HHS Public Access

Author manuscript

Conf Proc IEEE Eng Med Biol Soc. Author manuscript; available in PMC 2015 September 11.

Published in final edited form as:

Conf Proc IEEE Eng Med Biol Soc. 2010 ; 2010: 2939–2942. doi:10.1109/IEMBS.2010.5626247.

VLSI Implementation of a Template Subtraction-Based Algorithm for Real-Time Stimulus Artifact Rejection

Kanokwan Limnusun [Student Member, IEEE],

Electrical Engineering and Computer Science Department, Case Western Reserve University, Cleveland, OH 44106 USA

Hui Lu,

Department of Biology, Case Western Reserve University, Cleveland, OH 44106 USA

Hillel J. Chiel, and

Department of Biology, Case Western Reserve University, Cleveland, OH 44106 USA

Pedram Mohseni [Senior Member, IEEE]

Electrical Engineering and Computer Science Department, Case Western Reserve University, Cleveland, OH 44106 USA. Principal Investigator with the Advanced Platform Technology (APT) Center – A Veterans Affairs (VA) Research Center of Excellence, Cleveland, OH 44106 USA

Hillel J. Chiel: hjc@case.edu; Pedram Mohseni: pedram.mohseni@case.edu

Abstract

This paper presents very-large-scale integrated (VLSI) implementation of an infinite impulse response (IIR) temporal filtering technique for real-time stimulus artifact rejection (SAR) in bidirectional interfacing with the nervous system. The template subtraction-based, IIR-SAR algorithm is implemented on a μ W-level digital signal processing (DSP) unit that initializes its embedded 16b, 4K SRAM with the first recorded stimulus artifact to reduce the operation time in generating an accurate artifact template signal for subtraction. The DSP unit is integrated with spike-recording and microstimulation circuitry to create a functional, standalone, prototype system-on-chip (SoC) fabricated in AMS 0.35 μ m 2P/4M CMOS. The DSP unit functionality has been validated in benchtop tests using a prerecorded neural dataset from a laboratory rat, as well as in neurobiological experiments with isolated buccal ganglia of an *Aplysia californica* (a marine mollusk). The SoC can successfully remove mV-range stimulus artifacts with duration of <114.7ms from the contaminated neural data in real time and recover μ V-range extracellular neural spikes that occur on the tail end of the artifacts. The root-mean-square (rms) value of the pre-processed stimulus artifact is reduced on average by a factor of ~24–30 post-processing, with DSP unit power consumption of <25 μ W from 1.5V.

Index Terms

Closed-loop neuroprostheses; neural recording; neural signal processing; neurostimulation; stimulus artifact rejection; system-on-chip; template subtraction

I. Introduction

Stimulus artifact rejection (SAR) is an integral feature of the next-generation, bidirectional, neural interface microsystems that aim to combine electrical stimulation and neuroelectrical recording in a single device. Large stimulus artifacts often mask the neural activity of interest and either prevent any analysis of the stimulus-evoked recorded data [1], or limit the efficacy of activity-dependent stimulation in closed-loop operation [2], [3].

Blanking techniques in which the recording amplifier input is disconnected during stimulation have traditionally been effective in rejecting large, saturating stimulus artifacts at the expense of no viable recording during stimulation [4]–[6]. Using active feedback circuitry to provide a low-impedance discharge path for the stimulating electrode [7], [8], and careful stimulator design related to the isolation of stimulation channels and parasitic current injection [9] have also been shown to decrease the duration and amplitude of the artifacts. However, these approaches cannot fully eliminate them, often leaving behind considerable residual artifacts.

On the other hand, more complex subtraction techniques typically employ digital signal processing (DSP) algorithms to generate a high-fidelity template signal representative of the stimulus artifacts for subsequent subtraction from the contaminated neural data, resulting in artifact removal [10]–[12]. This is based on the premise that the overall shape, dynamic range and timing of the stimulus artifacts do not significantly vary over consecutive stimulation cycles. Hence, subtraction techniques have to rely on full-scale recording of the stationary stimulus artifacts for accurate template signal generation, but offer the advantage of retaining signal information during stimulation.

Since subtraction techniques require DSP algorithms for template signal generation, they have traditionally been implemented offline on a home-base computer post-data acquisition. To execute a subtraction-based SAR algorithm in real time (i.e., as the recording is taking place), a suitable template-generation technique should be realized in hardware after optimization and validated using real neural data. To that end, Wichmann and Devergnas have recently reported an artifact-removal device based on template subtraction, which employs a commercially available microcontroller as well as analog-to-digital and digital-to-analog converters [13]. The device employs digital averaging of the recurrent, stationary artifact waveforms [12] in order to generate and update a template signal that is subtracted from the contaminated data record on subsequent occurrences of the artifact.

We have previously assessed the feasibility of hardware implementation of a similar subtraction-based SAR algorithm using the well-established infinite impulse response (IIR) temporal filtering technique for template generation [14]. Further, we proposed memory initialization with the first recorded stimulus artifact as an elegant approach to reduce the IIR system response time [15], although our initial IIR system implementation suffered from >25dB of degradation in signal-to-noise ratio (SNR) for template signal generation due to a non-optimal architecture [15]. Hence, we optimized the IIR-SAR algorithm for hardware implementation and successfully realized it on a field-programmable gate array (FPGA) platform [16], [17].

In this paper, we present very-large-scale integrated (VLSI) implementation of the optimized IIR-SAR algorithm using a custom, low-power DSP architecture. The DSP unit is integrated with recording and microstimulating circuitry adapted from [3] to create a standalone, proof-of-concept, prototype system-on-chip (SoC) featuring low-noise neural recording, programmable neurostimulation and real-time stimulus artifact removal for bidirectional interfacing with the nervous system [18]. The DSP unit functionality is validated using a prerecorded neural dataset from a laboratory rat, as well as in neurobiological experiments with isolated buccal ganglia of an *Aplysia californica* (a marine mollusk), an experimentally tractable and well-studied invertebrate model system for stable, repeatable tests. The paper is organized as follows. Section II briefly describes the IIR-SAR algorithm. Section III presents the system architecture of the SoC with real-time SAR capability, and Section IV presents the integrated circuit architecture of the recording, microstimulating and signal-processing blocks of the SoC. Section V presents our measured results from the SoC, and Section VI finally draws some conclusions from this work.

II. IIR Implementation of SAR Algorithm

In this section, we briefly describe the template subtraction-based, IIR-SAR algorithm that was previously developed in MATLAB™ simulations [14] and optimized for hardware implementation [16]. Template signal generation is achieved by averaging a number of the properly shifted versions of the input neural data, which also contain the stimulus artifacts. This can be mathematically expressed as:

$$y(t) = \sum_{n=0}^{N-1} a(n) \cdot x(t - nT_{sti}), \quad (1)$$

where $y(t)$ is the estimated template signal, $x(t)$ is the input neural data, N is the number of stimulus artifact waveforms used for averaging, $a(n)$ are the averaging factors and T_{sti} is the stimulation period. In order for the artifact and its template to have equal amplitudes, the averaging factors should sum up to unity and can be all equal to $1/N$ for common averaging.

Fig. 1 depicts the architecture for IIR implementation of the SAR algorithm, depicting the number of bits for the internal nodes after considering dynamic range and fraction length requirements for optimum performance with a 10b recording front-end [16]. To generate a template signal representative of the stimulus artifacts, temporal filtering is employed in the form of an exponentially weighted moving average (EWMA) of several properly shifted versions of the input neural data containing the stimulus artifacts. This can be represented by:

$$y_n = (1-K) \cdot y_{n-1} + K \cdot x_n, \quad (2)$$

where y_n is the new artifact template signal, y_{n-1} is the previous template signal and x_n is the input neural data. Hence, the template signal is retained in a memory, and a new template signal is generated from the previous template signal and the input neural data according to (2), which is then subtracted from the input neural data. Factor $K < 1$ plays a similar role as

N in (1) and affects the IIR system response time and accuracy. It can be selected as $1/2^i$ (with i being an integer) to alleviate the use of multipliers in VLSI implementation.

III. SoC Architecture

Fig. 2 shows the system architecture of the proof-of-concept, bidirectional neural interface SoC that comprises a spike-recording front-end and programmable microstimulator adapted from [3], as well as an embedded DSP unit for real-time SAR. The SoC is designed to operate from 1.5V, except for the stimulating electrode driver stage in which 5V is used for enhanced voltage compliance.

The recording front-end performs ac amplification, dc baseline stabilization, highpass filtering and 10b digitization of the neural data with digitally programmable gain and bandwidth. The microstimulator delivers trains of charge-balanced monophasic or asymmetric biphasic current pulses followed by passive discharge, with digitally programmable current pulse amplitude and externally controllable current pulse duration, frequency and number within a stimulus train. The DSP unit performs additional highpass filtering digitally to remove any residual dc offset or low-frequency noise, and performs real-time stimulus artifact removal based on template subtraction, as previously described in Section II. The next section presents each SoC building block in further detail.

IV. Integrated Circuit Architecture

A. Recording Front-End

Fig. 3 depicts the circuit schematic of the spike-recording front-end, comprising a low-noise amplifier (LNA) with adjustable high cutoff frequency (HCF), highpass filter (HPF) with adjustable low cutoff frequency (LCF), variable-gain amplifier (VGA) with adjustable offset and 10b successive-approximation register analog-to-digital converter (SAR ADC) [3]. It features an overall bandpass frequency response with eight different gain values in 49–65.6dB at 1kHz, with the LNA providing fixed 32dB of ac gain via capacitive feedback and dc baseline stabilization via a MOS-bipolar resistor in parallel with the feedback capacitor [19]. With the ac gain nominally set to 60dB, the bandwidth can be programmed from <0.1Hz–12.3kHz to 390Hz–6.4kHz by digitally tuning the HPF and LNA bias currents. The SAR ADC is nominally clocked at 1MHz, with each ADC conversion cycle taking 28 clock cycles to relax current drive requirements of the preceding VGA for low-power operation, thus providing a nominal sampling frequency of 35.7kSa/s.

B. Microstimulator

The electrical microstimulator comprises a 6b current-based digital-to-analog converter (DAC), electrode driver, 12b parameter register, signal level shifter and bias circuitry. The electrode driver operates from 5V and integrates a pair of pMOS (anodic) and nMOS (cathodic) current sources with thick-oxide transistors and boosted output impedance (>100M Ω) via negative feedback for constant-current stimulation [3]. The DAC operates from 1.5V and outputs a programmable current in 0–2.2 μ A with 6b resolution, which is then amplified to generate a maximum current of ~100 μ A for the anodic phase. The cathodic-phase current is set to be 1/3 of that in the anodic phase via proper transistor sizing ratios to

generate an asymmetric biphasic current pulse for optimal use of the available voltage headroom (5V).

C. DSP Unit

The major building blocks of the DSP unit include a digital HPF, a digital control unit and a SAR processor, which in turn comprises a computational unit, SRAM (16b, 4K) and parity generator/checker to identify potential memory errors. A reset synchronizer generates an internal signal, *Sync Reset*, to asynchronously reset the DSP unit when the signal is low [20]. A 68b register is used to store user-set parameters such as bandwidth setting of the digital HPF and SAR coefficient, K , as well as memory-initialization, memory-length and output-blanking settings. The requisite memory length (i.e., # of 16b samples) depends on the sampling frequency and time duration over which the SAR algorithm operates, which should be at least equal to the stimulus artifact duration for proper operation. With a nominal sampling frequency of 35.7kSa/s in the recording front-end, the DSP unit can process stimulus artifacts $< \sim 115\text{ms}$ in duration.

Fig. 4 depicts the structure of the digital HPF and SAR processor in the DSP unit. The digitized neural sample (10b) from the ADC is first highpass filtered using a 1st-order filter with direct form II architecture and adjustable LCF. Factor K_1 is set as 1/16 or 1/8, resulting in LCF of 366 or 756Hz, respectively, with a system clock of 1MHz.

An overflow/underflow detector at the HPF output limits its dynamic range to 10b before feeding the SAR processor.

The SAR processor only operates for the duration of each stimulus artifact and uses fixed-point computation for simplicity, with 15b representation at its internal nodes to mitigate quantization noise effects on template signal accuracy [16]. The digitized/filtered sample at the HPF output (10b) is first converted to 15b and then multiplied by factor K_2 (same as K in (2), 1/16 or 1/32 in this work) stored in the 68b parameter register. Next, the memory data containing the previous template signal are read, multiplied by $(1-K_2)$ and added to $(K_2 \cdot x_n)$ to obtain the new template signal (15b), which is written back into the 16b memory for the next cycle after an even-parity bit addition. The new template signal is also converted back to 10b and subsequently subtracted from the 10b digitized/filtered input sample to produce the SAR output signal. Outside the duration of the artifact, the SAR processor is disabled, and the sample at the HPF output is directly routed to the output register. The three registers in Fig. 4 hold the data for processing in each stage with proper timing control. Instead of sharing the same global clock signal and using a local enable signal to synchronize the registers [16], clock gating is used for low power consumption.

Fig. 4 also depicts the DSP unit's timing operation for two consecutive stimulus artifacts assumed to be the very first and second artifacts. Using the *Path Control* signal from the digital control unit, the computational unit of the SAR processor (SAR-CU) initializes the memory with the very first recorded stimulus artifact, which is shown to significantly decrease the response time of the IIR-SAR algorithm in generating an accurate template signal, especially when stimulus artifacts are highly reproducible in consecutive stimulation cycles [15], [16]. Furthermore, using the *Blanking Control* signal from the digital control

unit, the SAR-CU can remove any residual stimulus artifacts in the output after template subtraction, especially at the rising and falling edges of the artifact where it rapidly changes with time.

Fig. 5 depicts the architecture of the digital controller of the DSP unit, comprising a timing generator and several finite-state machine (FSM) controllers to manage the SAR processor operation, memory operation and its initialization, as well as output blanking. The timing generator creates all internal timing signals for the digital control unit, as well as the gated clocks for the digital HPF and the three registers in Fig. 4. The SAR operation controller and its associated counter (C1) respectively generate the internal *SAR Enable* signal and the *Address* signal to access the memory between the start (A_0) and end (A_n) locations. Specifically, the SAR operation controller detects the rising edge of the *Stimulus Timing* signal upon which *SAR Enable* is activated and counter C1 is loaded with A_0 (12b). Next, it commands the counter to up-count from A_0 until the end memory location, A_n (12b), is reached upon which the counter sends the *Done_C1* signal back to the FSM controller to deactivate *SAR Enable*, clear the counter and wait for the next rising edge of the *Stimulus Timing* signal. Hence, the time duration over which the SAR processor operates (i.e., when *SAR Enable* is high) is programmable via the user-set memory-length setting, and should be ideally equal to the stimulus artifact duration (and not much longer) to save processing power. The SAR processor is disabled when *SAR Enable* is low (i.e., outside the stimulus artifact duration).

Using the *SAR Enable* signal, the memory controller generates *EN*, *RD* and *WR* signals to enable memory operation as well as its read and write functions, respectively. The read/write functions are requested at the rising edges of the *CS* signal, which is also generated by the memory controller.

The initialization controller generates the *Path Control* signal, which is used to initialize the memory with the very first recorded stimulus artifact, as stated previously. Specifically, if the user-set *Initialization Select* (1b) is high, *Path Control* is initially activated until the very first falling edge of *SAR Enable* is detected upon which *Path Control* is subsequently deactivated and remains low for the remainder of SAR processor operation (see Fig. 4).

The blanking controller and its two associated counters (C2 and C3) generate the *Blanking Control* signal, which is used to remove any residual stimulus artifacts in the output after template subtraction, as stated previously. Specifically, the blanking controller detects the rising edge of the *Stimulus Timing* signal upon which counter C2 is loaded with the *Rise-Edge Blanking Length (REBL)* parameter (11b). Next, counter C2 activates the *Rise-Edge Blanking* signal and starts to down-count from *REBL* until the counter value reaches zero upon which *Rise-Edge Blanking* is deactivated and counter C2 is cleared. Upon detecting the falling edge of the *Stimulus Timing* signal, the same process is repeated by the blanking controller and counter C3 to generate *Fall-Edge Blanking* using the *Fall-Edge Blanking Length (FEBL)* parameter (11b). By defining *Blanking Control* as the OR function of the *Rise-Edge Blanking* and *Fall-Edge Blanking* signals, the user can independently set the blanking duration from 0 (i.e., no blanking) to 2,047 data points (i.e., ~57ms given the

nominal sampling frequency), synchronized with the rising and falling edges of the *Stimulus Timing* signal (see Fig. 4).

V. Measurement Results

A prototype chip was fabricated in AMS 0.35 μ m 2P/4M CMOS, measuring $\sim 3.1 \times 3.1\text{mm}^2$ including the bonding pads. The DSP unit along with the 16b, 4K SRAM occupied 3.64mm² of silicon area. This section presents the measurement results from electrical benchtop characterization and neurobiological experiments in isolated buccal ganglia of an *Aplysia californica*.

A. Benchtop Characterization

The top plots in Fig. 6 depict the measured frequency response and input noise voltage spectrum of the analog recording front-end for three different bandwidth settings and with the mid-band ac gain nominally set to 60dB. The LCF was programmable from <0.1 to 390Hz, whereas the HCF was adjustable in the range of 6.4 to 12.3kHz. In the right plot, note how the HPF effectively removed the flicker noise contribution, when the LCF was digitally set to 260 and 390Hz (the HPF was bypassed for the case of the recording bandwidth of <0.1Hz–12.3kHz.) With the bandwidth of the analog recording front-end set to 390Hz–6.4kHz, the input noise voltage measured in 0.5Hz–50kHz was 3.42 μ V_{rms}, resulting in noise efficiency factor (NEF) of 2.74 for the LNA.

The bottom left plot depicts the measured microstimulator output current versus its output voltage for four different DAC input codes when sourcing and sinking current in anodic and cathodic phases, respectively. The microstimulator output voltage could reach at least 4.68V (going toward 5V) and 154mV (going toward 0V) when sourcing and sinking current, respectively, from 5V. The output impedance was measured to be >100M Ω (limited by the precision of our measurement setup), sufficiently high for constant-current stimulation.

To evaluate the DSP unit power consumption, two *prerecorded* neural datasets from a rat (sampled at $\sim 24.41\text{kHz}$ and recorded during 4-Hz cortical stimulation) and an *Aplysia californica* (sampled at 2kHz and recorded during 0.5-Hz stimulation) were used. These datasets were collected as part of experiments that were unrelated to this work and provided to us by two different research laboratories.

The system clock frequency was set to ~ 684 and 56kHz for the rat and *Aplysia* datasets, respectively (i.e., 28 $\times$ their sampling frequency). The bottom right plot in Fig. 6 shows the measured DSP unit power consumption versus requisite memory length, when processing each neural dataset for artifact removal. The measured power was in the range of 1.6–2.2 μ W for the *Aplysia* neural dataset and increased to 16.1–21.2 μ W for the rat neural dataset due to the higher system clock frequency. Within each dataset, the measured power also increased with the time duration over which the *SAR Enable* signal was high.

To further characterize the performance of the DSP unit, a 294-s window of the prerecorded rat neural dataset was directly fed to the input of the DSP unit after amplification in software by 520 ($\sim 54.3\text{dB}$). Memory initialization was enabled, and factors K_1 and K_2 in Fig. 4 were

both set to 1/16. The SAR processor was set to operate for 5ms upon receiving an indication of stimulation by the *Stimulus Timing* signal, and no output blanking was applied (i.e., *REBL* and *FEBL* were set to zero.) The left plot in Fig. 7(a) depicts the input neural data to the SAR processor of the DSP unit (i.e., after the digital HPF), consisting of neural spikes buried in large stimulus artifacts. The middle plot shows the generated artifact template signal with memory initialization as previously described. The right plot depicts the SAR processor output in which large stimulus artifacts were rejected and neural data recovered in real time. Fig. 7(b) depicts 5-ms snapshots of the waveforms at $t = \sim 208$ s, showing that the DSP unit was capable of recovering in real time a neural spike that occurred on the tail end of the artifact as close as within 0.5ms after the artifact spike (see the arrow). The measured results from the DSP unit were in excellent agreement with those previously obtained with an FPGA platform using the same rat neural dataset [16]. Table I summarizes the measured SoC performance.

B. Neurobiological Experiments

The SoC functionality was next verified in neurobiological experiments with isolated buccal ganglia of an *Aplysia californica* [21]. Custom-made hook electrodes were used for stimulating on branch A of buccal nerve 2 (BN2-A) and recording from both buccal nerve 2 (BN2) and buccal nerve 3 (BN3). The bandwidth of the analog recording front-end was nominally set to the minimum range. In the DSP unit, factors K_1 and K_2 were both set to 1/16, and memory initialization was enabled. Further, the DSP unit was set to operate for the duration of the stimulus artifact only, which was determined *a priori*, to save processing power. As shown in Fig. 8, an external stimulator was used for these experiments, since the current range of the on-chip microstimulator ($100\mu\text{A}$) was not enough for effective stimulation on BN2-A. A home-base computer was used for generating the *Stimulus Timing* signal, which was applied to both the external stimulator and the SoC, as well as for storing the serial output data from the SoC.

Case #1—In the first case, recording was performed from BN2, with the gain of the analog recording front-end nominally set to the minimum value to determine the largest stimulus artifact dynamic range that could be handled by the SAR processor. Stimulation was performed for 1 minute at 2Hz with a stimulus current pulse of $100\mu\text{A}$ in amplitude and 1ms in pulsewidth. The SAR processor was set to operate for 15ms upon receiving an indication of stimulation by the *Stimulus Timing* signal, and output blanking was applied for 0.5ms in synchrony with the rising and falling edges of the *Stimulus Timing* signal. Fig. 9 depicts the measured results in which plot numbers correspond to the node numbers in SAR-CU of Fig. 4. Specifically, plot #1 shows a 15-ms window of the input data to the SAR-CU, showing a total of 119 highpass-filtered stimulus artifacts superimposed. Plot #2 depicts the 119 artifact templates generated by the DSP unit and superimposed. Plot #3 shows the SAR-CU output signal without blanking, showing significant rejection of the stimulus artifacts after template subtraction, whereas plot #4 shows the SAR-CU output signal after template subtraction and residual blanking. This experiment demonstrated the SAR processor functionality in fully removing artifacts as large as $\sim 5\text{mV}_{\text{pp}}$ from the recorded data in real time, although no neural spikes could be detected due to subthreshold stimulation at $100\mu\text{A}$.

and the small gain of the analog recording front-end. The DSP unit power consumption was measured to be 23.3μW.

Case #2—In the second case, recording was performed from BN3, with the gain of the analog recording front-end nominally set to the maximum value to facilitate neural spike recording. Memory initialization with the first recorded artifact was performed using a single current pulse at 100μA for subthreshold stimulation to avoid the presence of stimulus-evoked neural spikes on the initial condition of the memory. This was then followed by 400-μA, 2-Hz stimulation for 1 minute, with the stimulus current pulsewidth always set to 1ms. The SAR processor was set to operate for 10ms upon receiving an indication of stimulation by the *Stimulus Timing* signal, and output blanking was applied for 2.5ms in synchrony with the rising edge of the *Stimulus Timing* signal.

Since memory initialization was performed based on subthreshold stimulation, the initial condition of the memory normalized to the steady-state artifact template signal was <1. Hence, prior to any further data analysis, we first determined the minimum number of stimulus artifacts required to generate an accurate template signal [15], [16]. The template error was defined as:

$$Error (\%) = \left| \frac{v_{rms,HPF} - v_{rms,template}}{v_{rms,HPF}} \right| \times 100, \quad (3)$$

where $V_{rms,HPF}$ and $V_{rms,template}$ are the root-mean-square (rms) values of the highpass-filtered stimulus artifact and corresponding template signal, respectively, within the 10-ms duration of SAR processor operation. Fig. 10 shows a plot of the template error that was calculated for each stimulus artifact within the 1-minute stimulation duration, except for the very first artifact and those with neural spikes present on their tail ends. As expected, the template error decreased with the number of stimulus artifacts, requiring 20 and 60 artifacts to reach 10% and 1% error, respectively.

Fig. 11 depicts the measured results from case #2 in neurobiological experiments after the template error reached 1%. Specifically, plot #1 shows a 10-ms window of the input data to the SAR-CU, showing a total of 62 highpass-filtered stimulus artifacts superimposed (artifact #60–121), with some (barely visible) neural spikes riding on their tail ends. Plot #2 depicts the 62 stimulus artifact templates generated by the DSP unit and superimposed. Plot #3 shows the SAR-CU output signal without blanking, showing significant rejection of the stimulus artifacts after template subtraction, whereas plot #4 shows the SAR-CU output signal with blanking in which the neural spikes were recovered in real time after template subtraction and residual blanking (see arrows). The DSP unit power consumption was measured to be ~24μW.

In order to assess the SAR performance of the SoC in a quantitative manner, a total of 216 stimulus artifacts (119 of 120 and 97 of 121 artifacts from case #1 and #2, respectively) were analyzed. Specifically, the mean and standard deviation of the rms values of the artifacts were computed pre- and postprocessing by the SAR processor of the DSP unit. The analysis excluded the very first artifact in each neural dataset and those artifacts in case #2

dataset that had neural spikes present anywhere in their 10-ms duration over which the SAR processor was operating. This ensured that the occasional presence of neural spikes did not confound the analysis. The same statistics were also obtained from segments of the SoC output that represented pure noise (i.e., absence of both neural spikes and artifact residuals).

Table II tabulates the results of this analysis. As can be seen, the rms value of the artifact on average was reduced by a factor of ~ 30 and ~ 24 in case #1 and #2 (template error $\sim 1\%$), respectively, resulting in post-processed rms values that were at the level of that for the output noise after template subtraction and residual blanking.

VI. Conclusion

This paper reported on the VLSI implementation of a neural signal-processing algorithm for real-time stimulus artifact rejection (SAR) based on template subtraction. Specifically, a low-power DSP unit generated a high-fidelity template signal representative of the stimulus artifacts via IIR temporal filtering and then subtracted it from the contaminated neural data to remove the artifacts. The DSP unit was integrated with previously developed spike-recording and microstimulation circuitry to create a functional, standalone, proof-of-concept, prototype SoC for bidirectional interfacing with the nervous system. The functionality of the SoC was demonstrated in neurobiological experiments with isolated buccal ganglia of an *Aplysia californica* by removing large stimulus artifacts in real time from the contaminated data record and recovering the extracellular neural activity that occurred on the tail end of the artifacts. The power consumption of the spike-recording front-end and DSP unit were measured to be ~ 27 and $< 25 \mu\text{W}$, respectively, from 1.5V.

One limitation of this approach is that it does not directly address the problem of amplifier saturation and hence is less effective with prolonged amplifier saturation, unless care is taken in the design of the recording and stimulating circuitry to prevent or minimize amplifier saturation by decreasing the duration and amplitude of the artifacts [7]–[9]. On the other hand, the major advantage of this approach as compared to blanking (i.e., disconnecting the recording amplifier input during stimulation) is that it has the potential to retain signal information during stimulation, while fully eliminating large stimulus artifacts from the contaminated data record in real time, as demonstrated in this work.

Acknowledgments

This work was supported by the Department of Defense Traumatic Brain Injury–Investigator-Initiated Research Award Program under Award W81XWH-10-1-0741 (to P. Mohseni) and National Institutes of Health grant NS047073 (to H. J. Chiel).

The authors would like to thank Jeffrey McManus, Case Western Reserve University, Cleveland, OH, for his assistance in conducting the neurobiological experiments, as well as Dr. David Guggenmos and Prof. Randolph Nudo, Kansas University Medical Center, Kansas City, KS, for providing the prerecorded rat neural dataset. The authors would also like to thank Dr. Meysam Azin, Qualcomm, San Diego, CA, and Prof. Marc Buchner, Case Western Reserve University, for helpful discussions that made this work possible.

References

1. McGill KC, Cummins KL, Dorfman LJ, Berlizot BB, Luetkmeyer K, Nishimura DG, Widrow B. On the nature and elimination of stimulus artifact in nerve signals evoked and recorded using surface electrodes. *IEEE Trans Biomed Eng.* Feb; 1982 BME-29(2):129–137. [PubMed: 7056556]
2. Guggenmos DJ, Azin M, Barbay S, Mahnken JD, Dunham C, Mohseni P, Nudo RJ. Restoration of function after brain damage using a neural prosthesis. *Proc Natl Acad Sci USA (PNAS)*. Dec; 2013 110(52):21177–21182. [PubMed: 24324155]
3. Azin M, Guggenmos DJ, Barbay S, Nudo RJ, Mohseni P. A battery-powered activity-dependent intracortical microstimulation IC for brain-machine-brain interface. *IEEE J Solid-State Circuits*. Apr; 2011 46(4):731–745.
4. Shahrokhi F, Abdelhalim K, Serletis D, Carlen PL, Genov R. The 128-channel fully differential digital integrated neural recording and stimulation interface. *IEEE Trans Biomed Circuits and Systems*. Jun; 2010 4(3):149–161.
5. Sharma V, McCreery DB, Han M, Pikov V. Bidirectional telemetry controller for neuroprosthetic devices. *IEEE Trans Neural Syst Rehab Eng.* Feb; 2010 18(1):67–74.
6. Nikolic ZM, Popovic DB, Stein RB, Kenwell Z. Instrumentation for ENG and EMG recordings in FES systems. *IEEE Trans Biomed Eng.* Jul; 1994 41(7):703–706. [PubMed: 7927392]
7. Brown EA, Ross JD, Blum RA, Nam Y, Wheeler BC, DeWeerth SP. Stimulus-artifact elimination in a multielectrode system. *IEEE Trans Biomed Circuits and Systems*. Mar; 2008 2(1):10–21.
8. Blum RA, Ross JD, Brown EA, DeWeerth SP. An integrated system for simultaneous multichannel neuronal stimulation and recording. *IEEE Trans Circuits and Systems – Part I*. Dec; 2007 54(12):2608–2618.
9. Hanson TL, Omarsson B, O'Doherty JE, Peikon ID, Lebedev MA, Nicolelis MAL. High-side digitally current-controlled biphasic bipolar microstimulator. *IEEE Trans Neural Syst Rehab Eng.* May; 2012 20(3):331–340.
10. Wagenaar DA, Potter SM. Real-time multichannel stimulus artifact suppression by local curve fitting. *J Neurosci Meth.* Oct; 2002 120(2):113–120.
11. Hashimoto T, Elder CM, Vitek JL. A template subtraction method for stimulus artifact removal in high-frequency deep brain stimulation. *J Neurosci Meth.* Jan; 2002 113(2):181–186.
12. Wichmann T. A digital averaging method for removal of stimulus artifacts in neurophysiologic experiments. *J Neurosci Meth.* May; 2000 98(1):57–62.
13. Wichmann T, Devergnas A. A novel device to suppress electrical stimulus artifacts in electrophysiological experiments. *J Neurosci Meth.* Sep; 2011 201(1):1–8.
14. Azin, M.; Chiel, HJ.; Mohseni, P. Comparisons of FIR and IIR implementations of a subtraction-based stimulus artifact rejection algorithm. *Proc. 29th Annu. Int. IEEE Eng. Med. Biol. Conf. (EMBC'07)*; Lyon, France. August 23–26, 2007; p. 1437-1440.
15. Limnusun, K.; Lu, H.; Chiel, HJ.; Mohseni, P. VLSI implementation of a template subtraction algorithm for real-time stimulus artifact rejection. *Proc. 32nd Annu. Int. IEEE Eng. Med. Biol. Conf. (EMBC'10)*; Buenos Aires, Argentina. August 31–September 4, 2010; p. 2939-2942.
16. Limnusun K, Lu H, Chiel HJ, Mohseni P. Real-time stimulus artifact rejection via template subtraction. *IEEE Trans Biomed Circuits and Systems*. Jun; 2014 8(3):391–400.
17. Limnusun, K.; Lu, H.; Chiel, HJ.; Mohseni, P. FPGA implementation of an IIR temporal filtering technique for real-time stimulus artifact rejection. *Proc. IEEE Biomed. Circuits and Systems Conf. (BioCAS'11)*; San Diego, CA. November 10–12, 2011; p. 49-52.
18. Limnusun, K.; Lu, H.; Chiel, HJ.; Mohseni, P. A bidirectional neural interface SoC with an integrated spike recorder, microstimulator, and low-power processor for real-time stimulus artifact rejection. *Proc. IEEE Custom Integr. Circ. Conf. (CICC'14)*; San Jose, CA. September 15–17, 2014; accepted for publication
19. Harrison RR, Charles C. A low-power low-noise CMOS amplifier for neural recording applications. *IEEE J Solid-State Circuits*. Jun; 2003 38(6):958–965.
20. Arora, M. *The Art of Hardware Architecture: Design Methods and Techniques for Digital Circuits*. New York, NY, USA: Springer-Verlag; 2012.

21. McManus JM, Lu H, Chiel HJ. An *in vitro* preparation for eliciting and recording feeding motor programs with physiological movements in *Aplysia californica*. J Vis Exp. Dec.2012 70

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1.
IIR architecture of the template subtraction-based SAR algorithm.

Fig. 2.
System architecture of the bidirectional neural interface SoC with real-time stimulus artifact rejection (SAR) capability.

Fig. 3.
Circuit schematic of the spike-recording front-end.

Fig. 5.

Architecture of the digital controller of the DSP unit along with its three finite-state machines (FSMs).

Fig. 6.

Top – Measured gain (left) and input noise voltage (right) of the analog recording front-end. Bottom – Measured microstimulator output current vs. output voltage in anodic and cathodic phases (left) and DSP unit power consumption vs. requisite memory length for processing neural datasets from a rat and *Aplysia californica* (right).

Fig. 7.

Measurement results using prerecorded neural data from a laboratory rat. a) Left plot shows a 294-s window of the input data to the SAR processor. Middle plot depicts the generated stimulus artifact template signal, whereas the right plot shows the IIR system output from the SAR processor. A 5-ms snapshot of the waveforms are shown at b) $t \approx 208$ s. The arrow points to a neural spike riding on and then recovered from the tail end of the stimulus artifact.

Fig. 8.
Illustration of the experimental setup for the neurobiological tests.

Fig. 9.
Measured results from case #1 in neurobiological experiments.

Fig. 10.
Error in template generation vs. number of stimulus artifacts.

Fig. 11.

Measured results from case #2 in neurobiological experiments. Plot #4 is shown after additional offline lowpass filtering in dc–2kHz. Note the smaller dynamic range of the Y-axis in the bottom plot after artifact removal and residual blanking.

Table I

Summary of Measured SoC Performance

RECORDING FRONT-END		MICROSTIMULATOR		
Max. BW	<0.1Hz – 12.3kHz (HPF bypassed)		Anodic	Cathodic
Min. BW	390Hz – 6.4kHz	Stimulus Waveform	Asymmetric biphasic & monophasic w/passive discharge	
AC Gain @ 1kHz	49 – 65.6dB			
RMS Input Noise (0.5Hz – 50kHz)	3.01μV (Max. BW) 3.42μV (Min. BW)	Max. I _{OUT}	100μA	34μA
		DAC Linearity	< ±1.9 LSB	
NEF	2.48 (Max. BW) 2.74 (Min. BW)	Voltage Compliance (Max. I _{OUT})	4.68V (of 5V)	4.846V (of 5V)
CMRR @ 1kHz	57.1dB			
PSRR @ 1kHz	63.7dB	DIGITAL SIGNAL PROCESSING UNIT (f _{CLK} = 1MHz)		
INL/DNL	< ±1.2 LSB			
ENOB	9b (f _s = 35.7kSa/s)	HPF LCF	366Hz (K _I = 1/16); 756Hz (K _I = 1/8)	
Power Consumption (Incl. SAR ADC)	31.5μW (Max. BW; f _{CLK} = 1MHz) 26.7 μW (Min. BW; f _{CLK} = 1MHz)	Stimulus Artifact Duration	<114.7ms	
		Blanking Duration	<57.3ms	
		SAR Coefficient (K ₂)	1/16, 1/32	
		Power Consumption	<25μW	
Total Power Consumption w/o Microstimulator (Min. BW; f _{CLK} = 1MHz) = <51.7μW				

Table II

Statistics of pre- and post-processed stimulus artifacts

CASE #1			
Template Error		Mean (μV_{rms})	SD (μV_{rms})
6% (119 of 120 SAs)	Pre-Processing	285.6	4.7
	Post-Processing (w/o Blanking)	13.2	5.9
	Post-Processing (w/Blanking)	9.6	1.6
	Output Noise	7.1	0.2
CASE #2			
10% (97 of 121 SAs)	Pre-Processing	143.1	0.7
	Post-Processing (w/o Blanking)	11.5	3.9
	Post-Processing (w/Blanking)	6.2	0.7
5% (90 of 121 SAs)	Pre-Processing	143.1	0.7
	Post-Processing (w/o Blanking)	10.9	3.2
	Post-Processing (w/Blanking)	6.1	0.6
1% (60 of 121 SAs)	Pre-Processing	142.9	0.4
	Post-Processing (w/o Blanking)	10.6	3.2
	Post-Processing (w/Blanking)	6.0	0.5
	Output Noise	7.2	0.6