Proceedings of the 2007 IEEE Symposium on Computational
Intelligence in Bioinformatics and Computational Biology (CIBCB 2007)

A Flexible Fully-Multiplicative Orthogonal-Group Based ICA Algorithm

Yalan Ye [1], Jing Wan [2], Zhi-Lin Zhang [1]
[1] School of Computer Science and Engineering,
University of Electronic Science and
Technology of China, Chengdu 610054, China
{yalanye,zlzhang,chenjia,wulei}@uestc.edu.cn

Chen Jia [1], Wu Lei [1]
[2] Institute of Image Processing and Pattern Recognition,
Shanghai Jiao Tong University
Shanghai 200240, China
awaning@sjtu.edu.cn

Abstract — **In this paper, we propose a flexible fully-multiplicative orthogonal-group based ICA (FlexibleOgICA) algorithm, which can instantaneously separate the mixture of sub-Gaussian and super-Gaussian source signals. It adopts a self-adaptive nonlinear function, which adjusts its parameter to achieve better performance based on the estimation of the kurtosis of super-Gaussian source signals. We also have successfully applied the algorithm to obtain the fetal electrocardiogram (FECG) signal, showing its fast convergence speed and high separation performance.**

I. INTRODUCTION

Over the past decade independent component analysis (ICA) [1][3][8] or blind source separation (BSS) has attracted a great deal of attention because of its potential applicability to a wide range of problems, spanning disciplines as diverse as communications, signal processing, feature extraction and biomedical signal processing [9][10][11]. The problem arises when multidimensional observations, generated when a set of signals are mixed by passage through an unknown medium, must be processed to recover the original sources without the benefit of any a priori knowledge about the mixing operation or the sources themselves.

Many different algorithms have been developed for the solution of ICA, which blindly separate the mixture of sub- and super-Gaussian [3][5][8]. An important work is done by Lee and Girolami [3]. They present an extend Infomax (ExtICA) algorithm that is able to blindly separate mixed signals with sub-Gaussian and super-Gaussian source distributions by using the nonlinear model switch technique. The application of the algorithm is prevalent in separating a wider range of source signals. Unfortunately, the convergence of the algorithm is relatively slow.

Recently, Fiori [4] presents a fully-multiplicative orthogonal-group ICA neural algorithm, which exploits the known principle of diagonalization of a tensor of a warped network's outputs. It has very fast convergence speed. However the algorithm is only separate sub-Gaussian source signals. In other words, it is unable to separate the mixture of pure super-Gaussian source signals, or the mixture of sub-Gaussian and super-Gaussian source signals. Thus the algorithm's application has been greatly limited.

To address the problem above, we propose the FlexibleOgICA algorithm, which exploits multiplicative orthogonal-group [4] and uses a self-adaptive nonlinear function controlled by a single parameter (Gaussian exponent) that is adjusted according to the estimated kurtosis value of algorithm output [5]. The algorithm has fast convergence speed and high separation performance. Computer simulations on artificially generated data and real-world ECG data have shown its better performance, compared with the classical extend Infomax algorithm [3].

II. PROPOSED FLEXIBLEOGICA ALGORITHM

The basis ICA model can be summarized as follows: assume that there exist mutually independent unknown source $s_i(i=1,\ldots,N)$, which have zero mean and unit variance. And also assume that the sources are linearly mixed with an unknown $M\times N(M\geqslant N)$ matrix $\mathbf{A}$:

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

where $\mathbf{s} = [s_1, s_2, \cdots, s_N]^T$ and $\mathbf{x} = [x_1, x_2, \cdots, x_M]^T$ are N-dimensional sources and M-dimensional mixed signals respectively. In independent component analysis, the basic goal is to find an $N\times M$ separating matrix $\mathbf{W}$ without knowing the mixing matrix $\mathbf{A}$, that is

$$\mathbf{y} = \mathbf{W}^T\mathbf{x} \tag{2}$$

such that $\mathbf{y}= [y_1, y_2, \cdots, y_N]^T$ is an estimate of $\mathbf{s}$ that each component of $\mathbf{s}$ may appear in any component of $\mathbf{y}$ with a scalar factor. For the sake of simplicity, we assume $M=N$.

According to [6], $\mathbf{W}$ belongs to the orthogonal group $\boldsymbol{O}(N)=\{\mathbf{A}\in \mathrm{R}^{N\times N}|\mathbf{A}^T\mathbf{A}=\mathbf{I}_N\}$. A class of ICA algorithm stems from the following well-known principle: Given two odd nonlinear functions $f(\cdot)$and $g(\cdot)$, the separating algorithm ought to diagonalize the matrix $E_{\mathrm{y}}[f(\mathbf{y})\, g(\mathbf{y})^T]$, where the symbol $E_{\mathbf{x}}[\ h(\mathbf{x})\ ]$ denotes statistical expectation of the function $h(\mathbf{x})$ over the distribution of the random-vector $\mathbf{x}$.

Let us define the matrix $\mathbf{G} = E_{\mathbf{y}}[f(\mathbf{y})\ g(\mathbf{y})^{\mathrm{T}}] \in \mathrm{R}^{N\times N}$ and define an $N\times N$ real, invertible, positive-define diagonal matrix $\mathbf{B}$. The above recalled diagonalization principle may be expressed in the following way which proves useful in the following development:

$$\mathbf{G}^{-1} = \mathbf{B}^{-1} \tag{3}$$

By pre-multiplying by $\mathbf{W}$ and post-multiplying by $\mathbf{B}$ both members of the above equation, we obtain:

$$\mathbf{W}\,\mathbf{G}^{-1}\,\mathbf{B} = \mathbf{W} \tag{4}$$

This expression readily suggests the following batch iterative learning algorithm:

$$\mathbf{W}_{n+1}^{+} = \mathbf{W}_n\,\mathbf{G}_n^{-1}\,\mathbf{B}, \quad n=0, 1, 2,\ldots \tag{5}$$

$$\mathbf{W}_{n+1} = (\mathbf{W}_{n+1}^{+}(\mathbf{W}_{n+1}^{+})^{\mathrm{T}})^{-1/2}\,\mathbf{W}_{n+1}^{+} \tag{6}$$

where $\mathbf{G}_n = E_{\mathbf{y}}[f(\mathbf{y}_n)\ g(\mathbf{y}_n)^{\mathrm{T}}]$, with $\mathbf{y}_n = \mathbf{W}_n^{\mathrm{T}}\mathbf{x}$, $\mathbf{W}_0 \in \boldsymbol{O}(N)$ is an initial value and n deno-tes the iteration index. The last expression (6) denotes the symmetric orthogonalization step that is necessary to project the update matrix $\mathbf{W}_{n+1}$ into the orthogonal group.

In order to separate instantaneous mixtures of sub- and super-Gaussian source signals, we choose a switching criterion, based on self-adaptive nonlinear function that controlled by a single parameter (Gaussian exponent a) according to the estimated kurtosis value of separating filter output [5]. So the FlexibleOgICA batch iterative learning algorithm:

$$\mathbf{W}_{n+1}^{+} = \mathbf{W}_n\,\mathrm{E}_{\mathbf{y}}[f(\mathbf{y}_n)\ g(\mathbf{y}_n)^{\mathrm{T}}]^{-1}\,\mathbf{B}, \quad n=0, 1, 2,\ldots \tag{7}$$

$$\mathbf{W}_{n+1} = (\mathbf{W}_{n+1}^{+}(\mathbf{W}_{n+1}^{+})^{\mathrm{T}})^{-1/2}\,\mathbf{W}_{n+1}^{+} \tag{8}$$

Some practical nonlinear functions are suggested for $f(\mathbf{y}_n)$ and $g(\mathbf{y}_n)$ [5] as following. The nonlinear function $g(\mathbf{y}_n)$ remains invariable whatever sub-Gaussian signals or super-Gaussian ones:

$$g(\mathbf{y}_n) = \mathbf{y}_n^{2}\ sign\,(\mathbf{y}_n) \tag{9}$$

where $sign(\mathbf{y}_n)$is the signum function of $\mathbf{y}_n$. And the typical examples of the other nonlinear function $f(\mathbf{y}_n)$ with different values of a is shown:

① $a = 4$ for $k_i < 0$ (i.e, $\mathbf{y}_n$ is sub-Gaussian.):

$$f(\mathbf{y}_n) = |\ \mathbf{y}_n\,|^{2}\ \mathbf{y}_n \tag{10}$$

② $a = 0.8$ for $k_i \geqslant 20$ (i.e, $\mathbf{y}_n$ is super-Gaussian whose kurtosis value is more than 20.):

$$f(\mathbf{y}_n) = |\ \mathbf{y}_n\,|^{a-1}\ sign(\mathbf{y}_n) \tag{11}$$

③ $a = 1$ for $0 < k_i < 20$ (i.e, $\mathbf{y}_n$ is super-Gaussian whose kurtosis value is less than 20.) :

$$f(\mathbf{y}_n) = \mathbf{y}_n / |\ \mathbf{y}_n\,| \tag{12}$$

where k_i is the estimated kurtosis value:

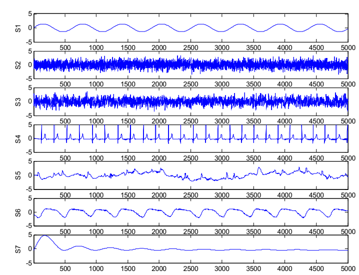
$$k_i = E[\mathbf{y}_i^{4}] / E^{2}[\mathbf{y}_i^{2}] - 3 \tag{13}$$

Note the above FlexibleOgICA batch iterative learning algorithm (7)-(12) can be directly transformed into an online FlexibleOgICA algorithm. In addition, it is clear to observe that, in contrast to other well-known on line or batch-type algorithms [2][3][5], the proposed FlexibleOgICA algorithm does not stem from first-order nor second-order (Newton-like) cost-function optimization and is expressed in a fully-multiplicative fashion in this paper. As a result, the FlexibleOgICA has fast convergence speed and high separation performance.

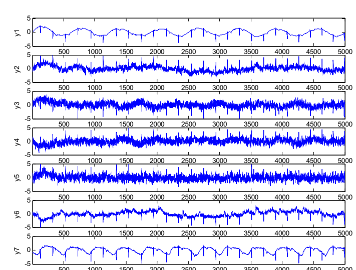
III. SIMULATIONS AND EXPERIMENTS

In the first simulation, we used the ABio7 data, which could be found in [12], shown in Fig. 1(a). Each signal had 5000 samples. Three signals s_1, s_5, s_6 were sub-Gaussian, two signal s_4, s_7 are super-Gaussian, and the others s_2, s_3 are Gaussian noisy. Their normalized kurtosis values were, respectively, -1.5, 0.04, 0.01, 37, -0.5, -1.05 and 11. On the other hand, we assumed the simplest choice for the real, invertible, positive-define diagonal matrix $\mathbf{B}=\mathbf{I}_N$ in the proposed algorithm [8]. All of the algorithms were offline versions. Their parameters were adjusted so that they obtained the best averaged performance.

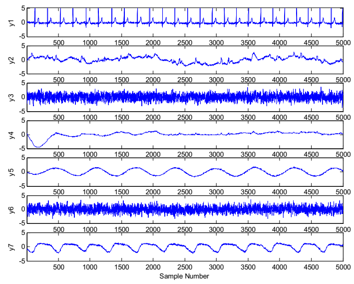
These source signals were randomly mixed. After whitening the mixed signals, we ran the original OgICA algorithm [4], the extended Infomax (ExtICA) algorithm [3] and the proposed FlexibleOgICA algorithm in this paper. The results are shown in Fig.1(b), Fig.1(c) and Fig.1(d), from which it is clear to see that the original OgICA algorithm in [4] can not separate the mixture of sub- and super-Gaussian because the algorithm itself was designed to separate the mixture of sub-Gaussian, while the other two algorithms correctly separated the seven source signals.



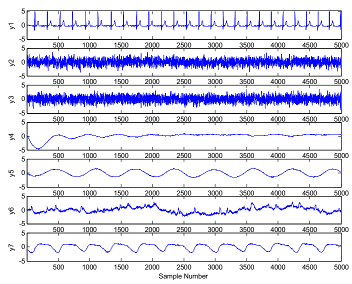
(a)



(b)



(c)



(d)

Fig.1 Simulation on artificially generated data. (a) The seven source signals. (b) The separated signals by the original OgICA algorithm in [4]. (c) The separated signals by the ExtICA algorithm in [3]. (d) The separated signals by the proposed FlexibleOgICA algorithm in this paper.

To evaluate the separation performance of the ExtICA algorithm in [3] and the proposed FlexibleOgICA algorithm in this paper that separated the correct signals, we adopted the following measure:

$$PI = \frac{1}{N-1}\left(\sum_{i=1}^{N}\frac{e_i^2}{\max_i e_i^2} - 1\right) \qquad (14)$$

where $\mathbf{e} = \mathbf{W}^T\mathbf{V}\mathbf{A} = [e_1,\ldots,e_N]$, $\mathbf{V}$ is the whitening matrix and $\mathbf{A}$ is the mixing matrix. The performance metric has the following features: (1) PI lies in [0,1] for any vector $\mathbf{e}$; (2) PI =1 if and only if $e_i^2=e_j^2$ for all i, j in the range [1, N]; (3) PI =0 if and only if $\mathbf{e}$ has only one non-zero element. When perfect signal separation is carried out, the performance index PI is zero. In practice, it is a very small number. The lower PI was, the better the performance was. The averaged performance over 200 independent trials of the original OgICA algorithm [4], the ExtICA algorithm in [3] and of the proposed FlexibleOgICA algorithm in this paper was shown in Fig.2. Clearly, it is shown that the original OgICA algorithm cannot separate the mixture of sub- and super-Gaussian signals as the performance index PI of the algorithm [4] is larger (here $PI \approx 0.55$). Furthermore, one can observe that the proposed FlexibleOgICA algorithm gives better separation performance and faster convergence speed (only 20 iterations are needed to achieve convergence) than the ExtICA algorithm in [3]. Faster convergence might be due to the fully-multiplicative orthogonal-group adopted in this paper. Also, the results of the FlexibleOgICA algorithm keep stable over time, showing that the algorithm didn't exhibit post-separation oscillations.

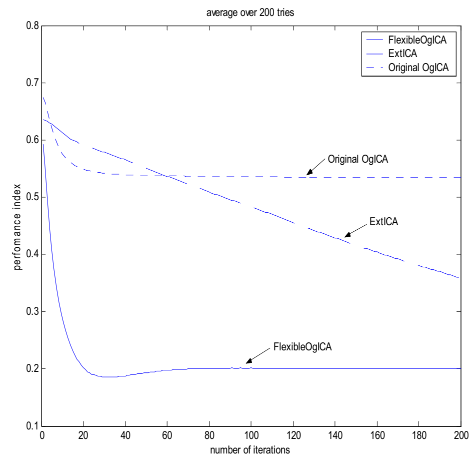
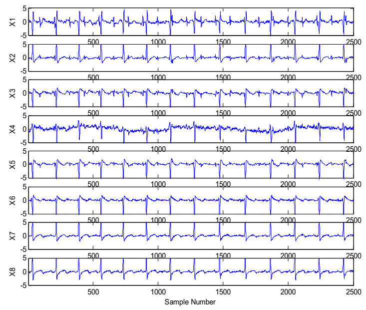


Fig.2 Averaged performance index over 200 independent trials: real line is for the proposed FlexibleOgICA algorithm in this paper, dashed line is for the ExtICA algorithm in [3], and dot line is for the original OgICA in [4].

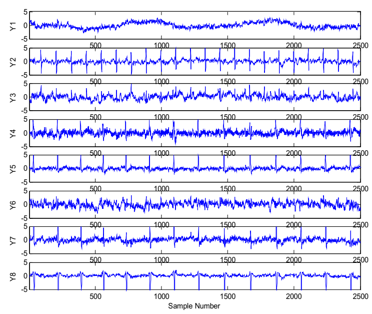
Next we used the real-world ECG data [13] (Fig. 3(a)). Our goal was to obtain the fetal ECG (FECG) signal by the proposed FlexibleOgICA algorithm in this paper, which was very weak and almost only visible in x_1. The result was shown in Fig. 3(b). One can observe that the weak FECG was well separated by the proposed FlexibleOgICA algorithm in this paper. Note the second node output signal y_2 corresponds to the FECG signal separated by the algorithm in Fig. 3(b). The rest of separated signals might be breathing artifact, the mother's ECG (MECG) and lots of noise.

IV. CONCLUSION

In this paper we propose a FlexibleOgICA algorithm, which not only can instantaneously separate the mixture of sub-Gaussian and super-Gaussian source signals, but also can separate super-Gaussian whose kurtosis value lies in a specific range. It converges quickly (only about 20 iterations are needed to achieve convergence), due to the use of full-multiplicative orthogonal-group. Furthermore, the algorithm has good separation performance because of adopting self-adaptive nonlinear function that controlled by a single parameter (Gaussian exponent a) according to the estimated kurtosis value of separating filter output. The validity and performance of the algorithms are confirmed by extensive computer simulations and experiments on real-world data.



(a)



(b)

Fig.3 Simulation on real-world ECG data. (a) The ECG data. (b) The separated FECG signal y_2 by the proposed FlexibleOgICA algorithm

V. ACKNOWLEDGEMENTS

The work was supported by the National Natural Science Foundation of China (Grant No. 60673142).

REFERENCES

[1] HyvÄarinen, A., Karhunen, J., Oja, E., Independent component analysis, NewYork: John Wiley & Sons, 2001

[2] Cichocki, A. and Amari, S., Adaptive Blind Signal and Image Processing, John Wiley & Sons, Inc., New York, 2003

[3] T.-W.Lee, M.Girolami, T.J.Sejnowski. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian source. Neural Compution, 1999, 11: 417-441

[4] FIORI, S. fully-multiplicative orthogonal-group ICA neural algorithm. Electronics Letters, 27th November 2003, Vol.39, No.24.

[5] Seungjin Choi, Andrzej Cichocki, Shunichi Amari. Flexible Independent Component Analysis. Neural Networks for Signal Processing VIII, 1998. Proceedings of the 1998 IEEE Signal Processing Society Workshop 31 Aug.-2 Sept. 1998 Page(s):83 – 92

[6] FIORI, S.: 'A theory for learning by weight flow on Stiefel-Grassman manifold', Neural Comput., 2001, 13, (7), pp. 16251647

[7] Cardoso.J.-F. Unsupervised adaptive filtering. IN S.Haykin (Ed.), Entropic contrasts for source separation. Englewood Cliffs, NJ:prentice Hall.

[8] Yalan Ye, Zhi-Lin Zhang, Shaozhi Wu, Xiaobin Zhou. Improved Multiplicative Orthogonal-Group based ICA for Separating Mixed Sub- and Super-Gaussian Sources. The IEEE International Conference on Communications, Circuits and Systems Proceedings. pp. 340-343

[9] Zhi-Lin Zhang, Zhang Yi, Extraction of a source signal whose kurtosis value lies in a specific range, Neurocomputing Volume: 69, Issue: 7-9, March, 2006, pp. 900-904

[10] Zhi-Lin Zhang, Zhang Yi, Extraction of temporally correlated sources with its application to non-invasive fetal electrocardiogram extraction. Neurocomputing Volume: 69, Issue: 7-9, March, 2006, pp. 894-899

[11] Zhi-Lin Zhang, Yalan Ye, Extended Barros's extraction algorithm with its application in fetal ECG extraction, The 2005 IEEE International Conference on Neural Networks and Brain, Beijing, 2005, pp.1077-1080

[12] Cichocki, A., Amari, S., Siwek, K., Tanaka, T. et al., ICALAB Toolboxes, http://www.bsp.brain.riken.jp/ICALAB

[13] D. De Moor(Ed.) Daisy: database for the identification of systems, available online at: (http://www.esat.kuleuven.ac.be/sista/daisy)