Journal of Biomedical Informatics 59 (2016) 1–14

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Resource efficient data compression algorithms for demanding, WSN based biomedical applications

Christos P. Antonopoulos *, Nikolaos S. Voros

Technological Educational Institute of Western Greece, Computer and Informatics Engineering Department, National Road Antiriou-Ioanninon, 30020 Antirio, Greece

ARTICLE INFO

Article history:
Received 21 January 2015
Revised 16 June 2015
Accepted 29 October 2015
Available online 7 November 2015

Keywords:
Biomedical signal processing
Data compression algorithms
Wireless sensor networks
Performance evaluation
ECG/EEG signal compression
Epilepsy monitoring

ABSTRACT

During the last few years, medical research areas of critical importance such as Epilepsy monitoring and study, increasingly utilize wireless sensor network technologies in order to achieve better understanding and significant breakthroughs. However, the limited memory and communication bandwidth offered by WSN platforms comprise a significant shortcoming to such demanding application scenarios. Although, data compression can mitigate such deficiencies there is a lack of objective and comprehensive evaluation of relative approaches and even more on specialized approaches targeting specific demanding applications. The research work presented in this paper focuses on implementing and offering an in-depth experimental study regarding prominent, already existing as well as novel proposed compression algorithms. All algorithms have been implemented in a common Matlab framework. A major contribution of this paper, that differentiates it from similar research efforts, is the employment of real world Electroencephalography (EEG) and Electrocardiography (ECG) datasets comprising the two most demanding Epilepsy modalities. Emphasis is put on WSN applications, thus the respective metrics focus on compression rate and execution latency for the selected datasets. The evaluation results reveal significant performance and behavioral characteristics of the algorithms related to their complexity and the relative negative effect on compression latency as opposed to the increased compression rate. It is noted that the proposed schemes managed to offer considerable advantage especially aiming to achieve the optimum tradeoff between compression rate-latency. Specifically, proposed algorithm managed to combine highly completive level of compression while ensuring minimum latency thus exhibiting real-time capabilities. Additionally, one of the proposed schemes is compared against state-of-the-art general-purpose compression algorithms also exhibiting considerable advantages as far as the compression rate is concerned.

1. Introduction

Epilepsy comprises a disease of profound social significance while it represents one of the most important medical challenges, troubling human kind throughout its history. Although significant research effort has been devoted for many decades, relatively little advancements can be reported in understanding, analyzing, identifying, categorizing and treating it. In this direction [1] offers a review on the results of the Global Campaign against Epilepsy aiming to shed light on the challenges and difficulties people with epilepsy and their families are facing. In addition, in [2] it is clearly depicted that it is also important to look into issues like stigmatization, social exclusion, medical or psychiatric comorbidities. Another aspect related to the economical side effects of epilepsy and respective treatments is discussed in [3–6].

Toward gaining a better understanding, advancements of Wireless Sensor Networks (WSN) are attracting increasing interest both by academia and industry in the area of Epilepsy study. On one hand this is indicated by the various relative projects using WSN advancements to study epilepsy [7] and home health care [8]. On the other hand, increased interest is noticed in lots of research efforts ranging from study of specific WSN medical applications [9] to performance study of WSN networks [10,11] and efficient handling of large volumes of data [12]. In that respect long term, non-intrusive monitoring of patients, during an extended period of time, has been used for many years to extract valuable conclusions and indications. In [13] extensive reports on intensive EEG/video monitoring are presented. Authors [14] EEG monitoring is used to offer a quantitative review of seizure risk in specific considerations, while [15] focuses on epileptic and non-epileptic disorder distinction. However, relative studies in diverse environments through adequate WSN equipment are expected to offer significant insights and advancements. In that context European Research

* Corresponding author.
E-mail address: cantonopoulos@teimes.gr (C.P. Antonopoulos).

http://dx.doi.org/10.1016/j.jbi.2015.10.015
1532-0464/

projects are devoted in Epilepsy study and monitoring using ultra low power wireless platforms [7,16]. Additionally, platforms specifically designed and implemented to wirelessly aggregate respective modalities are attracting intruding attention as indicated in efforts [17,18] where a wireless neural interface is the main objective. Also in respective efforts are presented focusing on EEG data aggregation [19–21]. Such implementations enable medical personnel to perform accurate, secure and non-intrusive monitoring and study of phenomena not possible through conventional approaches.

However, from an engineering point of view respective studies are based on the capability to acquire large volumes of data (digitized physiological measurements), for extended periods of time, which must be either stored locally or transmitted to an aggregation point. The two data types of paramount importance in epileptic seizure study, resulting into excessive amount of accumulated data, are *Electroencephalography* (EEG) and *Electrocardiography* (ECG) measurements. Typical acquisition devices produce samples represented as 16 bit numbers. Furthermore, a wired EEG setup is usually comprised of 64 sensors with sampling frequency up to 2.5 kHz, while ECG requires typically 4 sensors with adequate sampling frequency of a few hundreds of Hertz.

Thus, it can be easily deduced that a setup of 64 EEG sensors requires bandwidth of more than 2.5 Mbps (not considering packet headers and control data), thus posing a significant burden to WSN platforms, which typically offer extremely limited resources. This has been clearly indicated in [22], where the power consumption of critical components of a WSN node has been modeled in order to study its respective effect on the overall performance of the WSN; in [23] a relative study on the lifetime of a typical WSN node is evaluated. Relevant efforts include performance evaluation concerning time sensitive applications in specific topologies [24], and various environments like industrial test cases [25] and road tunnels [26]. The net result of all these efforts is the creation of specialized communication platforms like the ones presented in [27,28]. The above brief analysis reveals the necessity of effectively reducing the amount data that must be managed. Aiming at mitigating such inefficiencies, effective compression techniques can be valuable tools able to offer significant reduction of the data wirelessly transmitted or/and stored, without compromising information accuracy. Performance efficiency, with respect to data size reduction, is measured by compression rate corresponding to the reduction percentage achieved against the initial size. Furthermore, considering that sensor data are continuously acquired, compression must be executed on-the-fly in order to minimize CPU occupation and assure zero data loss. The latter can be caused from data overrun, which usually occurs due to the limited buffers offered by typical WSN platforms.

Another aspect that highlights the significance of this effort is the potential impact it can have at commercial level considering the respective state-of-art WSN platforms. Indeed nowadays the number of WSN platforms being able to acquire mainly ECG [29–31] and to a lesser degree EEG [31] modalities rapidly increase. However, although the number of features and algorithms offered increase (e.g. encryption algorithms is a quite common example) to the best of our knowledge none of them include efficient and specialized compression algorithms. Therefore, the respective implementations, algorithms' proposals and evaluation conducted in this paper can indeed be of high value for future development platforms.

From another perspective, similar approaches have received high research interest, as it is clearly indicated by the respective papers published in relative journals and conferences [32–35]. However, such efforts offer quite diverse functional characteristics and application domain suitability. In this paper, the main axes characterizing the proposed compression approaches are the following. On one hand design adequacy with respect to low resource communication and processing platforms is required, thus requiring low complexity and high efficiency. On the other hand, the respective design must be also adequate for epilepsy monitoring, which effectively means to exhibit high efficiency regarding compression performance of the respective demanding modalities i.e. EEG/ECG. Literature research has revealed a lack of proposals satisfying both these critical requirements' axes. Therefore following an elicitation process, which is analyzed in detail in the following section, a specific group of compression algorithms that adhere better to the aforementioned requirements has been selected. A significant contribution of this paper relates to the development of all selected approaches in Matlab environment and a consequent experimentally evaluation using real EEG/ECG medical datasets. This effort allows drawing important conclusions regarding the performance and behavior comparative analysis under common a common framework. However, the main contribution of this paper focuses on proposing novel compression schemes targeting at, on optimal "compression rate"–"compression latency" trade-off as well as maximum "compression rate". The evaluation of the proposed algorithms reveals critical advantages against already existing solutions. More specifically, the proposed research work proves to be highly efficient concerning "compression rate" as well as "compression latency". Furthermore, it exhibits critical advantages with respect to achieving the optimum trade-off between considered metrics, thus advocating the used of the proposed solutions over the already existing ones on real scenarios.

Lastly, another critical contribution significantly enhancing the added value of this paper compared to similar research efforts is the utilization of a variety of real experimental EEG and ECG datasets of high sampling frequency and high resolution, so as to achieve valid, objective and practical results. Specially [32,33] although proposing lossless approaches they don't focus on EEG/ECG signals characterized by specific attributes. Also in [34,35] although they offer significant information and background knowledge, respective evaluations are based on general data omitting the specificities of biomedical signals such as EEG/ECG. The datasets have been acquired by using actual WSN sensors as well as from publicly available databases. This aspect offers a critical advantage of this work over relative ones which to the best of the authors' knowledge do not base their performance evaluation on real data. In all cases the performance evaluation focuses on the following two metrics

- Compression rate = (compressed_data_size − uncompressed_data_size)/(uncompressed_data_size).
- Compression latency indicating the time interval required to compressed a specific sample of data.

The rest of the paper is structured as follows: Section 2 presents the rationale behind the main characteristics of the compression schemes considered. Section 3 outlines the theoretical background information focusing on selected compression schemes, while Section 4 describes the proposed extensions. Section 5 presents the experimental setup, while Section 6 presents and analyses the most valuable results and measurements. Section 7 offers a comprehensive yet concise comparative analysis. Finally, Section 8 discusses the main conclusions extracted from the aforementioned measurements and provides directions concerning potential future work.

2. Rationale

A critical categorization of the compression algorithms for the intended application domain could be *lossless* or *lossy*. *Lossless*