Data Quality Assessment of Capacitively-Coupled ECG Signals

Ivan Castro[1,2], Carolina Varon[1], Jonathan Moeyersons[1], Amalia Villa Gomez[1], John Morales[1], Margot Deviaene[1], Tom Torfs[2], Sabine Van Huffel[1,2], Robert Puers[1,2], Chris Van Hoof[1,2]

[1] KU Leuven, Leuven, Belgium
[2] imec, Leuven, Belgium

Abstract

Acquisition of capacitively-coupled ECG (ccECG) from daily life scenarios is limited by its sensitivity to motion and its variability in signal quality. 48 features, in combination with different classifiers, were evaluated for quality classification on a dataset of 10000 ccECG segments of 15 seconds. Feature subsets with potential high discriminatory power were identified and evaluated in multiple supervised models, for two classification problems with different tolerance to artefacts. This resulted in balanced accuracies of 94.02% and 92.4%, achieved using a Linear SVM and a fine KNN respectively. These models are useful tools for real-time and offline processing of ccECG signals recorded in real-life scenarios

1. Introduction

Long-term electrocardiography (ECG) recordings from real-life environments have been an important focus of recent research [1-4]. Capacitively-coupled ECG (ccECG) has been demonstrated as a technology that has the potential to achieve recordings from real-life environments and enable unobtrusive health monitoring [1,2], thereby improving the quality of life of people and lowering healthcare costs (e.g. by early diagnosis and timely treatment).

Despite the advantages of ccECG, it is highly susceptible to motion artefacts (MAs) [1–3], which are particularly problematic in recordings from real-life environments (e.g. while driving, while sleeping). This leads to a wide range of signal quality [4] and limits most of its use to experiments in controlled conditions. A promising approach to increase the robustness of extracted cardiac information is the use of signal quality indicators (SQIs) and quality-based classification models (CMs). These CMs enable MA handling methods, such as offline post-processing and real-time hardware adaptation methods [5,6], to be applied in real-life scenarios.

Since artefacts and noise in ccECG can be different than in contact ECG [3], conventional ECG SQIs may not be sufficient as an indicator of ccECG quality. Furthermore, the required signal quality depends on the specific intended use or application.

Some work on ccECG signal classification has been published in recent years. This includes the use of 'filter masks' to identify saturation, high frequency content and low signal power, with a reported balanced accuracy (BA) of 64.5% (44% sensitivity & 85% specificity) when evaluated in laboratory conditions [7]. Another approach [8] included a logistic regression model using pressure signals and an evaluation of signal saturation, which resulted in overall BA of 88.5% (93% sensitivity & 84% specificity) when evaluated in an airplane seat setting. An evaluation of driving monitoring [9] reported different quality-based classification algorithms. This resulted in a best-performing result with BA of 75% (54.7% sensitivity & 95.3% specificity).

In this work, the use of 48 SQI features was evaluated in different classification models. Different feature subsets were fed into multiple supervised models, and their BA was obtained as a performance metric.

2. Methods

2.1. Dataset

ccECG signals from diverse scenarios were used. These signals included data recorded from a system described in [5,10] as well as from the publicly available UnoVis dataset [11]. The data comprised 10000 randomly selected ccECG segments of 15 seconds, resulting in the distribution shown in Table 1. For each scenario data with floating sensors (e.g. no user present) was included, to allow a classification in this scenario. Because of this, the quality distribution of the data does not represent the signal coverage (i.e. percentage of good signal evaluated when user present) that is expected in practice. A coverage evaluation (e.g. in real-life in different scenarios) is out of the scope of this work.

 ISSN: 2325-887X DOI: 10.22489/CinC.2019.376

Table 1. Overview of all ccECG segments included in the evaluation of quality-based classification models.

Data source	Data type	Number of segments
Data recorded from system in (5,10)	Static car seat	2500
	Bed form factor	2500
	Office chair, normal working conditions	1520
	While driving a car	480
UnoVis database (11)	While driving a car	1000
	Bed form factor	1000
	Armchair form factor, induced MAs	1000

Five annotators with experience in ECG signal processing labelled the segments using three quality levels: 1. Useless or no ccECG; 2. ccECG with artefacts that may affect the detection of 2 to 5 heartbeats; 3. ccECG useful for heart rate variability (HRV) analysis and possibly morphology analysis. In total, 90 segments with strong annotation disagreement (i.e. labelled as both 1 and 3 by different annotators) were discarded, resulting in a 9910-segment dataset. The remaining segments had an agreement of at least 3 annotators and a Fleiss' Kappa of 0.80.

2.2. Feature selection & classification

The three quality levels were assigned to two binary classification problems: one with a 'low threshold' (level 1 vs level 2-3) -classifL- and another with a 'high threshold' (level 1-2 vs level 3) -classifH- (Table 2). This division allows to evaluate classifications in which signals with moderate artefacts are still considered useful as well as a stricter classification that only considers level 3 signals as useful. Different scenarios can benefit from these classifications (e.g. for -classifL- problem for HR and HRV extraction, -classifH- problem for morphology analysis). Each dataset was randomly divided in 70% training and 30% test, preserving the binary distribution ratio.

48 SQI features were extracted from each ccECG segment, including the features evaluated in [10]. Feature selection (FS) was performed on the training set by means of: 1. neigborhood component analysis (NCA) [12] available in the machine learning toolbox of Matlab®; 2. Random Forest (RF) classification as proposed in [13]; and 3. threshold-based one-level decision tree (DT) classification performance. The classification performance of different feature subsets identified by the FS methods was evaluated. This was done by training and validating 19 different supervised classifiers for both binary classification problems (i.e. -classifL- and -classifH-).

Table 2. Overview of the quality distribution for the two problems.

Classification problem	Binary quality grouping	Distribution (Bad vs Good)
'classifH'	{1,2} (Bad) vs 3 (Good)	80.5% vs 19.5%
'classifL'	1 (Bad) vs {2,3} (Good)	65.8% vs 34.2%

An important part of the data included floating sensors, which caused the class distribution to be unbalanced. Therefore, all the model training was done by setting a prior probability distribution of the classes as uniform (i.e. balanced distribution), so that the classification methods compensate for the dataset imbalance. In addition, the metric BA [14] was used to compare the models. This avoids an overestimation of the classification performance.

3. Results and discussion

For each problem, a feature subset was obtained from each of the three FS methods. Table 3 shows the best-performing SQI feature subset from each method for each problem, together with the best performing classifier and its corresponding BA. A brief description of the selected features is presented below.

corrSQI: Average correlation of the beats with a template extracted from the signal. The metric is obtained by averaging (corrSQImean) or obtaining the average without outliers (corrSQItrimmedmean), or the median (corrSQImedian) of QRST complexes from the ccECG window and computing the average of the individual correlations of the template with each of the beats. This was done by using the beat detector from [15]. This SQI was individually evaluated for ccECG by the authors [10] and previously proposed for contact signals [16] .
bSQI: Comparison of the beat detections from two algorithms ([17] and [18]). It is based on calculating the agreement rate of these. More details can be found in [19] and [20]. An initial evaluation of this metric by the authors on ccECG can be found in [10].
SDR: Ratio between power spectral density of band of interest and a broader band. (i.e. [5-14] Hz and [0-50] Hz). This was initially used in [19] for contact ECG for different limits, and evaluated by the authors for ccECG in [10].
msSQI: Modulation spectrum metric originally proposed in [21]. It consists of the windowed calculation of the frequency spectrum of the signal, followed by the spectrum of the spectral magnitudes. This results in a frequency-frequency representation used to extract the modulation energy of the signal. Details for its calculation can be found in [21].
bkSQI: Kurtosis-based metric using experimentally determined Kurtosis ranges of the mean Kurtosis from

each QRS beat. This metric receives a value of 1 when the per-beat Kurtosis is in the range (4.4-21) and a value of 0.5 for Kurtosis in the range (3.8-4.4) and (21-40). Other Kurtosis values are fixed to 0.3.

sKurt: Kurtosis calculated for each 15-second sement.

bSkewMod: Skewness-based metric. It replaces too-high values of the averaged Skewness calculated for each extracted QRS beat using experimentally determined limits. Average per-beat Skewness is kept for values lower than 3.5. A value of 0 is assigned to the metric for higher Skewness.

VrmsSQIper: Percentage of sub-windows in the (0.005 – 0.4) mVrms range. Sub-windows are 750 ms wide.

SD_b2b: Standard deviation of beat-to-beat HR values (extracted from the RR intervals).

SD_QSw: Standard deviation of Q-S durations measured in ms.

MedianAD_QSw: Median Absolute Deviation of Q-S durations measured in ms. Calculated from the beats in the window.

MeanAD_QSw: Mean Absolute Deviation of Q-S durations measured in ms. Calculated from the beats in the window.

MedianAD_QRd: Median Absolute Deviation of Q-R distances (Q-R trace) from the beats in the window.

It can be seen from Table 3 that the FS methods partially agreed on the selected features. Specifically, the corrSQI appears in all the feature subsets. This is in agreement with previous work [10], which concluded that this SQI has the highest performance when used as a stand-alone ccECG SQI.

Table 3. List of best-performing feature subsets for both datasets, with the corresponding classifiers and BAs.

Problem	Method	SQI Features	Classifier (BA)
'classifH'	NCA	{corrSQItrimmedmean, bSQI, SDR2, msSQI, bkSQI, SD_QSw}	Coarse Gaussian SVM (93.69%)
	RF	{corrSQItrimmedmean, SDR2, sKurt}	Linear Discriminant (93.71%)
	DT	{corrSQImean, VrmsSQIper, SD_b2b}	Linear SVM (94.02%)
'classifL'	NCA	{SD_b2b, bSQI, MedianAD_QSw, bkSQI, corrSQImean, corrSQImedian, sKurt, bSkewMod, MeanAD_QSw}	Fine KNN (92.4%)
	RF	{corrSQImedian, bSQI, SD_b2b, VrmsSQIper, MedianAD_QRd}	Fine KNN (90.84%)
	DT	{corrSQImedian, VrmsSQIper, SD_b2b}	RUS-Boosted Trees (91.74%)

The results of the best 5 classifiers (with BA > 80%) for each of the feature subsets are shown for the 'classifH' and 'classifL' problems in Figure 1 and Figure 2 respectively. In addition, the classification performance when using all the 48 features is included for reference purposes.

The resulting CMs achieved a maximum BA of 94.02% (95.19% sensitivity & 92.85% specificity) –for the 'classifH' problem, with a linear Support Vector Machine (SVM)-, and 92.4% (89.97% sensitivity & 94.84% specificity) – for the 'classifL' problem, with a fine K-nearest neighbors (KNN) classifier-. These accuracies are higher than previously reported ccECG classification

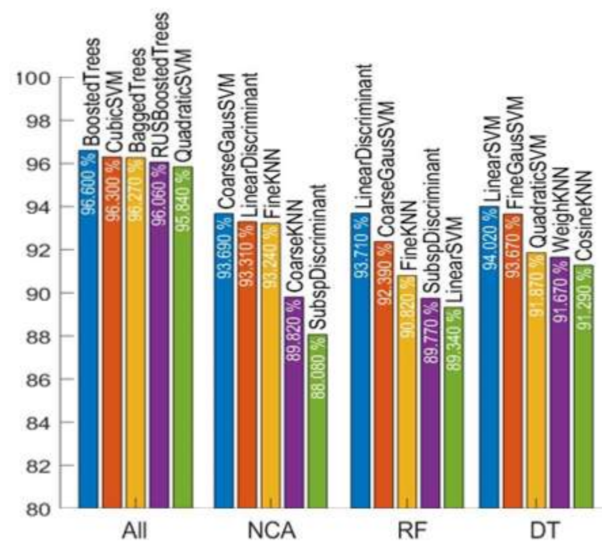

Figure 1. Results of the best 5 classifiers for each feature subset, for the 'classifH' problem.

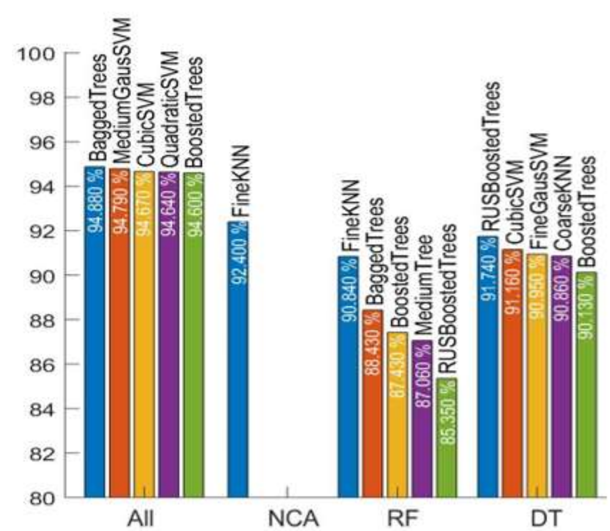

Figure 2. Results of the best 5 classifiers for each feature subset, for the 'classifL' problem. For NCA, only the FineKNN resulted in a BA higher than 80%.

literature mentioned in the introduction section, with max. BA of 88.5% [8].

The distinction of two classification problems allows to not only identify clean signals, but also signals with MAs that still contain ECG information, which is a more challenging classification problem.

Although classifiers using all the 48 features had slightly higher BAs than the presented CMs after FS, the latter allow to perform this classification with a reduced set of features. This significantly lowers the computational complexity, while keeping high BAs. Low-complexity CMs are useful in real-time artefact handling approaches and allow for fast post-processing approaches to improve the extracted information from unobtrusive, ubiquitous ECG monitoring.

4. Conclusions

This work presented CMs with high BA to be used in the automatic classification of ccECG signals from real-life environments. It was found that a DT-based feature subset with a linear SVM performs best for a 'classifH' problem, while an NCA-based subset with a KNN classifier performs best for a 'classifL' problem.

This type of classification is relevant not only as a post-processing tool, but also for real-time hardware adaptation approaches such as the modification of hardware settings [5] or the selection of electrodes from high-density arrays [6]. These tools are expected to result in increased coverage when acquiring signals from real-life scenarios and a reduction in the error of specific features of interest such as heart rate and heart rate variability. High-performance classification models and SQIs such as the ones presented in this work are key to enabling the use of ccECG collected from daily life, in order to allow health monitoring and long-term follow-up of patients.

Fine tuning of the classification cost of the models depending on the specific applications, and application-driven evaluations are necessary to further confirm the usefulness of these classification tools.

Acknowledgments

The authors would like to thank the creators of the UnoVis database, and the support from: imec funds 2017-2018, Agentshcap Innoveren en Ondernemen (VLAIO) 150466: OSA+ and O&O HBC 2016 0184 eWatch. Carolina Varon is a postdoctoral fellow from FWO Flanders.

References

[1] Chi et al. Dry-Contact and Noncontact Biopotential Electrodes: Methodological Review. IEEE Rev Biomed Eng. 2010;3:106–19.

[2] Lim YG, et al. Capacitive Measurement of ECG for Ubiquitous Healthcare. Ann Biomed Eng. 2014 Nov 23;42(11):2218–27.

[3] Ottenbacher et al. Motion Artefacts in Capacitively Coupled ECG Electrodes. IFMBE Proc. 2009. p. 1059–62.

[4] Czaplik et al.. The Reliability and Accuracy of a Noncontact Electrocardiograph System for Screening Purposes. Anesth Analg. 2012 Feb;114(2):322–7.

[5] Castro et al. Robust Wireless Capacitive ECG System with Adaptive Signal Quality and Motion Artifact Reduction. In: 2016 IEEE Int. Symp. on Medical Meas. & Appl. (MeMeA). IEEE; 2016. p. 1–6.

[6] Castro et al. Capacitive Multi-Electrode Array with Real-Time Electrode Selection for Unobtrusive ECG & BIOZ Monitoring. In: EMBC. Berlin, Germany; 2019.

[7] Eilebrecht et al.. A Capacitive ECG Array with Visual Patient Feedback. In: 2010 Annual Int. Conf. of the IEEE, EMBC. IEEE; 2010. p. 6539–42.

[8] Schumm et al. Automatic Signal Appraisal for Unobtrusive ECG Measurements. Int J Bioelectromagn. 2010;12(4):158–63.

[9] Wartzek et al. ECG on the Road: Robust and Unobtrusive Estimation of Heart Rate. IEEE Trans Biomed Eng. 2011 Nov;58(11):3112–20.

[10] Castro et al. Evaluation of a Multichannel Non-Contact ECG System and Signal Quality Algorithms for Sleep Apnea Detection and Monitoring. Sensors (Switzerland). 2018;18(2):1–20.

[11] Wartzek et al. UnoViS: The MedIT Public Unobtrusive Vital Signs Database. Heal Inf Sci Syst. 2015;3(2):1–9.

[12] Yang et al. Neighborhood Component Feature Selection for High-Dimensional Data. J Comput. 2012;7(1):162–8.

[13] Deviaene et al. Feature Selection Algorithm Based on Random Forest Applied to Sleep Apnea Detection. . In: EMBC. Berlin, Germany; 2019.

[14] Brodersen et al. The Balanced Accuracy and its Posterior Distribution. Proc Int Conf Pattern Recognit. 2010;3121–4.

[15] Romero et al. Robust Beat Detector for Ambulatory Cardiac Monitoring. In: 2009 Ann. Int. Conf. of the IEEE, EMBC. IEEE; 2009. p. 950–3.

[16] Orphanidou et al. Signal-Quality Indices for the Electrocardiogram and Photoplethysmogram: Derivation and Applications to Wireless Monitoring. IEEE J Biomed Heal Informatics. 2015;19(3):832–8.

[17] Hamilton et al. Quantitative Investigation of QRS Detection Rules Using the MIT/BIH Arrhythmia Database. IEEE Trans Biomed Eng. 1986 Dec;BME-33(12):1157–65.

[18] Zong et al. A Robust Open-Source Algorithm to Detect Onset and Duration of QRS Complexes. In: Computers in Cardiology, 2003. IEEE; 2003. p. 737–40.

[19] Li et al. Robust Heart Rate Estimation from Multiple Asynchronous Noisy Sources Using Signal Quality Indices and a Kalman Filter. Phys. Meas. 2008 Jan 1;29(1):15–32.

[20] Clifford et al. Signal Quality Indices and Data Fusion for Determining Clinical Acceptability of Electrocardiograms. Physiol Meas. 2012 Sep 1;33(9):1419–33.

[21] Tobon et al. Online ECG Quality Assessment for Context-Aware Wireless Body Area Networks. Can Conf Electr Comput Eng. 2015;2015-June(June):587–92.

Address for correspondence:
Ivan D. Castro
imec: Kapeldreef 75, 3001 Leuven, Belgium
ivand.castro@imec.be