Paper

Extraction of Fetal Electrocardiogram from Cutaneous Potential Recordings of a Pregnant Woman Using Time Delayed Decorrelation and Nonlinear Noise Reduction

Non-member Naoki Tanaka (Hitachi, Ltd.)
Non-member Masayoshi Naito (Kagoshima University)
Member Kazutomo Yunokuchi (Kagoshima University)

The fetal electrocardiogram is extracted from cutaneous potential recordings of a pregnant woman. The cutaneous potential recordings are noisy mixtures of fetal and maternal source signals. To separate these signals we use time delayed decorrelation, which is an algorithm for independent component analysis. The separated signals are noisy source signals. In order to reduce the noise, we apply nonlinear noise reduction method to the separated fetal component. The extraction of fetal component is sufficiently successful although the fetal signals are much weaker than the maternal ones.

Keywords: Electrocardiogram, Signal Separation, Independent Component Analysis, Nonlinear Noise Reduction

1. Introduction

In biomedical fields, noninvasive measurements are most desirable. The measurement of the fetal cardiogram is one of typical examples. When we obtain a cutaneous potential recording of a pregnant woman, the signal consists of a strong maternal component and a weak fetal one. Then the question is how to extract the fetal component.

In the present paper, we tackle this problem by combining nonlinear time series analysis (1) (2) with independent component analysis (ICA) (3) (4) on multi-probed recordings. When multi-channel-recorded signals can be regarded as linear superposition of several source signals, the estimation of the source signals based on the observed ones is called blind signal separation (5).

ICA is a method for solving blind signal separation problems under the assumption that source signals are independent each other. In usual cases, ICA problems are supposed to meet the condition that the number of observed mixture signals is greater than or equal to that of independent components in the mixture. However, this condition is not met usually, because the contamination of noise during measurement cannot be avoided. Each mixture signal includes as many components of noise as the number of sensors. Although several researchers have tried to tackle this problem (4), ICA method for noisy mixture has not been established yet.

We first show that the separated signals can be regarded as source signals contaminated with noise due to measurement when ordinary ICA method is formally applied to noisy mixtures. Although the noise components included in the formal solution of ICA are no longer independent each other, the formal solution can be regarded as an approximate one when the amplitude of noise is small. The noise contained in the separated components is reduced by using nonlinear noise reduction (2) (6). Most of biomedical signals have considerable nonlinearity, and hence the conventional linear method for noise reduction does not always work well.

Other approaches to extracting fetal ECG have been reported by several authors. A sophisticated multidimensional ICA approach was proposed by Cardoso (7). The obtained results are very nice, but the procedure is rather complicated. Another important approach is a full nonlinear method, a locally linear projection, which was proposed by Schreiber and Kaplan (8). Their method is based on the determinism appearing in the maternal ECG, and is not for multi-channel recorded signals. Our approach is a hybrid of linear and nonlinear ones, and is simple and efficient.

In the next section we describe our method and data for the present analysis. In section 3 we first discuss the validity of our approach based on the results for simple numerical data, and then present the results of the fetal ECG extraction.

2. Method and Data for Analysis

2.1 Estimation of source signals and mixing matrix based on ICA Suppose that we have obtained time series data with length of T from N sensors:

$$\boldsymbol{x}(1), \boldsymbol{x}(2), \cdots, \boldsymbol{x}(t), \cdots, \boldsymbol{x}(T-1), \boldsymbol{x}(T), \cdots (1)$$

where $\boldsymbol{x}(t) \in R^N$. We assume that the series is generated by another underlying time series

$$\boldsymbol{s}(1), \boldsymbol{s}(2), \cdots, \boldsymbol{s}(t), \cdots, \boldsymbol{s}(T-1), \boldsymbol{s}(T), \cdots (2)$$

components of which are independent each other, and that $\boldsymbol{x}(t)$ is described as a linear mixture of M independent components of $\boldsymbol{s}(t)$ $(\in R^M)$,

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t). \quad (3)$$

ICA provides the source signals $\{\boldsymbol{s}(t)\}$ and the mixing matrix $\boldsymbol{A}$ from the observed mixture signals $\{\boldsymbol{x}(t)\}$ based on the statistical independency of each component of source signals $\boldsymbol{s}(t)$. Actually, the time series resulting from ICA is not always $\{\boldsymbol{s}(t)\}$ itself. There are uncertainties in both the order of the independent components and a constant factor, and what we get is a series of

$$\boldsymbol{u}(t) = \boldsymbol{P}\boldsymbol{D}\boldsymbol{s}(t), \quad (4)$$

where $\boldsymbol{P}$ is a permutation matrix and $\boldsymbol{D}$ is a diagonal matrix. From Eqs. (3) and (4), the time series $\{\boldsymbol{u}(t)\}$ is connected with $\{\boldsymbol{x}(t)\}$ by a relation

$$\boldsymbol{u}(t) = \boldsymbol{W}\boldsymbol{x}(t), \quad (5)$$

where the matrix $\boldsymbol{W}$ is given by

$$\boldsymbol{W} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{A}^{-1}, \quad (6)$$

with $\boldsymbol{A}^{-1}$ being generalized inverse matrix of $\boldsymbol{A}$. The generalized inverse of $\boldsymbol{W}$, $\boldsymbol{W}^{-1}$, has the meaning of effective mixing matrix.

In the ICA, we determine the matrix $\boldsymbol{W}$ so that it makes the components of $\boldsymbol{u}(t)$ to be independent each other. To solve this problem, we use the time delayed decorrelation (TDD) algorithm [9][10], which imposes the linear independence between different components of $\boldsymbol{u}(t)$ and $\boldsymbol{u}(t+\tau)$ at several values of τ. The key step is to diagonalize several correlation matrices simultaneously. This diagonalization cannot be generally realized by any matrix operation. Instead, we minimize the off-diagonal elements of these correlation matrices:

$$\sum_{i \neq j} \langle u_i(t) u_j(t) \rangle^2 + \sum_{k=1}^{K} \sum_{i \neq j} \langle u_i(t) u_j(t+\tau_k) \rangle^2, \quad (7)$$

where $\{\tau_k\}$ are K different positive numbers. In this paper we set τ_k and K to be $5k\delta t$ and 20, respectively, where δt denotes sampling interval. As shown by Cardoso, [11] this minimization can be realized by diagonalizing 2×2 matrices† $N(N-1)/2$ times by using Jacobi method. By this procedure, we obtain $\boldsymbol{W}$ and $\boldsymbol{u}(t)$.

Letting observational noise to be $\boldsymbol{\eta}(t)$, the observed signals are given by

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{\eta}(t) \quad (8)$$

We can rewrite Eq. (8) as

$$\boldsymbol{x}(t) = \boldsymbol{A}\left\{\boldsymbol{s}(t) + \boldsymbol{A}^{-1}\boldsymbol{\eta}(t)\right\}. \quad (9)$$

Even though the components of the noise $\boldsymbol{\eta}(t)$ are independent each other, the components of $\boldsymbol{A}^{-1}\eta(t)$ are not independent because the matrix $\boldsymbol{A}^{-1}$ mixes the components of $\boldsymbol{\eta}(t)$. By this reason the solution obtained using ICA is not an exact solution. However, when $\|\boldsymbol{\eta}(t)\|$ is small, the solution can be regarded as an approximate one.

†Cardoso gave explicit form of 3×3 matrix to be diagonalized for a complex matrix. The expression for a real matrix can be easily derived from Eqs. (4) and (5) in Ref.(11).

2.2 Method for noise reduction The separated signals include computational artifacts due to ICA treatment in addition to the measurement noise. To reduce these noise and artifacts we use a nonlinear noise-reduction method proposed by Schreiber [6].

Suppose that we have obtained, by ICA, a noisy separated signal $y(t)$:

$$y(t) = u(t) + \eta'(t), \quad (10)$$

where $\eta'(t)$ denotes the noise component remaining in $y(t)$.

In Schreiber's method [6] we first reconstruct the state space from the observations $\{y(t)\}$ using a $2m+1$-dimensional embedding. The point $\boldsymbol{y}(t)$ corresponding to the observation $y(t)$ is given by

$$\boldsymbol{y}(t) = (y(t-m\tau), \cdots, y(t-\tau), y(t), y(t+\tau), \cdots, y(t+m\tau)). \quad (11)$$

Letting $\mathcal{N}_t(\epsilon)$ denote ϵ-neighborhood around the point $\boldsymbol{y}(t)$, we obtain the estimate $\hat{u}(t)$ as the central coordinate component of the weighted center of the neighborhood $\mathcal{N}_t(\epsilon)$:

$$\hat{u}(t) = \frac{1}{|\mathcal{N}_t(\epsilon)|} \sum_{\boldsymbol{y}(k) \in \mathcal{N}_t(\epsilon)} y(k), \quad (12)$$

where $|\mathcal{N}_t(\epsilon)|$ denotes the number of the points contained in the neighborhood $\mathcal{N}_t(\epsilon)$.

Unlike the noncausal moving average method this method can avoid degradation of the peak signals due to smoothing because this method averages the data locating near each other instead of the successive data. The method works well even when the noise amplitude is large or the number of data is relatively small. We below call this method phase space local averaging method or simply local averaging method.

The chronological order of applying ICA treatment and the noise reduction is important. One may think that the true separated signals are obtained by applying a noise reduction method to the observed mixture signals before ICA treatment. Mathematically, of course, this is the case. In most cases, however, the separated signals include undesirable computational artifacts due to ICA treatment. In addition it is difficult to determine the values of parameters for noise reduction before ICA treatment because it should be determined based on the relation between the amplitude of noise (and artifacts) and the signals to be separated. We believe that it is better to reduce the noise present in the separated signals after ICA treatment.

2.3 Data series for analysis

2.3.1 Numerical data We first use numerical data for $M = N = 2$ case in order to study the effects of observational noise on the results of ICA and appropriate values of parameter ϵ used in the noise reduction scheme. Source signals are given by

$$\begin{cases} s_1(t) = \sin 112\pi t \\ s_2(t) = \cos 214\pi t, \end{cases} \quad (13)$$

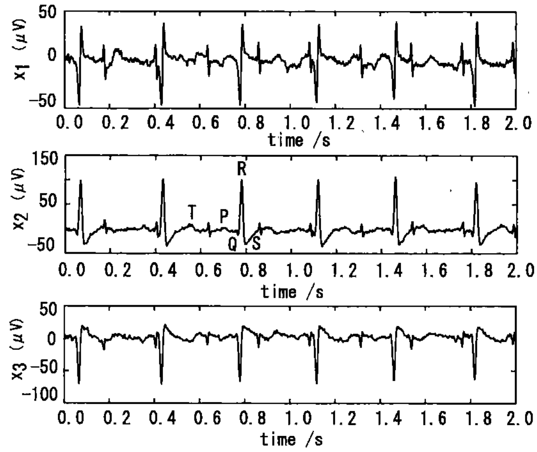

Fig. 1. Cutaneous potential recordings of a pregnant woman. x_1, x_2, x_3: data measured at different three points on the abdomen.

and the observed signals were generated from Eq. (8) using mixing coefficient matrix $\boldsymbol{A}$:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0.85 \\ 0.55 & 1 \end{bmatrix} \cdots\cdots\cdots \quad (14)$$

We used independent uniformly distributed random number as noise $\eta(t)$ and set the sampling frequency to be 4 kHz within range $t \in [0, 1]$.

2.3.2 Cutaneous potential recordings of a pregnant woman Our aim in the present paper is to extract fetal ECG from cutaneous potential recordings of a pregnant woman. The used data were downloaded from the website of DaISy [12]. The cutaneous potential was recorded at 8 points (5 points on the abdomen and 3 points on the thorax) with 500 Hz sampling for 5 seconds. We used three of the abdominal recordings for the extraction. We show them in Fig. 1. Electrocardiograms generally consist of major three structures, P-wave, QRS complex, and T-wave, which correspond to the sequential activation of the arteria, the ventricular depolarization, and ventricular reporlarization, respectively. Among them QRS complex has particularly large signal amplitude.

Although the fetal signals are maximized by positioning the electrode on the abdomen, the maternal component is still dominant in each recording. The aim of our analysis is to obtain the fetal electrocardiogram after reducing the noise included in the recordings and the artifacts generated by ICA treatment.

3. Results and Discussion

3.1 Noise effects on ICA and noise reduction

Before presenting the extraction of the fetal ECG, we show the results of ICA for a case where some amount of observational noise is superimposed on the mixture of source signals (section 2.3.1). We set the noise amplitude to be 10 % of the mixture signals. Figures 2(a) and (b) denote the source signals, Figures 2(c) and (d) denote the noisy mixture signals, and Figures 2(e) and (f) denote the independent components obtained by TDD algorithm. When we obtained the results in Figs. 2(e) and (f), we adjusted the signal amplitude based on the values of the diagonal elements in the estimated mixing matrix so that the amplitudes of both source and separated signals be equal. The estimate of the mixing matrix was then given by

$$\hat{\boldsymbol{W}}^{-1} = \begin{bmatrix} 1.000 & 0.871 \\ 0.531 & 1.000 \end{bmatrix} \cdots\cdots\cdots \quad (15)$$

The noise amplitudes of the separated signals were larger than those of the noisy mixtures in this case.

We applied the nonlinear noise reduction method to the separated signals shown in Fig. 2(e) and (f). The

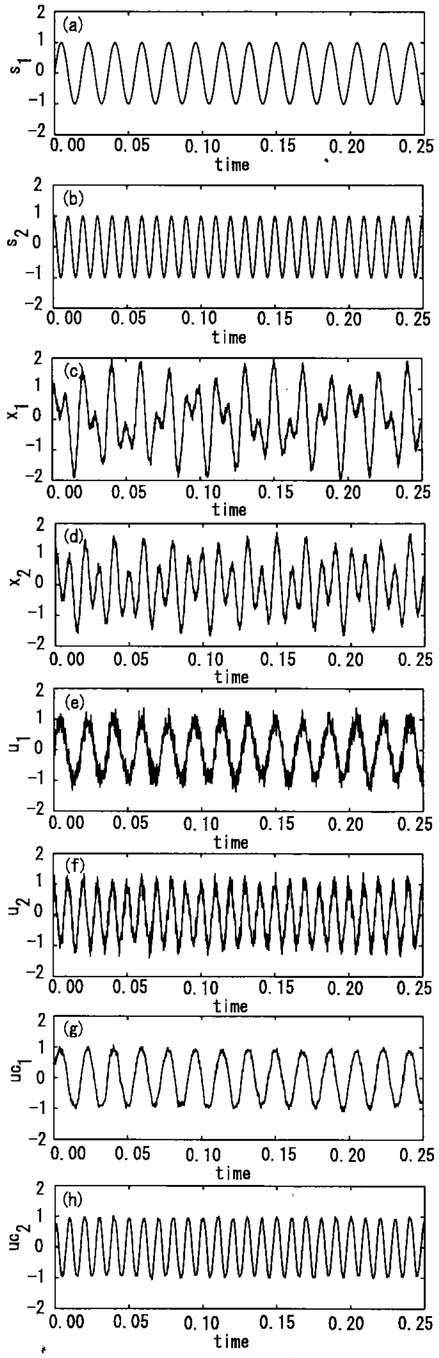

Fig. 2. Results of ICA for a noisy mixture of synthesized source signals. s_1, s_2: source signals, x_1, x_2: noisy mixture signals, u_1, u_2: separated signals, uc_1, uc_2: denoised separated signals.

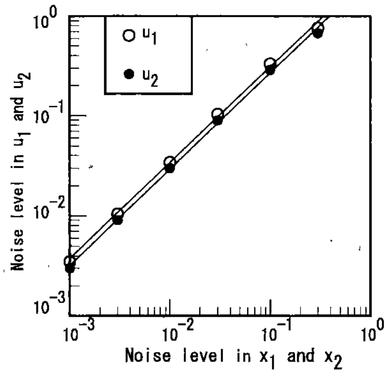

Fig. 3. Relationship between the noise levels in the mixture signals and in the separated signals.

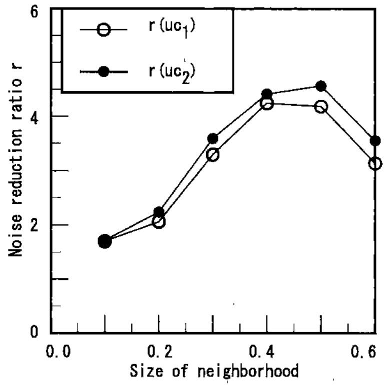

Fig. 4. Dependence of the noise reduction ratio r on the size of neighborhood ϵ for the separated signals.

results are shown in Figs. 2(g) and (h). To obtain these results we set ϵ to be 0.4. It is found in Figs. 2(g) and (h) that the nonlinear noise reduction method works well for this case.

We study the relationship between the amplitude of the noise added to each sensor and that of noise present in the separated signals. The results are shown in Fig. 3. In double logarithm plot, the relationship between them is plotted as a straight line with slope about one, which shows that the amplitude of the noise present in the separated signal is proportional to that of the noise added to the sensors. This shows that the distortion of the separated signals is small even for relatively strong noise η. Hence, the separated signals correspond to the source signals contaminated with noise.

In the nonlinear noise reduction process, the choice of a radius ϵ for neighborhood affects the reduction ratio of noise significantly. Letting the noise amplitude before and after reduction be σ_{before} and σ_{after}, respectively, the reduction ratio r of noise can be described as $r = \sigma_{\text{before}}/\sigma_{\text{after}}$. We calculated the dependence of r on ϵ for u_1 and u_2 shown in Fig. 3. The results for a constant $\sigma_{\text{before}}(= 0.1)$ are shown in Fig. 4. The figure shows that it is important to set the neighborhood radius ϵ to be several times as large as the noise amplitude σ_{before}. The appropriate value of ϵ is discussed also in ref. 6. In ref. 6 it is recommended that the value of ϵ be a few times as large as the noise amplitude. This noise reduction method works when the signal amplitude is sufficiently larger than the appropriate value of ϵ.

3.2 Extraction of fetal electrocardiogram from cutaneous potential recordings of a pregnant woman We show the results of ICA for the cutaneous potential recordings (Fig. 1) in Figs. 5 and 6. Figure 5 shows the results with TDD algorithm. Judging from higher heart beat rate in the component u_3 than in other two, the component u_3 corresponds to the fetal ECG. For comparison we show the results with the infomax algorithm [13][14] in Fig. 6. The infomax algorithm separates independent components from mixture signals based on the minimization of mutual information among the separated component signals. As for the learning rule we modified the rule proposed by Lee et al. [14].

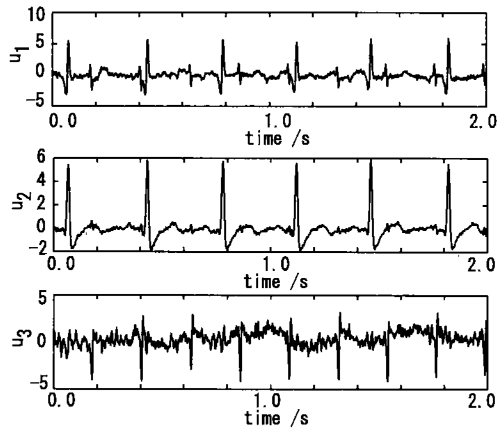

Fig. 5. Results of ICA with TDD algorithm for cutaneous potential recordings of a pregnant woman.

Comparing the results, the TDD algorithm separates the fetal ECG components more efficiently than the infomax algorithm. In ECG, the dynamics underlying signals is clear, and hence the signals have a definite structure in the time domain. The TDD algorithm conserves this information and so TDD seems to have led to better extraction of the fetal ECG from the cutaneous potential recordings.

The separated signal, u_3, corresponding to the fetal ECG (in Fig. 5) appears to include relatively strong noise. To reduce noisy character we applied the phase space local averaging method to the signal. The results are shown in Fig. 7. Figure 8 denotes phase space plots related to these results. Figure 7(a) shows the separated signal. The corresponding two-dimensional phase space plot is shown in Fig. 8(a). The large loop appearing in Fig. 8(a) corresponds to QRS complex. The center part of the orbits, where many state points locate, covers small loops corresponding P and T waves. Figure 7(b) and (c) denote the denoised signal and the removed noise, respectively. To obtain these results we set the dimension of reconstructed phase space at 9 ($m = 4$) and neighborhood diameter ϵ at 20 % of orbit diameter. It is found that the noisy component has been considerably removed, and the degradation of signals (e.g. QRS complex) is not recognizable. Figure 8(b)

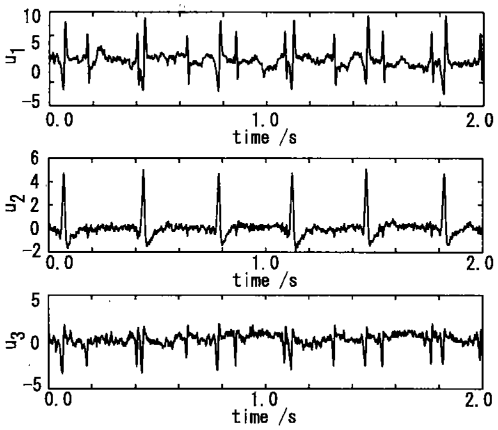

Fig. 6. Results of ICA with infomax algorithm for cutaneous potential recordings of a pregnant woman.

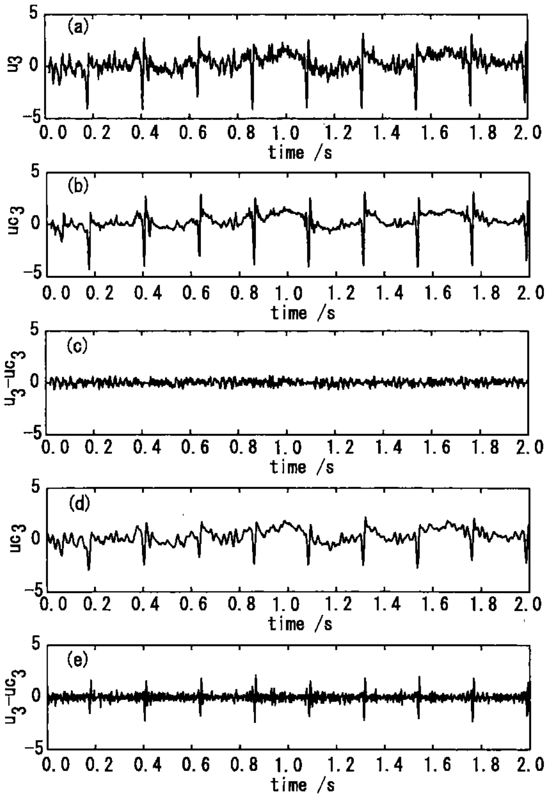

Fig. 7. Results of noise reduction for fetal component. (b)result with the present method. (d)result with the non-causal moving average.

denotes the corresponding orbits in the reconstructed state space, where the center part of high point density has shrunk and the main loop structure corresponding to QRS complex has been clarified.

For comparison, we show the result of smoothing by a non-causal moving average filter (window length = 5) in Fig. 7(d). Figure 7(e) denotes the noise removed by this smoothing. Comparing with the result presented in Fig. 7(c), we find that the signal is considerably degraded. In this smoothing, the QRS spikes are partly subtracted as noise.

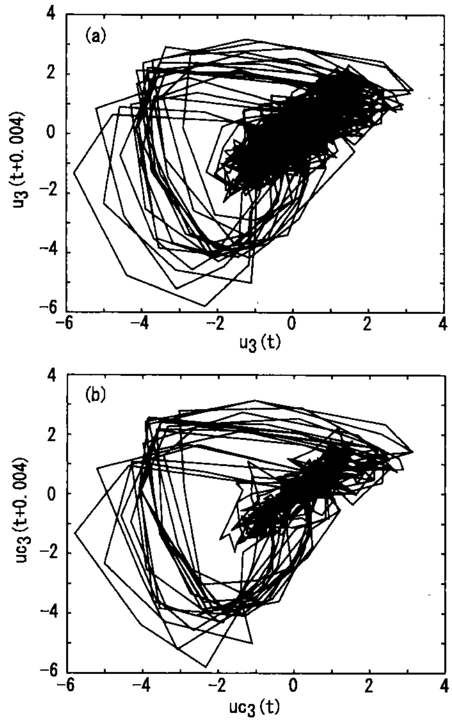

Fig. 8. Phase space plots related to the results presented in Fig. 7.

4. Conclusion

We presented the results of the extraction of fetal ECG from cutaneous potential recordings of a pregnant woman. Our approach to this extraction is combining nonlinear noise reduction with time delayed decorrelation. Although the fetal signals were much weaker than the maternal ones, it was able to extract the fetal component with little degradation. The approach we took in this paper is expected to be valid for other biomedical data obtained from multi-channel observations, because many of them can be regarded as noisy mixtures of nonlinear source signals contaminated with measurement noise.

(Manuscript received December 25, 2001, revised April 3, 2002)

References

(1) H. D. I. Abarbanel, R. Brown, J. J. Sidorowich and L. Sh. Tsimring: The analysis of observed chaotic data in physical systems, *Rev. Mod. Phys.*, **65**, 1331-1392 (1993).

(2) H. Kantz and T. Schreiber: Nonlinear Time Series Analysis, Cambridge Univ. Press, Cambridge, (1997).

(3) T.-W. Lee: Independent Component Analysis Theory and Applications, Kluwer, Boston (1998).

(4) M. Girolami (ed): Advances in Independent Component Analysis, Springer, London (2000).

(5) J.- F. Cardoso: Blind signal separation: statistical principles, *Proc. IEEE*, **9**, 2009-2025 (1998).

(6) T. Schreiber: Extremely simple nonlinear noise-reduction method, *Phys. Rev. E*, **47**, 2401-2404 (1993).

(7) J.- F. Cardoso: Multidimensional independent component analysis, *Proc. ICASSP'98* (1998).

(8) T. Schreiber and D. T. Kaplan: Signal separation by nonlinear projection: The fetal electrocardiogram, *Phys. Rev. E*, **53**, R4326-R4329 (1996).

(9) L. Molgedy and H. G. Schuster: Separation of a mixture of independent signals using time delayed correlations, *Phys. Rev. Lett.*, **72**, 3634-3637 (1994).
(10) A. Ziehe and K.-R. Mueller: TDSEP – an efficient algorithm for blind separation using time structure, *Proc. ICANN98*, 675-680 (1998).
(11) J.- F. Cardoso and A. Souloumiac: Jacobi angles for simultaneous diagonalization, *SIAM J. Matrix Anal. Appl.*, **17**, 161-164 (1996).
(12) De Moor B. L. D. (ed.): DaISy: Database for the Identification of Systems, Dept. of Electrical Engineering, ESTA/SISTA, K. L. Leuven, Belgium. URL: //www.esat.kuleuven.ac.be/sista/daisy/.
(13) A. J. Bell and T. J. Sejnowski: An information-maximization approach to blind separation and blind deconvolution, Neural Computation, **7**, 1129-1159 (1995).
(14) T.-W. Lee, M. Girolami and T. J. Sejnowski: Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian super-Gaussian sources, *Neural Computation*, **11**, 409-433 (1999).

Naoki Tanaka (Non-member) has worked for Hitachi, Ltd. since 1982. He received a Ph.D. degree in Engineering from University of Electro-communications in 2000. He is a senior research scientist at Advanced Research Laboratory of Hitachi, Ltd., and a visiting professor at the graduate school of Tokyo Institute of Technology. His research interest lies in nonlinear dynamical systems theory and its applications. He is a member of The Japan Society of Applied Physics and The Institute of Electronics, Information and Communication Engineers.

Masayoshi Naito (Non-member) received a Ph.D. degree in Physics from Tokyo Institute of Technology in 1972. After having worked for Hitachi, Ltd., he is currently professor at Department of Bioengineering, Kagoshima University. His research field covers nonlinear dynamics, nonlinear time series analysis, complex systems and neuronal information processing. He is a member of The Physical Society of Japan, The Japan Society of Applied Physics, The Biophysical Society of Japan, and The Institute of Electronics, Information and Communication Engineers.

Kazutomo Yunokuchi (Member) He worked at Francis Bitter Magnet Laboratory in M.I.T. as a visiting researcher from 1989 to 1991. He received Dr. Engineering degree from The University of Tokyo in 1993. He is a professor of Department of Bioengineering, Kagoshima University. His research interest concentrates on the information system of human body. He is a member of IEEE, Institute of Electronics, Information and Communication Engineers, Institute of Electrical Engineers of Japan, and The Japan Society of Medical Engineering.