ABSTRACT

Spike detection and sorting is a fundamental step in the analysis of extracellular neural recording. Here, we propose a combined spike detection-feature extraction algorithm that relies on a sparse representation space of the spike waveforms. The proposed method captures the wavelet footprint of the waveform, by calculating the power of the scale space vectors and finding an optimal detection threshold using histogram equalization techniques. Under the proposed scheme, a compact feature set is obtained simultaneously during detection, which eliminates the need for a separate feature extraction step for spike sorting. Our results demonstrate that this method yields improved performance, particularly in low SNRs, while preserving the separability between neuronal clusters in the feature space.

Index Terms—Sparse representation, wavelet footprint, spike detection, feature extraction

1. INTRODUCTION

Recording action potentials - or spikes - elicited by neurons is key to understanding the complex relationship between the physical features of the world and the brain’s interpretation and processing of those features. Spike detection refers to the determination of the arrival time of action potential waveforms produced by a neuron during active communication with other neurons in the nervous system. Since it is widely believed that spike arrival times carry most of the information about neuronal response to intrinsic or extrinsic inputs - and not the actual waveform shape - spike detection constitute a fundamental step in the analysis of neural recordings [1].

Many spike detection algorithms were proposed to overcome the inherent difficulties in detecting spikes embedded in background noise that is mostly neural. The simplest detection method is a spike amplitude threshold crossing method, but is highly sensitive to noise and often requires significant post detection analysis to discard spurious waveforms. Energy-based spike detectors, on the other hand, compare the local power of the signal with a threshold estimated from the noise power [2], which yields improved performance over the amplitude threshold method.

Wavelet-based methods have been motivated by the fact that the sparse representation of transient signals (such as neuronal spikes) enables superior identification of the information bearing properties of the waveforms. On one hand, they enable better separation of signal and noise, particularly in low signal-to-noise ratio (SNR) cases [3]-[5]. On another, they provide additional information about distinct properties of the signals such as discontinuities and smooth transitions, enough to capture the necessary information for compressive sensing [4].

Wavelet footprints have been proposed to detect and perfectly reconstruct piecewise linear signals [6]. Herein, we propose a spike detection and feature extraction algorithm based on this idea that enables assimilating information about spike transitions scattered across scale spaces. We further propose a threshold selection method based on histogram equalization to maximize separability between signal and noise. Under the proposed scheme, a compact feature set is obtained simultaneously during the detection, which further eliminates the need for spike extraction and alignment post-detection and prior to sorting the detected waveforms for multi-neuron analysis. We also demonstrate that the overall performance is not compromised as a result of this process relative to other known spike detection and feature extraction methods.

2. THEORY

2.1 Wavelet footprint

The projection of a continuous time signal $y(t)$ onto the subspace spanned by wavelet basis $\psi_{a,b}(t)$

$$y^{(a)}(t) = \int y(t) \psi_{a,b}(t) dt$$ (1)

where $y^{(a)}(t)$ is the sparse representation of the signal $y(t)$ at the $a^{th}$ scale [7]. Given a piecewise constant signal y with a single discontinuity at position k, the wavelet footprint $f^{(0)}_k$ is obtained by gathering the largest wavelet coefficient together from each subspace in the cone of influence (COI) of k [6]. The COI of k in the scale space is the set of points (a,b) such that k is included in the support of $\psi_{a,b}(k)$ [6]. The wavelet footprint $f^{(0)}_k$ can be written as

$$f^{(0)}_k = \max(y(t)) \left[ k - \frac{D_2}{2}; k + \frac{D_2}{2} \right], \ldots, \max(y(t)) \left[ k - \frac{D_2}{2}; k + \frac{D_2}{2} \right]$$ (2)

where $D_2$ is the duration of the COI of $1^{st}$ subspace at position k.

Typical spike waveform consists of three major phases: depolarization, repolarization, and hyperpolarization [4]. These phases create multiple discontinuities that can be modeled by multiple wavelet footprints. In theory, F+1 footprints are required to model piecewise polynomial signals with polynomials of maximum degree F [6].

The goal of this study is to detect and obtain a compact feature set for every spike without necessarily reconstructing the spike waveform. Therefore, it is not necessary to recover all F+1 wavelet footprints in the COI. Instead, we select the largest wavelet coefficient in each scale and construct a wavelet footprint feature set of each spike. To verify that this feature provides sufficient separability in a feature space, we calculated the scatter...
for the wavelet footprint $S_i$ by [8]

$$D_{ij} = \sum (m_i - f_j)^2$$  \hspace{1cm} (3)

where $m_i$ is the mean of sample wavelet footprints $S_i$ and $f_j$ is the wavelet footprint samples labeled $S_j$. Small $D_{ij}$ values mean that the samples within the class are less scattered, and the class is more compact in the feature space. Figure 1.b illustrates the scatter matrix of wavelet footprint samples from the five spike waveforms shown in figure 1.a. Spike 1 to 4 have distinguishable spike waveforms, while spike 4 and 5 are similar in shape. The $D_{ij}$ matrix illustrates that within-class scatter ($i=j$) is always smaller than between-class scatter ($i\neq j$). The smallest distance is between spike 4 and 5 due to their similarity.

### 2.2 Scale space vector detection model

In scale space detection, dyadic translation of wavelet transform poses a problem, because the truncated wavelet transform makes it difficult to define a scale space vector at time $t$. The Stationary Wavelet Transform (SWT), which we use here, overcomes this problem using its translation-invariance property [7]. By the linearity of SWT at scale $j$, $y$ can be expressed as

$$y^{(j)} = y^{(j),b} + y^{(j),s} = x^{(j)} + z^{(j)}$$  \hspace{1cm} (4)

where $y^{(j)} \in \mathbb{R}^{n \times N}$ expresses wavelet coefficients at $j$th subspace and $y^{(j),b}$, $y^{(j),s}$ is the corresponding wavelet base [8].

The detection problem can be expressed as a binary hypotheses testing problem, $H_0$ (no spike) and $H_1$ (spike present) of the form

- $H_0: y[n] = z[n]$ \hspace{1cm} $n = 0, 1, \ldots, N - 1$
- $H_1: y[n] = x[n] + z[n]$ \hspace{1cm} $n = 0, 1, \ldots, N - 1$

$$y[n] = [y^{(1)}[n], \ldots, y^{(L)}[n]]$$
$$z[n] = [x^{(1)}[n], \ldots, x^{(L)}[n]]$$
$$x[n] = [z^{(1)}[n], \ldots, z^{(L)}[n]]$$  \hspace{1cm} (5)

where $y[n]$, $x[n]$, and $z[n]$ are a scale space vector of the observation, the spikes, and the noise at time $n$.

When the signal is unknown, the optimal test is the generalized likelihood ratio test (GLRT) [4]

$$p(y|H_1) \overline{\triangleright} p(y|H_0)$$  \hspace{1cm} (6)

Under the assumption of a zero mean Gaussian noise, we have

$$p(y|H_1) \sim \mathcal{N}(\mu, \sigma^2)$$

and

$$p(y|H_0) \sim \mathcal{N}(0, \sigma^2)$$

where $\sigma$ and $\mu$ are the standard deviation and the mean of $y$ respectively. The GLRT can be simplified as

$$y^T \Sigma^{-1} y \geq R^2 \log \eta$$  \hspace{1cm} (7)

where $\Sigma$ is noise covariance.

The threshold $\eta$ in (6) is estimated considering the costs of false detection and a priori probability of $H_0$ and $H_1$ [9]. Since information of the true action potentials and the noise is unknown in unsupervised detection, $\eta$ is estimated as mean and variance of the noise [9]. The Median Absolute Deviation (MAD) is used to estimate the noise variance under Gaussian noise assumption [9]. The sufficient statistic

$$T = y^T \Sigma^{-1} y$$  \hspace{1cm} (8)

### 2.3 Modified histogram equalization

The sufficient statistic $T$ in (8) is $\chi^2$ distributed because it is the product of two Gaussian distributed variables. Figure 2.a illustrates the $\chi^2$ distribution of the sufficient statistic $T$, estimated from simulation data with three neurons. Bright bars represent histogram of noise and dark bars correspond to the signal. From this figure, there is no clear separation between the two distributions because the noise and the signal are $\chi^2$ distributed with very small shape parameter $v$.

To enhance this contrast, we used histogram equalization (HE) methods by linearizing the CDF of $T$ [10]. This process results in shifting $v$ from 1 to 3 or larger values where the distributions of the noise and the signal become more separable.

The modified HE (MHE) for $\chi^2$ distribution is expressed as

$$h'(k) = \begin{cases} 
 \text{cdf}(k) - \text{cdf}(k - 1) & k > 1 \\
 \text{cdf}(k) & k = 1
\end{cases}$$  \hspace{1cm} (10)

This modified histogram (MH) has $v$ greater than 1 and therefore the separation between the noise and signal is more likely. Figure 2.b shows the MH of $T$ with an improved separation between the noise and the signal distributions. The estimated threshold is located at a local minimum between the two peaks. This is similar to the gray-level histogram threshold selection in image processing [11]. After spikes are detected using the estimated threshold value, the wavelet footprint is extracted at each time stamp. Since the typical duration of single spike is 1-2 ms, the maximum range of the cone of influence was estimated to be 50 samples at a sampling rate of 25 kHz. The overall wavelet footprint detection and extraction process is illustrated in figure 3.
3. RESULTS

To evaluate the performance, simulated datasets with different SNRs were used. For comparison, the datasets were obtained from OSort, a spike detection and sorting package [12]. The proposed algorithm was fully implemented in NeuroQuest, a software package for neural data analysis, and all the results were obtained using NeuroQuest [13]. We used receiver operating characteristic (ROC) curves to examine the accuracy of each detection method. The SNR was calculated as the root-mean square (RMS) of the spike divided by the standard deviation of the noise [12]. Three commonly used spike detection methods, single amplitude detection (SAD) with a single threshold, absolute amplitude detection (AAD) with both positive and negative thresholds, energy-based detection (EBD), and the proposed wavelet detection (WD) were used for the comparison. For the WD, the observation signal was transformed into wavelet domain using 5-level SWT with a symlet4 wavelet base. The sufficient statistic $T$ was estimated from D2 to D4, where D denotes a detail node of wavelet transform, because typical action potentials have frequency range between 1 KHz to 5 KHz. As Figure 4 illustrated, SAD performed poorly in all cases, while WD performed better than the other methods especially in low SNR cases.

The effectiveness of MHE was evaluated by comparing the spike detection performance of WD with EBD. EBD was selected for the comparison because of its robust detection performance from the ROC test above and the $\chi$ distributed EBD samples that the MHE method can be applied to. Four different combinations of two detection methods and two threshold selection methods were tested. Dataset 3 from Osort was used for the comparison. FP rate is calculated by the number of false detection divided by the maximum false detection with a detection threshold value at 0. From the results shown in Table I, we concluded that WD is robust to noise and MHE yields low false detection rate for high positive detection rate.

Finally, we evaluated the feasibility of using wavelet footprint as a feature for spike sorting. Two feature sets, the wavelet footprint and the entire spike waveform, were compared to demonstrate spike sorting results. We applied the proposed method to extracellular recordings from the barrel cortex of an anesthetized rat. A total of 1849 spikes were detected, and the spike sorting results with the entire spike waveform and the wavelet footprints were compared. For display purposes, we selected only the largest two principal components (PCs) of each feature set. The spike sorting results were obtained using fuzzy c-mean clustering method.

The results of spike sorting are shown in figure 5. Three clear clusters were identified using the entire spike waveform with temp-
### Table I. Spike Detection Results

<table>
<thead>
<tr>
<th>Dataset3</th>
<th>SNR:5.4</th>
<th>SNR:2.7</th>
<th>SNR:1.8</th>
<th>SNR:1.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>EBD + 5 x MAD</td>
<td>TP:100%</td>
<td>TP:98.6%</td>
<td>TP:91.3%</td>
<td>TP:76.1%</td>
</tr>
<tr>
<td></td>
<td>FP: 0%</td>
<td>FP: 6%</td>
<td>FP: 12%</td>
<td>FP: 33%</td>
</tr>
<tr>
<td>EBD + MH</td>
<td>TP:98.6%</td>
<td>TP:93.5%</td>
<td>TP:64.1%</td>
<td>TP:54.4%</td>
</tr>
<tr>
<td></td>
<td>FP: 0%</td>
<td>FP: 2%</td>
<td>FP: 1%</td>
<td>FP: 2%</td>
</tr>
<tr>
<td>WD + 5 x MAD</td>
<td>TP:96.4%</td>
<td>TP:95.3%</td>
<td>TP:93.5%</td>
<td>TP:80%</td>
</tr>
<tr>
<td></td>
<td>FP: 1%</td>
<td>FP: 5%</td>
<td>FP: 48%</td>
<td>FP: 54%</td>
</tr>
<tr>
<td>WD + MH</td>
<td>TP:96.4%</td>
<td>TP:96.7%</td>
<td>TP:75.4%</td>
<td>TP:58.3%</td>
</tr>
<tr>
<td></td>
<td>FP: 1%</td>
<td>FP: 1%</td>
<td>FP: 1%</td>
<td>FP: 3%</td>
</tr>
</tbody>
</table>

Figure 5. Spike sorting results of spontaneous recordings from an anesthetized rat. 3 units could be isolated with Temporal PCA, whereas 4 units with wavelet footprints. Cluster 3 in temporal PCA was further separated into two clusters.

Oral PC, while four clusters were identified in the wavelet footprint. The identified clusters using the wavelet footprint were more scattered than those found in the temporal PC, which is not desirable. However, two clusters that were not separable in the temporal PCA domain could be separated.

### 4. Conclusion

Sparse representation of the extracellular recordings is key to achieve reliable spike detection in low SNR, and the wavelet transform exhibits advantages in obtaining this representation. We presented two contributions: 1) a novel spike detection/feature extraction method, capturing correlations across multiple scale subspaces; and 2) a threshold selection: modifying the distribution of the observation to maximize the separability between the noise and the signal. The proposed detection method was shown to outperform other methods in low SNRs in both simulated and experimental data.

Owing to the sparsity, the size of wavelet footprint feature set constitutes a small fraction of the entire spike waveform. Therefore a massive reduction in data size could be achieved. The proposed method also eliminates multiple processing steps in conventional methods, which is important for real time implementation in Brain Machine Interface applications [4].

### 5. References


