Biomedical Signal Processing and Control 11 (2014) 80–88

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Energy efficient telemonitoring of physiological signals via compressed sensing: A fast algorithm and power consumption evaluation

Benyuan Liu[a,*], Zhilin Zhang[b,*], Gary Xu[b], Hongqi Fan[a], Qiang Fu[a]

[a] The Science and Technology on Automatic Target Recognition Laboratory, National University of Defense Technology, Changsha, Hunan 410074, PR China
[b] Samsung Research America - Dallas, Richardson, TX 75082, USA

ARTICLE INFO

Article history:
Received 30 September 2013
Received in revised form 22 February 2014
Accepted 24 February 2014
Available online 25 March 2014

Keywords:
Low-power data compression
Compressed sensing (CS)
Block sparse Bayesian learning (BSBL)
Electrocardiography (ECG)
Electroencephalography (EEG)
Field programmable gate array (FPGA)

ABSTRACT

Wireless telemonitoring of physiological signals is an important topic in eHealth. In order to reduce on-chip energy consumption and extend sensor life, recorded signals are usually compressed before transmission. In this paper, we adopt compressed sensing (CS) as a low-power compression framework, and propose a fast block sparse Bayesian learning (BSBL) algorithm to reconstruct original signals. Experiments on real-world fetal ECG signals and epilepsy EEG signals showed that the proposed algorithm has good balance between speed and data reconstruction fidelity when compared to state-of-the-art CS algorithms. Further, we implemented the CS-based compression procedure and a low-power compression procedure based on a wavelet transform in field programmable gate array (FPGA), showing that the CS-based compression can largely save energy and other on-chip computing resources.

1. Introduction

Monitoring physiological signals via wireless sensor networks is an important topic in wireless healthcare. One major challenge of wireless telemonitoring is the conflict between huge amount of data collected and limited battery life of portable devices [1–3]. Data need to be compressed [4,3] before transmission. Most physiological signals are redundant, which means that they can be effectively compressed [3] using transform encoders such as discrete wavelet transform (DWT) based methods [5]. However, these methods consist of sophisticated matrix–vector multiplication, sorting and arithmetic encoding which subsequently drain the battery.

Compressed sensing (CS), [6], can recover a signal with less measurements given that the signal is sparse or can be sparse represented in some transformed domains. CS-based wireless telemonitoring technology [7–12] can thus be viewed as a lossy compression method. The block diagram of a typical CS-based wireless telemonitoring is shown in Fig. 1. Physiological signals are firstly digitalized (Nyquist sampling) via an analog to digital converter (ADC). Those digitalized samples are compressed by a simple matrix–vector multiplication and the results are transmitted via wireless networks. At the data central, a CS algorithm is used to recover original signals from the compressed measurements.

1.1. Overview of the compressed sensing

The basic goal of CS aims to solve the following underdetermined problem:

$$\min \|\mathbf{x}\|_1 \quad s.t. \quad \mathbf{y} = \boldsymbol{\Phi}\mathbf{x}, \tag{1}$$

where $\mathbf{x}$ is the samples, $\boldsymbol{\Phi}$ is the sensing matrix whose row number is smaller than column number, and $\mathbf{y}$ is the compressed measurements. $||\mathbf{x}||_1$ is the ℓ_1 norm penalty of $\mathbf{x}$, which prompts its sparsity.

In practice, physiological signals are not sparse in the time domain, therefore one often resorts to a transformed domain such that $\mathbf{x}$ can be expressed as $\mathbf{x} = \mathbf{D}\boldsymbol{\theta}$ where $\mathbf{D}$ is a dictionary matrix such that the representation coefficients $\boldsymbol{\theta}$ are much sparser than $\mathbf{x}$. The problem in (1) then becomes

$$\min \|\boldsymbol{\theta}\|_1 \quad s.t. \quad \mathbf{y} = (\boldsymbol{\Phi}\mathbf{D})\boldsymbol{\theta}, \tag{2}$$

The signal can be reconstructed afterwards using $\hat{\mathbf{x}} = \mathbf{D}\hat{\boldsymbol{\theta}}$ with the recovered coefficients $\hat{\boldsymbol{\theta}}$. Most CS-based telemonitoring systems [8,9] are build upon this model.

* Corresponding authors. Tel.: +86 13720726249.
E-mail addresses: liubenyuan@gmail.com, lbyoopp@163.com (B. Liu), zhilinzhang@ieee.org (Z. Zhang).

http://dx.doi.org/10.1016/j.bspc.2014.02.010
1746-8094/

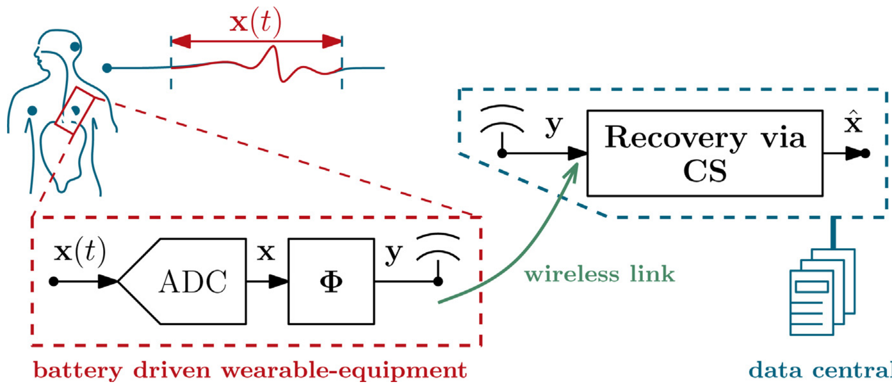

Fig. 1. The diagram of a compressed sensing (CS) based wireless telemonitoring system.

Recent advance in CS algorithms is to incorporate physical information [13–15] into the optimization procedure with the goal to achieve better reconstruction performance. One structure widely used is the block/group sparse structure [13,16–18], which refers to the case when nonzero entries of a signal cluster around some locations. Moreover, noticing intra-block correlation widely exists in real-world signals, Zhang and Rao [13,19] proposed the block sparse Bayesian learning (BSBL) framework. It showed superior ability to recover block sparse signals or even non-sparse raw physiological signals such as fetal ECG [11] and EEG signals [10].

1.2. Summary of contributions

BSBL algorithms [13] showed impressive recovery performance on physiological signals such as ECG and EEG. However, these algorithms derived so far are not fast and may limit their applications. The first contribution of our work is a fast implementation[1] of the BSBL framework using the fast marginalized (FM) likelihood maximization method [20]. Experiments conducted on real-life physiological signals showed that the proposed algorithm had similar recovery quality as BSBL algorithms, but was much faster.

Power consumption is a major concern in wireless telemonitoring systems. Traditionally, the power consumption was evaluated on a low-power microcontroller (MCU) [8]. However, MCU does not support fully parallel implementation and the power estimate is affected by the coding style. In this work, we analyzed the power consumption on field programmable gate array (FPGA). In FPGA, we can implement the compressor in parallel and control the overall activities. Only the logic cells related to the compression core are implemented and the rest are holding reset. Therefore the power estimate is more accurate. In the experiment, the CS-based compressor was compared to a low-power DWT-based compressor in terms of compression latency, the number of utilized on-chip resources and power consumption. We proved that the CS-based architecture was more suitable for low-power physiological telemonitoring applications.

1.3. Outline and notations

The rest of the paper is organized as follows. Section 2 presents the fast marginalized implementation of the BSBL algorithm and Section 3 provides the simulation setup and evaluation metrics. In Section 4 and Section 5, we conduct experiments on fetal ECG (FECG) and EEG signals. The extracted FECGs and the epileptic seizure classification results are used to evaluate the performance of CS. FPGA implementations and power consumption of the CS-based and the DWT-based compression methods are given in Section 6. Conclusion is drawn in the last section.

Throughout the paper, **Bold** letters are reserved for vectors $\mathbf{x}$ and matrices $\mathbf{X}$. $\mathrm{Tr}(\cdot)$ computes the trace of a matrix and $\mathrm{diag}(\mathbf{A})$ extracts the diagonal vector of the matrix $\mathbf{A}$. $(\cdot)^T$ is the transpose operator. $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a multivariate Gaussian distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$.

2. The fast implementation of the BSBL framework

2.1. Overview of the BSBL framework [13]

A block sparse signal $\mathbf{x}$ has the following structure:

$$\mathbf{x} = [\underbrace{x_1, \ldots, x_{d_1}}_{\mathbf{x}_1^T}, \ldots, \underbrace{x_1, \ldots, x_{d_g}}_{\mathbf{x}_g^T}]^T, \tag{3}$$

which means $\mathbf{x}$ has g blocks, and only a few blocks are nonzero. Here d_i is the block size for the ith block. The BSBL algorithms [13] exploit the block structure and the intra-block correlation by modeling the signal block $\mathbf{x}_i$ using the parameterized Gaussian distribution:

$$p(\mathbf{x}_i; \gamma_i, \mathbf{B}_i) = \mathcal{N}(\mathbf{x}_i; \mathbf{0}, \gamma_i \mathbf{B}_i). \tag{4}$$

with unknown deterministic parameters γ_i and $\mathbf{B}_i$. γ_i is a nonnegative parameter controlling the block-sparsity of $\mathbf{x}$ and $\mathbf{B}_i$ is a positive definite matrix modeling the covariance structure of $\mathbf{x}_i$. We assume that the blocks are mutually independent. Henceforth,

$$p(\mathbf{x}; \{\gamma_i, \mathbf{B}_i\}_i) = \mathcal{N}(\mathbf{x}; \mathbf{0}, \boldsymbol{\Gamma}), \tag{5}$$

where $\boldsymbol{\Gamma}$ denotes a block diagonal matrix with the ith principal block given by $\gamma_i \mathbf{B}_i$.

The measurement noise is assumed to be independent and Gaussian with zero mean and unknown variance β^{-1}. Thus the measurement model is

$$p(\mathbf{y}|\mathbf{x}; \beta) = \mathcal{N}(\mathbf{y}; \boldsymbol{\Phi}\mathbf{x}, \beta^{-1}\mathbf{I}). \tag{6}$$

Given the signal model (5) and the measurement model (6), the posterior $p(\mathbf{x}|\mathbf{y}; \{\gamma_i, \mathbf{B}_i\}_i, \beta)$ and the likelihood $p(\mathbf{y}|\{\gamma_i, \mathbf{B}_i\}_i, \beta)$ can be derived as follows:

$$p(\mathbf{x}|\mathbf{y}; \{\gamma_i, \mathbf{B}_i\}_i, \beta) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}), \tag{7}$$

$$p(\mathbf{y}|\{\gamma_i, \mathbf{B}_i\}_i, \beta) = \mathcal{N}(\mathbf{y}; \mathbf{0}, \mathbf{C}) \tag{8}$$

where $\boldsymbol{\Sigma} \triangleq (\boldsymbol{\Gamma}^{-1} + \boldsymbol{\Phi}^T \beta \boldsymbol{\Phi})^{-1}$, $\boldsymbol{\mu} \triangleq \boldsymbol{\Sigma}\boldsymbol{\Phi}^T \beta \mathbf{y}$ and $\mathbf{C} \triangleq \beta^{-1}\mathbf{I} + \boldsymbol{\Phi}\boldsymbol{\Gamma}\boldsymbol{\Phi}^T$. To estimate the parameters $\{\gamma_i, \mathbf{B}_i\}_i$ and β, the following cost function is used, which is derived according to the Type II maximum likelihood [13]:

$$\mathcal{L}(\{\gamma_i, \mathbf{B}_i\}_i, \beta) = -2 \log p(\mathbf{y}|\{\gamma_i, \mathbf{B}_i\}_i, \beta) \tag{9}$$

[1] The preliminary work of the developed algorithm is available in http://arxiv.org/abs/1211.4909.