

Efficient Implementation and Design of A New Single-Channel Electrooculography-based Human-Machine Interface System

J. F. Wu, A. M. S. Ang, Z. G. Zhang, K. M. Tsui, H. C. Wu, Y. S. Hung, and S. C. Chan

Abstract—This brief introduces an efficient hardware implementation of a new and practical human-machine interface (HMI) system based on single-channel electrooculography (EOG) signals, which are caused by eye movements including blinking and looking up. The EOG is acquired by a commercial wireless recording system, to convey users' intentions, and the simplicity of the system provides a comfortable and practical solution to HMI. Furthermore, to reduce the hardware complexity and power consumption of the signal processing modules of the EOG-based HMI system, a novel multiplier-less implementation is developed, where all the algorithms involved, such as wavelet filtering, feature extraction, and support vector machine, can be realized using limited number of adders and shifters only. Experimental results show that the proposed system offers a simple, practical and yet reliable EOG-based HMI with low complexity and power consumption.

Index Terms—electrooculography, human-machine interface, multiplier-less implementation, support vector machine, wavelet

I. INTRODUCTION

TO re-establish communication between people who are completely or severely paralyzed with the world is a central research problem in neural engineering [1]. Emerging human-machine interface (HMI) technologies have shown great promise to provide a direct communication path between human body and external machines without muscle activity by translating physiological signals (mainly, electroencephalography [EEG] and electromyography [EMG]) into control commands. Due to the very low signal-to-noise ratio (SNR) of intention-related physiological features, the performance of many HMI systems is highly dependent on complex signal processing algorithms which are usually difficult to be implemented efficiently in hardware. Low-power implementation of HMI is also desirable as it could be incorporated in portable or mobile applications, such as mind-controlled wheelchair and unmanned vehicles. Currently, efficient hardware realization of many commonly-used HMI algorithms is still under active research.

Recently, electrooculography (EOG)-based HMI, which can

infer users' intention from eye movements, such as blinking and looking up/down/left/right, has emerged as a promising technology. The advantages of EOG as a control signal for HMI include (1) EOG can be easily observed from most people who can move their eyes and the signal patterns are highly consistent and (2) eye movements are intentional and natural movements that do not introduce much uncomfortable experience to users. Therefore, EOG-HMI has been considered as a practical alternative to some conventional EEG-based HMIs such as motor-imagery EEG-based HMI systems that cannot work for specific cohorts [2] and visual evoked potentials-based HMIs that can easily fatigue and exhaust users [3]. However, most of existing EOG-HMI systems acquire EOG signals from multiple electrodes placed around the eyes and the setup is not convenient and user-friendly [4-6]. More importantly, the complicated signal processing algorithms, such as wavelet filtering and support vector machines (SVM), that are used to translate EOG into commands are usually hardware and power intensive in practice. This seriously impedes the practical applications of EOG-based HMI systems.

To overcome these difficulties in designing practical EOG-HMI systems, a new single-channel EOG-HMI system with efficient hardware implementation is developed in this brief. In the proposed EOG-HMI, EOG signals are recorded from a commercial single-channel EEG/EOG acquisition system, namely MindWave Mobile Headset (NeuroSky, CA, US) and users' intention is encoded by combinations of eye movements. The MindWave Mobile Headset places one single dry electrode on the forehead of the users to record mixed EEG/EOG activities and transmits the signals to computers or hardware systems via Bluetooth. Hence, the EOG-HMI equipped with the MindWave Mobile Headset is easy-to-use and has a high mobility.

Furthermore, this brief developed a hardware efficient system, which works in tandem with the MindWave Mobile Headset in the EOG-HMI system to translate encoded EOG into recognizable digital commands for external devices. The proposed system performs three different functions: pre-processing, feature extraction and classification. While the accuracy and reliability of these functions are always important for a HMI, particular emphasis has been paid to their implementation complexity and hence power consumption to facilitate its efficient implementation on any portable devices. For example, it can be conveniently attached to wheelchairs or

Manuscript submitted for review June 15, 2014. This study was supported by the University of Hong Kong CRCG Seed Fund (201209160022).

J. F. Wu, A. M. S. Ang, Z. G. Zhang, K. M. Tsui, H. C. Wu, Y. S. Hung, and S. C. Chan are with the Department of Electrical and Electronic Engineering, the University of Hong Kong, Pokfulam Road, Hong Kong (email: scchan@eee.hku.hk).

other related external systems so that the digital commands decoded could control the movement of the wheelchairs, etc. More precisely, we aim to replace all the expensive multipliers involved in our algorithms, such as filtering, wavelet denoising, SVM and so on, with hardwired shifters and adders. Such multiplier-less realization largely reduces the implementation complexity and power consumption, which applies to FPGA, ASIC or VLSI hardware implementations, as the hardware cost/power consumption of a general-purpose multiplier is generally much more expensive than that of a few shifters and adders. To demonstrate the usefulness and feasibility of the proposed approach, a system prototype is implemented and tested on Xilinx Zynq XC7Z020 FPGA. Simulation results show that the proposed system offer an attractive mean to realize a simple and reliable EOG-HMI system with low implementation complexity and power consumption.

II. SINGLE-CHANNEL EOG-BASED HMI

A. Encoding Paradigm

The proposed EOG-based HMI system records a single-channel EOG signal with a MindWave Mobile Headset. A user will encode his/her intentions with three different eye movements: double-blinking (DB), single-blinking (SB) and looking up (LU). The typical EOG signals are shown in Fig. 1.

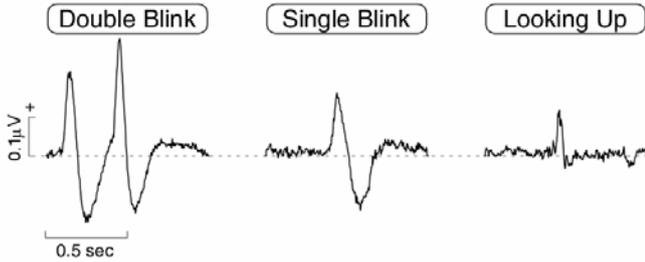


Fig. 1. Three types of EOG signals (one trial) used in the proposed HMI system.

To enhance the versatility of the system, the command inputted by the user is divided into two parts: (i) a DB action is used to indicate the start of a new control signal, and (ii) a combination of one and/or two eye movements in a two-second window after the DB in (i) is used to encode the control commands. In the present study, we will consider four different combinations including (1) <SB+LU>: one SB followed by one LU, (2) <LU+SB>: one LU followed by one SB, (3) <SB>: only one SB, and (4) <LU>: only one LU.

B. Decoding Paradigm

To translate encoded EOG signals into digital commands that can be recognized by external devices, the decoding module of the proposed system has three major functions: (1) pre-processing which includes three steps: bandpass filtering, segmentation, and wavelet denoising, (2) feature extraction, and (3) classification. Fig. 2 shows the block diagram of the whole system. The details of the three decoding functions will be introduced in the next section, along with their hardware implementation.

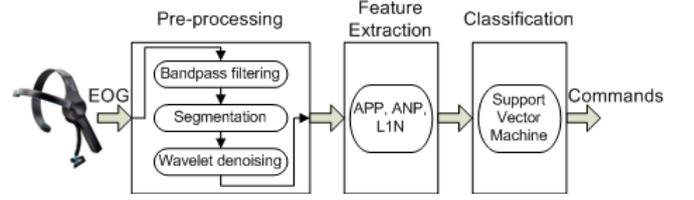


Fig. 2. Schematic diagram of the EOG-HMI system.

C. Experimental Design

In the experimental study, a subject was seated in a comfortable chair inside an unshielded laboratory with reasonable activities to simulate real-life situation. EOG signals were continuously recorded using the MindWave Mobile Headset with a sampling rate of 512 Hz. The experiment consisted of two sessions with the same experimental design: the first session is used for training the classifier and the second session is used for testing the classifier. In each session, the subject was instructed to gaze at the command (one of the four possible commands) randomly presented at the center of a monitor with a viewing distance of 100 cm. Immediately after seeing the command, the subject did a double-blink and then the corresponding eye movements as indicated by the command. Each control command was presented for 15 times and the time interval between two consecutive commands is 5 seconds. The signals recorded in the first session were processed in a computer to train a classifier for recognition of the commands as a later stage. The well-trained classifier and other associated hyper-parameters were then written in the FPGA board to test the online performance of the proposed system in the second session. We now describe the multiplier-less realization of the decoding system.

III. MULTIPLIER-LESS IMPLEMENTATION

In what follows, we shall describe the detail of the three decoding functions and their multiplier-less realization.

A. Pre-processing

A. 1. Bandpass filtering

To remove undesirable noises such as low-frequency drift, high-frequency EEG and muscle activities, EOG signals are firstly bandpass filtered at a bandwidth of 0.5-30Hz. Without loss of generality, the required bandpass filter is realized as:

$$F(z) = F_{LP}(z)F_{HP}(z), \quad (1)$$

where $F_{LP}(z)$ and $F_{HP}(z)$ are fifth order Butterworth lowpass (LP) and highpass (HP) filters, respectively. The cutoff frequencies of the LP and HP filters are at 30Hz and 0.5Hz, respectively. These IIR filters can be expressed as in the following recursive form:

$$F_r(z) = \frac{\sum_{n=0}^4 q_r[n]z^{-n}}{1 - \sum_{n=1}^4 p_r[n]z^{-n}}, \quad (2)$$

where the subscript r denotes either “LP” or “HP”. We note that the filter coefficients $p_r[n]$ and $q_r[n]$ are fixed and pre-designed. As will be shown later, the required multiplications with them can be realized efficiently using sum-of-power-of-two (SOPOT) representation.

A. 2. Segmentation

From continuous EOG recording, an online detection will be performed to find any DB action, which indicates the onset of a control command encoded in the subsequent two seconds. As seen in Fig. 1, a simple peak detector can be used to detect whether the signal contains two positive peaks in a short time period (1 second). Once a DB is detected, the SB or LU will be detected in the subsequent two seconds. Since it is very likely that the duration of both SB and LU activities would not exceed one second, we employ a simple peak detector to differentiate any potential candidate activity from background noise. This process is repeated for a sliding window until a suspected eye activity is detected. Then, we perform feature extraction and employ a SVM for classification if the activity found in the data segment belongs to SB or LU. After that, we continue the peak detection and invoke another SVM to classify the second potential candidate, if any, before the end of the two-second window. We can see that the four-class classification problem can be simplified to two separate binary classification problems. More details on the classifier will be discussed in Section III-C.

A. 3. Wavelet denoising

To improve SNR and classification accuracy, wavelet denoising is performed on data segments. More precisely, it will be hierarchically decomposed into a series of lower resolution components using discrete wavelet transform (DWT). Next, the DWT coefficients that are smaller than an appropriate threshold (which can be well trained offline) are set to zero. Finally, the data segment will be reconstructed from the thresholded DWT coefficients using inverse DWT (IDWT).

DWT can be performed by iterating a two-channel perfect reconstruction (PR) filter bank (FB) on its lowpass output, and the IDWT is an inverse operation that reconstructs the signal based on its synthesis bank. An efficient implementation structure of the class of biorthogonal FIR PR FBs is called lifting structure [7], which allows fast implementation of the FBs or the WTs because it makes use of similarities between the highpass and lowpass filters to speed up the calculation. One famous example of lifting structured FB commonly used in image compression is the 9/7 FB [8]. Fig. 3 shows a general lifting structure, which is parameterized by two scaling constants C_0 and C_1 , and M lifting coefficients s_m , $m = 0, 1, \dots, M-1$. Moreover, the z -transforms of the analysis filter pairs $[H_0(z), H_1(z)]$ and synthesis filter pairs $[F_0(z), F_1(z)]$ in the lifting structure can be written as follows:

$$H_0(z) = C_0 H^{(M-2)}(z) \text{ and } H_1(z) = C_1 H^{(M-1)}(z), \quad (3)$$

$$F_0(z) = H_1(-z) \text{ and } F_1(z) = -H_0(-z). \quad (4)$$

where $H^{(0)}(z) = z^{-1} + s_0 Q(z^2)$, and $s_m Q(z^2) H^{(m-1)}(z)$ with $Q(z) = (1+z^{-1})/2$ and $H^{(-1)}(z) = 1$ for $m = 1, 2, \dots, M-1$. An important advantage of the lifting structured FB is that the scaling constants and lifting coefficients can be arbitrarily quantized without affecting the PR condition.

We can see that the two basic denoising functions, bandpass filtering and WT, involves considerably number of multipliers, which are hardware and power intensive. For efficient

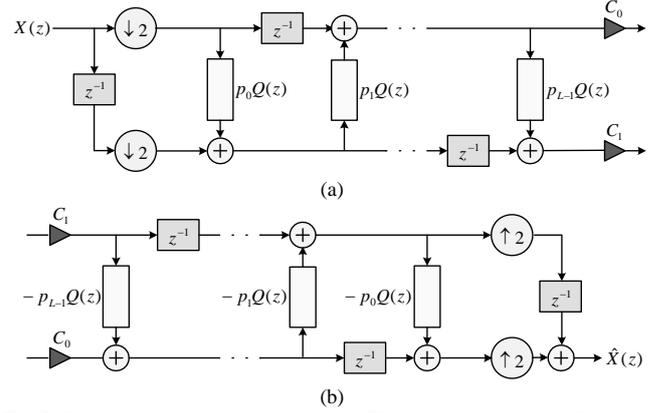


Fig. 3. Lifting structure: (a) Analysis bank (Forward transform). (b) Synthesis bank (Inverse transform).

hardware implementation, we explore that fact that all multipliers involved are constant and hence they can be expressed as SOPOT or canonical signed digit (CSD) [9]. Consequently expensive multipliers can be replaced with adders and shifters only, leading to multiplier-less realization. More precisely, suppose that these multipliers or coefficients are in the following SOPOT representation:

$$\lambda = \sum_{k=0}^{L-1} a_k \cdot 2^{b_k}, \quad (5)$$

where λ denotes the approximation of one of the fixed coefficients in either the bandpass filter (i.e. $p_{LP}[n]$, $q_{LP}[n]$, $p_{HP}[n]$ and $q_{HP}[n]$) or DWT (i.e., s_m , C_0 and C_1), L is the number of terms used in the coefficient approximation, $a_k \in \{-1, 0, 1\}$, $b_k \in \{-l_k, \dots, -1, 0, 1, \dots, u_k\}$, and l_k and u_k determine the dynamic range of each filter coefficient. The larger the numbers L , l_k and u_k , the more accurate the SOPOT approximation will be. In practice, the number of non-zero terms is usually kept to a small number, while satisfying given specifications, so that each multiplication can be implemented as limited number of shifts and additions (or subtractions). As an example, Table I summarizes the SOPOT coefficients of the lifting structured FB, which are determined by the algorithm proposed in [10]. We now present the proposed architecture to realize the multiplication with a SOPOT coefficient, which can be variable. Fig. 4 shows such a structure for multiplying an input with a SOPOT coefficient. The proposed structure, which is simply called the S-Unit, consists of two basic modules namely a scalable shifter and a variable shifter, which provide the required shift for a SOPOT term and an accumulator unit, which accumulates the partial sum of SOPOT terms. The purpose of decomposing the given shift by a series of fixed and larger shifts and a final stage of small variable shift is to avoid using a programmable shift with large shifting range, which may consume large hardware resources. In particular, the fixed shift is up to 4 bits, whereas the variable shifter offers variable shift up to 3 bits (both to the left and right). For each SOPOT term $a_k 2^{b_k}$, the exact operation of the S-Unit is controlled by a 7-bit control register \mathbf{sel}_k , which specifies the sign of a_k in $\mathbf{sel}_k(6)$, the polarity of

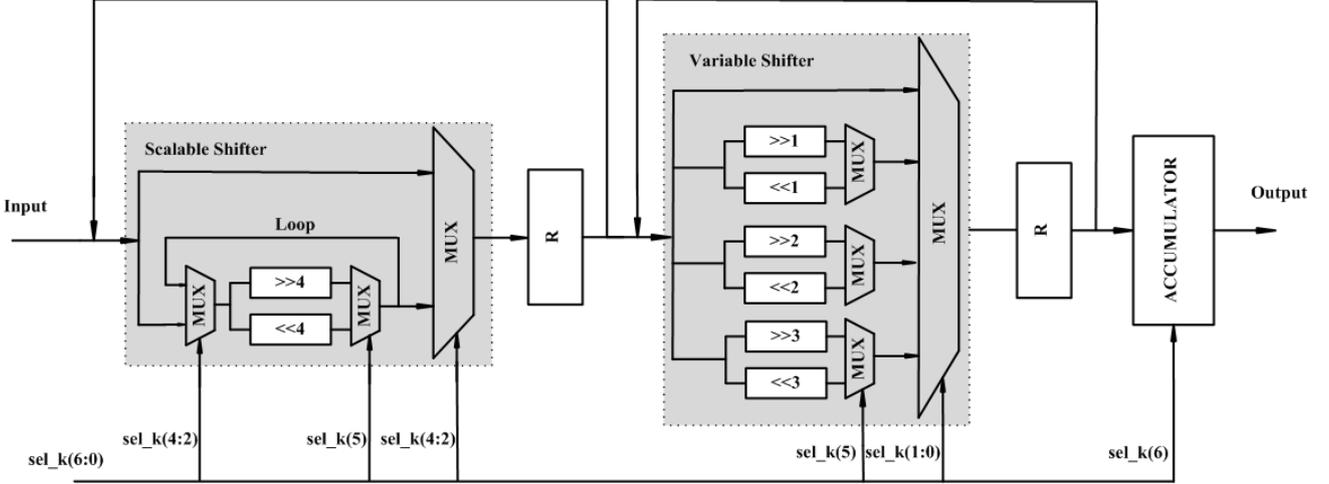


Fig. 4. The proposed structure of an SOPOT unit.

shifting direction, i.e. the sign of b_k , in $\text{sel}_k(5)$, the number of rounds used to realize the smallest factor of 4, i.e. $\lfloor |b_k| / 4 \rfloor$, in $\text{sel}_k(4..2)$, and the remaining amount of shift, i.e. $|b_k| \bmod 4$, in $\text{sel}_k(1..0)$. Hence, the overall coefficient λ in (5) can be computed by re-using one S-Unit L times with the help of an accumulator. Note, starting from input with zero shift, the SOPOT terms will be accumulated either in increasing order of their shifts (left and then right) so that only the incremental shift is required as each stage. The registers R are used to store the result of the previous stage for the next stage. As an example, $p_1 = -2^{-3} + 2^{-6} + 2^{-8} - 2^{-11}$ can be implemented in 4 cycles by shifting the input to the right by 3 bits, accumulate, then to the right by 3 bits again, accumulate, two bits to the right, accumulate, and finally 3 bits to the right and accumulate. Table II shows the configurations of the second lifting coefficient p_1 of the lifting structured FB. Configurations for other SOPOT coefficients can be derived similarly. However, details are omitted due to page limitation.

TABLE I
SOPOT COEFFICIENTS OF THE 9/7 FB

	9/7 filter bank [8]	SOPOT representation
p_0	-3.17226868408289	$-2^1 - 2^0 - 2^{-3} - 2^{-5} - 2^{-6} - 2^{-12} - 2^{-13}$
p_1	-0.10596023714940	$-2^{-3} + 2^{-6} + 2^{-8} - 2^{-11}$
p_2	1.76582215102264	$2^1 - 2^{-2} + 2^{-6} + 2^{-13} + 2^{-14} + 2^{-16}$
p_3	0.88701370409649	$2^0 - 2^{-3} + 2^{-7} + 2^{-8} + 2^{-12} + 2^{-15} + 2^{-16}$
C_0	0.81289306611600	$2^{-1} + 2^{-2} + 2^{-4} + 2^{-12} + 2^{-13} + 2^{-15}$
C_1	0.61508705245638	$2^{-1} + 2^{-3} - 2^{-7} - 2^{-9} - 2^{-13} - 2^{-15}$

TABLE II
CONFIGURATIONS OF THE CONTROL REGISTERS FOR THE SECOND LIFTING COEFFICIENT p_1 OF DWT

SOPOT Term	Control Register	Configuration
-2^{-3}	sel_0	11000011
2^{-6}	sel_1	01000011
2^{-8}	sel_2	01000010
-2^{-11}	sel_3	11000011

B. Feature Extraction

As mentioned earlier, we need to classify four combinations of actions, namely $\langle \text{SB} + \text{LU} \rangle$, $\langle \text{LU} + \text{SB} \rangle$, $\langle \text{SB} \rangle$ and $\langle \text{LU} \rangle$. Therefore, once the signal is filtered, segmented and denoised as described in Section III-A, several discriminative features should be chosen to enhance the classification accuracy of differentiating SB and LU. According to Fig. 1, we find that the following three features: (i) amplitude of positive peak (APP), (ii) amplitude of negative peak (ANP), and (iii) L_1 Norm (LIN) are effective in discriminating SB and LU while their computations do not require too much hardware cost. Due to page limitation, details are omitted here.

C. Classification (SVM)

To classify the extracted features, we first standardize the three features, $x_j^{\text{test}}, j=1,2,3$, as follows:

$$\bar{x}_j^{\text{test}} = (x_j^{\text{test}} - \mu_j) / \sigma_j, \quad (6)$$

where μ_j and σ_j denote the corresponding mean and standard deviation, respectively. Then the following linear classification or decision rule is used to classify the extracted feature and normalized feature $\bar{\mathbf{x}}^{\text{test}}$:

$$\begin{cases} \hat{y} = 1, & \mathbf{w}^T \bar{\mathbf{x}}^{\text{test}} - b \geq 0, \\ \hat{y} = -1, & \text{otherwise,} \end{cases} \quad (7)$$

where \hat{y} denotes the estimated class label for either SB or LU, $\bar{\mathbf{x}}^{\text{test}}$ denotes the standardized feature vector with its j -th given in (6), and \mathbf{w} and b denote the parameter of the SVM under consideration. To simplify the decoder hardware implementation, we assume that μ_j and σ_j in (6) and \mathbf{w} and b in (7) have been pre-computed from the training data of each user. In particular, the latter can be determined by solving the following optimization problem [11]:

$$\begin{cases} \min_{\mathbf{w}, b, \xi_t} & 0.5 \|\mathbf{w}\|_2^2 + C \sum_{t \in \Omega} \xi_t, \\ \text{s.t.} & y_t (\mathbf{w}^T \bar{\mathbf{x}}_t - b) \geq 1 - \xi_t \text{ and } \xi_t \geq 0, \forall t \in \Omega \end{cases} \quad (8)$$

where Ω denotes the training data set. Note the term

$C \sum_{\forall t \in \Omega} \xi_t$ is called the soft margin which is used to deal with possible mislabeled samples, and the associated parameter C can be chosen by cross-validation of the training samples. Since μ_j , σ_j , w and b are fixed, they can again be represented in the form of SOPOT coefficients defined in (5). Hence, the classification step in (7) can be similarly implemented without any multipliers.

IV. EXPERIMENTAL RESULTS

Experimental results on one representative subject showed that all 60 DB actions can be accurately detected while no DB was wrongly detected from ongoing spontaneous data. Overall, the classification accuracy for classifying four classes for this subject is 88.89%. Table III further showed the sensitivity and specificity for each class. The performance of the system was also assessed using the information transfer rate (ITR):

$$ITR = L \cdot [p \log_2(p) + (1-p) \log_2\left(\frac{1-p}{N-1}\right) + \log_2(N)], \quad (9)$$

where L is the number of decisions in one minute, and p is the accuracy of the subject in making decisions among N targets. In the present study, $L=12$, $p=88.89\%$, and $N=4$. Thus, the ITR of the EOG-based HMI system was 15.85 bits/min. In addition, the subject did not express any discomfortability of the EOG-HMI system after the long (>1 hours) experiment.

TABLE III
SENSITIVITY AND SPECIFICITY FOR EACH CLASS OF ONE SUBJECT

	<SB>	<LU>	<SB+LU>	<LU+SB>
Sensitivity	100%	100%	86.67%	80.00%
Specificity	95.56%	93.33%	100%	100%

The proposed architecture is implemented and tested on the Xilinx Zynq XC7Z020 FPGA [12]. All components in the proposed multiplier-less structures are specified using Verilog. The compilation and simulation were done using the ISE Design Suite 14.7 from Xilinx Corporation [13]. The hardware complexity in terms of number of logic slices (LS) used and the normalized power (NP) consumption at 50 MHz clock frequency of the main components [14] are summarized in Table IV, where we also demonstrate the two most expensive components (namely bandpass filter [BPF] and DWT) that require the most LS. In general, we can see that the proposed multiplier-less architecture requires less hardware complexity and achieves lower power consumption than the direct implementation using general-purpose multipliers. It is expected that more savings in hardware complexity and power consumption can be achieved in ASIC or VLSI realization.

TABLE IV
SAVINGS OF HARDWARE COMPLEXITY AND POWER CONSUMPTION FOR THE PROPOSED MULTIPLIER-LESS ARCHITECTURE

	Before SOPOT		After SOPOT	
	LS	NP	LS	NP
BPF	2896	2.01	1229	1.05
DWT	11903	7.21	4792	4.02
Total	17292	10.95	8288	6.38

LS: Logic Slices; NP: Normalized Power (mW)

V. CONCLUSION

The architecture and multiplier-less implementation of a new EOG-based HMI system using SOPOT representation have been presented. Experimental results show that the proposed EOG-HMI system offer satisfactory performance in terms of classification accuracy, implementation complexity and power consumption. This suggests that the proposed system may be an attractive alternative to other existing HMI systems in healthcare applications such as therapy. In addition, the low-power and hardware efficient architecture could benefit many other applications involving filtering, denoising, and classification.

REFERENCES

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, no. 6, pp. 767-791, Jun. 2002.
- [2] C. Vidaurre and B. Blankertz, "Towards a cure for BCI illiteracy," *Brain Topogr.*, vol. 23, no. 2, pp. 194-198, Jun. 2010.
- [3] Y. H. Tu, Y. S. Hung, G. Huang, L. Hu, and Z. G. Zhang, "An automated and fast approach to detect single-trial visual evoked potentials with application to brain-computer interface," *Clin. Neurophysiol.*, in press. DOI: doi: 10.1016/j.clinph.2014.03.028
- [4] Y. Punsawad, Y. Wongsawat, and M. Parnichkun, "Hybrid EEG-EOG brain-computer interface system for practical machine control," in *Proc. IEEE Conf. EMBS*, 2010.
- [5] E. English, A. Hung, E. Kesten, D. Latulipe, and Z. P. Jin, "EyePhone: A mobile EOG-based human-computer interface for assistive healthcare," in *Proc. IEEE Conf. EMBS NER*, 2013.
- [6] S. L. Wu, L. D. Liao, S. W. Lu, W. L. Jiang, S. A. Chen, and C. T. Lin, "Controlling a human-computer interface system with a novel classification method that uses electrooculography signals," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 8, pp. 2133-2141, Aug. 2013.
- [7] W. Sweldens, "The lifting scheme: a new philosophy in biorthogonal wavelet construction," *Proc. SPIE*, vol. 2569, pp. 68 -79 1995.
- [8] A. Cohen, I. Daubechies and J. C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 45, pp. 485 -560, 1992.
- [9] Y. C. Lim and S. R. Parker, "FIR filter design over a discrete power-of-two coefficient space," *IEEE Trans. ASSP-31*, pp. 583-591, April 1983.
- [10] S. C. Chan, K. M. Tsui, K. S. Yeung and T. I. Yuk, "Design and complexity optimization of a new digital IF for software radio receivers with prescribed output accuracy," *IEEE Trans. Circuits Syst. I*, vol. 54, no. 2, pp. 351-366, Feb. 2007.
- [11] C. Cortes, and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273 -297, 1995.
- [12] Zynq XC7Z020, [online] Available: http://www.xilinx.com/support/documentation/data_sheets/ds187-XC7Z010-XC7Z020-Data-Sheet.pdf
- [13] ISE Design Suite, [online] Available: http://www.xilinx.com/support/index.html/content/xilinx/en/supportNav/design_tools/ise_design_suite.html
- [14] P. K. Meher, and, S. Y. Park, "Critical-path analysis and low complexity implementation of the LMS adaptive algorithm" *IEEE Trans. Circuits and Syst. I*, vol. 61, no.3, pp. 778-788, 2014.