Removal of Motion Artifacts from Photoplethysmogram Data for Feature Based Diagnosis of Atrial Fibrillation

BS Electrical Engineering

Session: 2020 - 21

By:

Mohammad Afaq	21100214
Sheraz Hassan	21100285
Salaar Ali Assad	21100087

Supervised by: Dr. Muhammad Tahir,
Assistant Professor, EE

Department of Electrical Engineering

Lahore University of Management Sciences

Abstract

As a substitute of the traditional method of electrocardiogram (ECG) to monitor the heart activity of a patient, photoplethysmogram (PPG) has become an important alternative in the medical industry today. PPG holds immense advantages over the conventional ECG in terms of cost and flexibility. The data of a patient can be easily recorded through a non-invasive PPG sensor device without requiring the need of a medical supervision unlike ECG. Moreover, the recorded PPG data is useful in extracting vital cardiovascular parameters such as Heart Rate (HR), R-R interval, as well as respiration rate (RR) for identifying atrial fibrillation (Afib), a common cardiovascular disease showing symptoms of irregular heart behavior. However, the biggest challenge in extracting these parameters from the PPG data is the interference due to motion artifacts. These motion artifacts introduce noise in the data drastically affecting its quality prone to movement. This project will identify and address the above-mentioned challenges to remove noise due to motion artifacts from the corrupt PPG data, ensuring that the original nature of the PPG waveform is not lost. This noise free and cleansed data can then be used to extract vital cardiovascular parameters with accuracy. It will start off by first comparing the raw PPG data from a low-end device with that of the pre-processed data from a high end device with pre-implemented low frequency noise motion artifacts removal algorithm. Next, this project will propose a data processing algorithm to detect and eliminate motion artifacts from corrupted raw PPG signals by procuring data from both the PPG and its corresponding 3 axis accelerometer readings. Finally, it will process the cleansed PPG data and compute its cardiovascular parameters to accurately predict the possibility of whether the patient is exhibiting symptoms of Afib. Hence, the project we propose aims to create an end-to-end machine intelligence-based system that allows us to not only detect irregularities in the human heartbeat using a cleansed PPG data and data processing algorithms, but also monitor this behavior over a long time.

Acknowledgments

The authors wish to express sincere appreciation to Dr. Muhammad Tahir for his dedicated supervision during the course of this senior year project and in the preparation of this manuscript.

Originality Certificate

"We the undersigned certify that this submission is the original work of members of the group and meets the Faculty's Expectations of Originality":

1. Mohammad Afaq [2021-10-0214]
   Signed: __________________ 21100214@lums.edu.pk

2. Sheraz Hassan [2021-10-0285]
   Signed: __________________ 21100285@lums.edu.pk

3. Salaar Ali Assad [2021-10-0087]
   Signed: __________________ 21100087@lums.edu.pk

Advisor
Dr. Muhammad Tahir

tahir@lums.edu.pk

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

Disorders pertaining to the blood vessels and the heart are categorized as cardiovascular diseases (CVDs) such as; coronary heart disease (CHD), hypertension, angina, myocardial infarction, rheumatic heart disease and cardiac arrhythmias namely atrial fibrillation (Afib) etc., resulting from sedentary behaviour, poor dietary choices, smoking and pollution, brewing against a background of genetic susceptibility. Data from the World Health Organization states that CVDs are the primary cause of death world-wide among non-communicable diseases, and reasonable predictions state that the situation will become worse. In Pakistan, according to WHO statistics in 2014, approximately one fifth (19%) of the entire population was had CVDs, making it the leading non-communicable disease in the country [2]. In this scenario, a growing demand of medical assistance implies a large number of populations require assistance in hospitals where Electrocardiogram (ECG) signal recording remain the most prevalent clinical standards of care for cardiac health assessment and monitoring. However, ECG has not always been the first choice of patients due to its uncomfortable position of application, high cost and availability issues in the hospitals [3]. As a result, majority of the patients ignore the little to no symptoms of CVDs, resulting in a diagnosis only after a heart attack or a severe heart irregularity is experienced. Fortunately, with advancements in technology, the cardiovascular data of the patient can now be monitored real time without requiring any medical supervision. By having patients wear a wrist worn device which contains photoplethysmogram (PPG) sensors, data from the PPG signals can be collected and analyzed using data processing algorithms. Through this analysis, abnormal heart rhythms, which if left alone can evolve into major cardiac disorders, can be detected and addressed in the timely manner, potentially saving the lives of many.

1.2 Problem Statement

The precise nature of the proposed project is to deal with the problem of remote and ambulatory cardiac health monitoring and assessment, more specially for Afib, using data processing algorithms by utilizing PPG signals as recorded by a wrist-worn wearable prototype device. The challenge in PPG based classification framework is that, first of all, motion artifacts produced by the movement of the sensor with respect to skin surface corrupt the PPG data making it difficult to analyze, and secondly, the corrupted PPG data can super impose on the vital heart activity data that may denote the possibility of a cardiac event hence making it difficult to diagnose [4]. Hence, there is a dire need to address these issues and develope data processing algorithms for a robust diagnosis and classification of any troubling cardiac event [4].

1.3 General Block Diagram

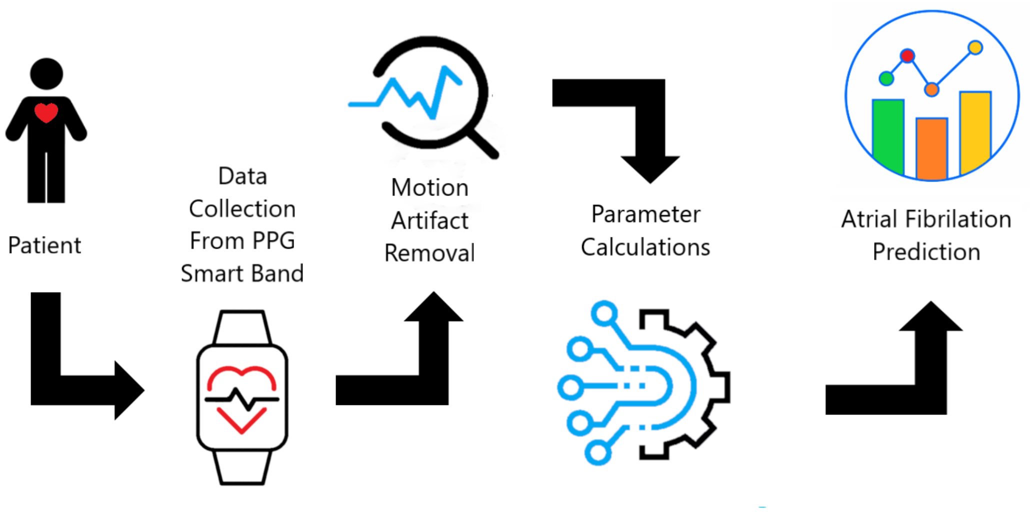

FIGURE 1.1: Block diagram showing different stages of our project pathway

Figure 1.1 gives us a brief overview of the different stages involved in our project, with the first stage comprising of PPG data collection. The data collected will be collected from two PPG smart bands, mentioned in Chapter 3. The second stage will be responsible for the extraction of the cleansed PPG signals from the corrupted raw signal by detecting and removing motion artifacts present in the signal using a data processing algorithm. This will be followed by parameter calculations so that the model can be trained to detect and differentiate between patients suffering from AFib and healthy individuals. The final stage involves the model accurately predicting whether a potential patient suffers from CVDs despite motion artifacts being introduced in the recorded PPG data.

1.4 Social Benefits And Relevance

As previously stated, statistics from WHO state that, among non-communicable diseases, CVDs are considered the primary cause of deaths. With the situation predicted to deteriorate as time passes, developing nations such as Pakistan are in dire need of a solution which can cater to this ever-increasing demand of cardiac medical assistance. This project, therefore, holds immense social benefits and relevance, as its application not only caters to local population but the national community as well.

Currently, medical resources, such as ECG signal recording instruments, are far too limited in Pakistan. They also require the patient to be present in the hospital hindering their daily life and, further straining doctors and medical staff at the hospital. But perhaps the most important aspect is that, despite the fact many cardiac disorders can be treated and possibly cured should there be timely intervention, the large medical expense such treatments incur act as a deterrent, particularly for those belonging to low income communities, making early detection of cardiac diseases difficult. Therefore, by providing an alternative that can detect early symptoms of AFib, is cheaper, requires minimal assistance from doctors and does not hinder the daily activities of the patients, this project can not only save lives, but also, minimize the expenses and time spent on hospital trips, medication, treatments, etc.

Furthermore, this project can also help doctors more accurately detect symptoms which would, under normal circumstances, go unnoticed as continuous monitoring of every patient at the hospital is impractical. The data of a patient during any hour of the day will be readily available for the medical staff hence allowing them to cater to a greater number of patients while at the same time offer reliable and accurate treatment.

1.5 Goals and Objectives

Our goals can be divided into three parts:

- Compare the differences between raw data from a low-end wearable PPG sensor device and processed data from a high-end PPG device.

- Understanding and analysis of the limits in data collections and its nature from low end PPG device such as PPG feature, artifacts, noise etc., with reference to data collected from high end device.
- Develop data processing algorithms that will take into account and adjust for motion artifacts present in the raw input data, such as motion artifacts.
- Create and use the cleansed PPG database to compute vital cardiac parameters using data science to identify potential cases of Afib.

1.6 Outcomes

Our project will present the following outcomes:

- Application of data processing algorithms on the raw PPG data to remove motion artifacts efficiently.
- Diagnosis of Afib on the cleansed PPG data by extracting and processing cardiovascular features
- If time permits, live simulation of the PPG data of an individual using a software algorithm that removes motion artifacts, extracts and computes PPG features and analyses them to predict red flags for any irregular heart activity under a time window T.

1.7 Timeline and Distribution of Work

Figure 1.2 shows the general timeline for our project. For Goal 1, in order to develop a firm understanding of these features and the nature of the raw data, an extensive review of existing literature began over the summer break and throughout September. Literature linking AFib and PPG signals and the limitations in its data accuracy and collection was also reviewed during this time and further research on these topics will continue throughout the current semester.

For the second goal, device comparison, the processed data from the high end PPG device will be compared to the raw data from a low-end PPG sensing device. This comparison will allow for a greater understanding of the limitations of PPG data collection in low-end devices, creating a more holistic view for the practical and commercial implementation of wearable PPG smart bands for AFib detection. Goal 1, which began in September, will be concluded in December 2020 for the Senior Project 1 report.

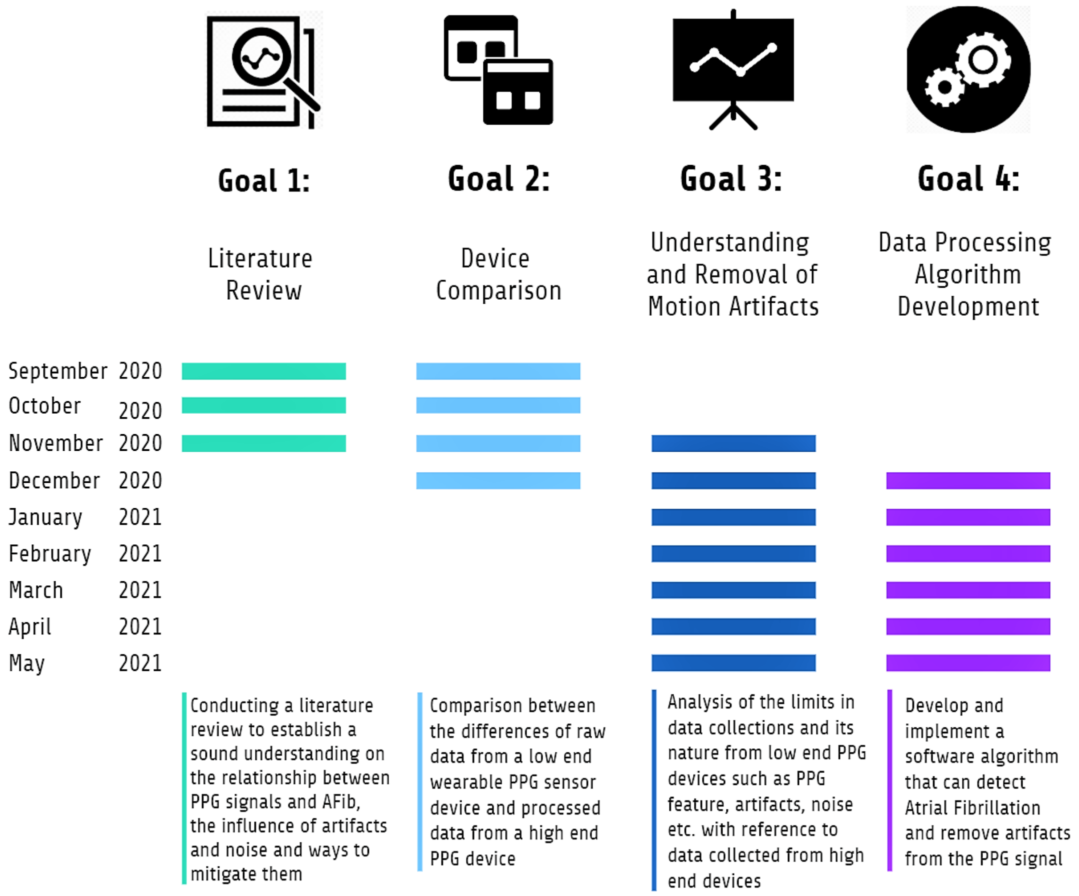

FIGURE 1.2: General timeline for our project

Goal 3 is the investigation of the nature of PPG features along with the various types of motion artifacts and noise that will be present in the raw data collected. The raw data will be analyzed to not only understand the challenges faced when collecting PPG signals, particularly from low-end devices, but also, determine potential solutions that can mitigate these problems. Goal 3 will begin in November and expand over both semesters till May 2021 where it will be concluded in the Senior Project 2 report.

The last goal will focus on the development and implementation of a software algorithm that can successfully detect and remove artifacts from the raw PPG signal, primarily motion artifacts, using data processing techniques. Moreover, the algorithm will also be able to correctly identify cases of individuals suffering from AFib. Research on exploring the various techniques previously used in existing literature has been carried out along with multiple attempts to simulate these results. The investigation and development of the software algorithm will start towards the end of the first semester and will conclude in May 2021.

Chapter 2

Background

2.1 Literature Review

2.1.1 Photoplethysmogram

Photoplethysmography (PPG) is an optical bio signal technique used to monitor the heart activity by measuring the variations in blood flowing through a tissue by observing the variations in the amount of light reflected or transmitted. It is measured using a PPG sensor at the skin surface and unlike ECG, is a low cost non invasive alternative which operates close to green and red infrared frequency. The way how a PPG sensor works is through LED and a photo diode. The LED first illuminates the tissue and a photo diode measures the variations in the light intensity related to blood volume flowing through the blood vessels since the light at infrared frequency is absorbed the most by the blood than bones and tissues. These volumetric changes of blood in the blood vessels produce variation in voltage through which the PPG waveform is produced with respect to time. The variation of the blood volume with respect to the heart activity forms the variation in the amplitude of the voltage hence forming the AC component of the waveform where as quantity of light reflected or transmitted by the tissues corresponds to the DC component of the waveform. As the PPG waveform is produced, the DC component forms the base of the signal on which the AC component varies to plot out the popular shape of the PPG waveform.

2.1.2 PPG Waveform and its Extractable Features

PPG waveform is the most popular clinical waveform as of today. A simple technology, PPG has dominate ECG in the past years due to its increased sensitivity and specificity in rapid detection and diagnosis of the heart activity [5]. It is considered to be relatively less complex than ECG. A lot of work has already been performed to measure heart rate and blood pressure PPG signals. PPG can give also give a good estimation of many other heart related diseases other than AF. Researchers are exploring the ways to use the power of modern machine learning techniques as well as deep learning to be used on PPG signals in order to increase the accuracy of CVDs detection and screening [6–8].

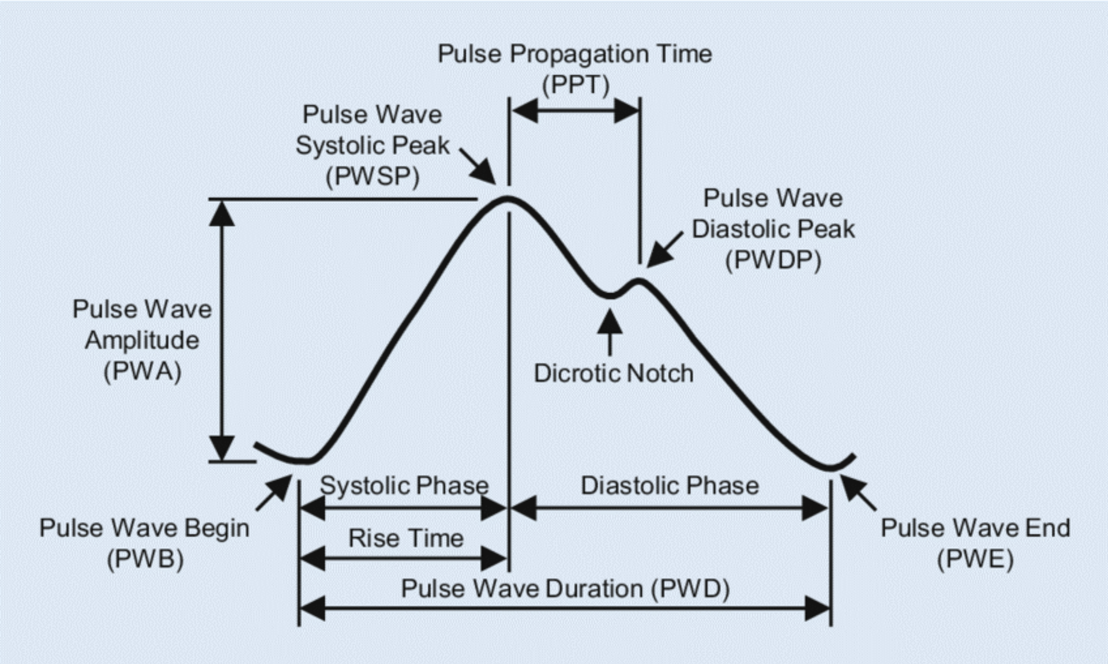

FIGURE 2.1: PPG waveform notation [1]

A cardiac cycle generated in the PPG waveform is shown in Figure 2.1. The signal is generated as the heart contracts and this contraction spreads through the vascular tree [9]. As the left ventricle contacts, the blood is pumped into the arterial tree, which is denoted as the positive gradient of the PPG waveform in figure 2.1. Further the closing of the aortic valves and the separation of the systolic and diastolic phase's results in a decrease in amplitude of the waveform [9]. These specific events which form part of the cardiac cycle together are a part of a PPG waveform for one heartbeat. PPG waveform is believed to hold a range of extractable feautures which are immensesfully resourceful in detecing Afib. Among them, the most relevant to our project is the R-R interval or Inter Beat Interval (IBI) which measures the difference in time between successive peaks (one heartbeat). Logically, with high metabolic activity, the R-R interval decreases as the heart rate increases and the peak magnitude remains somewhat constant. This means that the peaks of each successive PPG waveform gets closer together but this does not happen spontaneously. A certain pattern in the shifting of R-R interval is observed for a healthy individual. The case is opposite for a patient diagnosed with Afib.

2.1.3 Atrial Fibrillation Detection

For this project, our focus lies on detecting Afib by analyzing the difference of successive R-R intervals as well as the variation in the magnitude of the PPG signal since Afib diagnosed patients show the symptoms of ill regular rhythm of the heart activity lasting for more than 30 seconds [9]. Figure 2.2 and 2.3 shows the Blood Volume Pulse (BVP) data acquired from a PPG sensor of a healthy patient with that of a patient diagnosed with AF under a 30 second time window. We will constantly use the term BVP in this report to denote PPG data since it is commonly known as BVP in scientific literature.

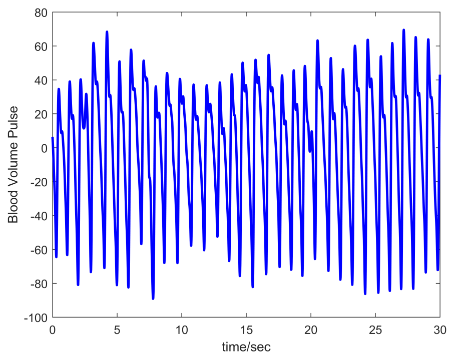

FIGURE 2.2: BVP for a healthy individual

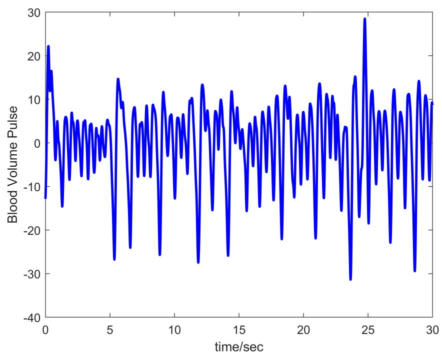

FIGURE 2.3: BVP for an Afib patient

Each peak in the BVP plot represents a heartbeat whereas the time difference in each successive peak represents the R-R interval. It can be clearly observed that for a patient diagnosed with AF in figure 2.3, the R-R interval are ill regular with random variations in the amplitudes of the BVP plot as well whereas for a healthy patient in figure 2.2 , both the parameters of R-R interval and amplitude remain constant throughout the 30 second window.

2.1.4 Motion Artifacts

Motion artifacts today remains the biggest challenge in recording reliable BVP data. Before reviewing literature on how such motion artifacts can be removed, it is vital to understand how noise due to motion artifacts is introduced in the signal. Motion artifacts occur not only by the disturbance in the contact made between the skin and the sensor, but also by the movement of the body which can be as small as movement through breathing [10]. Hence, making it a difficult task to obtain an accurate BVP reading through sensors. Motion artifacts and noise introduced into the readings by sensors make the BVP signal unreliable and inaccurate. Hence the parameters obtained through these readings cannot be relied on to detect AF.

With the introduction of motion artifacts, the recorded BVP signal is corrupted due to added noise making it difficult to analyze and extract vital information. These motion artifacts are random, meaning they are introduced by random movement of the user hence will be inevitably introduced in the BVP signal recording of a patient whose data is being taken simultaneously throughout the day. As a result, the signal obtained will be unsatisfactory and must be processed using data processing algorithms to eliminate these motion artifacts before accurate and reliable data can be obtained for analysis. To better understand the situation, it is important to visualise the BVP data from a healthy individual both with and without motion artifacts. To ensure that the BVP data was free from motion artifacts, the recording was taken with the individual at rest position.

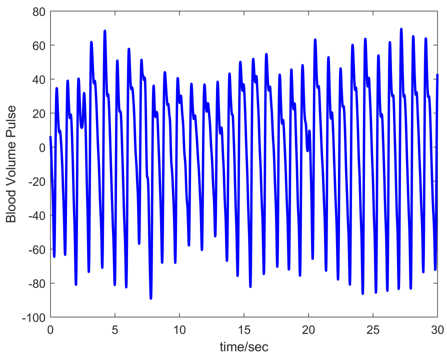

FIGURE 2.4: BVP for a healthy individual without motion artifacts

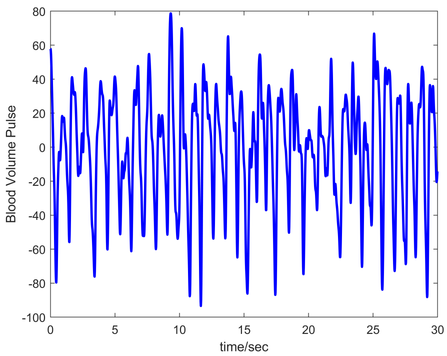

FIGURE 2.5: BVP for a healthy individual with motion artifacts

Figure 2.4 shows the recorded BVP of the individual without motion artifacts in a 30 second interval and figure 2.5 shows the BVP of the same individual but with motion artifacts in a 30 second interval. Figure 2.5 depicts how raw BVP signal for a patient will become noisy subject to motion artifacts hence making it difficult to analyze to extract information from it. Therefore, the major problem that we intend to solve in this project is to accurately remove the motion artifacts from the BVP signal such that no vital information is lost and the signal retains its shape which is extremely important for post analysis of the signal. This includes accurately detecting BVP signal cardiac parameters and using them to accurately predict the heart condition of the patient.

2.1.5 Existing Methods

As of today, there exists several existing methods which are capable of eliminating motion artifacts and cleansing BVP signal widely used in scientific research. Among these methods extensively studied in literature are independent component analysis (ICA), adaptive filtering and deep learning technique to extract physiological features from the BVP signal [10–12]. ICA algorithm is based on the probability statistical theory where the complex BVP signal is separated into its sub components corresponding to the BVP of the variation of blood volume in the vessels, motion artifacts and noise etc. These sub components corresponding to motion artifacts and noise can then be filtered out leaving behind noise free BVP signal [10, 13]. However, the ICA algorithm makes a strong assumption regarding the independence of these sub components of the corrupted BVP signal. As a result, the corrupted BVP signal usually does not meet this assumption of the ICA algorithm which may lead to inaccuracy of the results and lost of information which might otherwise prove to be useful [10, 14, 15].

Another widely used method is adaptive filtering which can achieve satisfactory results by suppressing in band frequencies of motions that might have been introduced in the corrupted BVP signal as motion artifacts [10, 16, 17]. This means that a correlation between the BVP and its corresponding 3 axis accelerometer data is assumed, highly depending on the quality of the desired reference signal to remove the motion artifacts effectively [10]. As a result, this method proposes several disadvantages e.g. a bad quality reference signal through which error is calculated will degrade the quality of the output PPG signal. Moreover, a complex PPG signal will make the corrupted signal more difficult to cleanse by only eliminating frequencies of motions through the accelerometer data. A sample plot of highly corrupted BVP along with its accelerometer data is shown in figure 2.6. Moreover, this method utilizes extensive computational resources and power, therefore not optimum for implementing on a wearable PPG sensor device which doesn't hold such high processors as a computer does.

Another proposed approach relies on deep learning technqiues to extract the physiological features from the corrupted BVP signal.The paper in [18] highlights that the difficulty in motion artifacts removal from corrupted BVP is attributed to the non linearity of the cleansed BVP signal and noise. It is this non-linear relationship that limits the effectiveness of other methods, such as adaptive filtering, due to the linear nature of the additive noise hypothesis, and signal decomposition algorithms. Therefore, this technique proposes a Signal-Noise Interaction modeling based algorithm for motion artifact removal (SniMA) for complex BVP

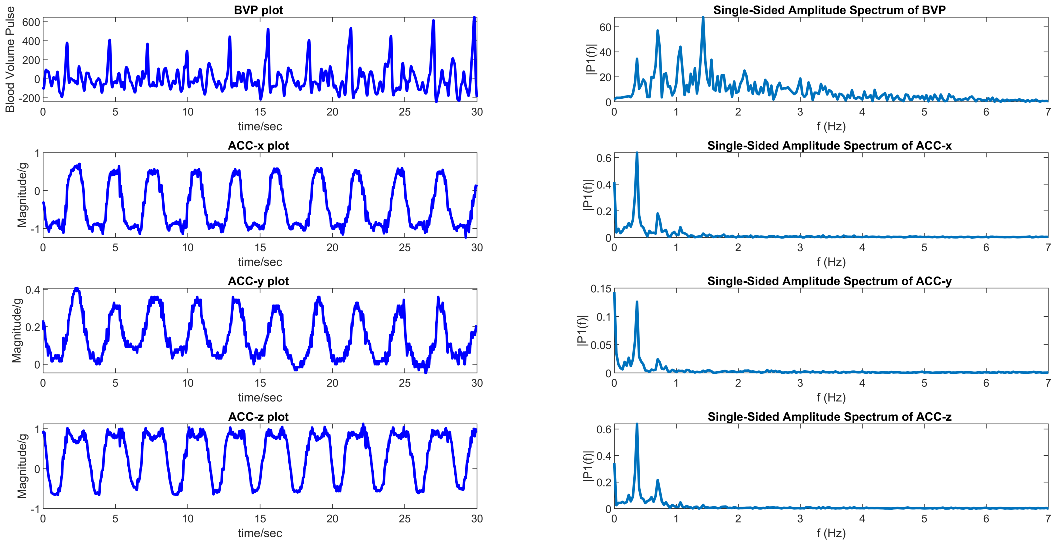

FIGURE 2.6: BVP and Accelerometer data for a healthy patient with complex motion artifacts. Frequencies of motions can be observed which aren't enough to fully restore the corrupted PPG signal through adaptive filtering

signals [18]. This algorithm will make use of Envelope Filtering (EF) as well as Time-Delay Neural Network (TDNN) in order to model signal-noise interaction as opposed to directly modeling PPG signals [18]. The EF algorithm is applied to normalize data from the PPG device and eliminate the training imbalances induced by respiration. Whereas, the TDNN aims to not only model the PPG signal noise interaction as a non-linear process, but also, introduce time-dependence information for artifact removal.

Hence, motion artifact removal from a corrupted PPG signal is the primary objective of this project. We will study the existing literature and propose a technique that able to remove the motion artifacts effectively without disrupting the nature of our original PPG signal such that vital cardio parameters can be estimated with accuracy. The methodology of our proposed technique will be discussed in the coming chapter.

2.1.6 Data Processing Algorithms

With the increased number of populations around the world, estimated to reach 8.5 billion by the end of 2030, meeting health care needs of its citizens has become a real challenge for every country. These significant numbers present significant amount of data available of patients which has rapidly increased the demand of machine intelligence in the health industry. Data science in medicine has proven to be immensely resourceful in identifying potential disease infections and can drastically improve the accuracy of diagnosis reducing the human error that presents itself at certain occasions. Hence, in this project, we intend to develop a prototype where heart activity of the patient is continuously being recorded through the PPG sensor, and is being monitored using our data processing algorithms to not only remove motion artifacts and cardiovascular parameters accuracy but to also predict any red flags or emergency situations a patient be subject to under a timely manner.

We will talk more in details about the different algorithms we used and evaluate the performance of each algorithm on our dataset in the coming chapters.

References

[1] C. Fischer, M. Glos, T. Penzel, and I. Fietze, “Extended algorithm for real-time pulse waveform segmentation and artifact detection in photoplethysmograms,” *Somnologie*, vol. 21, pp. 1–11, 05 2017.

[2] S. Naseem, U. K. Khattak, H. Ghazanfar, and A. Irfan, “Prevalence of non-communicable diseases and their risk factors at a semi-urban community, Pakistan,” *Pan Afr Med J*, vol. 23, p. 151, 2016.

[3] S. Lee, G. Ha, D. Wright, Y. Ma, E. Sen-Gupta, N. Haubrich, P. Branche, W. Li, G. Huppert, M. Johnson, B. Mutlu, K. Li, N. Sheth, J. Wright, y.-s. Huang, M. Mansour, J. Rogers, and R. Ghaffari, “Highly flexible, wearable, and disposable cardiac biosensors for remote and ambulatory monitoring,” *npj Digital Medicine*, vol. 1, 12 2018.

[4] S. P. Shashikumar, A. J. Shah, Q. Li, G. D. Clifford, and S. Nemati, “A deep learning approach to monitoring and detecting atrial fibrillation using wearable technology,” in *2017 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, 2017.

[5] A. Alian and K. Shelley, “Photoplethysmography,” *Best Practice Research Clinical Anaesthesiology*, vol. 28, 09 2014.

[6] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, “Deep learning for healthcare applications based on physiological signals: A review,” *Comput Methods Programs Biomed*, vol. 161, pp. 1–13, Jul 2018.

[7] A. Aliamiri and Y. Shen, “Deep learning based atrial fibrillation detection using wearable photoplethysmography sensor,” in *2018 IEEE EMBS International Conference on Biomedical Health Informatics (BHI)*, 2018, pp. 442–445.

[8] C. Yang, C. Garcia, J. Rodriguez-Andina, J. Farina, A. Iñiguez, and S. Yin, "Using ppg signals and wearable devices for atrial fibrillation screening," *IEEE Transactions on Industrial Electronics*, vol. PP, pp. 1–1, 01 2019.

[9] T. Pereira, N. Tran, K. Gadhoumi, M. Pelter, D. Do, R. Lee, R. Colorado, K. Meisel, and X. Hu, "Photoplethysmography based atrial fibrillation detection: a review," *npj Digital Medicine*, vol. 3, 12 2020.

[10] S. Nabavi and S. Bhadra, "A robust fusion method for motion artifacts reduction in photoplethysmography signal," *IEEE Transactions on Instrumentation and Measurement*, vol. PP, pp. 1–1, 07 2020.

[11] B. Sun, C. Wang, X. Chen, Y. Zhang, and H. Shao, "Ppg signal motion artifacts correction algorithm based on feature estimation," *Optik*, vol. 176, pp. 337 – 349, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0030402618313883

[12] A. Tharwat, "Independent component analysis: an introduction," *Applied Computing and Informatics*, 08 2018.

[13] B. Kim and S. k. Yoo, "Motion artifact reduction in photoplethysmography using independent component analysis," *Biomedical Engineering, IEEE Transactions on*, vol. 53, pp. 566 – 568, 04 2006.

[14] Jianchu Yao and S. Warren, "A short study to assess the potential of independent component analysis for motion artifact separation in wearable pulse oximeter signals," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, 2005, pp. 3585–3588.

[15] Y. Ye, Y. Cheng, W. He, M. Hou, and Z. Zhang, "Combining nonlinear adaptive filtering and signal decomposition for motion artifact removal in wearable photoplethysmography," *IEEE Sensors Journal*, vol. 16, no. 19, pp. 7133–7141, 2016.

[16] M. R. Ram, K. V. Madhav, E. H. Krishna, N. R. Komalla, and K. A. Reddy, "A novel approach for motion artifact reduction in ppg signals based on as-lms adaptive filter," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 5, pp. 1445–1457, 2012.

[17] P. Regalia, *Adaptive IIR filtering in signal processing and control.* Routledge, 2018.

[18] K. Xu, X. Jiang, and W. Chen, "Photoplethysmography motion artifacts removal based on signal-noise interaction modeling utilizing envelope filtering and time-delay neural network," *IEEE Sensors Journal*, vol. 20, no. 7, pp. 3732–3744, 2020.