Robust removal of short-duration artifacts in long neonatal EEG recordings using wavelet-enhanced ICA and adaptive combining of tentative reconstructions

M Zima[1,2], P Tichavský[2], K Paul[3], and V Krajča[4,5]

[1] Czech Technical University in Prague, Faculty of Nuclear Science and Physical Engineering, Trojanova 13, 120 00 Prague 2, Czech Republic

[2] Institute of Information Theory and Automation, Pod vodárenskou věží 4, P.O.Box 18,182 08 Prague 8, Czech Republic

[3] Institute for the Care of Mother and Child, Podolské nábřeží 157, Prague 4, Czech Republic

[4] Faculty Hospital Na Bulovce, Budínova 2, 182 00 Praha 8, Czech Republic

[5] Czech Technical University in Prague, Faculty of Biomedical Engineering, nám. Sítná 3105, 27201 Kladno, Czech Republic

E-mail: `zimamiro@fjfi.cvut.cz`,`tichavsk@utia.cas.cz`, `karelpaul@seznam.cz`,`vladimir.krajca@fbmi.cvut.cz`

Abstract. The goal of this paper is to describe a Robust Artifact Removal (RAR) method – an automatic sequential procedure which is capable of removing short-duration, high-amplitude artifacts from long-term neonatal EEG recordings. Such artifacts are mainly caused by movement activity, and have an adverse effect on automatic processing of long-term sleep recordings. The artifacts are removed sequentially in short-term signals using ICA transformation and wavelet denoising. In order to gain robustness of the RAR method, the whole EEG recording is processed multiple times. The resulting tentative reconstructions are then combined. We show results in a data set of signals from ten healthy newborns. Those results prove, both qualitatively and quantitatively, that the RAR method is capable of automatically rejecting the mentioned artifacts without changes in overall signal properties such as the spectrum. The method is shown to perform better than either the wavelet-enhanced ICA or the simple artifact rejection method without the combination procedure.

PACS numbers: 87.19.Ie

Submitted to: *Physiol. Meas.*

1. Introduction

One of the most important indicators used to study the maturation of the brain is an electroencephalogram (EEG). EEG describes the electrical activity of the brain and contains important information about the state of the patient's health. Visual analysis of the EEG activity is a difficult and tedious task; automatic quantitative methods of relevant signal parameters (other than spectrum or coherence analysis) are needed.

In previous studies, e.g., Gerla et al. (2009), methods have been developed that help to analyse different features obtained from neonatal EEGs. The major drawback of automatic methods is the fact that the neonatal EEG is almost always contaminated by various kinds of artifacts – see, e.g. Celka *et al* (2001). They may be caused by muscle activity (EMG artifacts), movement of the body, eye-induced artifacts (eye blinks and movements) etc. The amplitude of the artifacts can be quite large relative to the amplitude size of the cortical signals of interest. This is one of the reasons why an expert is needed to correctly interpret clinical EEGs, and why the artifact presence can damage an automatic EEG analysis. Because of this, an artifact-removing algorithm is much needed.

This work was first motivated by the fact that methods of the Independent Component Analysis (ICA) have been shown to be very useful in analysing biomedical signals, such as EEG and MEG, see Makeig *et al* (1996), Vigario *et al* (2002), Joyce *et al* (2004), James and Hesse (2005). These methods have an ability to separate artifacts which are statistically independent of useful biological signals, and have non-Gaussian probability density function or different spectra. In the EEG signal processing, the most widely studied ICA algorithms are Infomax (Bell and Sejnowski *et al* 1995), SOBI (Belouchrani *et al* 2002), and FastICA (Hyvärinen and Oja, 1997). While SOBI is based on second-order statistics, the other two algorithms use high-order statistics. In this paper, we use an algorithm BGSEP (Block Gaussian Separation, Pham and Cardoso, 2002) implemented through the algorithm of Tichavsky and Yeredor (2009). This method produces excellent separation performance and it is also very efficient computationally. A comparative study of several ICA methods can be found, e.g., in Delorme *et al* (2012) or Klemm *et al* (2009).

Performance of ICA can be enhanced by the Spatially Constrained ICA (scICA), first described in Ille (2001). ScICA not only extracts artifact-based independent components but it also incorporates prior knowledge about spatial topographies, for example of artifacts, into the ICA algorithm by means of constraints. In Hesse and James (2005), an efficient gradient-based algorithm was introduced to perform a spatially constrained ICA. It was also studied by Phlypo *et al* (2006) and De Vos *et al* (2011a). In Akhtar (2012), the spatially constrained ICA is combined with wavelet denoising.

An automatic artifact rejection method for the purpose of neonatal seizure detection was proposed recently by De Vos *et al* (2011b). The method was again based on ICA, and identification of artifact components relied on correlation with a simultaneously recorded polygraphic signal. The goal of that paper is somewhat different from ours.

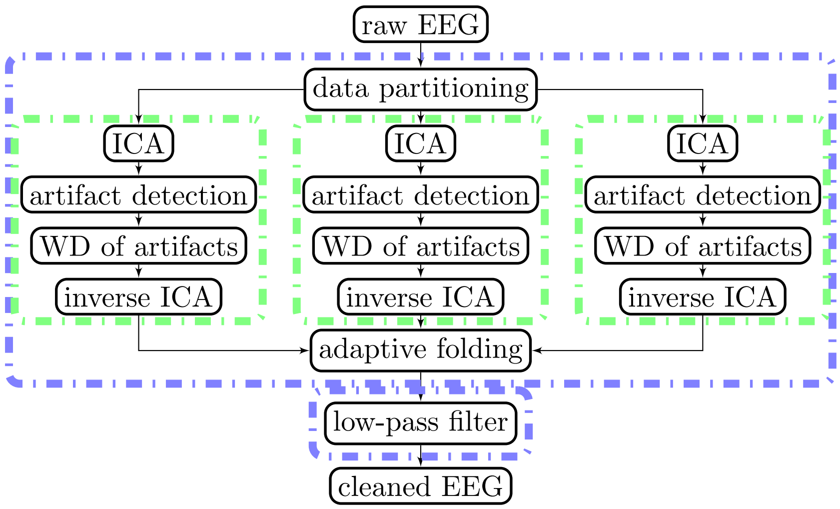

Figure 1. Steps of the RAR method. First, artifacts with too large amplitudes are removed (first blue block). This is performed by sequential usage of the wavelet-enhanced ICA (green blocks).

In this paper, we propose a Robust Artifact Removal (RAR) method for artifact rejection from an arbitrary-length signal. We are mainly interested in removal of short-duration artifacts characterised by a high amplitude. The main motivation is detection of sleep stages, which is difficult due to the frequent presence of artifacts. The method does not rely on polygraphic signals, but if these are available it is possible to utilise them as well, as is done in the De Vos paper.

The artifacts are removed sequentially: in a short-term signal, the ICA transformation of the signal is computed (subsection 2.1) and demixed artifacts are then thresholded by Wavelet Denoising (subsection 2.2). In order to achieve robustness within the RAR method, the whole EEG recording is processed multiple times and these tentative reconstructions are then combined (using a method presented in subsection 2.4). In order to reject high-frequency artifacts as well, the RAR method is completed by a standard low-pass filter. In the simulation section, we show results of processing the EEG recordings of ten healthy newborns. The results prove that the RAR method is capable of automatically rejecting the mentioned artifacts without changes in overall signal properties such as the spectrum. In particular it is shown to perform better than either the plain wavelet-enhanced ICA of Castellanos and Makarov (2006) or the simple artifact rejection method without the combination procedure.

2. Building blocks of the RAR method

In order to make the description of the RAR clearer, processing of an EEG record is schematically depicted in Figure 1. Details of the method are described in the following subsections.

2.1. ICA

The aim of ICA is to convert a multichannel signal X via an invertible linear transformation to so-called independent components S. Actually, the separated components may not be truly statistically independent, but they are as independent as possible according to certain criteria. Symbolically, the considered model is

$$X = AS \tag{1}$$

where S represents a $d \times N$ matrix, composed of d rows and N samples, so that each row denotes one independent component.

In this paper, we estimate the inverse of A using an algorithm BGSEP (Block Gaussian Separation) of Pham and Cardoso (2002) implemented through Tichavský and Yeredor, 2009. BGSEP is based on second-order statistics as is done in algorithm SOBI (Belouchrani *et al* 1998), but it uses the non-stationarity of separated signals. While SOBI is achieved by approximate joint diagonalisation (AJD) of a set of time-lagged covariance matrices of the signal (the mixture), BGSEP performs an AJD of zero lag covariance matrices in a partition of the signal. We use BGSEP because it is computationally very efficient and also produces better separation performance than other studied algorithms, e.g FastICA of Hyvärinen and Oja (1997) and Infomax (Makeig *et al* 1996). Comparison of BGSEP with other ICA methods can be found in Tichavský and Koldovský (2011).

In the context of the artifact removal, it is desirable to have unwanted signals concentrated in a small number of separated components. The original signal can be reconstructed without the artifact components (i.e., the components containing artifacts) using the estimated matrix A. An illustrative example is shown in Figure 2.

2.2. Wavelet-enhanced ICA

Dealing with real EEGs, estimated independent components capturing artifacts frequently contain a considerable amount of cerebral activity. Rejection of such components results in loss of a part of the cerebral activity and, consequently, distortion of the artifact-free EEG, see Figure 3 for example.

To mitigate this problem, we use the method of wavelet-enhanced ICA (wICA) proposed in Castellanos and Makarov (2006). This method uses Wavelet Denoising (WD), e.g., Quiroga *et al* (2003), on ICA components. The advantage of this approach is that it enables us to retain a residual neural signal in components containing artifacts.

In order to use WD for artifact removal, the partly separated component s is assumed to be composed of the high amplitude artifact $a(t)$ and a low amplitude residual neural signal $n(t)$, symbolically

$$s(t) = a(t) + n(t). \tag{2}$$

For removing artifacts without losing the residual neural signal $n(t)$, an estimate of $a(t)$ proposed by WD is subtracted from $s(t)$ and the inverse ICA transformation is

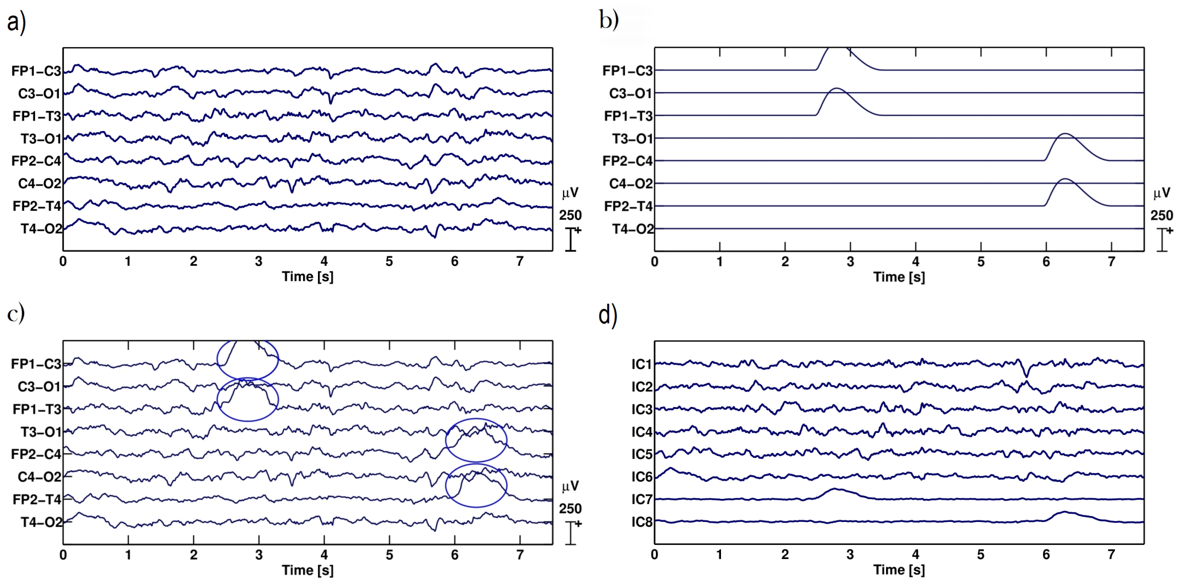

Figure 2. Short EEG with artificially added artifacts. The Figure contains: a) the original data, b) added artifacts, c) contaminated data and d) separated components provided by BGSEP. The artifacts have been separated into the last two components.

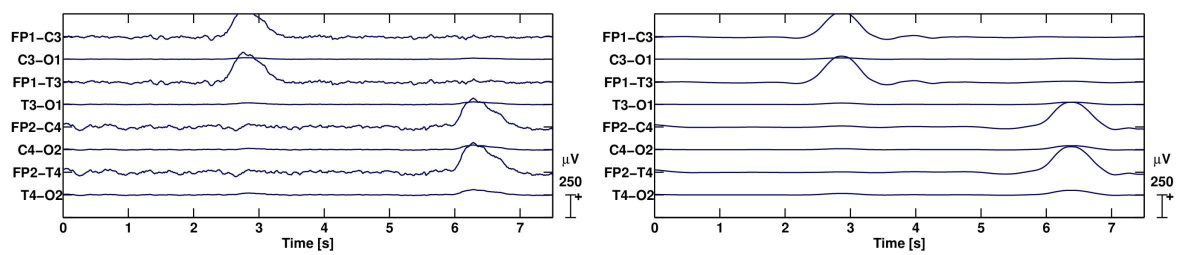

Figure 3. Artifacts added into the data in Figure 2 estimated by ICA are shown in the left part of this Figure. An estimate was computed via inverse ICA transformation after replacing all non-artifact components (the first six of them) by zeros. The estimate using wICA is shown in the right part.

performed using only $n(t)$ instead of $s(t)$. In particular, we apply level 7 decomposition with Daubechies wavelet ψ_{D6}, and a threshold $T = \sqrt{2\log(d)}$ for the denoising, where d denotes the number of samples in the segment‡. The WD we used can be described schematically

- compute the Discrete Wavelet Transformation (DWT) of $s(t)$, i.e., compute the wavelet coefficients $a_{j,k}$
- for all $a_{j,k}$ perform the soft thresholding

$$\hat{a}_{j,k} = \begin{cases} \operatorname{sgn}(a_{j,k})(|a_{j,k}| - T) & \text{if } |a_{j,k}| \geq T, \\ 0 & \text{if } |a_{j,k}| < T. \end{cases}$$

- compute the inverse DWT $\hat{a}(t)$ using wavelet coefficients $\hat{a}_{j,k}$.

‡ In later experiments, we used $d = 5000$, thus $T = 4.1273$.

component	1	2	3	4	5	6	7	8
sparsity	3.057	1.903	1.814	1.862	1.278	1.905	9.358	7.367

Table 1. Numerical values of the sparsity (3) computed for components in Figure 2.

Here, the $\hat{a}(t)$ approximates the artifact $a(t)$ without the neural signal $n(t)$.

In the original wICA of Castellanos and Makarov, the wavelet denoising is applied to *all* ICA components (without any selection). Each ICA component is decomposed into a sum of the noise and the rest. The "noise" is interpreted as the neural signal, and the rest is considered to be an artifact. The updated ICA components after removing the artifacts are multiplied by the estimated mixing matrix A to reconstruct the data. This procedure is capable of rejecting artifacts to some extent in our application, see Section 3 below. However, it appears to be more effective to apply the wavelet denoising only to those components that are classified to contain artifacts.

2.3. Automatic detection of artifact components

Correct identification of artifact components is crucial for the proposed method. In the spatially constrained ICA, the selection of the artifact component is performed jointly with the separation. It is also possible to utilise a simultaneously recorded polygraphic signal, if it is available, as is done in De Vos *et al* (2011b).

In this paper, we do not assume existence of the polygraphic signal and propose an ad hoc criterion that, although simplistic, is suitable in our application. In any case, the choice of the criterion is not crucial for the method: it can easily be replaced by another method of selecting the artifact component.

The criterion is based on the assumption that artifacts with high amplitude have one feature in common: their duration is short in comparison to the chosen frame length. Such signal components will be called *sparse* in the time domain. Sparse components have a large maximum absolute value (due to the presence of the artifact), and simultaneously the median of the absolute value close to zero relative to $\mathrm{std}[s_i^{(j)}]$, where "std" stands for a standard deviation. Thus, we propose the following definition of sparsity

$$\mathrm{sparsity}(s^{(j)}) = \frac{\max[|s_i^{(j)}|]}{\mathrm{std}[s_i^{(j)}]} \log\left(\frac{\mathrm{std}[s_i^{(j)}]}{\mathrm{median}[|s_i^{(j)}|]}\right), \tag{3}$$

where $s^{(j)} = (s_1^{(j)}, \ldots, s_N^{(j)})$ is the $j-$th component, i is the time index, and N is the number of samples in the frame.

The component is regarded to be sparse (artifact) if its sparsity exceeds some threshold. A higher value of the limit means a more conservative (a weaker) artifact reduction. For illustration, numerical values of the criterion on components from Figure 2 are shown in Table 1. In later computations, we use the threshold sparsity equal to 2.5.

Note that if the threshold sparsity is set to zero, it is assumed that each ICA component contains an artifact and the WD is performed in all of them. The resulting algorithm is equivalent to wICA of Castellanos and Makarov.

Another trivial artifact denoising procedure would be obtained if the wavelet denoising is applied to the original (raw) EEG data. Again, the "noise" is interpreted as the useful (cerebral) signal and the rest as the artifact. No ICA is needed at all in this procedure. Unfortunately, performance of this method appears to be even worse than performance of wICA; however, it can be expected.

2.4. Robust artifact rejection from long-term signal

The simplest way to cope with long-term signals is to take non-overlapping frames, and perform the artifact rejection in each of them separately. This simple sequential method will be denoted as the SAR (Simple Artifact Removal) method. The length of the frames should be selected so that each frame contains a sufficient amount of artifact-free signal. For example, in our case of eight channel EEG the number of artifacts should not exceed two or three artifacts per frame, each having a length of 1 to 2 seconds. If the number of artifacts is higher or if artifacts are longer, the artifact removal is not reliable.

If the number of channels forming the EEG record is higher, we assume that the method would work as well, or even better, because more information about the neural activity is available. However, some fine-tuning of the parameters might be necessary.

In this section, we propose a method that is better than SAR, namely in difficult scenarios where the artifact presence is frequent. In this method, called RAR (Robust Artifact Removal), the plain artifact removal is performed in multiple frames three times, each time with a different partitioning of the signal. Each partitioning gives one possible artifact-free reconstruction of the whole signal. These reconstructions are combined together in a special way so that the final reconstruction is generally smoother and more artifact-free than the tentative reconstructions. The advantage of using multiple processing becomes apparent in the experimental section.

2.4.1. Data partitioning Let N denote the length of one frame and L be the total length of the data. At first, the signal is divided into frames $[1 + (k-1)N, kN]$ where $k = 1 \dots n$, $n = \lfloor L/N \rfloor$. The second tentative reconstruction is done in a similar way with frames $[1 + N/3 + kN, N/3 + (k+1)N]$ for $k = 1 \dots n-1$. The third partitioning is $[1 + 2N/3 + (k-1)N, 2N/3 + kN]$ with $k = 1 \dots n-1$. For the second and third reconstructions, ICA is not performed at the beginning and end of the signal. Here, the first reconstruction is used as a final reconstruction instead.

The combination of three reconstructions into one proceeds sequentially, independently channel by channel, in segments of the length T which are generally shorter than N. Hence, segments have the form $[1 + (k-1)T, kT]$ for $k = 1 \dots \lfloor L/T \rfloor$. Division of the signal into frames and segments is shown schematically in Figure 4.

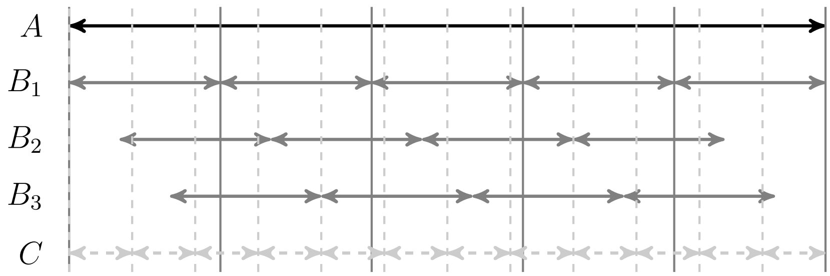

Figure 4. In three independent steps, the signal A is divided into frames B_i where the denoising is applied. After obtaining tentative reconstructions, they are combined channel by channel, segment by segment, into the final reconstruction. Locations of segments C are schematically shown.

2.4.2. Adaptive folding Let r_1, r_2 and r_3 denote three tentative reconstructions of a segment in a data channel. Let μ_i denote the maximum absolute value of elements in r_i. We assume that at least one tentative reconstruction is artifact-free (otherwise, there is no possibility of obtaining artifact-free reconstruction from their combination). Without any loss of generality we assume that $\mu_1 \leq \mu_2 \leq \mu_3$. Therefore, at least r_1 is artifact free. Let $\rho_{ij} = \|r_i - r_j\|^2$ denote the squared Euclidean norm of reconstructions and let ρ_r denote the average squared Euclidean norm $\|r\|^2$ of a segment r of the same length as r_i, randomly or systematically chosen from the entire available signal.

The final reconstruction r is obtained as the average of one, two, or all three tentative reconstructions depending on validity of the conditions:

$$\max(\rho_{12}, \rho_{13}, \rho_{23}) < 2\rho_r, \tag{4}$$

$$\max(\rho_{12}, \rho_{13}, \rho_{23}) \leq 2\min(\rho_{12}, \rho_{13}, \rho_{23}) \ . \tag{5}$$

The condition (4) indicates that there is probably no artifact in the segment. The condition (5) means that differences between the reconstructions are small. If any of these conditions is fulfilled, all three partial reconstructions are averaged to produce the final reconstruction. The complete procedure is summarised in Figure 5.

An illustrative example of the combination procedure is shown in Figure 6.

3. Experiments

In this section, performance of the RAR method is studied on a database of EEG recordings of ten different healthy newborns. Every recording has eight channels, about 70 min long, and was sampled at 256Hz under a bipolar montage. The recordings were processed by the RAR method with parameters $N = 5000$ samples (cca 19.5 s), $T = 256$ samples (1 s), BGSEP had an internal parameter of 10, sparsity threshold was 2.5, and the low-pass filter was the Butterworth type of the order 10 and cut-off frequency 50 Hz. Note that each processing (70 min. long recordings) takes approximately 30 s on an ordinary PC with a 2 GHz processor and 3 GB RAM in Matlab R2010b.

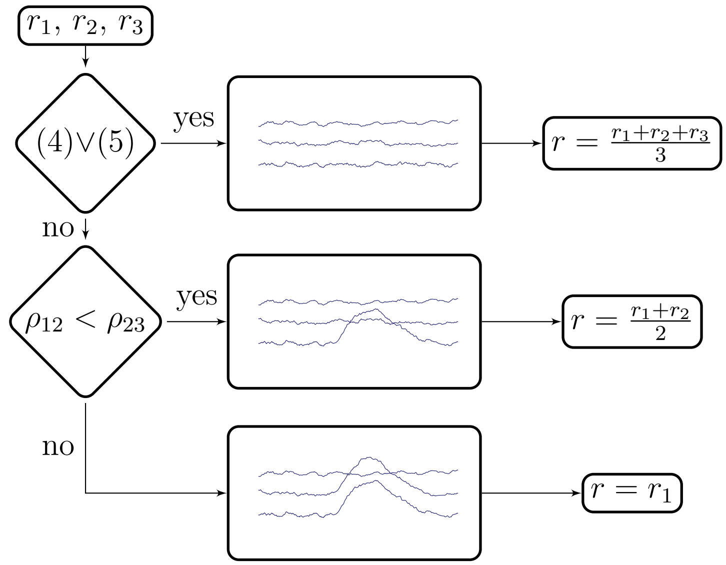

Figure 5. Scheme of combination of tentative reconstructions. The first decision means that there are not significant differences between r_1, r_2, and r_3. The second decision divides the cases according to whether r_2 contains the artifact or not (note that the r_1 is assumed to be artifact-free).

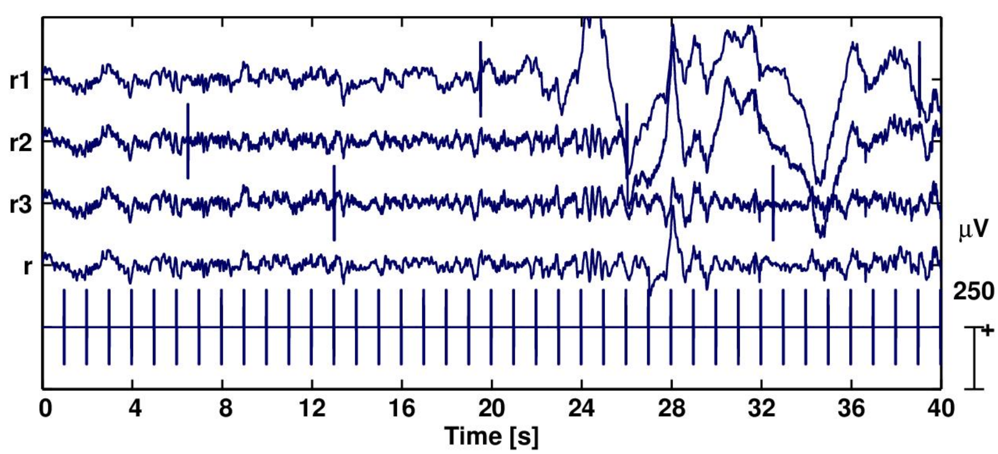

Figure 6. Real example of a combination procedure of possible reconstructions r_1, r_2, r_3 that still contain some artifacts. The final reconstruction r is in the fourth channel, vertical lines denote partitioning into frames and segments (shown in the bottom part).

no.	feature	method
1	standard deviation	$\mathrm{std}(x_t)$
2	amplitude of the signal	$\max(x_t) - \min(x_t)$
3	norm of PSD in the band 0.5-1.6 Hz	using $\mathrm{FFT}(x_t)$
4	norm of PSD in the band 1.6-3.0 Hz	using $\mathrm{FFT}(x_t)$
5	norm of PSD in the band 3.1-5.0 Hz	using $\mathrm{FFT}(x_t)$
6	norm of PSD in the band 5.1-8.0 Hz	using $\mathrm{FFT}(x_t)$
7	norm of PSD in the band 8.1-14.0 Hz	using $\mathrm{FFT}(x_t)$
8	mean absolute value of the first derivative	$\mathrm{E}(\lvert x_{t+1} - x_t\rvert)$
9	maximum of absolute value of the first derivative	$\max(\lvert x_{t+1} - x_t\rvert)$
10	maximum of absolute value of the second derivative	$\max(\lvert x_{t+1} - 2x_t + x_{t-1}\rvert)$

Table 2. Ten features that statistically characterise EEG signals. Reference values and std for falling asleep and REM sleep stages are displayed in Table 3.

.

3.1. Methodology

The main motivation for designing the artifact removal procedure was to develop a preprocessing tool for classification of sleep stages of newborns, which is often difficult because of artifacts. For this purpose, the signals (original and processed) were expertly divided into parts so that each part can be assigned to one of three possible classes: falling asleep stage, Non-Rapid Eye Movement (NREM) sleep (also known as quiet sleep) and Rapid Eye Movement (REM) sleep (also known as active sleep). Then, 20 s long parts corrupted by artifacts (expertly identified and denoted by Karel Paul) were selected from each EEG record from both the falling asleep and REM stages. The NREM sleep stages were excluded from further study because our data set was almost artifact-free in this domain. Moreover, our method does not cause any significant changes in the artifact-free signals, as we show in one of the later experiments.

In particular, we select twelve 20 s long parts, three corrupted by artifacts and three artifact-free from each of the two studied sleep stages and each of ten patients. Thus we have 60 parts of 20s long data samples containing artifacts and the same amount of artifact free signals.

In order to compare statistical properties of the processed and the original signals, every channel of signal x_t is described by ten features summarised in Table 2. These features are often used for various diagnostic purposes. Table 3 contains numerical values (mean and standard deviation) of these characteristics for healthy newborns in artifact-free parts for falling asleep and REM stages. Values were obtained by evaluating the statistics for every 20 s long part and taking their mean value and standard deviation across parts and channels.

	falling asleep				REM			
	artifact-free		contaminated		artifact-free		contaminated	
no.	orig	RAR	orig	RAR	orig	RAR	orig	RAR
1	45 ± 12	39 ± 11	132 ± 81	46 ± 17	34 ± 6	30 ± 6	150 ± 116	42 ± 18
2	281 ± 72	248 ± 69	914 ± 599	321 ± 98	218 ± 43	190 ± 42	977 ± 696	275 ± 81
3	148 ± 48	126 ± 46	345 ± 185	150 ± 56	107 ± 25	91 ± 25	324 ± 250	129 ± 67
4	92 ± 29	87 ± 27	118 ± 50	102 ± 37	62 ± 16	59 ± 15	96 ± 49	81 ± 34
5	56 ± 13	55 ± 12	62 ± 18	59 ± 16	43 ± 11	43 ± 10	55 ± 19	51 ± 14
6	40 ± 8	40 ± 8	42 ± 11	41 ± 9	35 ± 7	35 ± 7	39 ± 10	38 ± 8
7	26 ± 5	26 ± 5	35 ± 11	29 ± 7	23 ± 4	23 ± 4	32 ± 8	27 ± 5
8	2 ± 1	2 ± 1	7 ± 4	2 ± 1	2 ± 1	2 ± 1	6 ± 4	2 ± 1
9	11 ± 3	11 ± 3	104 ± 166	14 ± 7	9 ± 2	9 ± 2	68 ± 57	13 ± 6
10	2 ± 1	3 ± 1	119 ± 160	3 ± 1	2 ± 1	2 ± 1	82 ± 79	3 ± 1

Table 3. Comparison of the studied characteristics for the artifact-free signal and the signal contaminated by artifacts processed by the RAR method.

3.2. Results of the RAR method

Comparison of the original and processed signals is performed through the features shown in Table 3. The presented results prove that the RAR method significantly lowers overall artifact activity, both in amplitude and frequency and the resulting signals have nearly the same properties as the reference signal. In addition, the Table shows that the RAR method does not significantly affect the properties of the artifact-free signal.

3.3. Comparison with other techniques

In this subsection, the performance of RAR is compared with results of two simpler artifact rejection methods: SAR with different sparsity thresholds, and wavelet-enhanced ICA with a different denoising threshold. In the following, SAR(x) will denote the method with the sparsity threshold x, and wICA(T) will denote wICA with the denoising threshold T. Argument T is omitted if T is equal to the default threshold $T_0 = \sqrt{2\log(5000)} = 4.1273$.

The features of the signal processed by competitive methods are shown in the Table 4. In order to save space, we display only the first two features (standard deviation of the signal and maximum amplitude) in the falling asleep stage.

We note that SAR(2.5) is rather conservative in removing artifacts compared to RAR, because it compensates the presence of artifacts from 132 ± 81 to 53 ± 23 in place of 45 ± 12 in the case of the first characteristic. Note the twice larger variance of the characteristic compared to RAR. The larger variance is an indicator of the residual presence of artifacts in the cleaned data, which was observed by inspection of individual cases. The results for the second characteristic confirm the observed behaviour of the method. If the denoising threshold x is reduced from 2.5 to 2, the algorithm becomes more aggressive, but the variance still increases. Moreover, SAR(2.0) significantly affects the artifact-free signal.

	falling asleep							
	artifact-free				contaminated			
no	SAR(2.5)	SAR(2.0)	wICA	wICA(25)	SAR(2.5)	SAR(2.0)	wICA	wICA(25)
1	40 ± 11	35 ± 10	23 ± 5	32 ± 9	53 ± 23	49 ± 24	34 ± 12	45 ± 19
2	262 ± 69	225 ± 65	153 ± 39	218 ± 63	354 ± 145	334 ± 157	273 ± 121	362 ± 197

Table 4. The first two characteristics (std and amplitude) of the studied EEG signal in the falling asleep stage processed by SAR(2.5),SAR(2.0), wICA and wICA(25). The nominal (expected) characteristics obtained for artifact-free signals are 45 ± 12 and 281 ± 72, respectively (cf. Table 3).

The other simpler method, wICA with the default denoising threshold, is too aggressive and removes too much of the signal. If the denoising threshold is increased to $T = 25$, the mean value of the first characteristic is close to its expected value, but the other characteristic is spoiled. Apparently it is not possible to tune up both the first and second characteristics with the aid of a single tuning variable (T). Moreover, the method significantly affects the artifact-free signal.

These results prove that RAR outperforms SAR and wICA in removing artifacts of the considered type in neonatal EEG data.

4. Conclusions

In this article, the Robust Artifact Removal (RAR) method has been presented. The method has proved to be suitable for rejecting artifacts that stand out either in amplitude or in frequency (due to the standard low-pass filter). The artifact-free parts of the signal remain largely unaffected. RAR was shown to perform better than either the wavelet-enhanced ICA or the simple artifact rejection method (SAR). The RAR method can be used as a preprocessing step in identification of sleep stages of neonatal infants. If the purpose is different, we admit that the algorithm is not yet able to distinguish high voltage short-duration artifacts from a high voltage short-duration pathological activity. The algorithm allows us to at least indicate both kinds of events and separate them from other background EEG activity.

Matlab code of the RAR method has been posted on the Internet§.

Acknowledgments

This work was supported by Ministry of Education, Youth and Sports of the Czech Republic through Project 1M0572, and by the Czech Science Foundation through Project 102/09/1278.

§ http://si.utia.cas.cz/downloadPT.htm

References

Akhtar MT, Mitsuhashi W and James CJ 2012 Employing spatially constrained ICA and wavelet denoising, for automatic removal of artifacts from multichannel EEG data *Signal Processing* **92** 401-416

Bell AJ and Sejnowski TJ 1995 An information-maximization approach to blind separation and blind deconvolution. *Neural computation* **6** 1129-1159

Belouchrani A, Abed-Meraim K, Cardoso JF, Moulines E. 1997 A blind source separation technique using second-order statistics. *IEEE Tr. Signal Processing* **45** 434-444.

Castellanos NP, Makarov VA 2006 Recovering EEG brain signals: Artifact suppression and wavelet enhanced independent component analysis *J. Neurosci. Methods* **158** 300-312

Celka P, Boashash B, Colditz P 2001 Preprocessing and time-frequency analysis of newborn EEG seizures *IEEE Engineering in Medicine and Biology Magazine* **20** 30 - 39

Delorme A, Sejnowski T, Makeig S 2007 Improved rejection of artifacts from EEG data using high-order statistics and independent component analysis. *Neuroimage* **34** 1443-1449

Delorme A, Palmer J, Onton J, Oostenveld R, Makeig S (2012) Independent EEG sources are dipolar. *PLoS ONE* **7**(2): e30135. doi:10.1371/journal.pone.0030135

De Vos M, De Lathauwer L, Van Huffel S 2011a Spatially constrained ICA algorithms with applications in EEG processing *Signal Processing* **91** 1963-1972.

De Vos M, Deburchgraeve W, Cherian PJ, Matic V, Swarte RM, Govaert P, Visser GH, Van Huffel S 2011b Automated artifact removal as preprocessing refines neonatal seizure detection *Clin Neurophysiol.* **122** 2345-54

Gerla V, Paul K, Lhotska L, Krajča V 2009 Multivariate analysis of full-term neonatal polysomnographic data *IEEE Tr. on Inf. Technology in Biomedicine* **13** 104-110

Hesse CW, James CJ 2005 The FastICA Algorithm with spatial constraints *IEEE Signal Processing Letters* **12** 792-795

Hyvärinen A, Oja E 1997 A Fast fixed-point algorithm for independent component analysis *Neural Computation* **9** 1483-1492

Ille N 2001 Artifact correction in continuous recordings of the electro- and magnetoencephalogram by spatial filtering. Ph.D. dissertation, Univ. Mannheim, Germany

James CJ, Hesse CW 2005 Independent component analysis for biomedical signals *Physiol. Meas.* **26** R15-R39

Joyce CA, Gorodnitsky IF, Kutas M 2004 Automatic removal of eye movement and blink artifacts from EEG data using blind component separation *Psychophysiology* **41** 313-325

Klemm M, Haueisen J and Ivanova G 2009 Independent component analysis: comparison of algorithms for the investigation of surface electrical brain activity *Med Biol Eng Comput* **47** 413-423

Makeig S, Bell AJ, Jung TP, Sejnowski TJ 1996 Independent component analysis of encephalographic data *Adv. Neural Inf. Process. Syst.* **8** 145-151

Pham DT, Cardoso JF 2001 Blind separation of instantaneous mixtures of non stationary sources *IEEE Tr. Signal Processing* **49** 1837-1848

Phlypo R, Van Hese P, Hallez H, Boon P, Van de Walle R, D'Asseler Y and Lemahieu I 2006 Extracting the spike process from the EEG by spatially constrained ICA *Proc. IEEE EMBS, New York, USA* 5286 - 5289

Quiroga RQ, Garcia H 2003 Single-trial event-related potentials with Wavelet Denoising *Clin. Neurophysiology* **114** 376-390

Tichavský P, Yeredor A 2009 Fast approximate joint diagonalization incorporating weight matrices *IEEE Tr. Signal Proc.* **57** 878-891

Tichavský P, Koldovský Z 2011 Fast and accurate methods of independent component analysis: A survey. *Kybernetika* **47** 426-438