spike source identification
roberto a santiago
learning from data
2002 03 12

introduction

• In the treatment of Parkinson’s Disease, deep brain stimulator electrodes are employed to ablate or stimulate “malfucioning” neurons.

challenges to dbs procedures

• Success is highly dependent upon accurate placement of the electrode (probe).
• Currently, placement is determined by listening to the neurons near the probe tip taking into account depth and placement of probe.

microelectrode recording analysis

• The spiking of neurons near the tip of the probe can be recorded and analyzed (even in real-time).
• Through automated spike identification and spike sorting techniques, the spiking of individual neurons can be isolated.
• The set of temporal measurements indicating the time of each spike for an individual neuron is known as a spike train.
the challenge

• Does the spike train of an individual neuron contain enough information to identify what type of neuron it is?

• Constraints
  – 5 seconds or less
  – Identify as one of four types of neurons
    • Globus Pallidus Externus (GPE)
    • Globus Pallidus Internus (GPI)
    • Border (BRD)
    • Tremor (TRM)

data description

• An existing spike sorting algorithm was used on two sets of mer data to create three data sets.
• Each spike train in these data sets represents five seconds (or less) of individual neuron spiking
• Dirty Data – This data was recorded during a surgical procedure but was "labeled" after the procedure so its labeling lacks the benefits of probe placement and depth.
  – Train Data Set (47 spike trains)
  – Test Data Set (unknown)
• Starr Data – This data was recorded and labeled under optimal conditions and provides "perfect" examples of all cell types of interest.
  – Starr Data Set (47 spike trains)

the challenge

• WHERE ARE THE FEATURES?
  – Previous attempts to find a single statistic that identified the type of cell were not successful
    • Average Firing Rate
    • Average Interspike Interval (ISI)
    • Dominant Oscillation Frequency
  – Other feature extraction techniques have also failed to isolate unique identifiers
    • Histograms of ISIs
    • Spectral Estimation
    • Burst Analysis

some hints

• Spiking is the method through which neurons transmit information to one another. What do we know about the encoding of information through spikes?
  – Rate Coding Model – The number of spikes during a time period is most important.
  – Temporal Coding Model – The time between spikes is most important.

more hints

• Because of the physical properties of neurons, spiking behavior during a 5 to 10 ms time window can be mostly accounted for by incoming spikes and not the properties of the neuron.
• Repetitive spiking behavior from accountable from a neuron happens in a time window of no larger than 200 to 300 ms.
• A SET OF SPIKES IN A TIME WINDOW OF SIZE 10 TO 300 MS SHOULD REVEAL SOME SORT OF REGULAR PATTERN

spike trains

**temporal spike train (tst)** – provides the time (in seconds) for each spike

**digital spike train (dst)** – indicates whether a spike occurred in a given time interval. Each 1 and 0 represents an interval of time equal to the inverse of the sampling rate of the tst
new feature extraction method

RATE FEATURE EXTRACTION

00000000000000100001000100001000000000100000000000

Critical Factors: Sample Rate and Window Size

STRUCTURE FEATURE EXTRACTION

00000000000000100001000100001000000000100000000000

Count the number of ones and bin.

Convert the binary word into an integer and bin.

3

visualization method

RATE FEATURE EXTRACTION

000000000000000000000000100001000100001000000000100000000000

Critical Factors: Sample Rate and Window Size

STRUCTURE FEATURE EXTRACTION

00000000000000100001000100001000000000100000000000

Count the number of ones and bin.

Convert the binary word into an integer and bin.

1,2

Critical Factors: Sample Rate and Window Size

1,2
• The rate features were used to “train” a support vector machine (SVM) using a linear basis function.
• Because SVMs are computationally efficient, leave one out cross validation was used to test the classification method.

---

**Classification Method**

---

**Results on Starr Data**

<table>
<thead>
<tr>
<th>TOTAL</th>
<th>GPE</th>
<th>GPI</th>
<th>BRD</th>
<th>TRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPE</td>
<td>13</td>
<td>13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>GPI</td>
<td>19</td>
<td>0</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>BRD</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>TRM</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
</tbody>
</table>

---

**Results on Dirty Train Data**

<table>
<thead>
<tr>
<th>TOTAL</th>
<th>GPE</th>
<th>GPI</th>
<th>BRD</th>
<th>TRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPE</td>
<td>33</td>
<td>31</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>GPI</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>BRD</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>TRM</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>