A flexible PCA-based ECG-reconstruction algorithm with confidence estimation for ECG during exercise

Steffen A Mann[1], Reinhold Orglmeister[1]

[1] TU Berlin, Berlin, Germany

Abstract

Holter monitoring of the electrical activity of the heart is the key for analysis of many intermittent cardiac irregularities. The measurements are however prone to movement artifacts or poor skin-electrode contact, often making some parts of the recorded signals useless since important information might be hidden by these artifacts. In this paper a special ECG sensor node with active electrodes is used to record the 12-channel ECG during movement. Several methods for detection of degraded signal quality are combined and adapted to identify leads that have been compromised on a beat-to-beat level. A PCA-based reconstruction algorithm is then used to reconstruct leads with insufficient signal quality using only undisturbed leads. Furthermore this approach enables an estimation of the quality of reconstruction even in absence of an undisturbed reference.

1. Introduction

24-hour ECG recordings are common practice when suspected heart diseases are to be investigated. These holter ECGs are recorded in everyday situations where movement artifacts or deteriorated skin-electrode contact can compromise parts of these signals. Leads can also be lost due to electrodes that fell off during recordings sometimes requiring patients to re-take these tests. The automatic identification of low signal quality and the reconstruction of corrupted or unrecorded ECG leads can vastly improve automated signal processing. This paper presents a custom designed body-sensor-network (BSN) with 12-channel ECG and active electrodes and modifies existing approaches for detection of low signal quality. A novel flexible PCA-based reconstruction algorithm [2] enables dynamic, situation dependent reconstruction of whatever leads are currently not available or disturbed. In addition this paper also focuses on even severe motion artefacts during exercise like jogging and climbing stairs. It also gives an estimation of the reliability of its reconstruction, even though no artefact-free reference signal can be recorded, e.g. during running. This is done using only prior knowledge of the reconstructed lead and leads that were undisturbed.

Section 2 gives a brief overview over the developed hardware (BSN, sensor node & active electrodes). In section 3 first the algorithm for signal quality evaluation will be described, then the PCA-based reconstruction followed by the reconstruction quality estimation. Section 4 presents the results and section 5 concludes with a discussion and outlook.

2. Hardware

The sensor node for the acquisition of the 12-channel ECG was designed as a part of a mobile body-sensor-network (BSN) with multiple synchronous operating biosensors. These nodes comprise a Bluetooth interface for synchronisation, control and data streaming, a SD-card for data storage, a MSP430 microcontroller as the heart of each sensor node and sensor specific components, like analog frontends. The key features for this BSN are synchronism of all sensors (20μs jitter between samples on different nodes) and low energy consumption for continuous 24 hour data-acquisition.

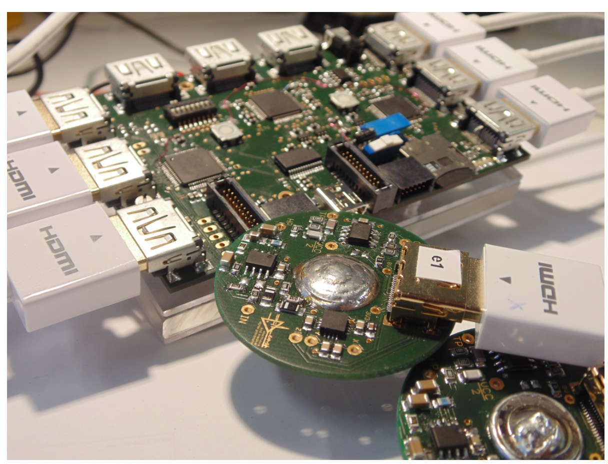

Figure 1. 12-channel ECG sensor-node with active electrodes and accelerometers on each electrode.

2.1. 12 - channel ECG - Sensor - Node

Apart from the common BSN components for synchronisation, storage and control the ECG sensor-node implements the 8-channel high-precision analog frontend ADS1298 from Texas Instruments for ECG acquisition. It offers programmable gain amplification of each channel and intern multiplexing for augmented leads, right-leg-drive while consuming only $0.75\frac{mW}{Channel}$. Further ADCs acquire accelerometer data from the active electrodes.

2.2. Active Electrodes

As the amplitudes of ECG leads are only several mV an early amplification is desirable. The electrodes used with the sensor node incorporate active filtering and amplification at the point of recording. They can be simply clipped to ordinary wet-gel ECG electrodes and possess additional 3-axes accelerometers for detection of movement.

3. Methods

The sensor described above was used to record 12-lead ECGs during everyday situations. These recordings include motion intensive activities such as climbing stairs and jogging. Beat-to-beat quality evaluation of ECG-leads was realised using the electrodes' accelerometric data together with selected ECG-signal features from the 'PhysioNet/CinC-Challenge-2011' (CinCC2011) [1]. As these features originally were designed to work on $10s$ segments they had to be adapted for beat-to-beat processing.
Reconstruction is achieved by transforming the first PCs of leads with high quality into those leads with lower quality using a reconstruction matrix **A**. **A** can be determined by principal component analysis (PCA) for different situations such as inhalation/exhalation, different heart-rates or even postures or activities, as classified from the electrodes acceleration data.
Lacking an undisturbed reference during exercise a confidence estimation was achieved by considering the reconstruction accuracy of the undisturbed leads in the same segment by means of the Mean Square Error (MSE) and Pearson Correlation, the overall segment quality as well as prior knowledge like R-wave amplitudes.
The following paragraphs will detail the features used for ECG-quality estimation, the PCA-based reconstruction and the final confidence estimation for the reconstructed leads.

3.1. Signal Quality Evaluatuion

In order to define the signal quality different approaches out of the CinCC2011 were analysed and adapted. The goal was to identify disturbances in ECG-signals like motion artefacts, insufficient skin-electrode contact as well as electric disturbances. From the many criteria used in the framework of this competition, four have been selected for closer analysis.

- **Overall range** between 0.2 and 15mV [7]
- **ratio of power** in frequency range 5-20Hz to total 0-62.5Hz [6]
- **self correlation**, correlation to averaged beat [5]
- **inter lead correlation**, correlation to other leads [5]

The above features were adapted to allow beat-to-beat analysis. Additionally the variance of the acquired accelerometer data was used as a further parameter for signal quality. The lead quality was evaluated on a beat-to-beat basis for all leads using a fixed length around all R-peaks detected using a QRS-detector based on filter banks.

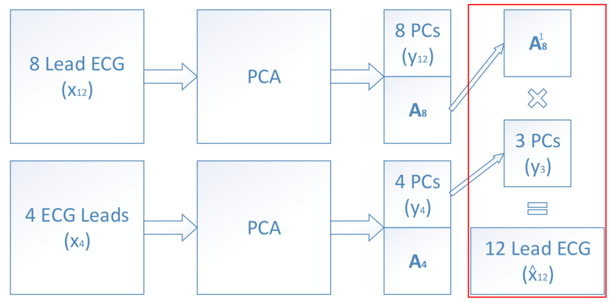

Figure 2. PCA based approach for ECG reconstruction.

3.2. ECG - Reconstruction

After disturbed P-QRS-T segments have been identified they have to be reconstructed. This is achieved by training a general or many situation-, posture- or activity-specific reconstruction matrices **A** using PCA. This section will give a short introduction to PCA, describe the selection of leads for training of **A** and elaborate on some special consideration when using PCA reconstruction.

3.2.1. Principal Component Analysis

Principal Component Analysis is a linear transformation

$$\mathbf{y} = \mathbf{A}\mathbf{x} \tag{1}$$

that transforms a set of input data **x** to reduce redundancy induced by correlation. This technique is often used for dimension reduction, feature extraction, data-decorrelation and whitening [4]. It can be described as the problem of finding an orthogonal transformation matrix **A** that decorrelates the elements of a centered vector **x**. The Transformation also maximises the variance of the principal components, which are the elements of **y**, and makes the covariance matrix of the transformed dataset **y** diagonal. The solution of maximizing the variance is given by the eigenvectors of the covariance matrix $\mathbf{C_x}$ of $\mathbf{x}$ [3]. The principal components were calculated using singular value decomposition.

3.2.2. Training

A linear transformation matrix $\mathbf{A_{Rec}}$ is used for reconstruction of disturbed segments. This matrix is obtained by PCA using all leads from undisturbed P-QRS-T segments or segments with minor disturbances. This training can be done either for the full P-QRS-T segments or certain parts of the ECG like the P- or T-wave or QRS interval. Different $\mathbf{A_{Rec}}$ can be used for different parts of P-QRS-T. As these segments have different origins within the heart this can increase the total reconstruction quality if a reliable ECG-segmentation can be achieved. Another possibility is the training for different postures and activities. As changes in posture lead to shifting of the body and the relative position of the electrodes the waveform and therefore also the shape of $\mathbf{A_{Rec}}$ can change. Training and reconstruction based on different posture and activity, derived from the accelerometer data, can increase reconstruction accuracy. A general or patient specific approach is possible.

3.2.3. Reconstruction approach

Whenever a lead of a P-QRS-T segment is found to have insufficient quality it is automatically reconstructed using other undisturbed leads in that same segment. First segments with high quality are selected and used to calculate the principal components. If too many leads have insufficient quality the reconstruction attempt is aborted. Otherwise the first principal components are used to calculate the reconstruction of the disturbed leads which are calculated by equation (2) using only the first n columns of $\mathbf{A_{Rec}^{-1}}$ and n PCs in $\mathbf{y}$.

$$\mathbf{x} = \mathbf{A_{Rec}^{-1}}\mathbf{y}. \tag{2}$$

This step assumes a high correlation between the recorded leads. Otherwise the PCs would change from their original shape and the reconstruction would be off. As the sign of the PCs can change the best reconstruction has to be found by changing the signs of every PC or some components of the ECG might be inverted. Also in segments where many leads are disturbed the order of PCs can change. Therefore all leads, the disturbed and undisturbed, are reconstructed, and the combination of PC signs and PC order is chosen, that gives the least mean square error for the undisturbed leads in the segment. The principal for the reconstruction is depicted in figure 2.

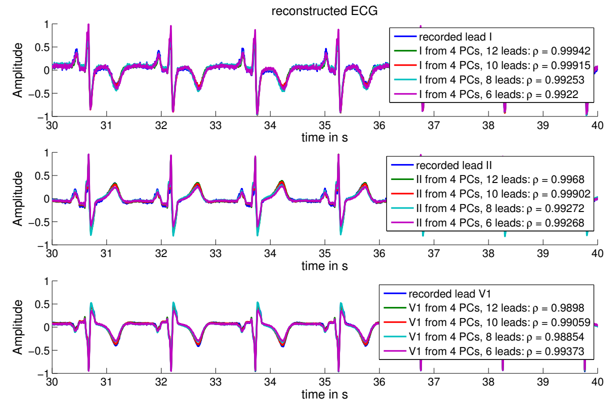

Figure 3. ECG reconstruction using different numbers of leads for reconstruction.

3.3. Reconstruction quality estimation

Common measures for reconstruction quality are mean square error (MSE) and correlation between reconstruction and original signal. As it is difficult or impossible to get a clean reference during exercise these measures can only be tested on artificially disturbed signals. On the other hand a high correlation or mean square error doesn't necessarily mean that all clinically relevant points are reconstructed correctly and that a diagnosis based on the reconstruction is equal to one based on the original signal. So on the one hand a substitute for the clean reference and on the other more clinically relevant measures have to be found. As a first step a confidence estimation was achieved by considering MSE and correlation for undisturbed leads and their reconstruction. The clinical equivalence was considered by calculating the mean and variance of the R-peak amplitude and comparing the deviance.

4. Results

In this study four hours of 12-channel ECG containing different activities including 96min of exercise and 36min of controlled breathing from six healthy subjects were evaluated, totalling in 17480 recorded heartbeats per lead. Table 1 gives an overview of the performance of the algorithm by showing the percentage of beats classified as disturbed, the percentage of disturbed beats that were reconstructed and the subsequent confidence estimation for different activities.

5. Conclusion

The algorithm is able to automatically detect disturbances in recorded ECGs on a beat-to-beat level. Leads with the lowest disturbances in a segment are automatically detected and used for reconstruction of all leads in that segment. The quality of the reconstruction can then

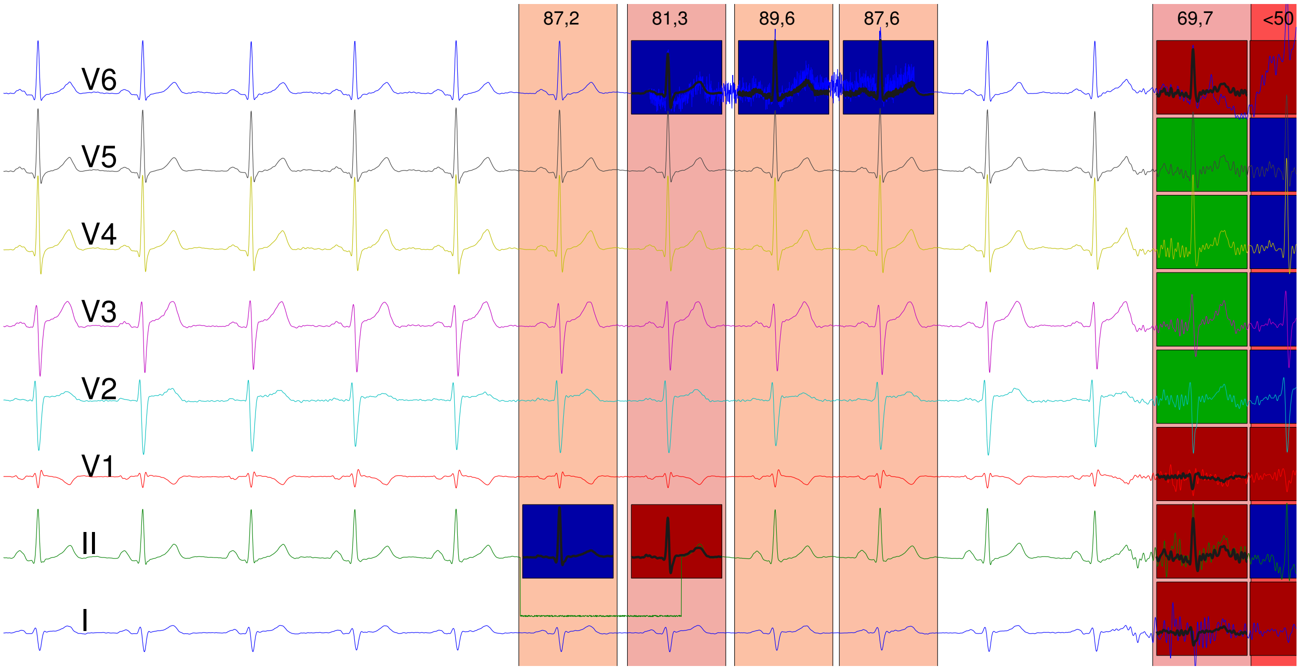

Figure 4. ECGs with induced noise (beats 8-10) and clipping (beats 7&8) and movement artefacts from running (last beats). Beats with sightly reduced quality are marked in green, medium quality in blue and low quality in red. The reconstructed signals are superimposed in black and the reconstruction confidence is noted on top of each reconstructed segment.

Table 1. Composition of test data and achieved results

Posture	Time (min)	percentage disturbed	percentage recon-structable	confi-dence
Rest	90	0.20%	95.0%	0.90
Controlled breathing	36	2.94%	100%	0.88
Selective distur-bances	18	39.85%	91.5%	0.87
Minor movement	36	13.54%	99.4%	0.73
Major movement	60	93.71%	76.3%	0.58

be assessed without the need for a clean reference. Good reconstruction results can be achieved as long as there are still multiple undisturbed leads available.

The reconstruction performance could be improved by posture and activity dependent reconstruction matrices using the accelerometer data for classification. Alternatively the reconstruction matrix could be trained for different parts of the ECG like Q- and T- wave or the QRS complex. More clinical measures for the confidence estimation should also improve the general evaluation of the reconstruction in absence of a clean reference.

References

[1] Silva I, et al: Improving the Quality of ECGs Collected Using Mobile Phones. The PhysioNet /Computing in Cardiology Challenge 2011. In: Comput. Cardiol. 38 (2011), S. 1273-76.

[2] Mann S, et al.: PCA-based ECG lead reconstruction. Biomedical Engineering / Biomedizinische Technik. ISSN (Online) 1862-278X, September 2013

[3] Hyvrinen A, Karhunen J, Oja E, Independent Component Analysis. Wiley, 2001

[4] Jolliffe I T, Principal Component Analysis, 2nd ed., Springer, 2002.

[5] Xia H, et al.: Computer Algorithms for Evaluating the Quality of ECGs in Real Time. Computing in Cardiology 2011;38:369372.

[6] Clifford GD, et al.:Signal Quality Indices and Data Fusion for Determining Acceptability of Electrocardiograms Collected in Noisy Ambulatory Environments. Computing in Cardiology 2011;38:285288.

[7] Moody B E, Rule-Based Methods for ECG Quality Control. Computing in Cardiology 2011;38:361363.

Address for correspondence:

Steffen A. Mann
Sekretariat EN 3
Einsteinufer 17
10587 Berlin
Steffen.Mann@tu-berlin.de