

Wearable Armband Device for Daily Life Electrocardiogram Monitoring

Jesús Lázaro, *Member, IEEE*, Natasa Reljin, *Member, IEEE*, Md-Billal Hossain , *Member, IEEE*, Yeonsik Noh , *Member, IEEE*, Pablo Laguna , *Fellow, IEEE*, and Ki H. Chon , *Senior Member, IEEE*

Abstract—A wearable armband electrocardiogram (ECG) monitor has been used for daily life monitoring. The armband records three ECG channels, one electromyogram (EMG) channel, and tri-axial accelerometer signals. Contrary to conventional Holter monitors, the armband-based ECG device is convenient for long-term daily life monitoring because it uses no obstructive leads and has dry electrodes (no hydrogels), which do not cause skin irritation even after a few days. Principal component analysis (PCA) and normalized least mean squares (NLMS) adaptive filtering were used to reduce the EMG noise from the ECG channels. An artifact detector and an optimal channel selector were developed based on a support vector machine (SVM) classifier with a radial basis function (RBF) kernel using features that are related to the ECG signal quality. Mean HR was estimated from the 24-hour armband recordings from 16 volunteers in segments of 10 seconds each. In addition, four classical HR variability (HRV) parameters (SDNN, RMSSD, and powers at low and high frequency bands) were computed. For comparison purposes, the same parameters were estimated also for data from a commercial Holter monitor. The armband provided usable data (difference less than 10% from Holter-estimated mean HR) during 75.25%/11.02% (inter-subject median/interquartile range) of segments when the user was not in bed, and during 98.49%/0.79% of the bed segments. The automatic artifact detector found 53.85%/17.09% of the data to be usable during the non-bed time, and 95.00%/2.35% to be usable

during the time in bed. The HRV analysis obtained a relative error with respect to the Holter data not higher than 1.37% (inter-subject median/interquartile range). Although further studies have to be conducted for specific applications, results suggest that the armband device has a good potential for daily life HR monitoring, especially for applications such as arrhythmia or seizure detection, stress assessment, or sleep studies.

Index Terms—Wearable devices, electrocardiogram (ECG), ECG denoising, electromyogram (EMG), artifact detection.

I. INTRODUCTION

ELECTROCARDIOGRAM is the basis for the diagnosis of most cardiac arrhythmias and other cardiac pathologies. Many of these pathologies produce paroxysmal symptoms in the ECG, *e.g.*, atrial fibrillation [1], which is associated with increased mortality and morbidity [2]. Therefore, continuous long-term ECG monitoring is desirable for such arrhythmia applications. Furthermore, ECG monitoring allows the measurement of the heart rate (HR) variability (HRV) which remains a powerful tool for autonomic nervous system (ANS) assessment [3]. This further expands the range of potential applications of long-term ECG monitoring, including epileptic seizure detection [4], stress assessment [5], and sleep studies [6], among others, which rely only on QRS detection.

Another interesting technology for long-term HR and HRV monitoring is the pulse photoplethysmographic (PPG) signal. The PPG has been receiving a lot of attention lately because it can be measured on the wrist by smartwatches, making it very convenient for daily life monitoring. However, the PPG signal is highly vulnerable to artifacts, and many data points have to be discarded [7]. Different studies using PPG in different settings report different amounts of usable data, including 14.76% [8], 24% [9], 25% [10], and 56% [11].

The only wearable continuous ECG monitoring options are Holter/event monitors and the more recently-developed patch devices. Holter and event monitors have some disadvantages, including that they are cumbersome devices with obtrusive leads, and they use hydrogel-based electrodes, which often lead to skin irritation [12] due to the use of hydrogel and adhesives needed to fix the position of the electrodes [13]. This makes Holter monitors usable only for short-term monitoring (<2 weeks). A patch monitoring device eliminates the electrode leads but still requires hydrogels, which often cause skin irritation since subjects wear

Manuscript received March 17, 2020; accepted April 10, 2020. Date of publication April 14, 2020; date of current version November 20, 2020. This work was supported in part by the European Union's Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Skłodowska-Curie Grant Agreement No. 745755. This work was also supported in part by the Government of Aragón and European Social Fund (EU) through BSICoS group (T39_20R), and in part by the by CIBER in Bioengineering, Biomaterials & Nanomedicine (CIBER-BBN) through Instituto de Salud Carlos III, and in part by the NSF SBIR Phase I (#1746589) and R43 HL135961. (Corresponding author: Jesús Lázaro.)

Jesús Lázaro is with the Biomedical Engineering Department, University of Connecticut, Storrs, CT 06269 USA, and also with the Biomedical Signal Interpretation and Computational Simulation (BSICoS) group, Aragón Institute of Engineering Research (I3A) IIS Aragón, University of Zaragoza, Spain (e-mail: jesus.lazaro@uconn.edu).

Natasa Reljin, Md-Billal Hossain, and Ki H. Chon are with the Biomedical Engineering Department, University of Connecticut.

Pablo Laguna is with the Biomedical Signal Interpretation and Computational Simulation (BSICoS) group, Aragón Institute of Engineering Research (I3A) IIS Aragón, University of Zaragoza and also with CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN).

Yeonsik Noh is with the College of Nursing Department of Electrical and Computer Engineering of the University of Massachusetts.

Digital Object Identifier 10.1109/TBME.2020.2987759

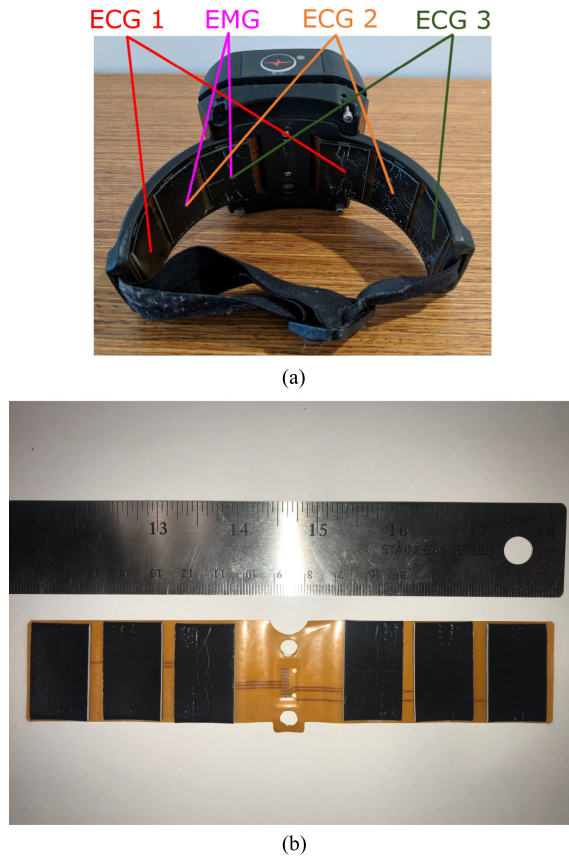


Fig. 1. (a) Armband device prototype and electrode configuration for the 3 ECG channels and the EMG channel. (b) Sensing part of the prototype.

the patch for a prolonged period of time. A wearable armband device aimed to monitor ECG during long periods, overcoming the limitations of Holter and patch devices, is being developed in our lab at the University of Connecticut. This armband is designed to be worn on the left upper arm, and incorporates three pairs of hydrophobic dry electrodes, which were also developed in our lab. Using these electrodes differentially, the armband can record 3 ECG channels simultaneously (see Fig. 1). A photo of the first prototype and another image of its sensing part can be observed in Fig. 1. The dimensions of the enclosed circuit board of this prototype are $65 \times 64 \times 28$ mm, while the dimensions of its sensing part are $170 \times 33 \times 0.1$ (1.5 when includes electrode thickness) mm. Note that this is a first prototype, created for assessing the feasibility of monitoring HR and HRV using dry electrodes over the upper arm. Although the prototype is currently large, the “sensing part” is thin, and the dimensions of the box can be considerably reduced in the final design.

Although the armband setup is much more convenient for long-term monitoring, it remains a more challenging scenario than the Holter setup mainly because of two reasons: not using hydrogel, and the electrodes’ location. Although the impedance matching (with the skin) of the electrodes is good [14], it is still not as good as that provided by hydrogels. With respect to the location of the electrodes, the armband device is located in the left upper arm while the Holter and patch devices use

electrodes over the chest. This leads to a lower signal power of the acquired ECG signal, and it is more susceptible to electromyogram (EMG) signals (mainly from the biceps and triceps). The EMG artifact contamination remains the most challenging obstacle for obtaining good fidelity armband ECG recordings. Some techniques for reducing the EMG artifact in ECG signals have been presented in the literature. As EMG overlaps with ECG in time and frequency, adaptive filtering techniques are usually used [15], [16], such as least mean squares (LMS) or its normalized version (NLMS) [17]. When multiple leads are available, the space diversity can be exploited. Principal component analysis (PCA) is an approach that exploits this, and has been proposed to attenuate the EMG artifact in the ECG [18]. The first principal component extracted by PCA is expected to be the component in which the noise has been reduced the most, especially for noises with muscular origin.

However, a pilot study showed that the quality of the armband-acquired ECG signals was high enough to obtain respiratory rate using ECG-morphology features during lab-controlled conditions with no movement [19]. Another pilot study showed promising results for 24 hours HR monitoring [20]. In this paper, the wearable armband device is evaluated as a 24 hour monitor during daily life. The study includes the application of PCA and NLMS signal processing techniques to deal with the EMG noise, the development of an automatic channel selector which selects the highest quality ECG signal at each time moment, and an automatic artifact detector which, significantly, discards noise-corrupted data largely due to EMG artifact. The methods are evaluated with a data set composed of 24-hour armband recordings during routine daily life, and simultaneously recorded Holter ECG signals.

II. MATERIALS AND METHODS

A. Data Acquisition and Preprocessing

The wearable armband records three ECG channels and one EMG channel with a sampling rate of $F_S = 1000$ Hz using three pairs of carbon-black dry electrodes [14]. In addition, the armband records tri-axial accelerometer channels with a sampling rate of 100 Hz. A picture of the armband and the configuration of multi-ECG channels is shown in Fig. 1. Moreover, Fig. 2 shows a picture of how the armband is designed to be worn on the upper left arm. Armband signals were continuously recorded from 16 healthy subjects aged 27.56 ± 8.82 years (mean \pm standard deviation) for 24 hours. The subjects were instructed to carry out their normal activities but without exercise. For reference purposes, three ECG channels were simultaneously recorded by a conventional commercially-available Holter: Rozinn RZ 153+ (Glendale, NY, USA).

The ECG signals from the armband were down-sampled to 256 Hz. Many of the potential applications for the armband are based on beat occurrences, and the value of this sampling rate is a trade-off between the time resolution and the computational cost. The value of 256 Hz was chosen because it is close to the minimum recommended for calculating the classical HRV indices [3], which are also based on beat occurrences. Furthermore, the ECG signals from the armband were found to be



Fig. 2. Subject wearing the armband device.

highly contaminated by noise, mainly due to the EMG from the local muscles when subjects moved their left arm. Thus, a strong band-pass filter was applied in order to remove much of the EMG artifact from ECG data. The low and high cut-off frequencies of this filter were set to 3 Hz and 25 Hz, respectively, based on the frequency bands used in the literature for QRS detection [21]. These filtered ECG signals are denoted $x_1(n)$, $x_2(n)$, and $x_3(n)$. Fig. 3 shows an example of these signals.

B. Channel Synthesis

Two ECG channels were synthesized from the armband data, to attenuate the effect of the EMG noise: one based on PCA [22], and another one based on the NLMS filter [17].

PCA channel: The first principal component extracted by PCA is expected to be the component in which the EMG noise has been most attenuated [18]. Similarly, the last principal component is expected to be the component in which the EMG noise is expected to be the most prominent. The armband data were split into 10-second segments, and PCA was applied segment by segment to all three ECG channels and the EMG channel. Then, the first component was normalized in amplitude with respect to its standard deviation, and inverted in case its minimum were greater than its maximum in absolute value. The concatenation of the resulting first principal component from the 10-s segments is denoted $x_{PCA}(n)$ in this paper, and it was considered as an additional ECG channel. In parallel, the concatenation of the last principal component from the 10-s segments is $x_N(n)$, and it was used as noise estimation for the NLMS adaptive filter.

NLMS filter: The NLMS filter can be seen as an adaptive Wiener filtering technique in which the filter is adapted based on the difference between the desired and the obtained filter. It can be used to attenuate the influence of a known corrupting noise on a corrupted signal, *e.g.*, to attenuate the EMG noise in the ECG [17]. An additional ECG channel was obtained by applying a NLMS filter using $x_{PCA}(n)$ as the corrupted signal, and $x_N(n)$ as an estimation of the corrupting noise. An example of the two synthesized ECG channels can be observed in Fig. 3, and a 10-s zoom in can be observed in Fig. 4.

C. Artifact Detection

An ECG artifact detection technique was developed based on the support vector machine (SVM) classifier with radial basis function (RBF) kernel [23]. It was designed to classify ECG segments of 10 seconds as normal or artifact, as this duration is enough to estimate the mean heart rate [24]. In addition, another criterion based on the level of signal power found in the accelerometers was used. A 10-second segment was considered as an artifact if either the SVM-based ECG artifact detector classified it as an artifact, or if a certain level of signal power was found in the accelerometers.

Features of the ECG artifact detector: Different signal quality indices (SQI) which are available in the literature were studied as potential features for the classifier. The SQI found in the literature can be divided into two groups: those based on fiducial features and those based on non-fiducial features [25]. When based on fiducial features, one detects the beats followed by the mean level and/or regularity of the resulting inter-beat intervals. However, abnormal values of mean level and, especially, of regularity of the inter-beat intervals are the key features for many potential applications, such as arrhythmia detection [1], epileptic seizure detection [4], stress assessment [5], and sleep studies [6]. The most valuable data for such applications may be considered of low quality by the SQI based on abnormal values of mean level and/or regularity of inter-beat intervals. Thus, no fiducial features were considered in this work. Nine other non-fiducial features were considered in this work:

- *Shannon entropy* (m_1) [26], which provides a quantitative measure of the average uncertainty present in a signal, quantifying how different its probability density function is from a uniform distribution. Thus, a clean ECG signal is expected to have a lower Shannon entropy than an EMG-corrupted ECG signal.
- *Multiscale entropy* (m_2) [27] is another measure of average uncertainty, in this case obtained from a sample entropy analysis of the signals for different time scale factors. As with the Shannon entropy, a clean ECG signal is expected to have a smaller multiscale entropy than an EMG-corrupted ECG signal has.
- *Ratio of powers* (m_3) [28], defined as the ratio between power in the frequency band 5–20 Hz with respect to the total power, which is expected to be higher when it is computed from a clean ECG than when it is computed from an EMG-corrupted ECG.
- *Self-correlation* (m_4) [29], defined as the autocorrelation at the highest peak, excluding the zero lag. As an ECG signal has a higher periodicity than an EMG signal, this self-correlation value is expected to be higher for a clean ECG than for an EMG-corrupted ECG.
- *Shannon entropy* (m_5), *mean* (m_6), and *variance* (m_7) of the first intrinsic mode function [30] are related to uncertainty and the power of the higher frequency components. Thus, they are expected to be lower for a clean ECG than for an EMG-corrupted ECG.
- *Skewness* (m_8) [31], which is expected to be further from zero for a clean ECG than for an EMG-corrupted ECG.

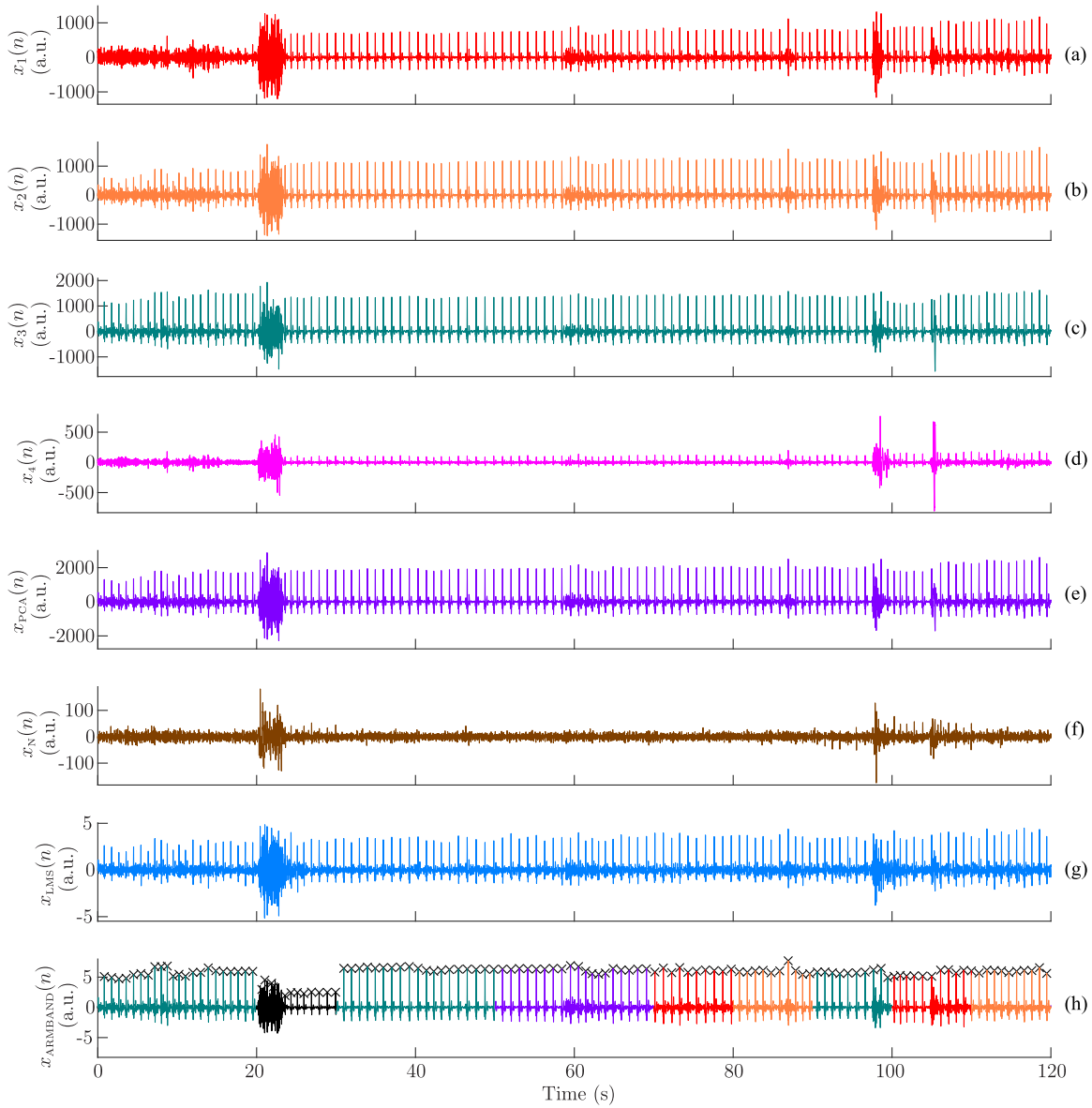


Fig. 3. A 2-minute segment of armband-recorded ECG channels (a) $x_1(n)$, (b) $x_2(n)$, and (c) $x_3(n)$, (d) recorded EMG channel $x_4(n)$, (e) ECG synthesized channel based on PCA $x_{PCA}(n)$ (f) EMG noise estimated by PCA $x_N(n)$, (g) ECG synthesized channel based on NLMS $x_{NLMS}(n)$, and (h) the ECG channel selected by the optimal channel selector $x_{ARMBAND}(n)$, where detected QRS complexes are represented with black 'X'. The color of $x_{ARMBAND}(n)$ corresponds to the color of the chosen channel at each instant. The segment from 20 to 30 seconds is represented in black because it was considered an artifact by the artifact detector.

- *Kurtosis* (m_9) [31], which is expected to be higher for a clean ECG than for an EMG-corrupted ECG.

Training set for the ECG artifact detector: The 10-second segments of the first hour from 5 subjects were labelled as artifact, normal, or neither. This labelling was based on the comparison of the mean HR estimated from the armband and the mean HR estimated from the Holter. The segment was considered as “artifact” if these mean HR estimations differed by more than 20%, whereas it was considered as “clean” if these mean HR estimations differed by less than 2%. 4,821 segments were labelled as normal, and 1,410 as artifact. Subsequently, the 9 features noted above were computed for distinguishing between artifact and normal segments, and 1,410 normal segments were

selected by a k-means algorithm in order to balance the groups and to obtain a good representation of the underlying distribution of the data. K-means was used to set 1,410 clusters in the “normal” class, and their centroids. Then, the element closest to each one of those centroids was selected. Therefore, a total of 2,820 segments (1,410 normal and 1,410 artifacts) were considered for the subsequent training of the SVM classifier. Feature selection was performed by a forward wrapper approach, which consisted of adding one feature at a time and selecting the one which provided the highest accuracy, and stopping when the obtained accuracy was lower than that obtained with one less feature. Subsequently, the SVM classifier was trained using only those features selected by the wrapper.

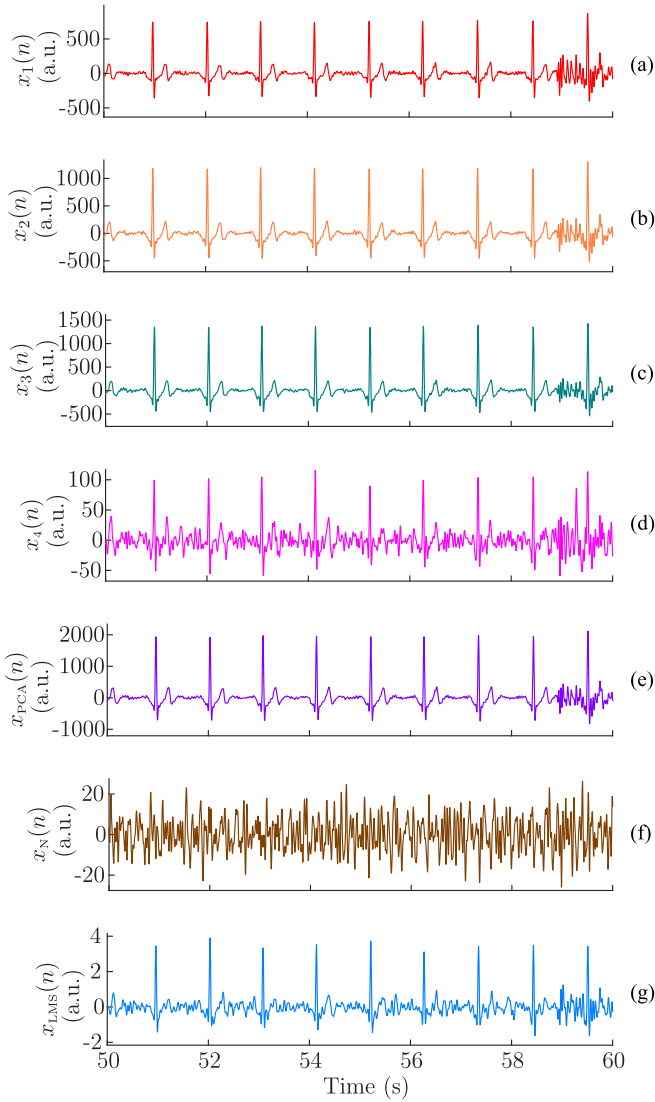


Fig. 4. A 10-second segment of armband recorded ECG channels (a) $x_1(n)$, (b) $x_2(n)$, and (c) $x_3(n)$, (d) recorded EMG channel $x_4(n)$, (e) ECG synthesized channel based on PCA $x_{PCA}(n)$ (f) EMG noise estimated by PCA $x_N(n)$, and (g) ECG synthesized channel based on NLMS $x_{NLMS}(n)$.

Test set for the ECG artifact detector: In order to assess the performance of the classifier, a test set was created by labelling the 10-second segments of the first hour from the 11 subjects who were not included in the training set. The criterion for labelling was exactly the same as that used for labeling the training set. The test set was composed of a total of 9,899 fragments (3,800 artifacts + 6,099 normal).

Accelerometer-based rule for artifact detection: The level of power in the accelerometers was defined as:

$$P_{ACC} = \text{var}(x_X(n) + x_Y(n) + x_Z(n)), \quad (1)$$

where $x_X(n)$, $x_Y(n)$, and $x_Z(n)$ denote the x, y, and z accelerometer channels, respectively, and $\text{var}(\cdot)$ denotes the variance. A 10-second segment was considered as artifact if P_{ACC} was higher than a certain threshold. This criterion was included

in order to detect some of the artifacts that were not detected from the ECG signal by the SVM-based detector, under the assumption that some may be related to movements. In order to set this threshold, P_{ACC} was computed from those segments of the training set which were classified as “normal” by the ECG-based artifact detector. The threshold was set to 0.02864 G^2 , which maximized the accuracy and did not discard more than 0.01% of (labelled as) “normal” segments.

D. Channel Selection

Among three ECG channels, the best signal fidelity channel was selected for every 10-second segment for further processing. Five channels were considered by the optimal channel selector: $x_1(n)$, $x_2(n)$, $x_3(n)$, $x_{PCA}(n)$, and $x_{NLMS}(n)$. The selection was based on the SVM classifier with RBF kernel used for artifact detection. The selected ECG channel was determined to be the one with the highest likelihood of belonging to the “normal” group. In the event that several channels obtained the same likelihood of belonging to the “normal” class, a similar signal quality was expected in those signals. The algorithm gives preference to the original channels $x_1(n)$, $x_2(n)$, and $x_3(n)$, in this order, and later to $x_{PCA}(n)$ and $x_{NLMS}(n)$, in this sequence. In this manner, a unique armband ECG signal $x_{ARMBAND}(n)$ was created by concatenating those selected segments at different time points. Fig. 3h shows an example of $x_{ARMBAND}(n)$.

E. Mean Heart Rate Measurement

The location of the QRS complexes of $x_1(n)$, $x_2(n)$, $x_3(n)$, $x_{PCA}(n)$, $x_{NLMS}(n)$, and $x_{ARMBAND}(n)$ were automatically detected by an algorithm based on variable frequency complex demodulation (VFCDM) and some adaptive threshold rules [32]. This algorithm was applied in segments of 20 seconds with 5 seconds of overlap, leaving 10 effective seconds at each segment. The fiducial point of each QRS complex, n_{QRS_i} , was set to that where the absolute value of the amplitude was maximum (R peak). Then, the instantaneous HR was computed from n_{QRS_i} as the inverse of the beat-to-beat intervals:

$$\hat{d}_{HR}^u(n) = F_S \sum_i \frac{1}{n_{QRS_i} - n_{QRS_{i-1}}} \delta(n - n_{QRS_i}), \quad (2)$$

where the superscript “u” denotes that the signal is unevenly sampled. A 4-Hz-evenly-sampled version of $\hat{d}_{HR}^u(n)$ was obtained by cubic-splines interpolation, and it is denoted as $\hat{d}_{HR}(n)$ in this paper.

For comparison purposes, the mean HR was calculated also from the Holter device by a similar procedure, obtaining the reference HR series $d_{HR}(n)$. The channel selector was not used in this case. Instead, the first channel was always used. The artifact detector was also not used. In order to identify those segments with artifacts in this channel, two different QRS detectors ([32] and [33]) were applied on the same signal. Those segments in which these QRS detectors offered a different output were considered artifacts and they were discarded for further analysis. The segments from the armband data identified as artifacts by the artifact detector were also discarded from further analysis. Note that this is a stricter criterion than that used for the

armband-ECG signals. This criterion requires that a rudimentary QRS detector performs as good as the sophisticated approach on the analyzed segment. This stricter criterion allows us to be sure that we only evaluate the armband performance when we have a reliable reference, without discarding too much data. However, this criterion cannot be applied to the armband data without discarding too many data segments as the armband signals are noisier in general, and the rudimentary QRS detector usually does not work for the armband.

The delay between $\hat{d}_{HR}(n)$ and $d_{HR}(n)$ series was estimated (and corrected) as the lag that maximized their cross correlation. Then, the percentage of 10-second segments where the mean HR estimated from the armband (mean of $\hat{d}_{HR}(n)$) differed by less than 10% from the mean HR estimated from the Holter (mean of $d_{HR}(n)$) was computed. This analysis was performed for non-bed time and for bed time independently. Bed time was reported by each subject, and further adjusted manually by visual inspection of the accelerometer signals. In addition, these percentages were computed also using only those 10-s segments that were determined to be usable according to the artifact detector described in Section II-C.

F. Heart Rate Variability Analysis

HRV analysis was performed in 5-min-length (overlapped 4 min) windows in which segments from $x_{\text{ARMBAND}}(n)$ were determined to be usable according to the artifact detector described in Section II-C. The length of 5 minutes was chosen since it is recommended in [3]. Four standard HRV parameters were computed: standard deviation of successive normal-to-normal beat intervals (SDNN), the root mean square of successive differences of normal-to-normal beat intervals (RMSSD), the power of $\hat{d}_{HR}(n)$ within the low frequency band [0.04 Hz, 0.15 Hz] (LF), and the power of $\hat{d}_{HR}(n)$ within the high frequency band [0.15 Hz, 0.4 Hz] (HF) [3]. These four parameters were computed also from $x_{\text{HOLTER}}(n)$ and used as the reference.

For each one of the 5-min-length windows, the relative error of the armband-derived parameters with respect to the Holter-derived parameters was computed. For each subject, the (intra-subject) median and IQR of this relative error was computed. In addition, the inter-subject median and IQR of those intra-subject medians and IQRs were also computed.

III. RESULTS

A. Artifact Detector

The forward wrapper selected all the 9 studied features for the SVM-based ECG artifact detector and channel selector, in the following order: $m_9, m_3, m_7, m_1, m_8, m_4, m_5, m_2, m_6$. The P_{ACC} -based threshold was set to 0.02864 G^2 , as it maximized the accuracy and did not discard more than a 0.01% of (labelled as) “normal” segments.

The resulting classifier obtained an accuracy of 90.79% in the test set, a sensitivity of 92.05%, a specificity of 90.00%, a positive predictive value of 85.15%, and a negative predictive value of 94.79%.

B. 24-hour Heart Rate Monitoring

Table I shows the median and interquartile range (IQR) of the percentage of segments where the heart rate was accurately estimated (less than 10% of relative error with respect to the Holter) from the armband device, during both non-bed time and bed time. In addition, Table I shows the median and IQR of the percentage of usable data obtained from the armband device according to the artifact detector as well as the percentage of usable data from the Holter, during both non-bed time and bed time. Note that these results were based on skipping the first hour of the 5 subjects that were used for training the artifact detector. A Bland-Altman plot illustrating the HR estimated from $x_{\text{ARMBAND}}(n)$ vs. the HR estimated from the Holter device is shown in Fig. 5. The obtained bias was 0.08 bpm, and the length of the limits of agreement was 6.58 bpm. Correlation between these measures was 0.95.

Table II shows the inter-subject median and IQR of intra-subject medians when estimating HRV parameters with respect to the Holter device, also based on skipping the first hour of the 5 subjects that were used for training the artifact detector. Note that a negative relative error corresponds to an underestimation of the studied parameter. These median values are illustrated in Fig. 6. The correlations between the HRV parameters estimated by the armband and those estimated from the Holter were 0.9479, 0.9142, 0.9989, and 0.9984 for SDNN, RMSSD, LF, and HF, respectively.

IV. DISCUSSION

ECG data from a wearable armband have been analyzed for HR monitoring during daily life. The armband simultaneously records three ECG channels and one EMG channel. However, the ECG channels were often contaminated with EMG artifacts during arm movements. In order to automatically discard the noisy data due to EMG corruption, a novel artifact detector was developed, based on the SVM classifier with RBF kernel and nine features that have been reported to be related to the ECG signal quality in the literature [25]. All nine features were selected by a forward-wrapper approach, suggesting that the features have complementary information about the signal quality, at least in part. The classifier was trained using the ECG segments of the first hour from 5 subjects, and tested with the ECG segments of the first hour from the remaining 11 subjects. The obtained accuracy in the test set was 90.79%, with a negative predictive value of 94.79%. This means that when a segment is detected as “normal,” the expectance of not having a clean ECG signal is 5.21%. Note that, in this case, a clean ECG signal is that from which the mean HR was estimated with an error lower than 1%. In addition, a segment was considered to contain artifacts if a certain level of power was found in the accelerometer signals. This accelerometer-based criterion was designed to find the artifacts that were not detected by the ECG-based artifact detector. A total of 85.31% of the detected artifacts were found by the ECG-based detector, while 45.52% of the detected artifacts were found by the accelerometer-based criterion. A total of 37.12% of the detected artifacts were found

TABLE I

MEDIAN AND INTERQUARTILE RANGE (MEDIAN / IQR) OF THE PERCENTAGE OF SEGMENTS WITH ACCURATE HEART RATE ESTIMATES (LESS THAN 10% OF RELATIVE ERROR WITH RESPECT TO THE HOLTER) FROM THE WEARABLE ARMBAND DEVICE; OF THE PERCENTAGE OF SEGMENTS WITH USABLE ARMBAND DATA ACCORDING TO THE ARTIFACT DETECTOR; AND OF THE PERCENTAGE OF SEGMENTS WITH HOLTER USABLE DATA; DURING NON-BED TIME AND DURING BED TIME

	Non-bed time		Bed time	
	Percentage of segments with		Percentage of segments with	
	usable data	accurate heart rate	usable data	accurate heart rate
$x_1(n)$	40.75% / 24.81%	98.99% / 0.71%	94.28% / 2.39%	99.28% / 0.24%
$x_2(n)$	45.02% / 17.60%	99.06% / 0.62%	94.49% / 4.20%	99.33% / 0.19%
$x_3(n)$	42.98% / 23.49%	98.72% / 0.61%	93.86% / 11.83%	99.22% / 0.19%
$x_{PCA}(n)$	39.34% / 24.04%	98.98% / 0.67%	91.84% / 9.72%	99.31% / 0.23%
$x_{NLMS}(n)$	36.83% / 17.58%	99.15% / 0.65%	84.29% / 17.04%	99.29% / 0.16%
$x_{ARMBAND}(n)$	53.85% / 17.09 %	98.54% / 0.99%	95.00% / 2.35%	99.25% / 0.21 %
$x_{HOLTER}(n)$	85.07% / 15.45%	-	97.14% / 5.08%	-

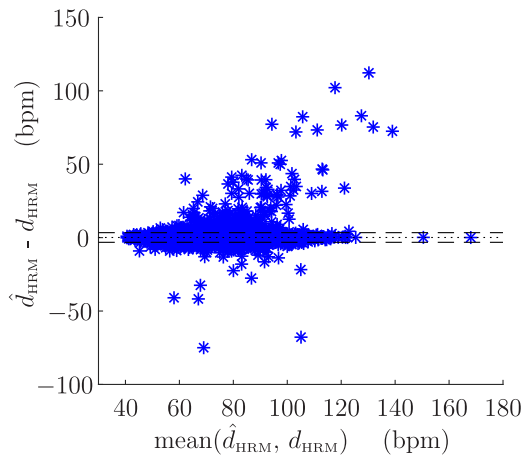


Fig. 5. Bland-Altman plot illustrating the HR estimated from $x_{ARMBAND}(n)$ (\hat{d}_{HRM}) with respect to the HR estimated from the Holter d_{HRM} . The obtained bias was 0.08 bpm, and the length of the limits of agreement was 6.58 bpm.

by both criteria. We learned that the accelerometer criterion should be applied first (as it is faster), and then is no need to compute the ECG-based criterion if an artifact is found.

The mean HR was estimated in every 10-second segment from each ECG channel of the armband $x_1(n)$, $x_2(n)$, and $x_3(n)$, and was compared to the mean HR estimated from the Holter monitor, which was taken as the Gold Standard reference data. The estimation was considered accurate if the mean HR differed by less than 10% from the estimation based on the Holter monitor. Note that this is a stringent criterion, as 10% in 10 seconds corresponds to an error of only one beat in a typical resting HR (around 60 per minute).

Results obtained from these armband ECG channels were not optimal during the non-bed time. The best channel in terms of median of percentage of segments with usable data was $x_2(n)$ (45.02%), obtaining an accurate HR estimate in a median of

TABLE II

INTER-SUBJECT MEDIAN AND IQR OF INTRA-SUBJECT MEDIANS OF OBTAINED RELATIVE ERROR WITH RESPECT TO THE HOLTER DEVICE WHEN ESTIMATING HRV PARAMETERS FROM $x_{armband}(N)$, FOR ALL 16 SUBJECTS, AND FOR ONLY THOSE 11 SUBJECTS THAT WERE NOT USED FOR TRAINING THE ARTIFACT DETECTOR

			Relative error Intra-subject median (median / IQR)	
			16 subjects	11 subjects
16 subjects	Non-bed time	SDNN	-0.04% / 0.25%	-0.07% / 0.15%
		RMSSD	-0.05% / 0.58%	-0.20% / 0.57%
		LF	-0.03% / 0.25%	0.03% / 0.43%
		HF	0.07% / 1.15%	-0.09% / 1.50%
	Bed time	SDNN	-0.09% / 0.15%	-0.11% / 0.13%
		RMSSD	-0.10% / 0.27%	-0.14% / 0.29%
		LF	0.16% / 0.21%	0.17% / 0.23%
		HF	-0.31% / 1.37%	-0.29% / 1.31%

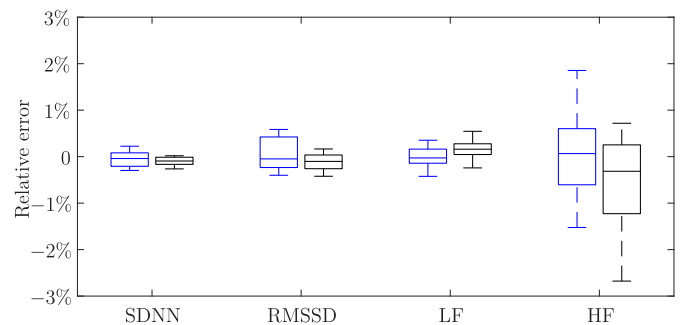


Fig. 6. Boxplot of intra-subject median of obtained relative error when estimating HRV parameters. Boxes on the left (in blue) are from non-bed time data, and boxes on the right (in black) are from bed time data.

99.06% of usable segments. Results were much better during bed time, when the movements of the subjects were greatly reduced. During bed time, the median of percentage of usable segments was 94.49%, and the mean HR was estimated accurately from 99.33% (in median) of these segments. However, none of these ECG channels was observed to be consistently the cleanest ECG signal for every subject at every time. A possible reason for this observation is that the electrode distribution (around the left upper arm) makes the recorded leads to be very dependent on the shape of the arm. Furthermore, the armband can slightly rotate during the recording, affecting the lead channels. Note that $x_1(n)$ could be $x_2(n)$ or $x_3(n)$ with the appropriate device rotation.

In order to clean the EMG noise, two signal processing techniques were applied: PCA [18] and NLMS filtering [17]. PCA was performed in segments of length as short as 10 seconds, making it robust against possible device rotations. $x_{\text{PCA}}(n)$ obtained usable data from a median of 39.34% during the non-bed time, and of 91.84% during bed time. This performance is lower than the performance of the original channels $x_1(n)$, $x_2(n)$, and $x_3(n)$. A possible reason for these results is that the electrodes of the armband are very close to each other, giving the leads a similar influence from the EMG, which is substantial, especially during arm movements. Furthermore, a strong EMG component would lower the automatic gain control which consequentially attenuates the ECG signal. This indicates that when local muscles (mainly left biceps and left triceps) contract, the EMG component becomes a principal component. This may be the reason why $x_{\text{NLMS}}(n)$ did not fare well (36.83% during the non-bed time and 84.29% during the bed time). The NLMS filter relies on the availability of good estimation of the noise, while our estimation of the noise (the last component from the PCA) often contained the desired signal (the ECG), albeit its magnitude was small when compared to the noise-free ECG case.

Similarly, for ECG channels $x_1(n)$, $x_2(n)$, and $x_3(n)$, none of the synthesized ECG channels were observed to be consistently the most clean ECG at all times, thus, they were treated as two additional ECG channels and a novel channel selector was developed in order to choose the best channel at each time. The channel selector was based on the SVM classifier with RBF kernel used for artifact detection. This SVM classifier was used to select the most clean ECG channel among the 5 eligible channels ($x_1(n)$, $x_2(n)$, $x_3(n)$, $x_{\text{PCA}}(n)$, and $x_{\text{NLMS}}(n)$), for every 10-second segment. A new signal $x_{\text{ARMBAND}}(n)$, composed of the cleanest segment at each time, was generated. Results obtained with $x_{\text{ARMBAND}}(n)$ outperformed those obtained from any of the ECG channels separately in terms of median percentage of usable data during both non-bed and bed time, while they were similar in terms of percentage of usable segments with accurate HR, demonstrating the advantage of best channel selection. During the non-bed time, a median of 53.85% usable data was obtained. Results during bed time were much better, obtaining usable data from a median of 95.00% of the segments. Note that this is very close to $x_{\text{HOLTER}}(n)$ performance during the bed time (97.14% usable data).

Furthermore, the median of the percentage of usable segments providing an accurate HR estimation was 98.54% during non-bed time, and 99.25% during bed time. The Bland-Altman plot in Fig. 5 did not show a dependence of the accuracy of the usable segments on the actual HR. Moreover, the HRV analysis in 5-min-length windows of continuous usable armband data obtained a relative error with respect to the Holter with inter-subject median and IQR of intra-subject medians not higher than 1.54% (see Table II), and they showed strong correlation. The analysis was repeated using only those 11 subjects that were not used for training the artifact detector in order to assess the possible bias of the results due to a possible overfitting. In this case, the mean HR was accurately estimated from a median of 98.55% of the segments detected as “clean” during the non-bed time, and a median of 99.29% during the bed time. These results are very similar to those obtained when using all subjects (98.54% and 99.25%, respectively), suggesting that the artifact detector did not overfit to those 5 subjects that were used for training. Furthermore, the relative errors obtained for the HRV analysis were also similar for both cases (See Table II), reinforcing this suggestion. These results suggest that the analysis for all of the subjects is not biased. A possible reason is that no differences were observed in the signals from the different subjects, thus, no overfitting to those 5 subjects is expected. Furthermore, the classifier was trained using only the first hour from those 5 subjects, and those data (from that first hour) were not used for further analyses, so no bias due to overfitting to those particular segments is expected.

Therefore, the obtained HR and HRV estimations were accurate for almost all the segments detected as usable by the artifact detector, suggesting that the segments automatically classified as clean are reliable for QRS detection. However, some of the segments that were automatically classified as artifact may be also reliable for this purpose (false positives), so the actual coverage of the armband device may be higher. In order to assess the actual coverage of the armband device, the percentage of segments of $x_{\text{ARMBAND}}(n)$ that offered an accurate estimation of mean HR was computed. This analysis revealed that the mean HR was accurately estimated from 75.25% / 11.02% (inter-subject median/IQR) of the armband-ECG segments during the non-bed time, and from 98.49% / 0.79% during bed time. This means that 21.40% (in median) of the segments during the non-bed time and 3.49% (in median) of the segments during the bed time were automatically classified as artifacts while they could provide an accurate estimation of the mean HR. Thus, the artifact detector is strict: those data that were automatically classified as clean are reliable at the expense of discarding a high amount of reliable data. Hence, there is room for improving this in the artifact detection, especially during the non-bed time.

However, these results are promising and suggest that the armband has a strong potential to be a wearable long-term ECG monitor, especially during bed time, which can potentially be used for overnight recordings such as sleep studies. During non-bed time, the armband records considerably less usable data per day than a conventional Holter monitor, but it causes no skin irritation and it is much more comfortable for the patient,

so it can be worn for months or even years when compared to the Holter. Therefore, the armband can provide much more usable data in total than the Holter even if it offers less usable data per day. Moreover, the obtained coverages were higher than those reported for PPG-based wearable devices. Only 134 of 908 segments (14.76%) were reported to be usable from PPG during lab-controlled conditions in [8]. A mean coverage of 76.34% was reported in [10], using PPG, where recordings from patients in bed during controlled-movement-restricted conditions were analyzed. In [9], 24% coverage from PPG was reported when requiring the same accuracy as a Holter device during 24-hour recordings. A higher mean coverage during 24-hour recordings was reported in [11], where 56% of data from PPG was considered to be usable. For comparison of these numbers with those obtained by the armband, it should be noted that subjects in [11] are patients aged 67.4 ± 12.1 years, so they may have much less active life styles than the subjects (healthy volunteers aged 27.56 ± 8.82 years) analyzed in the present study. It should also be noted that 24-hour recordings include both non-bed and bed time. The overall coverage of the armband during the 24-hour recordings in this study was $83.51\% \pm 8.00\%$.

These results suggest that the armband is very interesting for applications that may benefit from long-term ECG monitoring, such as paroxysmal arrhythmia detection, seizure detection, stress assessment, and monitoring of chronic respiratory patients. However, some specific problems may occur in different applications, such as patient movements that could lead to unusable data in the precise moment when it is most valuable. *E.g.*, a patient could move the arm every time that he or she has a short-period AF episode due to chest pain, or every time he or she has a seizure. Thus, further studies have to be conducted in order to evaluate the full potential of the armband in different applications and scenarios. Furthermore, in this paper the armband has been evaluated for QRS detection (through HR and HRV). Further studies have to be conducted for evaluating the armband device as a monitor for other ECG features which may be relevant in some applications, such as the ST elevation for ischemia.

V. CONCLUSION

The results suggest that the armband device is suitable for daily life HR monitoring, obtaining usable data approximately 3/4 of the non-bed time (median of 75.25%) and almost all the bed time (median of 98.49%). The automatic artifact algorithm found 53.85% (median) of data to be usable during the non-bed time, and 95.00% during the bed time. However, further studies must be conducted in order to assess the full potential of the armband for specific applications, such as arrhythmia detection, sleep studies, seizure detection, stress assessment, or monitoring of chronic respiratory patients.

ACKNOWLEDGMENT

The computation was performed by the ICTS NANBIO-SIS, specifically by the High Performance Computing Unit of CIBER-BBN at the University of Zaragoza.

REFERENCES

- [1] S. Dash, K. H. Chon, S. Lu, and E. A. Raeder, "Automatic real time detection of atrial fibrillation," *Ann. Biomed. Eng.*, vol. 37, no. 9, pp. 1701–1709, Sep. 2009, doi: [10.1007/s10439-009-9740-z](#).
- [2] S. H. Hohnloser et al., "Incidence of stroke in paroxysmal versus sustained atrial fibrillation in patients taking oral anticoagulation or combined antiplatelet therapy: An ACTIVE W Substudy," *J. Am. Coll. Cardiol.*, vol. 50, no. 22, pp. 2156–2161, Nov. 2007, doi: [10.1016/j.jacc.2007.07.076](#).
- [3] "Heart rate variability: Standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology," *Circulation*, vol. 93, no. 5, pp. 1043–1065, Mar. 1996.
- [4] M. Zijlmans, D. Flanagan, and J. Gotman, "Heart rate changes and ECG abnormalities during epileptic seizures: Prevalence and definition of an objective clinical sign," *Epilepsia*, vol. 43, no. 8, pp. 847–854, Aug. 2002.
- [5] R. Castaldo, L. Montesinos, P. Melillo, C. James, and L. Pecchia, "Ultra-short term HRV features as surrogates of short term HRV: A case study on mental stress detection in real life," *BMC Med. Inform. Decis. Mak.*, vol. 19, no. 1, p. 12, Jan. 2019, doi: [10.1186/s12911-019-0742-y](#).
- [6] C. Varon, A. Caicedo, D. Testelmans, B. Buyse, and S. Van Huffel, "A novel algorithm for the automatic detection of sleep apnea from single-lead ECG," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 9, pp. 2269–2278, Sep. 2015, doi: [10.1109/TBME.2015.2422378](#).
- [7] T. Pereira et al., "Photoplethysmography based atrial fibrillation detection: a review," *Npj Digit. Med.*, vol. 3, no. 1, pp. 1–12, Jan. 2020, doi: [10.1038/s41746-019-0207-9](#).
- [8] S. K. Basharet et al., "Atrial fibrillation detection from wrist photoplethysmography signals using smartwatches," *Sci. Rep.*, vol. 9, no. 1, pp. 1–10, Oct. 2019, doi: [10.1038/s41598-019-49092-2](#).
- [9] L. M. Eerikainen et al., "Comparison between electrocardiogram- and photoplethysmogram-derived features for atrial fibrillation detection in free-living conditions," *Physiol. Meas.*, vol. 39, no. 8, p. 084001, 08 2018, doi: [10.1088/1361-6579/aad2c0](#).
- [10] A. Tarniceriu et al., "The accuracy of atrial fibrillation detection from wrist photoplethysmography. A study on post-operative patients," *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2018, pp. 1–4, 2018, doi: [10.1109/EMBC.2018.8513197](#).
- [11] A. G. Bonomiet et al., "Atrial fibrillation detection using a novel cardiac ambulatory monitor based on photo-plethysmography at the wrist," *J. Am. Heart Assoc.*, vol. 7, no. 15, p. e009351, 07 2018, doi: [10.1161/JAHA.118.009351](#).
- [12] H.-C. Junget et al., "CNT/PDMS composite flexible dry electrodes for long-term ECG monitoring," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 5, pp. 1472–1479, May 2012, doi: [10.1109/TBME.2012.2190288](#).
- [13] Y. T. Tsukada et al., "Validation of wearable textile electrodes for ECG monitoring," *Heart Vessels*, vol. 34, pp. 1203–1211, 2019, doi: [10.1007/s00380-019-01347-8](#).
- [14] B. A. Reyes et al., "Novel electrodes for underwater ECG monitoring," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1863–1876, Jun. 2014, doi: [10.1109/TBME.2014.2309293](#).
- [15] M. AlMahamdy and H. B. Riley, "Performance study of different denoising methods for ECG signals," *Procedia Comput. Sci.*, vol. 37, pp. 325–332, Jan. 2014, doi: [10.1016/j.procs.2014.08.048](#).
- [16] Z. Wang, J. Zhu, T. Yan, and L. Yang, "A new modified wavelet-based ECG denoising," *Comput. Assist. Surg. Abingdon Engl.*, vol. 24, pp. 174–183, 2019, doi: [10.1080/24699322.2018.1560088](#).
- [17] F. R. Hashim, J. J. Soraghan, L. Petropoulakis, and N. G. N. Daud, "EMG cancellation from ECG signals using modified NLMS adaptive filters," in *2014 IEEE Conf. Biomed. Eng. Sci. (IECBES)*, 2014, pp. 735–739, doi: [10.1109/IECBES.2014.7047605](#).
- [18] F. Castells, P. Laguna, L. Sörnmo, A. Bollmann, and J. M. Roig, "Principal component analysis in ECG signal processing," *EURASIP J. Adv. Signal Process.*, vol. 2007, no. 1, p. 074580, Feb. 2007, doi: [10.1155/2007/74580](#).
- [19] J. Lázaro, R. Bailón, E. Gil, Y. Noh, P. Laguna, and K. H. Chon, "Pilot study on electrocardiogram derived respiratory rate using a wearable armband," presented at the *XLV Int. Conf. Comput. Cardiol.*, 2018, doi: [10.22489/CinC.2018.054](#).
- [20] J. Lázaro, N. Reljin, Y. Noh, P. Laguna, and K. H. Chon, "Feasibility of long-term daily life electrocardiogram monitoring based on a wearable armband device," *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2019, pp. 4314–4317, Jul. 2019, doi: [10.1109/EMBC.2019.8857219](#).

- [21] L. Sörnmo and P. Laguna, "ECG signal processing," in *Bioelectrical Signal Processing in Cardiac and Neurological Applications.*, Burlington, MA, USA: Elsevier Academic, 2005.
- [22] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. New York, NY, USA: Springer-Verlag, 2002.
- [23] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: [10.1007/BF00994018](https://doi.org/10.1007/BF00994018).
- [24] C. Orphanidou, T. Bonnici, P. Charlton, D. Clifton, D. Vallance, and L. Tarassenko, "Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 3, pp. 832–838, May 2015, doi: [10.1109/JBHI.2014.2338351](https://doi.org/10.1109/JBHI.2014.2338351).
- [25] U. Satija, B. Ramkumar, and M. S. Manikandan, "A review of signal processing techniques for electrocardiogram signal quality assessment," *IEEE Rev. Biomed. Eng.*, vol. 11, pp. 36–52, 2018, doi: [10.1109/RBME.2018.2810957](https://doi.org/10.1109/RBME.2018.2810957).
- [26] N. Selvaraj, Y. Mendelson, K. H. Shelley, D. G. Silverman, and K. H. Chon, "Statistical approach for the detection of motion/noise artifacts in Photoplethysmogram," *Conf. Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2011, pp. 4972–4975, 2011, doi: [10.1109/IEMBS.2011.6091232](https://doi.org/10.1109/IEMBS.2011.6091232).
- [27] Y. Zhang, S. Wei, Y. Long, and C. Liu, "Performance analysis of multiscale entropy for the assessment of ECG signal quality," *J. Electr. Comput. Eng.*, vol. 2015, Art. no. 563915, 2015, doi: [10.1155/2015/563915](https://doi.org/10.1155/2015/563915).
- [28] G. D. Clifford, D. Lopez, Q. Li, and I. Rezek, "Signal quality indices and data fusion for determining acceptability of electrocardiograms collected in noisy ambulatory environments," in *2011 Comput. Cardiol.*, 2011, pp. 285–288.
- [29] H. Xia *et al.*, "Computer algorithms for evaluating the quality of ECGs in real time," in *2011 Comput. Cardiol.*, 2011, pp. 369–372.
- [30] J. Lee, D. D. McManus, S. Merchant, and K. H. Chon, "Automatic motion and noise artifact detection in Holter ECG data using empirical mode decomposition and statistical approaches," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 6, pp. 1499–1506, Jun. 2012, doi: [10.1109/TBME.2011.2175729](https://doi.org/10.1109/TBME.2011.2175729).
- [31] J. Behar, J. Oster, Q. Li, and G. D. Clifford, "ECG signal quality during arrhythmia and its application to false alarm reduction," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 6, pp. 1660–1666, Jun. 2013, doi: [10.1109/TBME.2013.2240452](https://doi.org/10.1109/TBME.2013.2240452).
- [32] S. K. Bashar, A. J. Walkey, D. D. McManus, and K. H. Chon, "VERB: VFCDM-based electrocardiogram reconstruction and beat detection algorithm," *IEEE Access*, vol. 7, pp. 13856–13866, 2019, doi: [10.1109/ACCESS.2019.2894092](https://doi.org/10.1109/ACCESS.2019.2894092).
- [33] M. E. Nygård and L. Sörnmo, "Delineation of the QRS complex using the envelope of the e.c.g.," *Med. Biol. Eng. Comput.*, vol. 21, no. 5, pp. 538–547, Sep. 1983.