Int. J. Electron. Commun. (AEÜ) 69 (2015) 1265–1271

Contents lists available at ScienceDirect

International Journal of Electronics and Communications (AEÜ)

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

An efficient and robust algorithm for BSS by maximizing reference-based negentropy

Wei Zhao*, Yuehong Shen, Zhigang Yuan, Yimin Wei, Pengcheng Xu, Wei Jian

College of Communications Engineering, PLA University of Science and Technology, Houbiaoying Road No. 88, Qinhuai District, 210014 Nanjing, China

ARTICLE INFO

Article history:
Received 10 December 2014
Accepted 12 May 2015

Keywords:
Blind source separation
Cumulant
Reference-based contrast functions
Negentropy
FastICA
Kurtosis

ABSTRACT

A family of contrast criteria referred to as “referenced-based” has been recently proposed for blind source separation (BSS), which are essentially the cross-statistics or cross-cumulants between estimated outputs and reference signals. These contrast functions have an appealing feature in common: the corresponding optimization algorithms are quadratic with respect to the searched parameters. Inspired by this reference-based scheme, a similar contrast function is constructed by introducing the reference signals to negentropy, based on which a novel fast fixed-point (FastICA) algorithm is proposed in this paper. This new method is similar in spirit to the classical FastICA algorithm based on negentropy but differs in the fact that it is much more efficient in terms of computational speed than the latter, which is significantly striking with large number of samples. What is more, this new algorithm is more robust against unexpected outliers than those cumulant-based algorithms such as the FastICA algorithm based on kurtosis. The performance of this new method is validated through computer simulations.

1. Introduction

In recent decades, blind source separation (BSS) has found wide application in a variety of fields such as telecommunication [1], array processing [2], [3], biomedical signal detection [4], seismic exploration [5], and passive sonar [6], [7]. The task of BSS is to recover source signals using the observable signals. The noise-free instantaneous model of BSS can be described as follows:

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

where the mixing matrix $\mathbf{A} \in R^{M\times N}$ is unknown, $\mathbf{x}(t) \in R^M$ is the observed data vector and $\mathbf{s}(t) \in R^N$ is the unknown source vector. M and N are the number of observations and sources, respectively. When the sources are linearly and instantaneously mixed, BSS corresponds to independent component analysis (ICA) [8–10]. In this paper, we mainly study the efficient optimization problem of BSS in the ICA framework.

The BSS problem has found interesting solutions by maximizing non-Gaussian contrast criteria such as kurtosis and negentropy, which are the so-called contrast functions [11]. Based on these objective functions, a class of optimization algorithms has been proposed such as the gradient descent algorithms and FastICA algorithms [12]. Recently, a family of contrast criteria referred to as “referenced-based” has been proposed, which are essentially the cross-statistics or cross-cumulants between estimated outputs and reference signals [13–18]. These contrast functions have an appealing feature in common: the corresponding optimization algorithms are quadratic with respect to the searched parameters.

A maximization algorithm based on singular value decomposition (SVD) has been proposed in [13–15], and was shown to be significantly quicker than other maximization algorithms. However, the method often suffers from the need to have a good knowledge of the filter orders due to its sensitivity on the rank estimation [14]. The drawback of the SVD based method is well overcome when replaced by the gradient optimization method proposed in [16], in which the reference signals are fixed during the whole optimization process. Similarly, a relevant gradient algorithm with the reference signals updating after each one-dimensional optimization has been proposed in [17], which shows better performance. Based on the algorithms in [16] and [17], an improved method is proposed to adjust between performance and speed of it by introducing a new iterative updating parameter in [18], and simultaneously a detailed proof of the global convergence of the algorithm to a stationary point is performed.

Inspired from the reference-based schemes in [17] and [18], we construct a similar contrast function by introducing the reference signals into negentropy in [19], based on which a novel FastICA

* Corresponding author. Tel.: +86 15950592394.
E-mail address: gmajyie@126.com (W. Zhao).

http://dx.doi.org/10.1016/j.aeue.2015.05.008
1434-8411/

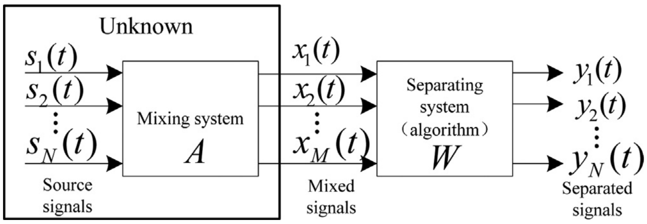

Fig. 1. Multi-input and multi-output system model.

algorithm is proposed. The main advantage of our new algorithm consists in two aspects:

(i) Because the reference signals are not directly involved in the iterative optimization process, the computational cost of our method is much lower than that of corresponding classical one, which is especially apparent when the number of samples is large. This means that our algorithm is much more efficient in terms of computational speed than the classical one in [19].
(ii) It is well accepted that the optimization algorithms based on higher-order cumulant such as kurtosis have some inherent drawbacks in practice, when its value has to be estimated from a measured sample. The main problem is that they can be very sensitive to outliers and their value may depend on only a few observations in the tails of the distribution, which may be erroneous or irrelevant observations. In other words, they are not robust against unexpected outliers. However, the widely used negentropy shows rather opposite properties to those of higher-order cumulant such as kurtosis, which is a well accepted concept. Therefore, the proposed algorithm in this paper is more robust than corresponding cumulant-based ones in [11], [12].

The remaining of this paper is organized as follows. Section 2 introduces the system model and assumptions. The reference-based negentropy and our proposed algorithm are presented in Section 3. Simulation results and analysis are shown in Section 4. Section 5 brings this paper to an end.

2. System model and reference signals

2.1. System model

The multi-input and multi-output (MIMO) system model considered in this paper is shown in Fig. 1

where the source signal vector and observation signal vector are denoted by $\mathbf{s}(\mathrm{t}) = [s_1(t), s_2(t), \ldots, s_N(t)]^T$ and $\mathbf{x}(\mathrm{t}) = [x_1(t), x_2(t), \ldots, x_M(t)]^T$ respectively. Note that we only consider the real-valued source signals in this paper, even though the signals can be real-valued and complex-valued. And the complex case will be considered in our latter work.

The mixing relationship is shown in (1), i.e., $\mathbf{x}(\mathrm{t}) = \mathbf{A}\mathbf{s}(\mathrm{t})$. Similarly, the separation operator is described in the form of

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

where the separating matrix $\mathbf{W}$ ($\in R^{M\times N}$) contains N column vector, i.e., $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_N]$. The estimation signal vector $\mathbf{y}(\mathrm{t}) = [y_1(t), y_2(t), \ldots, y_N(t)]^T$ is the approximate estimation of source signal vector $\mathbf{s}(\mathrm{t})$. Without loss of generality, we assume the number of sources and observations is equal in this paper, i.e., $M = N$.

For our system model in Fig. 1, we assume two assumptions on the sources: (1) the components of source signal vector are mutually statistically independent, i.e., $E\{s_i(t)s_j(t)\} = E\{s_i(t)\}E\{s_j(t)\}$, $i \neq j$; (2) the source signals have zero-mean and unit variance, i.e., $E\{s_i(t)\} = 0, \quad E\{s_i(t)^2\} = 1, \quad i = 1, 2, \ldots, N.$

2.2. Reference signals

As shown in [13–18], the reference signals are described in the form of

$$\mathbf{z}(\mathrm{t}) = \mathbf{V}^T\mathbf{x}(\mathrm{t}) \tag{3}$$

where $\mathbf{z}(\mathrm{t})$ and $\mathbf{V}$ have the similar form as $\mathbf{y}(\mathrm{t})$ and $\mathbf{W}$. They are denoted by $\mathbf{z}(\mathrm{t}) = [z_1(t), z_2(t), \ldots, z_N(t)]^T$ and $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N]$, respectively. As mentioned above, the reference signals are not directly involved in iterative optimization process. Actually, they are artificially introduced in the optimization algorithms for the purpose of facilitating the maximization of non-Gaussian contrast criteria [18]. More precisely, contrast criteria are reformulated by combining estimation signals and reference signals together, which results in the so-called reference-based contrast functions. Note that the reference signals have direct influence on the reference-based contrast function and their values do impact the optimization results, especially the initialization value.

In [16], the reference signals are initialized arbitrarily and kept the same during whole optimization process. In [17], the reference signals are indirectly involved in the iterative optimization process. In other words, the reference signals update following the estimated signals. More precisely, $\mathbf{V}$ updates following $\mathbf{W}$ in each loop iteration step. Then the separation quality of the algorithm in [17] is better than that in [16]. In [20] and [21], we have done some corresponding work to investigate the impact of reference signals, which is similar to [16] and [17]. Inspired by [17], we consider the case that the reference signals update circularly in this paper. Because our precious work is mainly based on kurtosis, our new algorithm in this paper will be more robust than those in [20] and [21]. Since no confusion is possible and for simplicity, in the following sections, we drop the time index and these signal vectors are denoted respectively by $\mathbf{s}$, $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$.

3. Contrast functions and algorithm

3.1. Reference-based negentropy

As mentioned above, the negentropy, or equivalently, differential entropy, shows better robustness as the non-Gaussian measure criterion, which mainly lies in that it is well justified by statistical theory. To use negentropy as the cost function requires knowledge of the pdf that may not be known, or online estimations of the pdf can be computationally complicated and difficult. To overcome these situations, nonlinear functions are used to approximate the negentropy and hence generate the higher order statistics implicitly [24]. In this paper, we use the approximate negentropy proposed in [19], which is denoted by

$$J(\mathbf{w}) = E\{(G(\mathbf{w}^T\mathbf{x}))^2\} \tag{4}$$

where $G : R \mapsto R$. It was shown through stability analysis in [19] that practically any non-quadratic even function can be used to construct a cost function for ICA through non-Gaussianity maximization. Therefore, we choose two nonlinear functions that are used and have been verified in [19]:

$$G_1 = a\sinh(x) = \ln\left[x + \sqrt{x^2 + 1}\right] \tag{5}$$

$$G_2 = \cosh(x) = \frac{e^x + e^{-x}}{2} \tag{6}$$

$$G_3 = x^{1.25} \tag{7}$$

It was shown clearly in [19] that these nonlinear functions match the pdf of sources properly, i.e., $p_G(y) = e^{-G(y)^2}$, in which p_G is the pdf associated with the nonlinearity G. More description in detail can be found in [19].