2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Cuff-less blood pressure estimation using wrist photoplethysmography

M. Pediaditis, E. G. Spanakis. G. Zacharakis, and V. Sakkalis, *Member, IEEE*

Abstract— **One of the most promising and at the same time rapidly growing sectors in healthcare is that of wearable medical devices. Population ageing constantly shifts towards a higher number of senior and elderly people with increased prevalence of chronic diseases often requiring long-term care and a need to decrease hospitalization time and cost. However, today most of the devices entering the market are not standardized nor medically approved, and they are highly inaccurate. In this work we present a system and a method to provide accurate measurement of systolic and diastolic blood pressure (BP) based solely on wrist photoplethysmography. We map morphological features to BP values using machine learning and propose ways to select high quality signals leading to an accuracy improvement of up to 33.5%, if compared against no signal selection, a mean absolute error of 1.1mmHg in a personalized scenario and 8.7mmHg in an uncalibrated leave-one-out scenario.**

I. Introduction

Photoplethysmography (PPG), as a method for non-invasive and low-cost sensing of vital signs, plays a significant role in wearable medical devices, meant to be used for continuous health monitoring systems. It is a well-established technique, already in use by a variety of wearable devices (e.g., smart bands, smart rings, smartphones) for the acquisition of different vital signs such as heart rate and pulse rate variability. Of great importance is the blood pressure (BP) estimation both for diagnosed patients at risk from cardiovascular diseases (CVDs) but also for healthy individuals as an early indicator of hypertension. In Europe CVDs are accused for almost four million deaths each year being the leading cause of mortality under 65 years [1].

Even if modern wearable devices offer the possibility of estimating BP, still the "golden standard" lies in traditionally used cuff-based approaches, mainly due to highly inaccurate measurements. Accuracy of such vital signs highly depends on the quality of the acquired signal and the presence/absence of artifacts generated by other sources such as motion, ambient light, sensor position etc. Especially PPG acquisition at the wrist is even more susceptible to noise due to poor sensor-to-skin contact as a result of arm movements. In addition, the personalization factor (calibration) is reported to be rather critical [2]. It should be stressed though that "cuff-less" wearables offer the unique capability of performing continuous measurement seamlessly throughout the day, offering a plurality of measurements while alleviating the *white coat syndrome* that can greatly influence the outcome.

Matthew Pediaditis, Emmanouil G. Spanakis, Giannis Zacharakis and Vangelis Sakkalis are with Traqbeat Technologies PC, Heraklion, Crete, Greece, 71305 GR ({pediaditis, spanakis, zacharakis, sakkalis} @traqbeat.com).

Conventionally, both ECG and PPG, or two synchronized PPGs in distant locations are needed in order to extract pulse wave velocity (PWV) and pulse transit time (PTT) further used as the main features in (BP) estimation. However, such approaches while adding to the cost, complexity and usability of designing such devices, they do not seem to achieve highly reliable measurements. A recent review examined the latest developments in terms of cuff-less BP estimation identifying the wrist to be a challenging body position for PPG acquisition [3]. This ten-year review in fact, identified very few (only 3) papers focusing on this type of signals for measuring BP and only one, Atomi et al. [4], presented a cuff-less BP estimation method using a wristwatch-type PPG sensor by applying multiple linear regression on features extracted from the PPG and its 2nd derivative. Besides not using the same features and ML methods, the main difference with our work is that only systolic blood pressure (SBP) was estimated. A further study also presented a smartwatch for accurately estimating BP in real time [5]. Their device, in comparison with ours, combines two pulse oximeters to collect two PPG (wrist & finger) signals, filter and cross-corelate them in order to obtain a PTT and then use a linear model to give an estimation of SBP and DBP (diastolic BP). Another relevant study of Sasso et al. [6] used a commercial sensor and estimated SBP and DBP using feature extraction and ML methods with comparable results in terms of accuracy to Atomi et al. [4].

Considering the assessment of PPG signal quality, Elgendi [7] tested eight quality metrics including perfusion, kurtosis, skewness, relative power, non-stationarity, zero crossing rate, entropy, and the matching of systolic wave detectors to differentiate brutally between excellent, acceptable and unfit PPG waveforms, labeled as such by three experts. As a result, skewness outperformed all other metrics in terms of differentiating between excellent PPG and acceptable. The work in [8] presented another quality index to separate bad from good signals using adaptive template matching as a measure to assess reliable heart rate obtained both from ECG and PPG. The work in [9] identified six morphological features in order to determine the quality index of a PPG waveform (positive/negative peak-to-peak components), using machine learning.

In this work we present a method to accurately estimate systolic and diastolic BP, based on wrist PPG only, using morphological features and machine learning (ML) methods. Furthermore, we investigate the possibility of reducing the estimation error by incorporating a combination of PPG signal quality metrics in order to remove erroneous signals.

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

II. METHODS

To capture optimal signal at wrist, we designed and developed a miniaturized sensor on a fully customized board, BLE communication, μ-Optoelectronics, multicolor LED sources, OEM optics and custom housing [10].

A. Wrist PPG sensor

Our reflective PPG sensor consists of two green (517nm) and two IR LEDs (950nm) and one large area photodiode (BPW34S). A lens positioned between the photodiode and the skin probes a larger tissue volume by efficiently coupling the reflected light. The LEDs are arranged in pairs around the Photodiode, offering a uniform tissue illumination. The distance between the light sources and the photodiode was optimized using light propagation models and a series of experiments to gain the best perfusion index. An analog-to-digital front-end (TI AFE4400) is used to drive the LEDs and convert the photodiode current at a sampling rate of 500Hz. The gain and timing parameters were optimized to achieve optimal signal to noise ratio. The captured signals are transferred in real-time to a mobile device or a PC using BLE 4.2. The sensor is battery powered and was integrated into a common wristwatch to model every-day usage (Figure 1).

B. PPG signal acquisition and SBP, DBP measurement

The PPG was acquired on the left wrist along with the SBP, DBP and pulse rate (PR) using the A&D UA-767 Plus BT-C medical grade cuff-based digital blood pressure monitor on the same arm and not on the opposite, since blood pressure deviation between both sides is common [11]. Acquisition was based on a simplified version of the protocol described in [12]. We first measured the blood pressure and heart rate, then loosened the cuff, waited one minute for the vasculature to relax and then measured the green PPG. Subsequently we measured the BP&PR again, loosened the cuff, waited for one minute and measured the IR PPG. Finally, a last BP&PR measurement was taken. The mean of those three measurements was taken as a reference for the evaluation. Both green and IR measurements consisted of 33 epochs each of 10s in duration. The left arm was resting on the surface of a desk, while the subject was sitting. In total 11 volunteers were measured 8 times in average over a few days. The study protocol complies with the Declaration of Helsinki. Informed consent was obtained from all participants. All procedures included anonymization processes to ensure that the study is executed in strict compliance with EU GDPR directive.

Figure 1, miniaturized wearable wrist PPG sensor.

C. Pre-processing

Ambient light measurements immediately captured after each LED illumination phase, were subtracted. Next, the DC component from each epoch was removed (but stored for the calculation of the perfusion index) and the signal was inverted and de-trended, followed by band-pass filtering (0.5-5Hz, 4th order). For each epoch the 1st and 2nd derivatives were computed and all signals were split into separate pulses, based on the absorption peaks (valleys) of the PPG.

D. Feature extraction

For each epoch, the pulse rate was calculated using the mean peak to peak time, as well as using a common FFT-based method. In addition, for each pulse a series of features that are based on key morphological elements, such as the dicrotic notch, pulse onset foot and zero crossings of the derivatives were extracted (TABLE I.) [10]. Pulses for which all features could not be extracted were rejected. Our dataset consisted of 20639 pulses in total.

TABLE I. LIST OF EXTRACTED FEATURES

F1a	Time interval between emission peak and dicrotic notch
F1b	Time interval between dicrotic notch and the following absorption peak (end of pulse)
F2a	Amplitude of the emission peak
F2b	Time interval between the foot and the emission peak
F3a	Amplitude of the absorption peak (end of pulse)
F3b	Time interval between the beginning and the end of the pulse (absorption peaks)
F4a	Time interval between foot and dicrotic notch
F4b	Amplitude of dicrotic notch
F5a	Pulse rate based on FFT
F5b	Pulse rate based on time between emission peaks
F5c	Pulse rate based on time between absorption peaks
F6	Time interval between max. positive peak of 1st derivative and the dicrotic notch
F7	Time interval between max. and min. peaks of 2nd derivative
F8a	Amplitude ratio of emission over absorption peak
F8b	Ratio of emission over absorption peak location (time)

E. Quality metrics

The following quality metrics were considered to further clean the input data by pre-selecting epochs containing all necessary morphological features:

- Signal-to-noise ratio (SNR), calculated relative to a real-valued sinusoidal, based on the periodogram of the signal.
- Skewness (SKW), a measure of the symmetry of a probability distribution, and has been found to be associated with corrupted PPG signals [13].
- Kurtosis (KRT), a statistical measure used to describe the distribution of observed data around the mean values, used for PPG signal quality assessment in [7].
- Zero crossing rate (ZCR), the rate of sign-changes in the signal, used for PPG signal quality assessment in [7].
- Short-time energy (STE), the energy of subsequent short signal segments, which can be used to describe abrupt changes in the signal, e.g. related to motion.
- RMS of AC component (actual pulse), describing the effect of the pulse wave to the propagating light attenuation through the probed tissue.
- Perfusion index (PERF), the ratio of the pulsatile blood flow to the non-pulsatile or static blood in peripheral tissue, one of the most important features for assessing PPG signal quality [14].

F. Analysis and validation

We examined the suitability of five regression methods in our attempt to estimate both the systolic and diastolic BP using green and IR wavelengths. We tested random forest

regression (RFR) without feature standardization, since it is a tree-based model. With feature standardization we tested SVM regression with three different kernels (rbf, linear, poly), as well as multi-layer perceptron (MLP) regression. Implementations were based on Python Scikit-learn package [15]. Three methods of validation were employed:

i) Random 10-fold cross-validation on the whole dataset. This approach was used to select the regression method, tune its hyper-parameters and finally select which quality metrics to use for rejecting epochs that do not provide the necessary information for accurate BP estimation. This was done by first training the regressor on the features of all pulses and then analyzing the errors of all pooled validation folds. The quality metrics of the outliers (using 25th and 75th percentiles) were then analyzed against those of the non-outliers to find any statistical difference between the distributions. Based on this information we selected a combination of quality metrics that can be used along with a threshold to filter out epochs that would return extreme errors;

ii) *Per subject train/test split.* Based on the results of the previous task we performed this split (66% train, 33% test) to test our method on a calibrated-like scenario, since data from the same subject were used for both training and testing; and

iii) *Leave-one-subject-out cross-validation.* This represents the least biased method to assess the performance in a non-calibrated scenario, where the data of the test subject has not been used for training at all.

III. Results

TABLE II. shows that RFR performed better than the others. Tuning of the hyper-parameters for RFR was achieved by performing an initial random search, followed by three cycles of grid search, gradually narrowing the search space, and concluding in the following: *n_estimators=300, max_features='auto', max_depth=30, min_samples_split=3, min_samples_leaf=1, bootstrap=True*.

TABLE II. Mean absolute error (mmHg) for each regression method (random 10-fold cross-validation)

Regression method	SBP Green	SBP IR	DBP Green	DBP IR
RFR	3.92	3.15	4.23	3.53
SVR (rbf)	5.91	5.69	6.38	6.61
SVR (lin)	6.97	6.54	7.39	8.16
SVR (poly)	16.76	7.64	17.04	7.85
MLP	5.59	5.36	5.87	6.17

We subsequently performed the analysis of the quality metrics of the regression error outliers (group A) against those of the remaining instances (group B). As a first step we tested the distribution of each quality metric, in each group, for normality using D'Agostino and Pearson's test [16]. Since all distributions resulted in not being normally distributed (alpha=0.01), we used the Kruskal-Wallis H-test to test for equality of the median of both groups. This resulted in the rejection of the null-hypothesis in all cases except for ZCR related to DBP estimation. Therefore, we analyzed the effect of removing "non-ideal" epochs based on the quality metrics proposed as the main criterion in terms of mean absolute error (MAE) reduction. As a threshold we used the median of group A (outliers). This was necessary because there was an overlap between the distributions of both groups and setting the threshold too close to the non-outlier median would result in removing too many epochs (c.f. Figure 2).

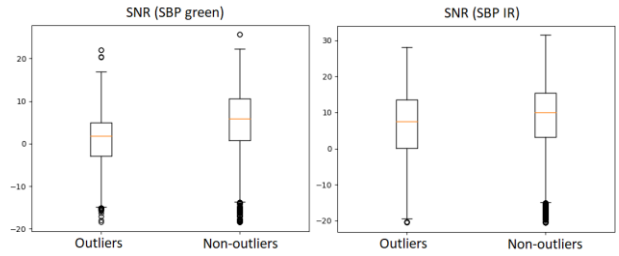

Figure 2, SNR distribution of the regression error outliers and non-outliers in estimating SBP for the green and IR colors, as an example.

The outcome in terms of reduced error is depicted in Figure 3. In order to assess the negative impact of removing epochs in terms of failed measurements we plotted the ratio of % removed pulses to % MAE reduction (Figure 4). Finally, based on these results we selected an optimal combination of the best performing quality metrics to further enhance the rejection performance. We combined the quality metrics with the lowest ratio according to Figure 4 with the logical "or" operator.

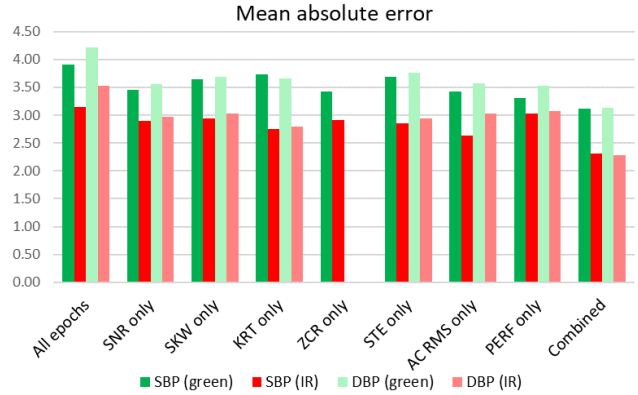

Figure 3, Mean absolute error (mmHg) as a result of removing erroneous epochs based on single quality metrics and a combination of them. The first set of bars shows the error without any removal.

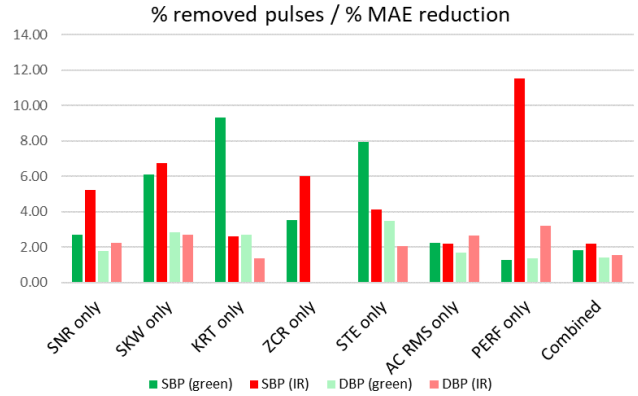

Figure 4, Ratio of % removed pulses to % mean absolute error reduction by either using one quality metric or a combination of multiple.

The results shown in the last set of columns of both previous Figures result from the following combinations: {SNR or AC RMS or PERF} for SBP green, {KRT or AC

RMS} for SBP IR, {SNR or AC RMS} for DBP green and {KRT or AC RMS} for DBP IR.

System evaluation was performed using a per-subject 10-fold cross-validation (personalized scenario) with, and without removing erroneous epochs (TABLE III.). The results from the final step, the leave-one-subject-out cross-validation are shown in TABLE IV.

TABLE III. MEAN ABSOLUTE ERROR (MMHG), PER SUBJECT 10-FOLD CROSS-VALIDATION. BEFORE AND AFTER REMOVING ERRONEOUS EPOCHS BASED ON A COMBINATION OF QUALITY METRICS.

	SBP Green	SBP IR	DBP Green	DBP IR
All epochs	1.33	1.20	1.26	0.98
Cleaned epochs	1.05	0.88	1.00	0.65
% reduced error	21.1	26.2	20.4	33.5

TABLE IV. MEAN ABSOLUTE ERROR (MMHG), LEAVE-ONE-SUBJECT-OUT CROSS-VALIDATION. BEFORE AND AFTER REMOVING ERRONEOUS EPOCHS BASED ON A COMBINATION OF QUALITY METRICS.

	SBP Green	SBP IR	DBP Green	DBP IR
All epochs	10.13	8.18	11.44	10.32
Cleaned epochs	7.65	7.97	9.45	9.81
% reduced error	24.5	17.4	2.6	5.0

IV. DISCUSSION

All quality metrics, if used independently for selecting erroneous epochs provide an improvement in error reduction, although some metrics reject too many epochs and pulses consequently (e.g., Figure 4, KRT for SBP green and PERF for SBP IR). By selecting a combination of quality metrics, we achieve a further MAE reduction, as well as a low percentage of rejected pulses. When applied on the personalized (calibrated) BP estimation scenario, an error reduction of up to 33.5% is achieved. In the leave-one-subject-out (uncalibrated) case we reach 24.5% less error for SBP although DBP shows less improvement, possibly reflecting the well-reported fact that the estimation of DBP from PPG is more challenging than SBP.

IR provides better results in all cases, while green is only marginally better in the non-calibrated cleaned data. Hence, our method of removing erroneous epochs is more efficient in the case of green PPG signal, which is most commonly used in the field.

Accuracy, in the existing literature, is most commonly calculated using the mean absolute error or the root mean square error as compared to clinically validated cuff-based BP monitoring devices adhering to standards setting the mean error to 5mmHg at most [17]. Our results for the calibrated scenario are well below this value. Considering the uncalibrated case, the MAE is comparable to that of Atomi et al. [4] and Sasso et al. [6]. Atomi et al. report a mean (not absolute) error of 1.58 mmHg (SBP only) after manual elimination of signals with measurement artifacts. Based on a provided figure one can estimate an MAE of 7.4 mmHg. Sasso et al. report a MAE of 8.79 mmHg for the SBP and 6.37 mmHg for the DBP in an uncalibrated case, with manual signal selection and longer time windows, as compared to our approach. Our MAE of 7.65 (SBP green) and 9.45 (DBP green) is comparable with these studies considering the respective accuracy limits. However, an advantage of our method is that selection of high-quality signals can be performed automatically. Moreover, our calculations are based on short 10s-epochs, a duration that is well below the average, even for finger PPG [3].

V. CONCLUSION

A method to accurately estimate SBP and DBP using IR and green PPG signals, is presented. Our approach provides an automatic mechanism for high-quality PPG epoch selection, based on a combination of different quality metrics.

REFERENCES

[1] A. Timmis, N. Townsend, C. P. Gale et al., "European society of cardiology: Cardiovascular disease statistics 2019," *Europ. Heart J*., vol. 41, no. 1, pp 12–85, 2020.

[2] W. Chen, T. Kobayashi, S. Ichikawa et al., "Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent calibration," *Med. Biol. Eng. Comput*., vol. 38, pp. 569–574, 2000.

[3] M. Hosanee, G. Chan, K. Welykholowa, et al., "Cuffless single-site photoplethysmography for blood pressure monitoring," *J. Clin. Med.*, vol. 9, no. 723, 2020.

[4] K. Atomi, H. Kawanaka, S. Md. Bhuiyan and K. Oguri, "Cuffless blood pressure estimation based on data-oriented continuous health monitoring system," *Comput. Math. Methods Med*., 2017.

[5] R. Lazazzera, Y. Belhaj and G. Carrault, "A new wearable device for blood pressure estimation using photoplethysmogram," *Sensors*, vol 19, no. 2557, 2019.

[6] A. M. Sasso, S. Datta, M. Jeitler et al., "HYPE: Predicting blood pressure from photoplethysmograms in a hypertensive population," in: Michalowski M., Moskovitch R. (eds) Artif. Intell. Med. (AIME), LNCS, vol. 12299, 2020.

[7] M. Elgendi, "Optimal signal quality index for photoplethysmogram signals," *Bioengineering*, vol. 3, no. 4, p. 21, 2016.

[8] C. Orphanidou, T. Bonnici, P. Charlton et al., "Signal-quality indices for the electrocardiogram and photoplethysmogram: Derivation and applications to wireless monitoring," *in IEEE J. Biomed. Health Inform*., vol. 19, no. 3, pp. 832–838, 2015.

[9] E. Sabeti, N. Reamaroon, M. Mathis et al., "Signal quality measure for pulsatile physiological signals using morphological features: Applications in reliability measure for pulse oximetry," *Inform. Med. Unlocked*, vol. 6, 2019.

[10] V. Sakkalis, M. Pediaditis, E. G. Spanakis, I. Zacharakis, "Wrist-watch back cases, systems and methods for measuring biomarkers," U.S. Patent 20200405233A1, Dec. 31, 2020.

[11] I. Weinberg, P. Gona, C. J. O'Donnell et al., "The systolic blood pressure difference between arms and cardiovascular disease in the Framingham Heart Study," *Am. J. Med.*, vol. 127, no. 3, pp. 209–215, 2014.

[12] G. S. Stergiou, et al., "A universal standard for the validation of blood pressure measuring devices: Association for the Advancement of Medical Instrumentation/European Society of Hypertension/ International Organization for Standardization (AAMI/ESH/ISO) Collaboration Statement," *Hypertension*, vol. 71, no. 3, pp. 368–374, 2018.

[13] R. Krishnan, B. Natarajan and S. Warren, "Two-Stage Approach for Detection and Reduction of Motion Artifacts in Photoplethysmographic Data," in *IEEE Trans. Biomed. Eng*., vol. 57, no. 8, pp. 1867–1876, 2010.

[14] H. Gehring, et al., "The effects of motion artifact and low perfusion on the performance of a new generation of pulse oximeters in volunteers undergoing hypoxemia," *Respir. Care,* vol. 47, pp. 48–60, 2002.

[15] F. Pedregosa, G. Varoquaux, A. Gramfort et al., "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[16] R. D'Agostino, and E. S. Pearson, "Tests for departure from normality", *Biometrika*, vol. 60, pp. 613–622, 1973.

[17] A. Coleman, S. Steel, P. Freeman et al., "Validation of the Omron M7 (HEM-780-E) oscillometric blood pressure monitoring device according to the British Hypertension Society protocol," *Blood. Press. Monit.*, vol. 13, no. 1, pp. 49–54, 2008.