This document is downloaded from DR-NTU (https://dr.ntu.edu.sg)
Nanyang Technological University, Singapore.

Design of a hybrid neural spike detection algorithm for implantable integrated brain circuits

Zeinolabedin, Seyed Mohammad Ali; Do, Anh Tuan; Yeo, Kiat Seng; Kim, Tony Tae-Hyoung

2015

Zeinolabedin, S. M. A., Do, A. T., Yeo, K. S., & Kim, T. T.-H. (2015). Design of a hybrid neural spike detection algorithm for implantable integrated brain circuits. 2015 IEEE International Symposium on Circuits and Systems (ISCAS), 794-797.

https://hdl.handle.net/10356/82912

https://doi.org/10.1109/ISCAS.2015.7168753

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: [http://dx.doi.org/10.1109/ISCAS.2015.7168753].

Downloaded on 13 Mar 2024 17:50:42 SGT

Design of a Hybrid Neural Spike Detection Algorithm for Implantable Integrated Brain Circuits

Seyed Mohammad Ali Zeinolabedin[1], Anh Tuan Do[1], Kiat Seng Yeo[1,2] and Tony Tae-Hyoung Kim[1]
[1]VIRTUS, School of EEE, Nanyang Technological University, Singapore 639798
[2]Engineering Product Development, Singapore University of Technology and Design, Singapore 138682
Email: seyedmoh001@e.ntu.edu.sg

***Abstract*— Real time spike detection is the first critical step to develop spike-sorting for integrated brain circuits interface applications. Nonlinear Energy Operator (NEO) and absolute thresholding have been widely used as the spike detection algorithms where NEO has a better performance measured by the probability of detection and false alarm. This paper proposes a hybrid spike detection algorithm incorporating both spike detection algorithms to reduce the power and to keep the detection rate the same as that of NEO. In the proposed algorithm, the absolute thresholding is performed first to detect a potential spike. Once a potential spike is detected, NEO is executed to check whether the detected spike by absolute thresholding is valid. Since NEO is conditionally conducted, this reduces the overall power consumption. The simulation shows that the proposed hybrid method improves the power consumption by 54.48% compared to NEO in 65 nm CMOS technology.**

Keywords—spike sorting, integrated brain circuits interface, CMOS, subthreshold.

I. INTRODUCTION

Multi-electrode intracranial recording technology offers exceptionally high spatial and temporal signal resolutions needed for neural prosthetic development and neuroscience research [1], [2]. In most existing designs, analog signals from different neurons surrounding electrodes are collected and amplified by low-noise amplifiers. After digitizing the analog signals, they are transmitted to a nearby computer for subsequent software processing such as spike sorting and neural encoding. However, this approach consumes a huge power and requires a large transmitter bandwidth when the number of electrodes is a few hundreds. For example, the data rate for 100 channels, 25kS/s recording system using 8 bit ADCs is 20 Mb/s. Furthermore, the neural prosthetic applications demonstrate that a software-based processing must be done online so that brain commands can be performed without significant time delay. A hardware realization of these processing units using ASIC have been proved to be faster and more power-efficient than that in the software domain. In addition, integrating whole or a part of these algorithms on-chip leads to a significant reduction of the data to be transmitted to the subsequent and more sophisticated software encoding blocks [1, 3]. The first step of neural prosthetic applications is called spike sorting, where each neural spike is detected and assigned to a specific neuron. The final output of the spike sorting process is a neuron ID which represents a physical neuron closest to the electrode. Fig. 1 illustrates a typical recording system where a multichannel analog front end receives brain signals directly from a 2D electrode array. After digitizing, the brain signals are sent to an on-chip, off-chip or software-based spike sorting processor. The spike sorting consists of several stages such as spike detection, alignment, feature extraction and clustering [4] as shown in Fig. 1.

This work proposes a hybrid spike detection algorithm that bases on both conventional absolute thresholding and NEO and implements it in 65 nm CMOS technology. Since the operating frequency of the detector is extremely low, a leakage-aware VLSI realization of the proposed algorithm is highly required to improve the power efficiency.

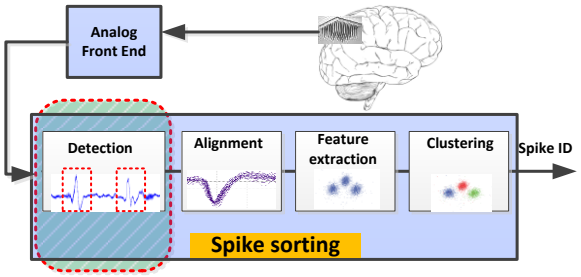

Figure 1. A typical brain signal recording system consists of an analog front-end and spike sorting back-end.

II. HYBRID SPIKE DETECTION ALGORITHMS

Spike detection is an essential step in the spike sorting flow that is used to separate the spikes from the neural signal [5]. All spike detection methods involve two stages of pre-emphasizing and thresholding. There are various pre-emphasis methods such as absolute thresholding, nonlinear energy operator (NEO) [6], and stationary wavelet transform (SWTP) [7], etc. After executing the pre-emphasis part, a corresponding threshold value is applied to decide whether a spike is present or not. If a spike is detected, a 3-ms window of the input signal is used to capture the whole spike [5]. From the hardware realization point of view, there are trade-offs between the detection accuracy, the power consumption and the hardware complexity from the other side [8]. Both absolute thresholding and NEO schemes are popular due to

the efficient structures and good detection accuracy [7], [9]. In this work, we adopt a hybrid spike detection algorithm combining the absolute thresholding and NEO methods [10].

A. Absolute thresholding

This is the simplest way of detecting the spike in which the absolute value of the input data (raw data) is considered as a spike if it is larger than the predefined threshold value [11] given below:

$$Thr = 4\sigma_N \quad \sigma_N = median\{\frac{|x(n)|}{0.6745}\} \tag{1}$$

where $x(n)$ is the input data at time n and σ_N is the standard deviation of noise over N points. This method just looks at the amplitude of the data, whereas a spike is basically described as a high frequency and abrupt energy change. So in order to address this problem, NEO was proposed [8], [9] to extract the high frequency and high energy data.

B. Nonlinear energy operator (NEO)

NEO is defined as (2). If the input data processed by NEO is not a spike, $\psi[x(n)]$ is a very small value and close to zero while if a spike is met, $\psi[x(n)]$ will be very large because a spike is high in power ($x(n)$ is large value) and high in energy (i.e. $x^2(n) - x(n-1) \times x(n+1)$ is large). Similarly, the threshold value is estimated during the training phase over N points as expressed in (3).

$$\psi[x(n)] = x^2(n) - x(n-1) \times x(n+1) \tag{2}$$

$$Thr\_NEO = 4\frac{1}{N}\sum_{1}^{N}\psi[x(n)] \tag{3}$$

C. Proposed hybrid spike detection algorithm

NEO operation requires two multipliers and one subtractor for each individual data while the absolute thresholding method just uses an absolute operator. Therefore the computational complexity of NEO-based detection (and thus its power consumption) is undoubtedly much more than absolute threshold-based detector [9]. As a result, the detection accuracy of NEO is much higher when compared to that of the absolute method. To achieve both good detection accuracy and low power consumption, we proposed a hybrid spike detection algorithm in which NEO complements the absolute method. Its block diagram is shown in Fig. 2. First the absolute threshold detector is utilized to filter out the noisy data before sending them to NEO. So NEO is disabled most of the time and is activated only when a potential spike is detected by the absolute threshold. NEO double-checks the data to make the final decision whether it is a high energy and high power input or not. Therefore, the hybrid spike detection algorithm theoretically has a similar accuracy as NEO and consumes less power than NEO [10].

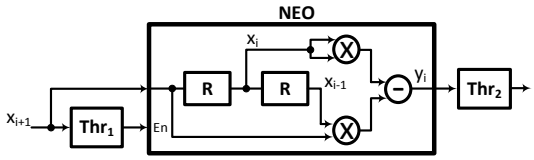

Figure 2. Hybrid spike detection algorithm.

III. Hardware Architecture for the Proposed Hybrid Spike Detection Algorithm

The hardware architecture of the proposed hybrid spike detection algorithm is shown in Fig. 3. The input data is first stored in the registers R1 and R2 to provide the last three data (i.e. x_{n+1}, x_n and x_{n-1}) for NEO operation. If the input data is larger than Thr_1 (absolute thresholding method), the control circuit will activate the NEO engine located in the bold rectangle. Then last three data are processed by NEO and its output will go high if the result is larger than Thr_2, the NEO threshold. In either of "Thr_1" and "Thr_2" blocks, the absolute value of input are calculated and then compared with the threshold value to generate the active (high) output.

To realize the NEO engine, all the input data are connected through the multiplexers to avoid any switching activity happening inside as long as the input data don't exceed Thr_1. Once the input is larger than Thr_1, the last three data are converted to unsigned numbers by "Signed to Unsigned Blocks" and then two multiplications are implemented by adder-based structure utilizing "Extension Blocks" and "Adders Banks" to further reduce the total power. Note that if the NEO engine is fully pipelined, the power consumption is increased due to the pipeline overhead circuitry and it violates the primary goal of reducing the power consumption by applying the hybrid spike detection algorithm. Moreover, our simulation showed that the internal power of synthesized multiplier used in NEO engine is larger than that of being implemented by adder-based structure. As a result, the multipliers are replaced by an extension block and an adder bank, as shown in Fig. 3.

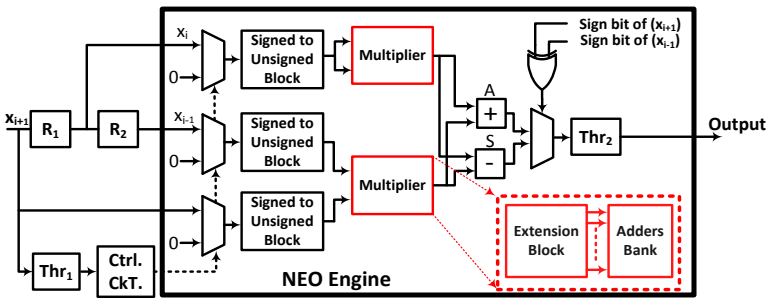

Figure 3. Hardware architecture of the proposed hybrid spike detection algorith.

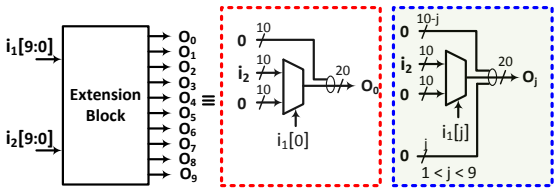

Figure 4. The Extension block introduced in Fig. 3.

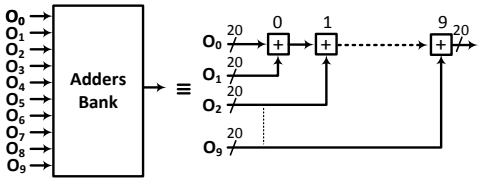

Figure 5. The adders Bank introduced in Fig. 3.

The extension block is depicted in Fig. 4. It receives two inputs as the multiplier operands and provides 10 extended outputs with the same width (20 bits) as partial products. For example, the first output (O_0) is defined as 20-bit zero if $i_1[0]$ is zero or else it would be a 20-bit number, including 10-bit zero as the most significant bits and the 10 remaining bits are replaced by i_2 (multiplicand). This is realized using MUX, as shown in Fig. 4. In the general case of $i_1[j]$ (j^{th} bit of multiplier), O_j is 20-bit zero if $i_1[j]$ is equal to zero, otherwise it consists of (10-j)-bit zero as the most significant bits, 10-bit i_2 and j-bit zero as the least significant bits. The j-bit zero inserted as the least significant bits of O_j implies the shifted partial product generated by the multiplier. Finally, multiplication is performed by 10 sequentially connected adders as shown in Fig. 5, where each one has 20-bit input and output width provided by the extension blocks.

The control circuit enables the NEO engine inputs and output once the input is larger than Thr_1 and keeps it enabled for five clock cycles because a spike may not happen exactly at the point which is singled out by the absolute thresholding and then NEO engine will be disabled. Therefore, compared to NEO method, the hybrid spike detection algorithm just activates NEO for five clock cycles, whereas it is always on in NEO method. Furthermore, it is only activated if a potential spike is detected by the simple threshold block. The final MUX at the output selects one between two results calculated by the subtractor (indicated by S) and adder (indicated by A) based upon the sign bits of input operands of x_{i+1} and x_{i-1}, simply implemented by XORing the sign bits. In other words, if both x_{i+1} and x_{i-1} are negative or positive numbers, the second term of (2) is positive and should be subtracted from the first term. If either of x_{i+1} and x_{i-1} is negative and another one is positive, the second term of (2) is then negative and it should be added to the first term. It has to be done because all the operands at the input of NEO engine are converted to unsigned numbers, whereas as given in (2) the output of the NEO may be negative.

IV. Simulation Results

In order to verify the proposed method and estimate the power consumption, we have implemented three structures. They are NEO, the proposed hybrid spike detection algorithm based on the architecture given in Fig. 3 and the proposed hybrid spike detection algorithm using two multipliers, all in 65 nm CMOS technology. The intended operating frequency is selected as 25 KHz which is the usual sampling rate of the ADCs in neural signal recording. All input data available from [11] is quantized to 10 bits. Functionally verified netlist generated by the Synopsys Design Compiler (DC) and accurate switching activity using the neural spike data from [11] are used to evaluate the power consumption of the designs by the Synopsys Primetime tool.

As tabulated in Table I, the proposed hybrid spike detection algorithm saves power by 54.48% when compared to NEO. In addition, it improves power by 26.87% when compared to the same hybrid spike detection algorithm implemented with two multipliers. Power breakdown of these implementation are detailed in Fig. 6. It can be seen that all three designs have similar leakage power (~4 nW), which indicates that they have similar hardware complexity. The dynamic power (switching and internal) of the proposed detection algorithms is much smaller than that of NEO due to the seldom activation of the NEO engine and its multiplierless architecture requiring less switching activity during the calculation of the NEO value.

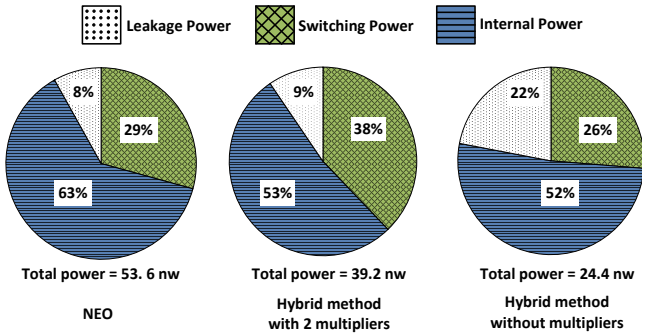

Figure 6. The power breakdown of the proposed hybrid spike detector and NEO.

TABLE I. Power Consumption Comparison

Spike Detection Algorithm	Total power for (nW)
NEO	53.6
Hybrid spike detection algorithm with 2 multipliers	39.2
Hybrid spike detection algorithm (Fig. 3)	24.4

Since the operating frequency of the detector is only 25 KHz, its leakage contributes a significant amount to the total power consumption if this component is not adequately managed. Fig. 7 shows the power breakdown of the proposed architecture realized by three different standard cell libraries (LPHVT: low power high-V_{th}, GPHVT: general power high-V_{th}, and LPLVT: low power low-V_{th}). It is apparent that all three options have similar dynamic power components but their leakages dramatically change from one option to another. Specifically, the LPHVT-based design has leakage 1367× and 202× smaller than that of GPHVT- and LPLVT-based implementation, respectively, which leads to more than 50× reduction in the total power consumption. As a result, we choose LPHVT device to maximize the energy efficiency of the design. All reported numbers are simulated at the supply voltage of 1 V and room temperature.

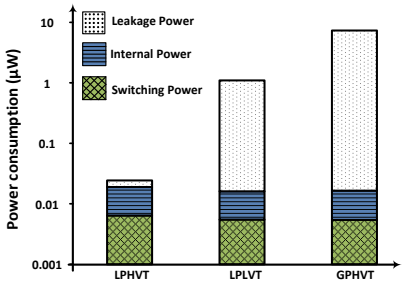

Figure 7. The power breakdown of the proposed hybrid spike detector with three different 65 nm standard cell libraries.

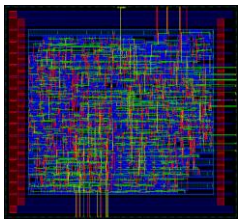

Figure 8. The layout of the proposed hybrid spike detection algorithm.

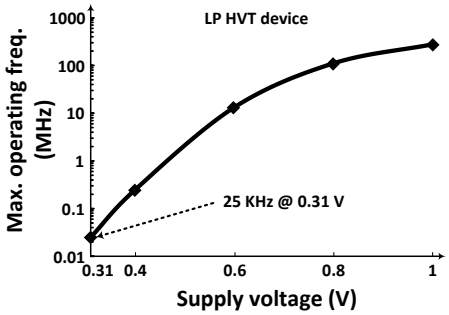

Figure 9. Operating frequecny vs. supply voltage.

A comparison of this work with other state-of-the-art single-detection chips is summarized in Table II. Area and area-power product of the proposed design are much lower than other works. The proposed architecture occupies the layout area of 0.006 mm^2.

Finally, we studied the scalability of the proposed design. Fig. 9 shows the maximum operating frequency of the proposed detector at different supply voltage levels. At 1 V, it can operate up to 330 MHz. For 25 KHz, the minimum required supply is 0.31 V. This indicates that our design can be scaled down to 0.31 V without affecting the speed requirement of the system with 9× power reduction for a single channel detection. This figure also suggests optimum supply voltage for multi-channel detection systems. For example, for a detector to support 128 channels, it must be able to operate at 3.2 MHz. It should then operate at 0.5V for the best energy efficiency.

V. CONCLUSION

We have presented a novel hybrid spike detection algorithm for spike sorting DSPs. It utilizes the absolute threshold algorithm and NEO to have the same detection accuracy as NEO and reduce the power. To achieve this, input data is first analyzed by the absolute thresholding, which conditionally activates the NEO block for double-checking. Since NEO is less frequently enabled compared to the previous NEO design, the power consumption of NEO can be significantly reduced without losing the detection accuracy. Simulation showed that the proposed design improves the power by 54.48% compared to NEO.

TABLE II. COMPARISON WITH PREVIOUS WORKS

	[12]	[13]	[14]	**This work**
Power (nW/channel)	400	752	25000	24.4
Area (mm^2/ch)	0.006	0.033	0.06	0.006
Area × Power (nW× mm^2)	2.4	24.82	1500	0.16
Process (nm)	130	90	90	65
Core Voltage	-	0.55	1.08	1.1

REFERENCES

[1] A. M. Kamboh and A. J. Mason, "Computationally Efficient Neural Feature Extraction for Spike Sorting in Implantable High-Density Recording Systems," Neural Systems and Rehabilitation Engineering, IEEE Transactions on, vol. 21, pp. 1-9, 2013.

[2] C. Moosung, L. Wentai, Z. Yang, C. Tungchien, K. Jungsuk, M. Sivaprakasam, and M. Yuce, "A 128-Channel 6mW Wireless Neural Recording IC with On-the-Fly Spike Sorting and UWB Tansmitter," in IEEE International Solid-State Circuits Conference, Digest of Technical Papers, pp. 146-603, 2008.

[3] L. Hongge, Y. Pan, and X. Tongsheng, "An efficient spike-sorting for implantable neural recording microsystem using hybrid neural network," in Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, pp. 5274-5277, 2012.

[4] C. Tung-Chien, M. Tsung-Chuan, C. Yun-Yu, and C. Liang-Gee, "Low power and high accuracy spike sorting microprocessor with on-line interpolation and re-alignment in 90nm CMOS process," in Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, pp. 4485-4488, 2012.

[5] S. Gibson, J. W. Judy, and D. Markovic, "Spike Sorting: The First Step in Decoding the Brain: The first step in decoding the brain," Signal Processing Magazine, IEEE, vol. 29, pp. 124-143, 2012.

[6] I. Obeid and P. D. Wolf, "Evaluation of spike-detection algorithms forabrain-machine interface application," IEEE Trans. Biomedical Engineering, vol. 51, pp. 905-911, 2004.

[7] S. Gibson, J. W. Judy, and D. Markovic, "Technology-Aware Algorithm Design for Neural Spike Detection, Feature Extraction, and Dimensionality Reduction," Neural Systems and Rehabilitation Engineering, IEEE Transactions on, vol. 18, pp. 469-478, 2010.

[8] S. Mukhopadhyay and G. C. Ray, "A new interpretation of nonlinear energy operator and its efficacy in spike detection," Biomedical Engineering, IEEE Transactions on, vol. 45, pp. 180-187, 1998.

[9] S. Gibson, J. W. Judy, and D. Markovic, "Comparison of spike-sorting algorithms for future hardware implementation," in IEEE Engineering in Medicine and Biology Society (EMBS), pp. 5015-5020, Aug 2008.

[10] A. T. Do, K. S. Yeo, "A Hybrid NEO-based Spike Detection Algorithm for Implantable Brain-IC Interface Applications," in IEEE International Symposium on Circuits and Systems (ISCAS), pp. 2393 - 2396, June 2014.

[11] R. Q. Quiroga, Z. Nadasdy, and Y. Ben-Shaul, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," Neural Computing, vol. 16, pp. 1661-1687, 2004.

[12] J. Drolet, H. Semmaoui, and M. Sawan, "Low-power energy-based CMOS digital detector for neural recording arrays," in Biomedical Circuits and Systems Conference (BioCAS), pp. 13-16, Nov 2011.

[13] Vaibhav Karkare, Sarah Gibson, and Dejan Markovic, "A 130-µw, 64-channel neural spike-sorting dsp chip," IEEE Journal of solid-state circuits,vol. 46, no. 5, pp. 1214–1222, 2011.

[14] T.C. Chen and et al., "128-channel spike sorting processor with a parallel-folding structure in 90nm process," ," in IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1253–1256, May 2009.