Continuous blood pressure estimation from PPG signal

Gašper Slapničar
Jožef Stefan Institute
Jamova cesta 39
1000 Ljubljana
gasper.slapnicar@ijs.si

Matej Marinko
Faculty of Math. and Physics
Jadranska cesta 19
1000 Ljubljana
matejmarinko123@gmail.com

dr. Mitja Luštrek
Jožef Stefan Institute
Jamova cesta 39
1000 Ljubljana
mitja.lustrek@ijs.si

ABSTRACT
Given the importance of blood pressure (BP) as a direct indicator of hypertension, regular monitoring is encouraged in people and mandatory for such patients. We propose an approach where photoplethysmogram (PPG) is recorded using a wristband in a non-obtrusive way and subsequently BP is estimated continuously, using regression methods based solely on PPG signal features. The approach is validated using two distinct datasets, one from a hospital and the other collected during every-day activities. The best achieved mean absolute errors (MAE) in a Leave-one-subject-out experiment with personalization are as low as 11.87 ± 12.31 / 11.09 ± 9.99 for systolic BP and 5.64 ± 5.73 / 6.18 ± 4.85 for diastolic BP.

Keywords
Photoplethysmography, blood pressure, regression analysis, m-health

1. INTRODUCTION
According to the World Health Organization (WHO), cardiovascular diseases were the most common cause of death in 2015, responsible for almost 15 million deaths combined [1]. Hypertension is a common precursor of such diseases and can be easily detected with regular blood pressure (BP) measurements.

Given the importance of BP, people should actively monitor its changes. This is not trivial as the traditional BP measurement method involves an inflatable cuff and a stethoscope, which should be placed directly above the main artery at approximately heart height. These requirements impose relatively strict movement restrictions on the patient and require substantial time commitment. Furthermore, when done by the patient himself, the process can cause stress, which in turn influences the BP values, so it is most commonly done by medical personnel. However, when BP is measured by medical personnel, this can again cause anxiety in the patient, commonly known as white coat syndrome.

Our work focuses on analyzing the photoplethysmogram (PPG) and then developing a robust non-obtrusive method for continuous BP estimation, which will be implemented and used in one such m-health system, based on a wristband with a PPG sensor.

2. RELATED WORK
Photoplethysmography is a relatively simple and non-expensive technique, which is becoming increasingly popular in wearables for heart rate estimation. Exploring its applications, we can see that it is also becoming more widely used in BP estimation.

PPG is based on illumination of the skin and measurement of changes in its light absorption. It requires a light source (typically a light-emitting diode – LED light) to illuminate the tissue (skin), and a photodetector (photodiode) to measure the amount of light either transmitted or reflected to the photodetector. Thus, PPG can be measured in either transmission or reflectance mode.

With each cardiac cycle the heart pumps blood towards the periphery of the body, thus producing a periodic change in the amount of light that is absorbed or reflected from the skin, as the skin changes its tone based on the amount of blood in it [6]. An example of this periodic signal as produces by Empatica E4 wristband [7] is shown in Figure 1.

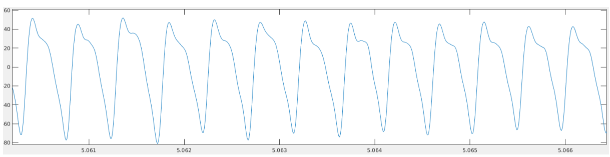

Figure 1: An example PPG signal as produced by Empatica E4 wristband.

It is being used in two common approaches: *1.)* BP estimation from two sensors (PPG + Electrocardiogram (ECG)) and *2.)* BP estimation using PPG only.

The first approach suggests the use of two sensors, typically an ECG and a PPG sensor, in order to measure the time it takes for a single heart pulse to travel from the heart to a peripheral point in the body. This time is commonly known as pulse transit time (PTT) or pulse arrival time (PAT) and its correlation with BP changes is well established [2].

The more recent approach is focused on PPG signal only, however the relationship between PPG and BP is only postulated and not well established, unlike the relationship between PTT and BP. This approach is the least obtrusive by far and PPG sensors have recently become very common in most modern wristbands.

One of the earliest attempts at this approach was conducted by Teng et. al. [3] in 2003. The relationship between arterial blood pressure and certain features of the photoplethysmographic (PPG) signals was analyzed. Data was obtained from 15 young healthy subjects in a highly controlled laboratory environment, ensuring constant temperature, no movement and silence. The mean differences between the linear regression estimations and the measured BP were 0.21 mmHg for SBP and 0.02 mmHg for DBP. The corresponding standard deviations were 7.32 mmHg for SBP and 4.39 mmHg.

A paper was published in 2013 in which authors used data from Multiparameter Intelligent Monitoring in Intensive Care (MIMIC) waveform database [4] to extract 21 time domain features and use them as an input vector for artificial neural networks (ANNs). The results are not quite as good as the linear regression model described earlier, however the data is obtained from a higher number and variety of patients in a less controlled environment, but was still measured in a hospital setting and an undisclosed subsample of all available data was taken. The results reached mean absolute difference between the estimation and the ground truth of less than 5 mmHg with standard deviation of less than 8 mmHg [5].

It is clear that the PPG only approach has potential, however a robust method that works well on a general case is yet to be developed.

3. METHODOLOGY

The workflow consists of two main parts, namely the signal pre-processing and machine learning part. In signal pre-processing, our PPG signal is cleaned of most noise and segmented into cycles, where one cycle corresponds to a single heart beat. Afterwards, features are extracted on per-cycle basis and fed into regression algorithms which build models that are further evaluated.

3.1 Signal pre-processing

When PPG is used in a wristband, the main problem comes from the contact between the sensor and the skin. During everyday activity, the patient moves his arm a lot, which in turn causes substantial movement artefacts in the signal. This is partially alleviated by the usage of green light, which is less prone to artefacts, however pre-processing is still required.

3.1.1 Cleaning based on established medical criteria

In first phase, both BP and PPG signal are roughly cleaned based on established medical criteria [9]. During this phase, parts of signals with systolic BP (SBP) > 280mmHg or diastolic BP (DBP) < 20mmHg or the difference between SBP and DBP < 20mmHg, are removed. This removes parts of signals for which the reference BP signal most likely contained an anomaly as such values indicate extreme medical condition and are not feasible in a common patient.

3.1.2 Peak and cycle detection

In order to do further cleaning and feature extraction, PPG cycle detection is mandatory. This is again not trivial, as substantial noise in the PPG signal poses a significant problem.

A slope sum function, which enhances the abrupt upslopes of pulses in the PPG signal is first created. Afterwards, a time-varying threshold for peak detection is applied [8]. After the peaks are detected, finding the cycle start-end indices is rather simple as the valleys between peaks must be found. An example of detected peaks and cycle locations is shown in Figure 2.

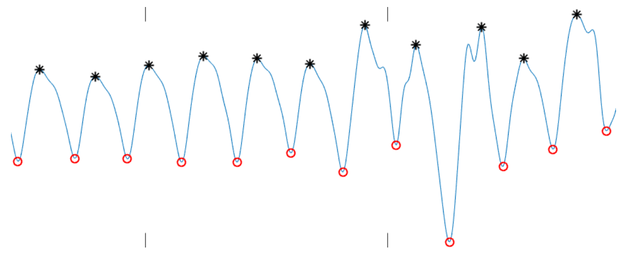

Figure 2: An example output of peak/cycle detection algorithm. Black asterisks correspond to a detected peak while red circles correspond to a detected cycle beginning.

Once cycles are detected, they are used for further cleaning and feature extraction.

3.1.3 Cleaning based on ideal templates

In the second cleaning phase, a sliding window of 30 seconds is taken and the mean of all cycles within this window is computed from the PPG signal. Presuming that the majority of cycles within a 30sec window are not morphologically altered, a good "ideal cycle template" is created. Each individual cycle is then compared to this ideal template and its quality is evaluated with three signal quality indices (SQIs). The most likely length of cycle L is always determined with autocorrelation analysis. The template is computed by always taking L samples of each cycle in the current window.

Signal quality indices are computed as follows. *SQI1*: First L samples of each cycle are taken, and each cycle is directly compared to the template using a correlation coefficient. *SQI2*: Each cycle is interpolated to length L and then the correlation coefficient is computed. *SQI3*: The distance between template and cycle is computed using dynamic time warping (DTW).

Finally thresholds for each SQI are determined and if more than half cycles in the given 30sec window are discarded, the whole window is considered too noisy and thus removed. Example of this cleaning is shown in Figure 3.

Once high quality signal is obtained, features can be extracted from each cycle.

3.2 Machine learning

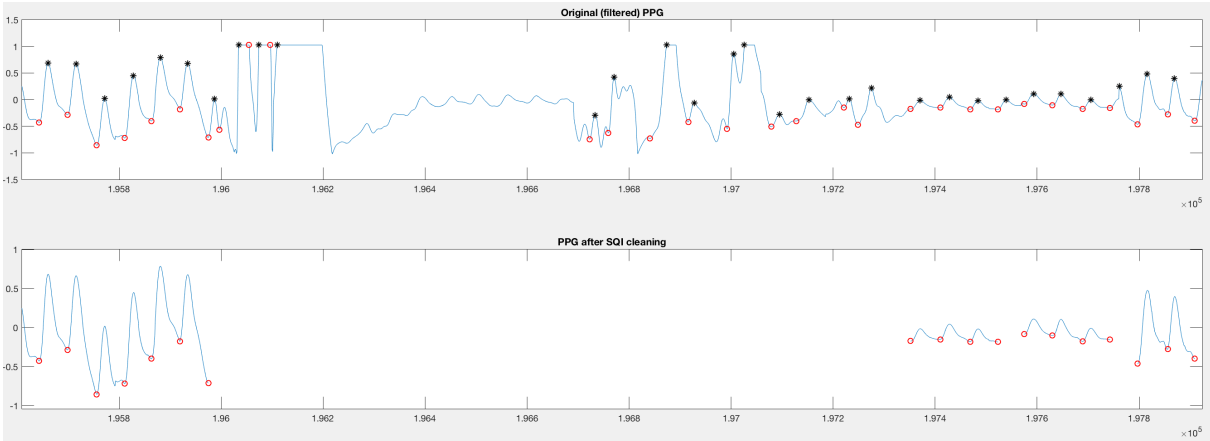

Figure 3: An example of the cleaning algorithm in the 2nd phase. Comparing the top (uncleaned) and bottom (cleaned) signal, we see that the obvious artefact period is discarded.

In accordance with related work [5] several time domain features were computed and the set of features was further expanded with some from the frequency domain [9]. These are shown in Figure 4.

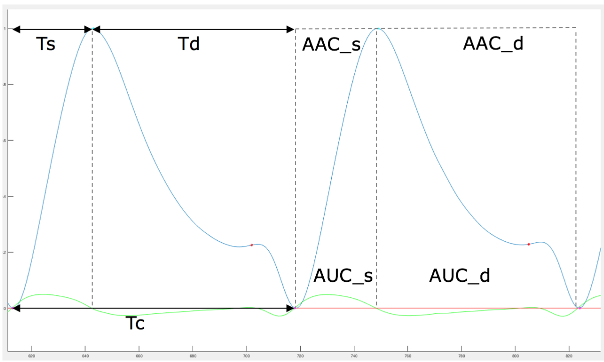

Figure 4: Time domain features which were used. Tc = cycle time, Ts = systolic rise time, Td = diastolic fall time, AAC = area above the curve and AUC = area under the curve for systolic and diastolic part of a cycle.

These features were extracted for each cycle and used in machine learning to derive a regression model for BP estimation.

4. EXPERIMENTS AND RESULTS

In an effort to make our method as general as possible, two datasets were considered for our experiment and all the data which had both PPG and BP signal were used.

4.1 Data

First is the publicly accessible MIMIC database set from which all the patients having both PPG and arterial BP (ABP) signal were taken. This results in 50 anonymous patients, each having on average several hours of both signals available. The data was collected in a hospital environment, using the hospital equipment.

Second is a dataset collected at Jozef Stefan Institute (JSI) using the Empatica E4 wristband for PPG and an Omron cuff-based BP monitor for the ground truth BP. In the first phase of data collection, 8 healthy subjects were considered, 5 male and 3 female. Each wore the wristband for several hours during every-day activities and measured their BP every 30min or more often. Finally, only parts of signals 3 minutes before and after the BP measurement were taken into consideration.

4.2 Experimental setup

Leave-one-subject-out experiment was conducted on each dataset, as it is the most suitable experiment to evaluate the generalization performance of the algorithms. Due to time and computational power restrictions, data was subsampled by taking 500 uniformly selected cycles.

During the initial attempt, a regression model was trained in each iteration on all subjects, except the left out. This yielded poor results, hinting at the fact, that most patients are unique in some way. This was confirmed by doing a cycle morphology analysis during which it was established that different subjects have different cycle shapes and that similar cycle shapes do not signify similar BP values. Thus, personalization of the trained models was considered.

In the second attempt, the regression models were again trained using all subjects except the left out, however they were further personalized using some data instances from the left out subject. The instances of the left out subject were first grouped by BP values and these groups were then sorted from lowest to highest BP. Afterwards, every *n-th* group ($n = 2,3,4,5,6$) of instances was taken from the testing data and used in training in order to personalize the model to the current patient. This ensures personalization with different BP values, as taking just a single group of instances gives little information, since the BP will be constant within this group. Given the fact that MIMIC data consists of roughly 5x the amount of patients compared to JSI collected data, the personalization data for it was multiplied 5 times, making it noticable within the large amount of training data from the remaining patients.

During both attempts, several regression algorithms were considered, as given in Figures 5 and 6. Mean Absolute Error (MAE) was used as a metric. All models were compared with a dummy regressor, which always predicted the mean BP value of the same combination of general and personalization data as the other models used to train themselves. Finally, the regressor with the lowest MAE was chosen.

4.3 Results

Due to low variations in BP, the dummy regressor often performs relatively well, however for MIMIC data with more BP variation, some improvements have been made as shown in Figure 5. The JSI collected data has proven to be more problematic, as there are only a low amount of different BP values in this phase of collection.

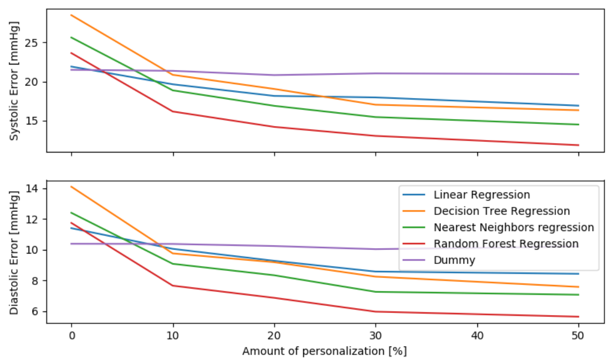

Figure 5: MAE for SBP and DBP for MIMIC dataset at different amounts of personalization.

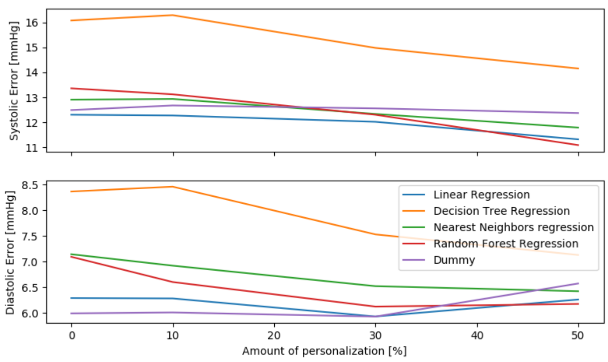

Figure 6: MAE for SBP and DBP for JSI collected dataset at different amounts of personalization.

The lowest error using MIMIC data was achieved by using the RandomForest regression algorithm, with the highest amount of personalization. The achieved errors were 11.87 ± 12.31 for SBP and 5.64 ± 5.73 for DBP.

Due to high amount of movement artefacts in JSI collected data, a lot of data was removed by the cleaning algorithm, leaving a very low amount of usable data with very low variations in BP. This further enhanced the performance of dummy regressor, while leaving little information for other algorithms. Best achieved errors of 11.09 ± 9.99 for SBP and 6.18 ± 4.85 for DBP are only slightly surpassing the mean predictions at maximum personalization, as shown in Figure 6.

5. CONCLUSION

We have developed a pipeline for BP estimation using PPG signal only and have evaluated its performance on two distinct datasets.

First part of the pipeline does signal pre-processing, removing most movement artefacts and detecting PPG cycles. The second part computes features on per-cycle basis and feeds them in regression algorithms. These were evaluated on hospital collected MIMIC database data as well as field collected data at JSI using a wristband. Due to low variations in subject's BP and high variation in their PPG, there is limited information about the correlation between the two, however promising results were obtained with best achieved mean absolute errors (MAE) in a Leave-one-subject-out experiment with personalization as low as 11.87 ± 12.31 / 11.09 ± 9.99 for systolic BP and 5.64 ± 5.73 / 6.18 ± 4.85 for diastolic BP.

Acknowledgement

The HeartMan project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 689660. Project partners are Institut Jozef Stefan, Unversita Degli Studi Di Roma La Sapienza, Universiteit Gent, Consiglio Nazionale Delle Richerche, ATOS Spain SA, SenLab, KU Leuven, MEGA Elektronikka Oy and European Heart Network.

6. REFERENCES

[1] The World Health Organization. "The top 10 causes of death", 2015.
[2] Geddes et. al. "Pulse transit time as an indicator of arterial blood pressure", 1981.
[3] Teng et. al. "Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach", 2003.
[4] Goldberger et. al. "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals", 2000.
[5] Lamonaca et. al. "Application of the Artificial Neural Network for blood pressure evaluation with smartphones", 2013.
[6] Allen. "Photoplethysmography and its application in clinical physiological measurement", 2007.
[7] Empatica Inc. "Real-time physiological signals - E4 EDA/GSR sensor". Accessed at `https://www.empatica.com/e4-wristband` in 2017.
[8] Lázaro et. al. "Pulse Rate Variability Analysis for Discrimination of Sleep-Apnea-Related Decreases in the Amplitude Fluctuations of Pulse Photoplethysmographic Signal in Children", 2014.
[9] Xing et. al. "Optical Blood Pressure Estimation with Photoplethysmography and FFT-Based Neural Networks", 2016.