Signal Processing 105 (2014) 175–184

Contents lists available at ScienceDirect

Signal Processing

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

Independent vector analysis with a generalized multivariate Gaussian source prior for frequency domain blind source separation

Yanfeng Liang*, Jack Harris, Syed Mohsen Naqvi, Gaojie Chen, Jonathon A. Chambers

Advanced Signal Processing Group, School of Electronic, Electrical and System Engineering, Loughborough University, Leicestershire LE11 3TU, UK

ARTICLE INFO

Article history:
Received 16 May 2013
Received in revised form
26 February 2014
Accepted 19 May 2014
Available online 28 May 2014

Keywords:
Independent vector analysis
Source prior
Multivariate generalized Gaussian distribution
Blind source separation

ABSTRACT

Independent vector analysis (IVA) is designed to retain the dependency within individual source vectors, while removing the dependency between different source vectors. It can theoretically avoid the permutation problem inherent to independent component analysis (ICA). The dependency in each source vector is retained by adopting a multivariate source prior instead of a univariate source prior. In this paper, a multivariate generalized Gaussian distribution is adopted as the source prior which can exploit frequency domain energy correlation within each source vector. As such, it can utilize more information describing the dependency structure and provide improved source separation performance. This proposed source prior is suitable for the whole family of IVA algorithms and found to be more robust in applications where non-stationary signals are separated than the one preferred by Lee. Experimental results on real speech signals confirm the advantage of adopting the proposed source prior on three types of IVA algorithm.

1. Introduction and relation to prior work

Blind source separation (BSS) has been widely researched over recent decades [1], as selecting specific signals from observed mixtures has potentially wide applications, such as in speech processing, biomedical signal processing, image processing and communication systems [2]. The classical BSS problem is the machine cocktail party problem, which was introduced by Cherry [3,4], wherein a target speech signal must be extracted from microphone measurements acquired in a room environment.

Independent component analysis (ICA) is a statistical signal processing based solution for the blind source separation problem. It can successfully address the instantaneous BSS problem, for example, by exploiting the non-Gaussianity of the signals [5]. However, in a real room environment, the problem becomes convolutive blind source separation (CBSS) due to reverberation. The length of the room impulse response is typically on the order of thousands of samples. Thus, time domain methods are generally not suitable for the CBSS problem due to the computational complexity [6]. In order to reduce the computational cost, frequency domain methods have been proposed [7]. The convolution operation in the time domain becomes multiplication in the frequency domain, so the computational cost reduces significantly [6]. When the mixtures are transferred into the frequency domain by

* Corresponding author.
E-mail address: Y.Liang2@lboro.ac.uk (Y. Liang).

http://dx.doi.org/10.1016/j.sigpro.2014.05.022
0165-1684/

using the discrete Fourier transform (DFT), the ICA method can be used in each frequency bin to separate the mixtures. However, the permutation ambiguity inherent to ICA becomes more pertinent due to the potential misalignments of the separated sources at different frequency bins. In this case, should the separated results be transformed back to the time domain, the separation performance will be poor. Therefore, different methods to mitigate the permutation problem have been proposed [6]. Most of which use extra information such as source geometry [8] or prior knowledge of the source structure [9,10]. However, pre- or post-processing is needed for all of these methods, which generally introduces delay and additional complexity.

Recently, the independent vector analysis (IVA) method has been proposed [11], which can theoretically avoid the permutation problem by retaining the dependency in each individual source vector while minimizing the dependency between different source vectors during the learning process [12,13]. Thus, it is a natural way to solve the permutation problem during the learning process without any pre- or post-processing. The main difference between ICA algorithms and IVA algorithms is the nonlinear score function. For ICA algorithms, the nonlinear score function is just a univariate function which only uses the data in each frequency bin to update the unmixing matrix. However, for IVA algorithms, a multivariate function is adopted as the score function. The multivariate function uses the data in all the frequency bins, therefore it can exploit the inter-frequency dependencies to mitigate the permutation problem.

There are three main types of IVA algorithm. The first one is the natural gradient IVA (NG-IVA) algorithm [12], which adopts the Kullback–Leibler divergence between the joint probability density function and the product of marginal probability density functions of the individual source vectors as the cost function. The natural gradient method is used to minimize the cost function. The second algorithm is the fast fixed-point IVA (FastIVA) algorithm [14], which uses Newton's method to update the unmixing matrix. The third one is called the auxiliary function based IVA (AuxIVA) algorithm [15]. By using the auxiliary function technique, the AuxIVA can converge quickly without introducing tuning parameters and can guarantee that the objective function decreases monotonically. There are also several other IVA algorithms which are based on these three basic IVA algorithms. The adaptive step size IVA algorithm which is based on the NG-IVA algorithm and can automatically select the step size to achieve a faster convergence in terms of iteration number [16]. The audio–video based IVA method which combines video information with FastIVA to obtain a faster and better separation performance in noisy and reverberant room environments [17]. Another advantage of the audio-video based IVA method is using smart initialization to overcome the problems in convergence due to the nature of the cost function, such as the presence of local minima. There are also IVA methods which exploit the source activity and dynamic structure to achieve improved separation performance [18,19].

The core idea of IVA algorithms applied to frequency domain BSS is preserving inter-frequency dependencies for individual sources. The nature of the score function used in the algorithm derivation is crucial in this process [12]. The nonlinear score function is derived from the source prior, therefore an appropriate source prior is needed. For the original IVA algorithms, a spherically symmetric distribution is adopted as the source prior, which implies that the dependencies between different frequency bins are all the same. However, the dependencies between frequency bins should be variable. In order to describe the dependency structure better, a chain-like overlapped source prior has been proposed [20]. More recently, a harmonic structure dependency model has been proposed [21]. Another possible source prior is the Gaussian mixture model, whose advantage is that it enables the IVA algorithms to separate a wider class of signals [22,23]. However, for all of these source priors, the covariance matrix of each source vector is an identity matrix because the Fourier basis is an orthogonal basis. This implies that there is no correlation between different frequency bins. Moreover, the higher order correlations between the elements of the source vectors are ignored when separating the mixtures with IVA algorithms. Recently, an IVA algorithm based upon a multivariate Gaussian source prior has been proposed to introduce the second order correlations in the time domain [24]. However, it is used in applications which have large second order correlations such as in functional magnetic resonance imaging studies. For the frequency domain IVA algorithms, other correlation information should be exploited.

In this paper, a particular multivariate generalized Gaussian distribution is adopted as the source prior. This proposed source prior has heavier tails compared with the original multivariate Laplace distribution. It can preserve the dependency across different frequency bins in a similar way as when the original multivariate Laplace distribution is used to derive an IVA algorithm. Moreover, the nonlinear score functions which are derived based on the proposed source prior additionally contain an expression in the form of frequency domain energy correlation between the elements of each source vector, thus they contain more information describing the dependency structure which can thereby better preserve the inter-frequency dependency to achieve an improved separation performance, as suggested by Hyvärinen [25].

The structure of this paper is as follows. In Section 2, the original independent vector analysis for frequency domain BSS is introduced. In Section 3, the particular multivariate generalized Gaussian distribution is proposed and analysed. The NG-IVA with the proposed source prior is discussed in Section 4. Sections 5 and 6 introduce respectively the FastIVA and the AuxIVA with the proposed source prior. The experimental results are shown in Section 7, and finally conclusions are drawn in Section 8.

2. Independent vector analysis

In this paper, we focus on the separation of statistically non-stationary speech signals in a real room environment. Due to reverberation, the problem becomes the CBSS problem. Time domain methods are not appropriate because of