Biomedical Signal Processing and Control 35 (2017) 1–7

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Respiratory rate estimation from the photoplethysmogram via joint sparse signal reconstruction and spectra fusion

Xiaorong Zhang[a,*], Quan Ding[b]

[a] School of Engineering, San Francisco State University, San Francisco, CA, USA
[b] The Home Depot Techshed, Foster City, CA, USA

ARTICLE INFO

Article history:
Received 28 October 2016
Received in revised form 8 January 2017
Accepted 3 February 2017

Keywords:
Photoplethysmogram
Respiratory rate estimation
Joint sparse signal reconstruction
Spectra fusion

ABSTRACT

Respiratory rate (RR) estimation from the photoplethysmogram (PPG) is a challenging problem due to the nonstationarity of RR and disturbance. In this work, we propose a novel approach to estimate RR from the PPG signal using joint sparse signal reconstruction (JSSR) and spectra fusion (SF). A window of PPG signal is segmented into multiple overlapped measurements. Sparse spectra of these measurements are estimated by JSSR using the regularized M-FOCUSS algorithm. The kurtosis of each spectrum is used to classify it into three signal quality categories, and spectra in the highest signal quality category are fused using Respiratory Rate Tracking (RRT) to estimate RR. Validated on a public benchmark database CapnoBase, our approach outperforms a state-of-the-art algorithm in accuracy and robustness in low signal quality conditions. This is the first time JSSR has been used for RR estimation from the PPG signal. In addition, our approach works well with a low sampling frequency of 10 Hz which has great potential to be used in low-cost wearable devices.

1. Introduction

Respiratory rate (RR), together with body temperature, heart rate, and blood pressure, constitute the four primary vital signs that indicate the status of the body's vital functions. Abnormal RR is often a sign of serious illness and can be used to predict clinical deterioration [1–7]. Therefore, accurately estimating RR in hospital settings is of great importance to both patients and health care providers.

RR is measured by the number of breaths a person takes per minute (breaths/min). Capnography monitors the concentration or partial pressure of carbon dioxide (CO_2) in the respiratory gases [8], and therefore provides the most accurate RR measurement. But due to its cumbersomeness, capnography is mainly used during anesthesia and intensive care. As an alternative solution, researchers have used electrocardiogram (ECG) to derive RR [9–11]. Despite the progress made, current hospital settings still suffer from inaccurate RR monitoring. It was recently observed that the ECG-derived respiratory waveform often appeared flat in ICU patients who were breathing adequately [12]. Furthermore, the ECG system is still bulky and requires trained professionals to operate.

More recently, pulse oximeter became an inexpensive solution for RR monitoring. Based on the principle of photoplethysmography (PPG), it uses the Beer-Lambert law to estimate a person's oxygen saturation (SpO_2) level by measuring the attenuation of light traveling through the tissue at finger or earlobe [13,14]. The PPG signal is modulated by respiration in three ways: (1) respiratory-induced intensity variation (RIIV), (2) respiratory-induced amplitude variation (RIAV), and (3) respiratory-induced frequency variation (RIFV), also known as baseline modulation, amplitude modulation, and frequency modulation, respectively [15–18]. Many algorithms have been developed to estimate RR from PPG signal using signal processing and machine learning techniques such as neural network [19], continuous wavelet transform [20,21], short-time Fourier transform [22], independent component analysis [23], autoregressive model [24,25], variable-frequency complex demodulation [26], particle filter [27], modified multi-scale principal component analysis [28], pulse width variability [29], smart fusion [16], correntropy spectral density [30], probabilistic approach [18], sparse signal reconstruction [31], multi scale independent component analysis (MSICA) [32], and singular spectrum analysis (SSA) [33]. All these methods aim to estimate RR from a window of PPG signal. However, for those spectral domain methods, they take the window as a whole and ignores its local nonstationarity. In more difficult situations when the respiratory modulation changes within the window or when disturbance occurs in only part of the window, the spectral domain methods may have downgraded performance.

* Corresponding author.
E-mail addresses: xrzhang@sfsu.edu, zxiaorong@ele.uri.edu (X. Zhang).

http://dx.doi.org/10.1016/j.bspc.2017.02.003
1746-8094/

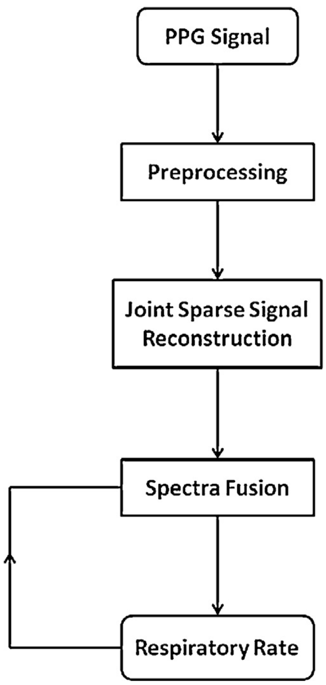

Fig. 1. Diagram of the JSSR-SF approach.

In this paper, we propose a novel approach termed "Joint Sparse Signal Reconstruction and Spectra Fusion" (JSSR-SF), which is a spectral domain method but works with PPG windows that are locally nonstationary. It consists of two major components: joint sparse signal reconstruction (JSSR) and spectra fusion (SF). JSSR aims to find a sparse representation of the PPG signal in the spectral domain by solving the multiple measurement vector (MMV) model [34]. We segment one PPG window into multiple overlapped measurements and use JSSR to obtain sparse spectrum for each segment by exploiting a common sparsity structure. Based on that, we propose two SF methods which estimates RR by fusing the sparse spectra. The JSSR-SF approaches are tested on a public benchmark database for respiratory signal analysis named CapnoBase [35] and outperforms a state-of-the-art algorithm in root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and number of reported RR estimates.

The paper is organized as follows. Section 2 describes the proposed JSSR-SF approach in detail. Section 3 presents the experimental results of JSSR-SF tested on CapnoBase. Section 4 discusses the contribution and limitation of this work. Finally, Section 5 draws the conclusion.

2. Joint sparse signal reconstruction and spectra fusion

The diagram of the proposed JSSR-SF approach is summarized in Fig. 1. It consists of a preprocessing step and two major steps: JSSR and SF. In this section, we first introduce the motivation behind the JSSR-SF approach and then describe the JSSR and SF steps in more details.

2.1. Motivation

Sparse signal reconstruction (SSR) is a signal processing technique aimed to find localized energy solutions from limited data. It is advantageous over traditional spectral estimation techniques such as periodogram because SSR has higher spectral resolution to separate frequency components that are close to each other (see Fig. 1 in [36] as an example). Recently, it was successfully used in the "Sparse Signal Reconstruction and Respiratory Rate Tracking" (S2R3T) approach for RR estimation from the PPG signal [31]. The sparse spectrum of the PPG signal is estimated by solving the single measurement vector (SMV) model as follows:

$$\mathbf{y} = \Phi \mathrm{x} + \mathbf{w} \quad (1)$$

where $\mathbf{y}$ is an $M \times 1$ vector of the PPG signal, $\boldsymbol{\Phi}$ is an $M \times N$ basis matrix, $\mathbf{x}$ is an $N \times 1$ solution vector to be estimated, and $\mathbf{w}$ is an $M \times 1$ noise vector. It is assumed that $M < N$ and $\boldsymbol{\Phi}_{m,n} = e^{j2\pi mn/N}$ so (1) can be thought of as a redundant discrete Fourier transform with additive noise. Since this is an underdetermined system, we need to have some constraints otherwise $\mathbf{x}$ will have an infinite number of solutions. For SSR, the constraint is that $\mathbf{x}$ is sparse, i.e., most elements in $\mathbf{x}$ are zero or close to zero, and only a few elements have large nonzero values. The "FOcal Underdetermined System Solver" (FOCUSS) algorithm [37] can be used to find the solution for (1).

The solution vector $\mathbf{x}$ represents the spectral energy of the PPG signal, which is preprocessed by a band-pass filter to remove baseline wonder, cardiac cycle modulation, and high-frequency artifacts. Respiratory modulation remains in the preprocessed PPG signal and is the major frequency component. As a result, the global or local maximum of the sparse spectrum $\mathbf{x}$ can be used to estimate RR, depending on the signal quality of the preprocessed PPG signal [31]. This step to estimate RR from $\mathbf{x}$ is termed "Respiratory Rate Tracking" (RRT) [31]. In RRT, kurtosis of the sparse spectrum is used to classify the signal quality into three categories: good, moderate, and poor. When signal quality is good, the global maximum in $|\mathbf{x}|$ is used to estimate RR. When signal quality is moderate, the local maximum in $|\mathbf{x}|$ near the previous RR is used to estimate RR because the global maximum may be from disturbance. In these two cases, RR is estimated as

$$\mathrm{RR} = \frac{n_{\max}}{N} F_s \times 60\,\text{breaths/min} \quad (2)$$

where $n_{\max}$ is the position of the global/local maximum, N is the length of $\mathbf{x}$, and F_s is the sampling frequency of the PPG signal. When signal quality is poor, no RR estimate is reported because the RR estimate is not reliable.

Note that S2R3T is a spectral domain approach, and it estimates the sparse spectrum from the whole PPG window. However, there are many cases when respiration and/or disturbance are locally nonstationary, which may bias the peak of the sparse spectrum. To overcome this drawback, we propose using JSSR to estimate the sparse spectra of the PPG signal by solving the MMV model, and then fusing the resulting spectra to estimate RR. We provide more details of each step in the next two subsections.

2.2. Joint sparse signal reconstruction

JSSR is an extension of SSR when multiple measurements are available and we want to estimate localized energy solution for each measurement using a common basis matrix. It can be expressed by the MMV model below:

$$\mathbf{Y} = \Phi \mathrm{X} + \mathbf{W} \quad (3)$$

where $\mathbf{Y}$ is an $M \times L$ matrix consisting of L measurements of the PPG signal, $\boldsymbol{\Phi}$ is an $M \times N$ basis matrix with $\boldsymbol{\Phi}_{m,n} = e^{j2\pi mn/N}$, $\mathbf{X}$ is an $N \times L$ solution matrix to be estimated, and $\mathbf{W}$ is an $M \times L$ noise matrix. Note that (3) can also be expressed as

$$\mathbf{y}_i = \boldsymbol{\Phi}\mathbf{x}_i + \mathbf{w}_i \quad \text{for } i = 1, 2, \ldots, L \quad (4)$$

where $\mathbf{Y} = [\mathbf{y}_1\ \mathbf{y}_2\ \ldots\ \mathbf{y}_L]$, $\mathbf{X} = [\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_L]$, and $\mathbf{W} = [\mathbf{w}_1\ \mathbf{w}_2\ \ldots\ \mathbf{w}_L]$. Therefore, the MMV model can be considered as L SMV models with a common basis matrix $\boldsymbol{\Phi}$, and each $\mathbf{x}_i$ is the spectrum of the ith measurement $\mathbf{y}_i$.