Biomedical Signal Processing and Control 31 (2017) 309–320

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

# Video-based human heart rate measurement using joint blind source separation

Huan Qi[a], Zhenyu Guo[a], Xun Chen[b,*], Zhiqi Shen[c], Z. Jane Wang[a]

[a] Department of Electrical and Computer Engineering, University of British Columbia, V6T 1Z4 Vancouver, Canada
[b] Department of Biomedical Engineering, Hefei University of Technology, 230009 Hefei, China
[c] School of Computer Engineering, Nanyang Technological University, Singapore 639798, Singapore

ARTICLE INFO

*Article history:*
Received 18 February 2016
Received in revised form 20 May 2016
Accepted 21 August 2016

*Keywords:*
Heart rate (HR)
Facial landmark
Joint blind source separation (J-BSS)
Non-contact

ABSTRACT

Remote (non-contact) measurements of human cardiopulmonary physiological parameters based on photoplethysmography (PPG) can lead to efficient and comfortable medical assessment, which is important in human healthcare. It was shown that human facial blood volume variation during cardiac cycle can be indirectly captured by common Red–Green–Blue (RGB) cameras. In this paper, we show that it is promising to incorporate data from different facial sub-regions to improve remote measurement performance. We propose a novel method for non-contact video-based human heart rate (HR) measurement by exploring correlations among facial sub-regions via joint blind source separation (J-BSS). To our knowledge, this is the first time that J-BSS approaches, instead of prevailing BSS techniques such as independent component analysis (ICA), is successfully applied in non-contact physiological parameter measurement. We test the proposed method on a large public database, which provides the subjects' left-thumb plethysmograph signals as ground truth. Experimental results show that the proposed J-BSS method outperforms previous ICA-based methodologies.



## 1. Introduction

Convenient measurements of human physiological parameters are of great potential for both clinic diagnosis and daily healthcare. In recent years, medical monitoring using non-contact detection techniques has been widely studied. Efforts have been made to develop non-contact monitoring methods without attaching any medical electrode or sensor to the subject. Some of them are clinically tested, such as vital signs monitoring during haemodialysis [1], neonatal intensive care unit [2], quantification of limb movement in epileptic seizure [3], and dynamic tissue phantoms evaluation [4]. Family healthcare also benefits a lot from non-contact detection techniques, especially with the rapid dissemination of smartphones [5]. Commercial apps such as Cardiio (Cardiio, Inc., San Francisco, CA, USA) and Vital Signs Camera [6] (Philips, Inc., Amsterdam, Netherlands) enable users to measure heart rates or respiratory rates using continuous recordings of their faces by front cameras on the phones.

Among many human physiological parameters, cardiovascular parameters are of great interest. Surveillance and prevention of cardiovascular disease (CVD) require regular medical assessment of heart rate and heart rate variability (HRV). As a prognostic factor and potential therapeutic target, HR has been verified in large epidemiological studies to be an independent predictor of cardiovascular and all-cause mortality for people with or without diagnosed cardiovascular disease [7]. The invention of stethoscope nearly 200 years ago was a milestone for heart rate monitoring (HRM). Currently, electrocardioscopes based on electrocardiogram (ECG) are widely used in HRM due to high reliability and low cost. Nowadays, commercial pulse oximeters can measure HR and blood oxygen saturation within a few seconds using sensors attached on a certain region of human body (usually finger tip or ear lobe). However, these are contact devices with adhesive electrodes or sensors. Placement and removal of these attachments can cause discomfort, stress and even epidermal stripping [2].

Photoplethysmography (PPG) is an indirect but effective technique to measure cardiovascular blood volume pulse (BVP). During cardiac cycle, the variations of tissue blood volume in certain human body segments modulate the transmission or reflection of visible light at those segments. Traditionally PPG sensors are used to capture the strongest BVP signal in dedicated light source, and heart rate can be estimated correspondingly by measuring time intervals between consecutive peaks of the signal [1]. Prevailing PPG

* Corresponding author.
*E-mail address:* xunchen@ece.ubc.ca (X. Chen).

http://dx.doi.org/10.1016/j.bspc.2016.08.020
1746-8094/

sensors (e.g., pulse oximeters) require contact with the subject. Many efforts have been made for non-contact HR measurement. Some of them used dedicated sensors such as Doppler wave sensor [8–10], and thermal imaging sensor [11]. The study in [12] showed, for the first time, that PPG signal can be remotely acquired from human face using consumer-level digital cameras in ambient light. Later Poh et al. [13] presented an efficient ICA framework to measure HR using a low-cost webcam and ambient light. A robust method [1] was devised to cancel out aliased frequency components from artificial light flicker and enable non-contact measurement of HR, RR and oxygen saturation even under strong fluorescent lights. Real-time measurement using continuous wavelet transform (CWT) can be achieved despite the existence of light and motion artifacts [14]. To overcome the frequency resolution limitation of traditional RGB sensors, Mcduff et al. [15] presented a modified five band (RGBCO) digital camera with cyan and orange frequency bands added to the original red, green and blue colour channels. However, aforementioned techniques extract face colour channel data by averaging over the entire facial region. Few studies consider potential variations among different facial sub-regions and how these variations might contribute to the measurements of physiological parameters such as HR, where a 'sub-region' is a region containing only one part of a human face. Guo et al. first proposed to use independent vector analysis (IVA) to jointly analyse colour signals from different facial sub-regions. Preliminary experiments showed more accurate measurement of HR compared to ICA-based method [16]. Kumar et al. recently reported that a weighted average over skin-colour variation signals from different facial sub-regions contributes to the improvement of signal-to-noise ratio (SNR) of camera-based vital sign monitoring [17]. Unlike their weighted signal combination methodology, in this paper we propose a novel method for non-contact HR measurement by exploring data correlations among specified facial sub-regions.

Conventional blind source separation (BSS) techniques such as ICA are originally designed for the decomposition of a single dataset, e.g., the colour channel data of the entire facial region, into constituent components, as illustrated in Fig. 1(a). When considering multiple datasets, e.g. colour channel data of multiple facial sub-regions, we are interested in recently proposed joint blind source separation (J-BSS) methods [18–21], as shown in Fig. 1(b). J-BSS methods have been successfully used in many biomedical applications, such as joint investigation of multi-modal datasets from a corticomuscular coupling experiment [22,23] and analysis of concurrent EEG and kinematic data [23]. The goal of J-BSS in these applications is to extract underlying sources within each dataset and meanwhile keep a consistent ordering of the extracted sources across multiple datasets. This can be extremely challenging for traditional BSS techniques. In the proposed method, we treat colour channel data of each facial sub-region as a multidimensional dataset, where the number of dimension depends on how many colour channels are used. Joint analysis of such multidimensional datasets allow us to identify reliable underlying sources corresponding to the BVP signal. We further explore whether an 'optimal' correlation combination of these multidimensional datasets exists so that the objective of J-BSS would be better achieved with respect to certain specific measures such as measurement accuracy.

In this paper, we focus on non-contact HR measurement using uncompressed facial video recording. The main contributions of this paper are threefold: (1) we propose a novel non-contact HR measurement method that jointly analyzes video data of different facial sub-regions via J-BSS methods; (2) we propose to exploit a max-margin multi-label (M3L) classifier to predict the optimal correlation combination of input datasets from different facial sub-regions to achieve higher HR measurement accuracy; (3) we test the proposed method on a large public database (DEAP [24]) and verify its performance.

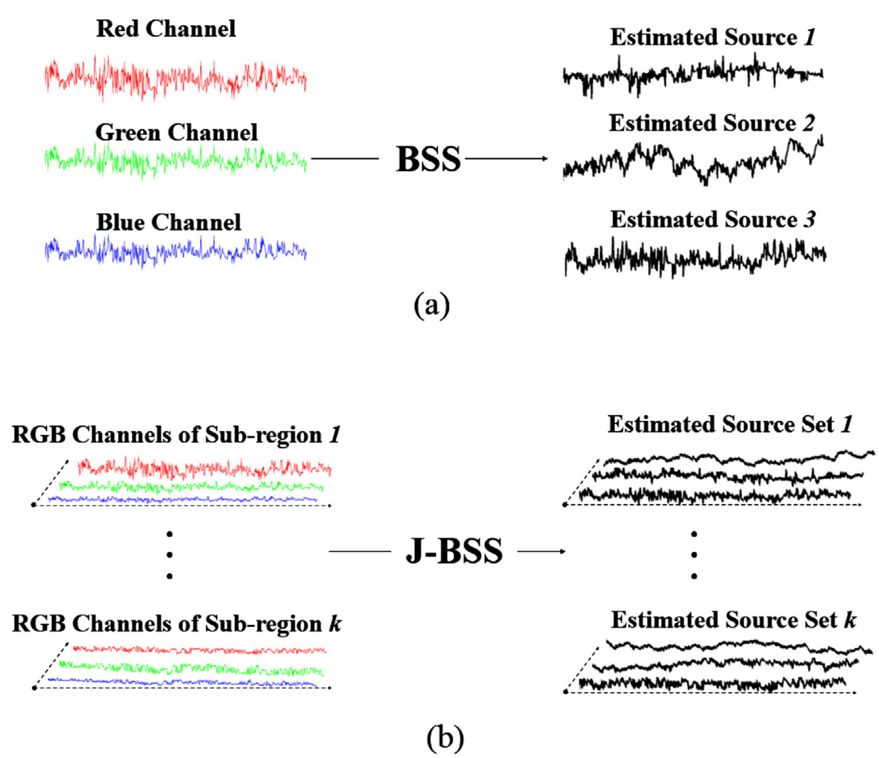


**Fig. 1.** A brief overview of two types of source separation methods, BSS and J-BSS. (a) The entire face colour values are averaged in three channels (RGB) and then calculated by a BSS algorithm into three estimated sources. (b) The input of J-BSS methods are colour channel datasets from different facial sub-regions. Each sub-region generates a dataset by averaging colour values over the sub-region. Details of J-BSS formulation is discussed in Section 2.2).