Adaptive Threshold Spike Detection using Stationary Wavelet Transform for Neural Recording Implants

Yuning Yang, Awais Kamboh and Andrew J. Mason
Department of Electrical and Computer Engineering, Michigan State University, East Lansing, Michigan, USA
{yangyuni, kambohaw, mason}@msu.edu

***Abstract*—Spike detection is an essential first step in the analysis of neural recording signals. A new spike detection hardware architecture combining absolute threshold method and stationary wavelet transform (SWT) is described. The method enables spike detection with 90% accuracy even when the signal-to-noise is -1dB. A noise monitoring block was implemented to automatically calculate the appropriate threshold value for spike detection, and the system then chooses either absolute threshold method or the SWT method to optimize power consumption. The system was designed in 130nm CMOS and shown to occupy 0.082 mm^2 and dissipate 0.45 μW for one channel.**

I. INTRODUCTION

With the advent of high density microelectrode arrays, the activity of thousands of neurons can be recorded for neuroscience research and treatment of neural disorders. Wireless transmission of the dense data generated by neural probe arrays eliminates the complications of wiring through the skull but requires high communication bandwidth and power consumption. To overcome these limitations, it is beneficial to implement a data compression system within the implant. Such a system must operate with very low power and occupy minimal area in order to facilitate implantation in living tissue. The advantage of spike detection is that it allows neuroprosthetic implants to transmit only a sparse collection of action potential waveforms instead of entire raw neural recording. This reduces the transmitting data rate and increases the number of recording channels within a limited bandwidth. Spike detection also permits the data rate to be further reduced by spike sorting after detection.

The primary challenge in detecting spikes is interference due to background noise, which is high impedance noise generated from microelectrode arrays because of their small electrode area. The main purpose of spike detection algorithms is to increase the intensity of spike signals relative to background noise. Several detection algorithms have been recently demonstrated [1], however, many of them are computationally excessive relative to the requirements of implanted circuits. A simple absolute threshold method [2] requires little circuitry, but its effectiveness is highly sensitive to the background noise. For example, when the noise level is comparable to neural signals, many noise peaks will be falsely detected as spikes. Alternatively, a nonlinear energy operator (NEO) has been introduced [3,4] that considers spikes as a sudden increase in both amplitude and frequency. It measures the instantaneous energy of spikes and provides a better performance for spike detection than simple thresholding. However, under low signal-to-noise ratio (SNR) or in the presence of high frequency noise, NEO becomes more sensitive to high noise peaks than to action potentials [5]. The stationary wavelet transform (SWT) has also been applied to spike detection [7], but the way in which SWT was applied requires complex hardware that is not suitable for an implantable implementation.

This paper presents a new SWT-based adaptive threshold (SAT) spike detection method that performs especially well at low SNR. A system approach capable of dynamically optimizing power consumption with signal characteristics is described, combining the SWT algorithm with noise level detection and adjustable thresholding. Finally, a hardware and power efficient VLSI architecture of the SWT-based spike detection system is presented.

II. STATIONARY WAVELET TRANSFORM ALGORITHM

The discrete wavelet transform (DWT) is a useful signal processing tool usually applied to denoise and compress signals in biomedical field [6]. Recently, DWT was utilized for spike detection [7] and for feature extraction in spike sorting [8].Generally, DWT is similar to the discrete Fourier transform used to analyze signals in frequency domain, but it decomposes signals into different frequency components at different scales. The DWT decomposition process can be described as

$$a_{j+1}(k) = \sum_n h\ (n-2k)a_j(k) \tag{1}$$

$$d_{j+1}(k) = \sum_n g_0(n-2k)a_j(k) \tag{2}$$

The DWT employs a half-band low-pass filter h_0 generating coarse *approximation* results a_j and a half-band high-pass filter g_0 generating *detail* results d_j. The filters outputs are decimated by two at each level of decomposition without any loss of information if the filters form a perfect reconstruction pair. DWT coefficients have large values for signals matching a chosen *wavelet basis* and small values for noise, facilitating the extraction of neural spikes from background noise. However, DWT suffers a major limitation in terms of shift variance because of decimation, and thus its performance is poor when the firing rate of spikes is high. In contrast, the SWT does not decimate signals after filtering at each level and provides information for more accurate analysis of signal properties. Fig. 1 illustrates this issue. An example spike with a copy shifted in time is shown in Fig. 1(a). Fig. 1(b) and (c) show the detail coefficients of DWT and SWT at the first level, respectively, for this spike pair. Notice that with DWT the coefficient values for the shifted copy are significantly different from those of the original spike while SWT correctly represents both spikes equally. However, SWT does necessitate more extensive computations, which requires careful

consideration at the circuit level to ensure it is suitable for implanted application.

The accuracy of spike detection using wavelet coefficients is closely associated with the wavelet basis. The 'symmlet4' wavelet basis has been reported to provide an optimal representation of neural signals [9], optimally enhancing spike signals and suppressing background noise. Because the symlet4 is both orthogonal and biorthogonal, it has high vanishing moments and is very robust in noisy environments. Using the *symlet4* basis, a study was conducted to analyze the spike signal strength within the each approximation and detail result at different decomposition levels. Defining SNR as [13]

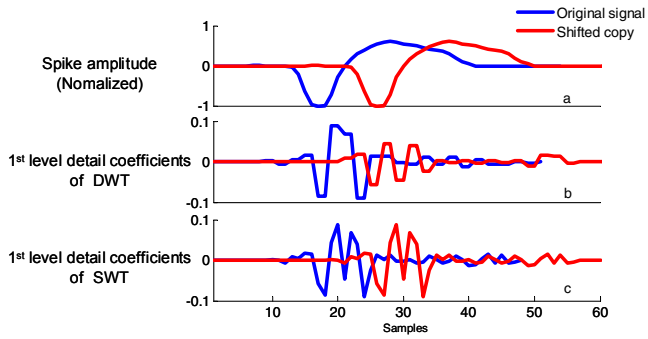

Fig. 1. (a) Original spike signal and shifted copy. (b) DWT coefficients of the spike and its shifted copy (c) SWT coefficients of the spike and its shifted copy.

$$SNR = 20\log\frac{\sigma_{spikes}}{\sigma_{noise}} \tag{3}$$

a simulated neural signal of 1 million data samples with 1,000 spikes at SNR= 0dB was analyzed. The results in Table I list the SNR of each coefficient relative to the SNR of the original signal for both SWT and DWT. Clearly, the SWT coefficients represent spikes stronger than the DWT coefficients. Of particular interest is the fact that the SWT 2nd level detail result consistently shows a strong spike signal relative to background noise. To further illustrate this concept, Fig. 2 plots a window of the original neural signal along with the 1st level approximation, 1st level detail and 2nd level detail SWT coefficients. Arrows mark the locations of spikes hidden within the noise in the original data. Both of the 1st level signals fail to amplify the spike signals, but the spikes are clearly visible in the 2nd level detail coefficients, consistent with the results in Table I. This feature is a function of the SWT characteristics acting on data that can be considered to contain band-limited Gaussian noise because of the analog band pass filtering that typically occurs before digitization in neural recording systems.

III. System Architecture

Results presented at the end of this paper will show that SWT can improve accuracy of spike detection, especially for signals with low SNR. However, for signals with high SNR, a simple absolute threshold (AT) method [2] provides adequate performance with significantly less power than computing SWT coefficients. Because power consumption is such a critical issue in neural implants, a spike detection system that combines the benefits of SWT with the power savings of AT has been explored. To minimize average power consumption, the methodology for the SWT-based system involves

TABLE I: COMPARISON OF SWT AND DWT COEFFICIENT SNRS FOR SEVERAL DECOMPOSITION LEVELS.

	Coefficient SNR relative to SNR of original signal (dB)					
	1st level		2nd level		3rd level	
	Approx.	Detail	Approx.	Detail	Approx.	Detail
SWT	0.2dB	2.2dB	0.3dB	6dB	0.2dB	0dB
DWT	-0.4dB	2.8dB	-1.79dB	3.8dB	-2.9dB	-3.2dB

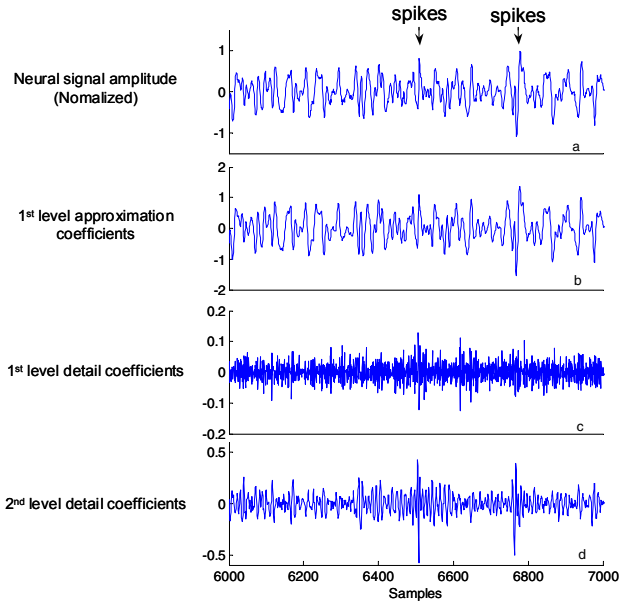

Fig. 2. (a) Noisy signal (SNR=0dB). (b) 1st level approximation coefficient values (c) 1st level detail coefficient values (d) 2nd level detail coefficient values. Arrows mark the locations of spikes within this data window.

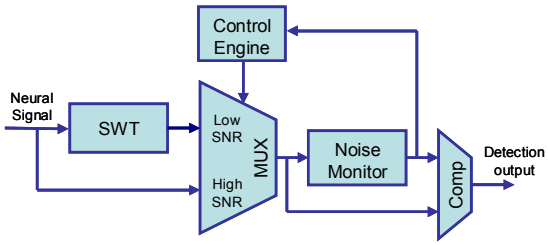

Fig. 3. Block diagram of SWT-based adaptive threshold (SAT) spike detection system.

monitoring the SNR of the input neural signal, using low-power AT when SNR is low, and enabling SWT only when SNR is below a user-set value. The system in Fig. 3 was designed to realize this approach and extend spike detection capabilities to signals with very low SNR. The SWT-based spike detector incorporates circuits to compute 2-level SWT coefficients, estimate and monitor input signal noise level and establish spike thresholds, and control the operation of the overall system to optimize power based on user-controlled parameters. A comparator determines when signals, either raw inputs or SWT coefficients, cross the spike threshold.

A. *Stationary Wavelet Transform Block*

SWT is a redundant algorithm because it does not decimate coefficients at each level. Direct implementation of SWT requires 16 multiplications and 16 additions, which is prohibitively hardware intensive for implantable applications. However, an area-power efficient VLSI structure developed

for lifting based DWT [11] can be adapted to SWT. Fig. 4 shows the data flow of a lifting-based SWT using the *symlet4* basis [10]. This data flow can be implemented using a single computation core with two adders and two multipliers. Calculations of the lifting-based transform can be accomplished by sequentially reusing the computation core as described in [11]. For 2 levels of decomposition, the SWT requires a second computation core, resulting in a total of only four adders and four multipliers for the entire 2-level SWT.

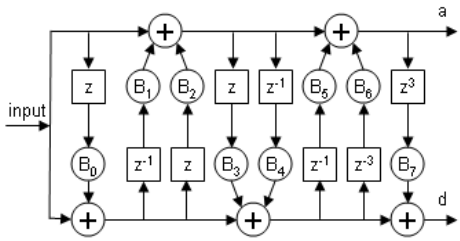

Fig. 4. Data flow of lifting-based 1-level SWT.

B. Noise Monitor Block

An accurate estimate of the input SNR is required permit the SWT-based spike detection system to optimize the tradeoff between detection accuracy and power consumption. Based on empirical studies of neural signals, it was determined that signal noise level provides a sufficient estimate of SNR. A noise monitor circuit has been designed to calculate the standard deviation of signals as a measure of noise level and, thus, an estimate of SNR. The noise monitor block is also responsible for setting the spike detection threshold that is passed to the comparator where spikes are identified. This threshold is scaled to the appropriate signal levels for either raw inputs or SWT coefficients, based on the active detection method chosen by the system. This design assumes the probability of detecting noise peaks as spikes is very low when the threshold value is set to four times the signal standard deviation.

A block diagram of the noise monitor circuit is shown in Fig. 5. The circuit estimates the standard deviation value, Std, of the input signal using a feedback loop. The loop first compares the input signal to the calculated Std value and generates a digital bit stream. A counter records the number of '1' values in the bit steam for 256 cycles. Because the probability that Gaussian noise exceeds its standard deviation is 0.159 [2], the counter value should be 256*0.159=42 when the Std value is correct. The low pass filter in the feedback loop adjusts the Std value until the counter reaches a constant value near 42. Under stable conditions, the Std value will vary slightly up and down to maintain a counter value close to 42.

The poles, zeros and gain of the digital low-pass filter have a critical effect on the stability and convergence time of the noise monitor feedback loop. The transfer function of the comparator, counter and subtractor can be considered as

$$F(z) = \frac{1}{1 - z^{-1}} \tag{4}$$

The transfer function of the digital low-pass filter can be expressed as

$$G(z) = K\frac{1 - \omega_z z^{-1}}{1 - \omega_p z^{-1}} \tag{5}$$

Thus, the loop is a second order system. To provide stability, the zero, pole and gain are set to be 0.5, 1, and 0.5, respectively. The selected parameters for gain, zero and pole do not require multiplication, and the filter thus requires little power.

Fig. 6 displays the behavior of the noise monitor block under different SNR levels as a function of input samples. The block can converge to the proper Std value quickly when the SNR is high and requires a little longer as the SNR decreases. Std values varies slightly once is has converged because of the finite size of the counter window.

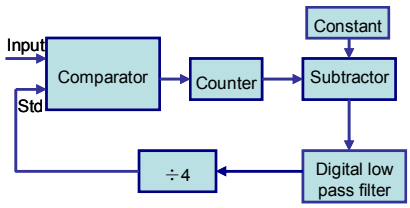

Fig. 5. Functional diagram of the noise monitor feedback block.

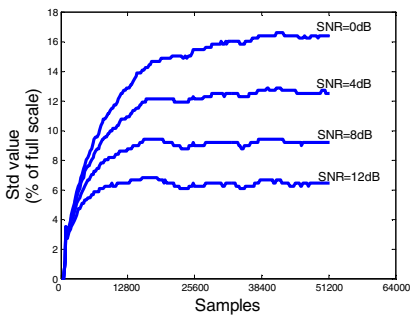

Fig. 6. Noise monitor Std value at different SNR levels.

C. Control Engine

The control engine selects the spike detection method, either AT or SWT, based on the Std value received from the noise monitor block as an estimate of the input signal SNR. Fig. 7 shows the system operation flow managed by the control engine. Initially, the control engine configures the noise monitor to calculate the standard deviation of the input signal, Std1. This value is then compared to a user-defined parameter, C1, that defines the switching point between AT and SWT detection methods. If Std1 is less than C1, the SNR is high enough to employ the AT method. and spikes are detected by comparing the input signal to a scaled value of Std1. Alternatively, if Std1 is greater than C1, the SNR is low and the SWT block is activated. In this case, noise monitor will calculate the standard deviation of the SWT coefficients, Std2. Std2 is then compared user-defined C2 that defines the point at which the system should switch back to AT method. Because SWT inherently improves SNR, the value of C2 will be different than C1 to maintain a stable switching point between detection methods.

IV. Results

A. Implementation Results

The architecture described in Section III was realized using Verilog hardware description language. The modules were synthesized and mapped to a 130nm CMOS standard cell library. Power and area were estimated from the synthesis

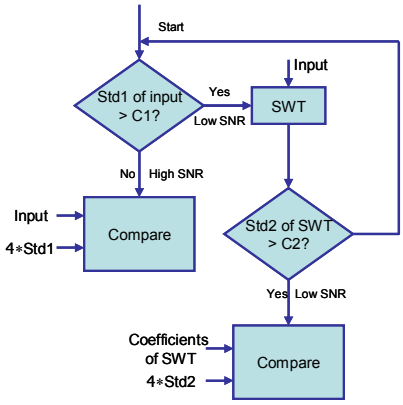

Fig. 7. The system operation flow controlled by the control engine.

Table II Power and area values for f_s=25 KHz

	SWT	Noise Monitor	Control Engine
Power (nW)	402	47	0.8
Area (μm^2)	77524	4095	119

CAD tools. For one channel at an operating frequency 25kHz, Table II show the total power consumption is only around 450nW and the circuit occupies a square of roughly 300µm on a side.

B. *Simulation Results*

Test signals were generated using a neural signal simulator with ten spike pattern templates which were obtained from [12]. Ten neural channels, nine of which are shown in Fig. 8, were generated with three spikes in each channel. The amplitude was normalized for each spike template with a length of 40 samples. Three detection methods were tested for comparison: absolute threshold (AT), non-linear energy operator (NEO) [3] and SWT-based adaptive threshold (SAT). As a criterion to measure performance, the detection accuracy is defined as

$$\text{Accuracy} = \frac{\#\text{ of total detection}}{\#\text{ of true spikes} + \#\text{ of false positives}} \qquad (5)$$

Fig. 9 shows that the SAT method outperforms both AT and NEO for SNR less than 4dB. It performs successfully (<10% error) even when the SNR less than 0dB. The simple AT method achieves reliable detection when the SNR is above 6dB. This data validates the design methodology of the SWT-based spike detection system described in this paper.

V. Conclusion

A new method for spike detection based on SWT was presented and shows good performance in detecting neural spikes even at low SNR. By combining SWT and absolute threshold based spike detection, a highly accurate and power efficient system can be realized for neural implants.

VI. Acknowlegement

Dr. K. Oweiss provided the neural spike templates.

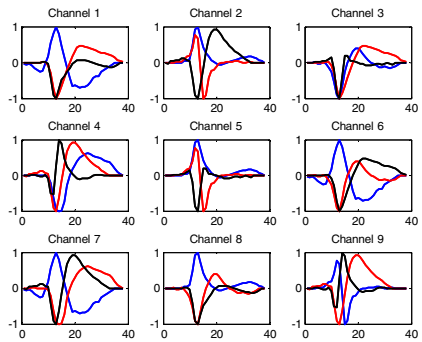

Fig. 8. Spike templates in different simulated channels

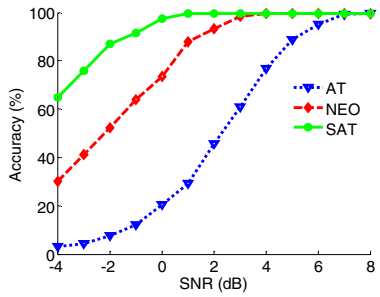

Fig. 9 Detection accuracy of SAT compared to AT and NEO

References

[1] I. Obeid and P. D. Wolf, "Evaluation of spike-detection algorithms for a brain-machine interface application," *IEEE Trans. Biomed.. Eng.,* vol. 51, pp. 905-911, 2004.

[2] R. R. Harrison, "A low-power integrated circuit for adaptive detection of action potentials in noisy signals," in *IEEE Eng. in Medicine and Biology Conf.*, pp. 3325-3328, 2003.

[3] K. Kim and S. Kim, "Neural spike sorting under nearly 0-dB signal-to-noise ratio using nonlinear energy operator and artificial neural-network classifier," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 1406–1411, Oct. 2000.

[4] B. Gosselin and M. Sawan, "Adaptive detecion of action potentials using ultra low-power CMOS circuits," *IEEE Conf. Biomedical Circuits and Systems*, 2008, pp 209-212

[5] C. Joon Hwan, J. Hae Kyung, and K. Taejeong, "A new action potential detector using the MTEO and its effects on spike sorting systems at low signal-to-noise ratios," *IEEE Trans. Biomed. Eng.*, vol. 53, pp. 738-746, 2006.

[6] M. Unser and A. Aldroubi, "A review of wavelets in biomedical application," *Proc. IEEE*, vol. 84, pp. 626–638, Apr. 1996.

[7] R. J. Brychta et al., "Wavelet methods for spike detection in mouse renal sympathetic nerve activity," *IEEE Trans. Biomed. Eng.*, vol. 54,. no. 1, pp. 82–93, Jan. 2007.

[8] R. Quian Quiroga, Z. Nadasdy, and Y. Ben-Shaul, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," *Neural Comp.*, vol. 16, no. 8, pp. 1661–1687, Aug. 2004.

[9] K. G. Oweiss, "A systems approach for data compression and latency reduction in cortically controlled brain machine interfaces," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 7, pp. 1364–1377, Jul. 2006.

[10] C. S. Lee, C. K. Lee and K. Y. Yoo, "New lifting based for undecimated wavelet transform," *Electronics Letters*, vol. 36, pp. 1894-1895, 2000.

[11] A. M. Kamboh, M. Raetz, K. G. Oweiss, A. Mason, "Area-power efficient VLSI implementation of multichannel DWT for data compression in implantable neuroprosthetics," *IEEE Trans. Biomed. Circuits and Systems*, vol. 1, no. 2, pp. 128-135, Jun. 2007.

[12] M. Aghagolzadeh, K. Oweiss, "Compressed and Distributed Sensing of Neuronal Activity for Real Time Spike Train Decoding," *IEEE Trans. on Neural Systems and Rehabilitation Engineering,* pp. 416-427, APRIL 2009.

[13] P. H. Thakur, "Automated optimal detection and classification of neural action potentials in extra-cellular recordings," *Journal of Neuroscience Methods,* pp 364-376, May 2007.