Noise Separation of BCI Data Using Machine Learning Technique

Rachana.M[1], Sayana.M[2], Dr.Venugopal.G[3]

PG Student [ICS], Dept. of EEE, Jawaharlal College of Engineering and Technology, Palakkad, Kerala, India[1]

Assistant Professor, Dept. of EEE, Jawaharlal College of Engineering and Technology, Palakkad, Kerala, India[2]

Assistant Professor, Dept. of ICS, NSS College of Engineering, Palakkad, Kerala, India[3]

ABSTRACT: Many different disorders can disrupt the neuromuscular channels through which the brain communicates with and controls its external environment. Amyotrophic lateral sclerosis (ALS), brainstem stroke, brain or spinal cord injury, cerebral palsy, muscular dystrophies, multiple sclerosis, and numerous other diseases impair the neural pathways that control muscles or impair the muscles themselves. Electroencephalography (EEG) is a measure of the brain's voltage fluctuations as detected from scalp electrodes, it reflect brain activity. These activities can be decoded by signal processing techniques. However, EEG recordings are always contaminated with artifacts which hamper the algorithmic process. Therefore, identifying and removing artifacts from EEG signal is an important prerequisite for Brain Computer Interface(BCI). A statistical model called Independent Component Analysis[ICA] can decompose EEG recording into artifact –related and event related potential (ERP)-related IC. Existing ICA -based artefact removal strategies concentrate only on the removal of a subset of artifacts, e.g. identifying eye movement artifacts only. In this paper we propose an automatic algorithm for the classification of general artifacts. The proposed algorithm consists of two parts: 1) Identifying artifacts which have physiological origins by using an event-related feature based clustering algorithm 2) Identifying non-biological artifacts by using the electrode-scalp impedance information. Here we compare the performance of our proposed method with other commonly used automatic artifact removal methods by the classification accuracies. The results illustrate the effectiveness of our proposed algorithm in identifying and removing artifacts.

KEYWORDS: Electroencephalogram; Independent component analysis; Event Related Potential.

I.INTRODUCTION

For many years people have speculated that electroencephalographic activity or other electrophysiological measures of brain function might provide a new non-muscular channel for sending messages and commands to the external world – a brain–computer interface (BCI). Over the past 15 years, productive BCI research programs have arisen. Encouraged by new understanding of brain function, by the advent of powerful low-cost computer equipment, and by growing recognition of the needs and potentials of people with disabilities, these programs concentrate on developing new augmentative communication and control technology for those with severe neuromuscular disorders, such as amyotrophic lateral sclerosis, brainstem stroke, and spinal cord injury. The immediate goal is to provide these users, who may be completely paralysed, or 'locked in', with basic communication capabilities so that they can express their wishes to caregivers or even operate word processing programs or neuroprostheses [3].

Present-day BCIs determine the intent of the user from a variety of different electrophysiological signals. These signals include slow cortical potentials, P300 potentials, and mu or beta rhythms recorded from the scalp, and cortical neuronal activity recorded by implanted electrodes. They are translated in real-time into commands that operate a computer display or other device. Successful operation requires that the user encode commands in these signals and that the BCI derive the commands from the signals. Thus, the user and the BCI system need to adapt to each other both initially and continually so as to ensure stable performance. Current BCIs have maximum information transfer rates up to 10–25 bits/min.[4]
BECAUSE of its high temporal resolution, electroencephalographic (EEG) recordings have been widely used to measure background activities of the brain as well as the specific activity for a cognitive task in brain computer

interface (BCI) [8]. However a major problem of EEG recordings is that they are highly susceptible to various artifacts. In other words, it is almost impossible to see any event-related potential (ERP), the typical electrophysiological response to an internal or external stimulus, in the raw EEG recordings due to the presence of artifacts. However, neuroscientists are often interested in visualizing the signals and their time domain ERPs such as N200 (a negative peak around 200ms after the excitation due to the stimuli) or P300 (a positive peak around 300ms after the excitation due to the stimuli) .Therefore, artifact identification and rejection is a crucial step in the ERP-related EEG-based BCI.

The artifacts can be divided into two categories: physiological and non-biological artifacts, based on their origins. Physiological artifacts arise from biological sources other than the brain such as eye blinking, eye movements or muscle movements, etc. Non-biological artifacts originate from outside the body due to factors such as high-impedance electrodes Over past decades, several approaches have been proposed to identify and remove these artifacts, especially for the physiological artifacts[1]. The most trivial of these approaches involves simply deleting the portions of the data in which the EEG activity exceeds some predefined thresholds. However, this may lead to a large loss of data, which in turn could mean the loss of relevant recorded ERP signals our proposed solution for identification and removal of general artifacts would be valuable for EEG researchers and BCI users. Firstly, the methods proposed represent a unified solution for all types of artifacts and not just ones caused by physiological phenomena. Secondly, a practical method must be applicable without the need of time-consuming preparations at the time of the experiment.

II.DEFINITION AND FEATURES OF A BCI

A BCI is a communication system in which messages or commands that an individual sends to the external world do not pass through the brain's normal output pathways of peripheral nerves and muscles. For example, in an EEG based BCI the messages are encoded in EEG activity. A BCI provides its user with an alternative method for acting on the world. BCIs fall into two classes: dependent and independent.

A dependent BCI does not use the brain's normal output pathways to carry the message, but activity in these pathways is needed to generate the brain activity (e.g. EEG) that does carry it. For example, one dependent BCI presents the user with a matrix of letters that flash one at a time, and the user selects a specific letter by looking directly at it so that the visual evoked potential (VEP) recorded from the scalp over visual cortex when that letter flashes is much larger that the VEPs produced when other letters flash (Sutter, 1992). In this case, the brain's output channel is EEG, but the generation of the EEG signal depends on gaze direction, and therefore on extra ocular muscles and the cranial nerves that activate them. A dependent BCI is essentially an alternative method for detecting messages carried in the brain's normal output pathways: in the present example, gaze direction is detected by monitoring EEG rather than by monitoring eye position directly. While a dependent BCI does not give the brain a new communication channel that is independent of conventional channels, it can still be useful (e.g. Sutter and Tran, 1992).

 In contrast, an independent BCI does not depend in any way on the brain's normal output pathways. The message is not carried by peripheral nerves and muscles, and, furthermore, activity in these pathways is not needed to generate the brain activity (e.g. EEG) that does carry the message. For example, one independent BCI presents the user with a matrix of letters that flash one at a time, and the user selects a specific letter by producing a P300 evoked potential when that letter flashes. In this case, the brain's output channel is EEG, and the generation of the EEG signal depends mainly on the user's intent, not on the precise orientation of the eyes.

A) Parts of a BCI
Like any communication or control system, a BCI has input (e.g. electrophysiological activity from the user), output (i.e. device commands), components that translate input into output, and a protocol that determines the onset, offset, and timing of operation. Fig. 1 shows these elements and their principal interactions.

Signals from the brain are acquired by electrodes on the scalp or in the head and processed to extract specific signal features (e.g. amplitudes of evoked potentials or sensorimotor cortex rhythms, firing rates of cortical neurons) that reflect theuser's intent.
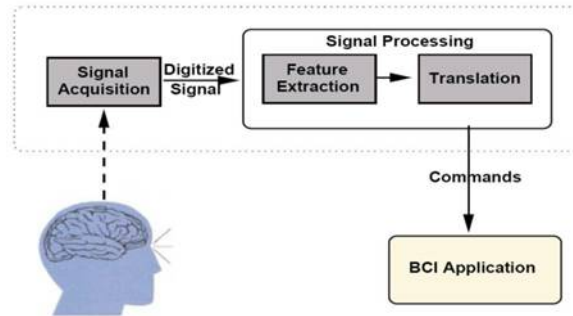
Fig. 1.Basic design and operation of BCI system.

These features are translated into commands that operate a device (e.g. a simple word processing program, a wheelchair, or a neuroprostheses). Success depends on the interaction of two adaptive controllers, user and system. The user must develop and maintain good correlation between his or her intent and the signal features employed by the BCI; and the BCI must select and extract features that the user can control and must translate those features into device commands correctly and efficiently.

III.EEG

EEG has been used mainly to evaluate neurological disorders in the clinic and to investigate brain function in the laboratory. People have speculated that the EEG could have a fourth application, that it could be used to decipher thoughts, or intent, so that a person could communicate with others or control devices directly by means of brain activity, without using the normal channels of peripheral nerves and muscles. This idea has appeared often in popular fiction and fantasy. However, EEG based communication attracted little serious scientific attention until recently, for at least 3 reasons.

First, while the EEG reflects brain activity, so that a person's intent could in theory be detected in it, the resolution and reliability of the information detectable in the spontaneous
EEG is limited by the vast number of electrically active neuronal elements, the complex electrical and spatial geometry of the brain and head, and the disconcerting trial to trial variability of brain function. The possibility of recognizing a single message or command amidst this complexity, distortion, and variability appeared to be extremely remote. Second, EEG-based communication requires the capacity to analyse the EEG in real-time, and until recently the requisite technology either did not exist or was extremely expensive. Third, there was in the past little interest in the limited communication capacity that a first generation EEG-based BCI was likely to offer.

Recent scientific, technological, and societal events have changed this situation. First, basic and clinical research has yielded detailed knowledge of the signals that comprise the EEG. For the major EEG rhythms and for a variety of evoked potentials, their sites and mechanisms of origin and their relationships with specific aspects of brain function, are no longer wholly obscure. Numerous studies have demonstratedcorrelations between EEG signals and actual or imagined movements and between EEG signals and mental tasks (e.g. Keirn and Aunon, 1990; McFarland et al., 2000a). Thus, researchers are in a much better position to consider which EEG signals might be used for communication and control, and how they might best be used. Second, the extremely rapid and continuing development of inexpensive computer hardware and software supports sophisticated online analyses of multichannel EEG. This digital revolution has also led to appreciation of the fact that simple 768 J.R. Wolpaw et al. / Clinical Neurophysiology 113 (2002) 767–791 communication capacities (e.g. 'Yes' or 'No', 'On' or 'Off') can be configured to serve complex functions (e.g. word processing, prosthesis control). Third, greatly increased societal recognition of the needs and potential of people with severe neuromuscular disorders like spinal cord injury or cerebral palsy has generated clinical, scientific, and commercial interest in better augmentative communication and control technology. Development of such technology is both the impetus and the justification for current BCI research. BCI technology might serve people who cannot use conventional augmentative technologies; and these people could find even the limited capacities of first-generation BCI systems valuable.

IV. METHOD

A) EEG Data Preprocessing

Raw EEG recordings were band-pass filtered from 0.5 Hz to 50 Hz.

B) Independent Component Analysis (ICA)

Independent component analysis (ICA) is a computational method for separating a multivariate signal into additive subcomponents. This is done by assuming that the subcomponents are non-Gaussian signals and that they are statistically independent from each other. ICA aims at finding linear projections of the data that maximize their mutual independence. It is assumed that we observe an array of electrodes that provide a vector of N electrode recordings $v=[v1, v2, \ldots, vN]T$ that are linear combinations of M unknown and statistically independent sources $s=[s1, s2, \ldots, sN]T$. The objective of the ICA algorithm is to find a separating matrix W, such that

$$\boldsymbol{s} = W \times \boldsymbol{v} \qquad (1)$$

C) Artifact-related ICs Identification

After the ICA decomposition, we chose to leverage some well-known features in order to best capture the behaviour of the ICs associated with the two different artifact classes. Here, we describe the features used for each artifact class.

V. PHYSIOLOGICAL ARTIFACT-RELATED ICS

A) Feature Extraction

Eye blinks, eye movements and muscle movements are the major sources of physiological artifacts. There are four kinds of features are extracted here to distinguish these artifacts from real brain signals:

a) Temporal features:

The amplitude of the artifact-related ICs will abruptly jump due to the presence of physiological artifacts like eye blinks, and show different temporal patterns compared to the normal brain-related ICs. This jump can be well captured by the kurtosis.

b) Spatial features:

The artifact-related ICs and the brain-related ICs are projected on different groups of electrodes. For instance, in our P300 experiment, the brain-related ICs are concentrated on the frontal and central electrodes (around Fz channel), while the eye blinks project most strongly on the far frontal site on the scalp [2]. To capture the spatial topography of artifact-related ICs, the median of each IC's topography weight is utilized here.

c) Spectral features:

The normal power of EEG signals are in delta band (0-4 Hz), theta band (4-8 Hz), alpha band (8-13 Hz), and beta band (13-30 Hz) and most of it falls in the range of 1–20 Hz [29]. However, the artifacts show dissimilar power distribution. For example, the spectrum of muscle artifacts is characterized usually by a high value in the 20-50 Hz range [30]. These differences can be highlighted by the average band power of delta, theta, alpha, beta and gamma bands (gamma is 30-50Hz).

d) Similarity over epochs:

The artifacts are random, unexpected, and usually only occur in some epochs. Thus, the epochs that contain artifacts have no common pattern and exhibit very low similarity with other epochs. On the other hand, the epochs with ERP-related ICs exhibit higher similarity with others. The correlation value is adopted to measure the similarity.

B) Autonomous data density clustering

A recently introduced data density based approach to clustering known as Data Density based Clustering has been presented which automatically determines the number of clusters. By using the Recursive Density Estimation for each

point the number of calculations is significantly reduced in offline mode and, further, the method is suitable for online use. The Data Density based Clustering method however requires an initial cluster radius to be entered.A different radius per feature/ dimension creates hyper-ellipsoid clusters which are axis-orthogonal. This results in a greater differentiation between clusters where the clusters are highly asymmetrical. In this paper we used the method which automatically derives suitable initial radii. The selection is data driven and requires no user input.[1]

C) Physiological Artifacts Identification

Hence, we build a template to guide the artefact identification for all the subjects based on this a priori knowledge:

Step 1: Calculate the back-projection value pi of each IC as follows:

$$pi = W{-}1(i) \times si \qquad (2)$$

pi is the back-projection value of IC i, W^-1 is the inverse of the unmixing matrix.

Step 2: Calculate contribution of each cluster to the desired signal of interest patterns: N200 ($\emptyset jn2$) and P300 ($\emptyset jP3$), separately, since they have different cognitive representations and impact factors related to the latency

Step 3: Finally, the cluster j which minimizes the difference between $\emptyset jP3$ and $\emptyset jn2$ is identified as being affected by artifacts and the ICs inside of this cluster are marked as artifact-related ICs.

VI. NON-BIOLOGICAL ARTIFACT-RELATED ICS

Poor scalp contact for a particular electrode that will produce consistently bad data for a long duration is the major source for non-biological artifacts. In order to identify this class of artifacts, the electrode-scalp impedance information is employed to guide the non-biological artifact related IC identification process.

A) Electrode-scalp Impedance

If there is any disruption between an electrode and the scalp, the reported results may not be accurate. Typically, the contact quality of an electrode to the scalp is evaluated by the impedance value between the electrode and the scalp. One method to measure the electrode scalp impedance would be to inject a current Ia at the signal electrode, This technique is called "Lead-off detection" When the applied current is a sinusoid at a known frequency fo then we have,

$$Vout,0 = Ia,f0 \times Zoverall \qquad (3)$$

The frequency response of $Vout$ at fo is dominated by the voltage drop across the overall impedance of the circuit due to the injected current,

$$Zoverall = Zelec, + ZL + Zelec,(4)$$

The overall impedance $Zoverall$ is the combination of $Zelec,$ (the impedance faced by the dry signal electrodes), ZL (the impedance of the length of the scalp between two electrodes), and $Zelec,$ (the impedance faced by the wet reference electrode). The power spectrum of the signal $Vout$ at fo is directly proportional to the impedance faced by the constant current.

B) Non-biological Artifacts Identification

In our study, we inject a sinusoidal signal with fo=30.5Hz as the constant current Ia and compute the magnitude of the power spectra of the output signals at 30.5Hz as a measure of the impedance between the electrode and the scalp. A higher magnitude at 30.5 Hz, measured at one output channel, implies higher impedance electrode and therefore poorer contact between that electrode and the scalp.

Some electrodes shows much higher impedance compared to the other electrodes, a Euclidean distance-based clustering procedure is employed on the magnitude values at 30.5 Hz.The non-biological artifacts can be identified in the following steps:

Step 1: Calculate the electrode-scalp impedance for each individual electrode and employ the clustering technique to find the electrodes which show extremely high impedance compared to others.

Step 2: Compare the contributions of all the electrodes for each IC. The ICs that represent the non-biological artifacts.

C) Artifacts Removal and Clean EEG Reconstruction

In the last step, the components labeled as artifact-related ICs are removed from the data. Then the artifact-free EEG data was reconstructed from the remaining ICs.

VII. RESULT AND ANALYSIS

Since the physiological and non-biological artifacts can randomly occur and are unexpected, it is difficult to generate a global model to identify them. Thus, instead of constructing a global template for artifact-related ICs, our approach is based on a global pattern that encapsulates models for signals of interest and identifies the artifact-related ICs by searching for the minimal contribution to this model.

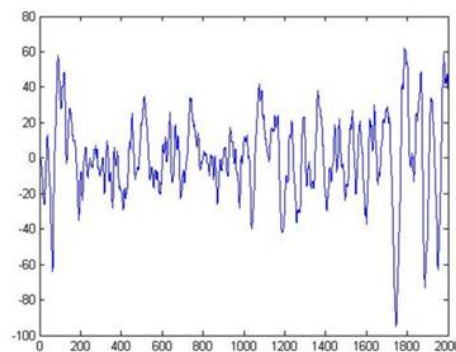In the fig 2, it shows the raw EEGwith different noises
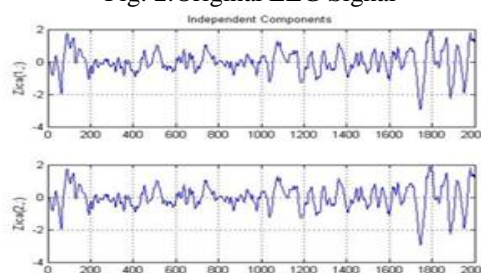

Fig. 2.Original EEG Signal

Fig. 3Independant Components

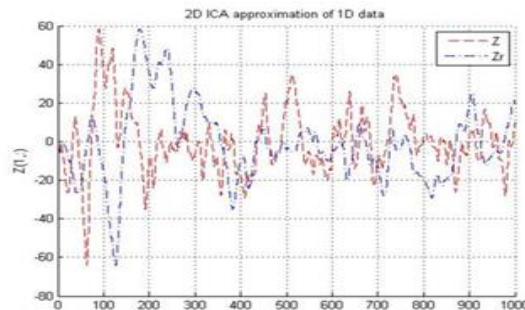In the fig 3, it shows the Independent Components of ERP related IC and Artifact related IC

Fig .4EEG Signal without Artifacts

In Fig 4, it shows the EEG signal with and without artifacts,the red dotted signal shows the signal without artifacts and the other signal shows the original signal input with noises
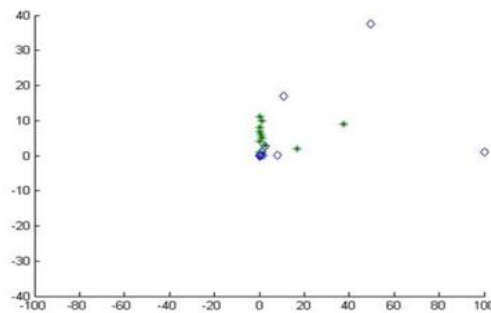

Fig.5.Autonomous Data Density Cluster Result

In Fig 5.It shows the result of autonomous data density clusterIn this a different radius per feature/ dimension creates hyper-ellipsoid clusters which are axis-orthogonal.

	KMeans CLUSTER	LDA CLASSIFIER	HIERARCHICAL CLUSTER	DATA DENSITY CLUSTER
1	84.8	81.7	86.3	92.1
2	70.9	82.9	74.5	83.5
3	76.2	75.9	83.9	86.4
4	69.5	70.9	72.1	89.4
Avg	75.35	77.85	79.2	87.85

Fig.6.Classification Accuracy in%

In Fig 6 it shows the Classification accuracy (in %) achieved by three commonly used automatic artifact identification and removal methods vs. our proposed method

VIII.CONCLUSION

A novel automated artifact-related ICs identification algorithm has been presented in this paper . The proposed methods take into account both physiological artifacts and non-biological artifacts. An ERP-related clustering method is proposed for physiological artifact-related ICs identification. A qualitative comparison of original EEG signals with reconstructed artifact-free signals shows that the proposed algorithm can effectively enhance the ERP quality for all subjects in the study. Electrode-scalp impedance information was employed for non-biological artifact-related ICs identification. Our proposed method compared to three commonly used automatic removal methods for noisy ICs.

REFERENCES

[1]   Richard Hyde, PlamenAngelov Data Science Group, School of Computing and Communications Lancaster University Lancaster, LA1 4WA, UK A Fully Autonomous Data Density Based Clustering Technique 2016

[2]   S. Silvoni, L. Konicar, M.A. Prats-Sedano , E. Garcia-Cossio, C. Genna , C. Volpato, M. Cavinato,Paggiaro , S. Veser, D. De Massari N. Birbaumer Tactile event-related potentials in amyotrophic lateral sclerosis (ALS):Implications for brain-computer interface 2015

[3]   Yuan Zou, Student Member, IEEE, Viswam Nathan, Student Member, IEEE, RoozbehJafari, Senior Member, IEEE Identification of Artifact-related Independent Components for Artifact Removal in EEG Recordings   2014

[4]   Mauro Marchetti , KonstantinosPriftis Brain–computer interfaces in amyotrophic lateral sclerosis:Ametanalysis 2014

[5]   Shiro Ikegami, Kouji Takano , Kiyohiko Kondo , Naokatsu Saeki, Kenji Kansaku A region-based two-step P300-based brain–computer interface for patients with amyotrophic lateral sclerosis 2014

[6]   YihengTuYeung Sam Hung , Li Hu, Gan Huang, Yong Hu, Zhiguo Zhang An automated and fast approach to detect single-trial visual evoked potentials with application to brain–computer interface 2014

[7]   Van Laghenhove M. Chance Spalding Assistive Device With Conventional, Alternative, and Brain-Computer Interface Inputs to Enhance Interaction With the Environment for People With Amyotrophic Lateral Sclerosis: A Feasibility and Usability Study 2014

[8]   YaldaShahriari,AbbasErfanian Improving the performance of P300-based brain–computer interface through subspace-based filtering 2013

[9]   L. Mayaud, M. Congedo, A. Van Laghenhove, D. Orlikowski, M. Figèrea, E. Azabou F. Cheliout-Heraut A comparison of recording modalities of P300 event-related potentials (ERP) for brain-computer interface (BCI) paradigm july2013

[10]   Meel Velliste1, SagiPerel, M. Chance Spalding, Andrew S. Whitford&   Andrew B. Schwartz Cortical control of a prosthetic arm for self-feeding 2008

[11]   Benjamin Blankertz, Guido Dornhege,MatthiasKrauledat, Klaus-Robert Müller,a,b   and Gabriel Curioc The non-invasive Berlin Brain–Computer Interface: Fast acquisition of effective performance in untrained subjects 2007

[12]   John P. Donoghue1 Bridging the Brain to the World: A Perspective on Neural Interface Systems 2007