Robust Assessment of Photoplethysmogram Signal Quality in the Presence of Atrial Fibrillation

Tania Pereira[1,2], Kais Gadhoumi[1,2], Mitchell Ma[1,2], Rene Colorado[3], Kevin J Keenan[3], Karl Meisel[3], Xiao Hu[1,2,4,5]

[1] Department of Physiological Nursing, University of California San Francisco (UCSF) , USA
[2] Center for Physiologic Research, UCSF, USA
[3] Department of Neurology, School of Medicine, UCSF, USA
[4] Department of Neurological Surgery, Affiliate Faculty of ICHS, San Francisco, USA
[5] Core Member of UCB/UCSF Joint Bio-Engineering Graduate Program, UCSF, USA

Abstract

A great deal of algorithms currently available to assess the quality of photoplethysmogram (PPG) signals is based on the similarity between pulses to derive signal quality indices. This approach has limitations when pulse morphology become variable due to the presence of some arrhythmia as in the case of atrial fibrillation (AFib). AFib is a heart arrhythmia characterized in the electrocardiogram mainly by an irregular irregularity. This arrhythmicity is reflected on PPG pulses by the presence of non-uniform pulses and poses challenges in the evaluation of the signal quality. In this work, we first test the performance of few algorithms from the body of methods reported in literature using a dataset of PPG records with AFib, and demonstrate their limitation. Second, we present a novel SVM-based classifier for PPG quality assessment in 30s-long segments of PPG records extracted from pulse oximetry data of 13 stroke patients admitted to the UCSF medical center neuro ICU. 40 time-domain, frequency domain and non-linear features were extracted from all segments. Using an independent test set, the classifier reached a 0.94 accuracy, 0.95 sensitivity and 0.91 specificity. These results demonstrate the robustness of the proposed method in properly evaluating PPG signal quality in the presence of atrial fibrillation.

1. Introduction

Atrial fibrillation (AFib) is the most common type of arrhythmia with an approximate prevalence 3% in adults older than 20 years [1]. AFib has a prevalence of 30 % in stroke patients, and additionally, is associated with a poorer neurological outcome than stroke patients without AFib [2]. AFib can occur occasionally and with short duration, a long-term ECG monitoring has been used for the diagnosis. The typical pattern of AFib in ECG is defined by irregular RR intervals; no discernible P waves; and at least 30 s of episode duration [1]. New tools that allow monitoring of heart activity in stroke patients will be important for risk stratification and has the potential to prevent stroke [3]. A great amount of effort had been made to develop wearable and mobile solutions for AFib detection [3]. Photoplethysmagraphy (PPG) signal is measured non-invasively in the peripheral areas of the body (as fingers or wrist) and can be incorporated into a wearable device. However, it is not a trivial task to acquire interference-free and clean PPG signals in real-world applications. The integrity of the signal is crucial for the pathological abnormalities identification and avoiding false alarms. The definition of good PPG quality is not straightforward since several factors need to be taken into considerations. As a continuous physiological signal, certain signal characteristics are expected to be stable over time. Accelerometry was introduced in some devices in order to identify periods with motion effect in the PPG [4], [5]. Also, the simultaneous electrocardiogram (ECG) signal synchronized with the PPG was used to identify the physiological beats from the contaminated PPG [6]. Taking account of these main aspects, the quality assessment of PPG should be based on: similarity between adjacent pulses [7]; absence of artifacts [6], [7]; clearly distinguishable the peaks in the morphology and presence of dicrotic notch [8]; absence of baseline fluctuations [7]; high signal-to-noise ratio; time aligned with ECG [6]; no correlation with accelerometer signal [9].

Several studies have been dedicated to develop algorithms that allow the detection of good and poor signals for the PPG signals. S. Asgari *et al.* developed a signal quality index (SQI) and used a 1336 ten-sec segments (18 472 beats) obtaining a true positive rate of

 ISSN: 2325-887X DOI: 10.22489/CinC.2018.254

99.06% and false positive rate of 7.69% [10]; A. Sukor *et al.* developed other SQI that showed an accuracy of 83 $\pm$ 11% using 104 60-sec PPG segments [6]; G. Clifford *et al.* obtained 95.2% accuracy with a new SQI algorithm [11]; W. Karlen *et al.* created an SQI that presented a sensitivity of 96.21% [7]; C. Orphanidou *et al.* developed a SQI that achieved the sensitivity of 91% and 95% of specificity [12]; C. Liu *et al.* presented an SQI with 90.79 % accuracy [13] and recently G. Papini *et al.* presented a different solution with a sensitivity higher than 90% [14]. Current signal quality assessment approaches compare the similarity between consecutive beats or using static evaluator algorithm that relies on thresholds derived from 'common-sense' physiology. They showed good performances in normal subjects but not in patients with AFib. The irregular irregularity of the rhythm in cardiac activity that characterizes the atrial fibrillation also produces differences between pulses that increases the difficulty to evaluate the quality of the signal (Figure 1), and can have impact in the performance of theses previous algorithms. In this work, we tested the previous SQI in order to assess the performance with AFib cases, and developed a novel method to overcome the limitations of the existing approaches. The solution proposed is based on two-class SVM approach classification using multi-domain features extracted from short duration PPG segments.

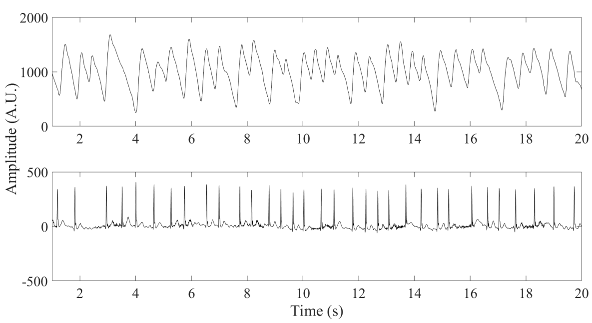

Figure 1. PPG and ECG signals for an AFib case.

2. Methods

2.1. Study Design

Inpatients with radiologically confirmed acute ischaemic stroke were recruited from intensive care unit (ICU) of UCSF medical center. Subjects provided written informed consent to protocols approved by the Institutional Review Board. Patients with acute ischaemic stroke, age higher than 18 years, and speaking English language were included. Patients with significant problems with attention, alertness, or cognitive function and inability to communicate were excluded from this study.

2.2. Data Collection

13 stroke patients participated in this study (age 19 to 91; median = 73.5), 6 patients with history of Afib. The ECG and PPG signals were acquired in continuous recordings using BedMasterEx (Excel Medical Inc, USA). Between 3h and 22h of recordings (median = 10.5h) were extracted at 240Hz sampling rate and were stored for offline analysis using $Matlab^{TM}$ tools (Mathworks Inc, USA). The signals were segmented into 30-sec strips (according to AFib guidelines [2]) without overlapping.

2.3. Annotation Process

In order to create a gold standard of signal quality assessment, four independent operators annotated the signals (8037 PPG recordings, 30-sec each) based on two classes: "Good" or "Bad". The classification as a good signal for a 30-sec segment of PPG should be based on three heuristic rules: the signal reflect the response of blood volume to the underneath pathophysiological characteristics of the cardiovascular system, irrespective of the particular shape of the pulse; artifacts-free; time aligned with ECG showing a correspondence for heart rate changes. The definition used in this work was very restrictive: all beats must be good quality for the 30-sec segment to be classified as good quality. In total, we analyze 8037 30-sec annotated segments. Cohen's kappa was determined in order to assess the inter-rater variability of the four annotators, using a small subset of 100 samples annotated by all operators. The other samples were annotated without having overlapped entries between annotators.

2.4. Machine Learning Approach

Features Extraction

The signals were parametrized by 40 features in the following subsets. Time-domain statistics: mean; median; standard deviation; variance; interquartile range; range; skewness; kurtosis; root mean square and entropy. Frequency domain statistics: first- to fourth-order moments in the frequency domain; median frequency; spectral entropy; total spectral power and peak amplitude in frequency band. Non-linear features, derived by the Poincare plot were used: SD1, standard deviation of the short-term beat to beat interval variability; the major axis SD2, the standard deviation of the long-term beat to beat interval variability and the SD1/SD2 ratio. Beat to beat analysis, were used four templates based on Gaussian waves to test the cross-correlation with each beat from the 30-sec segment, we determine the mean of the maxima list of cross-correlation results; standard deviation and range. Difference between beat to beat also were used, determined by the interquartile range for the differences of time domain statistics applied to each beat: mean; median; standard deviation; variance; interquartile range; range; skewness; kurtosis; root mean square and entropy. In beat to beat analysis, the mean of area under curve was determined; and the minimum period of a beat in the segment was used and maximum. The number of saturations in the segment (top and bottom of the signal) were also used as a feature. Due to the big difference in characteristics (amplitude and variation) of the feature components a normalization procedure was performed.

Two-class SVM Classifier

SVM classifier has been widely used in several biomedical problems, and is able to cope with aspects such as non-linearity and/or high-dimensionality of the physiological data [15], [16]. SVM classifier can distinguish two classes by finding a separating hyperplane with the maximal margin between two classes. The kernel function maps training data into a higher dimensional space. In this work, two different kernel functions were compared: linear and gaussian. In order to avoid overfitting and to test each model in a prospective setting, approximately 25% (15 to 35%) of the data samples was chosen as test set, which was never involved in the training phase. The remaining 75% (65 to 85%) of the data was used for learning the best model and determining the best parameters through 10-fold cross-validation, after being normalized in order to avoid within-subject differences in amplitude and variation among features. Training samples averaged values for each feature and corresponding standard deviation values were stored in order to normalize test feature sets, being therefore possible to map novel values into the training model features space. Performance analysis was conducted considering the accuracy (Ac), sensitivity (Sen) and specificity (Spe).

3. Results and Discussion

Agreement between four operators showed a kappa coefficient of 0.71 indicating substantial agreement. Annotated data showed that ratio bad/good signal by patients has a range between 0.02 (3 bad /137 good) to 16.7 (334/20), for the global dataset the ratio is 0.57 (2919 bad / 5118 good signals). Table 1 presents the results for the algorithms based on previously SQI that were tested with the annotated dataset. The performance of these algorithms is inferior to the results obtained with no AFib cases, as expected. Also, the algorithms were adapted and some thresholds were not described by the authors [6], in these cases the thresholds were empirically defined in order to optimize the performance. Only three algorithms showed a balance performance between sensitivity and specificity, [7], [12], [13]. The SQI used in these algorithms are based on a beat to beat analysis, and in the present work the objective is the segment classification, for this reason the final classification is good if 95% of the beats were considered of good quality.

Table 1. Performance of the SQIs, based on previous works, assessed with the current data.

SQI	Ac	Sen	Spe
A. Sukor [6]	0.8387	0.9760	0.5982
W. Karlen [7]	0.8677	0.8966	0.8171
S. Asgari [10]	0.6863	0.9945	0.1459
G. Clifford [11]	0.7429	0.9959	0.2994
C. Orphanidou [12]	0.8329	0.8015	0.8880
C. Liu [13]	0.7971	0.8134	0.7684
G. Papini [14]	0.4397	0.1360	0.9723

For the SVM classifier the performance results obtained for the models using different kernel functions are presented in the table 2, and showed higher performances for the classifier with the gaussian kernel.

4. Conclusions

Due to the prevalence of the AFib in the population, it is important to develop novel solutions that can properly assess the quality of the PPG signal when certain arrhythmia are present as in the case of atrial fibrillation. In this work, a new approach for robust PPG quality assessment based on SVM classification was proposed. While the classifier and features used in this work were used in previous studies, the key to the success of our approach is the adoption of a database that contains records of PPG signals recorded under AFib conditions and properly annotated. The distinction between waveforms with normal sinus rhythm and atrial fibrillation in the training of the classifier proved to improve the end performance of PPG quality assessment.

References

[1] P. Kirchhof, S. Benussi, D. Kotecha, A. Ahlsson, D. Atar, B. Casadei, M. Castella, H.-C. Diener, H. Heidbuchel, Hendriks, J. G. H. A. S. Manolis, J. Oldgren, B. A. Popescu, U. Schotten, B. V. Putte, and P. Vardas, “2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS,” *European Heart Journal*, vol. 37, no. 38, pp. 2893–2962, 2016.

[2] M. A. Christina Steger, Angelika Pratter, Monika Martinek-Bregel and C. S. Andreas Valentin, Jorg Slany, “Stroke patients with atrial fibrillation have a worse prognosis than patients without: data from the Austrian

Table 2. Performance of the model using random selection of the patients for the training and test.

			Gaussian			Linear		
Group	Patient	Percentage of Test	Ac	Sen	Spe	Ac	Sen	Spe
1	4, 5, 8, 10	26.91%	0.9464	0.9720	0.8721	0.9408	0.9683	0.8613
2	1, 2, 4, 11	22.15%	0.9315	0.9519	0.9126	0.9343	0.9332	0.9353
3	2, 7, 8, 12	19.55%	0.9542	0.9246	0.9789	0.9484	0.9176	0.9743
4	2, 5, 8, 10	27.24%	0.9529	0.9690	0.8998	0.9466	0.9714	0.8644
5	4, 8, 11, 12	25.58%	0.9314	0.8949	0.9521	0.9183	0.8720	0.9444
6	4, 6, 12, 13	30.22%	0.9510	0.9894	0.8631	0.9531	0.9882	0.8726
7	1, 2, 9, 10	32.29%	0.9595	0.9844	0.8876	0.9592	0.9813	0.8951
8	1, 6, 10, 12	28.06%	0.9570	0.9791	0.9052	0.9530	0.9715	0.9096
9	2, 4, 5, 11	25.91%	0.9299	0.9387	0.9207	0.9265	0.9453	0.9070
10	1, 4, 7, 8	16.59%	0.9355	0.9006	0.9712	0.9325	0.9036	0.9621
Mean	-	25.45%	0.9449	0.9505	0.9163	0.9413	0.9452	0.9126

Stroke registry," *European Heart Journal*, vol. 25, no. 19, pp. 1734 –1740, 2004.

[3] S. R. Steinhubl, R. R. Mehta, G. S. Ebner, M. M. Ballesteros, J. Waalen, G. Steinberg, P. Van Crocker, E. Felicione, C. T. Carter, S. Edmonds, J. P. Honcz, G. D. Miralles, D. Talantov, T. C. Sarich, and E. J. Topol, "Rationale and design of a home-based trial using wearable sensors to detect asymptomatic atrial fibrillation in a targeted population: The mHealth Screening to Prevent Strokes (mSToPS) trial," *American Heart Journal*, vol. 175, pp. 77–85, 2016.

[4] S. Kim, S. Im, and T. Park, "Characterization of quadratic nonlinearity between motion artifact and acceleration data and its application to heartbeat rate estimation," *Sensors*, vol. 17, no. 8, p. 1872, 2017.

[5] P. Gibbs and H. Asada, "Reducing motion artifact in wearable bio-sensors using MEMS accelerometers for active noise cancellation," *Proceedings of the 2005, American Control Conference*, pp. 1581–1586, 2005.

[6] J. A. Sukor, S. J. Redmond, and N. H. Lovell, "Signal quality measures for pulse oximetry through waveform morphology analysis," *Physiological Measurement*, vol. 32, no. 3, pp. 369–384, 2011.

[7] W. Karlen, K. Kobayashi, J. M. Ansermino, and G. A. Dumont, "Photoplethysmogram signal quality estimation using repeated Gaussian filters and cross-correlation," *Physiological Measurement*, vol. 33, no. 10, pp. 1617–1629, 2012.

[8] M. Pflugradt and B. M. R. Orglmeister, "OPRA: A fast on-line signal quality estimator for pulsatile signals," *IFAC-PapersOnLine*, vol. 28, no. 20, pp. 459–464, 2015.

[9] Y. Nam and Y. C. Nam, "Photoplethysmography signal analysis for optimal region-of-interest determination in video imaging on a built-in smartphone under different conditions," *Sensors (Switzerland)*, vol. 17, no. 10, 2017.

[10] F. Scalzo, P. Xu, S. Asgari, M. Bergsneider, and X. Hu, "Regression analysis for peak designation in pulsatile pressure signals," *Medical and Biological Engineering and Computing*, vol. 47, no. 9, pp. 967–977, 2009.

[11] Q. Li and G. D. Clifford, "Dynamic time warping and machine learning for signal quality assessment of pulsatile signals," *Physiological Measurement*, vol. 33, no. 9, pp. 1491–1501, 2012.

[12] 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 Journal of Biomedical and Health Informatics*, vol. 19, no. 3, pp. 832–838, 2015.

[13] C. Liu, Q. Li, and G. Clifford, "Evaluation of the accuracy and noise response of an open-source pulse onset detection algorithm on pulsatile waveform databases," *Computing in Cardiology*, vol. 43, pp. 913–916, 2016.

[14] G. B. Papini, P. Fonseca, X. L. Aubert, S. Overeem, J. W. Bergmans, and R. Vullings, "Photoplethysmography beat detection and pulse morphology quality assessment for signal reliability estimation," *Proceedings Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 117–120, 2017.

[15] J. Paiva, J. Cardoso, and T. Pereira, "Supervised learning methods for pathological arterial pulse wave differentiation: A SVM and neural networks approach," *International Journal of Medical Informatics*, vol. 109, 2018.

[16] S. M. Shan, S. C. Tang, P. W. Huang, Y. M. Lin, W. H. Huang, D. M. Lai, and A. Y. A. Wu, "Reliable PPG-based algorithm in atrial fibrillation detection," in *Proceedings - 2016 IEEE Biomedical Circuits and Systems Conference, BioCAS 2016*, 2016, pp. 340–343.

Address for correspondence:

Tania Pereira
Department of Physiological Nursing, School of Nursing
University of California, San Francisco
2 Koret Way, San Francisco CA 94143, USA
tania.pereira@ucsf.edu