

Research Article

A Fetal Electrocardiogram Signal Extraction Algorithm Based on Fast One-Unit Independent Component Analysis with Reference

Yanfei Jia^{1,2} and Xiaodong Yang^{1,3}

¹College of Information and Communication Engineering, Harbin Engineering University, Heilongjiang 150001, China

²College of Electrical and Information Engineering, Beihua University, Jilin 132012, China

³Collaborative Research Center, Meisei University, Tokyo 1918506, Japan

Correspondence should be addressed to Yanfei Jia; jia.yanfei@163.com

Received 10 July 2016; Accepted 8 August 2016

Academic Editor: Po-Hsiang Tsui

Copyright © 2016 Y. Jia and X. Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fetal electrocardiogram (FECG) extraction is very important procedure for fetal health assessment. In this article, we propose a fast one-unit independent component analysis with reference (ICA-R) that is suitable to extract the FECG. Most previous ICA-R algorithms only focused on how to optimize the cost function of the ICA-R and paid little attention to the improvement of cost function. They did not fully take advantage of the prior information about the desired signal to improve the ICA-R. In this paper, we first use the kurtosis information of the desired FECG signal to simplify the non-Gaussian measurement function and then construct a new cost function by directly using a nonquadratic function of the extracted signal to measure its non-Gaussianity. The new cost function does not involve the computation of the difference between the function of the Gaussian random vector and that of the extracted signal, which is time consuming. Centering and whitening are also used to preprocess the observed signal to further reduce the computation complexity. While the proposed method has the same error performance as other improved one-unit ICA-R methods, it actually has lower computation complexity than those other methods. Simulations are performed separately on artificial and real-world electrocardiogram signals.

1. Introduction

The fetal electrocardiogram (FECG) contains much important information about the health and possible diseases of the fetus, which reflects the complete view of the heart activities. Additionally, it is more sensitive than color Doppler ultrasound in the case of fetal acidosis and anoxia. With the analysis of the FECG, the doctor can discover fetal hypoxemia, umbilical abnormality, and other fetal abnormality situations timely and take early effective actions to ensure the health of the fetus. However, the FECG signal is considerably weaker than the maternal electrocardiogram (MECG) and is often embedded in the noise, MECG, baseline wandering, power line interference, and so forth. Accordingly, it is quite difficult to extract the FECG signal and further reduce the diagnostic accuracy.

A number of methods have been reported for extracting the FECG signal, like filtering method [1], singular value

decomposition [2], wavelet transform [3], independent component analysis (ICA) [4], and others [5]. Every method has its advantages and disadvantages [5, 6]. In this paper, we only consider the method based on ICA. The FECG extraction based on the ICA methods considers the extracting FECG signal as a blind source separation problem. They consider the unknown FECG signal, MECG, and other interference signals as source signals and the measured signals from the maternal abdomen and chest as mixed signals (measured signals). It actually can separate all source signals, including the FECG signal and MECG signal. Compared with the other methods, it has a simple structure and has been proved to be efficient in extracting FECG in many studies [4–9]. However, for extracting the FECG, the ICA method must first separate all the source signals from the measured signals and then artificially selects the FECG signal. When the number of measured signals is very large, the extraction of the FECG signal will consume plenty of time.

To overcome this problem, the independent component analysis with reference (ICA-R) is applied to extract the FECG [10–14], which separates the desired FECG signal by using some prior information about the desired FECG without separating all the source signals. Therefore, this FECG extraction method based on ICA-R is more efficient than those based on ICA. ICA-R was proposed and implemented by Lu and Rajapakse [15–17] by incorporating the prior information of the desired signal into the contrast function of the ICA [18]. To improve the ICA-R performance, some authors proposed improved ICA-R methods. For example, to reduce the computational complexity of ICA-R, Lin et al. [11] proposed a fast one-unit ICA-R by using prewhitening to deal with the observed signal. Compared with the original ICA-R, this method has shorter running time and the same error. Also, Huang and Mi [19] studied the gradient of inequality constraint function and corrected the gradient. This method has a higher success rate for extracting the desired signal. In addition, Li et al. [14] first divided the cost function of ICA-R into two parts that are the negentropy function and the closeness measure function and then directly used the first-order derivative of the negentropy contrast function to update the separating vector. This method has faster convergence rate and higher success rate to extract the desired signal than the previously reported method [19]. Moreover, Sun and Shang [20] studied the stability of ICA-R and proposed an initialization method of separating the matrix. This method has a higher accuracy and stability than the original ICA-R. Kavuri et al. [21] proposed a one-unit ICA-R by using evolutionary algorithm to optimize the cost function to find a global optimal solution. Although this method has a smaller error than the original ICA-R, it does consume a considerable amount of time. Additionally, Zhang et al. [22] proposed the ICA-R based on kurtosis and analyzed how to choose the reference signal. Wang [23] used the method that has been used in the fast ICA to propose a fixed-point ICA method. Rodriguez et al. [24] proposed a multiunit ICA-R method based on nonorthogonal decoupling of separated matrix. This method cannot be extended to one-unit ICA-R. Mi [25] proposed a strategy to detect future misconvergence to improve the probability of extracting the desired signal. Chen et al. [26] proposed an ICA-R that can be used in single channel by discrete wavelet transform.

The one-unit ICA-R is a special case of the multiunit ICA-R. It only separates one desired source signal every time. The only difference between the one-unit ICA-R and the multiunit ICA-R is that the former only extracts one desired signal, and the latter is based on the former and uses the orthogonalization method to extract multiple desired signals. If we use the deflationary orthogonalization method or symmetric orthogonalization method, which are widely used in ICA [18] to extract the signal in the one-unit ICA-R, the one-unit ICA-R can be easily extended to multiunit ICA-R. In the FECG extraction, we only want to extract the FECG with just one signal; thus, the one-unit ICA-R is enough to extract FECG.

Although the one-unit ICA-R and multiunit ICA-R have a little difference, they share the same cost function. Additionally, they both use the same non-Gaussian measurement

function that is used in the classical ICA algorithm to measure the non-Gaussian component of the extracted signal. The negentropy approximation is needed to compute the difference between the function of the extracted signal and the function of a Gaussian vector that has the same mean and variance as the extracted signal. This will consume plenty of time. The classical ICA algorithm has no information about the source signals, except that the non-Gaussian measurements and statistics are independent. However, the ICA-R has prior information about the desired signal, but the previous methods did not use such information to reduce the computational complexity of the negentropy approximation. The aim of this paper is to reduce the computational complexity for the one-unit ICA-R by taking advantage of the prior information of the desired ECG to simplify the negentropy approximation.

This paper is organized as follows. Section 2 reviews the previously reported algorithms including the traditional ICA and one-unit ICA-R. Section 3 analyzes in detail the non-Gaussian measurement function that is used in the one-unit ICA-R cost function. In Section 4, we simplify the non-Gaussian measurement function based on the analysis in Section 3 and, then, derive an improved one-unit ICA-R with lower computation complexity. In Section 5, we test Mi's method [25] and our proposed method on artificial ECG data and real-world ECG data to compare the performance of each method. Some conclusions are included in Section 6.

2. Previous Algorithms

2.1. Traditional ICA. The traditional ICA is a signal processing method for estimating the source signals from observed signals that are a mixture of source signals. The model of linear ICA is as follows:

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

where $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$ is the observed signal, \mathbf{A} is the mixing matrix with the size $(n \times m)$, and $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_m(t)]^T$ is the source signal. The aim of the ICA is to estimate the source signal $\mathbf{s}(t)$ from the observed signal $\mathbf{x}(t)$ by computing the separating matrix $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m]^T$ with size $(m \times n)$. The estimated source signal can be expressed as follows:

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

where $\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_m(t)]^T$ is the estimated source signal, which is the same as $\mathbf{s}(t)$ under ideal condition. However, under real condition, the shape of the waveforms of $\mathbf{y}(t)$ is close to $\mathbf{s}(t)$, except for the amplitude and sequence of the waveforms. The source signals $\mathbf{s}(t)$ are considered as non-Gaussian and mutually statistically independent. Accordingly, it has strong non-Gaussian properties. The mixing signals $\mathbf{x}(t)$ have strong Gaussian characteristics according to the central limit theorem. When the estimated source signal is close to the source signal $\mathbf{s}(t)$, it will have strong non-Gaussian properties. Thus, the ICA uses a non-Gaussian

measurement function of the estimated source signals as a cost function [18]:

$$J(y) \propto \rho [E\{F(y)\} - E\{F(v)\}]^2, \quad (3)$$

where ρ is a positive constant, y is the estimated signal (extracted signal), F is the nonquadratic function, and v is a Gaussian variable with the same mean and variance as y . By maximizing the cost function, we can obtain the separating matrix and recover the source signal.

2.2. One-Unit ICA-R. The ICA can separate all the source signals from the observed signals. However, if we only want to extract some or one of the desired source signals, we must use ICA to separate all source signals and further select the desired source signal from the estimated signals. When the dimension of observed signals is very large, especially for some biomedical signals, it will take a long period of time to extract the desired signal. The one-unit ICA-R was proposed by Lu and Rajapakse [15] to reduce computational time and avoid additional operations to select the desired source signal from the estimated signals.

The one-unit ICA-R method combines prior information of the desired source signal into the ICA cost function (3) and constructs a new cost function. The cost function for the one-unit ICA-R is as follows [15]:

$$\begin{aligned} \max \quad & J(y) \approx \rho [E\{F(y)\} - E\{F(v)\}]^2 \\ \text{s.t.} \quad & g(y) \leq 0 \\ & h(y) = E(y^2) - 1 = 0, \end{aligned} \quad (4)$$

where $y = \mathbf{w}^T \mathbf{x}$, \mathbf{w} is the separating vector, \mathbf{x} is the observed signals, y is the extracted signal, ρ is a positive constant, and v is a Gaussian vector with the same mean and variance as y . The inequality constraint term is $g(y) = \varepsilon(y, r) - \xi$, where r is the reference signal constructed by the prior information of desired signal, ε is used to measure the closeness between the estimated signal y and the reference signal r , and ξ is a threshold parameter to control the closeness level. A common and simple measure of closeness is the mean squared error (MSE) $\varepsilon(y, r) = E\{(y - r)^2\}$. $h(y)$ is used to restrict the estimated signal y to unit variance.

The above inequality constraints can be transformed into equality constraints by introducing the slack factor z . Thus, the cost function of the one-unit ICA with reference to equality constraint can be expressed as follows:

$$\begin{aligned} \max \quad & J(y) \approx \rho [E\{F(y)\} - E\{F(v)\}]^2 \\ \text{s.t.} \quad & g(y) + z^2 = 0 \\ & h(y) = E(y^2) - 1 = 0. \end{aligned} \quad (5)$$

The augmented Lagrange multiplier method is used to optimize the above cost function with equality constraint

and derive the augmented Lagrange function that can be expressed as follows:

$$\begin{aligned} L = J(y) - \frac{1}{2\gamma} \max((\gamma g(y) + \mu), 0)^2 + \frac{\mu^2}{2\gamma} - \lambda h(y) \\ - 0.5\gamma \|h(y)\|^2, \end{aligned} \quad (6)$$

where μ and λ are the Lagrange multipliers for the constrains $g(y)$ and $h(y)$, respectively, and γ is a scalar penalty parameter. A Newton-like learning algorithm is applied to optimize the above equation and get the updated separating vector. The updated formula is as follows [15]:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \eta \mathbf{R}^{-1} \frac{l}{\delta}, \quad (7)$$

where k is the current iterative number, η is the fixed learning rate, \mathbf{R} is the covariance matrix of the observed signals, and l and δ are the first and second derivatives of l , respectively, and can be expressed as

$$\begin{aligned} l = \widehat{\rho} E\{\mathbf{x}f(y)\} - 0.5\mu E\{\mathbf{x}g'(y)\} - \lambda E\{\mathbf{x}y\} \\ \delta = \widehat{\rho} E[f'(y)] - 0.5\mu E[g''(y)] - \lambda, \end{aligned} \quad (8)$$

where $\widehat{\rho} = 2\rho[E\{F(y)\} - E\{F(v)\}]$, $f(y)$ and $f''(y)$ are the first and second derivatives of $F(y)$ with respect to y , respectively, and $g'(y)$ and $g''(y)$ are the first and the second derivatives of $g(y)$ with respect to y , respectively.

The optimum multipliers μ and λ are updated by the following equations:

$$\mu_{k+1} = \max\{0, \mu_k + \gamma g(y)\} \quad (9)$$

$$\lambda_{k+1} = \lambda_k + \gamma h(y). \quad (10)$$

The same author Lu and Rajapakse proposed [17] another cost function for the one-unit ICA-R, as follows:

$$\begin{aligned} \min \quad & -J(y) \approx -\rho [E\{F(y)\} - E\{F(v)\}]^2 \\ \text{s.t.} \quad & g(y) \leq 0 \\ & h(y) = \{E(y^2) - 1\}^2 = 0. \end{aligned} \quad (11)$$

The learning rule of the separating vector is

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \eta \mathbf{R}^{-1} \frac{l}{\delta}, \quad (12)$$

where

$$\begin{aligned} l = \widehat{\rho} E\{\mathbf{x}f(y)\} - \mu E\{\mathbf{x}g'(y)\} \\ - 4\lambda (E\{y^2\} - 1) E\{\mathbf{x}y\} \\ \delta = \widehat{\rho} E[f'(y)] - \mu E[g''(y)] - 8\lambda. \end{aligned} \quad (13)$$

The symbols of the parameters in (12) have the same meaning as those in (7). Many studies about the ICA-R method are based on the cost function (11) [19, 20, 25].

3. Analysis of the Non-Gaussian Measurement Function in the One-Unit ICA-R Method

The non-Gaussian measurement function (5) that is also the cost function of the fast ICA can also be expressed as

$$J(y) = \rho |E\{F(y)\} - E\{F(v)\}|, \quad (14)$$

where $|\cdot|$ is the absolute value of (\cdot) . Maximizing (14) equals maximizing (3). In ICA-R, v is a Gaussian variable having zero mean and unit variance [17]. It actually has nothing to do with the separating vector \mathbf{w} and can be seen as a constant for \mathbf{w} . Accordingly, maximizing (14) equals maximizing or minimizing $E\{F(y)\}$. We denote it by $H(y)$; that is,

$$H(y) = E\{F(y)\}, \quad (15)$$

where $y = \mathbf{w}^T \mathbf{x} = \mathbf{w}^T \mathbf{A} \mathbf{s} = \mathbf{b}^T \mathbf{s}$, \mathbf{s} is the source signals, \mathbf{A} is the mixing matrix, and \mathbf{w} is the separating vector. In the one-unit ICA-R, we only extract one desired signal. We consider that the successfully extracted signal y equals s_i ; hence, the stability point of \mathbf{b} is $\mathbf{b}^* = (0_0, \dots, 0_{i-1}, 1, 0_{i+1}, \dots)^T$.

At the stability point, the first-order derivative of $H(\mathbf{b}^T \mathbf{s})$ with respect to \mathbf{b}^* is

$$\begin{aligned} \nabla H(\mathbf{b}^{*T} \mathbf{s}) &= E\{\mathbf{s} f(\mathbf{b}^{*T} \mathbf{s})\} = E\{\mathbf{s} f(s_i)\} \\ &= (0_1, \dots, 0_{i-1}, 1, 0_{i+1}, \dots)^T E\{s_i f(s_i)\}. \end{aligned} \quad (16)$$

The $E\{f(s_i) s_j\}_{i \neq j} = 0$, which means that the source signals with zero are mutually statistically independent, is used in (16). It is a fundamental assumption in ICA-R [17] and its second derivative is as follows:

$$\begin{aligned} \nabla^2 H(\mathbf{b}^{*T} \mathbf{s}) &= E\{\mathbf{s} \mathbf{s}^T f'(\mathbf{b}^{*T} \mathbf{s})\} = \text{diag}(E\{f'(s_i)\}, \\ &\dots, E\{s_i^2 f'(s_i)\}, \dots, E\{f'(s_i)\}). \end{aligned} \quad (17)$$

Adding a small fluctuation $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_{i-1}, \sigma_i, \sigma_{i+1}, \dots)$ into \mathbf{b}^* , we obtain

$$\mathbf{b}_{\text{new}} = \mathbf{b}^* + \boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_{i-1}, \sigma_i + 1, \sigma_{i+1}, \dots). \quad (18)$$

Taylor's series expansion of $H(\mathbf{b}_{\text{new}}^T \mathbf{s})$ at the point \mathbf{b}^* is

$$\begin{aligned} H\left[(\mathbf{b}^* + \boldsymbol{\sigma})^T \mathbf{s}\right] &= H(\mathbf{b}^{*T} \mathbf{s}) + \boldsymbol{\sigma} \nabla H(\mathbf{b}^{*T} \mathbf{s}) + \frac{1}{2} \boldsymbol{\sigma} \nabla^2 H(\mathbf{b}^{*T} \mathbf{s}) \boldsymbol{\sigma}^T \\ &\quad + o(\|\boldsymbol{\sigma}\|^2) \\ &= H(\mathbf{b}^*) + \boldsymbol{\sigma}_i E\{s_i f(s_i)\} \\ &\quad + \frac{1}{2} \left[E\{s_i^2 f'(s_i)\} \boldsymbol{\sigma}_i^2 + E\{f'(s_i)\} \sum_{n \neq i} \boldsymbol{\sigma}_n^2 \right] \\ &\quad + o(\|\boldsymbol{\sigma}\|^2). \end{aligned} \quad (19)$$

In the one-unit ICA-R, the extracted signal y is considered to have unit variance. When the desired signal has unit variance,

then $\|\mathbf{b}^* + \boldsymbol{\sigma}\| = \mathbf{I}$. That is, $\sigma_1^2 + \sigma_2^2 + \dots + (\sigma_i + 1)^2 + \dots = 1$. Thus, we obtain $(\sigma_i + 1) = \sqrt{1 - \sigma_1^2 \dots - \sigma_{i-1}^2 - \sigma_{i+1}^2 \dots}$ and its Taylor's series at point zero is

$$\sigma_i + 1 = 1 - \frac{1}{2} \sum_{n \neq i} \boldsymbol{\sigma}_n^2 + o\left(\sum_{n \neq i} \boldsymbol{\sigma}_n^2\right). \quad (20)$$

Substituting (20) in (19), we get

$$\begin{aligned} H\left[(\mathbf{b}^* + \boldsymbol{\sigma})^T \mathbf{s}\right] &= H(\mathbf{b}^{*T} \mathbf{s}) - \frac{1}{2} E\{s_i f(s_i)\} \sum_{n \neq i} \boldsymbol{\sigma}_n^2 \\ &\quad + \frac{1}{2} \left[E\{s_i^2 f'(s_i)\} \boldsymbol{\sigma}_i^2 + E\{f'(s_i)\} \sum_{n \neq i} \boldsymbol{\sigma}_n^2 \right] \\ &\quad + o(\|\boldsymbol{\sigma}\|^2) \\ &= H(\mathbf{b}^{*T} \mathbf{s}) + \frac{1}{2} [E\{f'(s_i)\} - E\{s_i f(s_i)\}] \sum_{n \neq i} \boldsymbol{\sigma}_n^2 \\ &\quad + o(\|\boldsymbol{\sigma}\|^2) \\ &\approx H(\mathbf{b}^{*T} \mathbf{s}) + \frac{1}{2} [E\{f'(s_i)\} - E\{s_i f(s_i)\}] \sum_{n \neq i} \boldsymbol{\sigma}_n^2. \end{aligned} \quad (21)$$

As the components of $\boldsymbol{\sigma}$ are very small, we can neglect its higher order term in (20) and (21). From the last item in (21), we can obtain that when $E\{f'(s_i)\} - E\{s_i f(s_i)\} > 0$ then $H(\mathbf{b}^*)$ is the minimum, and vice versa is the maximum.

4. Improve One-Unit ICA-R Method

As the previous one-unit ICA-R methods involve the computation of the $\hat{\rho} = 2\rho[E\{F(y)\} - E\{F(v)\}]$ in every iteration, it is a time consuming procedure. If the computation of $[E\{F(y)\} - E\{F(v)\}]$ is avoided in $\hat{\rho}$, the implementation of the ICA algorithm and ICA-R algorithm can be further speeded up. We wish to incorporate the prior information of the desired signal into non-Gaussian measurement function $J(y)$ to eliminate $E\{F(v)\}$ to avoid further computing $[E\{F(y)\} - E\{F(v)\}]$. $J(y)$ is a part of the cost function of the one-unit ICA-R. As a result, this will reduce computation complexity of the one-unit ICA-R.

4.1. Simplification of the Non-Gaussian Measurement Function. Since the nonlinear function F used in one-unit ICA-R is a positive even function, the terms $E\{F(v)\}$ and $E\{F(y)\}$ are both positive in (14), and v has nothing to do with the separating vector and thus can be seen as a constant. Accordingly, when $E\{F(y)\}$ has a maximum value, this equals $[E\{F(y)\} - E\{F(v)\}]$, which has a maximum value. On the contrary, when $E\{F(y)\}$ has a minimum value, this equals $-[E\{F(y)\} - E\{F(v)\}]$, which has maximum value.

Based on the analysis in Section 3, when $E\{f'(s_i)\} - E\{s_i f(s_i)\} < 0$, $E\{F(y)\}$ has maximum value. Therefore,

when $E\{f'(s_i)\} - E\{s_i f(s_i)\} < 0$, we can determine that $[E\{F(y)\} - E\{F(v)\}]$ has a maximum value based on the analysis described in the first paragraph of this subsection. Based on the above analysis, when $E\{f'(s_i)\} - E\{s_i f(s_i)\} < 0$, then (14) can be simplified as

$$J(y) = \rho [E\{F(y)\} - E\{F(v)\}]. \quad (22)$$

As v has nothing to do with the separating vector, we can further simplify the non-Gaussian measurement function (22) as

$$J(y) = \rho E\{F(y)\}. \quad (23)$$

On the other hand, when the $E\{f'(s_i)\} - E\{s_i f(s_i)\} > 0$, the non-Gaussian measurement function can be simplified as

$$J(y) = -\rho E\{F(y)\}. \quad (24)$$

Therefore, the negentropy approximation in (23) and (24) can be expressed as

$$J(y) = \rho \text{sign}(c) E\{F(y)\}, \quad (25)$$

where $c = E\{s_i f(s_i)\} - E\{f'(s_i)\}$.

$$\text{sign}(c) = \begin{cases} 1 & c > 0 \\ -1 & c < 0. \end{cases} \quad (26)$$

In [16], the following function $F(y)$ was proposed for extracting the super-Gaussian and sub-Gaussian signal, respectively.

$$F_{\text{sup}}(y) = \frac{1}{a} \log \cosh(ay) - \frac{a}{2} y^2 \quad (27)$$

$$F_{\text{sub}}(y) = \frac{b}{4} y^4, \quad (28)$$

where $a, b \in R^+$. In [25], Mi has proved that (27) should be corrected as

$$F_{\text{sup}}(y) = \frac{1}{a} \log \cosh(ay). \quad (29)$$

Accordingly, in this paper, we use (29) instead of (27). For simplicity, we set $a = 1$ in (29). The first and second derivatives are $\tanh(y)$ and $1 - \tanh^2(y)$, respectively.

If we use (29) as a nonlinear function, c in (25) can be expressed as follows:

$$\begin{aligned} c &= E\{s_i f(s_i)\} - E\{f'(s_i)\} \\ &= E\{s_i \tanh(s_i)\} - E\{(1 - \tanh^2(s_i))\} \\ &= E\left\{s_i^2 - \frac{s_i^4}{3} + o(s_i^6)\right\} - 1 \\ &\quad + E\left\{s_i^2 - \frac{2s_i^4}{3} + o(s_i^6)\right\}. \end{aligned} \quad (30)$$

In (30), we used Taylor's series expansion of $\tanh(s_i)$, which is described as

$$\tanh(s_i) = s_i - \frac{s_i^3}{3} + o(s_i^5). \quad (31)$$

Neglecting the higher order term in (30) and considering that the desired signal has unit variance, which is a very common hypothesis in ICA and ICA-R, we can obtain the following:

$$\begin{aligned} c &= E\left\{s_i^2 - \frac{s_i^4}{3}\right\} - 1 + E\left\{s_i^2 - \frac{2s_i^4}{3}\right\} \\ &= E\{2s_i^2\} - 1 - E\{s_i^4\} = 1 - E\{s_i^4\}. \end{aligned} \quad (32)$$

According to the kurtosis definition of a super-Gaussian signal, when s_i is a super-Gaussian signal with unit variance, $E\{s_i^4\} > 3$. Then $c < 0$ in (32). This means that if the desired signal is super-Gaussian, we can directly simplify (25) as

$$J(y) = -\rho E\{F(y)\}. \quad (33)$$

4.2. Improved One-Unit ICA-R for Extracting FECG. The one-unit ICA-R is used to extract one desired signal. The prior information that the desired signal is a super-Gaussian or a sub-Gaussian signal is easily obtained. The FECG and MECG are both super-Gaussian signals in general. Therefore, in the extraction of FECG, we can directly use (33) instead of (14) to obtain the non-Gaussian measurement of an extracted signal. Centering and whitening can reduce the computation complexity of one-unit ICA-R [11]. Thus, we use centering and whitening to process the observed signals. After preprocessing, the new cost function of the one-unit ICA-R for extracting FECG can be expressed as

$$\begin{aligned} \min \quad & -J(y) \approx \rho E\{F(y)\} \\ \text{s.t.} \quad & g(y) \leq 0, \end{aligned} \quad (34)$$

where $y = \mathbf{w}^T \mathbf{x}$ is the extracted signal, \mathbf{x} is the centered and whitened signal, and \mathbf{w} is the separating vector. ρ in (34) is different from $\hat{\rho}$ in (12). ρ is just a positive constant. Correspondingly, the augmented Lagrangian function $L(w, \mu)$ for the problem in (34) is expressed as follows:

$$L(w, \mu) = J(y) - \frac{1}{2\gamma} [\max^2\{rg + \mu, 0\} - \mu^2]. \quad (35)$$

To find the maximum of $L(w, \mu)$, Newton's method is used to optimize $L(w, \mu)$. The learning rule of the separating vector w is given by the following equation:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \eta \frac{l}{\delta}, \quad (36)$$

where η is the learning rate:

$$\begin{aligned} l &= \rho E[\mathbf{x}f(y)] + \mu E[\mathbf{x}g'(y)], \\ \delta &= \rho E[f'(y)] + \mu E[g''(y)], \\ g(y) &= \varepsilon(y, r) - \xi. \end{aligned} \quad (37)$$

The main difference between the improved algorithm in (36) and the original algorithm in (12) is that ρ in (36) is just a positive constant and $\hat{\rho}$ in (12) is $2\rho(E\{F(y)\} - E\{F(v)\})$. In the original algorithm, we have to additionally compute $(E\{F(y)\} - E\{F(v)\})$ in every iteration, which considerably consumes much more time compared with the improved algorithm.

The one-unit ICA-R uses Newton's method to optimize cost function. Newton's method is sensitive for the initial value. The initial separating vector directly affects the convergence of Newton's method. However, the original algorithm randomly initializes the separating vector. To avoid random initialization, inspired by previous a report [20], we use the following expression to initialize the separating vector:

$$\mathbf{w}_0 = (\mathbf{r}\mathbf{x}^{-1})^T, \quad (38)$$

where \mathbf{r} is the whitened reference signal, \mathbf{x} is the whitened signal, \mathbf{x}^{-1} is the inversion of \mathbf{x} , and \mathbf{w}_0 is the initialization separating vector. The closer to whitened desired signal \mathbf{r} is, the closer to the perfect separating vector \mathbf{w}_0 will be. This can speed up the algorithm convergence rate and improve the probability of global convergence compared with random initialization.

Therefore, the proposed one-unit ICA-R algorithm for extracting FECG can be described as follows:

- (1) Center the observed signals.
- (2) Whiten the centered signal.
- (3) Choose a proper scalar penalty parameter γ , convergence threshold ς , parameter ρ , and the leaning rate η ; generally let $\rho = 1$, $\eta = 1$.
- (4) Use (38) to initialize separating vector \mathbf{w}_0 .
- (5) Normalize $\mathbf{w}_0 \leftarrow \mathbf{w}_0 / \|\mathbf{w}_0\|$.
- (6) Choose an initial value for μ .
- (7) Update the Lagrange multiplier μ by (9).
- (8) Update the separating vector \mathbf{w} by (36).
- (9) Normalize the separating vector w by

$$\mathbf{w}_{k+1} \leftarrow \frac{\mathbf{w}_{k+1}}{\|\mathbf{w}_{k+1}\|}. \quad (39)$$

- (10) Until $\min(\|\mathbf{w}_{k+1} - \mathbf{w}_k\|, \|\mathbf{w}_{k+1} + \mathbf{w}_k\|) \leq \varsigma$, otherwise go back to step (7).

5. Simulation and Discussion

To demonstrate the effectiveness of the proposed one-unit ICA-R algorithm, both artificial ECG data and real-word ECG data are used in the following two experiments. To quantitatively compare the performance of the proposed one-unit ICA-R and Mi's one-unit ICA-R [25], we adopt the individual performance index (IPI) as an indicator function, as was used in a previous study [11]. It is defined as

$$\text{IPI} = \left(\sum_{i=1}^M \frac{|p_i|}{\max_j |p_j|} \right) - 1, \quad j = 1, \dots, M, \quad (40)$$

where $|\cdot|$ denotes the absolute value operator, p_i denotes i element of the global system vector $\mathbf{p} = \mathbf{w}^T \mathbf{B} \mathbf{A} = (p_1, p_2, \dots, p_M)$, \mathbf{A} is mixing matrix, \mathbf{B} is the whitened matrix, \mathbf{w} is the separating vector of mixed signals after whitening, $\mathbf{w}^T \mathbf{B}$ is the global separating vector, M is the number of mixing signals, and $\max_j |p_j|$ is to find the maximum valued among the elements in the vector $\mathbf{p} = (p_1, p_2, \dots, p_M)$. The extracted signal by proposed method can be described as

$$y = \mathbf{w}^T \mathbf{B} \mathbf{x} = \mathbf{w}^T \mathbf{B} \mathbf{A} \mathbf{s} = \mathbf{p} \mathbf{s} = \sum_{i=1}^M p_i s_i. \quad (41)$$

From (41), it can be easily seen that if $\mathbf{p} = (0, 0, \dots, p_k, \dots, 0)$, $p_k \neq 0$, we will have the extracted signal $y = p_k s_k$, which is just rescaled signals of s_k . It is obvious from (40) that we have $\text{IPI} \geq 0$ and that IPI equals zero if and only if \mathbf{p} is a rescaled canonical base vector that only one element of vector is not zero. Therefore, the closer to a rescaled canonical base vector the global system vector \mathbf{p} , the nearer to zero the IPI and thus the better the performance of one-unit ICA-R method.

In the experiments, we select $F(y) = \log \cosh(y)$ as the nonquadratic function and $\varepsilon(y, r) = E\{(y - r)^2\}$ as the closeness measure function. Additionally, $g(y) = \varepsilon(y, r) - \xi$. We also use the corrected first derivative of $g(y)$ as its first derivative. The corrected first derivative is given in [19] as

$$\nabla_{\mathbf{w}} g(y) = E\{\mathbf{x} g'(y)\} = 2E\{\mathbf{x}(y - r)\}. \quad (42)$$

In the following experiments, we set ρ , η , and γ all equal to one and perform 1000 trials for each data set.

5.1. Experiment with Artificial ECG Data. The artificial signals include the power line interference with 50 Hz, Gaussian noise, FECG, and MECG. The FECG and MECG are generated by using the ECG toolbox of Sameni [28]. The reference signals for the FECG and MECG are constructed with sign operation to roughly give the signs of the most data samples of the desired source signal [14, 15]. There are 5000 samples in the experiment. The mixing matrix is randomly generated as

$$A = \begin{bmatrix} 0.8925 & 0.0570 & 0.5044 & 0.9153 \\ 0.0169 & 0.0590 & 0.4364 & 0.4911 \\ 0.5165 & 0.4735 & 0.8193 & 0.7484 \\ 0.0418 & 0.3840 & 0.4448 & 0.3421 \end{bmatrix}. \quad (43)$$

Four artificial source signals are depicted in Figure 1, which are considered independently of each other. s1 is the power line interference with sub-Gaussian distribution. s2 is the random Gaussian noise with Gaussian distribution. s3 and s4 are FECG and MECG with super-Gaussian distribution, respectively. The mixtures are shown as mixed signals in Figure 2. The FECG extracted by the proposed method and its reference are shown in Figure 3. The MECG extracted by the proposed method and its reference are shown in Figure 4. Comparisons of the extracted FECG and the extracted MECG are shown in Figures 3 and 4 with the

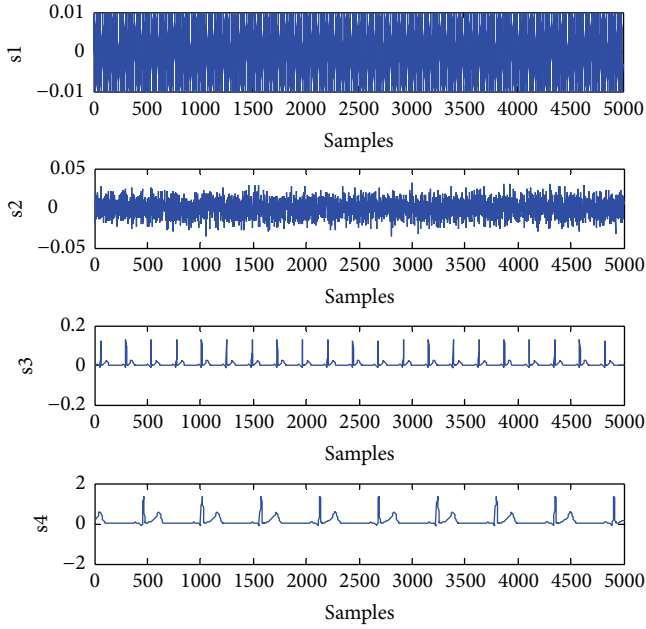


FIGURE 1: Four artificial source signals. (s1): power line interference. (s2): random Gaussian noise. (s3): FECG signal. (s4): MCEG signal.

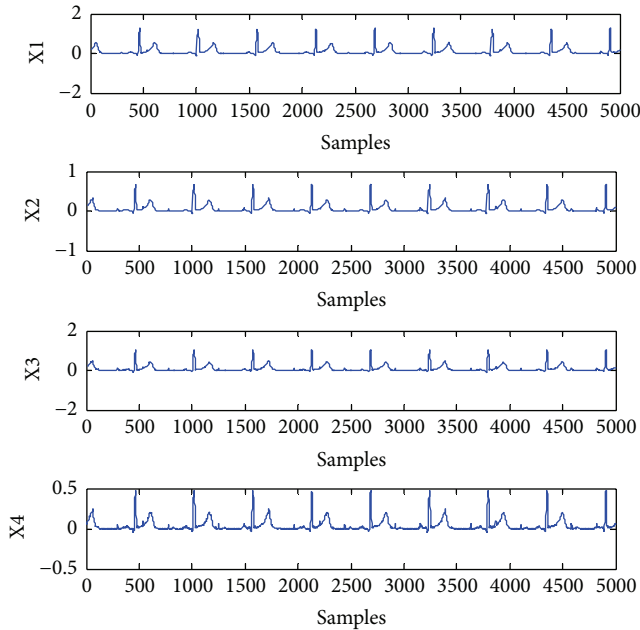


FIGURE 2: Mixed signals from sources in Figure 1.

desired signals s3 and s4 in Figure 1, respectively. It is evident that the waveforms of the extracted signals are most identical to the waveforms of the desired signals.

To quantitatively compare the performance of the proposed method and that of Mi's method [25], the IPI, defined in a previous report (40), are computed for the extraction results of both algorithms in the same experiment. Both algorithms use the same parameters and reference signal. The results are shown in Table 1, where it can be seen that the IPIs

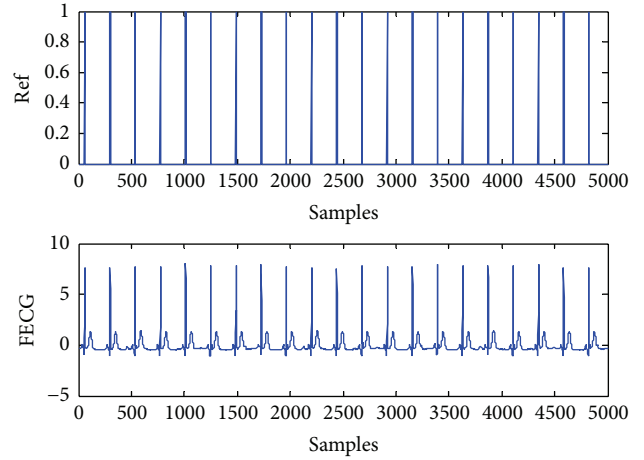


FIGURE 3: Reference signal and extracted artificial FECG signal. (Ref): reference signal for FECG signal. (FECG): extracted artificial FECG signal by the proposed method from mixed signals in Figure 2.

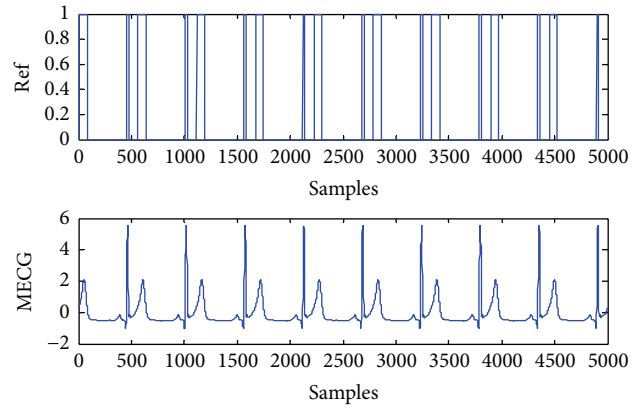


FIGURE 4: Reference signal and extracted artificial MCEG signal. (Ref): reference signal for MCEG signal. (MCEG): extracted artificial MCEG signal by the proposed method from mixed signals in Figure 2.

TABLE 1: Comparison of the IPI performance and time consumed by the proposed method and Mi's method.

Desired signal		FECG	MCEG
Proposed	IPI	0.0144	0.0788
Mi [25]	IPI	0.0144	0.0788
Proposed	Time (s)	0.010359	0.014047
Mi [25]	Time (s)	0.025578	0.033813

are the same for the proposed method and Mi's method. This means that both methods successfully extract the FECG and MCEG with the same low IPI. However, the running time consumed by both methods is quite different. The running time of the method proposed here is roughly half of that of Mi's algorithm.

5.2. Experiment with Real-World ECG Data. The real-world ECG data was contributed by De Moor [27] and has been

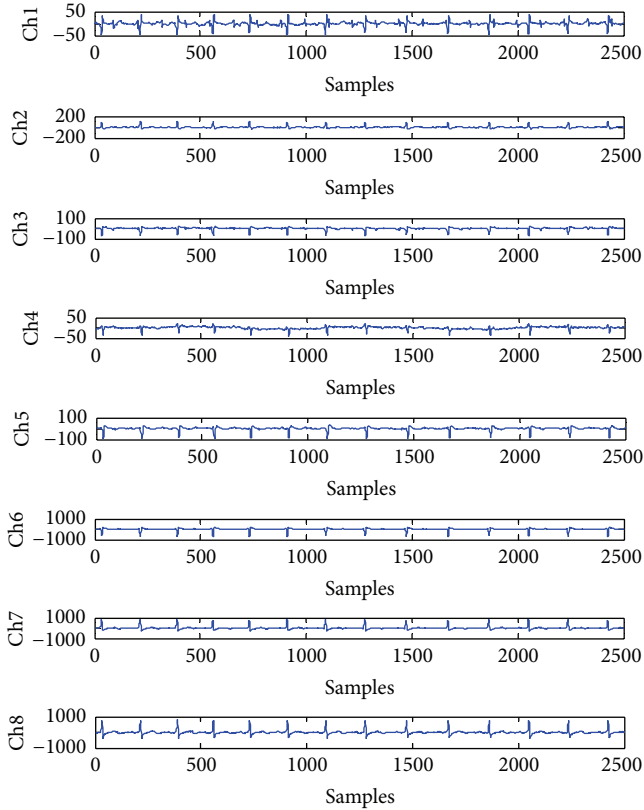


FIGURE 5: Real-world ECG recordings: the 8 channels of ECG recordings obtained from a pregnant woman with data length of 2500 samples and time of 10 s [27].

TABLE 2: Comparison of the time consumed by the proposed method and Mi's method for the real-word ECG data.

	Desired signal	FECG	MECG
Proposed	Time (s)	0.022828	0.022109
Mi [25]	Time (s)	0.050484	0.049234

widely used by other researchers [11–13, 25]. It was recorded over 10 s and sampled at 250 Hz by placing eight electrodes on different locations of a pregnant woman. The real-word ECG data are shown in Figure 5, where the signals Ch1–Ch5 were the recordings from five electrodes placed on the woman's abdomen. Accordingly, the FECG, respiratory motion artifacts, and the MECG were visible in these recordings. The signals Ch6–Ch8 were the recordings from 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.

In the Ch1 recording, the strong and slow heart belongs to the mother, while the weak and fast belongs to the fetus. Therefore, we can use the information about fetus in the Ch1 recording to construct reference signal for extracting the FECG. We use impulse series whose occurrence time is the same as that of the subpeaks in the first channel signal Ch1 as reference for extracting the FECG. The impulse series whose occurrence time is the same as the peaks of any channel signal were used as reference for extracting the MECG.

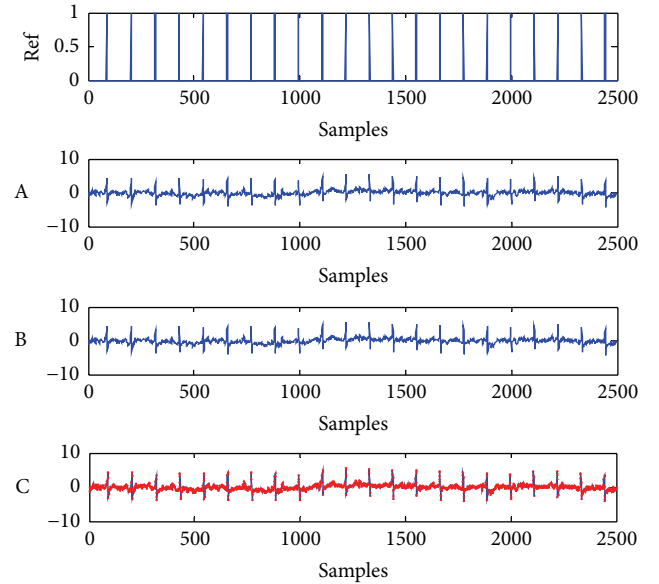


FIGURE 6: Reference signal and extracted FECG signals by both methods from the real-world ECG signals in Figure 5. Ref: reference signal for FECG signal. A-B: extracted artificial FECG signals by the proposed method and Mi's method, respectively. C: red and blue curves are exactly the same with the subfigures A and B, respectively.

The reference signal and the FECG signal extracted by both methods are shown in Figure 6, where the reference signal is denoted as Ref, the FECG extracted by the proposed method is denoted as A, the FECG extracted by Mi's method is denoted as B, and both extracted FECG signals are redescribed in C and represented with different color curves. The reference signal and the MECG signal extracted by both methods are shown in Figure 7, where the reference signal is also described as Ref, the MECG extracted by the proposed method is denoted as A, the MECG extracted by the previous method is denoted as B, and both extracted MECG signals are redescribed in C with different color curves. We can see in Figures 6 and 7 that the FECG and MECG are both successfully extracted by both methods. Since the mixing A and the pure FECG and MECG signals are not available for the real-word ECG recordings, the IPI performance cannot be computed. But we can compare the waveforms of extracted signals by both methods to estimate their relative error. Therefore, we redescribe the extracted FECG and MECG signal waveforms by both methods in the same subfigure C with different color curves in Figures 6 and 7, respectively. From the subfigure C in Figures 6 and 7, it can be seen that the different color curves are almost the same. This means that the proposed method and Mi's method have almost the same error. However, their consumed running time is considerably different, as shown in Table 2. The running time is approximately reduced to half of that of the Mi's algorithm.

6. Conclusion

In this paper a fast one-unit ICA-R algorithm for extracting ECG is proposed by simplifying the non-Gaussian measurement method of the extracted signal. In the proposed

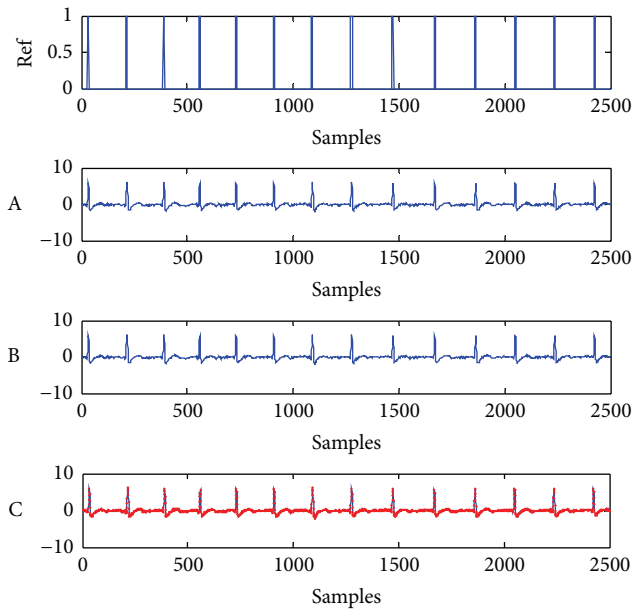


FIGURE 7: Reference signal and extracted MEGC signals by both methods from the real-world ECG signals in Figure 5. Ref: reference signal for MEGC signal. A-B: extracted artificial MEGC signals by the proposed method and Mi's method, respectively. C: red and blue curves are exactly the same with the subfigures A and B, respectively.

algorithm, the nonlinear function of the extracted signal is directly used as non-Gaussian measurement function. This avoids having to compute the difference between the function of the extracted signal and the function of the Gaussian random vector with the same mean and variance as the extracted signal. Centering and whitening are also used to preprocess the observed data to avoid computing the observed signal covariance matrix. As a result, the computation complexity for the one-unit ICA-R algorithm is greatly reduced.

The validity of the proposed algorithm is tested and compared to Mi's algorithm using artificial ECG data and real-world ECG data. Both experimental results demonstrate that the convergence rate of the proposed algorithm is two times faster than Mi's algorithm, while both methods have the same error.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61271115) and the Foundation of Jilin Educational Committee (2015235).

References

- [1] E. R. Ferrara and B. Widrow, "Fetal electrocardiogram enhancement by time-sequenced adaptive filtering," *IEEE Transactions on Biomedical Engineering*, vol. 29, no. 6, pp. 458–460, 1982.
- [2] P. P. Kanjilal, S. Palit, and G. Saha, "Fetal ECG extraction from single-channel maternal ECG using singular value decomposition," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 1, pp. 51–59, 1997.
- [3] A. H. Khamene and S. Negahdaripoure, "A new method for the extraction of fetal ECG from the composite abdominal signal," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 4, pp. 507–516, 2000.
- [4] L. de Lathauwer, B. de Moor, and J. Vandewalle, "Fetal electrocardiogram extraction by blind source subspace separation," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 5, pp. 567–572, 2000.
- [5] M. A. Hasan, M. I. Ibrahimy, and M. B. I. Reaz, "Techniques of FECG signal analysis: detection and processing for fetal monitoring," *WIT Transactions on Biomedicine and Health*, vol. 12, pp. 295–305, 2007.
- [6] S. Sargolzaei, K. Faez, and A. Sargolzaei, "Signal processing based for fetal electrocardiogram extraction," in *Proceedings of the International Conference on BioMedical Engineering and Informatics*, vol. 2, pp. 492–496, Sanya, China, May 2008.
- [7] Y. Sevim and A. Atasoy, "Performance evaluation of nonparametric ICA algorithm for fetal ECG extraction," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 19, no. 4, pp. 657–666, 2011.
- [8] G. Liu and Y. Luan, "An adaptive integrated algorithm for non-invasive fetal ECG separation and noise reduction based on ICA-EEMD-WS," *Medical & Biological Engineering & Computing*, vol. 53, no. 11, pp. 1113–1127, 2015.
- [9] Y. Li, W. Nie, F. Ye, and A. Li, "A fetal electrocardiogram signal extraction algorithm based on the temporal structure and the non-Gaussianity," *Computational and Mathematical Methods in Medicine*, vol. 2016, Article ID 9658410, 10 pages, 2016.
- [10] J. Lee, K. L. Park, and K. J. Lee, "Temporally constrained ICA-based foetal ECG separation," *Electronics Letters*, vol. 41, no. 21, pp. 1158–1160, 2005.
- [11] Q.-H. Lin, Y.-R. Zheng, F.-L. Yin, H. Liang, and V. D. Calhoun, "A fast algorithm for one-unit ICA-R," *Information Sciences*, vol. 177, no. 5, pp. 1265–1275, 2007.
- [12] Z.-L. Zhang, "Morphologically constrained ICA for extracting weak temporally correlated signals," *Neurocomputing*, vol. 71, no. 7-9, pp. 1669–1679, 2008.
- [13] Z. Yongjian, "Constrain component extraction techniques," in *Proceedings of the 3rd International Conference on Mechatronics, Robotics and Automation*, pp. 2796–2799, Shenzhen, China, 2015.
- [14] C. Li, G. Liao, and Y. Shen, "An improved method for independent component analysis with reference," *Digital Signal Processing*, vol. 20, no. 2, pp. 575–580, 2010.
- [15] W. Lu and J. C. Rajapakse, "ICA with reference," in *Proceedings of the 3rd International Conference on Independent Component Analysis and Blind Source Separation (ICA '01)*, pp. 120–125, San Diego, Calif, USA, 2001.
- [16] W. Lu and J. C. Rajapakse, "Approach and applications of constrained ICA," *IEEE Transactions on Neural Networks*, vol. 16, no. 1, pp. 203–212, 2005.
- [17] W. Lu and J. C. Rajapakse, "ICA with Reference," *Neurocomputing*, vol. 69, no. 16-18, pp. 2244–2257, 2006.
- [18] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

- [19] D.-S. Huang and J.-X. Mi, "A new constrained independent component analysis method," *IEEE Transactions on Neural Networks*, vol. 18, no. 5, pp. 1532–1535, 2007.
- [20] Z.-L. Sun and L. Shang, "An improved constrained ICA with reference based unmixing matrix initialization," *Neurocomputing*, vol. 73, no. 4-6, pp. 1013–1017, 2010.
- [21] S. S. Kavuri, J. M. Zurada, and J. C. Rajapakse, "Evolutionary approach to ICA-R," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '10)*, pp. 1–6, Barcelona, Spain, July 2010.
- [22] J. Zhang, Z. S. Zhang, W. Cheng et al., "Kurtosis-based constrained independent component analysis and its application on source contribution quantitative estimation," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 7, pp. 1842–1854, 2014.
- [23] Z. Wang, "Fixed-point algorithms for constrained ICA and their applications in fMRI data analysis," *Magnetic Resonance Imaging*, vol. 29, no. 9, pp. 1288–1303, 2011.
- [24] P. A. Rodriguez, M. Anderson, X.-L. Li, and T. Adali, "General non-orthogonal constrained ICA," *IEEE Transactions on Signal Processing*, vol. 62, no. 11, pp. 2778–2786, 2014.
- [25] J.-X. Mi, "A novel algorithm for independent component analysis with reference and methods for its applications," *PLoS ONE*, vol. 9, no. 5, Article ID e93984, 2014.
- [26] G. H. Chen, L. F. Qie, A. J. Zhang, and J. Han, "Improved CICA algorithm used for single channel compound fault diagnosis of rolling bearings," *Chinese Journal of Mechanical Engineering*, vol. 29, no. 1, pp. 204–211, 2016.
- [27] D. De Moor, "Daisy: database for the identification of systems," <http://www.esat.kuleuven.ac.be/sista/daisy>.
- [28] R. Sameni, <http://ee.sharif.edu/~ecg/Download/>.