Evaluation of Python HeartPy Tooklit for Heart Rate extraction from PPG

Hristina Mitrova[1], Bojana Koteska[1][0000-0001-6118-9044], Ana Madevska Bodganova[1], Fedor Lehocki[2,3], Beata Ondrusova[2], and Nevena Ackovska[1]

[1] Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University, Skopje, North Macedonia
[2] Institute of Measurement Science, Slovak Academy of Sciences, Slovakia
[3] Faculty of Informatics and Information Technologies, STU in Bratislava, Slovakia
hristina.mitrova@students.finki.ukim.mk, bojana.koteska@finki.ukim.mk, ana.madevska.bogdanova@finki.ukim.mk, fedor.lehocki@stuba.sk, umerondb@savba.sk, nevena.ackovska@finki.ukim.mk

Abstract. Handling the mass casualty emergency situations can be improved by introducing a chest patch sensor that is able to deliver the main vital parameters: Heart Rate (HR), Respiration Rate (RR), SPO2 and Blood Pressure. The START triage procedure requires both HR and RR parameters almost instantly.
In this paper we investigate the calculation of HR from a raw PPG signal, using appropriate functions from the Python HeartPy Tooklit, by comparing the calculated HR to the measured HR for the same patients, recorded at the same time as the PPG signal. By using several evaluation metrics, it was concluded that there is no significant difference between the measured and the calculated HR (MAE = 0,3, MSE=0,3, R^2 =0,99, Pearson's and the Spearman's coefficient of correlation, 0.99). This result is the same whether raw or filtered PPG signal was used for the HR calculation.

Keywords: Photoplethysmogram data · Signal processing · Heart rate analysis · Peak detection · Evaluation metrics

1 Introduction

Wearable real-time physiological status monitors constructed as patch-like devices capable to collect and analyse information on vital parameters such as respiration (respiration rate - RR), heartbeat (heart rate – HR), SPO2, ECG (electrocardiography), blood pressure (BP) or body temperature, can help first responders and remote personnel to rapidly intervene in time of critical events - military and civil scenarios as a result of terrorist attacks, IEDs' explosions or during rescue operations. A real-time analysis of the health status of a person in action (e.g. rescuers, emergency crews) and its prompt communication to a team leader can have critical impact on the outcome of crisis events. The mobile patch device is to be placed by emergency crews on the victims' chests as soon as possible.

One of the objectives of the ongoing project ("Smart Patch for Life Support Systems" - NATO project G5825) is to build a prototype for a patch sensor containing aforementioned vital parameters. In order to achieve this goal, the patch will contain ECG [1], PPG [13] and body temperature sensors. According to the START triage procedure [11], in order to assign the health status label (green, yellow, or red) to the injured person almost instantly, one needs to use his/her heart rate (HR) and respiratory rate (RR) vital parameters. After the triage labelling, the paramedics need the rest of the parameters – BP, SPO2 and body temperature to follow the health status of the injured person.

Our interest in this stage is to investigate the deliverance of the HR vital parameter. There are two most important technologies for measuring the heart rate: ECG and PPG (photoplethysmography). The difference between these two signal generators is as following: 1. ECG sensors measure the bio-potential generated by electrical signals that control the expansion and contraction of heart chambers. 2. PPG sensors use a light-based technology to sense the rate of blood flow, that is determined by the heart's working.

The ECG sensor is very common in the biosensor industry [1]. The integration of PPG sensor into chest-based patch device is primarily for measuring the SPO2 parameter.

It is more common to extract HR using the ECG signal, due to its power consumption, accuracy, ease of integration and richness of data [4]. Nevertheless, HR parameter can be extracted from the PPG signal, as well [11]. Following the fact that the operation of data collection from the patch device can take incorrect reading due to sensor displacement while rescuing the patient and since we have both of the sensors (ECG and PPG) on the same place (chest of the injured person), we developed an idea to use both of the signals, since it is very important to have the HR measured instantly and accurately.

In this paper we investigate the extraction of HR from the PPG signal, using Python HeartPy Tooklit. Our goal is to ensure the accuracy of calculating HR with the suitable Python Toolkit functions by comparing the calculated HR from the raw PPG signals and its correlation to the measured HR for the same patients recorded at the same time. We use the publicly available BIDMC PPG and Respiration Dataset available at Physionet.

The structure of the paper is as follows. Section II presents an overview on similar papers or related researches, Section III elaborates the process of database extraction and filtering, explanation of the used HeartPy functions and the used methodology. The results of the experiment and relevant discussion are presented in Section IV. The conclusion of the paper is given in Section V.

2 Related Work/Background

The PPG sensor measures the changes in the intensity of the light that passes through or is reflected by the examined tissue. The changes in the intensity of the transmitted or reflected light correspond with the changes in blood volume in the tissue that are caused by the periodic heart activity. Therefore, the PPG signal can be used to estimate the HR [2]. The HR from the PPG signal can be detected in the time or the frequency domain. In the time domain, the HR is estimated by the detection of the peaks in the PPG signal that occur with the heartbeat [20]. In the frequency domain, the HR is obtained by the analysis of the peaks in a spectrum where the most distinctive peak should correspond to the HR [12]. Wearable devices with the integrated PPG sensor placed e.g., on the wrist or a chest are used to monitor the physiological parameters of the subject in the hospital or during the emergency situations. Therefore, the recorded PPG signal can be contaminated with the artifacts caused by the subject movement and techniques for its elimination need to be implemented in order to obtain the most accurate estimate of the HR. Previously, analogue filters were used to obtain the required frequency range of the PPG signal followed by the digital filters that were used for the filtering and differentiation of the respiratory and heart signals from the measured PPG signal. Consequently, the heart signals were used to determine the HR by the zero-crossing method [15]. The results of the current research demonstrate that the signal decomposition techniques can be used to eliminate the motion artifacts in the PPG signal and thus determine the HR. An example of this approach is a general framework, named TROIKA, that is incorporating three features which are signal decomposition, sparse signal reconstruction and spectral peak tracking. The TROIKA shows a good accuracy in the determination of HR from the noisy PPG data obtained during running [21]. Another approach for the HR estimation from the noisy PPG signals involves a cascade of adaptive filters. First, the PPG signals are pre-processed and after that combination of adaptive filters is used to eliminate the moving artifacts. The HR is then estimated in the frequency domain from the periodogram. The results indicate that this method has a good result in the HR estimation for different datasets [3]. Further, the removal of motion artifacts can be done using Wiener filtering followed by a phase vocoder used for the HR estimation and its refinement. The system can be used also offline employing the Viterbi decoding algorithm [19]. Many other methods can be used for the determination of HR from the PPG signal contaminated with artifacts such as denoising using wavelets [17] or Kalman filtering [14]. At present, the use of deep learning techniques for the estimation of the HR from the PPG signal is being investigated. It was shown that a Multi-Layer Perceptron (MLP) neural network consisting of 3 layers and 22 neurons can be used to estimate the HR with acceptable accuracy and computational time [6]. The PPG signals obtained during various activities and obtained via measurements in clinical practice were used to test a framework named CorNET that uses Deep Neural Network (DNN). The DNN is composed of the Computational Neural Network (CNN) and the Long Short-Term Memory (LSTM) network, each in two layers and accompanied by a dense layer [5]. It follows from the above-mentioned information that various methods could be used for the estimation of the HR from the PPG signals from simpler to more sophisticated up to deep learning methods.

3 Materials and Methods

3.1 Database

In this research we used the BIDMC PPG and Respiration Dataset [16,10] available at `https://physionet.org/content/bidmc/1.0.0/`. BIDMC dataset is composed of signals extracted from the larger MIMIC II matched waveform Database, that also contains manual breath annotations. Total database size is 207.7 MB. This data was obtained from critically-ill patients at the Beth Israel Deaconess Medical Centre in Boston, MA, USA. There are total 53 patients in the dataset and for each patient the following physiological signals and physiological parameters are stored:

- Photoplethysmogram (PPG) sampled at 125 Hz;
- Electrocardiogram (ECG) sampled at 125 Hz;
- Impedance respiratory signal sampled at 125 Hz;
- Heart Rate (HR) sampled at 1 Hz;
- Respiratory Rate (HR) sampled at 1 Hz;
- Blood oxygen saturation level (SPO2) sampled at 1 Hz;

Each physiological signal is 8 minutes long. There is also information about age and gender of each patient.

The data was initially downloaded using a custom made Python script. The WFDB software package was used, in order to download the data in the wanted Python pickle format, specifically the wfdb.rdsamp function. The downloaded data was saved to a local storage.

We also performed filtering of the PPG signal, thus obtaining two databases: with raw and filtered PPG signal data. In order to filter the PPG signal, we conducted a cleaning procedure, for the waveform quality. First, the PPG signal was normalized to zero mean unit variance. Then, it was filtered with a 4th order Butterworth band-pass filter, with cutoff frequencies of 0.5 Hz and 8 Hz to remove the baseline wandering below 0.5Hz and high-frequency noise above 8Hz. Next, to remove the outliers of the PPG signal it was filtered using the Hampel filter. The choice of the filters was made according to the procedure proposed in [18].

3.2 Python HeartPy Algorithm

In order to calculate the Heart Rate from PPG, we used the Python HeartPy Toolkit [8]. The HeartPy Algorithm comes with different pre-processing options to clean up signals, including finite impulse response (FIR) filtering and outlier detection. To detect peaks, this algorithm uses an adaptive threshold to accommodate for morphology and amplitude variation in the PPG waveform, followed by outlier detection and rejection. Heartbeats identification was made by calculating the moving average using a window of 0.75 seconds on both sides of each data point. Then, the regions of interest (ROI) were computed between two points of intersection where the amplitude of the signal was larger than the moving average. This is a standard way of detecting peaks. This algorithms uses two methods of obtaining a peak's location. The first approach, the highest point in the marked ROI is taken as peak position. In the second approach a univariate spline is used to upsample and interpolate the ROI, which is then solved for its maximum [9].

A special case is the signal clipping that makes it difficult to find the accurate placement of a peak's position. This can happen due to different reasons. HeartPy algorithm detects the onset and end of these clipping segments, and uses spline interpolation to make a reconstruction of the waveform. The best solution is decided by lowering the standard deviation of peak-peak intervals. The instantaneous heart rate (beats per minute - BPM) is computed and evaluated together with the standard deviation of peak-peak intervals. The BPM value must be within a predefined range which can be set by user. The default setting is: 40 <= BPM <= 180 [7] .

3.3 Methodology

The methodology for evaluating the HeartPy algorithm for our purpose starts with the calculation of the Heart Rate for each patient, by processing the PPG signal using the Python toolkit HeartPy. For each patient there are 60001 records for the PPG signal, as for the HR there are 481 values, which verifies the fact that the data for PPG signal is sampled at 125 Hz frequency. Since the BIDMC database is not filtered, the HeartPy algorithm was tested on both raw and filtered data (Section 3.1). In order to process the unfiltered data more accurately, it was necessary to handle the missing data, using Python libraries Pandas and Numpy, that offer compatible data structures able to perform many kinds of data manipulation. The median from the available heart rate values of the given patient was placed in the fields where the data was missing.

After the described preprocessing part, for both of the data sets - raw and the filtered PPG signal, the HR was calculated by the HeartPy's Process function. By setting the frequency parameter at 125 Hz, we obtained the heart beats per minute (bpm) values. The second step is to compare these values with the average of 481 values, representing the measured Heart Rate for each patient available in the database.

Subsequently, the calculated HR values were compared with the measured HR values of each patient in the database, utilizing evaluation metrics for estimation of their difference - MAE (Mean Absolute Error), MSE (Mean Squared Error), RMSE (Root Mean Squared Error), RMSLE (Root Mean Squared Logarithmic Error) and R^2 (Squared Root), as well as Pearson's and Spearman's coefficient of correlation.

4 Results

4.1 Evaluation metrics

In this section, the used evaluation metrics are presented in more details. In the following formulas, HR parameters are presented as:

HRMi - measured HR for the i-th patient;

HRCi - calculated HR for the i-th patient;

n - number of data points (patients).

MAE is a simple metric - the absolute difference between measured HR and calculated HR, and is most robust to outliers.

$$MAE = \frac{\sum_{n=1}^{n}(|HRMi - HRCi|)}{n}$$

MSE - mean squared error, represents the squared difference between the measured HR and calculated HR.

$$MSE = \frac{1}{n}\sum_{n=1}^{n}(HRMi - HRCi)^2$$

This metric can be used as a loss function, but is sensitive to outliers and then the penalization is bigger that leads to calculating bigger MSE, decreasing the robustness compared to MAE. Another evaluation metric that can be used as a loss function as well, that is more frequently used with deep learning techniques, is RMSE.

$$RMSE = \sqrt{\frac{1}{n}\sum_{n=1}^{n}(HRMi - HRCi)^2}$$

By taking the log of the RMSE metric, another evaluation metric - RMSLE (Root Mean Squared Log Error) is produced. RMSLE value will only consider the relative error between the measured HR and calculated HR value, neglecting the scale of data.

$$RMSLE = \sqrt{\frac{1}{n}\sum_{n=1}^{n}\log\left(HRCi + 1\right) - \log\left(HRMi - 1\right))^2}$$

In contrast to the previously described metrics that depend on the context of the problem, we decided to use R squared metrics. This statistical measure represents the proportion of the variance for a dependent variable that is extracted by an independent variable. In this case, it is the variation between the measured HR and the calculated HR with HeartPy.

$$R^2 = 1 - \frac{\text{Variance expained by the HeartPy function}}{\text{Total variance}}$$

Table 1 presents the results obtained from the applied evaluation metrics on the raw PPG signal and Table 2 - on the improved, cleaned, i.e. filtered PPG signal. There is no difference between the pairwise elements.

Table 1. Evaluating the HeartPy performance on BIDMC database with different evaluation metrics on raw PPG signal

Evaluation metrics				
MAE	MSE	RMSE	RMSLE	R squared
0.30375	0.34617	0.58836	0.00771	0.99797

Table 2. Evaluating the HeartPy performance on BIDMC database with different evaluation metrics on filtered PPG signal

Evaluation metrics				
MAE	MSE	RMSE	RMSLE	R squared
0.30375	0.34617	0.58836	0.00771	0.99797

Another metrics that confirms the obtained results are the calculated Pearson's and the Spearman's coefficient of correlation between the same parameters, the measured and the calculated HR:

- Pearson: 0.99;
- Spearman: 0.99.

The obtained results from the evaluation metrics indicate a small difference between the measured and calculated values for HR, since normal Heart Rate for adults ranges from 60 to 100 beats per minute. Lower values of MAE, MSE, RMSE and RMSLE, imply high accuracy of the used HeartPy function. On the other hand, higher value of R^2 is considered desirable, which in our case is achieved. These results are verified by the high correlation coefficients.

This result is the same whether raw or filtered PPG signal was used for the HR calculation with the appropriate HeartPy function (Table 1 and Table 2). It can be concluded that HeartPy's algorithm is reliable for rapid and precise callculation of HR from a raw PPG signal.

5 Conclusion

Building a prototype of a chest patch sensor that can be used in mass casualty emergency situations means being able to deliver the main vital parameters: HR, RR, SPO2 and blood pressure, in real time. The START triage procedure requires both HR and RR parameters almost instantly, but to follow the hemodynamic stability of the injured person, paramedics need SPO2 and BP as well. The forementioned patch sensor will contain ECG, PPG and body temperature sensors.

The main goal of the paper was to explore the algorithms that can be used to calculate the Heart Rate only by utilizing the PPG signal from a patient, using HeartPy Toolkit [11]. It is very important to be able to rely on this fast HR calculation procedure, when trauma patient needs to be labeled in the Triage procedure (START triage system).

The measured HR from the used BIDMC database was compared to the calculated HR from the raw PPG signal for the same subject (patient), measured at the same time, utilizing the HeartPy Toolkit. In order to compare if the difference between the measured and calculated HR differs when the PPG signal is filtered, we conducted a cleaning procedure, taking into account the waveform quality. By using several evaluation metrics, it was concluded that there is no significant difference between the measured and the calculated HR (MAE = 0,3, MSE=0,3, R squared=0,99, Pearson's and the Spearman's coefficient of correlation, 0.99). This result is the same whether raw or the filtered PPG signal was used for the HR calculation with the appropriate HeartPy function. This is very important conclusion that enables us to use the unfiltered PPG signal directly from the sensor in order to get the HR value almost immediately with the help of the HeartPy function.

The success of this investigation leads us to an idea to verify the use of HeartPy as a tool to extract the Respiratory Rate and SPO2 from the raw PPG signal, as well. Another step is to compare these results with the reliability, accuracy and velocity when using ECG signal for calculating HR and RR using Python libraries.

Acknowledgment

This paper has been written thanks to the support of the "Smart Patch for Life Support Systems" - NATO project G5825 SP4LIFE.

References

1. Al-Hamadi, H., Gawanmeh, A., Al-Qutayri, M., Ismail, M.: A framework for the verification of an ecg biosensor algorithm. Analog Integrated Circuits and Signal Processing **90**(3), 523–538 (2017)
2. Allen, J.: Photoplethysmography and its application in clinical physiological measurement. Physiological Measurement **28**(3) (2007)
3. Arunkumar, K., Bhaskar, M.: Heart rate estimation from wrist-type photoplethysmography signals during physical exercise. Biomedical Signal Processing and Control **57**, 101790 (2020). https://doi.org/https://doi.org/10.1016/j.bspc.2019.101790, `https://www.sciencedirect.com/science/article/pii/S1746809419303714`
4. Becker, D.E., Reed, K.L.: Essentials of local anesthetic pharmacology. Anesthesia progress **53**(3), 98–109 (2006)
5. Biswas, D., e.a.: Cornet: Deep learning framework for ppg-based heart rate estimation and biometric identification in ambulant environment. IEEE Transactions on Biomedical Circuits and Systems **13**(2) (2019)
6. Boloursaz Mashhadi, M., Asadi, E., Eskandari, M., Kiani, S., Marvasti, F.: Heart rate tracking using wrist-type photoplethysmographic (ppg) signals during physical exercise with simultaneous accelerometry. IEEE Signal Processing Letters **23**(2), 227–231 (2016). https://doi.org/10.1109/LSP.2015.2509868
7. van Gent, P., Farah, H., Nes, N., Arem, B.: Analysing noisy driver physiology real-time using off-the-shelf sensors: Heart rate analysis software from the taking the fast lane project. (11 2018). https://doi.org/10.13140/RG.2.2.24895.56485

8. van Gent, P., Farah, H., van Nes, N., van Arem, B.: Heartpy: A novel heart rate algorithm for the analysis of noisy signals. Transportation research part F: traffic psychology and behaviour **66**, 368–378 (2019)
9. van Gent, P., Haneen, F., N, N., van Arem, Bart: Heart rate analysis for human factors: Development and validation of an open source toolkit for noisy naturalistic heart rate data pp. 173–178 (2018)
10. Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. circulation **101**(23), e215–e220 (2000)
11. Iserson, K.V., Moskop, J.C.: Triage in medicine, part i: concept, history, and types. Annals of emergency medicine **49**(3), 275–281 (2007)
12. Islam, Ahamed, M.Z.I., Yasar, S., Shahnaz, M., Fattah, C., S.: A time-frequency domain approach of heart rate estimation from photoplethysmographic (ppg) signal. Biomedical Signal Processing and Control **36**(5) (2017)
13. Kramer, M., Lobbestael, A., Barten, E., Eian, J., Rausch, G.: Wearable pulse oximetry measurements on the torso, arms, and legs: A proof of concept. Military medicine **182**(suppl_1), 92–98 (2017)
14. Lee, B., Han, J., Baek, H.J., Shin, J.H., Park, K.S., Yi, W.J.: Improved elimination of motion artifacts from a photoplethysmographic signal using a kalman smoother with simultaneous accelerometry. Physiological Measurement **31**(12), 1585–1603 (oct 2010). https://doi.org/10.1088/0967-3334/31/12/003, https://doi.org/10.1088/0967-3334/31/12/003
15. Nakajima, K., Tamura, T., Miike, H.: Monitoring of heart and respiratory rates by photoplethysmography using a digital filtering technique. Medical Engineering and Physics **18**(5), 365–372 (1996). https://doi.org/https://doi.org/10.1016/1350-4533(95)00066-6, https://www.sciencedirect.com/science/article/pii/1350453395000666
16. Pimentel, M.A., Johnson, A.E., Charlton, P.H., Birrenkott, D., Watkinson, P.J., Tarassenko, L., Clifton, D.A.: Toward a robust estimation of respiratory rate from pulse oximeters. IEEE Transactions on Biomedical Engineering **64**(8), 1914–1923 (2016)
17. Raghuram, M., Venu Madhav, K., Hari Krishna, E., Ashoka Reddy, K.: Evaluation of wavelets for reduction of motion artifacts in photoplethysmographic signals pp. 460–463 (2010). https://doi.org/10.1109/ISSPA.2010.5605443
18. Slapničar, G., Mlakar, N., Luštrek, M.: Blood pressure estimation from photoplethysmogram using a spectro-temporal deep neural network. Sensors **19**(15), 3420 (2019)
19. Temko, A.: Accurate heart rate monitoring during physical exercises using ppg. IEEE Transactions on Biomedical Engineering **64**(9) (2017)
20. Wójcikowski, M., Pankiewicz, B.: Photoplethysmographic time-domain heart rate measurement algorithm for resource-constrained wearable devices and its implementation. Sensors **20**(6) (2020), https://www.mdpi.com/1424-8220/20/6/1783
21. Zhang, Z., Pi, Z., Liu, B.: Troika: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. IEEE Transactions on Biomedical Engineering **62**(2), 522–531 (2015). https://doi.org/10.1109/TBME.2014.2359372