Robust artefact detection in long-term ECG recordings based on autocorrelation function similarity and percentile analysis*

Carolina Varon, *Member, IEEE*, Dries Testelmans, Bertien Buyse,
Johan A.K. Suykens *Senior Member, IEEE*, and Sabine Van Huffel, *Fellow, IEEE*

***Abstract*— Artefacts can pose a big problem in the analysis of electrocardiogram (ECG) signals. Even though methods exist to reduce the influence of these contaminants, they are not always robust. In this work a new algorithm based on easy-to-implement tools such as autocorrelation functions, graph theory and percentile analysis is proposed. This new methodology successfully detects corrupted segments in the signal, and it can be applied to real-life problems such as for example to sleep apnea classification.**

I. INTRODUCTION

The diagnostic capabilities of the electrocardiogram (ECG) can be reduced by the presence of artefacts in the signal. These contaminants can be caused by electrode motion, contact noise and electromyography among others [1]. Due to these undesired features, the output of different algorithms can be corrupted. For example when classifying segments of the signal, or deriving respiration from the ECG, an artefact can be interpreted as an anomaly. Therefore, in order to obtain accurate conclusions from any ECG analysis, it is very important to filter and eliminate any type of erroneous sections [2].

In principle this can easily be done when short signals are analysed, because the length of the measurements allows for manual inspection of the noise. However, in practice, long-term recordings of several patients are studied at the same time, which requires automated algorithms to detect and remove artefacts.

Different methodologies have been proposed to enhance the quality of the ECG, e.g. [3], [4] [5]. A disadvantage of those methods is that they do not provide information on the location of the corrupted sections. Therefore, even though the quality of the signal is enhanced, some artefacts are still present, and can have a negative influence on the analysis. An approach to find the specific location of disturbances is presented in [6], where the signal is evaluated with a moving window. One shortcoming of this method is the use of rules to identify contaminants. Furthermore, the method assumes that no abrupt alterations occur.

This paper addresses these problems, and proposes a new algorithm to identify the location of artefacts in long-term ECG recordings. The suggested methodology starts with characterizing each ECG segment by its autocorrelation function (ACF), and then uses graph theory [7] to identify the noise and contaminants. This is due to the fact that the ACF of a quasi periodic signal is significantly different from the one of a contaminated time series. The ACF has been used to analyse the properties of the ECG signal [8],[9], [10], and [11], but never to locate artefacts.

The rest of this document is organized as follows. Section II describes the data and elaborates more on the proposed algorithm. Results and discussion on the findings are presented in Section III, and section IV contains the conclusions of this work.

II. METHODOLOGY

A. Data

The dataset used in this study consists of 16 single lead ECG recordings extracted from polysomnographic measures of 16 different patients of the sleep laboratory at the University Hospital Leuven (UZ Leuven), Belgium. The signals were sampled at 200Hz and the recording time ranged from 4 up to 11 hours. In total 152 hours and 12 minutes were extracted. Each recording was manually annotated by a medical doctor who has experience with interpreting polysomnographic signals, including ECG. The annotations correspond to the occurrence of different types of sleep apnea events. This information is used in Section III to verify the algorithm when applied to the problem of detecting sleep apnea.

B. Segmentation and characterization of the ECG

The algorithm proposed in this paper was originally inspired by sleep apnea classification, where each signal is segmented and analysed on a minute-by-minute basis [12], [13],[14]. With this in mind, each ECG recording is first divided into segments of length N. Then, the signal is filtered by means of a band pass Butterworth filter with cut-off frequencies of 1Hz and 40Hz. The frequency bands are

*Research supported by Research Council KUL: GOA MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), PFV/10/002 (OPTEC), IDO 08/013 Autism, several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects: FWO G.0302.07 (SVM), G.0341.07 (Data fusion), G.0427.10N (Integrated EEG-fMRI), G.0108.11 (Compressed Sensing) G.0869.12N (Tumor imaging) research communities (ICCoS, ANMMM); IWT: TBM070713-Accelero, TBM070706-IOTA3, TBM080658-MRI (EEG-fMRI), PhD Grants; IBBT; Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011); ESA AO-PGPF-01, PRODEX (CardioControl) C4000103224; EU: RECAP 209G within INTERREG IVB NWE programme, EU HIP Trial FP7-HEALTH/ 2007-2013 (n 260777) (Neuromath (COST-BM0601); BIR&D Smart Care

C. Varon is with the KU Leuven, Department of Electrical Engineering ESAT, SCD-SISTA, and IBBT Future Health Department, Leuven, Belgium
e-mail: carolina.varon@esat.kuleuven.be

D. Testelmans and B. Buyse are with the UZ Leuven, Department of Pneumology , Leuven, Belgium

S. Van Huffel and J.A.K. Suykens are with the KU Leuven, Department of Electrical Engineering ESAT, SCD-SISTA, and IBBT Future Health Department, Belgium.

selected based on the low- and high-frequency limits of a diagnostic ECG [1], and on the fact that variations due to baseline drifts should not be considered.

After this preprocessing, the ACF of each ECG segment is obtained by computing the inverse Fourier transform of its corresponding power spectral density [15].

Note that the Fourier transform of a non-stationary signal is used, which normally should be avoided because of low frequency resolution. In other words, Fourier transform should not be used to study e.g. beat abnormalities or rhythm changes [1]. However, the goal of this research is to detect *artefacts* of the signal, rather than deviations from healthy ECGs. Therefore, in this particular application, the resolution of the Fourier transform is sufficient. This can be observed in Fig. 1, where the power spectra of two ECG segments are compared. The figure shows that a contaminated segment presents a broader spectrum when compared with a clean one. This behaviour is expected as the artefact should behave more like a random sequence with a broader frequency content. A similar assumption is made in [16], where motion artefacts are removed from photoplethysmographic data, using time and frequency analysis.

C. Weights Computation

After calculating the ACFs, the contaminated segments need to be localized. As mentioned in the introduction, when an ECG is interrupted by an artefact, a clear change in its autocorrelation is observed. This means that the ACFs that contain disturbances are "dissimilar" from the noise-free segments. To find these differences, a link to graph theory [7] is made, where the set of ACFs is represented on a mathematical structure known as a graph. Such a graph consists of vertices, i.e., the ACFs, that are pairwise connected by edges, i.e., their similarity. The length of these edges is determined by the inverse of a pairwise similarity between the ACFs, namely the cosine similarity which is quantified by

$$\cos\theta = \frac{A_1^{\mathrm{T}} A_2}{\parallel A_1 \parallel \parallel A_2 \parallel}, \tag{1}$$

where A_1 and A_2 are two vertices, θ is the angle between them, $\parallel . \parallel$ defines the two-norm and $A_1^{\mathrm{T}} A_2$ the dot product between the vectors. In this way, segments that are significantly different, i.e., artefacts, are represented by isolated vertices. Each of these vertices is then characterized by a lower *degree* value, which is computed by adding the pairwise similarities between that specific vertex, and the rest of the graph. When ECG segments are clean, their ACFs are similar, and the corresponding vertices get high degree values, i.e. they are strongly connected in the graph. This can be seen in Fig. 2, where two clean (left) and two contaminated minutes (right) are plotted together with their ACFs and their degree values.

This study proposes to use these degrees as "weights" to indicate how clean an ECG section is. In addition, the algorithm only retains the 95th percent of the most similar segments, which correspond to the vertices with the higher weights. In percentile analysis, the 95 percentile is an often used value to detect outliers. The proposed algorithm is summarized in Table I.

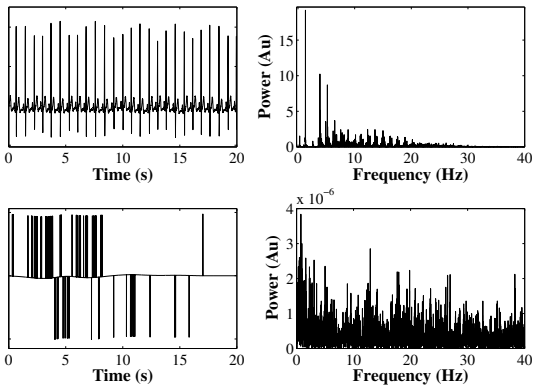


Fig. 1. ECG segments (left) and their corresponding power spectra (right). Note that the frequency content of an artefact is broader (bottom) compared to the one of a clean signal (top).

In the next section, the performance of the proposed algorithm is compared with two clustering methods, namely k-means [17] using the Euclidean distance, and kernel spectral clustering (KSC) [18] using a radial basis function.

III. RESULTS AND DISCUSSION

This section is divided in two parts, first the ability of the proposed method to identify artefacts in the signals is studied. Second, the algorithm is applied to a typical practical problem, namely a sleep apnea classifier.

A. Detection of artefacts

The total number of minutes in the dataset is 9132 and only 294 of them contain artefacts, which were manually labelled. This labelling allows to validate the method, and compare the performance with that of k-means and KSC for two cluster.

As mentioned in Table I, the ACF for each minute is computed. After that, the weights are calculated and the 95 percentile threshold is applied. This gives two groups: the normal segments and the contaminated ones. Fig. 3 shows the weights of the sections of one ECG signal, and the 95th

TABLE I
ARTEFACT DETECTION ALGORITHM

Input: ECG signal and length N. In this study we use $N = 60s$.
Output: Weight for each ECG segment and 95th percentile of the weights.
1: Segment the ECG into epochs of length N.
2: Compute the ACF of each segment using the inverse Fourier transform of its corresponding PSD.
3: Calculate the cosine similarity (1) between each pair of ACFs. This results in the similarity matrix between all ACFs.
4: Compute the degree values by summing all the columns of the similarity matrix. These degrees correspond to the weights.
5: Calculate the 95th percentile of the weights.
6: Label the segments with weights outside the 95th percentile, as contaminated.

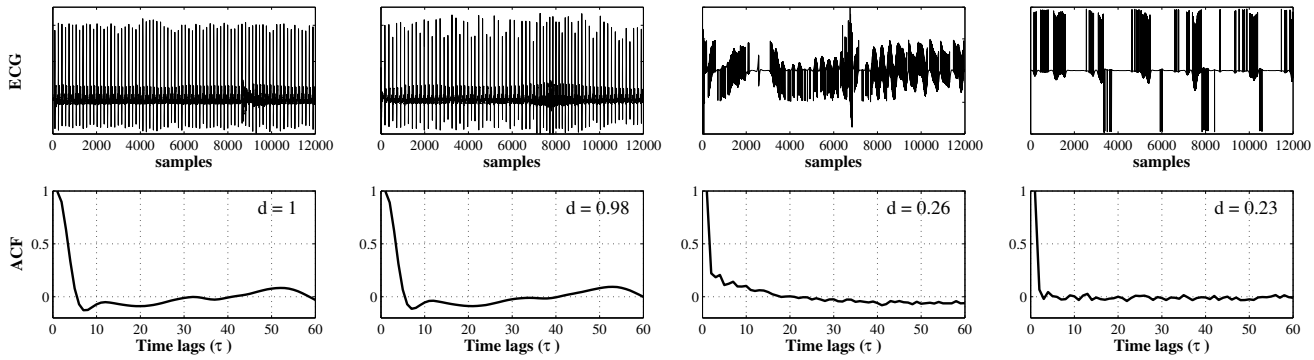


Fig. 2. Autocorrelation functions (bottom) of four different ECG segments of one minute (top). The first two (from left to right) correspond to normal minutes, while the other two to minutes with artefacts. The normalized degree values d, for the ACFs or vertices, are indicated. Note that the similarity between the contaminated segments and the rest is significantly low compared to the ones of normal minutes.

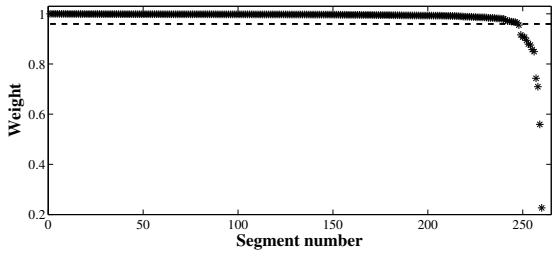


Fig. 3. Weights of one minute ECG segments of one patient. The dashed line indicates the 95th percentile. Note that the points below the threshold are clearly separated from the rest.

percentile. Note that the minutes with the lowest weights are clearly separated from the rest. Finally, the other two partition algorithms are also applied to the same set of ACFs.

Due to inter-subject variabilities in the ECG and ACF, the algorithms are applied to each ECG signal separately. To compare the performance of the algorithms, statistical measures often used in classification known as sensitivity, specificity, and accuracy are computed. These values are obtained for each algorithm and for each patient separately, and they are shown in Table II. It is observed that the proposed algorithm (95th) outperforms the other two, because accuracy, specificity and sensitivity are higher for all patients.

For completeness, all the one-minute segments are put together and the algorithms are applied to all the signals at the same time. The best results are again produced by the 95th percentile threshold and correspond to a sensitivity of 93.2% and specificity of 90.4%.

To determine if the duration of the segments affects the results, the experiments are repeated with N ranging from 5 to 60 seconds. The best results are obtained after analysing segments of 5 seconds, the lower bound of the epochs. This is no surprise, since as shorter segments are analysed, less information from normal ECG is included in the contaminated signals. In other words, when the selected epochs contain, proportionally to the window length, more artefact than normal ECG, its ACF differs more from that of a normal signal. Fig. 4 shows a one minute ECG segmented into epochs of 5 seconds, i.e., 1000 samples. These results indicate that when the window is reduced the resolution of the algorithm is improved. The ACFs of the epochs are also indicated in the figure, where it is possible to observe the differences between corrupted and clean signals.

There are reasons why the clustering algorithms do not perform as good as the 95th percentile threshold. On one hand, it is important to keep in mind that the problem deals with a highly unbalanced partition, since in most of the cases, the artefacts constitute a lower percentage of the signal. In other words, the number of segments in the "normal" group must be significantly higher than the amount of segments in the "contaminated" group. KSC uses a model selection criterion that takes into account these unbalanced problems [18]. However, special attention must be paid to the definition of the parameter that determines how balanced the clusters must be. A bad selection of this parameter can produce misleading results.

On the other hand, the k-means clustering finds globular structures in the data. When such structures are not present or are not clearly separable, this technique tends to perform poorly. In addition, this algorithm tends to converge to local minima, which is one of the well known disadvantages of k-means.

B. Sleep apnea classification

The algorithm is now used in a real-life problem. In [19], a method to classify sleep apnea is proposed and it makes use

TABLE II

COMPARISON OF THE PERFORMANCE OF THE PROPOSED ALGORITHM WITH k-MEANS AND KSC. MEAN VALUES AND STANDARD DEVIATIONS ARE INDICATED.

	KSC	k-means	95th
Sensitivity	0.92 ± 0.11	0.89 ± 0.13	**0.96 ± 0.05**
Specificity	0.75 ± 0.15	0.82 ± 0.13	**0.90 ± 0.03**
Accuracy	0.75 ± 0.14	0.82 ± 0.13	**0.90 ± 0.03**

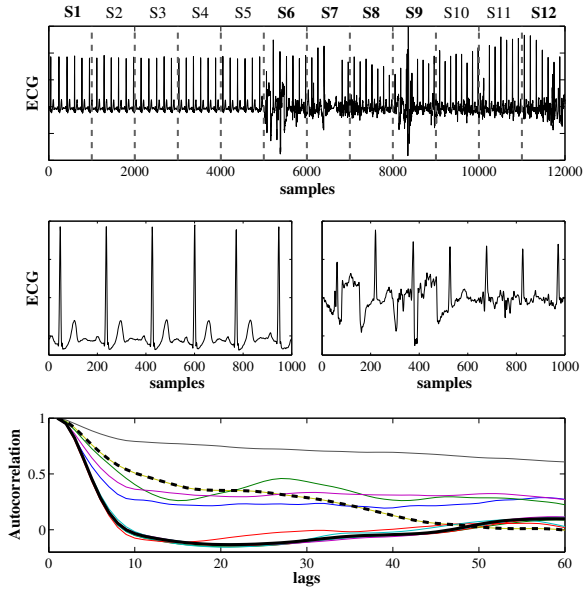


Fig. 4. One minute ECG segmented into epochs of 5 seconds (top). Note that segments S6, S7, S8, S9 and S12 are contaminated. Segments S1 (middle left) and S6 (middle right) are indicated. Note that the ACFs (bottom) are significantly different between S1 (tick solid line) and S6 (dashed line). The other curves (bottom) correspond to the ACFs of the other segments.

of a training and a validation set. However when artefacts are present in the ECG, these sets are influenced and the overall performance of the classifier is affected. This can be avoided by first applying the proposed methodology and then presenting only the clean ECG minutes to the classification algorithm.

The performances of the sleep apnea classifier on a validation set is evaluated before and after detecting the minutes with artefacts. The sensitivity and specificity of the classifier before applying the algorithm are 85.9% and 86.7% respectively. After removing the contaminated segments from the training and validation sets, the sensitivity and specificity are then 93.2% and 92.1% respectively. Hereby the ground truth provided by the clinician is used. Again, it is observed that the performance is improved when the training set is selected from the minutes with degrees within the 95th percentile.

IV. CONCLUSION

A new algorithm to detect artefacts by means of the autocorrelation function, in the ECG was presented. Moreover, a weighting methodology to distinguish normal from contaminated ECG signals is included, which outperforms other methods such as k-means and kernel spectral clustering.

The most important contribution of this study is that it is easy to implement and interpret. This is due to the fact that it uses well-known tools such as the autocorrelation function and percentile analysis. Furthermore, the algorithm does not require computationally expensive operations, e.g. eigenvalue decompositions.

Another advantage of the algorithm is that it allows to locate the contaminated segments in the ECG. This information is not provided by algorithms that enhance the overall quality of the signal. In addition, other artefact detection algorithms require more ECG channels or extra information, to identify and remove artefacts. This is not the case for the methodology presented in this paper, since it only requires a long term ECG signal.

A disadvantage of this method is that it can only be applied off-line. It is necessary to have the complete ECG recording before the artefacts can be detected.

REFERENCES

[1] G. D. Clifford, F. Azuaje and P. E. McSharry, Advanced Methods and Tools for ECG Data Analysis. Boston: Artech House, 2006, pp. 57.
[2] G. M. Friesen, T. C. Jannett, M. A. Jadallah, S. L. Yates, S. R. Quint and H. T. Nagle, A Comparison of the Noise Sensitivity of Nine QRS Detection Algorithms. IEEE Trans Biomedical Engineering, vol. 37, pp. 85-98, 1990.
[3] D. Benitez, P. A. Gaydecki, A. Zaidi and A. P. Fitzpatrick, The use of the Hilbert transform in ECG signal analysis. Computers in Biology and Medicine, vol. 31, pp. 399-406, 2001.
[4] K. Sharmila, E. H. Krishna, K. N. Reddy, and K. A. Reddy, Application of multiscale principal component analysis (MSPCA) for enhancement of ECG signals. Instrumentation and Measurement Technology Conference IEEE, 2011.
[5] R. Sameni, M. B. Shamsollahi, C. Jutten, and G. D. Clifford, A Nonlinear Bayesian Filtering Framework for ECG Denoising. IEEE Transactions on Biomedical Engineering, vol. 54, pp. 2172 - 2185, 2007.
[6] S. Bras, N. Ferreira and J. P. Silva, ECG Artefact Detection Algorithm: An Algorithm to Improve Long-Term ECG Analysis. Proceedings of the International Conference on Bio-Inspired systems and Signal Processing BIOSIGNALS, pp. 329-333, 2012.
[7] F.R.K. Chung, Spectral Graph Theory. Am. Math. Soc., 1997, ch. 1.
[8] B. Kulessa, T. Srokowski, S. Drozdz, Long-Time Autocorrelation Function of ECG Signal for Healthy versus Diseased Human Heart. Acta Physica Polonica B, vol. 34, pp. 3, 2003.
[9] T. Syeda-Mahmood, D. Beymer, and W. Fei, Shape-based Matching of ECG Recordings. IEEE EMBS, pp. 2012-2018, 2007.
[10] Z. Piotrowskia and K. Rózanowski, Robust Algorithm for Heart Rate (HR) Detection and Heart Rate Variability (HRV) Estimation. Acta Physica Polonica A, vol. 118, pp. 131-135, 2010.
[11] R. D. Throne, Detecting ventricular fibrillation using efficient techniques for computing a normalized autocorrelation. Computers in Biology and Medicine, vol. 23, pp. 317-325, 1993.
[12] C. W. Zywietz, V. von Einem, B. Widiger, G. Joseph, ECG Analysis for Sleep Apnea Detection. Methods Inf. Med., vol. 43, pp. 56-59, 2004.
[13] T. Penzel, J. McNames, P. de Chazal, B. Raymond, A. Murray and G. Moody, Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings. Medical & Biological Engineering & Computing, vol. 40, pp. 402-407, 2002.
[14] C. Maier, H. Wenz and H. Dickhaus, Steps toward subject-specific classification in ECG-based detection of sleep apnea. Physiological Measurement, vol. 32, pp. 1807-1819, 2011.
[15] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, Time Series Analysis: Forecasting and Control. 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1994, pp. 30-34.
[16] R. Krishnan, B. Natarajan, and S. Warren, Two-Stage Approach for Detection and Reduction of Motion Artifacts in Photoplethysmographic Data. IEEE Transactions on Biomedical Engineering, vol. 57, pp. 1867-1876, 2010.
[17] J. B. MacQueen, Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, pp. 281-297, 1967.
[18] C. Alzate and J. A. K. Suykens, Multiway Spectral Clustering with Out-of-Sample Extensions through Weighted Kernel PCA. IEEE Trans Pattern Analysis and Machine Intelligence, vol. 32, pp. 335-347, 2010.
[19] M. Bsoul, H. Minn, and L. Tamil, Apnea MedAssist: Real-time Sleep Apnea Monitor Using Single-Lead ECG. IEEE Trans Inf Technol Biomed, vol. 15, pp. 416-427, 2011.