Wavelet-Independent Component Analysis to remove Electrocardiography Contamination in surface Electromyography

Joachim Taelman, Sabine Van Huffel, *Senior Member, IEEE*, Arthur Spaepen

***Abstract*— Removing artifacts from biomedical signals, such as surface electromyography (sEMG), has become a major research topic in biomedical signal processing. In electromyography signals, a source of contamination is the electrophysiological signal of the heart (ECG signals). This contamination influences features extracted from the sEMG, especially during low-activity measurements of the muscles such as during mental stress. As the heart is a muscle, the frequency content of the heart signals overlaps the frequency content of the muscle signals, so basic frequency filtering is not possible. In this paper, we present the results of a recently developed algorithm: wavelet-independent component analysis. We compare these results with the widely described algorithm of ECG template subtraction for removing ECG contamination.**

I. Introduction

ELECTROPHYSIOLOGICAL signals are widely used to derive important information related to physiological and pathophysiological behavior of the human body. Electrocardiography (ECG) measures the electrical signals of the heart and is widely accepted as a tool to estimate the physiology and the pathophysiology of a human heart. Since the nineties, research has increasingly focused on electroencephalography (EEG) and electromyography (EMG), electrical signal measures of the brain and muscles respectively.

Information on muscle activity can be retrieved from the signal measured on the skin surface, the so-called surface EMG (sEMG), or from a muscle fiber within the muscle itself using a needle electrode. EMG records the electro-physiological signal of the muscle. Nerve cells and muscle fibers are depolarized when activated by a certain threshold voltage. This results in propagation of a depolarization wave along the nerve and muscle fiber. The electrical wave is a direct cause of muscular contraction. Muscles comprise several motor units, a single nerve unit, and the corresponding muscle fibers. The measured sEMG is the sum of the depolarization waves of several motor units in the environment of the electrode.

Electrophysiological signals of muscles hold information, so that certain features can be extracted from sEMG to predict muscle behavior. Currently the most frequently extracted features of sEMG are related to its amplitude (root mean square, RMS) and to its frequency content (mean and median power frequency, MPF and MdPF). These features are very sensitive to artifacts in the sEMG, especially with small signals. In the ConText-project [1], [2], we are trying to deduce mental stress from the muscle activity. It has been indicated that mental stress induces an increase in muscle activity, albeit subtle [3]. Artifacts cause a loss in resolution of the frequency content and amplitude of the signal, reducing accuracy; the number of artifacts therefore needs to be minimized.

Electrophysiological signals of the heart (ECG) are a considerable source of contamination in sEMG signals [4], [5], [6]. As the heart is a muscle, the frequency range of the ECG is similar to that of the EMG. It has been reported that ECG contamination induces huge errors with respect to the amplitude and frequency of the raw signal, especially for low-activity signals [7]. High pass filtering, which has been suggested as a possible solution [8], implies filtering out data, thus discarding potentially useful information.

In this paper, we use two algorithms to remove the ECG artifacts. One is the generally used algorithm known as template subtraction [4], [5], [6]. This is a simple algorithm: a calibrated template is subtracted from the measured signal of the heartbeat to remove the ECG contamination. The second technique is a recently developed algorithm based on wavelet-Independent Component Analysis (wICA) [9]. This algorithm combines wavelet analysis with a blind source separation technique to overcome the shortcomings of wavelet analysis (overlapping spectrum) and ICA (lack of redundancy in the number of channels, or signals, compared with the number of sources). We compare the results of these two algorithms.

II. Methods

A. EMG recordings

Earlier research suggests that mental tasks induce increased neck muscle activity [10]. To explore whether mental stress can be deduced from muscle activity, we studied the effect of mental tasks on muscle activity in the neck region. Our protocol distinguished four states: rest,

Manuscript received April 16, 2007. This work was partially sponsored by the European Commission under Grant IST-027291 (ConText).

J. Taelman is with the Department of Biomedical Kinesiology at the Katholieke Universiteit Leuven, 3001 Heverlee, Belgium (e-mail: joachim.taelman@ faber.kuleuven.be).

S. Van Huffel is head of the biomedical signal processing group at the Electrical Engineering Department at the Katholieke Universiteit Leuven, 3001 Heverlee, Belgium.

A. Spaepen is head of the ergonomic group at the Department of Biomedical Kinesiology at the Katholieke Universiteit Leuven, 3001 Heverlee, Belgium.

postural load, mental task, postural load and mental task. We measured the M. Trapezius pars descendens, M. Infraspinatus and the M. Deltoideus medius from the left and right shoulders, together with the subject's heart rate (ECG). The sEMG was measured differentially with pre-gelled contact electrodes (Ag-AgCl, 10 mm diameter, Nikomed, Denmark), positioned according to SENIAM European recommendations [11]. The signals were digitized at a sample rate of 1000 Hz.

We report results using the two algorithms described below, for data from a single subject (male, 23 years old).

B. Algorithms

1) ECG template subtraction: ECG template subtraction uses the ECG signal's periodic characteristics and the heart rate measured. An ECG template is subtracted from the EMG signal at each occurrence of the heartbeat. The moment when an ECG signal occurs can be localized very accurately from the QRS-complex in an EMG measurement using the Pan-Tompkins algorithm [12]; this is used as the moment of ECG contamination.

The ECG template subtraction algorithm involves two steps. In the first, a template of the ECG interference which is contaminating the EMG signal is derived. This is necessary because the nature of the ECG contamination depends on the position of the electrodes and differs for each EMG signal measured. The ECG template is calculated by averaging each of 500 EMG samples from 20 equal length data sets of the EMG signal in the rest state, at heartbeats in the ECG. This method assumes that EMG has a zero mean Gaussian distribution. This results in a template of ECG contamination comprising 500 (average) values, which is used as a subtraction template [6].

In the second step, the subtraction template is used to clean ECG contamination in the complete muscle signal. Each heartbeat is localized in the ECG signal, which thus allows detection of contamination occurrences. For each detected heartbeat, the ECG template is subtracted from the EMG signal.

With the template calibrated for the signal itself, the delay between the heart ECG signal and the ECG contamination in the EMG is taken into account.

These two steps need to be repeated for each channel measured. An ECG subtraction template can only be used for a particular channel measurement as the template varies for different channels.

2) Wavelet-Independent Component Analysis (wICA): wICA [6] combines two techniques, wavelet analysis and ICA, to overcome the shortcomings of each technique.

- The wavelet transform [13] is a time-frequency representation of a signal that was introduced to overcome the limitations in time and frequency resolution occurring with the classical Fourier transform and subsequently developed versions, such as the Short Time Fourier Transform (STFT). Instead of using a sine wave, which is the basic function for the Fourier Transform, a basic waveform is used. This basic waveform ψ can be modified to basic functions $\psi_{a,b}$ obtained from dilations and shifts of the basic waveform. The basic waveform is shown in (1), where *a* is the scaling parameter and *b* represents the translation parameter.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (1)$$

The continuous wavelet transform (CWT) of a given input signal x(t) is shown in (2) .

$$CWT(a,b) = \int x(t)\psi^{*}_{a,b}(t)\,dt \quad (2)$$

The representation of the CWT is redundant. The discrete wavelet transform (DWT) removes the redundancy of the CWT by using discrete steps for scale and translation.

DWT can be used to remove artifacts. Its application is based on a good spectral separation between the original signal and the artifact. DWT must be applied to a single channel recording as it separates this single channel into multiple channels.

- Independent Component Analysis (ICA) [11] is a method for solving the Blind Source Separation problem. This technique extracts statistically independent components, *S*, from a set of measured signals, *X*, as shown in (3).

$$X = A.S \quad (3)$$

The goal of ICA is to estimate the mixing matrix, *A*, and/or the source vector, *S*, from the measured data, *X*. ICA assumes that the different sources are statistically independent and based on this assumption estimates the mixing matrix recursively.

Unlike wavelet analysis, however, ICA can only be applied to multichannel recordings. Additionally, our analysis is applied to real data, with insufficient recorded channels to retrieve the hidden sources, due to noise.

- Wavelet-Independent Component Analysis combines these two techniques. In the first step, a wavelet analysis is applied for each channel. Each channel is split into a number of wavelet components, which are used as input for the ICA. With this improvement, more information from the original input signal is used to find the independent sources.

The second step of this algorithm is to find the original data, cleaned of the artifacts. When a source of contamination is found in one component of the ICA, this source can be set to zero. From the inverse step of the ICA, we are able to find the wavelet-components. From these components, the original, cleaned signal can be reconstructed.

C. Validation

We compare results from the cleaning algorithms with the original signals, which include the ECG contamination. As the template subtraction algorithm is widely described and used, these results are used as a reference to validate the results of the recently developed wICA algorithm. We use

visual validation to compare the results.

III. Results

Due to the large size of the entire data set, we are restricted here to demonstrating the results of the algorithms using only a small part of the data. In our research we are interested in small increases in muscle activity, and have therefore used a typically representative part (from second 180 to 230) of 6 EMG recordings. Figure 1 shows this data.

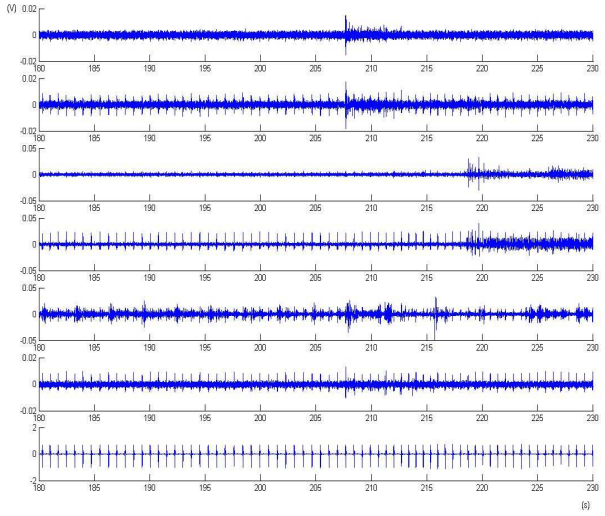

Fig. 1. Part of the signals recorded during the test. The first six signals are the sEMG of the M. Deltoideus medius left, M. Infraspinatus left, M. Trapezius pars descendens left, M. Trapezius pars descendens right, M. Infraspinatus right and the M. Deltoideus medius right respectively. The last channel represents the measured ECG of the subject during the test.

The contamination of ECG in the sEMG is clear in channel two, four and six. Traces of the contamination are apparent in channels three and five. The influence of the ECG artifact on the analysis of the measured muscle activity is thus evident from figure 1.

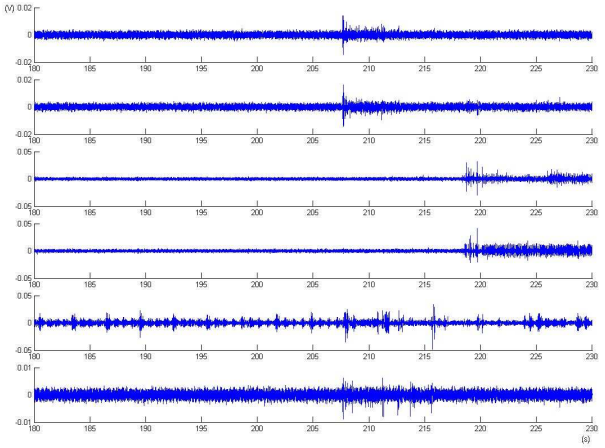

Fig. 2. Cleaned sEMG with template subtraction. The channels represent the sEMG of the M. Deltoideus medius left, M. Infraspinatus left, M. Trapezius pars descendens left, M. Trapezius pars descendens right, M. Infraspinatus right and the M. Deltoideus medius right, respectively.

The first algorithm used to clean the EMG signals was the template subtraction algorithm. This algorithm is performed on each muscle recording separately. This is necessary because the template differs for each muscle recording. A subtraction template was calibrated for all muscles measured. For each heartbeat detected in the ECG, the template was subtracted from the corresponding EMG. The results of this procedure are shown in figure 2. All the ECG artifacts, which are clearly visible in figure 1, were removed with this algorithm.

The second algorithm used to eliminate the ECG contamination was the wICA. Some preliminary decisions were necessary before applying the algorithm. The first choice was whether to include the ECG in the system: we tested with and without the ECG and concluded that the algorithm performed better with the ECG included. In the test without the ECG included, however, the algorithm was still able to reduce the contamination. In the test with the ECG included, the ECG and thus the ECG contamination of the other channels became dominant sources and separation of sources was more accurate. As we had measured the ECG during the test, we were able to include it as a signal.

Secondly we needed to decide whether to use DWT. We used the 'Daubechies 6' [10] as basic wavelet and six decompositions performed well.

Applying this algorithm to the original data resulted in six dominant ECG sources and eight other relevant sources. The reconstruction of the cleaned signal is shown in figure 3.

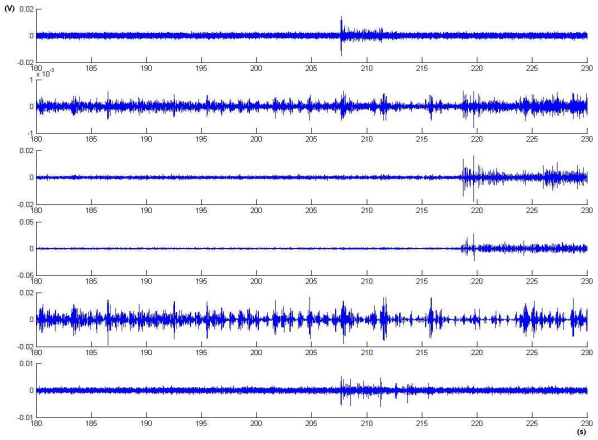

Fig. 3. Cleaned sEMG with wICA. The channels represent the sEMG of the M. Deltoideus medius left, M. Infraspinatus left, M. Trapezius pars descendens left, M. Trapezius pars descendens right, M. Infraspinatus right and the M. Deltoideus medius right, respectively.

Comparing the reconstructed signal with the result after template subtraction (which we used as the reference), clearly shows that the algorithm performed well on channels one, three, four and six. The reconstruction of signals two and five (the measurements on the M. Infraspinatus right and left) differs somewhat compared with the reference (shown in figure 2). This indicates that the wICA algorithm was unable to find the real sources of these two signals.

In a subsequent test we split the original signal into two

test sets: the first test set comprised the original channel one, three, four and six, and the second test set comprised signals two and five. We completed the wICA algorithm separately on these two data sets and the results are shown in figure 4. The reconstructions of the first data set are similar to the results of the whole data set, but the reconstructions of the other two signals were improved over the results in figure 3.

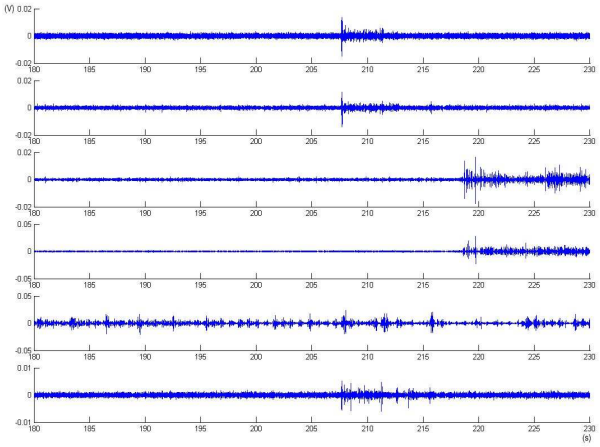

Fig. 4. Cleaned sEMG with wICA. The channels represent the sEMG of the M. Deltoideus medius left, M. Infraspinatus left, M. Trapezius pars descendens left, M. Trapezius pars descendens right, M. Infraspinatus right and the M. Deltoideus medius right, respectively.

This result indicates that the sources of channels one, three, four and six are dominant, compared with the sources of channels two and five. We conclude that the wICA algorithm is capable of cleaning ECG contamination from the sEMG.

IV. Discussion

wICA is a recently developed algorithm to eliminate artifacts in biomedical signals. In this paper we have tested this algorithm for the elimination of ECG contamination. The results were compared with a known and robust template subtraction algorithm. The wICA algorithm was capable of eliminating the ECG artifacts, which is a promising result. Further, algorithm performance was improved when the original data of six channels was split into two data sets of two and four channels.

We must point out that we included the ECG as raw data; this procedure outperformed the results of applying the wICA algorithm without inclusion of the ECG to the data set.

Regarding wICA, arbitrary decisions were made regarding the wavelet type and the number of decompositions. Further research is required to accurately inform such decisions.

Acknowledgment

This work was partially sponsored by the European Commission under Grant IST-027291 (ConText).

We thank the European partners of ConText: Philips Research (The Netherlands), TNO (The Netherlands), Clothing Plus (Finland), TITV (Germany), and TU Berlin (Germany).

References

[1] ConText project website; http://www.context-project.org.

[2] J. Taelman, “Contactless emg sensors for continuous monitoring of muscle activity to prevent musculoskeletal disorders,” in *Proc. of the first Annual Symposium of the IEEE/EMBS Benelux Chapter*, Brussels, Belgium, Dec 2006, pp 223-226.

[3] U. Lundberg, “Psychological stress and EMG activity of the Trapezius Muscle”, *International Journal of Behavioral Medicine*, 1(4), 354-370.

[4] K. Mekhora and L. Straker, “Elimination of electrocardiograph noise in neck muscle electromyography,” in *Ergonomics Australia*, 1st. vol. 13, R. Burgess-Limerick, 1999, pp. 15–21.

[5] P. Zhou and T. A. Kuiken, “Eliminating cardiac contamination from myoelectric control signals developed by targeted muscle reinnervation,” in *Physiological Measurement*, vol. 27, Institute of physics publishing, 2006, pp. 1311–1327.

[6] A. Bartolo, C. Roberts, R. Dzwonczyk and E. Goldman, “Analysis of diaphragm EMG signals: comparison of gating vs. subtraction for removal of ECG contamination,” in *J. Appl. Physiol*, vol. 80, 1996, pp. 1898–1902.

[7] E.A. Clancy, E.L. Morin, R. Merletti, “Sampling, noise-reduction and amplitude estimation issues in surface electromyography,” in *Journal of Electromyography and Kinesiology*, vol. 12, 2002, pp. 1-16.

[8] M.S. Redfern, R.E. Hughes, D.B. Chaffin, “ High-pass filtering to remove electrocardiographic interference from torso EMG recordings, “ in *Clinical Biomechanics*, vol. 8, no. 1, 1993, pp. 1993.

[9] B. Azzerboni, M. Carpentieri, F. La Foresta and F.C. Morabito, “Neural-ICA and Wavelet Transform for artifacts removal in EMG,” in *Proc. Of the IEEE International Joint Conference on Neural Networks*, vol. 4, 2004, pp. 3223–3228.

[10] R.H. Westgaard and R. Bjørklund, “Generation of muscle tension additional to postural muscle load,” in *Ergonomics*, 6th ed. vol. 30, 1987, pp. 911–923.

[11] H. Hermens, “*SENIAM, European Recommendations for Surface ElectroMyoGraphy* ,” Roessingh Research and Development, 1999

[12] J. Pan and W.L. Tompkins, “A Real-Time QRS Detection Algorithm,” in *IEEE Trans. Biomed. Eng.*, vol. 32, 1985, pp. 230–236.

[13] I. Daubechies, *Ten Lectures on wavelets*, CMBS, SIAM Publ., 1992.

[14] A. Hyvärinen and E. Oja, “Independent Component Analysis: Algorithms and Applications,” in *Neural Network.*, 4th – 5th ed vol. 313, 2000, pp. 411–430.