Turk J Elec Eng & Comp Sci, Vol.19, No.4, 2011, © TÜBİTAK
doi:10.3906/elk-0912-311

Performance evaluation of nonparametric ICA algorithm for fetal ECG extraction

Yusuf SEVİM*, **Ayten ATASOY**
Department of Electrical and Electronics Engineering, Engineering Faculty,
Karadeniz Technical University, 61080 Trabzon-TURKEY
e-mails: ysevim@ktu.edu.tr, ayten@ktu.edu.tr

Received: 16.12.2009

Abstract

Fetal electrocardiograms (FECG) contain important indications about the health and condition of the fetus. In this respect, it is crucial to apply a robust algorithm to ECG data for extraction of the FECG signal. Most of the independent component analysis (ICA) algorithms used for this purpose rely on simple statistical models. Such algorithms can fail to separate desired signals when the assumed statistical model is inaccurate. Statistical models can be estimated accurately using kernel density estimation methods. Therefore, the kernel density estimation method was used in this paper for building an ICA algorithm (nonparametric ICA: NpICA) and the algorithm was applied to abdominal recordings to separate the FECG signals, which had not been implemented before. Checking of the separation quality of the NpICA algorithm was applied to synthetic ECG signals and real multichannel ECG recordings obtained from a pregnant woman's skin. The test results showed that the NpICA algorithm outperformed other known ICA algorithms such as FastICA and JADE. The superior performance of the NpICA algorithm was especially evident in recordings with high signal length. This indicates that the NpICA method is more robust than other classical ICA algorithms for FECG extraction.

Key Words: *ECG, FECG, ICA, JADE, FastICA, BSS*

1. Introduction

FECG extraction is an interesting and difficult problem in biomedical engineering. Doppler ultrasounds are routinely used for measurement of fetal heart rate (FHR) during pregnancy and delivery [1]. Nevertheless, FHR variability, an important parameter to assess fetal distress, is the only parameter obtained by the Doppler ultrasound, while classical ECG measurements can provide additional information such as morphological and temporal parameters of the FECG during gestation. These temporal parameters are associated with the level of fetal oxygenation [2]. Another advantage of the FECG is the long-term monitoring of high risk pregnancies at home. The Doppler ultrasound is not suitable for long-term ambulant monitoring due to its high sensitivity

*Corresponding author: Department of Electrical and Electronics Engineering, Engineering Faculty, Karadeniz Technical University, 61080 Trabzon-TURKEY

to movements and the need for frequent repositioning of the ultrasound probe [3,4]. The ultrasound probe is also problematic and uncomfortable, while the procedure involves launching a 2 MHz signal toward the fetus [5]. Nevertheless, besides many advantages, the abdominal ECG measurement also has some disadvantages. One of them is that the FECG is generated from a very small heart, so the signal amplitude is low and noise from electromyographic activity affects the signal due to its low voltage. Another noise source is the maternal ECG (MECG), which can be 5-1000 times higher in terms of intensity and ability to induce surface potentials [6]. The MECG affects all of the electrodes, so there is no place to put an electrode on the mother's skin to receive just the FECG signal without the mother's signal. In each probe, the MECG is higher than the FECG in magnitude. Therefore, extracting the FECG from the recorded signal is very important.

Due to the fact that the MECG, the FECG, and the mixing environment (mother's abdomen) are not known, the FECG extraction can be assumed as a typical blind source separation (BSS) problem. The first application of BSS techniques to FECG extraction was done by De Lathauwer [7]. Many different methods [8-16] have been developed to detect the FECG for abdominal recordings. Pani et al. developed a real time application using the OL-JADE algorithm [17]. Among these algorithms, the most promising algorithms are ICA algorithms, but the standard ICA algorithm still has not been improved for FECG extraction, and statistical modeling of probability density functions (pdfs) is a challenging problem for ICA algorithms. Many existing methods rely on simple assumptions about the source statistics and are characterized by well-assessed convergence and consistency properties [18]. When such hypotheses hold strictly, most conventional ICA algorithms are capable of quickly and efficiently achieving the desired source separation. However, such algorithms can perform suboptimally or even fail to produce the desired source separation when the assumed statistical model is inaccurate [19]. As an example, FastICA [20], which requires the selection of a contrast function according to the pdfs of the sources, can be given. To tackle this problem, the kernel-based model, in which the pdfs are directly estimated from the data using a kernel density estimation technique [21-23], can be used. These methods are called nonparametric ICA algorithms (NpICA) [24,25]. The NpICA simultaneously estimates the unknown pdfs of the source signals and the linear operator that allows the separation of the mixed signals. The resulting algorithm is nonparametric and data-driven, and does not require the definition of a specific model for the density functions.

In this paper, the separation of ECG signals is conducted using ICA approaches. A review of the ICA algorithms is addressed in Section II. The actual algorithmic implementation of the proposed technique is given in Section III. Simulations were conducted to demonstrate the performance improvement obtained with the NpICA, described in Section IV. Section V offers concluding remarks.

2. ICA algorithm

ICA is a mathematical technique for recovering latent source signals from only observed signal mixtures. Depending on the mixing process, ICA can be modeled as linear, nonlinear, or convolutive. In this work, the mixing process is considered as instantaneous linear mixing:

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

where the sources $s = [s_1, ..., s_m]^T$ are mutually independent and A_{mxm} is an unknown mixing matrix. The aim is to find matrix W only from observations, x, such that the output

$$y = Wx \tag{2}$$

is an estimate of the possible sources. Due to the fact that source signals are not known, estimating the demixing matrix W in the closed form is not possible. For this reason, different cost functions have been proposed in the literature [19,26,27]. The solutions to ICA algorithms are found at the minima or maxima of these cost functions. Minimization of these functions, possibly under some constraints on the solutions, is the subject of optimization theory.

3. NpICA algorithm

Most ICA algorithms use mutual information (MI) between the reconstructed signals as a cost function [28]:

$$I\left(y_1, .., y_N\right) = \sum_{i=1}^{N} H(y_i) - H\left(y\right), \tag{3}$$

where $H\left(y_i\right)$ are entropies of the marginal pdfs of the reconstructed signals, and $H\left(y\right)$ is the entropy of the joint pdf of the reconstructed signals. $H\left(y_i\right)$ and $H\left(y\right)$ can also be seen as the lengths of codes for y_i and y, respectively. MI is always nonnegative and equal to zero if the y_i values are independent. This makes the MI function the cost function for ICA algorithms.

For the optimization step, Eq. (3) can be written as:

$$I = \sum_{i=1}^{N} H(y_i) - \log\left|\det W\right| - H\left(x\right), \tag{4}$$

where $\{\log\left|\det W\right| + H\left(x\right)\}$ is the linear transformation of $H\left(y\right)$. $H\left(x\right)$ is the entropy of the observed signals and it is a constant with respect to W, so the objective function is reduced to:

$$L\left(W\right) = \sum_{i=1}^{N} H\left(y_i\right) - \log\left|\det W\right|. \tag{5}$$

This reduction makes the MI calculation very easy, so the estimation of the joint pdf is very difficult for high-dimensional data. Using the entropy definition [29], Eq. (4) can be written as:

$$L\left(W\right) = -\sum_{i=1}^{N} E\left[\log p_{y_i}\left(w_i x\right)\right] - \log\left|\det W\right|, \tag{6}$$

where w_i is the ith row of matrix W. To calculate Eq. (6), the pdfs of the reconstructed signals are needed. Cardoso shows in [30] that incorrect assumptions about pdfs can result in bad estimation or in complete failure for source separation. Donoho and Cardoso further showed that MI is characterized by having minimum asymptotic variance [31], and it is also equivalent to the maximum likelihood principle when the source distributions are known [30,32]. To handle this problem, a nonparametric model was proposed [25], in which the pdfs are directly estimated from the data using a kernel density estimation technique [21,22]. Using this approach, there is no need for separating the optimization step from the reestimation of the cost functions as in [33,34]. Given M sample data, the pdf of the reconstructed signal is:

$$P_{y_i}\left(y_i\right) = \frac{1}{Mh} \sum_{m=1}^{M} \varphi\left(\frac{y_i - Y_{im}}{h}\right), \ i = 1, ..., N, \tag{7}$$

where h is the kernel bandwidth and φ is the kernel. This kernel function must satisfy the condition $\int \varphi(t)dt = 1$. The kernel is usually a symmetric pdf, and generally a standard normal density is used:

$$\varphi(v) \triangleq \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}. \tag{8}$$

From the definition of kernel density estimation, it can be seen that estimation of pdfs inherits all of the properties of the kernel function, such as continuity and differentiability. These properties are very important for the optimization step.

Using the expectation in Eq. (6) for a batch of data of size M, the cost function approximates its ergodic average:

$$L(W) \approx \frac{1}{M} \sum_{i=1}^{N} \sum_{k=1}^{M} \log p_{y_i} \left(w_i x^{(k)} \right) + \log |\det W|, \tag{9}$$

where $x^{(k)}$ is the kth column of the data matrix. Evaluating the estimates of the pdf of the source signals at the data points, and replacing marginal pdfs p_{y_i} with the kernel density, the estimate cost function becomes:

$$L(W) \approx \frac{1}{M} \sum_{i=1}^{N} \sum_{k=1}^{M} \log \left[\frac{1}{Mh} \sum_{m=1}^{M} \varphi \left(\frac{w_i \left(x^{(k)} - x^{(m)} \right)}{h} \right) \right] + \log |\det W|, \tag{10}$$

where parameter h controls the smoothness of the function. A small value of h yields detailed estimation, while a large value of h yields less detailed estimation. Its optimal value is a function of the signal length $\left(h = 1.06 M^{-1/5} \right)$ [21,23], and this equation was used in this study for calculating h.

The final step for ICA algorithms is to minimize the cost functions. In this respect, the BFGS [35] algorithm derived from Newton's method is utilized in optimization. In BFGS, the Hessian matrix of the second derivative of the cost function to be minimized does not need to be computed at any stage. The Hessian matrix is updated by analyzing successive gradient vectors. The BFGS algorithm can be summarized as follows:

Step 1. Given $x_1 \in R^n$, $B_1 \in R^{nxn}$ positive definite, $f(x)$ is a function,
compute $g_1 = \nabla f(x_1)$. If $g_1 = 0$, stop; otherwise, set $k := 1$.

Step 2. Set $d_k = -B_k^{-1} g_k$.

Step 3. Carry out a line search along d_k, getting $\alpha_k > 0$,
$x_{k+1} = x_k + \alpha_k d_k$, and $g_{k+1} = \nabla f(x_{k+1})$.
If $g_{k+1} = 0$, stop.

Step 4. Set $B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k}$,

where

$$s_k = \alpha_k d_k,$$
$$y_k = g_{k+1} - g_k.$$

Step 5. $k := k + 1$; go to Step 2.

4. Performance evaluation

To validate the NpICA algorithm using the nonparametric method as a pdf estimation method, synthetic datasets and a real dataset were adopted as input mixtures.

In order to produce the test data, synthetic datasets were linearly mixed by 1000 randomly generated mixing matrices with signal lengths ranging between 512 and 4608. The source signals and one of the synthetic datasets are shown in Figures 1 and 2, respectively. As a noise source, 1000 Gaussian random signals were used in this synthetic mixing process. The NpICA algorithm was compared with classical FastICA [20,27] and JADE [36,37] algorithms on these datasets. The algorithms were quantitatively compared using only one measure, the median signal-to-noise ratio (SNR) of the separated signals. The equation describing the SNR is:

$$SNR = 10\log_{10}\left(\frac{E\left[s(t)^2\right]}{E\left[n(t)^2\right]}\right) \quad , \tag{11}$$

where $E\left(.\right)$is the mean of the arguments, $s(t)$is the desired ECG signal, and $n(t) = s(t) - \hat{s}(t)$ is the noise, indicating an undesired signal. Here $\hat{s}(t)$ is the estimated source signal, and $s(t)$ and $\hat{s}(t)$are at the same energy.

In Figure 1, the source signals are seen before the mixing process for synthetic datasets. In Figures 1a and 1b, the FECG and the MECG signals are presented with heart beats of 90 and 60, respectively. Figure 1c presents the Gaussian noise source signal; it was changed randomly for every mixing process. In Figure 2, one mixing process is seen, and these signals represent the synthetic ECG recordings.

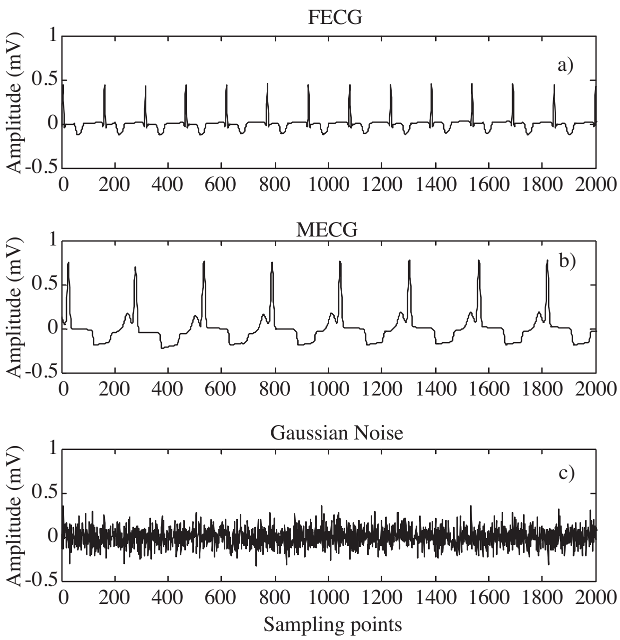

Figure 1. Source signals.

Figures 3 and 4 show that increasing the signal length affects the separation quality of algorithms. When the signal length is altered, the non-Gaussianity of the signal is affected by the alteration. Depending on the increasing of the signal length, kurtosis of the MECG signals ranges between 6.46 and 6.84, and kurtosis of the FECG signals ranges between 13.25 and 14.26. For some signal lengths (e.g. 1536 and 2048 samples for mother, and 3072 sample for fetus) the non-Gaussianity of the signals are larger than the other signal lengths. As a result, the NpICA algorithm is able to separate signals accurately [25]. Another advantage of increasing signal length is the possibility of comparing all overlap conditions between the fetal QRS and maternal QRS.

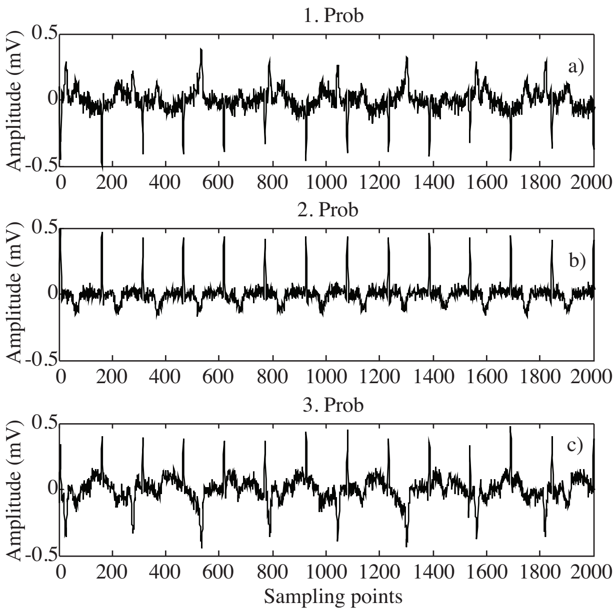

Figure 2. Synthetic ECG recordings.

The results of this experiment are presented in Figures 3 and 4. These results clearly show the performance gain obtained with the NpICA algorithm for different signal lengths.

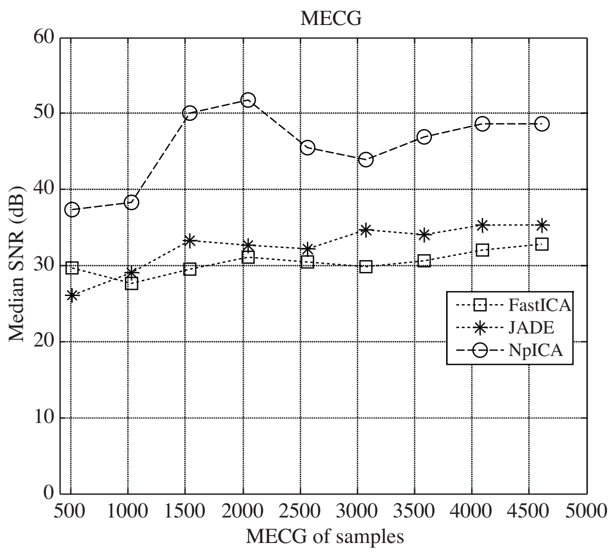

Figure 3. The results of attempting the separation of the MECG signal for various ICA algorithms (averaged over 1000 simulations).

As a second experiment, only the NpICA algorithm was applied on the real cutaneous ECG recordings [7]. These signals were selected among 9 different skin electrodes located on different points of a pregnant woman's skin. In Figure 5, 4 real ECG recordings are seen.

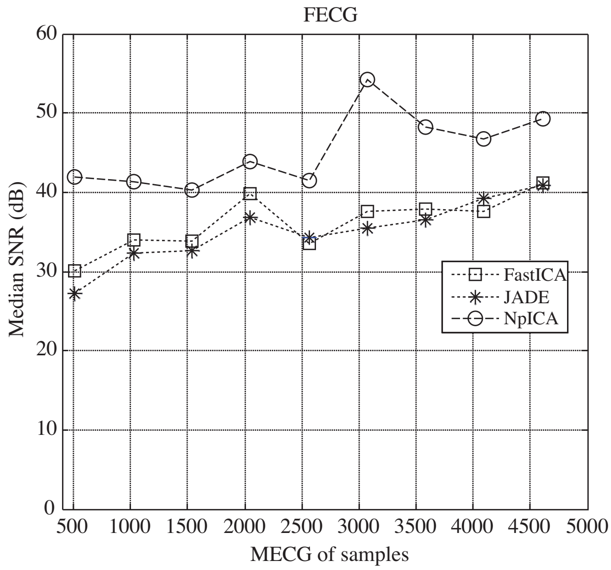

Figure 4. The results of attempting the separation of the FECG signal for various ICA algorithms (averaged over 1000 simulations).

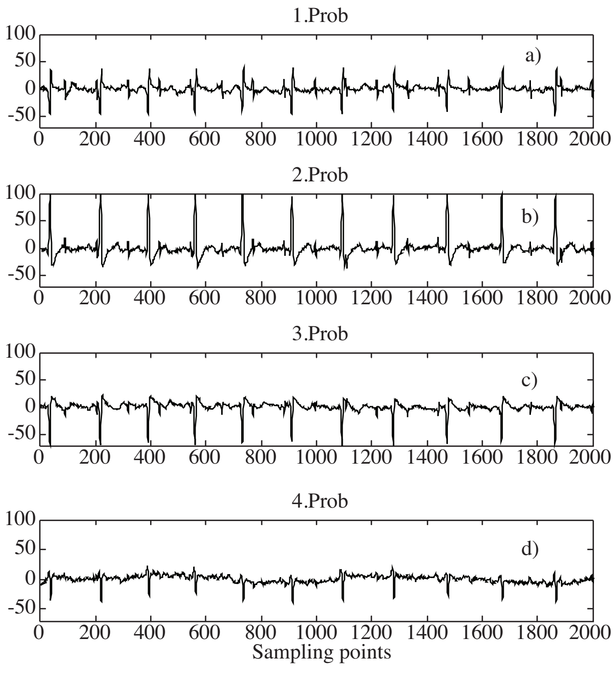

Figure 5. Real cutaneous recordings from a pregnant woman.

Figure 6 illustrates 4 signals obtained after applying the NpICA algorithm. Figure 6a includes signals coming from both the FECG and MECG signals. In Figure 6b, baseline wandering and residual noise can be seen clearly, and the FECG the MECG signals can be noticed in Figures 6c and 6d, respectively.

It is a well known fact that, before the mixing process, the real MECG and FECG are not known. Performance comparison is not possible with the other algorithms by using Eq. (11). It is only possible observing the algorithm results by eye [38,39].

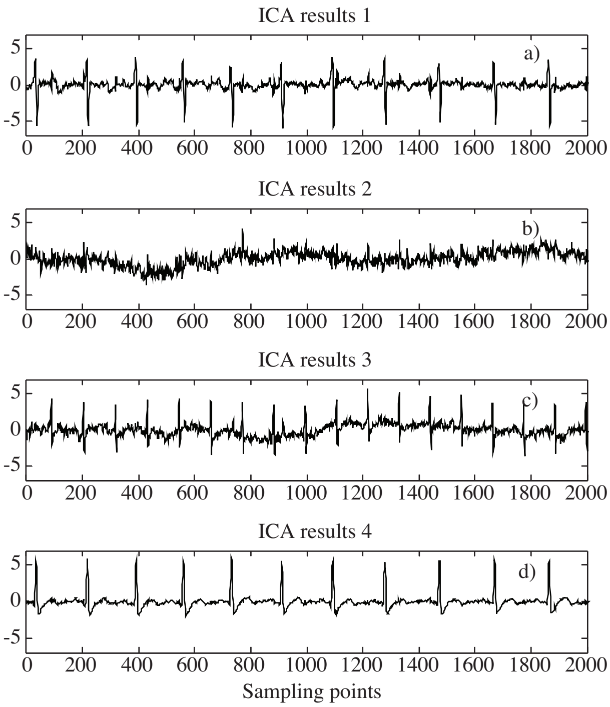

Figure 6. Separated FECG, MECG, and residual noise signals.

5. Conclusions

In this paper, a NpICA algorithm was proposed for FECG separation. Within this study, the results of 2 classical ICA algorithms were compared with the NpICA algorithm results on synthetic datasets, and only the nonparametric method was applied on the real ECG recordings to see the extraction ability of the NpICA algorithm. The comparison of algorithm results is not possible for the real ECG signals. The results of the synthetic datasets demonstrate that the proposed approach is capable of separating the ECG signals, with a noticeable performance improvement when compared to the classical ICA algorithms. The NpICA method has a reliable performance while the classical ICA methods may perform poorly at relatively small signal lengths. Nevertheless, the superior performance is attained at the expense of increased computational complexity and makes it possible to compare all overlap conditions between the fetal QRS and maternal QRS. The achieved FECG separation quality offers promising prospects for the use of the NpICA technique in prenatal medical diagnosis.

Acknowledgments

This study was supported by Karadeniz Technical University Scientific Research Projects Unit, Project No: 2007.112.04.2.

References

[1] M.J. Lewis, “Review of electromagnetic source investigations of the fetal heart”, Med. Eng. Phys., Vol. 25, pp. 801-810, 2003.

[2] E.M. Symonds, D. Sahota, A. Chang, Fetal Electrocardiography, London, UK, Imperial College Press, 2001.

[3] J.A. Crowe, A. Harrison, B.R. Hayes, “The feasibility of long-term fetal heart rate monitoring in the home environment using maternal abdominal electrodes”, Physiol. Meas., Vol. 16, pp. 195-202, 1995.

[4] M. Ungureanu, J.W.M. Bergmans, M. Mischi, S.G. Oei, R. Strungaru, “Improved method for fetal heart rate monitoring”, 27th Annual International Conference of the Engineering in Medicine and Biology Society, IEEE-EMBS, Shanghai, China, pp. 5916-5919, 2005.

[5] E.C. Karvounis, M.G. Tsipouras, D.I. Fotiadis, K.K. Naka, “An automated methodology for fetal heart rate extraction from the abdominal electrocardiogram”, IEEE Trans. Inf. Technol. Biomed., Vol. 11, pp. 628-638, 2007.

[6] D. Adam, D. Shaut, “Complete foetal ECG morphology recording by synchronized adaptive filtration”, Medical & Biological Eng. & Computing, Vol. 28, pp. 287-292, 1990.

[7] L. de Lathauwer, B. de Moor, J. Vandewalle, “Fetal electrocardiogram extraction by blind source subspace separation”, IEEE Trans. Biomed. Eng., Vol. 47, pp. 567-572, 2000.

[8] T. Oostendorp, Modelling the Fetal ECG, PhD dissertation, K.U. Nijmegen, The Netherlands, 1989.

[9] P.P. Kanjilal, S. Palit, G. Saha, “Fetal ECG extraction from single channel maternal ECG using singular value decomposition”, IEEE Trans. Biomed. Eng., Vol. 47, pp. 51-59, 1997.

[10] J.F. Cardoso, “Multidimensional independent component analysis”, Proceedings of the IEEE International Conference on Acoustics, Speech, and Sig. Proc. (ICASSP’98), Vol. 4, pp. 1941-1944, 1998.

[11] F. Vrins, V. Vigneron, C. Jutten, M. Verleysen, “Abdominal electrodes analysis by statistical processing for fetal electrocardiogram extraction”, in Proc. 2nd IASTED Int. Conf. on Biomedical Eng. (Biomed 2004), Innsbruck, Austria, pp. 244-249, 2004.

[12] M. Kotas, “Projective filtering of time-aligned beats for foetal ECG extraction”, Bulletin of Polish Academy of Sciences, Vol. 55, pp. 331-339, 2007.

[13] D.V. Prasad, R. Swarnalatha, “Extraction of fetal ECG from abdominal signal”, BIOSIGNAL 2009, Portugal, pp. 245-248, 2009.

[14] P. Gao, E.C. Chang, L. Wyse, “Blind separation of fetal ECG from single mixtures using SVD and ICA”, Proceedings of ICICS-PCM 2003, Singapore, pp. 15-18, 2003.

[15] Y. Ye, Z.L. Zhang, J. Chen, “A robust and non-invasive fetal electrocardiogram extraction algorithm in a semi-blind way”, IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences, Vol. E91-A, pp. 916-920, 2008.

[16] J. Tsao, C.Y. Hsu, M.T. Lo, “A supervised ICA algorithm for fetal ECG extraction”, ISSNIP 2008, Australia, pp. 327-330, 2008.

[17] D. Pani, S. Argiolas, L. Raffo, “A DSP algorithm and system for real-time fetal ECG extraction”, Computers in Cardiology, Vol. 35, pp. 1065-1068, 2008.

[18] S.I. Amari, T.P. Chen, A. Cichocki, "Stability analysis of learning algorithms for blind source separation", Neural Networks, Vol. 10, pp. 1345-1351, 1997.

[19] S. Amari, S. Cichocki, H. Yang, "A new learning algorithm for blind source separation", Advances in Neural Information Processing Systems, Vol. 8, pp. 757-763, 1996.

[20] A. Hyvarinen, "Survey on independent component analysis", Neural Computing Surveys, Vol. 2, pp. 94-128, 1999.

[21] B.W. Silverman, Density Estimation for Statistics and Data Analysis, New York, Chapman and Hall, 1985.

[22] M.C. Jones, The Projection Pursuit Algorithm for Exploratory Data Analysis, PhD dissertation, School of Mathematics, Univ. of Bath, 1983.

[23] W.L. Martinez, A.R. Martinez, Computational Statistics Handbook with MATLAB 2e, Chapman and Hall, 2008.

[24] R. Boscolo, H. Pan, V.P. Roychowdhury, "Non-parametric ICA", Proceedings of the Third International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2001), San Diego, pp. 13-18, 2001.

[25] R. Boscolo, H. Pan, V.P. Roychowdhury, "Independent component analysis based on non-parametric density estimation", IEEE Transactions on Neural Networks, Vol. 15, pp. 55-65, 2004.

[26] A. Bell, T. Sejnowski, "An information maximization approach to blind source separation and blind deconvolution", Neural Computing, Vol. 7, pp. 1129-1159, 1995.

[27] A. Hyvarinen, "Fast and robust fixed point algorithms for independent component analysis", IEEE Trans. on Neural Networks, Vol. 10, pp. 626-634, 1999.

[28] H.B. Barlow, "Unsupervised learning", Neural Computation, Vol. 1, pp. 565-608, 1989.

[29] T.M. Cover, J.A. Thomas, Elements of Information Theory, John Wiley & Sons, 1991.

[30] J.F. Cardoso, "Blind signal separation: statistical principles", Proceedings of the IEEE, Special Issue on Blind Identification and Estimation, Vol. 9, pp. 2009-2025, 1998.

[31] D. Donoho, "On minimum entropy deconvolution", Proceedings of the Second Applied Time Series Symposium, D.F. Findley, Ed., pp. 565-608, 1981.

[32] J.F. Cardoso, "Infomax and maximum likelihood for source separation", IEEE Letters on Signal Processing, Vol. 4, pp. 112-114, 1997.

[33] N. Vlassis, Y. Motomura, "Efficient source adaptivity in independent component analysis," IEEE Trans. Neural Networks, Vol. 12, pp. 559-566, 2001.

[34] J. Karvanen, J. Eriksson, V. Koivunen, "Pearson system based method for blind separation", Proceedings of Second International Workshop on Independent Component Analysis and Blind Signal Separation, pp. 585-590, 2000.

[35] E.K.P. Chong, S.H. Zak, An Introduction To Optimization, John Wiley & Sons, 2001.

[36] J.F. Cardoso, "High-order contrasts for independent component analysis", Neural Computation, Vol. 11, pp. 157-192, 1999.

[37] J.F. Cardoso, A. Souloumiac, "Blind beamforming for non-Gaussian signals", IEEE Proceedings Part F, Vol. 140, pp. 362-370, 1993.

[38] C. Salustri, G. Barbati, C. Porcaro, "Fetal magnetocardiographic signals extracted by 'signal subspace' blind source separation", IEEE Transactions on Biomedical Engineering, Vol. 52, pp. 1140-1142, 2005.

[39] V. Zarzoso, J.M. Roig, A.K. Nandi, "Fetal ECG extraction from maternal skin electrodes using blind source separation and adaptive noise cancellation techniques", Computers in Cardiology, Vol. 27, pp. 431-434, 2000.