ICA-based Reduction of Electromyogenic Artifacts in EEG Data: Comparison With and Without EMG Data

Florian Gabsteiger[1], *Student Member, IEEE, EMBS*, Heike Leutheuser[1], *Student Member, IEEE, EMBS*, Pedro Reis[2], Matthias Lochmann[2], and Bjoern M. Eskofier[1], *Member, IEEE, EMBS*

***Abstract*— Analysis of electroencephalography (EEG) recorded during movement is often aggravated or even completely hindered by electromyogenic artifacts. This is caused by the overlapping frequencies of brain and myogenic activity and the higher amplitude of the myogenic signals.**

One commonly employed computational technique to reduce these types of artifacts is Independent Component Analysis (ICA). ICA estimates statistically independent components (ICs) that, when linearly combined, closely match the input (sensor) data. Removing the ICs that represent artifact sources and re-mixing the sources returns the input data with reduced noise activity. ICs of real-world data are usually not perfectly separated, actual sources, but a mixture of these sources.

Adding additional input signals, predominantly generated by a single IC that is already part of the original sensor data, should increase that IC's separability. We conducted this study to evaluate this concept for ICA-based electromyogenic artifact reduction in EEG using EMG signals as additional inputs.

To acquire the appropriate data we worked with nine human volunteers. The EEG and EMG were recorded while the study volunteers performed seven exercises designed to produce a wide range of representative myogenic artifacts.

To evaluate the effect of the EMG signals we estimated the sources of each dataset once with and once without the EMG data. The ICs were automatically classified as either 'myogenic' or 'non-myogenic'. We removed the former before back projection. Afterwards we calculated an objective measure to quantify the artifact reduction and assess the effect of including EMG signals.

Our study showed that the ICA-based reduction of electromyogenic artifacts can be improved by including the EMG data of artifact-inducing muscles. This approach could prove beneficial for locomotor disorder research, brain-computer interfaces, neurofeedback, and most other areas where brain activity during movement has to be analyzed.

I. INTRODUCTION

Analysis of brain activity during exercise can be used for neurofeedback systems to assist athletes in their training [1], [2]. The system could provide instant and customized feedback and thereby increase training efficiency. Clinical and research applications include improved insight into the relation between locomotion and neural activity, e.g. to control a robotic prosthesis or help understand neurological diseases, as well as analysis of cognitive processes with respect to motivated natural behavior [3], [4], [5].

[1]Digital Sports Group, Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Erlangen, Germany

[2]Institute of Sport Science and Sport, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Erlangen, Germany

Corresponding author: F. Gabsteiger:
florian.gabsteiger@fau.de

Electroencephalography (EEG) measures the electrical potentials inside the brain using scalp electrodes. The signals at each electrode are the result of a linear combination of the electric activity of the brain's neurons. EEG is the most often used, non-invasive, technique to measure brain activity during movement. Other popular functional brain imaging methods, like single-photon emission computed tomography, depend upon the patient to abide in a resting position and expose the patient to ionizing radiation. EEG equipment on the other hand is light enough to be used during movement. EEG also satisfies the temporal resolution requirements to accurately measure neural activity [4].

EEG measurements are susceptible to many different types of artifacts and are easily masked because of their low amplitude (10 µV to 100 µV) [1]. Eye movements, eye blinks, power-line interference, cardioballistic artifacts, and artifacts caused by cable movement occur in most EEG recordings (often even more so during movement) [1], [4], [6].

The EEG electrodes unintentionally measure the electrical signals produced by some muscles simultaneously to the brain activity due to their spatial proximity and the electrode's necessarily low sensitivity [7]. If the measurement is done while the patient is exercising these interferences obscure most of the neural activity as their amplitude is significantly higher (100 µV to 1000 µV) [4]. The frequency band of electromyogenic, or muscle, artifacts (0 Hz to 200 Hz) overlap the band of neural activity (0 Hz to 30 Hz) [7].

Independent Component Analysis (ICA) algorithms estimate the independent components (ICs) from a linear mixture of these sources [8]. This concept can be applied to artifact reduction in EEG data. Removing the ICs representing the myogenic activity and back projecting the residual ICs returns the EEG data with reduced muscle artifacts [9].

Including the electromyograpic (EMG) signal of the artifact-inducing muscles into the ICA decomposition could improve the model accuracy, as the ICA algorithm could potentially produce a clearer distinction between neural and electromyogenic activity. The ICs could have a more distinctive assignment to either neural or myogenic activity and less components would be a mixture of the two. Clustering the muscle activity into fewer components would simplify the artifact reduction process, remove more myogenic artifacts, and keep more neural sources intact.

In order to objectively determine if including the EMGs improves the artifact reduction we conducted a study with nine volunteers. We chose seven different exercises designed to induce a wide range of myogenic artifacts. These exercises

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

should be representative for most muscle artifacts occurring in EEG recordings. The results of this study should therefore be beneficial for other applications as well.

II. METHODS

A. Data acquisition

We used a QuickAmp-72 amplifier (Brain Products GmbH, Gilching, Germany) to record the EEG and the EMG simultaneously. To achieve consistent electrode placement we worked with an actiCAP 64 Channel (Brain Products GmbH, Gilching, Germany) EEG cap. To acquire the real positions of the electrodes the ELPOS (zebris Medical GmbH, Isny im Allgäu, Germany) and its software Electrode Guide ElGuide (zebris Medical GmbH, Isny im Allgäu, Germany) were used. We placed the EMG electrodes on the left and right sternocleidomastoid muscle and to the left and right sagittal plane of the trapezius muscle (Fig. 1).

Nine (four male, five female) healthy volunteers (age 25 $\pm$ 2 years, mean $\pm$ standard deviation(SD)) participated in the study. All participants were in good physical condition and gave written informed consent. The Ethics Committee of the University Erlangen-Nuremberg reviewed and approved the study design beforehand.

During the four isometric contraction exercises the participants were asked to force their head against an immovable object towards different directions: forward, backward, rotation to the left, and rotation to the right. Each recording consisted of eight repetitions with a duration of 15 s, interrupted by 30 s pauses. The weight lifting was done at a chest press machine with slowly increasing weight until a sub-maximum lift was achieved. The participants ran for two minutes on a h/p/cosmos treadmill (h/p/cosmos sports & medical GmbH, Nussdort-Traunstein, Germany). The treadmill was set to 2.316 m s^{-1}, which is 20 % above the average speed people transition from walking to running stance, to ensure that the person was running instead of walking [10]. As the running was performed indoors the absence of air resistance was compensated for by setting the inclination of the treadmill to 1 % [11]. Comparing the ergometer and running exercise required the step and cycling frequencies to be proportionate. We used an instrumented treadmill with a build in FDM-T zebris force plate (zebris Medical GmbH, Isny im Allgäu, Germany) to calculate the mean step frequency and used this frequency for the ergometer exercise. The resistance of the ergometer bike, a sanabike 250F (MESA Medizintechnik GmbH, Benediktbeuern, Germany), was set to 50 W.

B. Preprocessing

We used the BrainVision Analyzer 2 (Brain Products GmbH, Gilching, Germany) to preprocess the data. We removed high (>70 Hz) and low (<0.5 Hz) frequency noise with a band-pass filter. Power-line interference at 50 Hz was reduced by a notch filter. Ocular artifacts were treated using a regression model in the time domain [13]. Cardioballistic artifacts were removed using a template averaging method [14]. We removed the pause phases in the isometric contraction and weight lifting exercises. Afterwards the data was exported to EEGLab [15], a widely used Matlab (MathWorks Inc., Natick, USA) toolbox for EEG data, for further analysis.

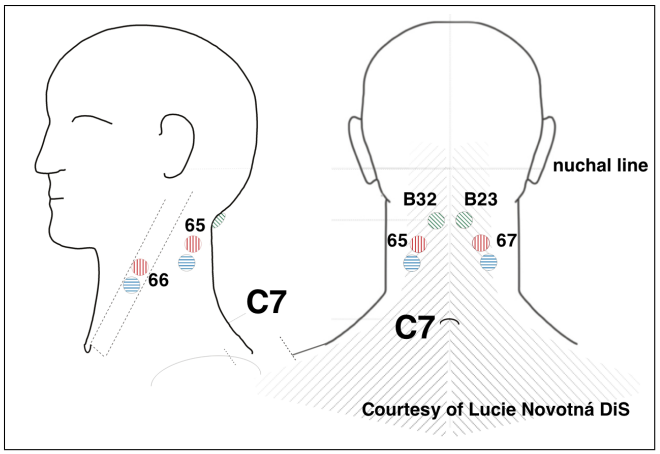

Fig. 1. EMG electrode positions. B32 and B23 are EEG electrodes. Electrodes 65 and 67 record trapezius muscle activity. Electrodes 66 and 68 (not visible) record sternocleidomastoid muscle activity. C7 is the approximate location of the seventh cervical vertebra. First shown in [12].

C. ICA decomposition

Applying ICA on EEG data containing electromyogenic artifacts returns, in the ideal case, ICs exclusively containing the brain activity and ICs representing the electrical activity of the muscles [9]. Setting the latter to zero and back projection yields the EEG without the myogenic artifacts [9].

Most ICA algorithms run Principle Component Analysis (PCA) beforehand to decorrelate, i.e. sphere, the input data. The usual choices for ICA-based artifact reduction in EEG data are the Information Maximization (InfoMax) algorithm [8] and the FastICA algorithm [16]. A more recently proposed method is the Adaptive Mixture of Independent Component Analyzers (AMICA) algorithm [17]. AMICA employs a maximum likelihood estimate for mixture models of ICs and an asymptotic Newton method for the optimization.

We chose the AMICA algorithm for this study, as it was shown to be superior to the InfoMax algorithm for removing electromyogenic artifacts in EEG data [12]. We used the AMICA implementation by Palmer et al. [18]. We ran AMICA on all datasets using default parameters: one ICA model, three mixture models for the ICs, and a maximum of 2000 iterations.

D. ICA component rejection

Selecting which and how many ICs to reject, i.e. setting which columns of the mixing matrix to zero, greatly impacts the result of the artifact reduction process. In order to achieve an objective comparison between the two ICA decompositions (once with and once without the appended EMG data) we fixed the number of ICs to reject to five. We furthermore adjusted the PCA algorithm to retain 64 principle components. We therefore removed 5 out of 64 ICs in both cases to allow a fair comparison. We worked with an automatic classifier that was specifically designed

to discriminate between electromyogenic and neural ICA components to reduce user-dependent factors [19].

E. Evaluation methodology

For the assessment of the effect that the EMGs have on the artifact reduction an objective measure was necessary. In [12] we introduced a novel measure to calculate the artifact reduction by measuring features on resting state data and to compare them to features from the data before and after the artifact reduction. The steps for calculating the improvement are shown in Fig. 2. The signal characteristics were computed on a empirically defined window size of 2000 samples.

We improved the previously shown measure by including new signal characteristics and performing a best subset feature selection. The candidate feature set was compiled from generic features for biosignal classification as well as expert features, i.e. features adapted to the problem at hand. The generic features were statistical moments and signal characteristics [20]. As expert features we calculated the normalized power above 35 Hz, the peak (position) in the power spectral density (PSD) (measured from 0 Hz to 50 Hz), the mean value of the squared derivative, and the maximum value of the squared derivative. These characteristics should be able to quantify the degree of EMG contamination in EEG recordings [7]. Additionally we added features based on the autocorrelation of the signal, as it has been used successfully for muscle artifact detection. These different feature combinations were compared on data of known improvement or artifact reduction [21]. These datasets were created by overlaying a resting state EEG recording with simulated myogenic activity. The myogenic artifacts were artificially created by generating a time series from an extracted PSD of noisy data (of the same participant).

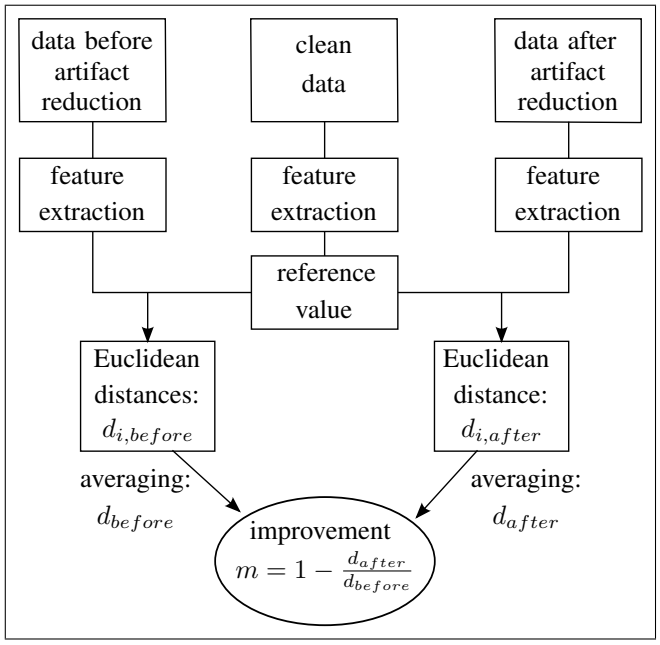

Fig. 2. Flow chart of the evaluation methodology procedure. First shown in [12].

Using this simulation tool we created data with known artifact reductions, denoted m_{SNR}, in the range of 0 % to 100 % [12]:

$$m_{\text{SNR}} = 1 - \frac{\text{SNR}_{\text{before}}}{\text{SNR}_{\text{after}}} \tag{1}$$

The feature selection process was initialized with an empirically defined SNR for the artificially created 'before artifact reduction' dataset. The SNR for all other datasets, e.g. 30 % improvement ($m_{\text{SNR}} = 0.3$) from the fixed starting point, was calculated by reformulating (1) and solving for $\text{SNR}_{\text{after}}$. Each dataset of known improvement was evaluated with the 'before artifact reduction' dataset using the objective measure (with all possible subsets of our candidate features). The feature combination that showed the lowest mean absolute error over the complete percentile range was selected. Afterwards we tested this feature selection on a disjunct dataset from a different study participant.

III. RESULTS

The feature selection process for the objective measure returned the following best subset of features:

1) mean
2) standard deviation
3) kurtosis
4) normalized power from 35 Hz to 50 Hz
5) mean value of the autocorrelation

The validation of these features showed a mean absolute error of 1.20 %.

One measurement had to be completely removed from the evaluation as it contained mostly non-myogenic artifacts (presumably from cable movement). The remaining 62 recordings, respectively the resulting 124 datasets (each recording once with and once without the EMG data), were evaluated after the artifact reduction process. Two datasets (both without EMG data) decreased in their quality, i.e. the evaluation returned negative values, after the artifact reduction. For the averaged improvement rates we assigned these to be zero.

The overall average artifact reduction for all exercises was 41.3 % with and 37.9 % without the EMG data. The artifact reduction was improved in every exercise except the ergometer test by including the EMG into the ICA decomposition (Fig. 3).

IV. DISCUSSION

We increased the accuracy of our previously introduced objective improvement measure by performing a feature selection on a larger, more diverse candidate set. The absolute error was reduced from 3.18 % to 1.20 %. The $\text{SNR}_{\text{before}}$ was determined empirically and fixed during the feature selection. We therefore assumed a specific degree of myogenic artifacts as a starting point. In future research we will perform the feature selection also with different SNR starting points to guarantee the accuracy of the objective measure for other applications.

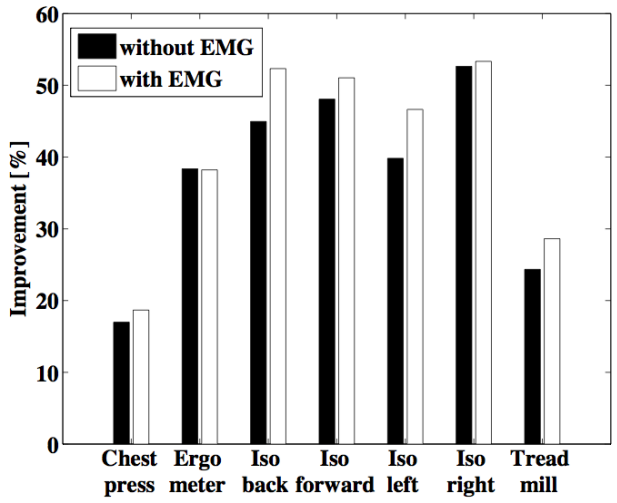

Fig. 3. Averaged improvement for each exercise - with and without the EMG data.

The results of our study showed an averaged increase in artifact reduction of −0.2 % to 7.4 %, depending on the exercise, if the EMG data is included. The largest improvements were in the isometric contraction exercises. The artifacts in these datasets were mostly caused by the muscles we monitored with the EMG. This indicates that including the EMG data of the muscles that most interfere with the EEG signal in the ICA decomposition raises the accuracy of the ICA model and makes the ICs more distinguishable.

We fixed the number of ICs that were rejected to achieve an unbiased comparison. Rejecting a different number of components, or not limiting the number at all, could produce different results. The error of the classifier used for selecting the ICs to reject also propagates to the results of this study.

The EMGs were placed to record the activity of the sternocleidomastoid and the trapezius muscle because of their close proximity to the EEG electrodes and negligible signal strength of muscles farther away [22]. However, propagation of electrical signals through the body is a highly complex problem, e.g. because of the different electric resistivity of different tissue types [23]. In future studies the influence of including the EMG of different muscles should be evaluated.

Our study showed that the ICA-based reduction of electromyogenic artifacts can be improved by including the EMG data of artifact-inducing muscles. This approach could prove beneficial for locomotor disorder research, brain-computer interfaces, neurofeedback, and most other areas where brain activity during movement has to be analyzed.

ACKNOWLEDGMENT

This work was supported by the Bavarian Ministry for Economic Affairs, Infrastructure, Transport and Technology and the European fund for regional development.

REFERENCES

[1] T. Thompson, T. Steffert, T. Ros, J. Leach, and J. Gruzelier, "EEG applications for sport and performance," *Methods*, vol. 45, no. 4, pp. 279–288, Aug. 2008.

[2] D. Vernon, "Can neurofeedback training enhance performance? An evaluation of the evidence with implications for future research," *Appl Psychophysiol Biofeedback*, vol. 30, no. 4, pp. 347–364, 2005.

[3] S. Makeig, K. Gramann, T.-P. Jung, T. J. Sejnowski, and H. Poizner, "Linking brain, mind and behavior," *Int J Psychophysiology*, vol. 73, no. 2, pp. 95–100, 2009.

[4] J. T. Gwin, K. Gramann, S. Makeig, and D. P. Ferris, "Electrocortical activity is coupled to gait cycle phase during treadmill walking," *NeuroImage*, vol. 54, no. 2, pp. 1289–1296, Jan. 2011.

[5] ——, "Removal of Movement Artifact From High-Density EEG Recorded During Walking and Running," *J Neurophysiol.*, vol. 103, no. 6, pp. 3526–3534, Jun. 2010.

[6] P. Reis, F. Hebenstreit, F. Gabsteiger, V. von Tscharner, and M. Lochmann, "Methodological aspects of eeg and body dynamics measurements during motion." *Front Hum Neurosci.*, vol. 8, no. 156, 2014.

[7] I. I. Goncharova, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "EMG contamination of EEG: spectral and topographical characteristics," *Clin Neurophysiol.*, vol. 114, no. 9, pp. 1580–1593, Sep. 2003.

[8] A. Bell and T. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comput.*, vol. 7, no. 6, pp. 1129–1159, 1995.

[9] S. Makeig, A. Bell, T. Jung, and T. Sejnowski, "Independent component analysis of electroencephalographic data," *Adv Neural Inf Process Syst.*, pp. 145–151, 1996.

[10] A. Hreljac, R. T. Imamura, R. F. Escamilla, and W. B. Edwards, "When does a gait transition occur during human locomotion," *J Sports Sci Medicine*, vol. 6, no. 1, pp. 36–43, 2007.

[11] A. M. Jones and J. H. Doust, "A 1% treadmill grade most accurately reflects the energetic cost of outdoor running," *J Sports Sci.*, vol. 14, no. 4, pp. 321–327, 1996.

[12] H. Leutheuser, F. Gabsteiger, F. Hebenstreit, P. Reis, M. Lochmann, and B. Eskofier, "Comparison of the AMICA and the InfoMax algorithm for the reduction of electromyogenic artifacts in EEG data," in *Conf Proc IEEE Eng Med Biol Soc*, Osaka, Japan, 2013, pp. 6804–6807.

[13] G. Gratton, M. Coles, and E. Donchin, "A new method for off-line removal of ocular artifact," *Electroencephalogr Clin Neurophysiol.*, vol. 55, no. 4, pp. 468–484, 1983.

[14] P. J. Allen, G. Polizzi, K. Krakow, D. R. Fish, and L. Lemieux, "Identification of eeg events in the mr scanner: the problem of pulse artifact and a method for its subtraction," *NeuroImage*, vol. 8, no. 3, pp. 229–239, 1998.

[15] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J Neurosci Methods*, vol. 134, no. 1, pp. 9–21, 2004.

[16] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Comput.*, vol. 9, no. 7, pp. 1483–1492, 1997.

[17] J. Palmer, K. Kreutz-Delgado, and S. Makeig, "AMICA: An Adaptive Mixture of Independent Component Analyzers with Shared Components," unpublished.

[18] J. Palmer, "AMICA - Adaptive Mixture ICA," http://sccn.ucsd.edu/~jason/amica_web.html, Accessed on March 25th, 2014.

[19] F. Gabsteiger, H. Leutheuser, P. Reis, M. Lochmann, and B. M. Eskofier, "SVM for Semi-automatic Selection of ICA Components of Electromyogenic Artifacts in EEG Data," in *15th Int Conf Biomedical Eng.*, Singapore, 2014, pp. 132–135.

[20] U. Jensen, M. Ring, and B. Eskofier, "Generic Features for Biosignal Classification," in *Proc Sportinformatik 2012*, 2012, pp. 162–168.

[21] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical Correlation Analysis Applied to Remove Muscle Artifacts From the Electroencephalogram," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2583–2587, Sep. 2006.

[22] K. Gramann, J. T. Gwin, N. Bigdely-Shamlo, D. P. Ferris, and S. Makeig, "Visual evoked responses during standing and walking," *Front Hum Neurosci.*, vol. 4, no. 202, 2010.

[23] T. J. C. Faes, H. A. van der Meij, J. C. de Munck, and R. M. Heethaar, "The electric resistivity of human tissues (100 hz-10 mhz): a meta-analysis of review studies," *Physiol Measurement*, vol. 20, no. 4, p. R1, 1999.