in figure 4(A). In simulated data set 2, which demonstrated a high ACC (i.e. similar spike waveforms in the simulated data), results indicate CA (%) of consistently under 37%, despite the increased number of feature coefficients from WT-MV, PCA-KS and PCA-MV, as shown in figure 4(B). Nevertheless, the best results were obtained with the WT-KS in simulated data set 2. Results from simulated data set 3 and 4 are similar to those in simulated data set 2, as shown in figures 4(C) and (D), respectively. These results suggested that with similar spike waveforms, classification accuracy was not improved by increasing the number of feature coefficients.

3.1.2. Contrasting WT feature space and PCA feature space.

As presented in table 2, spike sorting results showed prominent clustering with GSLC. Different numbers of feature coefficients from the WT and the PCA feature spaces yielded different results when selecting among the KS test and MV. The WT-KS method attained a 100% of CNN (%) and a maximum CA (%) of 94.87% at six feature coefficients. The best result with the WT-MV method was attained with ten

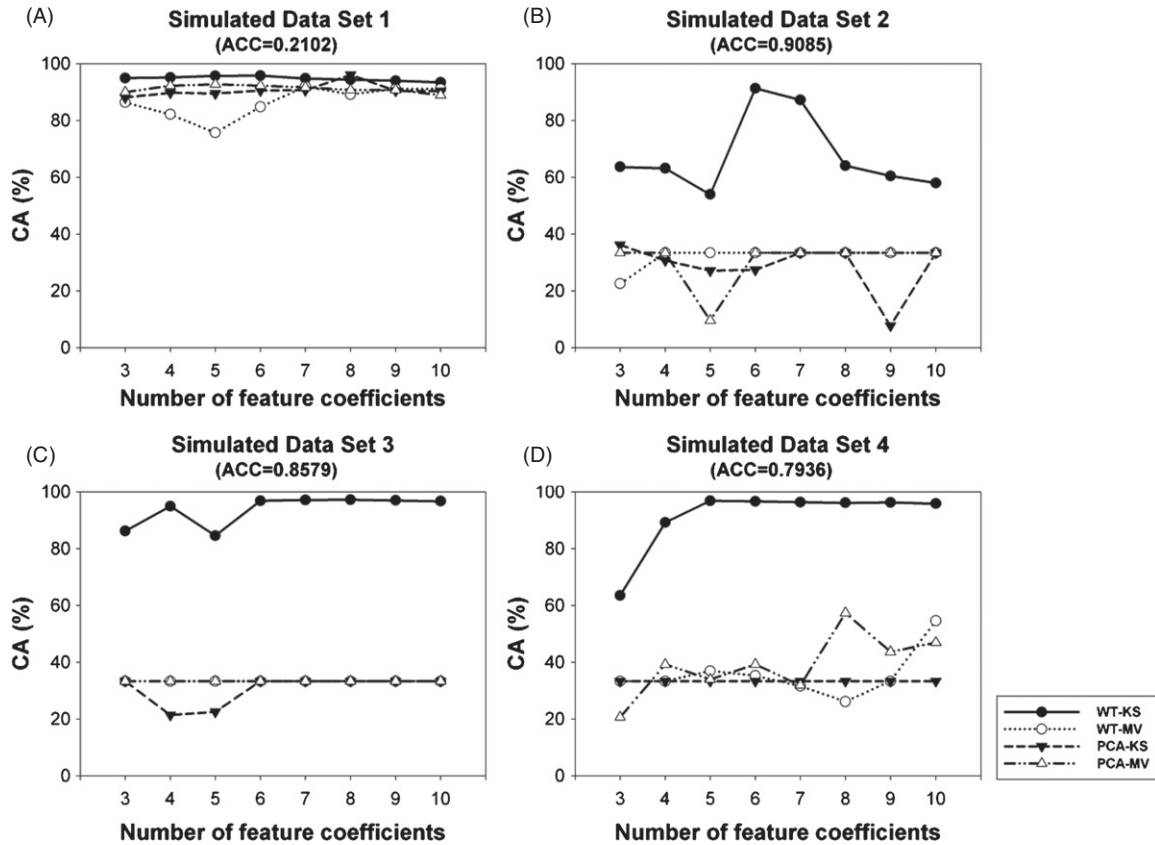


Figure 4. The CA (%) of the four feature extraction methods is plotted with an increasing number of feature coefficients. Simulated data sets 1, 2, 3, and 4 at 0.15 noise levels are presented in (A), (B), (C), and (D), respectively. (A) There are three dissimilar spike waveforms in simulated data set 1; thus, high CA (%) are obtained due to the smaller number of feature coefficients used with the four feature extraction methods. (B), (C) and (D) show that the CA (%) of the simulated data with similar spike waveforms is not increased with the implementation of WT-MV, PCA-KS and PCA-MV to increase the number of feature coefficients.

Table 2. CNN (%) and CA (%) of all simulated data sets using WT-KS, WT-MV, PCA-KS and PCA-MV methods and different numbers of feature coefficients, respectively.

		Number of feature coefficients							
		3	4	5	6	7	8	9	10
WT-KS	CNN (%)	81.25	93.75	93.75	100*	93.75	93.75	100*	100*
	CA (%)	84.94	89.12	89.39	94.87*	92.86	90.06	89.18	88.84
WT-MV	CNN (%)	43.75	37.5	50	50	43.75	50	50	56.25*
	CA (%)	53.15	52.92	53.57	55.73	57.73	58.3	59.15	62.63*
PCA-KS	CNN (%)	37.5	25	31.25	43.75*	25	25	31.25	25
	CA (%)	50.33	49.15	50.7	51.34	53.06	58.49*	51.88	52.98
PCA-MV	CNN (%)	50	56.25	56.25	62.5	56.25	68.75*	62.5	56.25
	CA (%)	58.32	60.81	57.87	63.35	62.15	65.14*	62.92	64.57

* The best performance in the row.

feature coefficients: the maximum CNN (%) was 56.25% and the maximum CA (%) was 62.63%. These results illustrated that the KS test was better than the MV for selecting feature coefficients in the WT feature space. With PCA, CNN (%) was maximum (43.75%) when the KS test was used to select six principal components. CA (%) was highest (58.49%) when the first eight principle components were selected. At eight feature coefficients, the PCA-MV method acquired a maximum CNN (%) and CA (%) of 68.75% and 65.14%, respectively. Conversely, these results showed that the MV was better than the KS test for selecting feature coefficients in

the PCA feature space. To summarize, optimum clustering was attained with the WT-KS method when six feature coefficients were used in WT feature space; optimum performance was attained by MV with eight feature coefficients in the PCA feature space. In order to compare the performances of the WT feature space and the PCA feature space among all spike sorting methods, six wavelet coefficients by WT-KS and eight principal components by PCA-MV were used.

The three-dimensional (3D) projections and distributions of the three best feature coefficients, determined by WT-KS, WT-MV, PCA-KS and PCA-MV, are shown in figure 5.

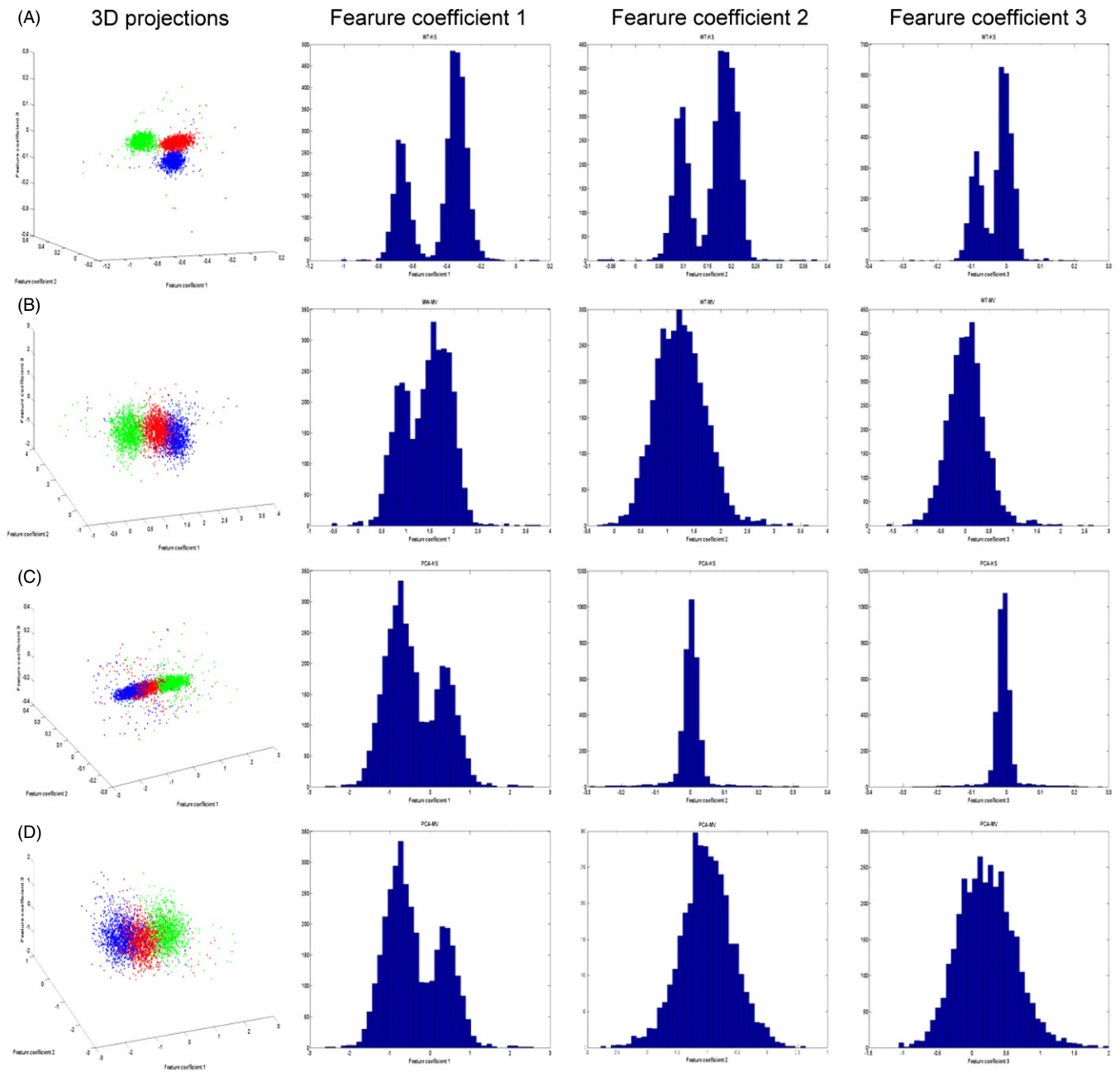


Figure 5. The three-dimensional (3D) projections and distributions of the best three feature coefficients determined with the wavelet coefficients selected by (A) the KS test (WT-KS), (B) the maximum variance (WT-MV), the principal components selected by (C) the KS test (PCA-KS), and (D) the maximum variance (PCA-MV). Simulated data set 2 with a 0.15 noise level is used in this example.

The best three wavelet coefficients selected by the KS test present a multimodal distribution for each feature coefficient, as shown in figure 5(A). Our results indicated that the three clusters were clearly identifiable with WT-KS. In contrast, only one projection of the first feature coefficient shows a slightly multimodal distribution, whereas both the second and the third feature coefficients display a uniform distribution in WT-MV, PCA-KS and PCA-MV (figures 5(B)–(D)). These results demonstrated that the WT-KS selected crucial feature coefficients for spike cluster separation.

3.2. Comparing the classification accuracy using GSLC with SLC, SPC, K-means, MWHC, EToS and HMMs in simulated data

In order to illustrate the advantage of GSLC, classification accuracies were compared between those obtained from GSLC and those from SLC, SPC, K-means, MWHC, EToS and HMMs. Taking into account the effects of noise levels, classification results from GSLC, SLC, SPC, MWHC, K-means, MWHC, EToS and HMMs, with different noise levels, are shown in table 3.

Table 3. CNN (%) and CA (%) of all simulated data sets using ISLC, SLC, SPC and K-means clustering, respectively.

	Noise levels	GSLC		SLC		SPC		K-means				
		WT-KS	PCA-MV	WT-KS	PCA-MV	WT-KS	PCA-MV	WT-KS	PCA-MV	MWHC	EToS	HMM
CNN (%)	0.05	100.00	100.00	100.00	100.00	100.00	91.67	N/A	N/A	N/A	91.67	100.00
	0.1	100.00	100.00	100.00	66.67	100.00	91.67	N/A	N/A	N/A	91.67	100.00
	0.15	100.00	66.67	100.00	50.00	100.00	83.33	N/A	N/A	N/A	100.00	100.00
	0.2	100.00	50.00	91.67	50.00	91.67	91.67	N/A	N/A	N/A	100.00	100.00
	Mean	100.00*	79.17	97.92	66.67	97.92	89.58	N/A	N/A	N/A	95.83	100.00*
CA (%)	0.05	94.25	87.88	94.63	82.63	94.45	63.65	93.35	97.45	92.35	78.56	99.90
	0.1	95.85	81.48	94.40	44.60	93.93	62.28	87.53	94.23	95.07	79.99	99.91
	0.15	95.13	53.65	92.20	47.73	93.98	59.93	80.03	80.20	94.74	86.04	99.80
	0.2	94.25	47.43	81.10	40.65	84.20	63.98	77.15	70.45	94.68	84.13	99.07
	Mean	94.87	67.61	90.58	53.90	91.64	62.46	84.51	85.58	94.21	82.18	99.67*

* The optimum performance in the row.

'N/A' Note that the CNN (%) using K-means clustering and MWHC were unavailable because the number of clusters must be determined beforehand by manual operation.

K-means is a supervised clustering method; the number of clusters was set as 3 in each simulated data. In addition, K-means is also a partition method and thus all spikes were categorized into different clusters. In K-means clustering, the CNN (%) must be 100%; thus, the CNN (%) is labelled 'N/A' in table 3. The mean CA (%) of WT-KS and PCA-MV were 84.51% and 85.58%, respectively. MWHC was implemented by using multiwavelet transform to extract feature coefficients from each spike, and using hierarchical clustering to classify spikes. The number of clusters was determined by cutting the dendrogram manually. Thus, the CNN (%) must be 100% and labelled 'N/A' in table 3. The mean CA (%) of MWHC was 94.21%.

As the data illustrate, as the noise level increased, CNN (%) decreased. This trend is particularly significant for the clustering algorithms using PCA-MV, as shown in table 3. In general, the CA (%) using PCA-MV was lower than when using WT-KS in all simulated data sets, especially at high noise levels. Both HMMs and GSLC on WT-KS attained a CNN (%) of 100% at all noise levels. HMMs obtained a maximum mean CA (%) of 99.67% while GSLC on WT-KS resulted in the second highest mean CA (%) of 94.87%. Furthermore, the mean CNN (%) and mean CA (%) for SLC on WT-KS was 97.92% and 90.58%, respectively. Such results indicated that SLC clustering was improved with the implementation of GSLC. The application of SPC on WT-KS resulted in a mean CNN (%) of 97.92% and a mean CA (%) of 91.64%. The CA (%) of GSLC was significantly higher than those with SLC, SPC and K-means applied on WT-KS, especially at high noise levels (0.2 noise level). In addition, a new alternative approach, the EToS, obtained a mean CNN (%) of 95.83% and a mean CA (%) of 82.18%.

3.3. Comparing the quality of spike sorting using GSLC with SLC, SPC, and T-Dist E-M in electrophysiological data

The number of clusters and the J -measure using GSLC-WT, GSLC-PCA, SLC-WT, SLC-PCA, SPC-WT, SPC-PCA and T-Dist E-M in the electrophysiological data recorded from eight

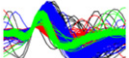
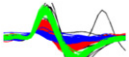
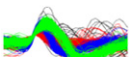
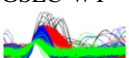

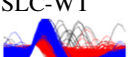


patients with Parkinson's disease were presented in table 4. The maximum number of total clusters was 20 with GSLC-WT, GSLC-PCA, SPC-PCA and T-Dist E-M (subject nos 1, 2, 3 and 4 with three clusters; subject nos 5, 6, 7 and 8 with two clusters). With SPC-WT, the total minimum number of clusters was 13, which was significantly less than the other methods, as shown in table 4. For only one cluster, J -measure was zero with SPC-WT, as shown in table 4 (subject nos 4, 6, 7 and 8). In comparison to GSLC-PCA, SLC-WT, SLC-PCA, SPC-WT, SPC-PCA and T-Dist E-M, higher J -measure values were obtained with the GSLC-WT method in six patients (subject nos 1, 2, 3, 4, 5 and 8). In addition, spike sorting results were ranked according to J -measure. The minimum summarization of rank was 14, indicating that the best quality of spike sorting was obtained with the GSLC-WT method, as shown in table 4. The optimum sorted spike waveforms with the highest J -measure is shown in the right column of table 4.

4. Discussion

In this study, an unsupervised algorithm is established for automatic spike sorting in neuronal signal recordings. The GSLC is an automatic spike sorting procedure with satisfactory classification accuracy; it also reduces manual labour and subjectivity. The classification accuracy of the proposed algorithm is evaluated on simulated data sets, at different noise levels and different similarity degrees of spike waveforms. The algorithm is then applied in spike sorting of electrophysiological data. Results show higher classification accuracy for the GSLC than existing methods in both the simulated data and the electrophysiological data (Murtagh 1983, Lewicki 1998, Jain *et al* 1999, Kaneko *et al* 1999, Shoham *et al* 2003, Quiroga *et al* 2004, Herbst *et al* 2008, Geng *et al* 2010, Takekawa *et al* 2010).

Results indicate that using WT for spike feature extraction is superior in conditions of heavy noise and similar spike waveforms. Consistent with previous research (Hulata *et al* 2002, Quiroga *et al* 2004), the results show that WT is superior

Table 4. Spike sorting quality comparisons for the electrophysiological data of eight patients with Parkinson's disease.

Parkinsonian patients		GSLC-WT	GSLC-PCA	SLC-WT	SLC-PCA	SPC-WT	SPC-PCA	T-Dist E-M	Best result
Subject 1	No of clusters	3	3	3	3	3	3	3	GSLC-WT
	<i>J</i> -measure	18.222 ¹	12.810 ⁴	16.172 ²	11.741 ⁶	13.865 ³	9.517 ⁷	12.377 ⁵	
Subject 2	No of clusters	3	3	2	2	2	3	3	GSLC-WT
	<i>J</i> -measure	43.502 ¹	22.310 ²	13.041 ⁶	21.123 ³	3.146 ⁷	20.266 ⁵	20.716 ⁴	
Subject 3	No of clusters	3	3	3	3	2	3	3	GSLC-WT
	<i>J</i> -measure	16.597 ¹	14.197 ³	14.938 ²	13.747 ⁴	3.111 ⁷	11.021 ⁶	12.008 ⁵	
Subject 4	No of clusters	3	3	3	3	1	3	3	GSLC-WT
	<i>J</i> -measure	20.105 ¹	18.448 ³	19.464 ²	18.261 ⁴	0 ⁷	16.085 ⁶	17.618 ⁵	
Subject 5	No of clusters	2	2	2	2	2	2	2	GSLC-WT
	<i>J</i> -measure	27.826 ¹	21.879 ²	14.647 ⁴	14.43 ⁵	6.798 ⁷	13.839 ⁶	16.202 ³	
Subject 6	No of clusters	2	2	2	2	1	2	2	SLC-WT
	<i>J</i> -measure	5.882 ⁴	5.963 ³	6.495 ¹	6.303 ²	0 ⁷	5.508 ⁶	5.722 ⁵	
Subject 7	No of clusters	2	2	1	2	1	2	2	SLC-PCA
	<i>J</i> -measure	3.556 ⁴	3.888 ²	0 ⁶	9.313 ¹	0 ⁶	3.479 ⁵	3.697 ³	
Subject 8	No of clusters	2	2	2	2	1	2	2	GSLC-WT
	<i>J</i> -measure	5.241 ¹	3.543 ²	2.017 ⁶	3.387 ³	0 ⁷	3.075 ⁵	3.188 ⁴	
Total clusters	20	20	18	19	13	20	20		
Summarization of rank	14	21	29	28	51	46	34		

Numbers in superscript indicate the performance rank in the row.

Note: Kruskal–Wallis one-way analysis obtained $p < 0.0001$.

to PCA in noise rejection (figure 3). Although PCA has been considered a feature extracting method with satisfactory classification accuracy (Hu *et al* 2005, Adamos *et al* 2008), results from PCA are affected by spike waveform variance. As noise levels are directly related to the amplitude of spike waveforms, the directions of principal components in PCA may be affected, which remains a major disadvantage of PCA estimations (Adamos *et al* 2008). On the other hand, spike amplitudes are not taken into account for spike feature extraction with WT; thus, WT is minimally influenced by noise levels. In addition, spike waveforms are decomposed into different frequency bands and time scales with WT (figure 4(B)) and therefore variations among similar spike waveforms can be easily distinguished (Hulata *et al* 2002, Quiroga *et al* 2004, Pavlov *et al* 2007, Takekawa *et al* 2010).

Although WT is a potentially more powerful tool for feature extraction, feature coefficients must be selected prior to spike clustering. Therefore, accurate feature selection is crucial for optimal sorting. Crucial feature coefficients are obtained by WT with the KS test, which is utilized for discriminating between different clusters (figure 5(A)). Although the use of WT for spike sorting has been

recommended (Hulata *et al* 2002, Quiroga *et al* 2004, Brychta *et al* 2007, Pavlov *et al* 2007, Geng *et al* 2010, Takekawa *et al* 2010), researchers have used different selection methods based on empirical procedures (Letelier and Weber 2000), multiple-procedures (Takekawa *et al* 2010) and complicated information theory (Hulata *et al* 2002). The use of the aforementioned methods may not guarantee optimum classification accuracy; in addition, computing load and complexity may be exacerbated. Wavelet coefficients are selected by comparing the empirical distribution function to the cumulative distribution function, based on the maximum deviation. As distribution is estimated, the KS test does not have prior assumptions pertaining to data distribution (D'Agostino and Stephens 1986); thus, it is applicable for the estimation of multi/uni-modal non-Gaussian distributions with high classification accuracy.

In general, the identification of different spike waveforms is facilitated by more accessible features. However, the results show reduced classification accuracy with the use of increased wavelet coefficients (table 2). Although increasing the number of feature coefficients may improve classification accuracy at high noise levels, the use of additional feature coefficients may

result in over-fitting. Classification accuracy is affected due to irrelevant information from redundant feature coefficients (Saul and Roweis 2003, Kim 2006, Sanguinetti 2008).

Instead of the Euclidean distance, GRG is used as a measure of distance to estimate the level of similarities in waveforms for spike sorting in GSLC. Results show that spike sorting quality is improved with the current clustering method, as compared with SLC in the simulated data (table 3) and the electrophysiological data (table 4, J -measure). The use of an appropriate distance measure for the estimation of similarities is important for any clustering algorithm that could be affected by the shape of the clusters. That is, the distance between spikes may vary, depending on the distance measure used. In the clustering algorithm of a previous study (Thakur *et al* 2007), a Mahalanobis distance, rather than a Euclidean distance, was used due to non-Gaussian feature vector distributions. In the current study, GRG is used for series arrangement and for a comparison of relative intensity between series (Chang and Yeh 2005). Since the overall distribution of feature vectors can be estimated with GRG (Lin *et al* 2009), GRG was adopted as a measure of similarity for the feature vectors of spikes.

Determining the appropriate number of clusters has always been a challenge in the development of an automatic clustering algorithm. The number of clusters is automatically determined by locating the elbow with the maximum number of clusters and spikes. A maximum number of clusters indicate the most number of classes in the data, whereas a maximum number of spikes represent the least unclustered spikes. Accordingly, the number of clusters is determined by the maximum CNN (%) and CA (%). The elbow criterion is used in the graph of agglomerate coefficients to determine the number of clusters, but more than one elbow may exist (Aldenderfer and Blashfield 1984). Although the elbow criterion has been used to determine the location with an obvious curvature of the within-class variance in previous studies, the corresponding number of clusters does not necessarily attain the highest classification accuracy (Goutte *et al* 1999). The concept of the elbow criterion has been adopted as a universal method to determine the number of clusters, even though it may not always be unequivocally identified (David and Christopher 1996). In this study, variable linkage thresholds provide the means to merge adjacent clusters, resulting in the gradual variations of clusters and spikes. The unequivocal and unique position of the elbow is represented in the threshold-cumulative distribution. The proposed regulation is an effective method of quantification for determining the appropriate number of clusters.

As shown in the data, GSLC with WT-KS showed at least 3.23% improvement in classification accuracy from unsupervised clustering methods, such as SLC, SPC, and ETOS. In terms of the classification accuracy, the HMM performed better than the GSLC in the simulation data sets. Furthermore, the HMM algorithm combines the learning procedure (i.e. learning spike templates) and the reconstruction procedure (i.e. sorting subsequent spikes), of which neither relies on spike detection (Herbst *et al* 2008). The absence of a spike-detection step in HMMs may reduce the number

of missed spikes (false negative errors) in low signal-to-noise conditions, which resolves problems encountered by traditional spike detection during spike bursts or synchronous spiking events. However, unknown spike-onset probability, spike template, and noise variance may cause uncertain parameter learning; reconstruction error may result from the manually determined spike-onset times. The most important issue is constraints on memory and computation time, since the run of the learning and reconstruction procedures of HMM grows exponentially with the number of neurons (Herbst *et al* 2008). Given the unstable classification accuracy and potential learning problems, HMMs may not be feasible for on-line and real-time system implementation.

The proposed GSLC is an unsupervised spike sorting method using WT-KS for feature extraction, and GRG for the identification of similarities between spikes for clustering. Our strategy uses the elbow criterion for an accurate determination of the number of clusters. Results show an optimal performance of the clustering algorithm in simulated data and electrophysiological data. In conclusion, the GSLC attains high classification accuracy despite high noise levels. Thus, it is applicable in low-dimensional feature spaces for fully automated spike sorting. In future studies, the GSLC spike sorting method may be applied to on-line spike sorting to reduce manual labour and human errors.

Acknowledgments

The authors would like to thank the volunteers for their generous assistance with this study. They are also very grateful to Dr Yi-Chun Du for his expertise, valuable advice and availability. His feedback has made a significant contribution to the completion of this study. This study was supported by grant NSC 98-2221-E-009-142 from the National Science Council of the Republic of China (Taiwan) and grant VGHUST 99-P4-16 from the Veterans General Hospitals University System of Taiwan Joint Research Program, Tsou's Foundation.

Appendix. The quantitative determination of the maximum segment parameter

The CA (%) of spike sorting is best performed by the GSLC algorithm with a proper linkage threshold (v_q defined in equation (8) in section 2.4). This algorithm, however, is generic in the sense that the number of linkage thresholds is dependent upon the maximum segment parameter (Q). Sixteen simulated data sets (Quiroga *et al* 2004) were tested by the GSLC algorithm with various maximum segment parameters from 1 to 40. The relation between the empirical value of the maximum segment parameter and the CA (%) of the spike sorting was investigated and is illustrated in figure A1.

The experimental results obtained from simulated data sets ($N = 16$) indicate that the mean \pm S.D. in the CA (%) curve increased and decreased, respectively, in relation to increasing values of the maximum segment parameter Q . It is worth noting, however, that they reached a plateau (there was no further improvement in accuracy) when the maximum

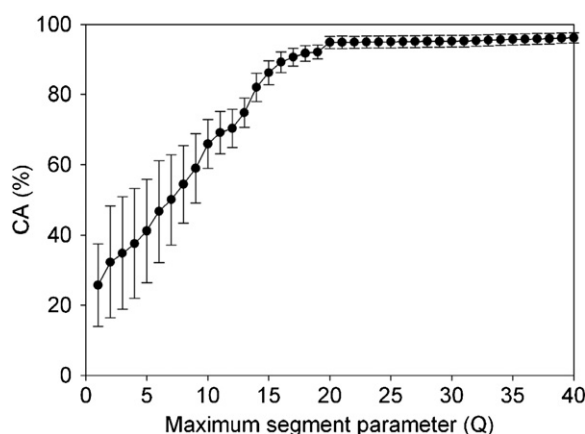


Figure A1. The CA (%) curve showing the relationship between the maximum segment parameter Q (x-axis) and the correct classifications as a percentage (y-axis). The mean CA (%) increases with increasing values of Q until it reaches an apparent plateau at $Q = 20$. Error bar: \pm standard derivation (S.D.) for this study.

segment parameter Q exceeded 20. Based on this data, an empirical Q value was determined as 20 and a novel reformulation of GSLC with the proper linkage threshold criterion was proposed.

References

- Adamos D A, Kosmidis E K and Theophilidis G 2008 Performance evaluation of PCA-based spike sorting algorithms *Comput. Methods Programs Biomed.* **91** 232–44
- Aharon B H, Spiro A and Stark E 2006 Spike sorting: Bayesian clustering of non-stationary data *J. Neurosci. Methods* **157** 303–16
- Aldenderfer M S and Blashfield R K 1984 *Cluster Analysis* (Beverly Hills, CA: Sage)
- Anil K J, Robert P W D and Jianchang M 2000 Statistical pattern recognition: a review *IEEE Trans. Pattern Anal. Mach. Intell.* **22** 4–37
- Azzeh M, Neagu D and Cowling P 2010 Fuzzy grey relational analysis for software effort estimation *J. Empirical Softw. Eng.* **15** 60–90
- Bankman I, Johnson K and Schneider W 1993 Optimal detection, classification, and superposition resolution in neural waveform recordings *IEEE Trans. Biomed. Eng.* **40** 836–41
- Brychta R J, Tuntrakool S, Appalsamy M, Keller N R, Robertson D, Shiavi R G and Diedrich A 2007 Wavelet methods for spike detection in mouse renal sympathetic nerve activity *IEEE Trans. Biomed. Eng.* **54** 82–93
- Carlos V I and Donoghue J P 2007 Automated spike sorting using density grid contour clustering and subtractive waveform decomposition *J. Neurosci. Methods* **164** 1–18
- Chandra R and Optican L M 1997 Detection, classification, and superposition resolution of action potentials in multiunit single-channel recordings by an on-line real-time neural network *IEEE Trans. Biomed. Eng.* **44** 403–12
- Chang K C and Yeh M F 2005 Grey relational analysis based approach for data clustering *IEE Proc.—Vis. Image Signal Process.* **152** 165–72
- Chapin J K 2004 Using multi-neuron population recordings for neural prosthetics *Nat. Neurosci.* **7** 452–5
- D'Agostino R and Stephens M 1986 *Goodness-of-Fit Techniques* (New York: Dekker)
- Daubechies I 1992 *Ten Lectures on Wavelets* (Philadelphia: SIAM)
- David J K and Christopher L S 1996 The application of cluster analysis in Strategic Management Research: an analysis and critique *Strateg. Manage. J.* **17** 441–58
- Delescluse M and Pouzat C 2006 Efficient spike-sorting of multi-state neurons using inter-spike intervals information *J. Neurosci. Methods* **150** 16–29
- Deng J L 1982 Control problems of grey systems *Syst. Control Lett.* **1** 288–94
- Donoho D L and Johnstone J M 1994 Ideal spatial adaptation by wavelet shrinkage *Biometrika* **81** 425–55
- Fee M S, Mitra P P and Kleinfeld D 1996 Variability of extracellular spike waveforms of cortical neurons *J. Neurophysiol.* **76** 3823–33
- Geng X, Hu G and Tian X 2010 Neural spike sorting using mathematical morphology, multiwavelets transform and hierarchical clustering *Neurocomputing* **73** 707–15
- Goutte C, Toft P, Rostrup E, Nielsen F A and Hansen L K 1999 On clustering fMRI time series *NeuroImage* **9** 298–310
- Gower J C and Ross G J S 1969 Minimum spanning trees and single linkage cluster analysis *J. R. Stat. Soc. C* **18** 54–64
- Hariharan G and Kannan K 2010 Haar wavelet method for solving FitzHugh–Nagumo equation *Int. J. Math. Stat. Sci.* **2** 281–5
- Herbst J A, Gammeter S, Ferrero D and Hahnloser R H R 2008 Spike sorting with hidden Markov models *J. Neurosci. Methods* **174** 126–34
- Hu J, Si J, Olson B P and He J 2005 Feature detection in motor cortical spikes by principal component analysis *IEEE Trans. Neural Syst. Rehabil. Eng.* **13** 256–62
- Hulata E, Segev R and Eshel B J 2002 A method for spike sorting and detection based on wavelet packets and Shannon's mutual information *J. Neurosci. Methods* **117** 1–12
- Hulata E, Segev R, Shapira Y, Benveniste M and Ben-Jacob E 2000 Detection and sorting of neural spikes using wavelet packets *Phys. Rev. Lett.* **85** 4637–40
- Jain A K, Murty M N and Flynn P J 1999 Data clustering: a review *ACM Comput. Surv.* **31** 264–323
- Kaneko H, Suzuki S, Okada J and Akamatsu M 1999 Multineuronal spike classification based on multisite electroderecting, whole-waveform analysis, and hierarchical clustering *IEEE Trans. Biomed. Eng.* **46** 280–90
- Kim K H 2006 Improved algorithm for fully-automated neural spike sorting based on projection pursuit and Gaussian mixture model *Int. J. Control Autom. Syst.* **4** 705–13
- Kim K H and Kim S J 2000 Neural spike sorting under nearly 0-dB signal-to-noise ratio using nonlinear energy operator and artificial neural-network classifier *IEEE Trans. Biomed. Eng.* **47** 1406–11
- Letelier J C and Weber P P 2000 Spike sorting based on discrete wavelet transform coefficients *J. Neurosci. Methods* **101** 93–106
- Lewicki M S 1998 A review of methods for spike sorting: the detection and classification of neural action potentials *Netw. Comput. Neural Syst.* **9** 53–78
- Lin Y H, Lee P C and Chang T P 2009 Practical expert diagnosis model based on the grey relational analysis technique *Expert Syst. Appl.* **36** 1523–8
- Morán J, Granada E, Míguez J L and Porteiro J 2006 Use of grey relational analysis to assess and optimize small biomass boilers *Fuel Process. Tech.* **87** 123–7
- Murtagh F 1983 A survey of recent advances in hierarchical clustering algorithms *Comput. J.* **26** 354–9
- Pavlov A, Makarov V, Makarova I and Panetsos F 2007 Sorting of neural spikes: when wavelet based methods outperform principal component analysis *Nat. Comput.* **6** 269–81
- Pouzat C, Mazar O and Laurent G 2002 Using noise signature to optimize spike-sorting and to assess neuronal classification quality *J. Neurosci. Methods* **122** 43–57
- Press W H, Teukolsky S A, Vetterling W T and Flannery B P 2002 *Numerical Recipes in C: the Art of Scientific Computing* (New York: Cambridge University Press)

- Quiroga R Q, Nadasdy Z and Ben-Shaul Y 2004 Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering *Neural Comput.* **16** 1661–87
- Roh S E, Choi J H and Kim T 2008 A new action potential classifier using 3-Gaussian model fitting *Neurocomputing* **71** 3631–4
- Rutishauser U, Schuman E M and Mamelak A N 2006 Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, *in vivo* *J. Neurosci. Methods* **154** 204–24
- Sanguinetti G 2008 Dimensionality reduction for clustered data sets *IEEE Trans. Pattern Anal. Mach. Intell.* **30** 535–40
- Saul L K and Roweis S T 2003 Think globally, fit locally: unsupervised learning of low dimensional manifolds *J. Mach. Learn. Res.* **4** 119–55
- Shoham S, Fellows M R and Normann R A 2003 Robust, automatic spike sorting using mixtures of multivariate t-distributions *J. Neurosci. Methods* **127** 111–22
- Stewart C M, Newlands S D and Perachio A A 2004 Spike detection, characterization, and discrimination using feature analysis software written in LabVIEW *Comput. Methods Programs Biomed.* **76** 239–51
- Takekawa T, Isomura Y and Fukai T 2010 Accurate spike sorting for multi-unit recordings *Eur. J. Neurosci.* **31** 263–72
- Thakur P H, Lu H, Hsiao S S and Johnson K O 2007 Automated optimal detection and classification of neural action potentials in extra-cellular recordings *J. Neurosci. Methods* **162** 364–76
- Venkatesh S, Ayyaswamy S and Hariharan G 2010 Haar wavelet method for solving initial and boundary value problems of Bratu-type *Int. J. Math. Stat. Sci.* **2** 64–7
- Vollgraf R and Obermayer K 2006 Improved optimal linear filters for the discrimination of multichannel waveform templates for spike-sorting applications *IEEE Signal Process. Lett.* **13** 121–4
- Wheeler B C 1999 *Methods for Neural Ensemble Recordings* ed M Nicolelis (Boca Raton, FL: CRC Press) pp 61–76
- Wood F, Black M J, Vargas I C, Fellows M and Donoghue J P 2004 On the variability of manual spike sorting *IEEE Trans. Biomed. Eng.* **51** 912–8
- Yang Z, Zhao Q and Liu W 2009 Improving spike separation using waveform derivatives *J. Neural Eng.* **6** 046006
- Yi D and Pin G 1992 *Gray Theory and Methods* (Beijing: Petroleum Industry)
- Zouridakis G and Tam D C 2000 Identification of reliable spike templates in multi-unit extracellular recordings using fuzzy clustering *Comput. Methods Programs Biomed.* **61** 91–8