Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational and Applied Mathematics 223 (2009) 409–420
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Semi-blind source extraction algorithm for fetal electrocardiogram based on generalized autocorrelations and reference signals

Hongjuan Zhang [a], Zhenwei Shi [b], Chonghui Guo [c,*], Enmin Feng [a]

[a] Department of Applied Mathematics, Dalian University of Technology, Dalian 116024, PR China
[b] Image Processing Center, School of Astronautics, Beijing University of Aeronautics and Astronautics, Beijing 100083, PR China
[c] Institute of Systems Engineering, Dalian University of Technology, Dalian 116024, PR China

Received 15 November 2007; received in revised form 17 January 2008

Abstract

Blind source extraction (BSE) has become one of the promising methods in the field of signal processing and analysis, which only desires to extract "interesting" source signals with specific stochastic property or features so as to save lots of computing time and resources. This paper addresses BSE problem, in which desired source signals have some available reference signals. Based on this prior information, we develop an objective function for extraction of temporally correlated sources. Maximizing this objective function, a semi-blind source extraction fixed-point algorithm is proposed. Simulations on artificial electrocardiograph (ECG) signals and the real-world ECG data demonstrate the better performance of the new algorithm. Moreover, comparisons with existing algorithms further indicate the validity of our new algorithm, and also show its robustness to the estimated error of time delay.

Keywords: Independent component analysis; Semi-blind source extraction; Constrained optimization; Generalized autocorrelation; Fixed-point algorithm

1. Introduction

Over the past decades the problem of blind source separation (BSS) [8,9,11,18] has received much research attention because of its potential applicability to a wide range of problems, such as communications signals and biomedical signals analysis and processing, geophysical data processing, data mining, speech analysis, image recognition, texture modelling and so on [1,2,4,5,7,12,15,16]. In BSS problems, the multidimensional observations must be processed to recover the original sources without the benefit of any a priori knowledge about the mixing operation or the sources themselves. Generally, classical BSS methods consider the simultaneous recovery of all the independent components from their linear mixtures. However, in practice, extracting all the source signals from a large number of observed sensor signals, for example, a magnetoencephalographic (MEG) measurement which may output hundreds of recordings, could take a long time and in these signals only a very few are desired with given characteristics. In this case it is more practical to recover a subset of the source only. This is known as blind source extraction (BSE). When combined with a deflation procedure, BSE algorithms can be viewed as the methods

* Corresponding author. Tel.: +86 411 84708007; fax: +86 411 84748135.
E-mail address: guochonghui@tsinghua.org.cn (C. Guo).

0377-0427/$ - see front matter
doi:10.1016/j.cam.2008.01.031

of sequential extraction of all the independent sources [10]. BSE may have several advantages over simultaneous BSS [8]. For example, only "interesting" source signals need to be extracted; signals can be extracted in a specific order according to some features of source signals; lots of computing time and resources can be saved. Therefore, it has become a promising method in various fields such as biomedical signal processing and analysis, data mining, speech and image processing, and so on [3,8,11,18].

Nowadays, many source extraction algorithms have been developed through optimization of different cost functions, generally based on high-order statistics [1,6,8,13,18] to extract a source signal. And these methods have been used successfully in many fields. However, it is computationally expensive to exploit high-order statistics. Thus the trend is to develop second-order statistics based extraction algorithms using the priori knowledge about source signals, such as sparseness [29], high-order statistics [27], smoothness or linear predictability [3,8], or time structure [3,7,14,26]. Recently, Lu et al. [20–22] proposed a good candidate, that is ICA with reference (ICA-R), for extracting several source signals from a large number of observed signals. It was formed by minimizing the less-complete ICA objective function and makes the best of the traces of the interesting sources referred to reference signals which carry some prior information to distinguish the desired components but are not identical to the corresponding sources. This method has become an efficient approach utilizing prior information and it has been successfully used for fMRI data analysis etc.

Moreover, in many applications, such as ECG extraction, the desired source signal is periodic or quasi-periodic. So the period property can be used as prior information to extract the desired source signal. Barros and Cichocki [3] provided a simple batch learning algorithm (simplified "BCBSE algorithm") for semi-blind extraction of the desired source signal, which can extract the desired source as long as they are decorrelated and show a temporal structure. However, this method is only carried out the constrained minimization of the mean squared error, which can not accurately describe the probability distribution of the innovation of the signals. It is a possible reason why this algorithm is very sensitive to the estimation error of the time delay and cannot reliably cancel noise contamination in the recovered signals [24]. Recently, Shi and Zhang [24] developed a semi-blind source extraction method (simplified "SemiBSE algorithm"), which based on the non-Gaussianity and the autocorrelations of the desired source signal. This method has been successfully used for fetal electrocardiogram (FECG) extraction, and its advantage in its tolerance to large estimate errors of the period has been pointed in [24]. However, its performance was not entirely satisfactory because the recovered signals often included noise contamination. In [25], authors addressed the semi-blind source extraction problem when the desired source signals have linear or nonlinear autocorrelations. Based on the generalized autocorrelations of the primary sources, a BSE algorithm (called "GABSE algorithm") was proposed. It has been shown that the GABSE algorithm has good stability and convergence, moreover it possesses a higher accuracy of extraction. However, this algorithm is not very robust to the estimation error of the time delay. The width of estimate errors of the period is limited to only about ten time delays, which is smaller than that of SemiBSE algorithm in [24].

In order to improve extraction performance and the tolerance to the estimated error of the time delay, we develop an objective function for extraction of temporally correlated source in this paper, which based on generalized autocorrelations and reference information of desired signals. Maximizing this objective function, we propose a semi-blind source extraction fixed-point algorithm. This algorithm incorporates more priori information of the desired signal, which can be viewed as a refined substitution of the GABSE algorithm. It is able to extract decorrelated periodic source, which is the closest one, in some sense, to the reference signal when the closeness measure is properly chosen. The following simulation experiments show that our new algorithm outperforms many existing algorithms, such as the BCBSE algorithm, SemiBSE algorithm and GABSE algorithm.

This manuscript is organized as follows. In Section 2, we provide a new cost function which is based on the generalized autocorrelations and reference signals of the desired sources, after which we derive the semi-blind source extraction fixed-point algorithm. Section 3 demonstrates the present technique with experiments using artificial ECG signals and real-world ECG data. Conclusions are drawn in the final section.

2. Proposed algorithms

2.1. Objective function

Denote the observed sensor signals $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))^{\mathrm{T}}$ described by matrix equation

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \tag{1}$$