10th Symposium on Neural Network Applications in Electrical Engineering, NEUREL-2010
Faculty of Electrical Engineering, University of Belgrade, Serbia, September 23-25, 2010

Independent Component Analysis (ICA) Methods for Neonatal EEG Artifact Extraction: Sensitivity to Variation of Artifact Properties

Nadica Miljković, Vladimir Matić, Sabine Van Huffel, and Mirjana B. Popović

***Abstract*—Independent Component Analysis (ICA) is becoming an accepted technique for artifact removal. Nevertheless, there is no consensus about appropriate methods for different applications. This study presents a comparison of common ICA methods: RobustICA, SOBI, JADE, and BSS-CCA, for extraction of ECG artifacts from EEG signal. Algorithms were applied to the data created by superimposing artifact free real-life neonatal EEG and synthetic ECG. Their sensitivity to variation of noise property was compared: we examined variability of Spearman correlation coefficients (SCC) for various Heart Rates (HR) in each of ICA methods. Results show that SOBI and BSS-CCA methods were less sensitive than RobustICA and JADE to artifact alterations (mean SCCs were 0.85 and 0.85 compared to 0.80 and 0.73, respectively) being quite successful in source signal extraction.**

***Index Terms*—ICA, artifact extraction, neonatal EEG**

I. Introduction

INDEPENDENT Component Analysis (ICA) is a promising technique [1] for artifact removal and potential EEG reconstruction. ICA algorithms may be used for EEG decomposition in order to extract ECG artifact [2] or to examine separate dynamic inter-relationships of different cortical areas [3].

One of possible application is detection of epileptic seizures from infant EEG [4]. Epileptic components are morphologically very similar to ECG artifact in EEG. ECG artifacts occur as a consequence of cardiac electrical field that affects the surface potential near the scalp [5]. Therefore it is important to distinguish between ECG artifact and epileptic components by extracting ECG artifact from EEG signal. There are several ICA based methods for artifact removal [6], but there is no consensus about the appropriate method for ECG extraction [4]. Running of inadequate ICA algorithm might lead to removal of important EEG components, false or missed epileptic detection. In this paper we present a comparison of four commonly used ICA methods for artifact detection: RobustICA, SOBI, JADE and BSS-CCA. Sensitivity of each algorithm to various Heart Rates (HR) of infant ECG is presented as a Spearman correlation coefficient (SCC) dependence of HR.

This work was supported in part by Ministry of Science, Serbia. V. Matić and S. Van Huffel are supported by Research Council KUL: GOA-MANET, FWO: G.0341.07 (Data fusion), DWTC: IUAP P6/04 (DYSCO).

N. Miljković is with University of Belgrade, Faculty of Electrical Engineering and Fatronik Serbia, Bulevar kralja Aleksandra 73, Belgrade, Serbia (phone: +381-11-321-83-48; e-mail: nadica.miljkovic@etf.rs).

V. Matić is with Department of Electrical Engineering (ESAT), Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, 3001 Leuven-Heverlee, Belgium (e-mail: vladimir.matic@esat.kuleuven.be).

S. Van Huffel is with Department of Electrical Engineering (ESAT-SCD), Katholieke Universiteit Leuven, Kasteelpark Arenberg 10, 3001 Leuven-Heverlee, Belgium (e-mail: sabine.vanhuffel@esat.kuleuven.be).

M. B. Popović is with University of Belgrade, Faculty of Electrical Engineering, Bulevar kralja Aleksandra 73, Belgrade, Serbia and with Center for Sensory-Motor Interaction (SMI), Aalborg University, Fredrik Bajers Vej 7D-3, 9220 Aalborg, Denmark (e-mail: mpo@smi.auc.dk).

II. Method

Neonatal EEG signals with polygraphy, were recorded continuously using a Nervus™ monitor (Taugagreining hf, Reykjavik, Iceland). Seventeen scalp electrodes were placed according to the full 10-20 International System EEG data was recorded at the Sophia Children's Hospital – part of Erasmus MC, the University Medical Center in Rotterdam, the Netherlands. EEG signals were recorded on neonates within 24 hours following the birth with sampling frequency of 256 Hz. We used 30 seconds long EEG segments for our simulations. Neurophysiologists confirmed that those segments do not contain any artifact of significant noise level nor epileptic activity [4].

A. Synthetic ECG Dataset

We used embedded Matlab R2007b (The MathWorks, Natick, MA, USA) function for ECG generation. We designed program that allows HR selection with the sampling rate of 256 samples per second that matches the EEG sampling rate.

From the literature on [7] we found that neonatal HR ranges from 92 bpm (Beats Per Minute) to 157 bpm, mean 122.18 bpm. We varied HR for a slightly wider range, from 50 bpm to 160 bpm.

B. Creation of Synthesized Data

We superimposed one randomly chosen channel from acquired EEG data [4] with synthetic ECG dataset. Literature on [1] indicates that artifact data should range in size from about three hundred times smaller (- 50 dB) to the size of the EEG data (0 dB). We set size to be within this range: five times (-14 dB) smaller artifact size than EEG. One example of synthesized (noisy) EEG *vs.* original EEG is shown in Fig. 1.

978-1-4244-8819-3/10/$26.00 ©2010 IEEE

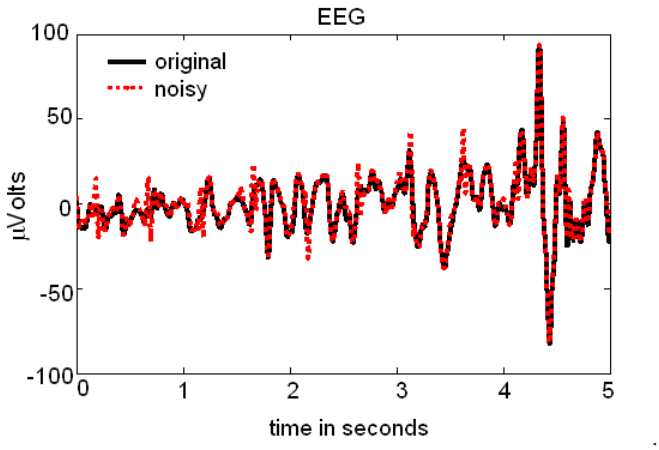

Fig. 1. Original and noisy EEG data for 122 bpm and five times smaller size of ECG artifact compared to EEG data

C. ICA Algorithms

A fundamental problem in neural network research, as in many other disciplines, is to find a suitable representation of multivariate data. ICA is recently developed method and its goal is to find representation of nongaussian data so that the components are statistically independent or as independent as possible [6]. For reasons of computational and conceptual simplicity [6], we assumed the basic linear ICA model that is defined by the following equation:

$$x(t) = As(t) \tag{1}$$

Mutually independent sources *s* are mixed by mixing matrix *A*, creating observable multichannel signals *x*. The goal of ICA methods is to estimate signal *s*, without any *a priori* information about either source components *s* or mixing matrix *A*. In order to accomplish this, ICA algorithms estimate an inverse mixing matrix *W*:

$$\hat{s}(t) = Wx(t) \tag{2}$$

Depending on the defined criteria function a numerous ICA algorithms have been developed.

We compared four ICA algorithms that were available on the Internet in public domain. Precondition for algorithms is to be written in Matlab with the open source code, [8]–[11]. RobustICA, SOBI, JADE and BSS-CCA were implemented in our Matlab R2007 code. We applied synthesized data to four algorithms and compared their sensitivity to the variation of one property of noise and that is HR.

The first algorithm that we applied was RobustICA (Robust Independent Component Analysis). It is based on the FastICA (Fast Independent Component Analysis) algorithm and its criteria function optimizes kurtosis contrast applying exact line search optimization method [9].

SOBI (Second Order Blind Identification) algorithm is very effective when independent sources are mutually uncorrelated, but correlated individually. It calculates second order statistics – covariance matrices, which are later jointly diagonalized [10].

JADE (Joint Approximation Diagonalization of Eigen matrices) algorithm uses fourth order moments (cumulant matrices) to separate sources from the mixed signals. It employs joint approximate diagonalization to create cumulant matrices as diagonal as possible [8].

BSS–CCA (Blind Source Separation – Canonical Correlation Analysis) algorithm assumes mutually uncorrelated sources that are maximally correlated with a predefined function. This predefined function is specified as a delayed version of a mixed signal *x*. It should be emphasized that BSS–CCA assumes uncorrelated sources, whereas other ICA methods assume statistical independence which is more rigorous criterion.

D. Comparison of ICA Algorithms: Sensitivity to HR

Spearman correlation coefficient (SCC) is often used method [12] for comparison of the original source and the independent component, since it is not dependent on the absolute signal amplitude. It is only dependent on the signal shape.

We calculated Spearman correlation coefficient between original ECG noise (that is superimposed to EEG) and extracted independent component (that is the most correlated to ECG signal) for HR in the range of 50 to 160 bpm.

For the extraction of independent component that corresponds to ECG, we used linear correlation method between normalized extracted component, and normalized original ECG noise with mean value 0 and with standard deviation 1 [4]. The highest linear correlation gave corresponding ECG independent component.

TABLE I
SCC FOR DIFFERENT ICA ALGORITHMS

Values for	RobustICA	SOBI	JADE	BSS-CCA
Mean	0.80	0.85	0.73	0.85
STD	0.04	0.04	0.11	0.03
1st quartile	0.78	0.83	0.65	0.83
2nd quartile	0.80	0.86	0.78	0.86
3rd quartile	0.82	0.88	0.82	0.88
4th quartile	0.87	0.92	0.87	0.90

Mean, STD and quartiles calculated for four ICA algorithms due to Spearman correlation coefficient (SCC) variation for different HR. STD stands for standard deviation.

III. RESULTS

Table I presents summary results of presented sensitivity analysis. Mean values with standard deviations, as well as quartile values for Spearman correlation coefficients for four

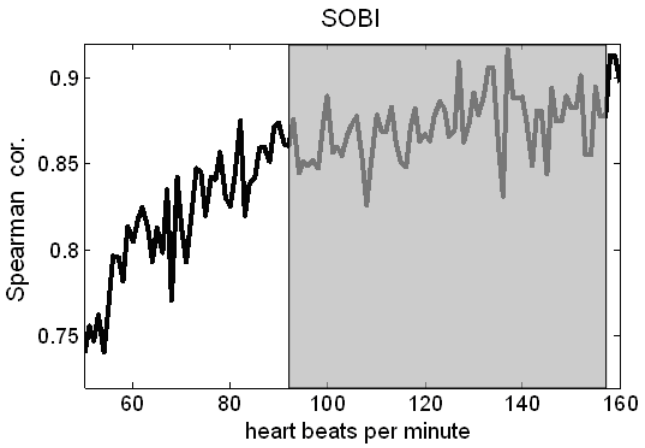

Fig. 2. Spearman correlation coefficient variation for various HR of ECG data in the synthesized EEG applied to SOBI algorithm. Gray area shows the range for neonatal HR typical values

ICA algorithms are given. The lowest mean SCC value is for JADE while the highest values are for SOBI and BSS-CCA algorithms. Standard deviation is noticeably higher for JADE algorithm compared to the three others.

Fig. 2 shows Spearman correlation coefficient variation due to different HR values of ECG synthesized in recorded EEG when applied to SOBI algorithm. Gray area in Fig. 2 represents range of typical HR values in neonatal ECG [7]. It is noticeable that correlation values go to saturation for higher HR, thus demonstrating decrease of algorithm sensitivity with HR increase. Spearman correlation coefficient goes from 0.74 for HR of 50 bpm to 0.9 for HR of 160 bpm, while for the neonatal ECG range of 92 to 157 bpm SCC is in the range of 0.86-0.88 as shown in Fig. 2

IV. Discussion

While the difference in mean SCC between SOBI and JADE algorithms is 0.12 it is quite interesting to analyze their quartile values. The SCC difference between SOBI and JADE is 0.18 in the 1st quartile, while in the 4th they differ for 0.05. Another example is the difference between RobustICA and JADE in 1st quartile of +0.13, while in the 4th quartile SCC decreases to 0. To conclude, maximal difference in SCC for the four ICA algorithms is 3.6 smaller in the 4th quartile compared to the 1st. This analysis implicates that higher HR causes decrease in SCC differences among four algorithms. For neonatal HR variation in SCC is even less (gray area in Fig. 2).

It has been established that effectiveness of different ICA algorithms varies for different artifacts [9]. Here, we presented that effectiveness varies due to distinct properties of ECG artifact. We studied neonatal, artifact free EEG, contaminated with synthetic ECG that provides more realistic effects than simulated blinks and sinus wave [4], [13]. SOBI and BSS-CCA methods have been more successful in ECG extraction than JADE and RobustICA according to Table I. Since ECG is of periodic nature, the differences among SCCs might be due to the fact that SOBI and BSS-CCA are algorithms based on temporal coherence [6]. Algorithms based on maximization of non-gausianity properties (RobustICA and JADE) showed lower mean SCC. In our study SOBI algorithm showed the best properties due to HR variations of ECG.

What all four algorithms have in common is that their performance changes due to different HR. Since ICA method is based on criterion that mixed signal must be statistically independent or noncorrelated [6], it is possible that changes in SCC originate from statistical parameter variations between EEG and ECG when HR varies.

V. Conclusion

Presented results indicate that selection of ICA for the specific application must be done carefully in respect to the properties of proposed method.

Our study has several limitations. We have not examined other noise variable parameters such as signal amplitude (it was a constant). Linear ICA representation seems to capture the essential structure of the data in many applications as in EEG artifact extraction, [4], [6], although it does not correspond to the true nature of the processes.

For detection of the independent component that corresponds to ECG, we used linear correlation method. It might be more appropriate to use neural network, since it was recently reported that neural networks have higher classifying performance when used with ICA [14].Once trained neural network on ICA decomposed components might be used for classifying seizure and nonseizure (ECG) EEG segments. It could lead to early and effective epileptic seizure detection in infant and to early medication treatment [4].

Acknowledgment

Authors would like to thank the Department of Neonatology, Sophia Children's Hospital, Erasmus MC, University Medical Center Rotterdam, the Netherlands, for providing selected segments of neonatal EEG signals.

References

[1] S. Devuyst, T. Dutoit, P. Stenuit, M. Kerkhofs, and E. Stanus, "Cancelling ECG Artifacts in EEG Using a Modified Independent Component Analysis Approach," *EURASIP J. on Advances in Signal Processing*, vol. 2008, pp. 1-13, 2008.

[2] A. Delorme, T. Sejnowski, and S. Makeig, "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis", *NeuroImage*, vol. 34, no. 4, pp. 1443-1449, Feb. 2007.

[3] A. Delorme, and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis", *Neurosci Methods J.*, vol. 134, pp. 9-21, 2004.

[4] V. Matic, W. Deburchgraeve, and S. Van Huffel, "Comparison of ICA algorithms for ECG artifact removal from EEG signals.", *Proc. of the 4th Annual symposium of the IEEE-EMBS Benelux Chapter. (IEEE-EMBS)*, Twente, The Netherlands, pp. 137-140, Nov. 2009.

[5] G. Dirlich, L. Vogl, M. Plaschke, and F. Strian, "Cardiac field effects on the EEG", *Electroencephalography and Clinical Neurophysiology J.*, vol. 102, no. 4, pp. 307-315, Apr. 1997.

[6] A. Hyvarinen, and E. Oja, "Independent component analysis: algorithms and applications", *Neural Networks J.*, vol. 13, no. 4-5, pp. 411-430, Jun. 2000.

[7] E. Longin, T. Schaible, T. Lenz, and S. König, "Short term heart rate variability in healthy neonates: Normative data and physiological observations", *Early human development J.*, vol. 81, no. 8, pp. 663-671, Aug. 2005.

[8] JADE Matlab code, avaliable: http://sig.enst.fr/~cardoso/stuff, Aug. 2010

[9] RobustICA Matlab code, avaliable: http://www.i3s.unice.fr/~zarzoso/robustica, Aug. 2010

[10] SOBI Matlab code, avaliable: http://sccn.ucsd.edu/eeglab/allfunctions/acsobiro, Aug. 2010

[11] BSS-CCA Matlab code, avaliable: http://www.imt.liu.se/~magnus/cca, Aug. 2010

[12] M. Klemm, J. Haueisen and G. Ivanova, "Independent component analysis: comparison of algorithms for the investigation of surface electrical brain activity". *Med Biol Eng Comput J.*, vol. 47, pp. 413–423, 2009.

[13] K. A. Glass, G. A. Frishkoff , R. M. Frank, C. Davey, J. Dien, A. D. Malony, D. M. Tucker, "A Framework for Evaluating ICA Methods of Artifact Removal from Multichannel EEG". *Proc. Of the 5th International Conference on Independent Component Analysis and Blind Signal Separation*, Granada, Spain, pp. 1–8, Sep. 2004.

[14] Z. Nouir, B. Sayrac, B. Fourestie, W. Tabbara, F. Brouaye. "Independent Component Analysis for Radio Network Prediction Enhancement". *Proc. of the 20th European Conference on Modelling and Simulation*, Bonn, Germany, pp. 1-5, 2006.