Journal of Computational and Applied Mathematics 223 (2009) 908–915

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

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

A fixed-point algorithm for blind source separation with nonlinear autocorrelation☆

Zhenwei Shi *, Zhiguo Jiang, Fugen Zhou

Image Processing Center, School of Astronautics, Beijing University of Aeronautics and Astronautics, Beijing 100083, PR China

ARTICLE INFO

Article history:
Received 8 December 2007
Received in revised form 8 March 2008

Keywords:
Blind source separation (BSS)
Independent component analysis (ICA)
Nonlinear autocorrelation
Fixed-point algorithm

ABSTRACT

This paper addresses blind source separation (BSS) problem when source signals have the temporal structure with nonlinear autocorrelation. Using the temporal characteristics of sources, we develop an objective function based on the nonlinear autocorrelation of sources. Maximizing the objective function, we propose a fixed-point source separation algorithm. Furthermore, we give some mathematical properties of the algorithm. Computer simulations for sources with square temporal autocorrelation and the real-world applications in the analysis of the magnetoencephalographic recordings (MEG) illustrate the efficiency of the proposed approach. Thus, the presented BSS algorithm, which is based on the nonlinear measure of temporal autocorrelation, provides a novel statistical property to perform BSS.

1. Introduction

Recently, blind source separation (BSS) [6,11], as an increasingly popular data analysis technique, has received wide attention in various fields such as biomedical signal processing and analysis, speech and image processing, wireless telecommunication systems, data mining, etc. The main task of BSS is to recover original sources from their mixtures using some statistical properties of original sources.

The blind source separation problem has been studied by researchers in applied mathematics, neural networks and statistical signal processing. Several methods for BSS using the statistical properties of original sources have been proposed, such as non-Gaussianity (or equivalently, independent component analysis, ICA) [1,3,5–8,11–13,24,34], linear predictability or smoothness [2,6], linear autocorrelation [4,19,31], coding complexity [9,26,27,29], temporal predictability [23], nonstationarity [10,18,21], sparsity [14,15,25,35], energy predictability [28], nonlinear innovation [30] and nonnegativity [20,22], etc.

In this paper, we present a way using nonlinear autocorrelation of source signals for BSS. We propose a fixed-point algorithm for BSS based on nonlinear autocorrelation and analyze its stability conditions. We show that the BSS problem can be solved by maximizing nonlinear temporal autocorrelation of the sources. When sources have square temporal autocorrelation, we demonstrate that the efficient implementation of the method. Also, the proposed algorithm can extract the most interesting and meaningful sources in brain magnetoencephalography (MEG) data analysis. Thus, the presented BSS algorithm, which is based on the nonlinear measure of temporal autocorrelation, provides a new statistical property to perform BSS.

The structure of the paper is as follows. The objective function based on the nonlinear autocorrelation of the desired source signals, and a fixed-point algorithm for optimizing the objective function are proposed in Section 2. Furthermore, we analyze the stability conditions of the fixed-point algorithm in Section 3. In Section 4, experiments on different datasets are presented. Some conclusions are drawn in Section 5.

☆ The work was supported by the National Science Foundation of China (60605002, 60776793, 10571018).

* Corresponding author. Tel.: +86 10 627 96 872; fax: +86 10 627 86 911.
E-mail address: shizhenwei@mail.tsinghua.edu.cn (Z. Shi).

0377-0427/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2008.03.009

2. Proposed algorithm

2.1. Objective function

We observe sensor signals $\mathbf{x}(t) = (x_1(t), \ldots, x_n(t))^{\mathrm{T}}$ described in the matrix notation:

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

where $\mathbf{A}$ is an $n \times n$ unknown mixing matrix, $\mathbf{s}(t) = (s_1(t), \ldots, s_n(t))^{\mathrm{T}}$ is a vector of unknown zero-mean and unit-variance primary sources.

The basic problem of BSS is then to estimate both the mixing matrix $\mathbf{A}$ and the source signals $s_i(t)$ using only observations of the mixtures $x_i(t)(i = 1, \ldots, n)$, i.e., to find an $n \times n$ separating matrix $\mathbf{W} = (\mathbf{w}_1, \ldots, \mathbf{w}_N)^{\mathrm{T}}$ such that the unmixed signals $\mathbf{y}(t) = (y_1(t), \ldots, y_N(t))^{\mathrm{T}}$,

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t) \tag{2}$$

are the estimated source signals. The sources are recovered up to scaling and permutation.

A useful preprocessing strategy in BSS is to first whiten the observed variables [11]. This means that we transform the observed vector $\mathbf{x}$ linearly so that we obtain a new vector $\tilde{\mathbf{x}}$ which is white, i.e., its components are uncorrelated and their variances equal unity. In other words, $\mathbf{x}$ is linearly transformed into a vector

$$\tilde{\mathbf{x}}(t) = \mathbf{V}\mathbf{x}(t) = \mathbf{V}\mathbf{A}\mathbf{s}(t) \tag{3}$$

whose covariance matrix equals the identity matrix: $E\{\tilde{\mathbf{x}}(t)\tilde{\mathbf{x}}(t)^{\mathrm{T}}\} = \mathbf{I}$, where $\mathbf{V}$ is a whitening matrix.

If we want to estimate a desired source signal, for this purpose we design a single processing unit described as

$$\tilde{y}_i(t) = \mathbf{w}_i^{\mathrm{T}}\tilde{\mathbf{x}}(t), \tag{4}$$

$$\tilde{y}_i(t - \tau_k) = \mathbf{w}_i^{\mathrm{T}}\tilde{\mathbf{x}}(t - \tau_k), \tag{5}$$

where $\mathbf{w}_i = (w_{i1}, \ldots, w_{in})^{\mathrm{T}}$ is the weight vector which corresponds to the estimate of one row of $(\mathbf{VA})^{-1}$, and $\tilde{y}_i(t)$ is the output signal which corresponds to the estimate of the source signal $s_i(t)$, and τ_k is a certain time delay.

We present the following constrained maximization problem based on several time delay nonlinear autocorrelation function of the desired source:

$$\begin{aligned} \max_{\|\mathbf{w}_i\|=1} \Psi_1(\mathbf{w}_i) &= \sum_{k=1}^{M} E\{G(\tilde{y}_i(t))G(\tilde{y}_i(t - \tau_k))\} \\ &= \sum_{k=1}^{M} E\{G(\mathbf{w}_i^{\mathrm{T}}\tilde{\mathbf{x}}(t))G(\mathbf{w}_i^{\mathrm{T}}\tilde{\mathbf{x}}(t - \tau_k))\}. \end{aligned} \tag{6}$$

G is a differentiable nonlinear function, which measures the nonlinear autocorrelation degree of the desired source. Examples of choices are $G(u) = u^2$ or $G(u) = \log\cosh(u)$.

Sometimes, we could use only one time lagged nonlinear autocorrelation to obtain the desired source signal:

$$\begin{aligned} \max_{\|\mathbf{w}_i\|=1} \Psi_2(\mathbf{w}_i) &= E\{G(\tilde{y}_i(t))G(\tilde{y}_i(t - \tau_k))\} \\ &= E\{G(\mathbf{w}_i^{\mathrm{T}}\tilde{\mathbf{x}}(t))G(\mathbf{w}_i^{\mathrm{T}}\tilde{\mathbf{x}}(t - \tau_k))\}. \end{aligned} \tag{7}$$

Thus, the constrained maximization problem (7) is a special case of the problem (6) when only one time delay can be used. In Section 3, we give some theoretical properties about the maximization problems (6) and (7).

2.2. Learning algorithms

Maximizing the objective function in (6), we derive a fixed-point blind source separation (BSS) algorithm. The gradient of $\Psi_1(\mathbf{w}_i)$ with respect to $\mathbf{w}_i$ can be obtained as

$$\frac{\partial \Psi_1(\mathbf{w}_i)}{\partial \mathbf{w}_i} = \sum_{k=1}^{M} E\{g(\tilde{y}_i(t))G(\tilde{y}_i(t - \tau_k))\tilde{\mathbf{x}}(t) + G(\tilde{y}_i(t))g(\tilde{y}_i(t - \tau_k))\tilde{\mathbf{x}}(t - \tau_k)\}. \tag{8}$$

The function g is the derivative of G. According to the Kuhn–Tucker conditions [17], we note that at a stable point of the optimization problem (6), the gradient of $\Psi_1(\mathbf{w}_i)$ at $\mathbf{w}_i$ must point in the direction of $\mathbf{w}_i$, thus we can optimize the objective function in (6) by a fixed-point algorithm [11]. This means that we have

$$\begin{aligned} \mathbf{w}_i &\leftarrow \sum_{k=1}^{M} E\{g(\tilde{y}_i(t))G(\tilde{y}_i(t - \tau_k))\tilde{\mathbf{x}}(t) + G(\tilde{y}_i(t))g(\tilde{y}_i(t - \tau_k))\tilde{\mathbf{x}}(t - \tau_k)\}, \\ \mathbf{w}_i &\leftarrow \mathbf{w}_i / \|\mathbf{w}_i\|. \end{aligned} \tag{9}$$