Chapter 6

Review of Artifact Rejection Methods for Electroencephalographic Systems

Suguru Kanoga and Yasue Mitsukura

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/68023

Abstract

Technologies using electroencephalographic (EEG) signals have been penetrated into public by the development of EEG systems. During EEG system operation, recordings ought to be obtained under no restriction of movement for routine use in the real world. However, the lack of consideration of situational behavior constraints will cause technical/biological artifacts that often mixed with EEG signals and make the signal processing difficult in all respects by ingeniously disguising themselves as EEG components. EEG systems integrating gold standard or specialized device in their processing strategies would appear as daily tools in the future if they are unperturbed to such obstructions. In this chapter, we describe algorithms for artifact rejection in multi-/single-channel. In particular, some existing single-channel artifact rejection methods that will exhibit beneficial information to improve their performance in online EEG systems were summarized by focusing on the advantages and disadvantages of algorithms.

Keywords: electroencephalographic signal, artifact rejection, blind source separation, regression, filtering, signal decomposition, non-negative matrix factorization

1. Introduction

Variegated branching patterns and trends of sympathetic neurons for realizing the brain function/dysfunction have yet to be completely definitized so far. A functional neuroimaging technique of the human brain has established itself as a trustworthy visible tool to definitize indeterminate patterns and discover new functions [1]. Indeed, visualized information through neuroimaging techniques has contributed building intuitive understanding and relative quantification of brain functions [2, 3].

Key benefits of the electroencephalographic (EEG) modality hold over other neuroimaging techniques (e.g., local field potential, near infrared spectroscopy, and electrocorticogram) are the high

temporal resolution on the order of milliseconds, the small installation space for operating systems, and its usability in noninvasive recording [4]. Although the spatial resolution and specificity are low because it observes the volume conduction effects in brain network [5], this has been attracted attention as a viable and inexpensive modality to study kaleidoscopic functional states of the cerebral cortex: where, when, how, and under what our brain functions come into being [6]. Therefore, providing a capacity to adapt EEG systems to real environments is always a major challenge for neuroscientists and neuroengineers on the final stretch of constructing systems.

Using an extremely small number of electrodes (the single-electrode case would be an extreme case) for signal acquisition should result in better practical application in daily life. Recently, specialized (headband type or headset type) devices, which are endowed with small number of electrodes less than gold standard devices having 16, 32, 64, or more channels, have been developed as for compact, portable, and feasible EEG systems to use themselves in the real environments [7]. The devices are usually implemented with dry electrodes and wireless sensor network technology for recordings. These can diminish the burden on the user caused by oppressive feeling in the head, eliminate the discomfort from conductive gel or paste, and improve degree of freedom of movements by doing away with wires plugged into an amplifier [8].

However, technical/biological artifacts, such as active power line interference, eyeblink, and muscle activity caused by recording mistake, good conductivity of the scalp, and so on, are often mixed with EEG signals whether the type of device is gold standard or specialized. They ingeniously disguise themselves as EEG components in observed EEG signals and cause a discrepancy between research motivation and system realization. Removing mimetic components (artifacts) or extracting intrinsic EEG components from observed EEG signals will become a more important process in all EEG systems for practical use even if single electrode is integrated with data acquisition module by a specialized device.

Disclosing the meaning of electric signals comprising various neuronal populations (sources) breaks down the EEG inverse (blind source separation (BSS)) problem [9]. It is well known that the enormous indeterminacies in brain make the BSS problem ill-posed; however, statistical natures lead to restoring the well-posedness of the problem in a biosignal processing. By the properties, theoretically multivariate statistical analysis approaches like independent component analysis (ICA) can separate observed EEG signals into spatially and temporally distinguishable components effectively, and then, estimated components will be identified as neuronal or artifactual sources by hard/soft threshold to reconstruct artifact-free EEG matrix [10, 11]. Whereas there are several reviews on artifact rejection methods including overall procedure (signal separation, component identification, and signal reconstruction) for multi-channel EEG signals [12–16], we have never seen review of artifact rejection methods for single-channel EEG signals. In this chapter, we therefore describe algorithms for artifact rejection in multi-/single-channel EEG signals.

2. Concise description of technical/biological artifacts

2.1. Technical artifacts

Technical artifacts such as power line interference, impedance fluctuation, and wire movement superimpose their energy on observed EEG signals because of faults in setting conditions

http://dx.doi.org/10.5772/68023

[18, 19]. These can be precluded from easy ways, detaching a charging AC adapter from the recording device, carefully attaching electrodes to the scalp, and using appropriated electrode wires or adhesive tapes to stabilize wires shown in **Figure 1**. The cross mark in the figure indicates detaching the source of technical artifact from the setting conditions.

2.2. Biological artifacts

Biological artifacts, which are discharged potentials of internal organs, diffuse their energy over the head and reach each electrode attaching on the surface of the scalp as observed EEG signal. They contaminate observed signals due to the iron accumulation in the brain and good conductivity of the scalp can be broadly separated into four categories: (i) muscular, (ii) cardiac, (iii) eyemovement, and (iv) eyeblink. EEG devices capture comprehensive electric field which was reached at an electrode even if the potential contains information of electrophysiological actions except neuronal one (see **Figure 2**). Because all electrical potentials will be equally and blindly treated, recording information including only EEG components from electrodes placed on the scalp is hardly realized. Furthermore, frequency characteristics of biological artifacts and neuronal oscillations could be overlapped. That means that shunning contact with biological artifacts may seem hopelessly difficult compared with technical

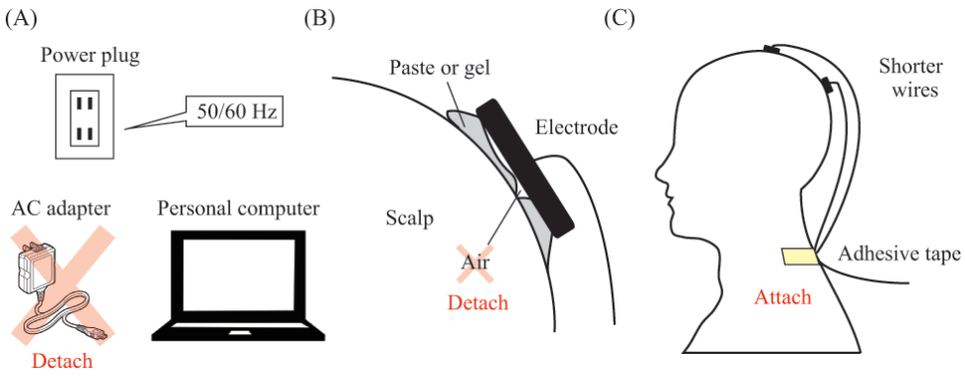

Figure 1. Ways of precluding technical artifacts [17]. (A) Power line interference. (B) Impedance fluctuation. (C) Wire movement.

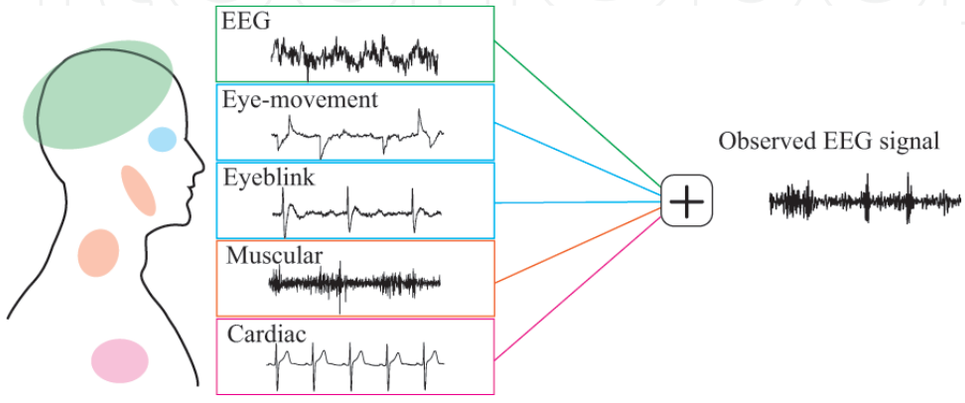

Figure 2. Configuration of an observed EEG signal including biological artifacts.

artifacts. If contaminated epochs are found in visual or quantitative analysis, the EEG system has to ignore them before deciding control commands. Otherwise, the operator will make a fatal mistake in its system by counterfeit EEG patterns [12, 17].

Alternatively, signal processing techniques can extract EEG components from observed signals. Through this process, EEG systems would provide correct outputs for their unique and beneficial interface. Even today, many works for detection, classification, and removal of artifacts within observed EEG signals have been reported [20–22].

3. Review of existing methods on artifact rejection

In this section, the standard assumptions of observed cerebral signal for spatially and temporally separating components are described before introduction of artifact rejection methods to reach deep understanding of the statistical framework. Then, methods of multi-/single-channel artifact rejection (principal component analysis (PCA), independent component analysis (ICA), regression, filtering, ICA-based signal decomposition, and nonnegative matrix factorization) are presented. Each algorithm has specialized approaches for calculating demixing matrix, identifying separated components, and denoising the artifactual components to complete source separation. We have focused on the advantages and disadvantages of approaches.

3.1. In multi-channel signals

3.1.1. Standard assumption of sources

The first thing that all artifact rejection methods have to do is calculating demixing matrix $\boldsymbol{W}$ under the standard assumption of sources regardless of the target object. In EEG signal processing, the observed cerebral signal $x(n)$ is considered as the sum of the cerebral source (local-field) activity $s(n)$ and the noise/artifact $d(n)$. Neuronal cells have limited their connection ability to short-range order (less than 500 μm) [23]. Besides, synchrony in local-field activities diffuses through a contiguous cortical area rather than jump between distant and weakly connected cortical areas [24].

Therefore, an assumption that cerebral sources and non-cerebral sources are linearly combined, allows the following formulation of the underlying biophysics of the signal generation and propagation of the potential [25]:

$$\boldsymbol{x}(n) = \boldsymbol{A}\boldsymbol{s}(n) + \boldsymbol{d}(n), \tag{1}$$

where: $\boldsymbol{x}(n) = [x_1(n),\ x_2(n), \ldots,\ x_P(n)]^{\mathrm{T}}$ is the observed P-channel EEG data at the n-th point (superscript T means the transpose of a vector or matrix); $\boldsymbol{s}(n) = [s_1(n),\ s_2(n), \ldots,\ s_Q(n)]^{\mathrm{T}}$ is the Q unknown source data, in which each row means cerebral or non-cerebral source; $\boldsymbol{A}$ is the $P \times Q$ full-rank unknown mixing matrix; and $\boldsymbol{d}(n) = [d_1(n),\ d_2(n), \ldots,\ d_P(n)]^{\mathrm{T}}$ is the P additive zero-mean noise data. In real scenarios, there are likely to be more sources than observations ($Q > P$); however, handing the number of sources the same as the number of observations

http://dx.doi.org/10.5772/68023

(Q = P) does not normally become a fatal problem. Thus, most algorithms extract a linear combination of sources belonging to the same subspace [26, 27].

All algorithms have a common disadvantage that they can only handle over-determined mixture for the inverse process while having no priori information on the characteristics of the sources. Additional three assumptions are reluctantly accepted: (i) the noise/artifact is spatially uncorrelated with the observed data ($\mathbb{E}[\boldsymbol{A}\boldsymbol{s}(n)\boldsymbol{d}(n)^{\mathrm{T}}] = \boldsymbol{0}$, where $\mathbb{E}[\cdot]$ is the expectation operator), and temporally uncorrelated ($\mathbb{E}[\boldsymbol{d}(n)\boldsymbol{d}(n+\tau)^{\mathrm{T}}] = \boldsymbol{0}$, where τ is lag time and $\forall\tau > 0$); (ii) the number of sources is equal to or less than the number of observations ($Q \leq P$); and (iii) the mixing matrix $\boldsymbol{A}$ is stationary [28].

3.1.2. Blind source separation algorithms

Under aforementioned assumptions, BSS approaches estimate sources $\hat{\boldsymbol{S}} = [\hat{\boldsymbol{s}}(1), \ldots, \hat{\boldsymbol{s}}(N)]$ from observed EEG data $\boldsymbol{X} = [\boldsymbol{x}(1), \ldots, \boldsymbol{x}(N)]$. Unsupervised learning methods such as PCA and ICA jointly estimate demixing matrix $\boldsymbol{W}$ $(= \boldsymbol{A}^{-1})$:

$$\hat{\boldsymbol{s}}(n) = \boldsymbol{W}\boldsymbol{x}(n). \tag{2}$$

Each unsupervised learning method has an algorithm that is subject to various indices: uncorrelatedness, independence, non-Gaussianity, instantaneous propagation, and linearity [29]. Linear mixture concept of blind EEG source separation is shown in **Figure 3** that presents a demixing matrix $\boldsymbol{W}(=\boldsymbol{W}_1\boldsymbol{W}_2)$ as two-step estimator because some methods firstly decorrelate an observed matrix by $\boldsymbol{W}_1$ and then demix it by $\boldsymbol{W}_2$. Given a mixing matrix $\boldsymbol{A}$ is composed of the three blind cerebral sources $\boldsymbol{s}(n)$ and provides the same number of observations $\boldsymbol{x}(n)$ in the figure.

PCA converts the observed matrix of possibly correlated variables into values of linearly uncorrelated variables (principal components (PCs)) with the first-and second-order statics [30]. This algorithm conducts the eigenvalue decomposition to get the directions $\boldsymbol{u}$ of greater

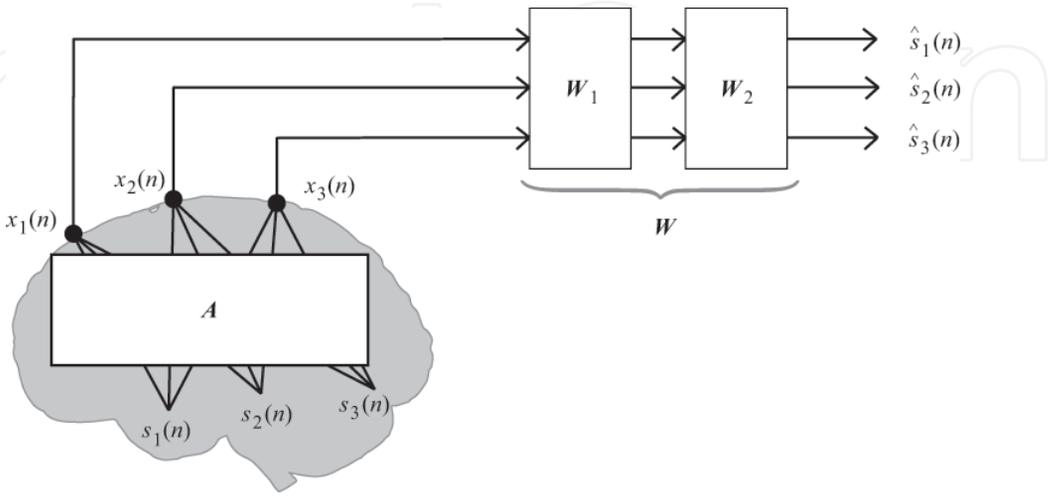

Figure 3. Linear mixture concept of blind EEG source separation [15, 17].

variance in the input space of the EEG data $\boldsymbol{X}$ based on assumptions that data are jointly normally distribution, and the sources are uncorrelated. In order to satisfy the assumptions, obtained matrix $\boldsymbol{X}^{\text{old}}$ should be standardized to decorrelate samples of the same dimension ($\mathbb{E}[\boldsymbol{x}(n)\boldsymbol{x}(n+\tau)^{\mathrm{T}}] = \boldsymbol{0}$) and to uniform unit ($\mathbb{V}[\boldsymbol{X}_p] = 1$).

In PCA algorithm, the first PC, which has the largest variance in the standardized input space, is a linear combination of $\boldsymbol{X}$ defined by weights $\boldsymbol{u}_1 = [u_1, \ldots, u_P]^{\mathrm{T}}$:

$$\mathrm{PC}_1 = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{u}_1, \tag{3}$$

$$\mathbb{V}[\mathrm{PC}_1] = \mathbb{V}[\boldsymbol{X}^{\mathrm{T}}\boldsymbol{u}_1] = \boldsymbol{u}_1{}^{\mathrm{T}}\sum \boldsymbol{u}_1, \tag{4}$$

where $\sum (= \boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}/(N-1))$ is covariance matrix of $\boldsymbol{X}$. Therefore, this algorithm formulates the given problem in an optimization problem:

$$\max \boldsymbol{u}_1{}^{\mathrm{T}}\sum \boldsymbol{u}_1, \tag{5}$$

$$\text{subject to } \boldsymbol{u}_1{}^{\mathrm{T}}\boldsymbol{u}_1 = 1. \tag{6}$$

It can be solved by Lagrange multiplier method:

$$L(\boldsymbol{u}_1, \lambda_1) = \boldsymbol{u}_1{}^{\mathrm{T}}\sum \boldsymbol{u}_1 + \lambda_1(1 - \boldsymbol{u}_1{}^{\mathrm{T}}\boldsymbol{u}_1), \tag{7}$$

$$\frac{\partial L(\boldsymbol{u}_1, \lambda_1)}{\partial \boldsymbol{u}_1} = 2\sum \boldsymbol{u}_1 - 2\lambda_1\boldsymbol{u}_1 = 0, \tag{8}$$

$$\boldsymbol{u}_1{}^{\mathrm{T}}\sum \boldsymbol{u}_1 = \lambda_1\boldsymbol{u}_1{}^{\mathrm{T}}\boldsymbol{u}_1 = \lambda_1. \tag{9}$$

The covariance matrix $\sum$ is sequentially decomposed into eigenvector $\boldsymbol{u}_p$ and eigenvalue λ_p by an assumption that the PCs are orthogonal. The eigenvector $\boldsymbol{u}_p$ is similar to the column of the inverse demixing matrix $\boldsymbol{W}^{-1}$. PCA-based methods have an advantage over stationary data; however, satisfying their assumption for EEG data is difficult [31]. On the other hand, PCA algorithm is often incorporated into a first decorrelation or whitening step of some ICA algorithms [32].

ICA is the most famous and prevalent unsupervised learning algorithm to decompose multi-channel EEG data $\boldsymbol{X}$ into independent components (ICs) $\hat{\boldsymbol{S}}$ with high-order (spatial) moments, beyond the second-order statics used in PCA, whereas some algorithms use the statics as well as PCA [4]. A state-of-the-art topical review published on 2015 reported that second order blind interference (SOBI) and information maximization (InfoMax) are the most commonly used algorithm for EEG signal processing [15]. In this chapter, we describe InfoMax algorithm.

The fundamental problem tackled by InfoMax ICA is how to minimize the mutual information (MI) of the output vector $\hat{\boldsymbol{s}}$,

http://dx.doi.org/10.5772/68023

$$\mathrm{MI}(\hat{\boldsymbol{s}}) = \sum_{p=1}^{P} H(\hat{\boldsymbol{s}}_p) - H(\hat{\boldsymbol{s}}). \tag{10}$$

Probability density functions of observed signal $p(\boldsymbol{x})$ and estimated signal $p(\hat{\boldsymbol{s}})$ have following relationship:

$$p(\hat{\boldsymbol{s}})d\hat{\boldsymbol{s}} = p(\boldsymbol{x})d\boldsymbol{x}, \tag{11}$$

$$d\hat{\boldsymbol{s}} = J(\boldsymbol{x})d\boldsymbol{x} = |\boldsymbol{W}|d\boldsymbol{x}, \tag{12}$$

$$p(\hat{\boldsymbol{s}}) = p(\boldsymbol{x})d\boldsymbol{x} = p(\boldsymbol{W}^{-1}\hat{\boldsymbol{s}})|\boldsymbol{W}|^{-1}, \tag{13}$$

where $J(\boldsymbol{x})$ is Jacobian matrix. The estimating entropy $H(\hat{\boldsymbol{s}})$ is given by:

$$\begin{aligned} H(\hat{\boldsymbol{s}}) &= -\int p(\hat{\boldsymbol{s}}) \log p(\hat{\boldsymbol{s}}) d\hat{\boldsymbol{s}} \\ &= -\int \Big(\log p(\boldsymbol{x}) - \log|\boldsymbol{W}|\Big) p(\boldsymbol{x}) d\boldsymbol{x} \\ &= -\int p(\boldsymbol{x}) \log p(\boldsymbol{x}) d\boldsymbol{x} + \log|\boldsymbol{W}| \\ &= H(\boldsymbol{x}) + \log|\boldsymbol{W}|. \end{aligned} \tag{14}$$

Therefore, the MI can be rewritten as following:

$$\mathrm{MI}(\hat{\boldsymbol{s}}) = \sum_{p=1}^{P} H(\hat{\boldsymbol{s}}_p) - H(\boldsymbol{x}) - \log|\boldsymbol{W}|. \tag{15}$$

By partially differentiating this index on parameters $\boldsymbol{W}$, optimized solution for source separation will be obtained.

$$\frac{\partial \mathrm{MI}(\hat{\boldsymbol{s}})}{\partial \boldsymbol{W}} = \sum_{p=1}^{P} \frac{\partial\left(-\int p(\hat{\boldsymbol{s}}) \log p(\hat{\boldsymbol{s}}_p) d\hat{\boldsymbol{s}}\right)}{\partial \boldsymbol{W}} - (\boldsymbol{W}^{\mathrm{T}})^{-1} = -\mathbb{E}[\boldsymbol{\varphi}(\hat{\boldsymbol{s}})\boldsymbol{x}^{\mathrm{T}}] - (\boldsymbol{W}^{\mathrm{T}})^{-1}, \tag{16}$$

where

$$\varphi(\hat{s}_p) = \frac{d \log p(\hat{\boldsymbol{s}}_p)}{d\hat{\boldsymbol{s}}_p}. \tag{17}$$

As analytical computation of equation as mentioned above is difficult, this algorithm uses a gradient update rule based on the natural gradient [33] and learning rate η that is a positive constant:

$$\boldsymbol{W} \leftarrow \boldsymbol{W} + \eta \Delta \boldsymbol{W}, \tag{18}$$

$$\Delta \boldsymbol{W} = \left(\mathbb{E}[\boldsymbol{\varphi}(\hat{\boldsymbol{s}})\boldsymbol{x}^{\mathrm{T}}] + (\boldsymbol{W}^{\mathrm{T}})^{-1} \right) \boldsymbol{W}^{\mathrm{T}} \boldsymbol{W} = \left(\mathbb{E}[\boldsymbol{\varphi}(\hat{\boldsymbol{s}})\hat{\boldsymbol{s}}^{\mathrm{T}}] + \boldsymbol{I} \right) \boldsymbol{W}. \tag{19}$$

3.1.3. Component identification after source separation

After source separation, estimated sources $\hat{\boldsymbol{S}}$ have to be continuously identified as neuronal or artifactual sources to reconstruct artifact-free EEG matrix $\hat{\boldsymbol{X}}$. Visual inspection of scalp topography and empirical judgment was given the credit for identification of components [10, 14]. The overused techniques are still examined in an expedient manner for checking the results. That leads to increase in workload; therefore, hard/soft-threshold function, probability approach, and machine learning algorithm with features of the prepared material have been used for automatically identifying artifacts in estimated sources to reduce the workload and to get more repeatable labels [34, 35]. Proposing automatic and unsupervised component identification algorithm to characterize more precisely and flexibly has still been an active research area [36, 37]. Once estimated sources are identified, they advance to next step called denoising step, and then an underlying EEG matrix will be reconstructed using inverse linear demixing process (see **Figure 4**).

3.2. In single-channel signals

3.2.1. Discrepancy among standard assumptions about multi-/single-channel data

We can easily imagine that single-channel data do not always satisfy the assumptions for BSS techniques. Calculating demixing matrix $\boldsymbol{W}$ is especially difficult with single-channel artifact rejection methods (see **Figure 5**), so that researchers are forced to select whether to add information by using the reference channel before applying a method or to separate data by using only one-channel.

3.2.2. Regression

Regression algorithm was most frequently used to remove artifact up to the mid-1990s [38, 39]. In this algorithm, an observed EEG signal $x(n)$ can be expressed as

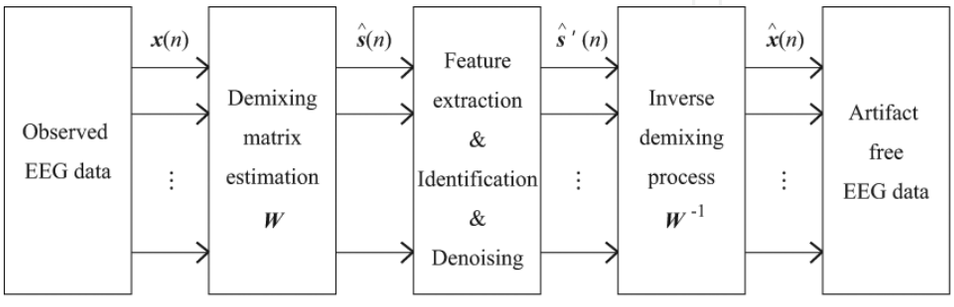

Figure 4. Block diagram of the blind source separation [11].

http://dx.doi.org/10.5772/68023

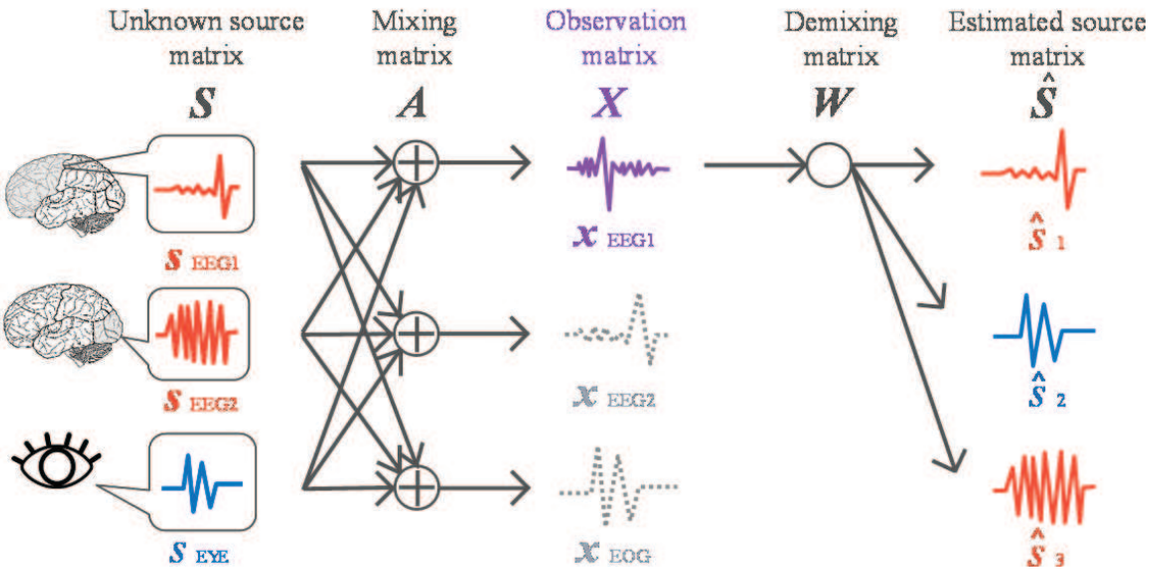

Figure 5. Procedure of signal separation in single-channel artifact rejection methods.

$$x(n) = x_{\mathrm{EEG}}(n) + x_{\mathrm{Art}}(n) + d(n), \tag{20}$$

where $x_{\mathrm{EEG}}(n)$, $x_{\mathrm{Art}}(n)$, and $d(n)$ are intrinsic EEG data, artifact, and noise. It is assumed that the expected value of $d(n)$ is 0.

The artifact would be corrected by calculating propagation factors to estimate the relationship between the reference signal $x_{\mathrm{Ref}}(n)$ and the observed EEG signal and subtracting the regressed portion [40]. The rationale of the procedure is as follows:

Step 1. Separately average over observed EEG and reference signals of T trials to estimate the artifact waveform related variation for the channels:

$$\overline{x}(n) = \frac{1}{T}\sum_{t=1}^{T} x_t(n), \tag{21}$$

Step 2. Subtract the averages from every trial data to obtain deviations:

$$\boldsymbol{x}'(n) = \boldsymbol{x}(n) - \overline{\boldsymbol{x}}(n), \tag{22}$$

where $\overline{\boldsymbol{x}}(n)$ is duplicated $T \times 1$ matrix of the observed EEG average,

Step 3. Calculate the propagation factor C by linear least-square regression whereby the observed EEG data are considered as a dependent variable and the reference data are considered as the independent variable:

$$\boldsymbol{X} = C(\boldsymbol{X}_{\mathrm{Ref}}), \tag{23}$$

where

$$\boldsymbol{X} = [\boldsymbol{x}'(1), \ldots, \boldsymbol{x}'(t), \ldots, \boldsymbol{x}'(T)]^{\mathrm{T}}, \tag{24}$$

$$\boldsymbol{x}'(t) = [x'\Big(1 + N(t-1)\Big), \ \ldots, \ x'(tN)], \tag{25}$$

Step 4. Correct the observed EEG data by subtracting the reference data scaled by the propagation factor C:

$$\hat{x}(n) = x(n) - C\Big(x_{\mathrm{Ref}}(n)\Big). \tag{26}$$

Because averaging operator emphasizes a time-locked activity in observed EEG signals, this method requires a reference channel and is powerful only if the operating system treats event-related brain potentials. Cerebral activities are usually not time-locked that means that important nontime-locked components will be lost by the averaging operation. Furthermore, this method does not take bidirectional contamination into account and cancels the cerebral information from each observed EEG signal upon linear subtraction [41]. Despite its disadvantages, regression is still used as the "gold standard" method to which the performance of any artifact rejection algorithms may be compared.

3.2.3. Filtering

Band-pass is one of the classical and simple separation attempts to remove artifacts from an observed EEG signal. This method is effective if the spectral distributions of the EEG component and artifact do not overlap, and there are small band artifacts such as power line noise (50/60 Hz interference) [42]. However, fixed-gain filtering is not effective for biological artifacts because it will attenuate EEG component and change both amplitude and phase of signal if the filtering keeps doing that [43]. Some adaptive algorithms try to adapt the filter parameters $\boldsymbol{w}$ to minimize the error between the artifact-free EEG signal $\hat{\boldsymbol{x}}(n)$ and the desired original signal $\boldsymbol{x}(n)$ to suppress the limitations of this method.

Adaptive filtering assumes that the intrinsic EEG signal and artifact are uncorrelated; therefore, the artifact is considered to be an additive noise within the observed signal:

$$x_t(n) = s_t(n) + n_{0t}(n), \tag{27}$$

where $x_t(n)$ is the observed EEG signal of t-th trial, $n_0(0)$ is the additive noise to offset and is uncorrelated with intrinsic EEG signal $s_t(n)$. The filter parameters $\boldsymbol{w}$ are iteratively adjusted by a feedback (recursive) process designed to make the output as close as possible to some desired response with an additive noise interference [44, 45]. **Figure 6** shows the noise canceller system using adaptive filtering. In this system, the primary input $x_t(n)$ and the reference input $x_{\mathrm{Ref}t}(n)$ are the observed EEG and reference signals. A reference input $x_{\mathrm{Ref}t}(n) = n_{1t}(n)$ which is a noise correlated with $n_{0t}(n)$ and uncorrelated with intrinsic EEG signal $s_t(n)$, adds information to minimize the error $e_t(n)$ between the response $y_t(n)$ and the desired response.

http://dx.doi.org/10.5772/68023

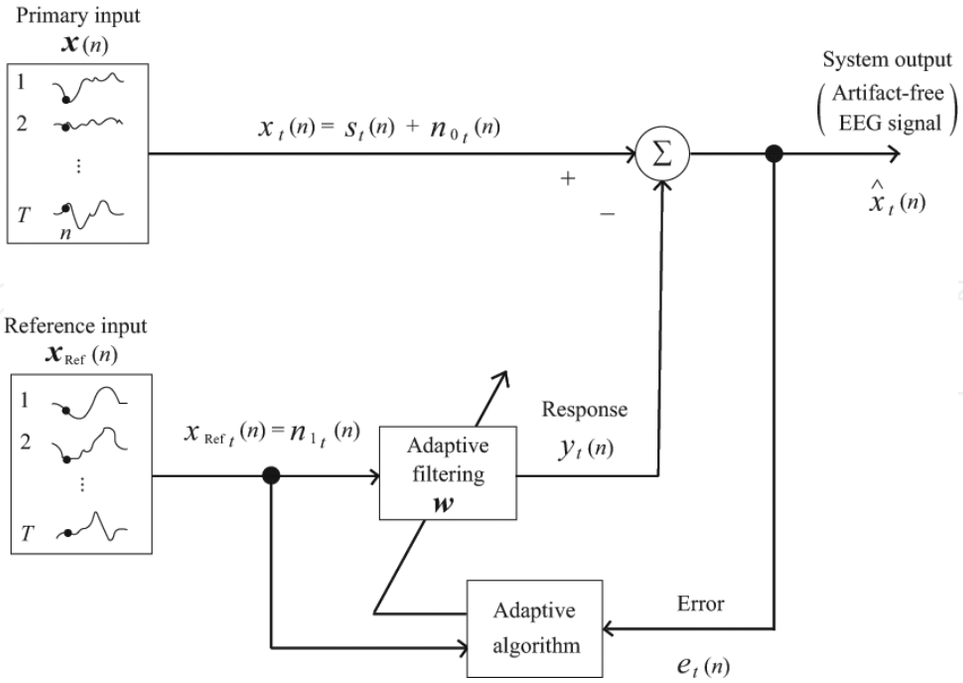

Figure 6. Noise canceller system using adaptive filtering [17, 47].

Recursive least squares (RLS)-based adaptive filtering presents a superior performance than least mean squares-based one [46]. The algorithm can be implemented using the following equations:

$$\boldsymbol{g}(n) = \frac{\boldsymbol{R}(n-1)\boldsymbol{x}_{\mathrm{Ref}}(n)}{\lambda + \boldsymbol{x}_{\mathrm{Ref}}{}^{\mathrm{T}}(n)\boldsymbol{R}(n-1)\boldsymbol{x}_{\mathrm{Ref}}(n)}, \tag{28}$$

$$\boldsymbol{e}(n) = \boldsymbol{x}(\mathrm{n}) - \boldsymbol{y}(\mathrm{n}), \tag{29}$$

$$\boldsymbol{y}(n) = \boldsymbol{w}(n)\boldsymbol{x}_{\mathrm{Ref}}(n), \tag{30}$$

$$\boldsymbol{R}(n) = \frac{\boldsymbol{R}(n-1) - \boldsymbol{g}(n)\boldsymbol{x}_{\mathrm{Ref}}{}^{\mathrm{T}}(n)\boldsymbol{R}(n-1)}{\lambda}, \tag{31}$$

$$\boldsymbol{w}(n) = \boldsymbol{w}(n-1) + \boldsymbol{g}(n)\boldsymbol{e}(n), \tag{32}$$

where $\boldsymbol{g}(n)$ and $\boldsymbol{w}(n)$ are the gain vector and the filtering parameters. The initial value of cross-correlation $\boldsymbol{R}(0)$ is $\delta\boldsymbol{I}$, where δ and $\boldsymbol{I}$ are some sufficiently large positive value and identity matrix. The updated filter parameters lead to output artifact-free EEG signal.

Consequently, adaptive filtering approach has a potential to recover "pure" EEG signal more rapidly and accurately than linear regression for ocular and cardiac artifacts [48]. However, it is rather difficult to converge to the solution of filtering parameters if muscular and vibration artifacts have contaminated in the observed EEG signal. In that situation, the algorithm sometimes does not converge because of their convulsive burst.

Optimal filtering like Kalman filtering can capture non-stationary properties of artifacts. The framework has flexibility for non-linear system due to approximating the probability density function that might lead to more effective artifact rejection method. Many works on filtering algorithms have developed this approach for more useful module in real-time applications [49, 50].

3.2.4. ICA-based signal decomposition

ICA will achieve an artifact rejection with an outstanding performance if the number of independent sources is equal to or lower than observations. Unfortunately, this method is only applicable to multi-channel data; however, some works extended the idea to single-channel data to unmix a set of observed signals (components) into intrinsic sources [51–53]. These methods decompose a single-channel into multiple components by dividing into a sequence of blocks or different spectral modes before applying ICA so that we call these methods ICA-based signal decomposition approaches (see **Figure 7**).

Single-channel ICA is the oldest method for single-channel data under an assumption that stationary sources are being disjoint in the frequency domain [54]. An observed signal $x(n)$ is split up into K short segments $\boldsymbol{X}$, a sequence of contiguous blocks of length L which is to be handled as a set of observations.

$$\boldsymbol{X} = [\boldsymbol{x}(1), \ldots, \boldsymbol{x}(k), \ldots, \boldsymbol{x}(K)]^{\mathrm{T}}, \tag{33}$$

$$\boldsymbol{x}(k) = [x\Big(L(k-1)+1\Big), \ldots, x(kL)]^{\mathrm{T}}, \tag{34}$$

where k is the block index. A standard ICA algorithm than performs to the matrix $\boldsymbol{X}$ to derive the demining matrix $\boldsymbol{W}$. The artifacts overlap with EEG components and EEG signal has non-periodic components; therefore, this method can be applied within limited situations. Wavelet transform (WT)-based and empirical mode decomposition (EMD)-based ICA have already been reported successful in removing artifacts for solving the similar problem than single-channel ICA [42].

WT-based ICA transforms an observed signal into components of disjoint spectra (a matrix) instead of signal (a vector) via discrete WT [55].

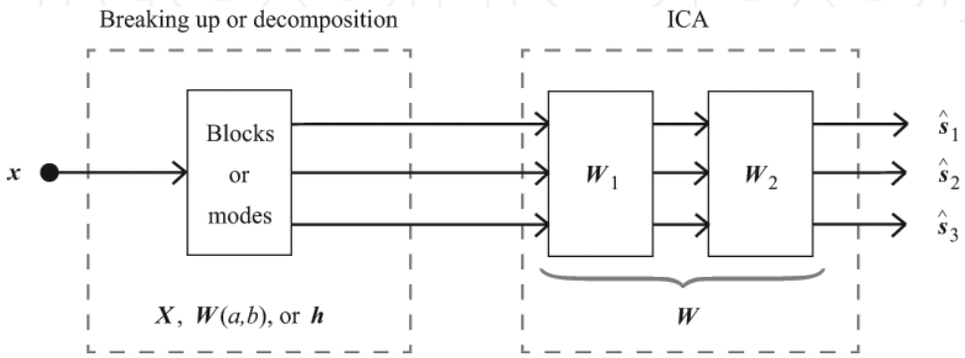

Figure 7. Procedure of separation method using ICA-based signal decomposition.

http://dx.doi.org/10.5772/68023

$$W(a,\ b) = \frac{1}{\sqrt{a}} \int x(n)\psi_{a,b}(n)dn, \tag{35}$$

$$\psi_{a,b} = \psi\left(\frac{n-b}{a}\right), \tag{36}$$

where $W(a, b)$ and $\psi_{a,\ b}$ denote that the wavelet representation of $x(n)$ and the mother wavelet with a and b defining the time-scale and location. The decision of parameters is hard if the user does not have a priori knowledge of the signal of interest. Each IC using wavelet coefficients is, respectively, identified as either neuronal or artifactual by manually. The artifactual ICs are replaced their values with arrays of zeros and then reconstructed to wavelet components. Finally, artifact-free signal is acquired by inverse discrete WT.

EMD-based ICA decomposes an observed signal into a number of K intrinsic mode functions (IMFs) $h_k(n)$,

$$x(n) = \sum_{k=1}^{K} h_k(n) + d(n), \tag{37}$$

where $d(n)$ is a residue of the original data and a nonzero mean slowly varying function with only a few or no extreme [56]. This method can remove artifacts without a priori knowledge regarding characteristics of the signal embedded in the data [57]. Each IMF has monocomponent of the original data and is estimated by an iterative process called "shifting process":

Step 1. Find the local maxima and minima in $x_k(n)$,

Step 2. Connect all of the local maxima and minima by cubic splines to form an upper and a lower envelope,

Step 3. Calculate the mean of the two envelopes, respectively,

Step 4. Obtain improved IMF $h_{k+1}(n)$ by subtracting the mean of the two envelopes from the current IMF $h_k(n)$,

Step 5. Go to Step 1 until the residue is below a stopping criterion.

This decomposition is based on the three conditions: (i) the number of extreme and the number of zero-crossing must be equal or up to plus/minus one; (ii) zero mean; and (iii) all the maxima and all the minima of IMF will be positive and negative everywhere. Each IC using IMFs is, respectively, identified as either neuronal or artifactual by manually as well as WT-based ICA. The artifactual ICs are replaced their values with arrays of zeros. Finally, reconstructed IMFs are summed simply together to acquire artifact-free signal.

WT-based and EMG-based ICA have been reported as superb methods for artifact rejection [51, 58, 59]. Therefore, a certain number of researchers tends to select them over recent years. However, separating intrinsic EEG components and artifacts are not successfully completed by this approach because frequency characteristics of biological artifacts and EEG components

could be overlapped. In addition, a presence of similar oscillations in different modes or a presence of disparate amplitude oscillations in the same mode, named "mode mixing" makes the performance of artifact rejection worse [60]. Signal distortion or attenuation typically occurs according to the above-mentioned methods by excessive interference. Thus, these approaches are not suitable for real-time applications.

3.2.5. Nonnegative matrix factorization

In linear regression, filtering, and ICA-based signal decomposition approaches, parameters $\boldsymbol{W}$ cannot often converge to a solution for perfectly demixing the mixtures. This implies that partially restricting the active space should be determined for single-channel signals.

Meanwhile, non-negative matrix factorization (NMF) [61] has recently attracted attention as effective algorithms to remove artifacts from single-channel signals because it can find the latent features underlying the interactions between EEG components and artifacts. An M-dimensional non-negative data vector $\boldsymbol{x}_n$ is placed in the column of $M \times N$ matrix $\boldsymbol{X}$, where N is number of data vectors. The matrix $\boldsymbol{X}$ is based on short-time Fourier transform and approximately factorized into an $M \times K$ nonnegative matrix $\boldsymbol{H}$ and a $K \times N$ nonnegative matrix $\boldsymbol{W}$ where K is the number of "basis" which is optimized for linear approximation of the input vectors. It can be represented by the following equation:

$$\boldsymbol{x}_n \approx \boldsymbol{y}_n = \sum_{k=1}^{K} \boldsymbol{h}_k w_{k,n}, \tag{38}$$

where an $\boldsymbol{h}_k$ and a $w_{k,n}$ denote an entry of $\boldsymbol{H}$ and $\boldsymbol{W}$. This equation means that respective non-negative EEG feature (power spectrum or amplitude spectrum) vector is approximated by linear combination of the basis vector $\boldsymbol{h}_k$ weighted by the component of $w_{k,n}$. Therefore, it can be rewritten as

$$\boldsymbol{X} \approx \boldsymbol{H}\boldsymbol{W}. \tag{39}$$

Some works reported that the supervised NMF could effectively factorize the observed EEG signals into the brain activity components and the artifacts if the user has artifact data in advance [62, 63]. Before applying supervised learning, template matrix $\boldsymbol{X}_{\mathrm{Art}}$ has been factorized into $\boldsymbol{H}_{\mathrm{Art}}$ and $\boldsymbol{W}_{\mathrm{Art}}$. The matrix $\boldsymbol{X}$ is continuously factorized into $\boldsymbol{H}$ and $\boldsymbol{W}$ where $\boldsymbol{H}$ contains the elements of matrix $\boldsymbol{H}_{\mathrm{Art}}$. The matrix $\boldsymbol{H}_{\mathrm{Art}}$ has no relation to the elements of $\boldsymbol{H}$ while using standard NMF algorithm because the initial values are set randomly and updated by multiplicative rules. In supervised learning algorithm, the matrix $\boldsymbol{H}_{\mathrm{Art}}$ is used as a fixed value that will partially restrict the active space. By contrast, activity components in the matrix $\boldsymbol{W}_{\mathrm{Art}}$ are variable values. For this constraint, the matrix $\boldsymbol{H}$ can attempt to express EEG components in the matrix $\boldsymbol{X}$ with the remaining based K'. EEG components will be stored in the bases (see **Figure 8**).

After these processing, non-negative data of artifact-free EEG are reconstructed from the following equation:

$$\hat{\boldsymbol{X}} = \boldsymbol{X} * \sum_{k=K_{\mathrm{Art}}+1}^{K} \sum_{n=1}^{N} \frac{\boldsymbol{H}_k W_{k,n}}{HW}. \tag{40}$$

http://dx.doi.org/10.5772/68023

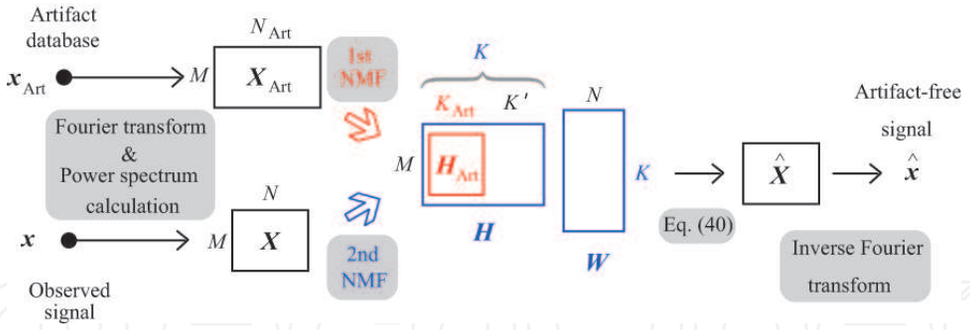

Figure 8. Procedure of supervised NMF.

Eq. (40) and inverse Fourier transform make it possible to acquire artifact-free signal. Supervised NMF is still in its infancy, showed high performance for artifact rejection. However, epoch detention step, which is not part of normal procedures in artifact rejection, must be embedded in the epoch-based method. This leads to increase the computational cost inevitably. Some low-cost (real-time) artifact detection algorithms for single-channel EEG signal [64, 65] are a silver lining in a dark cloud.

4. Conclusions

By the properties of artifacts, theoretically multivariate statistical analysis approaches such as PCA and ICA, which separate multi-channel EEG signals into spatially and temporally distinguishable components, are useful for extracting EEG components from the scalp recordings. In particular, ICA is a powerful tool for separating observed EEG signals into maximally independent activity patterns derived from cerebral or non-cerebral (artifactual) sources. However, ICA is unsuitable for analyzing EEG signals recorded by specialized EEG device because of mismatching of its assumption in the single (or few) channel case. Thus, proposing a removal method of artifact from single-channel EEG signals is currently a major challenge in EEG signal processing for the widespread use of systems as a conventional technology.

In this chapter, we tried to summarize some existing artifact rejection algorithms (PCA, ICA, regression, filtering, ICA-based signal decomposition, and NMF) focusing on the advantages and disadvantages of algorithms, which would provide beneficial information to improve their performance in online EEG systems. Last but not least, muscular artifacts reflecting body actions are natural enemies of EEG systems. The inevitable encounter must be solved by artifact rejection techniques. During real-time EEG system operation using specialized devices, unsupervised learning algorithms cannot separate observed signal into EEG and EMG components so far. Neuroscientists and neuro-engineers should carefully analyze the characteristics of artifacts and integrate them in a supervised learning algorithm for effective rejection of artifacts or extraction of intrinsic EEG components from observed EEG signals without altering the underlying brain activity to routinely use EEG systems in the future.

Acknowledgements

This work was supported in part by NEDO (New Energy and Industrial Technology Development Organization) SIP Number YYN6022-111123.

Author details

Suguru Kanoga* and Yasue Mitsukura

*Address all correspondence to: kanouga@mitsu.sd.keio.ac.jp

Keio University, Kanagawa, Japan

References

[1] Friston KJ. Modalities, modes, and models in functional neuroimaging. Science. 2009;**326** (5951):399-403. DOI: 10.1126/science.1174521

[2] Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. Nature. 2001;**412**(6843):150-157. DOI: 10.1038/35084005

[3] Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. Magnetoencephalography—theory, instrumention, and applications to non-invasive studies of the working human brain. Reviews of Modern Physics. 1993;**65**(2):413-505. DOI: 10.1103/RevModPhys.65.413

[4] Rampil IJ. A primer for EEG signal processing in anesthesia. The Journal of the American Society of Anesthesiologists. 1998;**89**(4):980-1002

[5] Bear MF, Connors BW, Paradiso MA. Brain rhythms and sleep. In: Neuroscienece: Exploring the brain. 3rd ed. USA: Lippincott Williams and Wilkins; 1996. pp. 585-616

[6] Pfurtscheller G, Allison BZ, Brunner C, Bauernfeind G, Solis-Escalante T, Scherer R, et al. The hybrid BCI. Frontiers in Neuroscience. 2010;**4**(30):42. DOI: 10.3389/fnpro.2010.00003

[7] Liao LD, Chen CY, Wang IJ, Chen SF, Li SY, Chen BW, et al. Gaming control using a wearable and wireless EEG-based brain-computer interface device with novel dry foam-based sensors. Journal of Neuroengineering and Rehabilitation. 2012;**9**(5)DOI: 10.1186/1743-0003-9-5

[8] Chi YM, Jung TP, Cauwenberghs G. Dry-contact and noncontact biopotential electrodes: Methological review. IEEE Reviews in Biomedical Engineering. 2010;**3**:106-119. DOI: 10.1109/RBME.2010.2084078

[9] Xu P, Yao D. A novel method based on realistic head model for EEG denoising. Computer Methods and Programs in Biomedicine. 2006;**83**(2):104-110. DOI: 10.1016/j.cmpb.2006.06.002

[10] Jung TP, Makeig S, Humphries C, Lee TW, Mckeown MJ, Iragui V, et al. Removing electroencephalographic artifacts by blind source separation. Psychophysiology. 2000;**37**(2):163-178. DOI: 10.1111/1469-8986.3720163

[11] Shoker L, Sanei S, Chambers J. Artifact removal from electroencephalograms using a hybrid BSS-SVM algorithm. IEEE Signal Processing Letters. 2005;**12**(10):721-724. DOI: 10.1109/LSP.2005.855539

[12] Fatourechi M, Bashashati A, Ward RK, Birch GE. EMG and EOG artifacts in brain computer interface systems: A survey. Clinical Neurophysiology. 2007;**118**(3):480-494. DOI: 10.1016/j.clinph.2006.10.019

[13] Daly I, Nicolaou N, Nasuto SJ, Warwick K. Automated artifact removal from the electroencephalogram: A comparative study. Clinical EEG and Neuroscience. 2013;**44**(4):291-306. DOI: 10.1177/1550059413476485

[14] Romero S, Mañanas MA, Barbanoj MJ. A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: A simulation case. Computers in Biology and Medicine. 2008;**38**(3):348-360. DOI: 10.1016/j.compbiomed.2007.12.001

[15] Urigüen JA, Garcia-Zapirain B. EEG artifact removal—state-of-the-art and guidelines. Journal of Neural Engineering. 2015;**12**(3):031001. DOI: 10.1088/1741-2560/12/3/031001

[16] Ghaderi F, Kim SK, Kirchner EA. Effects of eye artifact removal methods on single trial P300 detection, a comparative study. Journal of Neuroscience Methods. 2014;**221**:41-47. DOI: 10.1016/j.jneumeth.2013.08.025

[17] Kanoga S. An accurate removal of eyeblink artifact from single-channel electroencephalogram by supervised tensor factorization [dissertation]. Yokohama: Keio University; 2016. 159p. Available from: http://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO50002002-20164510-0003

[18] Teplan M. Fundamentals of EEG measurement. Measurement Science Review. 2002;**2**(2):1-11

[19] Huhta JC, Webster JG. 60-Hz interference in electrocardiography. IEEE Transactions on Biomedical Engineering. 1973;**20**(2):91-101. DOI: 10.1109/TBME.1973.324169

[20] Goncharova II, McFarland DJ, Vaughan TM, Wolpaw JR. EMG contamination of EEG: Spectral and topographical characteristics. Clinical Neurophysiology. 2003;**114**(9):1580-1593. DOI: 10.1016/S1388-2457(03)00093-2

[21] Olund T, Duun-Henriksen J, Kjaer TW, Sorensen HBD. Automatic detection and classification of artifacts in single-channel EEG. In: Engineering in Medicine and Biology

Society; 26-30 Aug.; Chicago, Illinois, USA. IEEE; 2014. pp. 922-925. DOI: 10.1109/EMBC.2014.6943742

[22] Hu J, Wang CS, Wu M, Du YX, He Y, She J. Removal of EOG and EMG artifacts from EEG using combination of functional link neural network and adaptive neural fuzzy inference system . Neurocomputing. 2015;**151**(1):278-287. DOI: 10.1016/j.neucom.2014.09.040

[23] Onton J, Makeig S. Information-based modeling of event-related brain dynamics. Progress in Brain Research. 2006;**159**:99-120. DOI: 10.1016/S0079-6123(06)59007-7

[24] Freeman WJ. Origin, structure, and role of background EEG activity. Part 1. Analytic amplitude. Clinical Neurophysiology. 2004;**115**(9):2077-2088. DOI: 10.1016/j.clinph.2004.02.029

[25] Sarvas J. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Physics in Medicine and Biology. 1987;**32**(1):11-22. DOI: 10.1088/0031-9155/32/1/004

[26] Picton TW, Bentin S, Berg P, Donchin E, Hillyard SA, Johnson R, et al. Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. Psychophysiology. 2000;**37**(2):127-152

[27] Safieddine D, Kachenoura A, Albera L, Birot G, Karfoul A, Pasnicu A, et al. Removal of muscle artifact from EEG data: Comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches. EURASIP Journal on Advances in Signal Processing. 2012;**2012**:127. DOI: 10.1186/1687-6180-2012-127

[28] James CJ, Hesse CW. Independent component analysis for biomedical signals. Physiological Measurement. 2005;**26**(1):R15. DOI: 10.1088/0967-3334/26/1/R02

[29] Lee SY. Blind source separation and independent component analysis: A review. Neural Information Processing Letters and Reviews. 2005;**6**(1):1-57

[30] Berg P. Scherg M. Dipole models of eye movements and blinks. Electroencephalography and Clinical Neurophysiology. 1991;**79**(1):36-44. DOI: 10.1016/0013-4694(91)90154-V

[31] Dien J. Addressing misallocation of variance in principal components analysis of event-related potentials. Brain Topography. 1998;**11**(1):43-55. DOI: 10.1023/A:1022218503558

[32] Vigário RN. Extraction of ocular artefacts from EEG using independent component analysis. Electroencephalography and Clinical Neurophysiology. 1997;**103**(3):395-404. DOI: 10.1016/S0013-4694(97)00042-8

[33] Amari SI. Natural gradient works efficiently in learning. Neural Computation. 2006;**10**(2):251-276. DOI: 10.1162/089976698300017746

[34] Lagerlund TD, Sharbrough FW, Busacker NE. Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition. Journal of Clinical Neurophysiology. 1997;**14**(1):73-82

[35] Nicolaou N, Nasuto SJ. Automatic artefact removal from event-related potentials via clustering. The Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology. 2007;**48**(1):173-183. DOI: 10.1007/s11265-006-0011-z

[36] Mognon A, Jovicich J, Bruzzone L, Buiatti M. ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology. 2011;**48**(2): 229-240. DOI: 10.1111/j.1469-8986.2010.01061.x

[37] Mahajan R, Morshed B. Unsupervised eye blink artifact denoising of EEG data with modified multiscale sample entropy, kurtosis, and wavelet-ICA. IEEE Journal of Biomedical and Health Informatics. 2015;**19**(1):158-165. DOI: 10.1109/JBHI.2014.2333010

[38] Woestenburg JC, Verbaten MN, Slangen JL. The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain. Biological Psychology. 1983;**16**(1-2):127-147. DOI: 10.1016/0301-0511(83)90059-5

[39] Klados MA, Papadelis C, Braun C, Bamidis PD. REG-ICA: A hybrid methodology combining blind source separation and regression techniques for the rejection of ocular artifacts. Biomedical Signal Processing and Control. 2011;**6**(3):291-300. DOI: 10.1016/j.bspc.2011.02.001

[40] Gratton G, Coles MG, Donchin E. A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology. 1983;**55**(4):468-484. DOI: 10.1016/0013-4694(83)90135-9

[41] Croft RJ, Barry RJ. EOG correction of blinks with saccade coefficients: A test and revision of the aligned-artefact average solution. Clinical Neurophysiology. 2000;**111**(3):444-451. DOI: 10.1016/S1388-2457(99)00296-5

[42] Sweeney KT, Ward TE, McLoone SF. Artifact removal in physiological signals—practices and possibilities. IEEE Transactions on Information Technology in Biomedicine. 2012;**16**(3): 488-500. DOI: 10.1109/TITB.2012.2188536

[43] Gao J, Sultan H, Hu J, Tung WW. Denoising nonlinear time series by adaptive filtering and wavelet shrinkage: A comparison. IEEE Signal Processing Letters. 2010;**17**(3):237-240. DOI: 10.1109/LSP.2009.2037773

[44] Feintuch PL. An adaptive recursive LMS filter. Proceedings of the IEEE. 1976;**64**(11):1622-1624. DOI: 10.1109/PROC.1976.10384

[45] Rupp M. A family of adaptive filter algorithms with decorrelating properties. IEEE Transactions on Signal Processing. 1998;**46**(3):771-775. DOI: 10.1109/78.661344

[46] Peng H, Hu B, Shi Q, Ratcliffe M, Zhao Q, Qi Y, et al. Removal of ocular artifacts in EEG—an improved approach combining DWT and ANC for portable applications. IEEE Journal of Biomedical and Health Informatics. 2013;**17**(3):600-607. DOI: 10.1109/JBHI.2013.2253614

[47] Rahman FA, Othman MF, Shaharuddin NA. A review on the current state of artifact removal methods for electroencephalogram signals. In: 2015 10th Asian Control Conference;

31 May–3 Jun.; Kota Kinabalu, Malaysia. IEEE; 2015. pp. 1-6. DOI: 10.1109/ASCC.2015.7244679

[48] He P, Kahle M, Wilson G, Russell C. Removal of ocular artifacts from EEG: A comparison of adaptive filtering method and regression method using simulated data. In: 27th Annual International Conference of the IEEE Engineering; 17-18 Jan.; Shanghai, China. IEEE; 2006. pp. 1110-1113. DOI: 10.1109/IEMBS.2005.1616614

[49] Morbidi F, Garulli A, Prattichizzo D, Rizzo C, Rossi S. Application of Kalman filter to remove TMS-induced artifacts from EEG recordings. IEEE Transactions on Control Systems Technology. 2008;**16**(6):1360-1366. DOI: 10.1109/TCST.2008.921814

[50] Navarro X, Porée F, Beuchée A, Carrault G. Denoising preterm EEG by signal decomposition and adaptive filtering: A comparative study. Medical Engineering and Physics. 2015;**37**(3):315-320. DOI: 10.1016/j.medengphy.2015.01.006

[51] Mijovic B, De Vos M, Gligorijevic I, Taelman J, Van Huffel S. Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis. IEEE transactions on biomedical engineering. 2010;**57**(9):2188-2196. DOI: 10.1109/TBME.2010.2051440

[52] Lindsen JP, Bhattacharya J. Correction of blink artifacts using independent component analysis and empirical mode decomposition. Psychophysiology. 2010;**47**(5):955-960. DOI: 10.1111/j.1469-8986.2010.00995.x

[53] Khatun S, Mahajan R, Morshed BI. Comparative analysis of wavelet based approaches for reliable removal of ocular artifacts from single channel EEG. In: IEEE International Conference on Electro/Information Technology; 21-23 May; DeKalb, IL, USA. IEEE; 2015. pp. 335-340. DOI: 10.1109/EIT.2015.7293364

[54] Davies ME, James CJ. Source separation using single channel ICA. Signal Processing. 2007;**87**(8):1819-1832. DOI: 10.1016/j.sigpro.2007.01.011

[55] Lin J, Zhang A. Fault feature separation using wavelet-ICA filter. NDT & E International. 2005;**38**(6):421-427. DOI: 10.1016/j.ndteint.2004.11.005

[56] Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. 1998;**454**(1971):903-995. DOI: 10.1098/rspa.1998.0193

[57] Flandrin P, Goncalves P. empirical mode decompositions as data-driven wavelet-like expansions. International Journal of Wavelets, Multiresolution and Information Processing. 2004;**2**(4):477-496. DOI: 10.1142/S0219691304000561

[58] Sweeney KT, Ayaz H, Ward TE, Izzetoglu M, McLoone SF, Onaral B. A methodology for validating artifact removal techniques for physiological signals. IEEE Transactions on Information Technology in Biomedicine. 2012;**16**(5):918-926. DOI: 10.1109/TITB.2012.2207400

[59] Mourad N, Niazy RK. Automatic correction of eye blink artifact in single channel EEG recording using EMD and OMP. In: The 2013 European Signal Processing Conference; 9-13 Sept.; Marrakech, Morocco. IEEE; 2013. pp. 1-5

[60] Wu SD, Chiou JC, Goldman E. Solution for mode mixing phenomenon of the empirical mode decomposition. In: IEEE 3rd International Conference on Advanced Computer Theory and Engineering; 20-22 Aug.; Chengdu, China. IEEE; 2010. pp. 500-504. DOI: 10.1109/ICACTE.2010.5579417

[61] Lee DD, Seung HS,. Learning the parts of objects by non-negative matrix factorization. Nature. 1999;**401**:788-791. DOI: 10.1038/44565

[62] Damon C, Liutkus A, Gramfort A, Essid S. Non-negative matrix factorization for single-channel EEG artifact rejection. In: IEEE International Conference on Acoustics, Speech and Signal Processing; 26-31 May; Vancouver, BC, Canada. IEEE; 2013. pp. 1177-1181. DOI: 10.1109/ICASSP.2013.6637836

[63] Kanoga S, Mitsukura Y. Eye-blink artifact reduction using 2-step nonnegative matrix factorization for single-channel electroencephalographic signals. Journal of Signal Processing. 2014;**18**(5):251-257

[64] Majmudar CA, Mahajan R, Morshed BI. Real-time hybrid ocular artifact detection and removal for single channel EEG. In: IEEE International Conference on Electro/Information Technology; 21–23 May; DeKalb, IL, USA. IEEE; 2015. pp. 330-334. DOI: 10.1109/EIT.2015.7293363

[65] Chang WD, Cha HS, Kim K, Im CH. Detection of eye blink artifacts from single prefrontal channel electroencephalogram. Computer Methods and Programs in Biomedicine. 2016; **124**:19-30. DOI: 10.1016/j.cmpb.2015.10.011

IntechOpen