A lossless data reduction technique for wireless EEG recorders and its use in selective data filtering for seizure monitoring

Chengliang Dai, Christopher Bailey

***Abstract*— This paper presents a time-domain based lossless data reduction technique called Log2 Sub-band encoding, which is designed for reducing the size of data recorded on a wireless electroencephalogram (EEG) recorder. A data reduction unit can help to save power from the wireless transceiver and from the storage medium since it allows lower data transmission and read/write rates, and then extends the life time of the battery on the device. Our compression ratio(CR) results show that Log2 Sub-band encoding is comparable and even superior to Huffman coding, a well known entropy encoding method, whilst requiring minimal hardware resource, and it can also be used to extract features from EEG to achieve seizure detection during the compression process. The power consumption when compressing the EEG data is presented to evaluate the system′s overall improvement on its power performance, and our results indicate that a noticeable power saving can be achieved with our technique. The possibility of applying this method to other biomedical signals will also be noted.**

I. Introduction

Electroencephalography (EEG) signals are known to have great potential in multiple applications such as Brain-Computer Interface, Prosthetic Devices, Physical Activity Monitors, etc. Since these signals can reflect various biological changes inside human′s body, caused by either diseases or external stimuli, they are widely used to diagnose or to study certain diseases. Researchers have been using EEG for decades in epilepsy[1], Parkinson's disease[2] and other neurological disorder studies. However, the long process of collecting biomedical data sometimes causes inconvenience to the patients or research subjects since it usually restricts their mobility. The wearable wireless signal recorder is an effective solution to this problem, because it frees users from the inpatient environment and promises better real-time healthcare from caregivers to patients with neurological or other diseases.

Nevertheless, there are some challenges when designing a wireless biomedical signal recorder, and power consumption is no doubt one of the critical issues. To give a better user experience, an EEG recorder has to be small, which will inevitably make the space for a battery even smaller. In order to satisfy some long-term recording requirements, the small battery has to last as long as possible, and that is often difficult to achieve.. On a wireless EEG monitor, the transceiver is always one of the biggest contributors to the power consumption, therefore reducing the transmitted data size might be a significant answer to the power issue.

Moreover, reducing the data size can further help to save power where an onboard storage medium, such as flash memory, is used as high read/write rates may be reduced.

On the other hand, a data reduction unit certainly consumes extra power on a device, such that there are some trade-offs we have to consider.

Log2 Sub-band is a data compression algorithm designed to be used on wireless EEG recorders and possibly on other biomedical signal recorders. This algorithm gives satisfying compression results without consuming too much hardware resource, and it also inherently offers some potential on seizure detection. A comprehensive analysis of the compression performance of Log2 Sub-band is given in[3], and its power analysis and seizure detection potential will be discussed in this paper.

II. Log2 sub-band encoding

Biomedical signals are highly non-Gaussian, non-stationary and non-linear, and with strict clinical requirements, many data compression techniques are ruled out despite the outstanding CRs they can give, and lossless techniques are always desirable. Many well-known compression algorithms have been tested on EEG signals, including entropy encoding[4][5], predictive coding[4][6], transform-based encoding[7], etc. Most of these algorithms are able to produce excellent compression results, but some of them are either lossy or incapable of compressing the signal in real-time, and those very sophisticated techniques might be suitable for compressing real-time EEG signals theoretically, except for their high complexity of design at the system level. For these reasons, Log2 Sub-band is proposed to tackle some of the mentioned issues, and provides lossless compression with minimal hardware.

lossless + not too complex

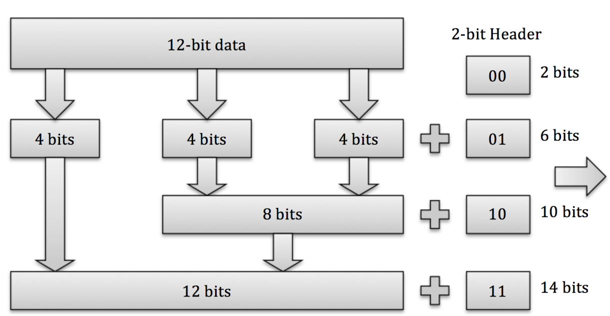

Fig. 1. Log2 Sub-band encoding

On most recorders, biomedical signals are usually digitized into 8-12 bit samples, and Log2 Sub-band is a basic technique that works on time-domain. As shown in Fig. 1, for instance, a channel of EEG signal is converted into 12 bits

978-1-4244-9270-1/15/$31.00 ©2015 Crown

per sample after Analog-to-Digital Conversion(ADC), then in this case every data sample is divided into three bands, and each band is a 4-bit nibble, and nibbles in each current sample will be compared with the same part of previous sample. Instead of transmitting or storing all nibbles of every sample, only the bands that are different from previous sample are sent to the transceiver or the memory. A header is given to every sample to indicate how many nibbles are left after this process, and it could be 3 nibbles, 2 nibbles, 1 nibble or none in the best case, so a 2-bit header is required.

The scheme of Log2 Sub-band resembles a simple prediction encoding technique known as Differential Pulse-Code Modulation (DPCM), but with some advantages. DPCM removes short-term redundancies by taking the difference of adjacent data samples, but its performance deteriorates quickly if more rapid changes occur in the signal, and causes the CR it gets from signals like seizure to be sometimes disappointing. Even with Huffman encoding, the CR can be slightly better, but the algorithm is still very sensitive to the nature of input signals, and to effects of scale and level offsets. Comparing with DPCM and Huffman coding, Log2 Sub-band is more adaptive to these biomedical signals, as will be demonstrated later.

III. Experimental EEG data

The EEG signals used in data reduction simulations are from two sources: (a) research on non-linear deterministic patterns of brain electrical signals at the University of Bonn[8] and (b) research on seizure detection at the Massachusetts Institute of Technology[9].

The data from (a) were originally categorised into different groups, and each group contains one hundred data segments of 23.6 sec duration of EEG signals. The data segments were selected and cut out from continuous multichannel EEG recordings after visual inspection and removal for artifacts such as EOG. Three groups of data from this source are chosen for the simulation, which are from five healthy people(eyes closed), five epilepsy patients during seizure-free time and seizure signal.

The EEG from (b) were collected from 24 different subjects, and signals are mostly divided into many uncategorised one-hour long data segments. Two one-hour segments from one subject that contain both non-seizure signals and a seizure event in each segment are used in the following test.

The EEG from both sources were converted into 12-bit, and retained sampling rates of 173.6 Hz and 256 Hz respectively.

IV. Compression Ratio Results

The two-stage DPCM and Huffman coding technique mentioned earlier is used here to set a benchmark for Log2 Sub-band, and the results are shown in Table I. CR is given as:

$$CR = Size_{Original}/Size_{Compressed} \tag{1}$$

The Log2 Sub-band compresses the data with a three-band '4,4,4' encoder as mentioned in our earlier example, due to the 12-bit width of target signals. The results indicate

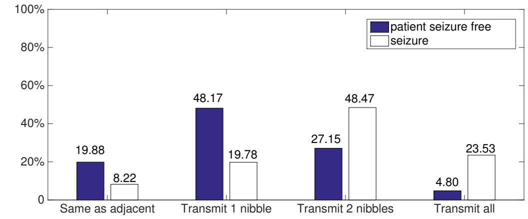

Fig. 2. Seizure-free/seizure Band Distribution[8]

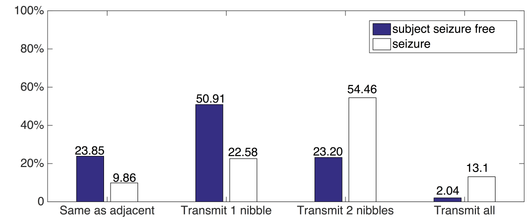

Fig. 3. Seizure-free/seizure Band Distribution of Mixed Signal 1[9]

TABLE I

Compression Ratio results

Data Type	CR of DPCM +Huffman	CR of Log2 Sub-band	
Healthy people with eyes closed[8]	1.82	1.94	O
Patients′ seizure free[8]	2.20	2.58	F
Seizure[8]	1.51	1.66	S
Mixed Signal 1[9]	2.13	2.77	
Mixed Signal 2[9]	1.91	2.32	

that Log2 Sub-band clearly has a superior compression performance, and even though both the techniques give lower CRs whilst compressing seizure signals, unlike Huffman coding′s scheme, which requires a pre-generated codebook based on the frequency of occurrence of every sample, Log2 Sub-band is more adaptive and flexible.

During the compression process, different signals show various behaviours, and signals that can be largely compressed are clearly more likely to be represented with less bands under Log2 Sub-band scheme. Therefore, the distribution of the occurring frequency of each band can be used to describe some of the features of EEG signals. The band distributions of patients′s seizure-free and seizure signal from[8][9] are given in Fig. 2 and Fig. 3 respectively.

According to Fig. 2 and Fig. 3, seizure signals are mostly encoded with two or three bands whilst seizure-free signals are encoded with one or two bands. This will be shown to be useful and will be used to identify seizure in the next section.

V. Seizure detection

Most seizure detection systems involve two basic stages, and the first is to extract the features from the EEG signal, whilst the second is to set a threshold to identify the seizure based on the extracted features. As previous noted, Log2

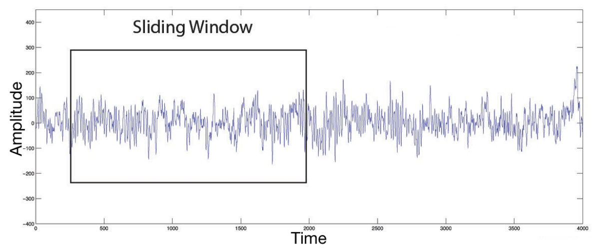

Fig. 4. Healthy Human EEG

Sub-band provides a band distribution whilst compressing each sample, and we suggest that this distribution is one of the signal features that can be applied to seizure detection. To carry out real-time seizure detection, a sliding window scheme(Fig. 4) is proposed.

A sliding window allows Log2 Sub-band encoding to analyse the band distribution of samples within the given window during the compression, and if the distribution within the window matches the seizure criteria, then the current segment can be considered as the seizure candidate. Various methods can be used to set the threshold or criteria with the help of the band distribution, and one straight-forward method will be presented here. A threshold will be set based on the total probabilities of a sample encoded with zero and one nibble, and this method should be rather effective according to Fig. 2 and Fig. 3.

EEG signals from datasets (a) and (b) are used, and data from (a) are pre-categorised, hence 20 seizure data segments were randomly picked and then inserted into patients′ seizure free data to create a simulated signal for seizure detection test. The detection performance of Log2 Sub-band on the simulated data is measured in terms of sensitivity(SEN), false positive rate(FPR) and speed of detection(SoD), which are percentage of successfully detected seizure, percentage of false positive detection, and average time cost to identify the seizure. The window size is set to 100 samples, which is typically around 580 ms. The detection process is repeated with different threshold settings from 30 to 45, and if the percentage of samples encoded with zero and one nibble in a current window falls below the given threshold, samples will be identified as seizure cases. The results when threshold is set to 30, 35, 40 and 45 are given in Table II.

The EEG signals from[9] are uncategorised and therefore are used as the clinical signal input in the experiment. There is only one seizure event in each one-hour′s signal, and the first mixed signal was used for feature extraction(Fig. 3), then based on the acquired feature, the threshold was set for the second one-hour signal to detect the seizure. The performance of Log2 Sub-band derived detection in this case is measured merely with the SoD under difference threshold settings, and the highest SoD achieved is 0.9 sec.

The algorithm we present here is of course somewhat less sophisticated than more advanced techniques such as[10][11]. However those techniques often use computationally expensive methods, with an implication for high power cost when used continuously. We envisage the Log2 band signature technique being useful as a pre-filter, identifying possible seizure candidates with low power and then triggering a more sophisticated analysis on demand, or recording for later analysis. In many epilepsy diagnosis studies, researchers are more interested in seizure signals, and discontinuous recording has been more often applied as a result[12], and our technique is a good option for these studies since a balance between power cost and seizure detection accuracy may be facilitated by this technique with minimal hardware cost.

VI. Power Analysis

The hardware design for this case of three-band '4,4,4' encoder is conducted in 65-nm CMOS technology. We have modelled this and produced trial layouts, and estimated power whilst compressing the EEG data. The results will be given in next section. A silicon test chip is currently being fabricated.

A. Power Saving from Data Compression

The extra power consumed whilst compressing the data[8] is given in Table III. The total power consists of static and dynamic power. Obviously, static power dominates the total power consumption with the typical EEG data sample rates, which in this case is 1 kHz, but dynamic power will increase when more channels of signals are needed.

According to the results, it is quite clear that the circuit consumes more dynamic power when compressing seizure signals, as compare to non-seizure signals. A regression analysis between CR and the dynamic power is given in Fig. 5.

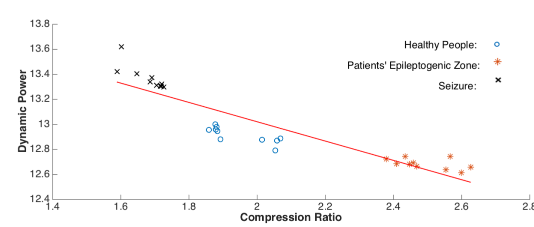

Fig. 5. regression analysis

In order to estimated the possible power saving after data compression, a simplified EEG recorder that only contains amplifier(s), ADC(s), a transceiver and a data compression unit is given[13]. The system power consumption P_{sys} can be roughly modeled as[14]:

$$P_{sys} = P_{amp} + P_{ADC} + P_{comp} + (D * P_{Tx})/CR \quad (2)$$

where P_{Amp} and P_{ADC} are the power consumption of amplifier(s) and ADC(s), and P_{Tx} and P_{Comp} are the power consumption of the transceiver and the data compressor.

Furthermore D is duty cycle rate which is defined such that:

$$D = R_{Orig}/R_{Tx}. \quad (3)$$

where R_{Orig} is the required data rate before compression, and R_{Tx} is the maximum data rate of transceiver.

TABLE II

SEIZURE DETECTION RESULTS ON[8]

Threshold	30	35	40	45
SEN	95%	100%	100%	100%
FPR	0	9.1%	31.0%	37.5%
SoD(sec)	2.9	2.5	1.4	0.6

TABLE III

POWER CONSUMPTION OF LOG2 SUB-BAND

Data Type	CR	Static Power(nW)	Dynamic Power(nW)	Total
Healthy people with eyes closed	1.94	147.23	12.92	160.15
Epileptogenic zone	2.58	147.29	12.68	159.97
Seizure	1.66	147.29	13.37	160.66

TABLE IV

POWER SAVING

Data Type	Power Saving(μW)	Extra Circuit Cost(nW)	Total Saving(μW)
Healthy people with eyes closed	127.91	160.15	127.75
Epileptogenic zone	161.67	159.97	161.51
Seizure	104.96	160.66	104.80

A typical 22mW off-the-shelf Bluetooth transceiver nRF8001 is chosen[15] in this case, and the maximum data rate is 1Mbps, and therefore we assume that to transmit 1 bit of data, it consumes 22nJ′s power.

Combining all the known quantities with (2), we derived the estimated power saving results in Table IV. We see for instance, if we transmit one channel of healthy subject EEG signals, a system with the proposed data compression unit could save $127.91\mu W'$s power from the transceiver with only an extra cost of 160.15 nW. More details of this analysis is given in[14].

B. Power Saving from Seizure Detection

Saving power from conducting seizure detection is based on the assumption that only the seizure signals are of interest, therefore by sending or storing seizure signals only, or applying lossy compression technique on non-seizure signals, a great power saving can be achieved and also a reduction in data storage or transmission. The one-hour signal we used to test seizure detection contains 90 sec seizure signal, and with only 90 sec′s data are transmitted, a total 90% transceiver power could be saved even if that data is uncompressed. If we also use lossless Log2 Sub-band compression on these seizure data segments, then we can project power savings of over 94% (assuming 10% of data is seizure and CR = 1.66). Even though some false positive seizure signals might be encountered, Log2 Sub-band could still help to save a great amount of power.

VII. CONCLUSION

Based on all the results presented, Log2 Sub-band encoding is clearly a promising data reduction candidate for wireless EEG recording devices. However it also has potential on seizure detection. Our on-going research also suggests that our encoder could be applied to other biomedical signals such as electromyogram(EMG) and electrocardiogram(ECG) after making some slight changes on its scheme such as altering the width of each band to attune to new signals. Our research team have recently prototyped the described compression scheme in 65nm test chip which is currently being fabricated for Apr, 2015 delivery.

VIII. ACKNOWLEDGEMENTS

The 65nm test chip is being fabricated as part of a UK Technology Strategy Board founded project in collaboration with Cybula Ltd, grant number 26172-182148.

REFERENCES

[1] G. E. Chatrian, C.-M. Shaw, and H. Leffman, "The significance of periodic lateralized epileptiform discharges in eeg: an electrographic, clinical and pathological study," Electroencephalography and clinical neurophysiology, vol. 17, no. 2, pp. 177–193, 1964.

[2] R. Soikkeli, J. Partanen, H. Soininen, A. Pääkkönen, and P. Riekkinen, "Slowing of eeg in parkinson's disease," Electroencephalography and clinical neurophysiology, vol. 79, no. 3, pp. 159–165, 1991.

[3] C. Dai and C. Bailey, "A time-domain based lossless data compression technique for wireless wearable biometric devices," in SENSORCOMM 2013, The Seventh International Conference on Sensor Technologies and Applications, 2013, pp. 104–107.

[4] G. Antoniol and P. Tonella, "EEG data compression techniques." IEEE transactions on bio-medical engineering, vol. 44, no. 2, pp. 105–14, Feb. 1997.

[5] D. O'Shea and R. McSweeney, "Efficient Implementation of Arithmetic Compression for EEG," IET Proceedings of Irish, 2011.

[6] N. Sriraam, "A high-performance lossless compression scheme for EEG signals using wavelet transform and neural network predictors." International journal of telemedicine and applications, vol. 2012, p. 302581, Jan. 2012.

[7] W. Optimization, L. Brechet, M.-f. Lucas, C. Doncarli, and D. Farina, "Compression of Biomedical Signals With Mother Wavelet Packet Selection," IEEE transactions on bio-medical engineering, vol. 54, no. 12, pp. 2186–2192, 2007.

[8] R. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," Physical Review E, vol. 64, no. 6, p. 061907, Nov. 2001.

[9] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in Proceedings of the 27th International Conference on Machine Learning (ICML-10), 2010, pp. 975–982.

[10] T. Kalayci, O. Ozdamar, and N. Erdol, "The use of wavelet transform as a preprocessor for the neural network detection of EEG spikes," Proceedings of SOUTHEASTCON '94, pp. 1–3, 1994.

[11] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis of EEG records in an epileptic patient using wavelet transform," Journal of Neuroscience Methods, vol. 123, no. 1, pp. 69–87, 2003.

[12] P. Valenti, E. Cazamajou, M. Scarpettini, A. Aizemberg, W. Silva, and S. Kochen, "Automatic detection of interictal spikes using data mining models," Journal of Neuroscience Methods, vol. 150, no. 1, pp. 105–110, 2006.

[13] D. C. Yates and E. Rodriguez-Villegas, "A Key Power Trade-off in Wireless EEG Headset Design," 2007 3rd International IEEE/EMBS Conference on Neural Engineering, pp. 453–456, May 2007.

[14] C. Dai and C. Bailey, "Power analysis of a lossless data compression technique for wireless wearable biometric devices," submitted.

[15] "nRF8001 Bluetooth® low energy Connectivity IC," http://www.nordicsemi.com/eng/Products/Bluetooth-Smart-Bluetooth-low-energy/nRF8001, 2014, [Online; accessed 09-Dec-2014].