PCA and ICA applied to Noise Reduction in Multi-lead ECG

I Romero

IMEC, Eindhoven, the Netherlands

Abstract

The performance of PCA and ICA in the context of cleaning noisy ECGs in ambulatory conditions was investigated. With this aim, ECGs with artificial motion artifacts were generated by combining clean 8-channel ECGs with 8-channel noise signals at SNR values ranging from 10 down to -10 dB. For each SNR, 600 different simulated ECGs of 10-second length were selected. 8-channel PCA and ICA were applied and then inverted after selecting a subset of components. In order to evaluate the performance of PCA and ICA algorithms, the output of a beat detection algorithm was applied to both the output signal after PCA/ICA filtering and compared to the detections in the signal before filtering.

Applying both PCA and ICA and retaining the optimal component subset, yielded sensitivity (Se) of 100% for all SNR values studied. In terms of Positive predictivity (+P), applying PCA, yielded to an improvement for all SNR values as compared to no cleaning (+P=95.45% vs. 83.09% for SNR=0dB; +P=56.87% vs. 48.81% for SNR=-10dB). However, ICA filtering gave a higher improvement in +P for all SNR values (+P=100.00% for SNR=0dB; +P=61.38% for SNR=-10dB).

An automatic method for selecting the components was proposed. By using this method, both PCA and ICA gave an improvement as compared to no filtering over all SNR values. ICA had a better performance (SNR=-5dB, improvement in +P of 8.33% for PCA and 22.92% for ICA).

1. Introduction

In the recent years, new ambulatory cardiac monitors have been developed for continuous ECG monitoring. These devices are portable with an autonomy which is increasing with the improvement of low-power micro-electronics. Integration of microprocessors allows performing some signal processing and automatic interpretation. However, in ambulatory conditions, noise increases with higher levels of activity. Motion artifact could reduce signal quality significantly, making ECG interpretation very difficult.

Several methods for noise reduction and motion artifact removal have been proposed in the literature. Traditional denoising techniques were based on time averaging [1] and frequency analysis such as filter banks [1] or the wavelet transform [2]. In adaptive filtering, a filter is applied after adjusting its parameters in time to a time varying noise. This is particularly useful when the noise is non-stationary as it is the case in ambulatory motion artifacts. However, a reference signal has to be additionally recorded together with the ECG. Several adaptive filtering approaches have been proposed to obtain an adequate reference signal such as measurement of skin-electrode impedance [3, 4], skin stretching measured with optical sensors [4, 5] or accelerometers [6, 7].

Blind source separation (BSS) techniques could be used for separating ECG and noise, as these signals are uncorrelated [8, 9]. In order to apply those methods, a multi-lead ECG recording is required and the different recorded leads should be linearly independent. The literature describing the use of BSS techniques for ECG denoising is scant. Principal Components Analysis (PCA) has been used for reducing noise in both single lead ECG segmented in time intervals [10] and multi-lead ECGs [11]. A combination of PCA and Independent Component Analysis (ICA) for ECG denoising was also proposed by Chawla [12].

In this work, the performance of PCA and ICA are investigated in the context of motion artifact reduction in ECG signals.

2. Methods

2.1. PCA and ICA

PCA and ICA are techniques that are commonly used in multivariate statistical analysis to reduce the number of dimensions from a numerical measurement of several variables. With this dimensional reduction, these techniques look for simplifying a statistical problem with the minimal loss of information. These methods are also used in signal processing for separating a linear combination of signals generated from sources that are statistically independent. This is performed by representing the data with a new coordinate system [13]. Applying PCA or ICA to *n* ECG leads that are statistically independent gives *n* new signals or principal/independent components. These transformations can be reverted after selecting a subset of components in order to filter out part of the original information.

2.2. Data collection

Clean ECG signals were obtained by recording 8-unipolar lead ECGs from 5 healthy subjects. For each subject 8 sets of 10 seconds were obtained while the subject was at rest. 8-channel noise recordings were obtained by placing 9 electrodes on the back of the subjects at the height of the lumbar curve where ECG signals were negligible. Then, the subjects were asked to move randomly. For each subject, 8 sets of 10 seconds were recorded. Signals were obtained with a sampling frequency of 1000 Hz, using a generic biosignal acquisition system from g.Tec (g.USBamp). All recordings were filtered by a high pass filter with cut-off frequency of 0.5 Hz and a 50 Hz notch filter. Each 8-channel noise signal was multiplied for a gain factor and added to each 8-channel clean ECG in order to obtain a specific SNR. SNR values ranging from 10 to -10 dB were considered. For each SNR value, 600 different combinations of clean ECG and noise were selected. Figure 1 shows one example of a clean ECG, a pure noise and a combination of both signals.

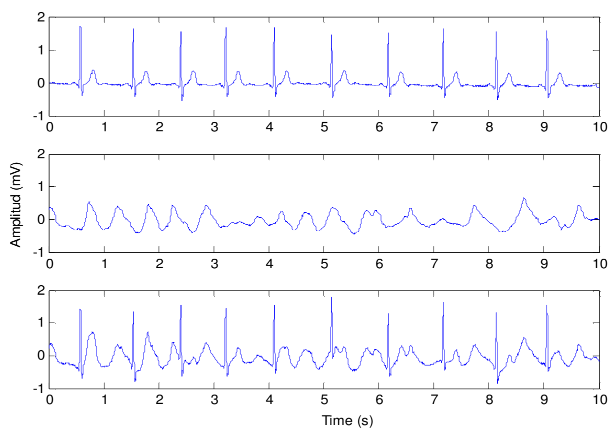

Figure 1. Extract of a clean ECG (upper panel), a noise (middle panel) and the sum of both signals (lower pannel). The SNR of the combined signal is of 0 dB.

2.3. Evaluation criteria

PCA or ICA was applied and then inverted after selecting a subset of principal/independent components. In order to evaluate the performance of PCA and ICA algorithm, the output of a beat detection algorithm [14] was used. The beat detector was applied to both the output signal after PCA and ICA filtering and to the signal before filtering. The detections were compared with the annotations obtained before adding the noise in order to calculate the Sensitivity (Se) and Positive Predictivity (+P). From these two parameters, +P is more sensitive to noise due to the fact that high voltage peaks in the ECG caused by motion artifacts, might be confused with QRS complexes and wrongly detected as such (false positive). The median over the 8 ECG leads was considered as representative value for each signal and the median over all signals was considered as representative value for each SNR. As median is more robust to outlier values, it was preferred to the arithmetic mean and standard deviation.

3. Results

SNR values from 10 to -10 dB in steps of 1 dB were considered. Both the noisy signal and the output signal from PCA/ICA filtered were then compared with the clean ECG in order to study the improvement due to the algorithm. Figure 2 shows an example of a clean ECG (before adding noise), a noisy ECG (when noise was added to a SNR=-10dB) and filtered ECG using ICA. The beat detection had a significant improvement after filtering as compared to the noisy signal.

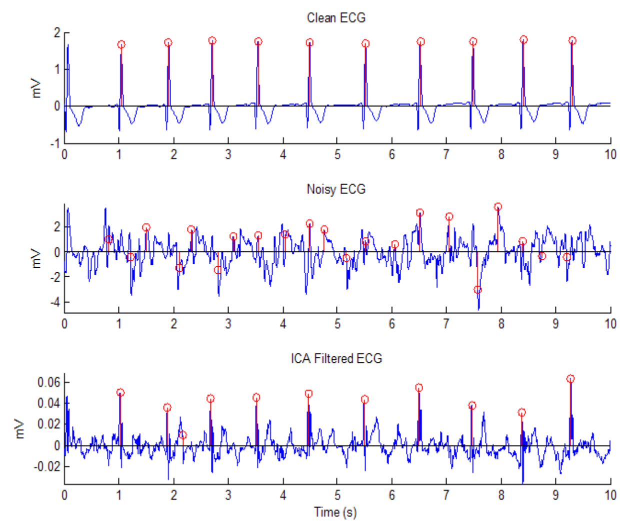

Figure 2. Extract of a clean ECG, a noisy ECG (SNR=-10dB) and filtered ECG using ICA. The beat detection had significantly improvement in performance after ICA filtering as compared to no filtering.

3.1. Optimal Component Selection

Initially, the optimal component subset was considered. The optimal component subset was defined as the one (from all possible combinations) that gave the highest correlation coefficient between the clean ECG signal and the output of the inverted PCA or ICA. Results are plotted in Figure 3.

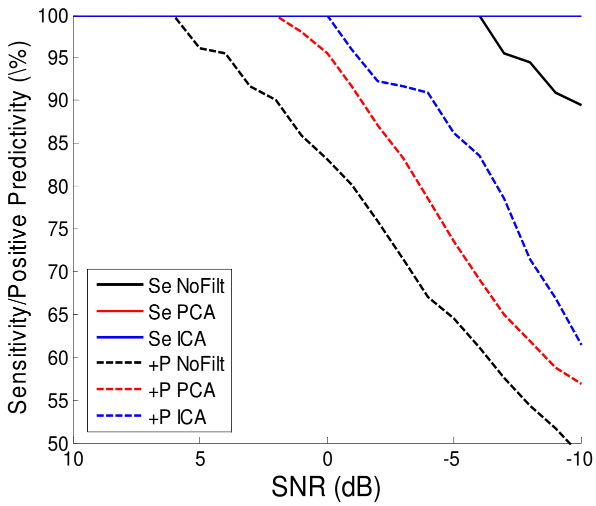

Figure 3. Beat detection output after applying PCA and ICA and retaining the optimal component subset, and with no filtered.

In terms of Se, the beat detector had a good performance on the non-filtered signal with Se=100% down to -6dB. When +P was considered, +P was 100% down to 6dB. Below that value, it dropped down to 83.09% at 0dB and 48.81% with SNR=-10dB.

Applying both PCA and ICA and retaining the optimal component subset yielded a Se of 100% for all SNR values studied.

Considering +P, applying PCA, yielded to an improvement for all SNR values (+P=95.45% for SNR=0dB; +P=56.87% for SNR=-10dB). However, ICA filtering gave a higher improvement in +P for all SNR values (+P=100.00% for SNR=0dB; +P=61.38% for SNR=-10dB).

3.2. Automatic Component Selection

In order to use PCA or ICA for denoising of ECG signals in an automatic algorithm, the selection of principal and independent components without human intervention is necessary [15]. A method for automatic selection of components was also investigated. Half of the dataset, i.e. 300 signals for each SNR value, was used as a training dataset for designing a selection method. This was subsequently evaluated in the other half of the dataset (evaluation dataset).

Kurtosis was calculated to identify which components correspond to ECG information. As the ECG signal is super-Gaussian as compared to motion artifact noise, components that had a kurtosis over a fixed threshold were selected while components below this threshold were rejected. The optimal threshold on the training dataset was found to be 50. When no components had a kurtosis over this threshold, then the component with the maximum kurtosis was selected.

This automatic method was used on the evaluation dataset (i.e. half of the signals from the original dataset) and compared with both no filtering and PCA/ICA when the optimal subset was retained. The median sensitivity obtained was 100% for all SNR values when using both PCA and ICA and in both situations: selecting the optimal component subset and using the automatic component selection (Figure 4).

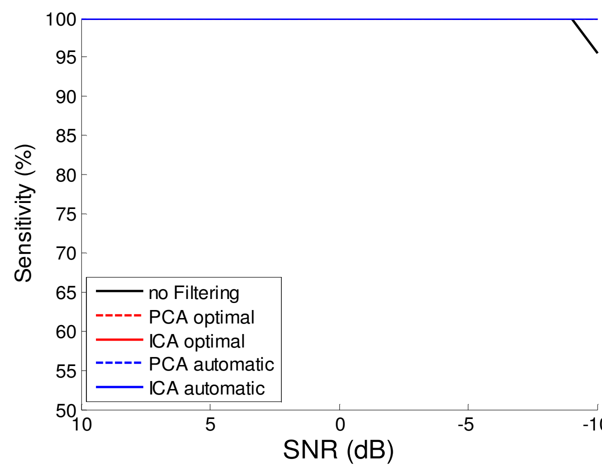

Figure 4. Sensitivity of beat detection after applying PCA and ICA and retaining the optimal and automatic component subset, and with no filtered.

Considering positive predictivity, the median +P was of 100% when no filtering was used for SNR inputs down to 3 dB. However, below this SNR, +P dropped considerably down to 57.43% when SNR=-10dB. Applying PCA and selecting the components automatically, gave a higher +P than not filtering for all SNR below 3 dB (+P=100% for SNR down to 0dB, SNR=-10dB, +P=58.82%). PCA with optimal component selection outperformed PCA with automatic selection for all SNR values below 0dB. ICA and automatic component selection gave a higher +P below 0dB as compared to PCA and automatic component selection (+P=100% for SNR down to -4dB, SNR=-10dB, +P=59.69%). Applying ICA and retaining the optimal component subset gave higher +P than both versions of PCA and automatic ICA for SNRs below -5dB. It is interesting to note that between -3 and -5dB, automatic ICA had a higher +P as compared to optimized ICA. This is due to the fact that the optimal component subset was identified by choosing the one that maximized the correlation coefficient between the clean and the filtered signals. The +P results for PCA and ICA with both automatic and optimal component selection are plotted in Figure 5.

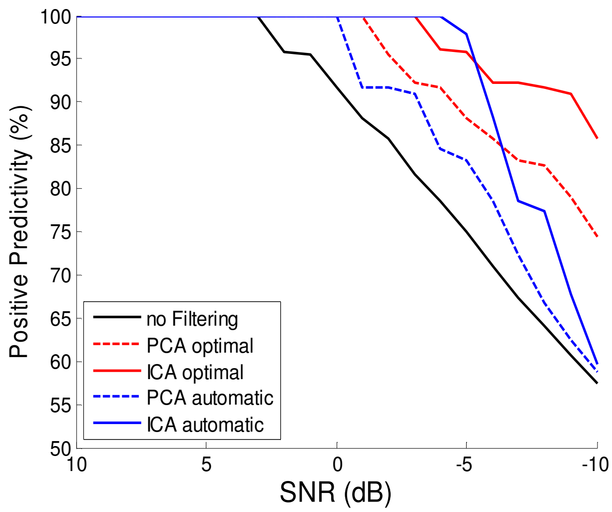

Figure 5. Positive predictivity of beat detection after applying PCA and ICA and retaining the optimal and automatic component subset, and with no filtered.

4. Conclusions

This work investigated the performance of PCA and ICA in denoising ECG signals recorded in ambulatory conditions. A simulated database formed by the combination of clean ECG signals with noise scaled to different levels of energy was developed for evaluation.

Regarding sensitivity, the beat detection algorithm gave a performance of 100% for SNR down to -6dB. Below this value, filtering with the use of PCA and ICA gave median Se of 100% for all SNR values considered in this study. Positive predictivity is more sensitive to noise and with no filtering had a 100% in median for SNR values down to 6dB. Below this value, +P dropped fast when SNR decreased. Applying PCA gave a higher +P than no filtering for any SNR value below 6dB. ICA gave equal +P than PCA for SNR down to 2dB and outperformed PCA for SNRs below that value.

An automatic method based on kurtosis for component selection was proposed. Filtering using this method yielded a higher performance in beat detection as compared to non-filtered signals, especially when the noise level was high. However, optimal selection of components yet obtained a higher performance. This suggests that other method for automatic component selection could lead to a better performance.

As a limitation of this study, it should be noted that some stationarity has been assumed as signals were of 10 seconds length. The performance under shorter duration noise was not studied. In addition, a beat detection algorithm was used in order to evaluate the signal quality. This method is very specific to a small part of the ECG signal (QRS complex) and does not give any information about the denoising performance in the remaining wave of the ECG.

References

[1] Afonso V, Tompkins W, Nguyen T, Michler K, Luo S. Comparing stress ECG enhancement algorithms. IEEE Eng Med Biol Mag 1996;15:37–44.

[2] Augustyniak P. Separating cardiac and muscular ECG components using adaptive modelling in time-frequency domain. Proc. of the WACBE World Congress on Bioengineering 2007.

[3] Devlin PH, Mark RG, Ketchum JW. Detection electrode motion noise in ecg signals by monitoring electrode impedance. Computers in Cardiology 1984; 51-56.

[4] Hamilton P, Curley M, Aimi R, Sae-Hau C. Comparison of methods for adaptive removal of motion artifact. Computers in Cardiology 2000; 27: 383– 386.

[5] Hamilton P, Curley M, Aimi R. Effect of adaptive motion-artifact reduction on QRS detection. Biomed Instrum Technol 2000; 34: 197–202.

[6] Tong D, Bartels K, Honeyager K. Adaptive reduction of motion artifact in the electrocardiogram. Proc. Second Joint EMBS/BMES Conf 2002; 2: 1403–1404.

[7] Raya M, Sison L. Adaptive noise cancelling of motion artifact in stress ECG signals using accelerometer. Proc. Second Joint EMBS/BMES Conf 2002; 2: 1756–1757.

[8] Hyvärinen A, Oja E. Independent component analysis: Algorithms and applications. Neural Networks 2000; 13: 411–430.

[9] Castells F, Cebrián A, Millet J. The role of independent component analysis in the signal processing of ECG recordings. Biomed Tech (Berl) 2007; 52(1): 18–24.

[10] Palaniappan R, Khoon TE. Uni-channel PCA for noise reduction from ECG signals. Proc 1st International Bioengineering Conference 2004; 436-439.

[11] Romero I, PCA-based Noise Reduction in Ambulatory ECGs. Computing in Cardiology 2010; 37:677–680

[12] Chawla MPS. A comparative analysis of principal component and independent component techniques for electrocardiograms. Neural Comput & Applic 2009; 18: 539-559.

[13] Jolliffe IT. Principal Component Analysis. Springer Series in Statistics 2002. New York: Springer.

[14] Romero I, Addison PS, Reed MJ, Grubb N, Clegg GR, Robertson CE, Watson JN, Continuous wavelet modulus maxima analysis of the electrocardiogram: Beat characterisation and beat-to-beat measurement. Int J Wavelets, Multiresolution Inf. Process 2005; 3(1): 19-42.

[15] Romero I, Component Selection for PCA-based Extraction of Atrial Fibrillation. Computing in Cardiology 2006; 33:137-140.

Address for correspondence.

Iñaki Romero.
High Tech Campus 31
5605 KN Eindhoven
The Netherlands
romero@imec-nl.nl