Signal Processing 92 (2012) 401–416

Contents lists available at SciVerse ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Employing spatially constrained ICA and wavelet denoising, for automatic removal of artifacts from multichannel EEG data

Muhammad Tahir Akhtar [a,*,♣], Wataru Mitsuhashi [b], Christopher J. James [c]

[a] The Center for Frontier Science and Engineering (CFSE), The University for Electro-Communications, 1-5-1 Chofugaoka, Tokyo 182-8585, Japan
[b] Department of Information and Communication Engineering, The University for Electro-Communications, 1-5-1 Chofugaoka, Tokyo 182-8585, Japan
[c] Institute of Digital Healthcare, International Digital Laboratory, WMG, University of Warwick, Coventry, UK

ARTICLE INFO

Article history:
Received 16 March 2011
Received in revised form
26 June 2011
Accepted 9 August 2011
Available online 19 August 2011

Keywords:
EEG
Artifact removal
Independent component analysis (ICA)
Wavelet denoising
Spatially constrained ICA

ABSTRACT

Detecting artifacts produced in electroencephalographic (EEG) data by muscle activity, eye blinks and electrical noise, etc., is an important problem in EEG signal processing research. These artifacts must be corrected before further analysis because it renders subsequent analysis very error-prone. One solution is to reject the data segment if artifact is present during the observation interval, however, the rejected data segment could contain important information masked by the artifact. The independent component analysis (ICA) can be an effective and applicable method for EEG denoising. The goal of this paper is to propose a framework, based on ICA and wavelet denoising (WD), to improve the pre-processing of EEG signals. In particular we employ concept of the spatially constrained ICA (SCICA) to extract artifact-only independent components (ICs) from the given EEG data, use WD to remove any cerebral activity from the extracted-artifacts ICs, and finally project back the artifacts to be subtracted from EEG signals to get clean EEG data. The main advantage of the proposed approach is faster computation, as it is not necessary to identify all ICs. Computer experiments are carried out, which demonstrate effectiveness of the proposed approach in removing focal artifacts that can be well separated by SCICA.

1. Introduction

Ocular artifacts (OAs) (eye movements and eye blinks), muscle noise, heart signals, and line noise often produce large and distracting artifacts in electroencephalographic (EEG) recordings [1]. Rejecting EEG segments with artifacts offers an easy solution, however, the amount of data lost to artifact rejection may be unacceptable. The EEG signals contain neural information below 100 Hz (in many applications the information lies below 30 Hz) [1], and conventional filtering methods can be used to remove the line noise and other higher frequency components. The main problem is OAs and muscle artifacts that have a spectral overlap with the underlying EEG and cannot be removed using conventional filtering [2]. Several methods have been developed to cope with the problem of OAs. The most popular approach is the correction of OAs by means of regression analysis [3,4]. The regression based approaches require a good regression channel, e.g., electrooculographic – EOG – channel (giving recording of mainly OAs), to estimate the influence of OAs on the signals recorded by scalp electrodes and to remove it from the EEG recording. Furthermore, there is a cross-contamination between EOG and EEG channels. Other approaches documented for OAs removal include trial rejection, eye

* Corresponding author. Tel.: +81 42 443 5885; fax: +81 42 443 5287.
E-mail addresses: akhtar@ice.uec.ac.jp, akhtar@ieee.org (M.T. Akhtar), mit@ice.uec.ac.jp (W. Mitsuhashi), C.James@warwick.ac.uk (C.J. James).
♣ EURASIP Member.

0165-1684/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2011.08.005

fixation, principle component analysis (PCA) [5], blind source separation (BSS) [2,6,7], robust beamforming [8], hybrid BSS-SVM approach [9], and methods based on H^{∞} adaptive [10] and spatial filters [11].

It has been shown by many researchers that independent component analysis (ICA) [12–16] can be used to efficiently separate the distinct artifactual processes from EEG data. Various researchers have conducted a comparative study between regression and components based methods [17–20]. Though results in [17] supported the use of regression based and PCA-based OAs correction, recently [18] demonstrated the effectiveness of ICA-based methods for eye movement removal even when EOG recordings were not available or when data length was short. [19] also showed that ICA yields better performance as compared with the regression method, supporting the results of [20] which reported that ICA preprocessing led to a 10–20% increase in detection performance for all ICA algorithms tested therein. In the most of existing methods, after ICA, independent components (ICs) corresponding to artifacts are manually selected using visual inspection and requires experts view. This manual selection and classification of artifact ICs is time-consuming and not well suited for real-time applications, and is subject to human biases. To overcome these problems, many procedures for automatic identification of artifact ICs have been developed [21–26]. In [21], Hurst exponent is suggested for this purpose with the assumption that that this is constant for all artifacts of the same type. One idea, for example, is to utilize information on localization, frequencies or other various properties of source signals [23,24]. Yet, another possibility is to use constrained ICA [22] to extract a single IC that is constrained to be similar to some reference signal. The method was used for eye blink removal, with the reference signal being obtained by recording a pulse when EEG channel Fp1 exceeds a predefined threshold. In almost all of these approaches, the ICs identified as artifacts; either manually or automatically; are rejected and the remaining ICs are used to reconstruct a clean EEG data. If some cerebral activity is leaked to artifact ICs, rejecting these components results in loss of some desired information.

Recently discrete wavelet transform (DWT) has been used for enhancing artifact suppression in EEG signals [27–30]. Probably [27] is a first attempt to incorporate DWT with ICA for artifact removal from EEG. The wavelet enhanced ICA (wICA) of [27] combines ICA with wavelet denoising (WD), and makes use of wavelet thresholding for denoising of the demixed ICs. The thresholding allows conservation of the time-frequency structure of artifacts and recovering of the cerebral activity "leaked" into the components [27]. This method does not require manual identification of artifact ICs, as all ICs are wavelet denoised. However, extracting and processing all ICs poses a problem mainly due to long processing time, and WD of all ICs might result in possible removal of some desired information as well. The method proposed in [28] is a one step modification of wICA: after applying DWT to all ICs, various measures (see [28] for details) are computed for ICs and wavelet ICs and a flag for each measure is associated with ICs, and finally ICs with at least four flags are identified as artifact and rejected before reconstructing the clean EEG data. This method requires reference EOG signals; for computing flags on the basis of comparison between wavelet components of ICs and reference signals; which might not be available in all situations. As compared with wICA, Mammone et al. [29] take a different approach: DWT partitions each channel of the original EEG into the four major bands of brain activity, each rhythm of each channel is represented by a wavelet component, the wavelet components linked to artifactual events are automatically identified by means of a quantitative measure and passed through ICA in order to concentrate the artifactual content in a few ICs, and finally artifactual components are rejected before employing inverse ICA and inverse DWT to reconstruct clean EEG data. Authors of [30] also employ DWT directly to the EEG data, where they use a (variable) multilevel decomposition technique for wavelet analysis, to ensure that artifacts are completely removed from the EEG data. It is worth mentioning that in [29,30], applying DWT directly to the EEG data requires great care and extensive experiments in choosing appropriate wavelet basis and denoising threshold levels, and above-all increases the computational complexity.

In this paper our objective is to realize an efficient approach for artifact removal from EEG data meeting the following properties: (1) automatic identification of artifactual components, (2) simplify parameter selection for WD, and (3) reduced loss of desired information from originally artifact-free segments. The proposed approach is based on spatially constrained ICA (SCICA) [31,32] which incorporates reference or constraint topographies in the ICA algorithms. This allows to extract only desired ICs – that is artifacts in our case – and hence we do not need to run a complete ICA to extract all sources. We then apply WD to artifact ICs to remove cerebral activity leaked to these ICs and get artifact-only signals. In contrast to the existing approaches [27–30] which use standard DWT, we employ stationary wavelet transform (SWT) for its desirable properties, as explained later. Finally these artifact-only signals are projected back, and subtracted from EEG data to get the clean EEG signals. This proposed approach is faster in computation, as not all ICs are extracted and processed. The rest of the paper is organized as follows: Section 2 introduces the basic ICA model for the multichannel EEG data, and gives an overview of existing approaches (investigated in this paper) for artifact removal. Section 3 presents proposed approach for EEG denoising. The computer experiments are detailed in Section 4 followed by the concluding remarks presented in Section 5.1.

2. Basic ICA model and wavelet-enhanced ICA

As shown in Fig. 1(a), the EEG data is assumed to be generated according to the model given as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{v}(t), \tag{1}$$

where $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_M(t)]^T$ are a linear mixture of N sources $\mathbf{s}(t) = [s_1(t), s_2(t), \ldots, s_N(t)]^T$, $\mathbf{A}$ is $M \times N$ mixing matrix, and $\mathbf{v}(t) = [\nu_1(t), \nu_2(t), \ldots, \nu_M(t)]^T$ is additive noise