8th International IEEE EMBS Conference on Neural Engineering
Shanghai, China, May 25 - 28 , 2017

Use of Wavelet Transform Coefficients for Spike Detection for a Robust Intracortical Brain Machine Interface*

Gary C. F. Lee[1], Camilo Libedinsky[2], Cuntai Guan[3] and Rosa So[1]

***Abstract*— A common problem in Brain-Machine Interface (BMI) is the variations in neural signals over time, leading to significant decrease in decoding performance if the decoder is not re-trained. However, frequent re-training is not practical in real use case. In our work, we found that a temporally more robust system may be achieved through the use of wavelet transform in feature extraction. We used wavelet transform coefficients as means to detect spikes in neural recordings, in contrast to conventional amplitude threshold methods. Using offline data as the preliminary testbed, we showed that decoding based on firing rates determined from four levels of wavelet transform decomposition resulted in a decoder with 6-12% improvement in accuracy sustained over four weeks after training. This strategy suggests that wavelet transform coefficients for spike detection may be more temporally robust as features for decoding, and offers a good starting point for further improvements to tackle nonstationarities in BMI.**

I. INTRODUCTION

Brain-Machine Interfaces (BMI) has the potential to be developed into assistive technologies for patients with severe motor disabilities by translating their thoughts to control external devices. Such devices, which include computers [1], [2], robotic arms [3], [4] or wheelchairs/mobile platforms [5], [6], can give these patients the ability to accomplish tasks that they would otherwise be unable to. In a typical BMI system, there are five key components, namely, 1) the signal acquisition system, 2) signal pre-processing, 3) feature extraction, 4) decoding algorithm and 5) control interface to the external device [7]. Using existing neural recording devices, our work primarily focuses on the feature extraction and decoding algorithm (3 and 4) to achieve optimal performance and robustness in the BMI decoder.

In our work, we focused on intracortical BMI system, where signals are recorded from the motor cortex. The development of multi-electrode arrays has given us access to intracortical recordings, thereby allowing us to record single-unit spiking activity of neurons. Many BMI algorithms have used neural firing rates as inputs, given that there has been strong foundation in the neural tuning between firing rates and directions [8].

However, BMI systems may not be temporally stable. This variability in recordings may occur for a variety of reasons, from variations in behavior, micro-movements in the implant to changes in electrode impedance during chronic implantation [9]. As a result, the BMI decoder would result in degraded performance over time in the absence of re-training [11]. However, constant recalibration and re-training is not practical during actual deployment as it causes inconvenience and frustration to the user.

One strategy is to investigate the extraction of features that are temporally more stable. Given that neural recordings contain spikes against large background noise, a robust yet simple spike detection would help in BMI performance. In this paper, we propose the use of wavelet transform as a simple feature extraction method. This paper presents our initial findings on the relevance of wavelet transform coefficients (WTCs) for spike detection in intracortical BMI decoders. In our offline tests based on pre-existing dataset collected, we found that decoders using firing rates determined from spikes in WTCs may be temporally more robust than using firing rates from conventional amplitude threshold methods, and can sustain a higher performance across weeks after training.

II. PRELIMINARIES

A. Wavelet Transform

Wavelet transform is a time-frequency analysis tool that decomposes signals into a set of wavelets, which are waveforms with compact support. Such a decomposition provides information on different frequency sub-bands while preserving its temporal information. This information is obtained through the convolution of the signal and a wavelet function, which is a scaled and shifted version of the wavelet basis (also termed as mother wavelet). Scaling by geometric series of 2 at each level of decomposition is commonly used, particularly for discrete-time signals.

In this respect, wavelet transform resembles short-time Fourier transform in providing frequency information over time; however, due to the scaling of wavelets at each decomposition level, wavelet transform provides high temporal resolution for high frequency components and high spatial resolution for low frequency components. In this work, we are focusing on the high frequency components. Furthermore, the convolution with a well-selected wavelet would aid in highlighting neural spikes and suppressing background noise. We chose 'symmlet4' as our wavelet basis, as it has been described to be the optimal representation of neural signals amongst well-established wavelet types [12].

In traditional discrete wavelet transform, the outputs are decimated (i.e., down-sampled) by two at each level without causing any loss in information. However, such a scheme results in the loss of the shift-invariant property of the

*This work was supported by Agency for Science, Technology and Research (A*STAR), Singapore.

[1]Authors are with the Institute for Infocomm Research, Singapore rosa-so@i2r.a-star.edu.sg
[2] Author is with SiNAPSE, National University of Singapore, Singapore camilo@nus.edu.sg
[3] Author is with Nanyang Technological University, Singapore ctguan@ntu.edu.sg

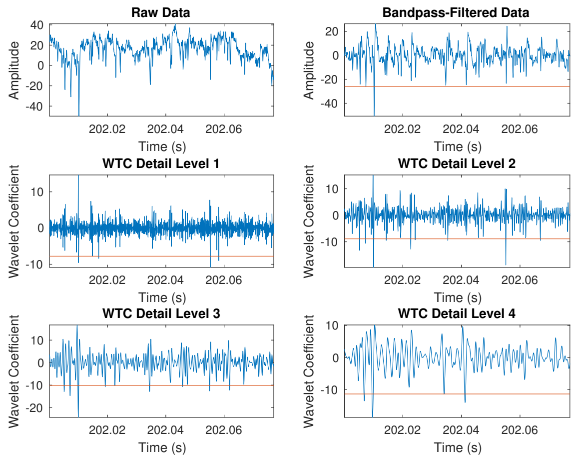

Fig. 1. Comparison of spikes detected in bandpass-filtered data to those detected in WTCs detail levels 1 to 4 (high frequency components)

wavelet transform, which would affect the accurate temporal localisation of spikes. Hence, we adopt the use of stationary wavelet transform (SWT), where the output is not decimated. This results in redundancy in data, but is shift-invariant and thus a more reliable representation of the spikes present.

The use of wavelet transforms is not new to neuroscience and BMI. In neuroscience, wavelet transforms have been used for spike detection [12], [13], spike sorting [14] and denoising of neural signals [15]. In BMI, it has been used with EEG and ECoG recordings to capture oscillatory components with good temporal resolution [16], [17]. However, the use of WTCs in intracortical BMI has not been well studied. The key contribution of our work is in showing the improvement WTC-based spike detection can offer to intracortical BMI performance and robustness, without the need for further intensive processing such as spike sorting or signal reconstruction from the WTCs.

B. Feature Extraction with Wavelet Transform

We propose the use of wavelet transform for spike detection due to its ability to localise high frequency information in the signal. At lower detail levels (higher frequency bands), the coefficients correspond to spiking activity, wherein the presence of a spike (similar in shape to the wavelet) corresponds to a large coefficient. Therefore, we can deploy a threshold on the WTCs to detect spikes and obtain their respective spike timings.

In this work, we took the first four detail levels of WTCs. This results in four times as many features. Given that each detail level corresponds to the convolution with differently scaled wavelets, differently shaped spikes in the raw signal may be captured as spikes in certain detail levels but not in others. Figure 1 is a representative window of the spikes detected through WTCs, compared to conventional bandpass-filtering and amplitude threshold. Note that this is not a robust spike sorting process based on matching waveforms or features, but would provide a quick and simple way of detecting spikes and differentiating distinctly different waveforms from a single recording channel. However, the relationship between spikes in raw recording and the profile in its four detail levels may require further study and characterization to better understand the kinds of features WTCs can capture.

In addition, we report that the use of WTCs can provide spike timings without the need for reconstruction for each frequency sub-band. Firing rates based on spike timings determined by WTCs can be used for decoding with comparable performance, and may be more robust to any inter-day variations in the raw signal.

III. PROPOSED DECODING ALGORITHM

A. Dataset

The performance of the decoding algorithm was assessed through offline tests with datasets collected from one macaque monkey. Details of the experiment are described in [6]. Briefly, the intracortical neural data was obtained from a macaque, implanted with multi-electrode arrays and seated on a robotic platform. The animal was trained to control the robotic platform with a 3-direction joystick to follow one of four commands: turn right for 90°, turn left for 90°, move forward for 2m, or stay still for at least 5s (right, left, forward and stop respectively). The data used was collected over seven different days within a two-month span. The data from each experimental day contained four to five sessions, each session with 20 successful trials (5 in each direction). All sessions were conducted with the monkey controlling the robotic platform using a joystick.

For our offline analysis, our goal was to develop a high performing classifier that can translate the monkey's neural data into its control intent at every 100ms timepoint. The ground truth of the monkey's intent was obtained from the joystick position.

B. Workflow

The offline decoding was tested on MATLAB (MathWorks Inc, Massachusetts, USA). The raw data was pre-processed to obtain the corresponding neural recording signals (100 channels), the accompanying joystick position data and the experimental session data (such as trial start and end times).

First, we establish the conventional method of calculating firing rates. The raw signals are filtered with a 2nd order elliptic bandpass filter with passband frequencies from 300 to 3000Hz, 0.1dB passband ripple and 40dB of stopband attenuation. Spikes are detected through a simple amplitude threshold using data from the first 30s from the respective sessions, wherein the threshold value for the channel i is determined using the following expression [14]:

$$Thr_i = 5 * median\{|x_i|\}/0.675 \quad (1)$$

where x_i are the values in the recorded signal from channel i. The time associated with negative crossings from the threshold (i.e., spikes with amplitude less than negative of the threshold) was determined, and the number of spikes detected in each 100ms window gave the firing rates. We also calculated the firing rate in 500ms moving window at every 100ms. The latter strategy has been empirically shown to perform better than decoding with non-overlapping 100ms bins. It is believed that the inclusion of such long history may be more resistant to noise and provide more accurate information on the directional intent. However, we note that there is a compromise in refresh rate of the control commands, and that the response time to changes may be slowed due to the effects of longer history in the features used. Additionally, noting that not all channels contain information for the decoding, we only used features that have non-zero in-class variance for each of the four classes.

For the new method proposed here, the key modification from the conventional method is the replacement of the elliptic filter with an SWT. We performed SWT and obtain the first four detail levels (discarding the remaining approximate level). The undecimated fast wavelet transform function `ufwt` from the LTFAT toolbox [18], [19] (also listed as SWT in the accompanying description) was used to filter the raw neural signals. Since the transform assumes periodic extension, the signal was de-meaned and zero-padded with half the signal length before and after the signal prior to taking the wavelet transform. Subsequently, we detected the spikes through thresholding, using the same equation above, and calculated the "firing rates" per detail level for each channel at every 100ms non-overlapping window.

With the conventional firing rates (for 100ms non-overlapping and 500ms moving windows) and the WTC-based firing rates (for 100ms non-overlapping bins), we used linear discriminant analysis (LDA) to perform classification. Since we are investigating the temporal robustness, our training set comprises of all observations from sessions recorded on the first three experimental days, and the testing set is each session for the remaining four experimental days after, which span across one month.

IV. RESULTS AND DISCUSSION

Accuracy was calculated by the percentage of observations at each 100ms timepoint that were correctly classified, based on the ground truth from the joystick position. Figure 2 shows the comparison of 4-class decoding accuracy for each untrained experimental day; the error bars correspond to the standard deviation of accuracy scores across the sessions in the given experimental day. The classifier performance on trained days is also reported using a 10-fold cross-validation.

Generally, on the untrained days, the use of WTC-based firing rates resulted in improved decoding accuracy over conventional firing rates for both 100ms and 500ms windows. The accuracy with our algorithm is about 60 to 70%. As mentioned in our previous study [6], we require roughly 70% accuracy in decoded commands directed towards the target direction for good control of the mobile platform. We recognize that, despite the improvement in decoding accuracy when WTCs were used, the performance for untrained days falls slightly short of 70%. One of the contributing factors may be the nonstationarity in one of the classes, the 'stop' command, as discussed in [20].

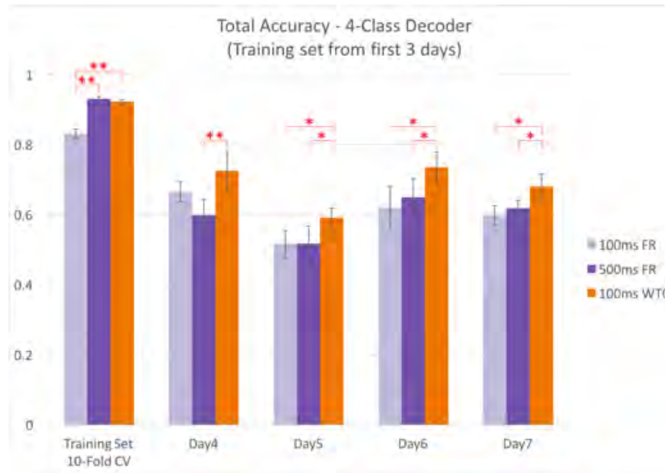

Fig. 2. Comparison of 4-class decoding accuracy between the use of conventional firing rates (100ms and 500ms windows) and WTC-based firing rates (100ms bins). Significant difference in mean accuracy across sessions in a given day is denoted with (*) for $p < 0.05$ and (**) for $p < 0.01$.

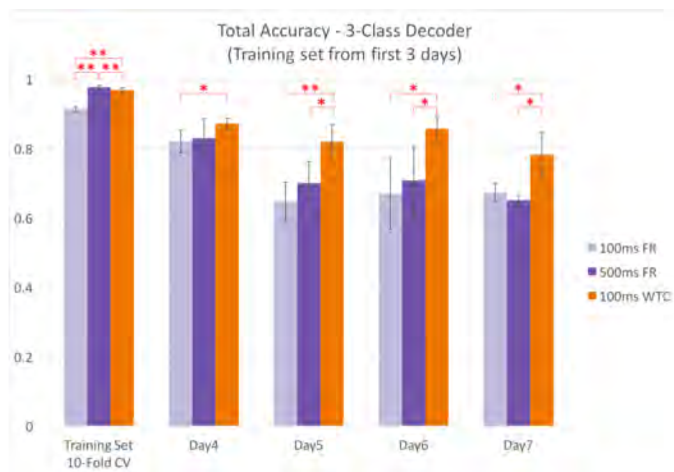

Fig. 3. Comparison of 3-class decoding accuracy, by removing 'stop' observations, between the use of conventional firing rates and WTC-based firing rates. Significant difference in mean accuracy across sessions in a given day is denoted with (*) for $p < 0.05$ and (**) for $p < 0.01$.

Therefore, we re-investigated the decoder performance for decoding the 3 directions (leaving out observations corresponding to the 'stop' command). Figure 3 shows the comparison of the 3-class decoding accuracy. The improvement when WTCs were used was significant, wherein the untrained days show an average of approximately 80% accuracy that is sustained across the four experimental days. The performance is a significant improvement over using conventional firing rates, particularly in the later experimental days. However, the omission of any class other than 'stop' did not result in similar improvements in decoding accuracy.

We noted that LDA with WTC-based firing rates is capable of classifying the three different directions ('left', 'right' and 'forward'), but performs worse when the fourth command ('stop') is included. Hence, we believe the use of WTC-based spike detection and LDA could serve well as a 3-class decoder for active movement commands, but the decoding of ‘stop’ may require a different set of features or strategy, such as the inclusion of state classifier in addition to a directional decoder [21], presenting an opportunity for further work.

This demonstration has also shown that we can use WTCs for decoding without the need for spike sorting methods or reconstruction of signal for each level. The potential benefit of using WTC, as demonstrated in this offline analysis, is in a more prolonged performance that lasts for up to four weeks. This improvement could be due to a more precise spike detection, particularly on some detail levels, which may be more robust against noise and artifacts. However, further characterization of how WTCs help with spike detection is required to understand the underlying mechanism for the decoder’s improvement. We also recognize the need for more data across a longer period to assess the extent of robustness, as well as the need for more subjects to demonstrate the generalizability of our algorithm.

One of the key concerns using WTC-based spike detection is the processing time of the signals. As there are four times as many features to process and different filter functions used, preliminary measurements of execution time shows that feature extraction may take about nine times as long (approximately 1.1s to process 1s worth of a session’s data with WTCs, versus about 0.12s to process 1s data with conventional methods). Nevertheless, we have not made attempts to optimize the code for speed and efficiency, especially for operations like spike detection and filtering. Furthermore, in the offline tests, we processed data as a whole, rather than treating them as realtime streams. Other strategies, both algorithmic and hardware solutions, can potentially be adopted to ensure fast performance. Some design changes to the algorithm is also needed to ensure filter causality and realtime processing for implementation in online BMI.

All in all, this work demonstrated significant improvements to performance through the simple introduction of WTC-based spike detection. The inclusion of wavelet transform can easily replace the filtering step or be used in conjunction with current processing workflow. We believe that this could be one of key strategies for processing of raw data that will ensure long-term high performing decoder for BMIs, while subsequent strategies such as semi-supervised learning and other adaptive algorithm can be adopted to further improve decoding performance.

ACKNOWLEDGMENT

The authors would like to thank Clement Lim for helping with animal training and data collection.

REFERENCES

[1] M. D. Serruya, N. G. Hatsopoulos, L. Paninski, M. R. Fellows, and J. P. Donoghue, Brain-machine interface: Instant neural control of a movement signal, Nature, vol. 416, no. 6877, pp. 141142, Mar. 2002.

[2] S.-P. Kim, J. D. Simeral, L. R. Hochberg, J. P. Donoghue, and M. J. Black, Neural control of computer cursor velocity by decoding motor cortical spiking activity in humans with tetraplegia, J. Neural Eng., vol. 5, no. 4, pp. 455476, Dec. 2008.

[3] M. Velliste, S. Perel, M. C. Spalding, A. S. Whitford, and A. B. Schwartz, Cortical control of a prosthetic arm for self-feeding, Nature, vol. 453, no. 7198, pp. 10981101, Jun. 2008.

[4] L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. van der Smagt, and J. P. Donoghue, Reach and grasp by people with tetraplegia using a neurally controlled robotic arm, Nature, vol. 485, no. 7398, pp. 372375, May 2012.

[5] S. Rajangam, P.-H. Tseng, A. Yin, G. Lehew, D. Schwarz, M. A. Lebedev, and M. A. L. Nicolelis, Wireless Cortical Brain-Machine Interface for Whole-Body Navigation in Primates, Sci. Rep., vol. 6, p. 22170, Mar. 2016.

[6] C. Libedinsky, R. So, Z. Xu, T. K. Kyar, D. Ho, C. Lim, L. Chan, Y. Chua, L. Yao, J. H. Cheong, J. H. Lee, K. V. Vishal, Y. Guo, Z. N. Chen, L. K. Lim, P. Li, L. Liu, X. Zou, K. K. Ang, Y. Gao, W. H. Ng, B. S. Han, K. Chng, C. Guan, M. Je, and S.-C. Yen, Independent Mobility Achieved through a Wireless Brain-Machine Interface, PLoS One, vol. 11, no. 11, p. e0165773, Nov. 2016.

[7] M. B. Khalid, N. I. Rao, I. Rizwan-i-Haque, S. Munir, and F. Tahir, Towards a Brain Computer Interface using wavelet transform with averaged and time segmented adapted wavelets, in 2009 2nd International Conference on Computer, Control and Communication, 2009, pp. 14.

[8] D. M. Taylor, S. I. Tillery, and A. B. Schwartz, Direct Cortical Control of 3D Neuroprosthetic Devices, Science (80-.)., vol. 296, no. 5574, pp. 18291832, Jun. 2002.

[9] C. W. Lu, P. G. Patil, and C. A. Chestek, Current Challenges to the Clinical Translation of Brain Machine Interface Technology, 2012, pp. 137160.

[10] A. M. Green and J. F. Kalaska, Learning to move machines with the mind, Trends Neurosci., vol. 34, no. 2, pp. 6175, Feb. 2011.

[11] R. D. Flint, Z. A. Wright, M. R. Scheid, and M. W. Slutzky, Long term, stable brain machine interface performance using local field potentials and multiunit spikes, J. Neural Eng., vol. 10, no. 5, p. 56005, Oct. 2013.

[12] Y. Yang, A. Kamboh, and J. M. Andrew, Adaptive threshold spike detection using stationary wavelet transform for neural recording implants, in 2010 Biomedical Circuits and Systems Conference (BioCAS), 2010, pp. 912.

[13] R. J. Brychta, S. Tuntrakool, M. Appalsamy, N. R. Keller, D. Robertson, R. G. Shiavi, and A. Diedrich, Wavelet Methods for Spike Detection in Mouse Renal Sympathetic Nerve Activity, IEEE Trans. Biomed. Eng., vol. 54, no. 1, pp. 8293, Jan. 2007.

[14] R. Q. Quiroga, Z. Nadasdy, and Y. Ben-Shaul, Unsupervised Spike Detection and Sorting with Wavelets and Superparamagnetic Clustering, Neural Comput., vol. 16, no. 8, pp. 16611687, Aug. 2004.

[15] E. S. G. Carotti, V. Shalchyan, W. Jensen, and D. Farina, Denoising and compression of intracortical signals with a modified MDL criterion, Med. Biol. Eng. Comput., vol. 52, no. 5, pp. 429438, May 2014.

[16] O. Aydemir and T. Kayikcioglu, Wavelet Transform Based Classification of Invasive Brain Computer Interface Data, Radioengineering, vol. 20, no. 1, pp. 3138, 2011.

[17] N. Robinson, A. P. Vinod, C. Guan, K. K. Ang, and T. K. Peng, A Wavelet-CSP method to classify hand movement directions in EEG based BCI system, in 2011 8th International Conference on Information, Communications and Signal Processing, 2011, pp. 15.

[18] Z. Pra, P. L. Sndergaard, N. Holighaus, C. Wiesmeyr, and P. Balazs, The Large Time-Frequency Analysis Toolbox 2.0, 2014, pp. 419442.

[19] P. L. Sndergaard, B. Torrsani, and P. Balazs, The Linear Time Frequency Analysis Toolbox, Int. J. Wavelets, Multiresolution Inf. Process., vol. 10, no. 4, p. 1250032, Jul. 2012.

[20] Zhiming Xu, R. Q. So, K. K. Toe, Kai Keng Ang, and Cuntai Guan, On the asynchronously continuous control of mobile robot movement by motor cortical spiking activity, in 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2014, pp. 30493052.

[21] M. Velliste, S. D. Kennedy, A. B. Schwartz, A. S. Whitford, J.-W. Sohn, and A. J. C. McMorland, Motor Cortical Correlates of Arm Resting in the Context of a Reaching Task and Implications for Prosthetic Control, J. Neurosci., vol. 34, no. 17, pp. 60116022, Apr. 2014.

[22] B. P. Christie, D. M. Tat, Z. T. Irwin, V. Gilja, P. Nuyujukian, J. D. Foster, S. I. Ryu, K. V Shenoy, D. E. Thompson, and C. A. Chestek, Comparison of spike sorting and thresholding of voltage waveforms for intracortical brainmachine interface performance, J. Neural Eng., vol. 12, no. 1, p. 16009, Feb. 2015.