Constrain Component Extraction Techniques

Yongjian Zhao

Information Engineering Institute, Shandong University (Weihai), China

jian123cn@sdu.edu.cn

Keywords: Extraction, Mixture, Component, Signal, Function

Abstract. Component extraction is wide used in biomedical signal processing. Traditional independent component analysis (ICA) techniques aim to estimate all underling components from their mixtures simultaneously. To overcome the drawbacks, a constrained separation approach is presented for extraction of one or a few desired source signals. Its performance is demonstrated by computer simulations on biomedical signals.

Introduction

Independent component analysis (ICA) is a statistical and computational technique for revealing hidden factors that underlie sets of measurements [1,2]. Generally speaking, ICA defines the following model for the observed multivariate data, which is typically given as a large database of samples:

$$x = As , \tag{1}$$

where $s = (s_1, \cdots, s_m)^T$ denotes the original independent sources, $x = (x_1, \cdots, x_n)^T$ denotes mixtures of original sources and A is an unknown mixing matrix. The assumptions of an ICA problem are that each component of S has 0 as its mean value, mutually independent and drawn from a probability distribution which is not a normal distribution except for at most one component [3]. For technical simplicity, m is also supposed to be smaller than or equal to n. In fact, this assumption is necessary for the existence of a linear solution. Without knowing the source signals and the mixing matrix, ICA technique aims to recover independent components from their observed mixtures using some statistical properties of original sources [4]. However, traditional ICA technique only copes with the simultaneous recovery of source signals from their mixtures [5]. In most cases especially in biomedical signal processing, the number of original sources is often large while only one or a few are desired.

In many ICA applications such as biomedical signal processing, there are more or less additional requirements along with prior information [6]. In other words, someone can know specific prior information about the desired source signal in advance. In this paper, we introduce such prior information to a negentropy based contrast function so as to construct a constrained independent component analysis (cICA) problem. Thus the ill-posed ICA is converted to a better-posed problem, facilitating more applications. Solving the cICA problem, a flexible cICA algorithm is developed for extraction of one or a few desired source signals exclusively. Computer simulations on biomedical signals demonstrate its validity and usefulness.

Constrained ICA Techniques

To formulate the ICA problem, one must construct a reliable contrast function, which is greatly dependent on the parameters of the specified neural network model in (1). Maximization/Minimization of such a contrast function should make the output of the model satisfy the statistical conditions of stochastic independence and other inherent properties, depending on the specific problem. For instance, the BCBSE algorithm [2] proposed by Barros and Cichocki is performed by making the output independent and uncorrelated for different time delays.

What we do for cICA problem is to find a separating or demixing vector w so that $y = w^T x$,

where y is equal to the desired source signal.

For extracting one source signal, a classical negentropy based contrast function [4,5] is

$$J(y) = \rho [E\{G(y)\} - E\{G(v)\}]^2, \tag{2}$$

where v is a Gaussian variable with zero mean and unity variance, ρ is a positive constant, and $G(*)$ can be any non-quadratic function, such as

$$G_1(y) = \log coash(a_1 y) / a_1, \tag{3}$$

$$G_2(y) = \exp(-\frac{1}{2} a_2 y^2) / a_2, \tag{4}$$

$$G_3(y) = y^4 / 4, \tag{5}$$

where $1 \le a_1 \le 2$ and $a_2 \approx 1$. G_1 is a good general purpose function; G_2 may be better when the independent components are highly super-Gaussian, or when robustness is very important; G_3 is justified on statistical grounds only for estimating sub-Gaussian signals when there are no outliers.

The gradient of $J(y)$ with respect to w can be deduced as

$$\nabla_w J(y) \propto 2\gamma E\{xg(w^T x)\} \tag{6}$$

where $\gamma = E\{G(w^T x)\} - E\{G(v)\}$. Thus one can deduce the famous FastICA algorithm [4,5]:

$$w \leftarrow 2\gamma E\{xg(w^T x)\}. \tag{7}$$

As a typical one-unit ICA algorithm, FastICA has many desirable properties [4-6]. However, maximization of (2) will give any of source signals. When one is only interested in a specific source signal, FastICA has to estimate source signals one by one, thus introducing a large amount of additional computational burden. FastICA is of little use while only one source signal is desired.

In most cases of biomedical signal processing, one may know specific prior information about a desired source signal in advance [5,6]. In this paper, we introduce such reference to modify the one-unit algorithm (6), thus ensuring the output is necessarily the desired one.

The closeness measure between an estimated output y and its corresponding reference r is denoted by $\xi(y,r)$. Further suppose that w_m or w^* is the demixing vector for the desired signal and w_i $(i = 1, \cdots, m-1)$ are the demixing vectors for undesired components. Note that the desired source signal is the only one closest to the reference signal r, satisfying the following inequality relationship:

$$\varepsilon(w^{*T}\tilde{x}, r) \le \varepsilon(w_1^T\tilde{x}, r) \le \cdots \le \varepsilon(w_{m-1}^T\tilde{x}, r). \tag{8}$$

In fact, the extracted signal $w^{*T}\tilde{x}$ is equal to y^* if and only if the following inequality relationship is satisfied:

$$g(w) = \varepsilon(w^T\tilde{x}, r) - \xi \le 0, \tag{9}$$

where $\xi \in [\varepsilon(w^{*T}\tilde{x}, r), \varepsilon(w_1^T\tilde{x}, r))$ is a threshold parameter.

Through introducing (9) into the negentropy contrast function (2) as a constrained term, we can deduce the following constrained optimization problem:

$$\begin{cases} \max & J(y) = \rho[E\{G(y)\} - E\{G(v)\}]^2 \\ s.t. & g(w) \le 0, h(w) = E(y^2) - 1 = 0 \end{cases}, \tag{10}$$

where the equality constraint aims at bounding $J(y)$ and w.

The inequality constraints of (10) can be transformed into equality ones by introducing a vector of slack variables, so the optimal solution can be obtained by the Lagrange multipliers method [1,6]. Subsequently, through finding the maximum of the augmented Lagrange function, we can deduce the following learning rule for w:

$$w_{t+1} = w_t - \eta R_z^{-1} \Psi_1 / \psi_2, \tag{11}$$

$$\psi_1 = \bar{\rho} E\{z G_y'(y)\} - 0.5\mu E\{z g_y'(y)\} - \lambda E\{zy\}, \tag{12}$$

$$\psi_2 = \bar{\rho}E\{zG''_{y^2}(y)\} - 0.5\mu E\{zg''_{y^2}(y)\} - \lambda, \tag{13}$$

where t is the iteration count, $R_z = E\{zz^T\}$, $\bar{\rho} = \rho \cdot sign(EG(y) - EG(v))$, $G'_y(y)$ and $g'_y(y)$ are the first derivatives of $G(y)$ and $g(y)$ with respect to y, and $G''_{y^2}(y)$ and $g''_{y^2}(y)$ are the corresponding second derivatives.

Simulation Results

To confirm validity and usefulness of the proposed algorithm in this paper, we performed computer simulations on real ECG data. The data set shown in Fig.1 was distributed by De Moor [7], which was a well-known electrocardiogram measured from a pregnant woman.
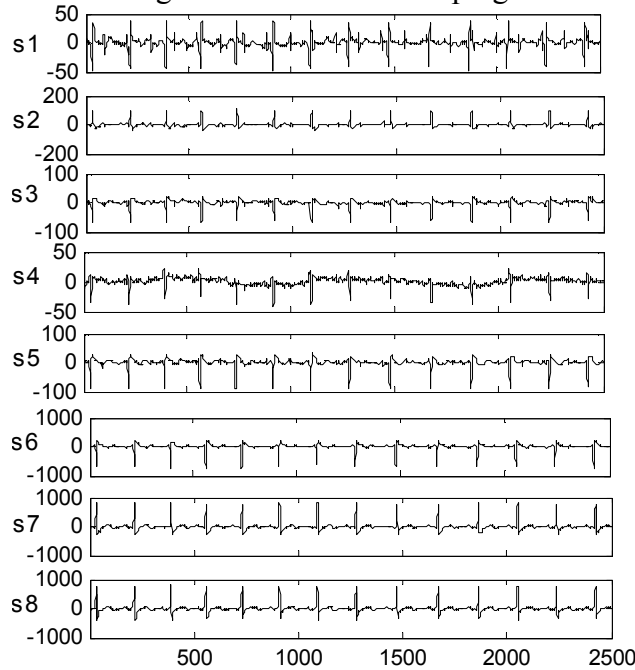

Fig.1. The ECG recording obtained from a pregnant woman

These measurements were recorded over 10s and sampled at 250 Hz with 8 electrodes located at abdomen and thorax of a pregnant woman. Signal S1-S5 were the recordings by five electrodes placed on the woman's abdomen. Therefore the FECG, respiratory motion artifacts as well as the maternal ECG (MECG) were visible in these recordings. Signal S6-S8 were the recordings by three electrodes placed on the woman's thorax. In these thoracic measurements, the FECG was invisible because of the distance between the fetus and the chest leads. Generally speaking, one can see the heart beating of both the mother (stronger and slower) and the fetus (weaker and faster).

In contrast to the conventional diagnostic techniques merely on the basis of the fetal cardiac rate estimation [3,4], the non-invasive extraction of the FECG from the maternal cutaneous potential recordings offers more detailed information about the fetus's heart. It must be mentioned that the measured FECG is always corrupted by various noise, such as the maternal electrocardiogram (MECG) with extremely high amplitude, the mother's respiration, and the thermal noise due to electronic equipments [5]. Fortunately, these components are statistically independent. Therefore, one can use ICA technique to extract the desired FECG. Barros and Cichocki [2] formally defined the prerequisites to extract the desired FECG using only prior information about its autocorrelation function. This algorithm needs specific prior information about the desired source signal in advance, namely, a suitable time delay at which the autocorrelation of the desired source signal has a peak whereas those of other source signals have low values. Since the desired FECG satisfies the constrained condition, the BCBSE algorithm can be used for FECG extraction. However, the extracted FECG often contains noise contributions, such as the mother's breathing artifact [4,5].

798

The reference signal need not be a perfect match to the desired signal (if it were, one would argue that there would be little point in introducing the constrained ICA method at all). Although the BCBSE algorithm [2] cannot extract an elaborate FECG, the extracted signal is a rough template of the desired FECG. So we may consider the output of the BCBSE algorithm as the reference signal. The BCBSE algorithm requires prior information about the desired signal's autocorrelation information and needs to estimate a time delay. The true time delay for extraction of the FECG is 112 [2,3]. Here we set the optimal time delay as $\tau^* = 112$. The extraction result is shown as BCFECG in Fig.2. One can see that the extracted BCFECG is still contaminated by various noise even if the true time delay is adopted.

Fortunately, the morphology of the desired FECG can be obtained from the BCFECG. In other words, the BCFECG can be utilized as a reference signal. We ran the algorithm proposed in this paper using the BCFECG as a reference signal for extraction of the desired FECG. The extraction results were described as FECG1 in Fig.2. To make comparison, we also ran other typical algorithms [3,4] for FECG extraction. ICA-R and FECG2 shown in Fig.2 are the corresponding results of algorithms in [3,4] and [5].
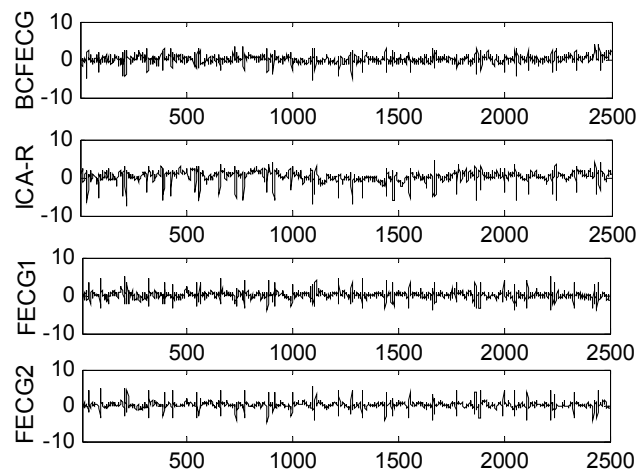

Fig.2. The extracted FECGs by different algorithms

From these simulation results, one can see that the extracted signal by the ICA-R algorithm [3,4] is still mixed with noise. The algorithm in [5] can obtain good results. However, its computation load is still large. The desired FECG can be extracted efficiently with the proposed method in this paper, and that the noise, which is present in the reference signal, is almost invisible here. Generally speaking, the algorithm proposed in this paper has better properties in contrast to other ones.

References

[1] A. Hyvärinen, J. Karhunen, E. Oja, Independent Component Analysis, New York: Wiley, 2001.
[2] A. K. Barros, A. Cichocki, Extraction of specific signals with temporal structure, Neural Computation, vol.13, 2001, pp.1995-2003.
[3] W. Lu, J. C. Rajapakse, ICA with reference, Neurocomputing, vol. 69, 2006, pp.2244-2257.
[4] W. Lu, J. C. Rajapakse, Approach and applications of constrained ICA, IEEE Transactions on Neural Networks, vol.16, 2005, pp.203-212.
[5] Z. L. Zhang, Morphologically constrained ICA for extracting weak temporally correlated signals, Neurocomputing, vol.71, 2008, pp. 1669-1679.
[6] D. S. Huang, J. X. Mi, A new constrained independent component analysis method, IEEE Transactions on Neural Networks, vol.18, 2007, pp.1532-1535.
[7] Http://www.cis.hut.fi/projects/ica