Removing ECG-contamination from EMG signals

1. Introduction
As is well known the heart also creates strong electromagnetic fields. The electric potential from the heart can be picked up as the electrocardiographic signal (ECG). Since this electric potential is conducted to all parts of the body it may interfere with the electromyographic (EMG) signals from the muscles that are detected using surface electrodes. This may be a minor problem when recording EMG e.g. from leg muscles, but for vaginal EMG and EMG from the region of the thorax the ECG-signal may be quite strong. It is therefore desirable to have a method to remove the ECG-part from the contaminated EMG-signal without significantly distorting the EMG-part. There are two basic methods for this kind of problems of signal separation: those which employ reference signals, and those which only have recourse to the contaminated signal itself. A leading representative among the former methods is the adaptive filter method whose basic version is the Least-Mean-Square algorithm (LMS) by Widrow and Hoff (1960). In the later category feature the time-frequency methods of which the most common ones are the bandpass filters.

![Fig. 1. The figure shows vaginal EMG contaminated by strong ECG-pulses. Data provided by Dr Peter Konrad.](image.png)

2. Adaptive filters
Suppose we have system with an output $x$ depending on some other variables $y$. In the present case $x$ is the contaminated EMG-signal, and $y$ is the reference signal, a nearly pure sample of ECG. The linear adaptive cleaning filter is based on the assumption that the contaminated EMG-signal can be written as a sum.
of pure EMG-signal $\xi$ and the disturbing ECG-part $\zeta$ and that $\zeta$ can be estimated by a linear function $z$ of the reference ECG-signal ($m$ is the length of the filter)

$$z_i = \sum_{k=1}^{m} a_k y_{i-m+k}$$

Equ(2) describes how the ECG-signal at the EMG-electrode site depends on the the ECG reference signal from a different site. In a detailed physical model we would have to set up differential equations for the electrical potential describing how it propagates through the body and how the potential values at one point are correlated with its values at another point. Equation (2) may be regarded as a partial discretized version of such a set of differential equations. Given equ (2) then the pure EMG-signal $\xi$ can be estimated by taking the difference

$$e_i = x_i - z_i$$

The idea of an adaptive filter is that the filter-coefficients $a_k$ are to be automatically adjusted through a feed-back loop. If we assume that the pure EMG-signal and the reference ECG-signal are uncorrelated, then the feed-back loop can be based on the objective to minimize the squared value of the difference (3). The idea is that the "cost function"

$$E(a) = \frac{1}{2} (x_i - z_i)^2$$

will reach a minimum exactly when the filter-coefficients $a_k$ are such that (2) will remove the ECG-contamination from the EMG-signal. Indeed, suppose we do several runs and take the average over the runs, then we get 

---

1 The difference $e_i$ is commonly called the "error" because in many applications the goal is not to clean the signal, but to try to predict $x$ from $y$ via a linear function (2); $e_i$ then describes the error of the prediction. For instance, the muscle force ($x$) can be quite successfully predicted from its EMG-signal ($y$) by using an adaptive linear filter (Herzog et al., 1995).
\[ \langle E(a) \rangle = \left\langle \frac{1}{2} (\xi_i + \zeta_i - z_i)^2 \right\rangle = \frac{1}{2} \left( \langle \xi_i^2 \rangle + 2 \langle \xi_i \zeta_i \rangle + \langle \zeta_i^2 \rangle \right) = \frac{1}{2} \left( \langle \xi_i^2 \rangle + \langle (\zeta_i - z_i)^2 \rangle \right) \]

The last equality follows due to the assumption that the pure ECG- and EMG-signals are uncorrelated, which seems to be a quite reasonable assumption. From (5) we see that when the cost function reaches a minimum (on the average) this will imply that

\[ \left\langle (\zeta_i - z_i)^2 \right\rangle \approx 0 \]

whence the estimation \( z \) (2) must approach the pure ECG-value. The feed-back loop in the cleaning algorithm will thus be based on calculating the cost (4) and adjust the coefficients \( a_k \) such as to reduce the cost in the next step. Let \( a(n) \) denote the filter coefficients at time step \( n \) arranged as a vector,

\[ a(n) = (a_1, \ldots, a_n) \]

and similarly for the reference signal \( y(n) \)

\[ y(n) = (y_{n-m+1}, \ldots, y_n) \]

Thus, using the vector notation we may write (2) as \( z_n = a(n) \cdot y(n) \). The change in the cost function (4) during a single step may be approximated by

\[ \Delta E(a(n)) \approx \frac{\partial E}{\partial a} \Delta a(n) = -e_n y(n) \Delta a(n) \]

It follows that if the correction \( \Delta a \) to the filter parameters is chosen according to

\[ \Delta a(n) = \eta e_n y(n) \]

for a positive stepsize parameter \( \eta \), then the filter parameters will be adjusted such as to reduce the cost (4). Inserting (10) in (9) we obtain

---

F Borg (borgbros@netti.fi), Mika Herrala. Biosignals Project, Jyväskylä University, Chydenius Institute.
These points suggest the following adaptive cleaning algorithm:

\[ e_n = x_n - a(n) \cdot y(n) \]
\[ a(n+1) = a(n) + \eta e_n \cdot y(n) \]

That is, at time step \( n \) we calculate the cleaned signal \( e_n \) according to (12a) and update the filter parameters \( a(n+1) \) for step \( n + 1 \) according to (12b). A graphical presentation of the algorithm is given in the next figure.

![Schematic presentation of the adaptive algorithm.](image)

The linear filter (2) acts basically as a low-pass filter with a time constant inversely proportional to the step parameter \( \eta \) (Haykin 1999). If \( \eta \) is very small it takes longer for the filter to converge. A bigger value accelerates the convergence but there is the risk for big \( \eta \)-values that the updating procedure (12a) diverges. If we insert (12a) into (12b) and take the averages (over many runs) we get

\[ \langle a(n+1) \rangle = \langle a(n) \rangle - \eta \langle [a(n) \cdot y(n)] \cdot y(n) \rangle + \eta \langle x_n \cdot y(n) \rangle \]

Suppose that for large \( n \) we have a convergence

\[ \langle a(n) \rangle \to c_\infty \text{ as } n \to \infty \]
If we assume that \( a(n) \) and \( y(n) \) are statistically independent then we get from (13) in the limit (14),

\[
(15) \quad x_n \cdot y(n) = R c_\infty
\]

with the *correlation matrix* \( R \) defined by

\[
(16) \quad R_{jk} = \langle y(n)_j \cdot y(n)_k \rangle
\]

Defining next

\[
(17) \quad \Delta(n) = \langle a(n) \rangle - c_\infty
\]

and using (16) in (13) we get finally (Hänsler 1997),

\[
(18) \quad \Delta(n+1) \approx \Delta(n)(1 - \eta R(n))
\]

This suggests that (13) indeed converges if we have

\[
(19) \quad 0 \leq \eta \leq \frac{2}{\lambda_{\text{max}}}
\]

(i.e. the last factor on the rhs of (18) must be less than 1) where \( \lambda_{\text{max}} \) is the biggest eigenvalue of \( R \). The assumptions for this argument are not likely to be valid for the ECG-EMG case but (16) hints that the step parameter should be scaled as the inverse of the square of the characteristic amplitude of the reference signal \( y \). Since the average of the reference ECG-signal \( y \) is zero we can use the standard deviation \( \sigma \) of \( y(n) \) as a measure of its size. The rule (12b) is then replaced by

\[
(20) \quad a(n+1) = a(n) + \frac{2\mu}{m(\sigma_y)^2} y(n)
\]

where \( \mu \) is a new step parameter which is expected to be in the interval \( (0, 1) \) and \( m \) is

---

F Borg (borgbros@netti.fi), Mika Herrala. Biosignals Project, Jyväskylä University, Chydenius Institute.
the length of the filter. The specific form (20) is based on a conservative upper bound estimate for $\lambda_{\text{max}}$ given by

$$\lambda_{\text{max}} \leq \sum_k \lambda_k = \text{Trace}(R) = \sum_k \langle y_k^2 \rangle = m \sigma_y^2 \quad (21)$$

![Figure 3. EMG from the region of the chest. The ECG-reference signal obtained from an electrode close to the heart. Data sampled at 2000 S/s.](image)

We have applied the above LMS-algorithm to a contaminated surface EMG recorded from the chest, and used the ECG-signal from an electrode close to heart as a reference signal.

One starts by setting the initial filter coefficients to zero. The figure below shows the result of applying the above LMS-algorithm using $\mu = 0.025$ in (17) and the filter length $m = 25$. 

F Borg (borgbros@netti.fi), Mika Herrala. Biosignals Project, Jyväskylä University, Chydenius Institute.
The LMS-algorithm has been extensively used in solving problems similar to the present one so alone on that basis one would expect to be able to make the algorithm work here too\(^3\). The LMS-algorithm can be considered as a special case of neural networks. We have implemented the LMS-algorithm as a realtime version in a pilot software that has been tested successfully with similar setups as in the above example.

\(^3\) For a similar application of the adaptive cancellation to eliminate muscle sound interference in vibroarthography see the paper by Zhang et al. (1994) and Rangayyan (2002).
3. Time-frequency methods and a "hybrid" algorithm

In the case we do not have a reference signal we cannot use the standard LMS-algorithm. A commonly used method then is to investigate whether the disturbing signal can be separated from the contaminated signal in the frequency space. The spectrum of the ECG-spectrum extends to about 60 Hz whereas for surface EMG one typically studies a frequency band from 20 Hz to 500 Hz.

As the mean frequency of EMG can be around 60 Hz there may be a significant spectral overlap with the ECG-signal. Still, Bruce (2001 p. 281) describes an example where ECG is removed using a 2nd order Butterworth high-pass filter with a cutoff frequency of 70 Hz. Later a somewhat improved algorithm is presented (Bruce 2001pp. 337-8) using a digital equiripple FIR-filter of order 91. After the cleaning the ECG-pulses are still quite prominent. The filtering procedure will also affect the regions of the EMG-signal between the pulses. Thus there does not seem to be much to gain by this filtering procedure. In case one has silent EMG-periods where the ECG comes clearly through, like the beginning in figure 1, one might try to extract the typical shape of an ECG-pulse and use it to track and subtract the ECG-pulses in the EMG-burst sections.4

4 As Bruce himself points out: "A markedly more successful approach to the difficult problem of ECG removal from EMG signals involves the application of a topic beyond the current text, namely, adaptive filtering" (Bruce 2001 p. 282).

5 This seems to be the basic idea tested by the Noraxon software group (e-mail from Ilia Sakhenko).
Fig. 7. Short Time Fourier Transformation of a contaminated EMG-signal. The ECG-pulses are clearly visible. Frequency (Hz) along the vertical axis.

Fig. 8. Mexican wavelet transform (width parameter $w = 81$ ms) applied to a section of EMG with an ECG-pulse. The vertical axis represents the scales; in terms of frequency it covers the range from about 5 Hz to 160 Hz. The ECG-pulse is clearly visible in the wavelet plot as the dominant vertical stripes around $x = 1250$.

In figure 10 we have used the ECG-pulses in the quiet section in order to calculate an average ECG-pulse (PULSE) in figure 9.

F Borg (borgbros@netti.fi), Mika Herrala. Biosignals Project, Jyväskylä University, Chydenius Institute.
Fig. 9. The average of four ECG-pulses in the quiet section of the EMG. The location of the peak values have been used for aligning the pulses.

From this we construct an artificial ECG-signal $ECGA$ by adding a number of such pulses,

$$ECGA(t) = \sum_k \text{PULSE}(t-t_k)$$

based on the time instants $t_k$. The time instants $t_k$ are determined from the contaminated EMG-pulse e.g. by using a low-pass filter and extract the local maxima in subintervals centered on points where the signal dives below a given threshold value (signifies the presence of an ECG-pulse).

Fig. 10. This shows the effect of simply subtracting an averaged ECG-pulse from the contaminated EMG-signal. Because of the strong ECG-interference the pulses are easy to locate in the EMG-signal.
This might work for short EMG-bursts where the ECG-shape and size are not expected to change much and the ECG-signal has a bigger amplitude than the EMG. We can add an adaptive feature to this procedure by resizing the ECG-pulses in the artificial ECGA-signal by the local ECG-amplitude found in the contaminated EMG-signal. We can carry this one step further and make an "hybrid" adaptive filter by using the ECGA-signal as a reference signal in the LMS-algorithm described in the first section. The figure 11 shows the result of using this "hybrid" algorithm to the same data as in figure 10. As expected this seems to work quite well too. As described earlier, the LMS-algorithm can match the "hidden" ECG-signal with a transformed version of the reference signal. Still the convergence could perhaps be accelerated a bit by resizing the pulses in the ECGA-signal with the local ECG-amplitude extracted from the EMG. However, the most important aspect is to be able to locate the ECG pulses along the time axis in the contaminated signal so that we can construct the reference signal ECGA. Of course, if the ECG-signal is entirely hidden in the EMG then there is little point in attempting any "cleaning" to begin with.

![Fig. 11. The LMS-filter procedure applied to contaminated EMG using the "artificial" ECG-signal (ECGA) as a reference signal.](image)
References