Proceedings of the 3rd International
IEEE EMBS Conference on Neural Engineering
Kohala Coast, Hawaii, USA, May 2-5, 2007

Neural Data Compression with Wavelet Transform: A Vocabulary Based Approach

Seetharam Narasimhan, Massood Tabib-Azar, Hillel J. Chiel[1], and Swarup Bhunia
Case Western Reserve University, Cleveland, OH, USA
Email: {sxn124, mxt7, hjc, skb21}@case.edu

Abstract—**An algorithmic approach to develop the vocabulary of the nervous system and to use the vocabulary to communicate with the outside world is presented. The vocabulary is constructed using wavelet analysis of the recorded waveforms. Spikes of different frequency and amplitude from different channels are identified to construct unique signatures and relate them to physiological activities. A vocabulary-based communication of recorded action potentials renders two major advantages: a) it allows transmission of recorded data with large compression, thus, saving power and communication bandwidth of the integrated telemetry device; b) it helps easy mapping of alphabets in the vocabulary to muscular dynamics, which facilitates micro-stimulation based neural prostheses. In this work, we study the effectiveness of the proposed approach in neural data compression. Simulation results on pre-recorded data from the buccal nerves of a sea-slug shows that the proposed approach results in up to 80X compression.**

I. Introduction

Researchers and scientists are working hard towards understanding the central nervous system and performing surgical interventions to manipulate its activity. With the advances in bio-MEMS and microelectronics, interpreting and engineering the activity of the central nervous system are likely to make dramatic progress using miniaturized implantable biomedical microsystems [1-13]. These implants offer new possibilities in terms of medical diagnosis and therapy [2, 12-13]. Neural rehabilitation using micro-stimulation appears promising for possible treatment of severe disorders such as blindness, deafness, epilepsy, paralysis and Parkinson's disease.

In late 1950s, Gesteland et al. used microelectrodes together with electronic recording and signal processing to investigate the operation of the central nervous system at a cellular level [4]. Since then, numerous efforts [5-14] have been pursued in this direction. Currently, it is possible to precisely measure the activity of very few neurons (e.g. using intracellular electrodes); or to obtain an overall measure of activities in large neural assemblies (using techniques such as EEG or fMRI) [10].

Neuronal action potentials typically appear as a spike in the recorded waveform, where a spike means a concentrated short-time signal with high frequency band [8]. In order to use the recorded information containing action potentials mixed with background noise, we need to perform spike detection and alignment of the action potential signals. A typical neuronal recording setup with animal or human subjects requires high-bandwidth communication between the recording electrodes and a backend computer, where spikes are processed. When hundreds of electrodes are employed for parallel recording, transmission resource of the telemetry device (e.g. bandwidth) become insufficient and power-hungry [2, 13]. Therefore, it is extremely urgent to use on-chip electronics for pre-processing and data compression, which requires efficient signal-processing algorithms and special-purpose customized hardware to implement them on-chip.

Large numbers of algorithms have been proposed to date for automatic spike detection, alignment and sorting. In an unsupervised spike detection and classification, an adaptive thresholding scheme is often used for detection, while principal component analysis [11] and fuzzy c-means (FCM) clustering are commonly used for spike classification. Another detection technique based on the Teager energy Operator (TEO) is reported to show relatively good performance at low signal to noise ratios (SNRs) [8]. Wavelet transform based spike analysis algorithms have also been investigated earlier, primarily in the context of off-line action-potential detection [7].

In this paper, we develop a novel classification algorithm for action potential signal using wavelet analysis based *"vocabulary"* development. In the wavelet coefficient space, we define distinct hyper-clusters that constitute the elements of our vocabulary. The spike detection process effectively de-noises the signal, thereby removing the biological noise (that includes action potentials from neural cells far from the measuring electrodes) and other electrical noise [8]. Next, the spikes are identified from the wavelet coefficients of de-noised signal and matched with the existing set of spikes (which we refer as "*alphabets*") in the vocabulary. The process of creating the vocabulary and representing the spikes as alphabets results in significant compression of recorded neural signal.

Our proposed setup will consist of an implanted chip which monitors the neuronal action potentials in a sea slug. Outputs of multiple sensors go to the bio-telemetry system, which performs spike detection, wavelet analysis based classification and vocabulary construction. Next, it encodes the recorded signals in terms of alphabets in the vocabulary before transmitting wirelessly to the outside world using the built-in telemetry device.

The organization of the paper is as follows. In the next section, we describe the algorithm employed for neural data compression. In section III, the simulation results are presented. Finally, we conclude in section IV.

[1]*Dr. Chiel was supported by NIH grant NS047073.*

II. THE VOCABULARY-BASED APPROACH

The initial approach towards implementing the proposed idea of vocabulary based neuronal signal analysis and resultant data compression can be explained with the flow diagram depicted in Fig. 1. Input to the algorithm is the recorded analog neural data, which is amplified and converted to digital form. The output is the compressed neural data encoded as packets containing wavelet domain information about the recorded signal. Major data processing steps of the algorithm are described below.

i) Multi-Resolution Wavelet Analysis: In the first stage, the neural signal is segmented into blocks of equal size and wavelet transformation is applied to each block. The wavelet transform is similar to Fourier analysis, with the exponential (or sinusoidal) matching function replaced by a finite-time waveform, with a particular shape, called the "mother wavelet". This transformation is performed by multi-resolution Discrete Wavelet Transform (DWT), using Daubechies-3 (db3) basis function [17]. Here, we chose the basis function as one of the built-in mother wavelets in the wavelet toolbox provided by MATLAB, for ease of simulation. For better reconstruction quality, we propose to design a custom mother wavelet based on the shape of a typical action potential. Wavelet transform decomposes the recorded signal in both time and frequency domain [7, 10], which turns out to be very useful in feature extraction [15-16]. It generates a set of coefficients corresponding to the low-frequency components, called the "approximations" and another set of coefficients corresponding to the high-frequency components, called the "details", while preserving the timing information. If we consider only the approximations for reconstructing the neural signal by using an inverse wavelet transform, then we achieve our first level of data compression, along with de-noising. The parameters of importance are the mother wavelet used for the wavelet decomposition and the number of levels of decomposition performed on the signal. The former determines the efficiency of matching with the action potentials, which increases if the mother wavelet resembles a typical spike in shape. The latter signifies the amount of data compression achieved in this first step, without having to lose important information about the recorded signal.

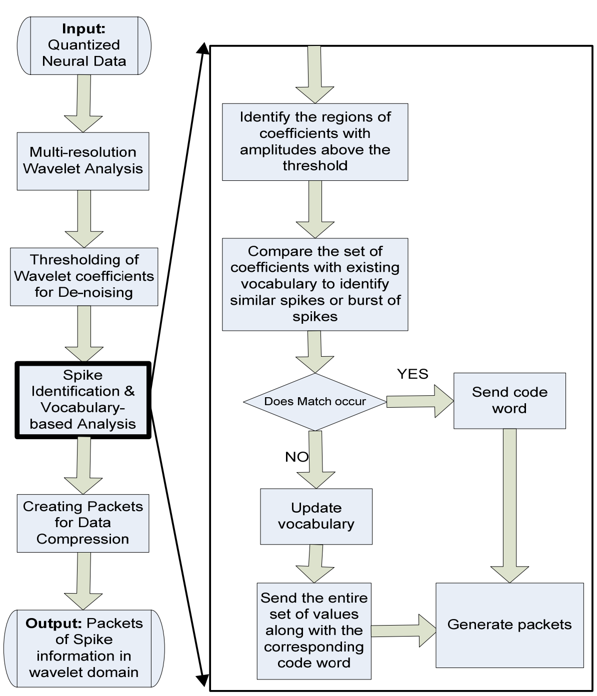


Figure 1: Flowchart of the proposed vocabulary-based neural data compression algorithm.

To implement wavelet transform in hardware, appropriate architecture for low-power discrete wavelet transform (DWT) has to be investigated. There is a number of emerging VLSI architectures for hardware implementation of the wavelet transform algorithm, such as liftoff-based forward and reverse wavelet transform, which requires fewer computations and less memory than previous approaches. Existing hardware architectures for wavelet transform can be exploited to develop performance and power-efficient VLSI implementation of the proposed algorithm.

ii) Thresholding: Once the wavelet coefficients corresponding to the "windowed" signal for a particular mother wavelet and a certain level of decomposition are generated, we can perform a thresholding step. This helps in removing most of the noise from the wavelet domain, by preserving only the most significant coefficients. Further, since we perform the thresholding in the wavelet domain, the timing information corresponding to the location of spikes is maintained with negligible error. The parameter of importance here is the thresholding level. It can be either fixed or variable. If variable, then it can be recomputed for each window separately or once in a while. The thresholding step is a compromise between information retention and data compression. If we keep a high threshold, the latter will be very high, but we might miss out on some spikes. On the other hand, if the threshold is too low, we end up with lot of false positives and a less compression ratio. Since it is important to avoid false negatives, the threshold values must be chosen optimally.

Corresponding to each window of the signal, we have a set of wavelet coefficients with some non-zero values, which correspond to the spikes of the neural signal. We intend to create data packets from these coefficients such that we need to send only the information regarding location, duration and amplitudes of a set of consecutive non-zero approximation coefficients. This "packetizing" algorithm is aimed at further data compression.

iii) Vocabulary Based Analysis: From each set of consecutive non-zero approximation coefficients, which can be considered as the wavelet analogue of a spike or a burst of spikes in the time domain, we construct an alphabet of the vocabulary. This vocabulary is aimed at distinguishing between different spikes occurring on the same channel and also at identifying the similar spikes. The similarity is noted on the basis of the duration and the amplitudes of the approximation coefficients. As the vocabulary is constructed online, we are basically generating a signature for each spike or burst of spikes. As the spikes are detected in the wavelet domain, they are compared with existing signatures. If a match is found, we need to send only the signature and not the entire set of amplitude values. Thus we attain a significant amount of data compression by transmitting the recorded data using vocabulary.

Note that one can employ the vocabulary concept on the de-noised time-domain data. The implementation of the vocabulary concept with wavelet coefficients exploits two advantages: 1) number of data points is reduced and 2) high-frequency noise from the spike is removed, both of which help in spike analysis.

iv) Creating Data Packets: A typical packet consists of the location, duration, signature and the amplitudes corresponding to the non-zero approximation coefficients. We also keep one bit for identifying whether a match for the identified spike has been found in the existing vocabulary or not. If yes, then we do not need to send the amplitudes again. If it is a newly generated signature, then the corresponding information has to be sent. Hence, the data packets have variable length, based on the duration and also on whether the signature is repeated or not. A frame identification bit is needed to signify the beginning of each window of the signal. This bit is also useful for synchronization purpose.

In the initial stage, when the vocabulary is being constructed and updated, there will not be considerable data compression. Once the warm-up period is over i.e. when the vocabulary is built up, the data compression is considerable. It can be increased by pre-computation of all possible signatures and embedding the vocabulary inside the chip and only updating the vocabulary whenever necessary.

III. Simulation Results

Simulations are carried out on pre-recorded neural data from the buccal nerves of a sea-slug (*Aplysia californica*). The data contained intracellular recordings of the stimulus signal generated from the cell body as well as the corresponding extracellular recordings from an electrode near the end of its axon. This particular data is used to show the results so that we can identify which of the reconstructed spikes correspond to "real" spikes. Simulations were also carried out on extracellular data recorded from different neurons to check the effectiveness of the algorithm. The simulation tool used is MATLAB [17]. Fig. 2 shows the different observations. The entire waveform is plotted in Fig. 2(a), which shows that all the spikes are captured, with excellent de-noising. Single spikes are shown in Fig. 2(b) to demonstrate the quality of reconstruction. Though the timing is almost accurate, there is observable mismatch in the amplitudes of the reconstructed and recorded spikes. A better reconstruction quality is achieved by including some detail coefficients. A series of spikes are shown in Fig. 2(c) including a spike whose positive peak was not captured. The only false spike which showed up is depicted in Fig. 2(d).

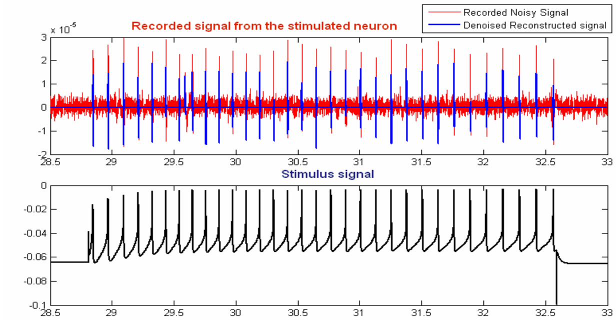
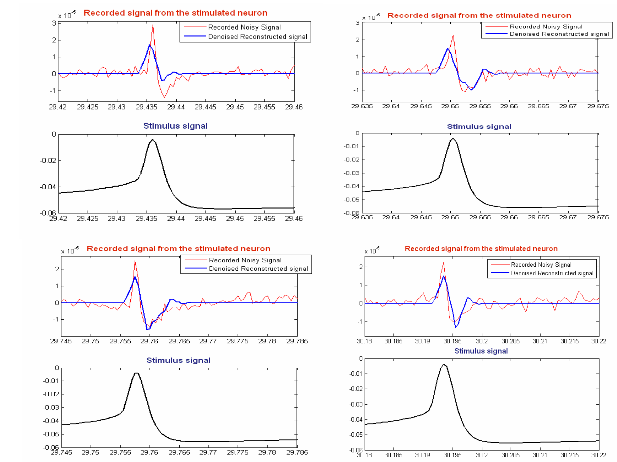
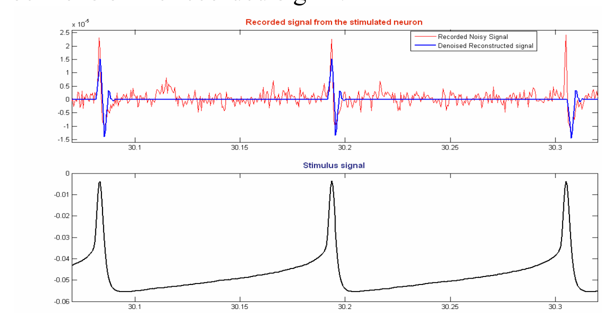


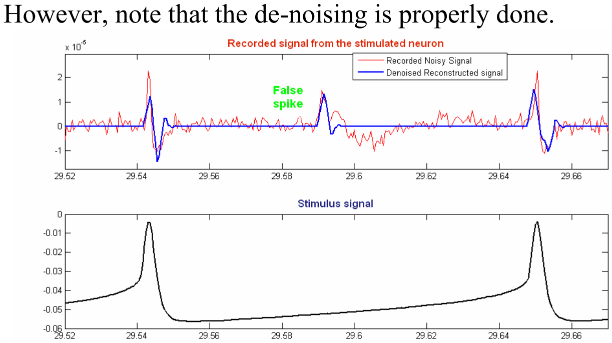
Figure 2: Simulation results showing the original and reconstructed waveforms from an extracellular neural recording along with the corresponding stimulus signals. a) The full waveform (top) with stimulus (below).



b) Comparison of original and reconstructed spike waveform for four individual spikes from different locations of the recorded signal.



c) Part of the reconstructed waveform. The third spike is not faithfully reconstructed (positive part is missing). However, note that the de-noising is properly done.



d) The second spike in the reconstructed waveform is a false positive (the only case of mismatch).

TABLE 1: Performance of the proposed algorithm

Sample window	Data compression		Percentage of matches with existing alphabets in the vocabulary	
	W/o warm-up period	W/ warm-up period	W/o warm-up period	W/ warm-up period
1	30.85X	31.87X	84.20%	89.00%
2	34.21X	35.70X	79.24%	85.85%
3	80.62X	83.29X	90.17%	95.62%
4	76.62X	80.29X	89.34%	96.75%

Simulations with the different built-in mother wavelets showed that the best reconstruction quality is obtained with the *'db3'* mother wavelet, since it resembles, to some extent, the action potential. The number of levels of wavelet decomposition is chosen as two, since the reconstructed signal quality is degraded beyond that value. The threshold for de-noising was chosen such that we obtain the best reconstruction (only one false spike and no missed spikes).

The compression ratio was computed as follows. Considering 8 bits for representing the amplitude of each data point, the total number of bits in the uncompressed signal is 72000, since 9000 sampling time instants have been considered. Using the proposed algorithm, we can send the information in wavelet domain using 42 data packets, with the total number of bits being 1029, which yields 69.97X compression. Here we considered 6 bits for the location of the "spike", 3 bits for the duration of the "spike", 4 bits for the amplitude of a non-zero value and 5 bits for the signature or the alphabet. We also considered 1 bit for frame synchronization purpose.

If we consider fixed-size (2^5=32) of the vocabulary then we need to replace existing alphabets with newly acquired alphabets and resend the information. Hence, we keep 1 additional bit to signify whether the "spike" matches an existing alphabet or not. This corresponds to the "page table" concept used in computer systems. The replacement algorithm used is Least Frequently Used (LFU).

The compression ratios using the proposed vocabulary-based approach were noted for different simulation runs and found to vary between 30X to 80X, depending on whether we allow a warm-up period or not and also on the spike density in a particular sample window. Results for extracellular recording on buccal nerve B1 of a sea-slug are tabulated in Table 1. Two cases were considered for each of the 4 samples of duration 66.66 sec (2×10^5 sample points): 1) with a fresh vocabulary being generated each time (i.e. without any warm-up period); and 2) with a previously built-up vocabulary (i.e. with a warm-up period).

IV. Conclusion

The algorithm presented in this paper is not only effective from the point of view of the data compression achieved, but also has far-reaching significance in neural engineering. The classification of spikes is done online along with their detection. Thus, by pushing more signal-processing circuitry on-chip, we can attain data compression along with information expansion, since the savings in bandwidth can be utilized for recording from more channels.

From the point of view of neurologists, by studying the nature of a spike or burst of spikes, including their shapes and frequency of occurrence, one can make important deductions regarding the physiological activity going within the animal or human body. The wavelet transform coefficients of these spikes encoded as alphabets and used to build the "vocabulary" of the nervous system, can help in correlating behavior of the neural system with muscular dynamics, which can subsequently help in stimulation-based neural rehabilitation.

The future work is aimed at getting near-perfect spike reconstruction by using a customized mother wavelet, which can be generated based on the spike characteristics from a recorded signal. A possible inclusion of variable length encoding can increase the currently achieved compression.

References

[1] A. L. Nicolelis, "Actions from thoughts", *Nature*, pp. 403-407, 2001.

[2] K.D. Wise et al., "Wireless implantable microsystems: High-density electronic interfaces to the nervous system", *Proceedings of the IEEE*, 92 (1), pp. 76-97, Jan 2004.

[3] B. Gosselin, V. Simard and M. Sawan, "Low-power implantable microsystem intended to multichannel cortical recording", *IEEE Circuits and Systems*, Vol. 4, 23-26, May 2004, pp. IV - 5-8.

[4] R. C. Gesteland, B. Howland, J. Y. Lettvin and W. H. Pitts, "Comments on microelectrodes", *Porc. IRE,* Vol. 47, pp. 1856-1862.

[5] R.H. Olsson III and K.D. Wise, "A three-dimensional neural recording microsystem with implantable data compression circuitry", *IEEE Journal of Solid-State Circuits,* Dec. 2005, pp. 2796-2804.

[6] Wilson & Emerson, "Spike detection: a review and comparison of algorithms", *Clinical Neurophysiology*, 113, pp. 1873-1881, 2002.

[7] K.H. Kim and S.J. Kim, "A wavelet-based method for action potential detection from extra cellular neural signal recording with low signal-to-noise ratio", *IEEE Trans. on Biomed. Eng.,* pp. 999-1011, 2003.

[8] J.H. Choi, H.K. Jung and T. Kim, "A new action potential detector using the MTEO and its effects on spike sorting systems at low signal-to-noise ratios", *IEEE Transactions on Biomedical Engineering,* 53 (4), pp. 738-746, Apr 2006.

[9] A. Snellings, D.J. Anderson and J.W. Aldridge, "Improved signal and reduced noise in neural recordings from close spaced electrode arrays using independent component analysis as a preprocessor", *Journal of Meuroscienece Methods,* 150 (2), pp. 254-264, 2006.

[10] X. Hu, Z.Z. Wang and X.M. Ren, "Classification of surface EMG signal using relative wavelet packet energy", *Computer Methods and Programs in Biomedicine*, 79 (3), pp. 189-195, Sep 2005.

[11] A. Zviagintsev, Y. Perelman and R. Ginosar, "Algorithms and architectures for low power spike detection and alignment", *Journal of Neural Engineering,* 3 (1), pp. 35-42, Mar 2006.

[12] T. Stieglitz, "Implantable microsystems for monitoring and neural rehabilitation – Part I", *Med Device Technol*, 2001, pp. 16-8, 20-1.

[13] M. Sawan, H. Yamu and J.Coulombe, "Wireless smart implants dedicated to multichannel monitoring and micro stimulation", *IEEE Circuits and Systems Magazine*, Vol. 5, Issue 1, 2005, pp. 21 – 39.

[14] G.P. Sutton, E.V. Mangan, D.M. Neustadter, R.D. Beer, P.E. Crago and H.J. Chiel, "Neural control exploits changing mechanical advantage and context dependence to generate different feeding responses in *Aplysia*", *Biol Cybern* 91, pp. 333-345.

[15] I. Daubechies, "Ten Lectures on Wavelets", PA: *SIAM*, 1992.

[16] R. Rao and A. Bopardikar, "Wavelet Transforms: Introduction to Theory and Applications", *Reading, MA*: *Addison-Wesley*, 1998.

[17] Matlab Wavelet Toolbox, Version 2.1. Mathworks, Natick, MA. [Online] http://www.mathworks.com/products/wavelet