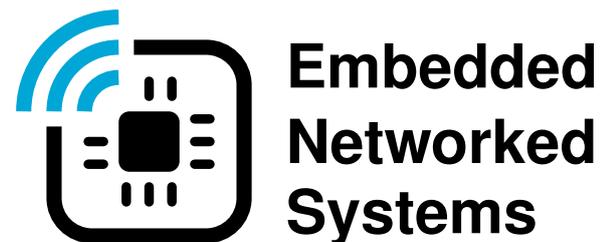


Delft University of Technology
Master of Science Thesis in Embedded Systems

Adaptive Real-Time PPG Signal Qualification and Pulse Segmentation

Marek Łukasz Vette



Adaptive Real-Time PPG Signal Qualification and Pulse Segmentation

Master of Science Thesis in Embedded Systems

Embedded and Networked Systems Group
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands

Marek Lukasz Vette
m.l.vette@student.tudelft.nl
marek.vette@gmail.com

1 October 2020

Author

Marek Lukasz Vette (m.l.vette@student.tudelft.nl)

Title

Adaptive Real-Time PPG Signal Qualification and Pulse Segmentation

MSc Presentation Date

9 October 2020

Graduation Committee

dr. ir. Marco Antonio Zúñiga Zamalloa (chairman) Delft University of Technology

dr. ir. Qing Wang Delft University of Technology

dr. ir. Jan S. Rellermeyer Delft University of Technology

Abstract

Ever since its invention in the 1930s, photoplethysmography (PPG) is a widespread technique used for health-monitoring. Via illumination of the human skin with a light source and capturing the light, an estimate of important physiological properties such as the heart rate can be made. This is commonly done with dedicated medical equipment, but studies from the last decade have shown that the smartphone could also be an adequate sensor, using the flash light and camera as transmitter and receiver respectively. Small, mobile and off-the-shelf, the smartphone provides many advantages over the conventional sensor. However, the quality of the PPG signals measured is a point of concern, as these could lead to incorrect estimates and several challenges need to be addressed.

Firstly, user movement disturbs the contact area between the skin and sensor, introducing distortions in the PPG signal. Especially subtle motion artifacts like finger contact pressure impact the PPG signal quality. Secondly, the smartphone camera has a wide range of settings that can be used, but a proper analysis was lacking in literature. Thirdly, recent research has shown that PPG signals from an individual are unique. There is no common morphology. This means that algorithms developed for PPG need to account for unknown characteristics of PPG signals. Fourthly, analysis on PPG signals from the smartphone is mainly done *offline* and as such, a real-time implementation is desired.

This thesis introduces a smartphone application that tackles these challenges and provides PPG signals of high quality. A real-time multi-stage PPG qualification pipeline combined with a pulse segmentation algorithm is proposed. Furthermore, analysis of the camera settings and finger detection algorithm resulted in PPG signals with 12.63% higher quality than a dedicated PPG sensor.

Preface

The human body is more ‘talkative’ than most people know. There are so many things that can be inferred by simply listening, looking and measuring the signals it produces. All that is needed, is a method to detect these signals, which in this day and age is possible for anyone with the right tools. Especially in a world where monitoring our health has become important to prevent rather than cure an illness. This forms the basis of this thesis, a beautiful fusion of biology and technology. However, the available tools are expensive or offer limited quality, which hinders further research. The great musical minds of the previous centuries were all limited to the same set of instruments, yet composed many different pieces that are still played today. The smartphone in the 21st century is what these classical instruments were in the past: tools accessible to many people such that only greatness can ensue. I hope to support people in achieving their goals with the system I have developed.

Regarding great people, I would like to thank my supervisor, Marco Zuniga, for all the guidance he has provided me within the past 11 months. Secondly, I would like to thank Eric Wang for introducing me to this research area as it is different from conventional topics related to embedded systems. During our face-to-face meetings and later online meetings, I have learned a lot and hope I have returned the favor. Next, I would like to thank the people of our group, for the high-level weekly discussions we held. I would also like to thank my dearest brother, Maksym, for his valuable input and support during my research and thesis writing.

Finally, I would like to thank my family for their unconditional support and faith. Without them, I would not be the person I am today.

Marek Lukasz Vette

Delft, The Netherlands
1st October 2020

Contents

Preface	v
1 Introduction	1
1.1 Motivation	2
1.2 Contributions	3
1.3 Thesis Organisation	3
2 Background	5
2.1 PPG signal	5
2.1.1 Introduction	5
2.1.2 Reflective and Transmissive Modes	7
2.2 Physiological Interpretation Cardiac Pulse	7
2.3 Red and Green Wavelength Comparison	8
3 Related Work	11
3.1 Related Studies that Use PPG	11
3.2 Related Smartphone-based Studies	13
3.2.1 Trivial Physiological Features	13
3.2.2 Advanced Physiological Features	13
3.2.3 Novel Application of PPG	15
3.2.4 Summary	15
4 System Overview	17
4.1 System Design	17
4.2 System Diagram	19
5 Camera Settings	21
5.1 PPG Evaluation Metrics	21
5.2 Smartphone Camera Settings	22
5.2.1 Video Recording	22
5.2.2 Light Exposure	23
5.2.3 Image-Processing	24
5.2.4 Region of Interest Selection	26
5.2.5 Summary	29
5.3 Evaluation	30
5.3.1 Frame Resolution	30
5.3.2 Frame Jitter	31
5.3.3 Light Exposure	32
5.3.4 PPG Signal Generation	33
5.3.5 Conclusions	34
6 Raw PPG Signal Analysis	35
6.1 Fingertip Detection	35
6.1.1 Methods	35

6.1.2	Evaluation	36
6.1.3	Solution	36
6.2	Finger Pressure Detection	37
6.2.1	Methods	37
6.2.2	Method Description	39
6.2.3	Limitation	39
6.2.4	Results	41
6.3	Conclusions	41
7	Pulse Segmentation	43
7.1	State of the Art	43
7.2	Proposed Method	46
7.2.1	De-Noising	46
7.2.2	Systolic Peak Detection	47
7.2.3	Valley Detection	48
7.2.4	Peak & Valley Decision Logic	48
7.3	Conclusions	50
8	Pulse Qualification	51
8.1	Pulse Normalization	51
8.2	Double Pulse Problem	51
8.2.1	Detection	52
8.2.2	Separation	52
8.3	Systolic Pulse Part	53
8.3.1	Rise-time Correction	53
8.3.2	Discarding Incorrect Pulses	54
9	Evaluation	55
9.1	Measure of Pulse Variation	55
9.2	Acceptance Rate	56
9.3	Impact of Individual System Stages	57
9.3.1	Experiment Setup	57
9.3.2	Summary of Improvements	59
9.4	PPG Sensor versus Smartphone App	60
9.4.1	Experiment Setup	60
9.4.2	Results	60
9.4.3	Summary of Improvements	63
10	Conclusions	65
10.1	Conclusions	65
10.2	Future Work	66
A	Image Processing Pipeline	67

Chapter 1

Introduction

The introduction of the first smartphone in 2007, the iPhone, would revolutionize the world. Its ease of use, high mobility and compactness, while packing good computational performance made its applications endless, without even mentioning the large number of sensors embedded into the device. Many types of sensors such as cameras, accelerometers, gyroscopes, thermometers, air humidity sensors, barometers, fingerprint sensors, compass, you name it and the smartphone has it. Besides the original intent of these sensors, such as cameras for photography, their applications are also extended to other domains. Accelerometers and gyroscopes are used for e.g. steering in racing games or counting steps, the microphone can be used to interact with the device for voice command or speech-to-text, but for this Master thesis, the camera is the point of focus as it can be used for more than only photography.

In the 1930s, a method was developed for obtaining the heart rate. For this method, a light source is placed on a part of the skin, usually the fingertip, and is illuminated, whereafter a photoreceiver next to the light source captures the remainder of the light after absorption due to the skin and underlying tissue. The idea was that, as blood pumps through the finger at the rate of our heart beat, the amount of light being absorbed and reflected would change at the same rate, which is observed by the photoreceiver. This method is called *Photoplethysmography*, abbreviated as PPG. This simple yet elegant method can also be used with a smartphone, which was discovered by researchers in 2010. The smartphone is equipped with a camera (the photoreceiver) and a flash light (light source) that is often located next to the camera. By placing the finger over the camera and flash light, PPG can be performed similarly as was done at the time of its discovery. Besides the heart rate, several other health related features such as blood pressure, glucose level, blood oxygen and respiration can be extracted from the PPG signal.

Another important discovery related to PPG is the shape of the signals that are extracted. Essentially, a PPG signal consists of a continuous repetition of the same shape, *the pulse*, at the same rate as the heart rate. This shape, called the *pulse waveform*, has a somewhat general shape but contains *unique* features for individuals. What this means is that everyone has a slightly different and unique waveform, which elevates the field of PPG sensing. Not only do health-feature algorithms have to take this into account, but it also opens up a field not related to health monitoring, security, since the pulse waveform stays ‘relatively’

constant over long periods of time. Similarly to other *biometric* features such as the fingerprint and iris, the PPG pulse carries information that allows for authentication/identification of individuals.

However, both health-monitoring and security applications need clean PPG signals in order to function correctly. Obtaining clean PPG signals is challenging. Medical measuring devices at hospitals use expensive but extremely accurate sensors, whereas the smartphones use off-the-shelf equipment, using sensors that are not initially intended for health measurements and often suffer from noise.

1.1 Motivation

As has been emphasized, the quality of the PPG signal makes or breaks the application. Several effects play a direct or indirect role in the signal quality.

Known influences are related to motion and noise artifacts (MNAs), such as body movement, that disturb the contact between camera and measurement site. The contact pressure between the camera and the finger is often named as a crucial artifact, but is not properly explored. Originally, medical PPG sensors use finger-clips to alleviate these movement and pressure variations, but this is not available for the smartphone. Many studies [49],[4],[7] introduce add-on devices for the smartphone such as pressure sensors, light sources, photodetectors and more to improve their results, but the addition of hardware completely defeats the idea of using the smartphone. As such, *the first challenge is to develop a system that overcomes these MNAs using only the bare smartphone*. Here the focus is on MNAs that are more subtle, such as finger pressure, which are hard to regulate for users by themselves. A user can be instructed to not move during measurements, but still minor movements will be present and controlling the level of applied finger pressure is difficult.

Another challenge is related to the smartphone itself. There are so many different devices, all with different cameras and hardware that can affect the PPG signal extraction process. Furthermore, there is such a wide range of possible camera settings that can be used, but that is not properly explored in literature. *Thus the impact of the camera settings on the PPG signal needs to be explored.*

The third challenge was already introduced, which is related to *the uniqueness of the pulse waveform for individuals*. The system is developed for unknown users of whom no data is available, yet the developed algorithms need to be robust and work for any person. An extensive study on the pulse waveform is required to understand what general features are shared by all individuals.

Fourthly, the system needs to operate in real-time. Many studies in literature record videos of the PPG signal with a smartphone, transfer the video to a PC and process it in an *offline* environment to evaluate their methods. However, these signals are obtained under circumstances that do not mimic the situations for which the method is intended.

To conclude, *the goal of the system is to obtain clean PPG signals under real-time conditions*. In order to achieve this, the aforementioned challenges need to be overcome.

1.2 Contributions

In this thesis, a real-time smartphone application is proposed that tackles the aforementioned challenges. Besides building on existing studies, the following contributions are made:

- An analysis of the available smartphone camera settings and their impact on the PPG signal (Chapter 5).
- A fingertip detection method to automate the signal extraction process (Section 6.1).
- A pulse segmentation algorithm to extract the cardiac pulses from the PPG signal (Chapter 7).
- A pulse qualification algorithm to improve the pulse quality and discard corrupt pulses (Chapter 8).
- A real-time smartphone application that implements all the above.

1.3 Thesis Organisation

In order to understand the developed system, first an understanding of the concepts related to PPG is required, which will be discussed in Chapter 2. Chapter 3 discusses the related methods developed in literature. Then, in Chapter 4, an overview of the developed system and smartphone application is given. Hereafter, the modules are explained in detail in the following chapters, starting with the evaluation of the different possible camera settings in Chapter 5. This is followed by Chapter 6, covering the first evaluation of the quality of the PPG signal as it comes from the camera. Hereafter, Chapter 7 discusses the extraction of the PPG pulses from the signal which then need to be tested on their quality, covered in Chapter 8. In Chapter 9 the proposed system is evaluated, followed by Chapter 10 which summarizes the observations, conclusions and future work. Appendix A goes into detail on the image processing pipeline, which is related to the camera settings of the smartphone.

Chapter 2

Background

This chapter serves as an introduction to the fundamental concepts needed to understand the proposed system. The first section (Section 2.1) describes the different aspects related to the generation of the PPG signal, the different types of applications and how it has been used since its discovery. Hereafter, the physiological meaning of the contents of the PPG signal are discussed (Section 2.2). The chapter is concluded with a comparison of PPG signals obtained from the green and red light spectrum in Section 2.3, with a revision of the PPG model as it has been known to date.

2.1 PPG signal

2.1.1 Introduction

Photoplethysmography (PPG) is a technique where changes in blood volume in the micro-vascular tissue bed can be monitored via an optical light source and detector. Light in the human body is transmitted, reflected or absorbed, depending on multiple factors such as the properties of the different skin layers, the presence of capillaries, arteries or veins. PPG has been around for a long time, ever since the term has been coined by A.B. Hertzman in 1937 [16]. The relationship between blood volume change and light had already been demonstrated in 1935 by K. Matthes, father of the pulse oximeter [36], but the actual term PPG as it is used today originates from Hertzman's work. Hertzman demonstrated that the PPG signal is divided into two components: A relatively static (DC) component and a dynamic (AC) component due to blood volume changes in the arteries. As the heart pumps blood into the arteries, the blood volume at a point under observation changes. This is illustrated with figure 2.1. The blood volume present at a certain time, is inversely proportional to the amount of light received at the detector. This is because when more blood is present, it absorbs more light and thus less light is transmitted, scattered or reflected to the detector.

Over decades of research, the principles of PPG have stayed the same, but the sensor implementations have changed continuously. Hertzman used a single 'pencil flashlight bulb' and photoelectric cell [17]. Matthes [36] used two light sources with green and red wavelengths to obtain information about the blood-oxygen levels. Later works by T. Aoyagi in 1972 [43] demonstrated that com-

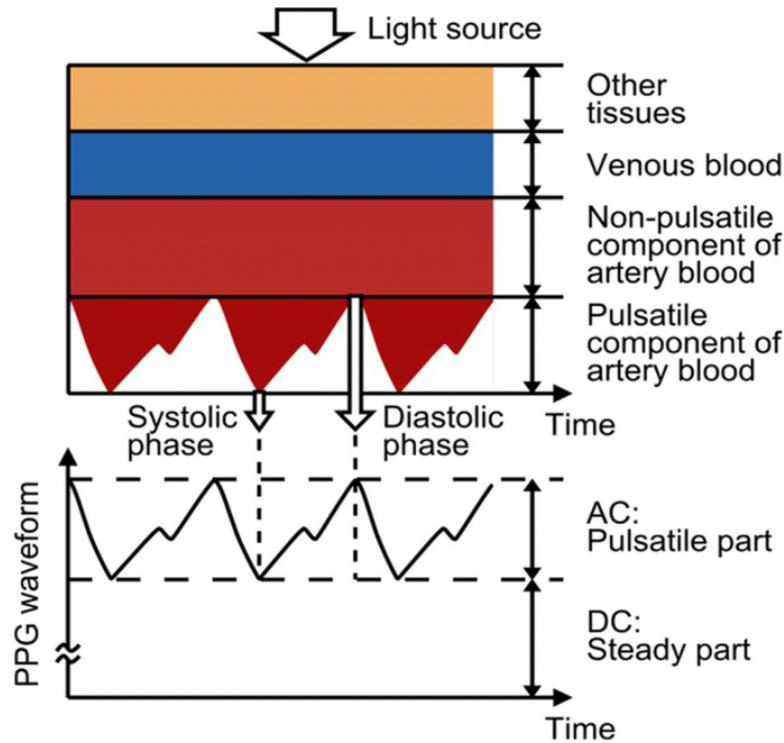


Figure 2.1: The differentiation of the AC and DC components of the PPG signal, together with their sources of origin [47].

paring the difference in blood absorption by red and infrared light would prove to be more accurate, which became the foundation for commercially-used pulse oximeters as they are known today. In 2010, Pelegris et. al. proposed the first PPG heart rate detection method using a smartphone [37]. After this discovery, a surge in interest for mobile PPG sensing was seen, but several challenges had to be overcome.

From the PPG signal, information can be extracted that can give an estimate of a wide range of clinical features such as heart rate, respiration and blood oxygen levels but also helps in the detection of arterial diseases or even aging. The medical applications of PPG have been extensively studied, but more recently other types of applications such as security have become popular. Liu et. al. showed that the biometric information carried by the PPG signal can be used to differentiate between people [30]. Although the types of applications vary, a ‘clean’ PPG signal that is noise-free is desired in all cases. If the quality of the PPG signal cannot be guaranteed, it could have disastrous consequences for the application.

Since light consists of different wavelengths that interact differently with the skin and tissue of our human body, an evaluation is required on which light source and receiver should be used. These transmitter and receiver related properties will be discussed in the following sections.

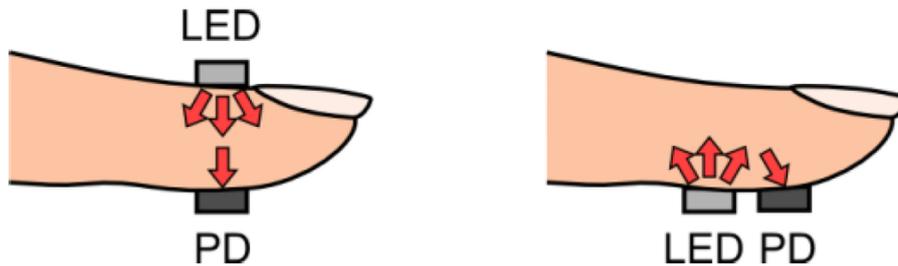


Figure 2.2: Transmissive (left) and reflective (right) PPG signal measuring modes [48].

2.1.2 Reflective and Transmissive Modes

The light emitted by the transmitter propagates through the skin, tissue and blood of the human body and is either reflected, absorbed or transmitted. There exist two modes for capturing the light after it has propagated: reflective mode and transmissive mode. In reflective mode, the transmitter and receiver are on the same side of the skin and the receiver observes the light that has passed through the human body and is reflected back to the skin surface. In transmissive mode, the transmitter and receiver are opposite of each other, such that the receiver mainly observes the light capable of propagating through the human body. Figure 2.2 showcases both applications. Transmissive mode is usually limited to thin areas such as fingers or earlobes due to the fact that large amounts of tissue would absorb all light. Reflective mode on the other hand can be applied to any body area and has the additional advantage of being on the same side, which means they can be incorporated in a single device. An important dual relationship exists between the chosen mode and wavelength for the light emitter. In transmissive mode the light needs to penetrate e.g. the entire finger, which can only be accomplished with a large wavelength such as infrared (IR) light or with a powerful light source.

2.2 Physiological Interpretation Cardiac Pulse

As has been mentioned previously, the captured PPG signal represents the volumetric blood changes within the veins and arteries of the finger. These changes in blood volume are a result of the systolic and diastolic pressure, exerted by the heart as it pumps. Since the pumping of the heart happens periodically with the same sequence of physiological steps, it is reflected in the PPG signal obtained by the smartphone camera. The sequence of physiological steps can be fitted in a single period: the cardiac pulse or cycle. These cardiac pulses contain valuable information. It is important to understand the meaning and origin of the captured PPG signal. Figure 2.3 identifies the key characteristics of a single pulse.

At the start of a cardiac pulse, the heart pumps and the blood-volume rapidly increases up until a maximum has been reached: the systolic peak. Hereafter, the blood-volume slowly declines until the diastolic notch. At this moment, the pressure wave that accompanied the blood pump is reflected back and the blood-

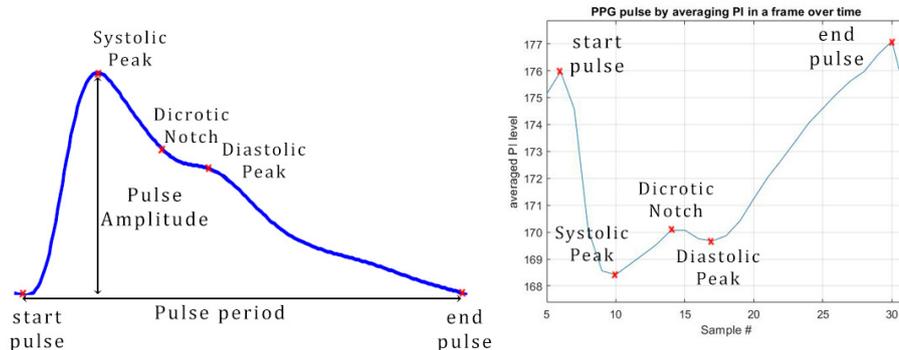


Figure 2.3: **(Left) Theoretical PPG pulse with characteristics. (Right) PPG pulse extracted from the smartphone camera.**

volume slightly increases till the diastolic peak. After the reflected pressure wave declines, the blood-volume decreases accordingly, until a minimum is reached. The process is repeated with the next heart-pump.

An important aspect of figure 2.3 is that the left pulse represents the blood volume present at the measurement location over time. This is the general morphology used in literature. In contrast, the right figure shows a PPG signal captured with a smartphone camera. In this case, the PPG signal represents the averaged pixel intensity (PI). At first sight, it seems that the PPG signal is mirrored. Recall the different physical quantities the left and right images in figure 2.3 represent. As the heart pumps blood into the aorta, blood will flow into the finger, which is being observed. As the blood volume increases, more light will be absorbed. This is visible in the right figure where the pixel intensity decreases. Once the blood starts flowing and the blood volume decreases, the pixel intensity increases which confirms our hypothesis that blood volume and pixel intensity are inversely related. This is a fundamental aspect to keep in mind when considering the physiological properties of the PPG signal. Some papers, such as [49], incorrectly identify the peaks in the captured PPG signal from the smartphone camera as systolic or diastolic, because they do not account for the relation between pixel intensity and blood volume.

2.3 Red and Green Wavelength Comparison

As was mentioned in Section 2.1 all that is needed for obtaining PPG signals is a light emitter and receiver. Furthermore, a wide range of wavelengths is available that can be used for the light emitter. Two important factors play a role in the wavelength selection:

- Penetration depth
- Skin color

The red wavelength penetrates more deeply than green, as it is a longer wavelength, but also due to lower absorption by skin and tissue. The skin and tissue consist of multiple layers with arteries, veins, arterioles and capillaries, that all

have an effect on the PPG signal [31]. Red light reaches these deeper layers and thus contains much information whereas green light only reaches the superficial layers. Skin color is mainly determined by a chromophore called melanin which is a strong absorber of green light. Darker skin contains more melanin, and thus very little light reaches the blood vessels, whereas red light is much less absorbed [29]. Each wavelength carries different pieces of information.

That’s why medical equipment such as pulse oximeters use specific LEDs with the desired wavelengths for e.g. red (670nm) and IR (779nm) light. After scattering and reflection, the light detector measures the changing absorbance at each of the wavelengths over time. By comparing the PPG signals, the blood volume change of only the arteries can be deduced, excluding the blood volume changes in the capillaries and arterioles, and constant factors such as skin, muscle, bone and fat.

In the smartphone implementation of PPG signal extraction, things work differently. Firstly, there is a (usually white) LED with a wide range of emitted wavelengths, differing per device. Secondly, the camera is used as a detector and the captured light is mapped to three channels with a range of [0-255] values: the red, green and blue (RGB) channels. At the time of the introduction of the smartphone, multiple studies have tried to implement and design methods to detect heart rate, measure blood pressure, blood oxygen etc. During this time, a comparison was made for the different PPG signals obtained from each RGB channel. Some studies use Signal-to-Noise Ratios (SNR) [35] or AC-DC ratios [34] to identify the best performing channel. They reported the green channel as the best channel without accounting for the information each channel carries nor giving an explanation for scoring better on their defined metric. Due to the limited penetration of the green light, few disturbances are collected, resulting in a ‘good’ SNR value. For the red wavelength, the signal reaches much deeper and captures more information, but is more ‘chaotic’ and thus a higher signal value is accompanied with more noise. Table 2.1 summarizes the observations with regards to used wavelengths.

Wavelength	Penetration depth	SNR	Information captured
Red	High	Low	Large
Green	Medium	High	Small
Blue	Low	High	Small

Table 2.1: **Red light penetrate deeper and thus obtains more information, yet is more chaotic and scores low on SNR metrics.**

For heart rate measurements, only the peak in the PPG signal is of interest, which means that the remainder of the signal can be as noisy as possible, as long as the peaks can be extracted. Furthermore, any other information related to the skin, vessel, bone or muscle structure is simply irrelevant, which is why the earlier smartphone PPG methods used green light.

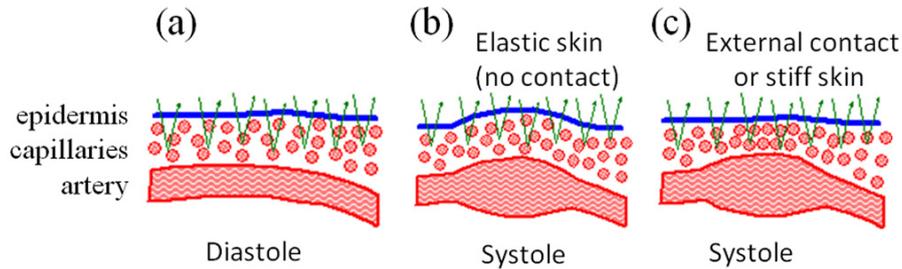


Figure 2.4: **Origin of the PPG waveform for green light [23].**

A fundamental paradox remains with the PPG signal obtained from the green channel, since it cannot reach the arteries yet produces a ‘stronger’ PPG signal. This paradox was explored by [23], where they proposed a new physiological model of the green PPG signal, displayed in figure 2.4. Recall from Section 2.2 that the blood is pumped into the arteries during the systole phase and as such the blood volume at the measurement site increases, expanding not only the diameter of the artery, but also the skin as can be seen in 2.4(b). After the blood flows out of the artery it decreases in size. The PPG signal obtained from the green channel does not originate from these blood-volume pulsations because of the short penetration. In the case of smartphone PPG signal extraction, the finger is pressed against the lens and the situation as is illustrated in figure 2.4(c) occurs. Due to the pressure from the blood during systole and the external pressure from the finger, the capillaries situated in the dermis are compressed. Due to this compression, the density of the capillaries changes and as such the absorption and scattering coefficients change. Thus the PPG signal captured by the green channel is an indirect result of the pulsating pressure of the arteries and external pressure of the finger [23], which explains why the heart rate can be derived.

In conclusion, the red channel carries a lot of information with regards to the blood-volume changes in the arteries but also the structure of the human tissue, blood vessels, bone, muscles and fat. These are all necessary for applications such as identification and/or authentication via PPG. Therefore, the remainder of this thesis focuses mainly on the red channel.

Chapter 3

Related Work

This chapter serves as a guide to understand the relevance and contributions of this thesis to the mobile PPG sensing field. First, Section 3.1 gives an overview of all the different types of studies that have been done, and the different aspects on which they differ. Then in Section 3.2 a more detailed overview is given of the specific sub-field that this thesis fits in and all the comparable state of the art (SoA) work that has been done. The relevant studies are shortly discussed and their shortcomings are listed, which this thesis tries to tackle and improve. Throughout this thesis, these studies should be kept in mind, because some of them are used as benchmarks to quantify the improvement of the methods developed here.

3.1 Related Studies that Use PPG

Because photoplethysmography (PPG) is so simple to apply, you only need a light source and a photoreceiver, its applications are wide-spread and there are multiple ways to classify them. Some of these aspects are inter-dependent, but for now they are considered as separate categories. The first few aspects are related to sensing itself, starting with the **application method**. PPG signals can be obtained in a *remote* setting, for instance using a camera or a webcam that is monitoring someone's face [38]. The other method is using a PPG sensor set-up that requires *contact* with the person [4],[49],[7],[50]. Since both methods differ greatly they also offer different challenges. For example, contact-based methods need to consider the pressure and movement between sensor and skin, whereas remote-based methods need to consider ambient light and distance or even privacy given that it is essentially camera monitoring.

The second aspect is related to the **photoreceiver**. Originally, *photodiodes* were used to detect changes in light intensity, which are converted to an electric current [16]. Later, the *camera* present in smartphones or webcams became an interesting alternative, since they are omnipresent in the current digital era [37]. The challenge here, is that many factors need to be considered regarding the extraction and the quality of PPG signals. Instead of an electric current, pixel values in consecutive images are evaluated. There are many of them as opposed to one in the photodiode case and each pixel measures a different light intensity, which makes combining their values challenging. Furthermore, the large number

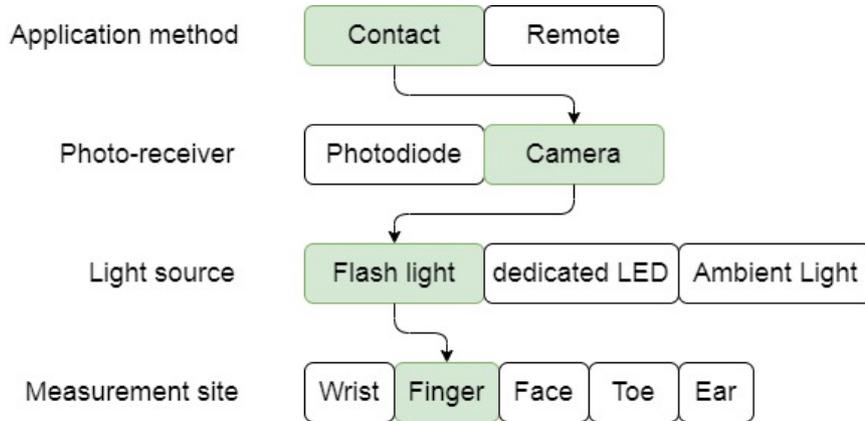


Figure 3.1: Classification of the PPG research field with aspects relevant to this thesis highlighted in green.

of different cameras can affect the PPG sensing differently, making it difficult to develop a single method that works for any device.

The third property is related to the used **light source**. The skin and underlying tissue interacts differently based on the wavelength of light that is projected on to it. It determines the penetration depth, rate of absorption, reflection or transmission and as such, the information collected by the photoreceiver is light-spectrum dependent. To tackle this, conventional medical equipment uses light sources with a *single dedicated wavelength* such as red or infra-red. However, smartphones use the flash light usually located next to the back-end camera, which is normally white and consists of a wide range of unknown wavelengths with different power distributions. Similarly to the camera, smartphones are equipped with different types of flash lights making it challenging to establish a relationship between light source and photoreceiver. The previously mentioned light sources are active, whereas there also exist passive light source PPG techniques that use ambient light, but this is limited to remote sensing [38].

Slightly related to the sensor hardware is the fourth property, which considers the **measurement site** of the human body where PPG signals are extracted. Remote sensing techniques are mainly limited to the *face and forehead* [38], whereas contact-based methods can be applied to any place of the body although the PPG signal quality and information differ. The most common methods measure at the *finger* [49], *wrist* [44] or *ear* [5] as they are located on the periphery of the human body, are easily accessible and the blood vessels are located close to the skin surface.

The diagram in figure 3.1 gives a summary of the classification aspects and highlights the aspects related to this thesis. *The focus will lie on contact-based methods using the smartphone camera and flash to obtain PPG signals from the finger.*

3.2 Related Smartphone-based Studies

In the sub-field of PPG sensing outlined by figure 3.1, several studies have a large overlap with the goal outlined in the introduction (Chapter 1), namely the extraction of high quality PPG signals. Albeit this goal is the same for all studies, it is not a goal on its own. By having a method that provides high quality signals, the measurement of physiological features such as heart-rate, blood pressure and blood oxygen can be more accurately estimated. Other studies require high quality signals to differentiate between individuals for authentication and identification methods. As such, the challenge of obtaining clean PPG signals is the same yet achieved differently due to the application goals being different. The most important works will be discussed in the following sections and are divided based on the degree of PPG quality. Trivial physiological features such as heart rate estimations are more robust and allow for more noise and distortion as opposed to advanced features such as blood oxygen or glucose estimations. The last category is the non-medical variant related to security, which requires equal or higher quality PPG signals than the advanced physiological features. At the end of this chapter, a summary of the most important studies is presented.

3.2.1 Trivial Physiological Features

The early works on smartphone PPG sensing after its invention in 2010 were focused on obtaining accurate heart rate estimations [37]. This is most commonly done via a time or frequency analysis of the PPG signal. For time analysis, the systolic peaks are identified with a peak detection algorithm. Then, the time between subsequent peaks is used to estimate the heart rate [41]. For frequency analysis, a Fast Fourier Transform (FFT) is applied on the PPG signal [37]. The general idea is that the most dominant component in the spectral density distribution corresponds to the heart rate, as the PPG signal is a repetition of the same pulse with the same period as the heart rate. Another trivial feature is the respiration rate, which was also done by [41]. Due to the triviality and leniency towards noise and distortions, the methods proposed by these studies cannot be used in more quality-demanding domains. Furthermore, these studies are from a time when PPG sensing with the smartphone was novel and as such not everything was properly explored.

3.2.2 Advanced Physiological Features

After the initial studies on the trivial features, more interest was put into implementing methods that obtain other physiological features, normally done with medical equipment. Examples of these are:

1. Hemoglobin estimations with HemaApp [49] - 2016
2. Blood oxygen level estimations with PhO2 [4] - 2017
3. Blood pressure estimations by Chandrasekhar et. al. [7] - 2018
4. Blood glucose estimations by Zhang et. al. [50] - 2020

HemaApp is a smartphone application that estimates hemoglobin concentrations by evaluating the PPG signals obtained at different wavelengths, which all represent different biometric information as was explained in Section 2.1. From these PPG signals, 5 features need to be extracted and provided to a regression model (SVM) to produce an estimate of the hemoglobin levels. However, the authors propose an LED attachment instead of using the flashlight present on the smartphone, because their method requires IR light (970nm) to function. Furthermore, some camera settings such as white balance, exposure and camera sensitivity are only superficially explored and no concrete parameters are given.

Another smartphone application, PhO2 [4], was developed to measure blood oxygen levels which are referred to as SpO2 levels. PhO2 is very similar to HemaApp as they also use an attachment for the smartphone and have a complicated algorithm to determine the SpO2 levels. PhO2 recognizes the problem of using IR light in smartphone applications such as HemaApp, as the flash light is unable to produce this wavelength. To accomplish this, an attachment is introduced to differentiate between red light (670nm to 690nm) and near IR (NIR) light (700nm to 779nm). For this, two optical filters are designed with a divider that splits the red and NIR light which are then mapped by the camera to two separate sides of an image. Two separate PPG signals are obtained from the two areas, representing the red and NIR light intensity respectively. They use a region of interest (RoI) technique to only select a part of the area for evaluation, to preserve the quality of the PPG signal. Moreover, PhO2 is one of the first to propose a finger pressure detection that only uses the information of the PPG signal. No additional pressure sensors are used and as such their method will be evaluated in this thesis. On the other hand, evaluation of camera settings is largely neglected, but then again the camera is circumvented by using the attachment to do the bulk of processing.

The third application that will be discussed is a blood pressure estimation technique by Chandrasekhar et. al. [7]. Although it also uses an attachment just like the previous 2 methods, it is quite different. HemaApp mainly focused on tackling the noisy quality and limitations of the smartphone camera and flash, which form the *camera sensor*. PhO2 does the same but also considers finger pressure as an important factor. Chandrasekhar et. al. completely ignore the smartphone camera and flash and use a separate PPG sensor, only using the smartphone to display data. However, they use a force sensor to detect the finger pressure applied by the user. Via visual feedback, the user is instructed to obtain an optimal contact pressure. Furthermore, the user is guided in the placement of the finger by the attachment. These proposed solutions in addition to the finger pressure detection by PhO2 try to tackle the *human related artifacts*.

The previous methods try to overcome the smartphone limitations with a custom hardware attachment, because the quality of the signal needs to be similar to medical devices. This defeats the entire point of using a smartphone, namely, its ease of use, mobility and ubiquitous presence. By introducing additional hardware, users need to go out of their way to get this hardware, whereas smartphone-exclusive based solutions require no additional work from the user besides downloading the app. The apps can only be used in combination with

the hardware, which the user is required to carry with them. That is why Zhang et. al. [50] introduced a method that can differentiate between three glucose levels, using PPG signals obtained with the bare smartphone camera. The goal was to achieve the same accuracy as with a regular PPG sensor. They use an algorithm to detect the presence of a fingertip to initiate measurements and propose a PPG extraction algorithm that combines information from the RGB channels in images. By not adding an attachment to improve the sensor quality, the settings of the smartphone camera become more important than previously was the case. However, the authors completely neglect the camera settings to possible boost the PPG signal quality. The consequences of this will be discussed in 6.1.

3.2.3 Novel Application of PPG

Whereas the previous SoA methods were focused on conventional health monitoring applications, the work here introduce a new field for PPG signal usage. CardioCam [30] by Liu et. al. proposes a method to identify and authenticate humans based on their PPG signals. Using large amounts of features extracted from a PPG pulse, a unique biometric cluster can be created to distinguish between individuals. The authors of CardioCam use a fingertip detection algorithm to automate the start of measurements. In order to achieve similar results as with PPG signals obtained from medical devices, they evaluated the impact of several camera settings. However, CardioCam only records a video with only two dynamic camera settings: ISO and flash intensity. The other camera settings such as light exposure or white balanced are locked to unknown values. There is no explicit mentioning that the system is a smartphone application.

3.2.4 Summary

Based on the previously described studies, there seems to be a lack of research in an area that covers *online methods* that tackle both the *camera sensor* and *human artifacts* without the usages of attachments. The methods that do not use *attachments* have only done superficial research on the camera settings or operate in an offline setting. The SoA works from this section and the proposed method in this thesis can then be classified based on these metrics as shown by table 3.1:

Author	Year	Description	Meth.	Attach.	Camera Sensor	Human Artifact
Wang [49]	2016	HemaApp: Hemoglobin	online	Yes	Yes	No
Bui [4]	2017	PhO2: blood Oxygen Level	online	Yes	Yes	Yes
Chandrasekhar [7]	2018	Blood Pressure	online	Yes	Yes	Yes
Zhang [50]	2020	Blood Glucose	offline	No	Yes	No
Liu [30]	2019	CardioCam: Id+Auth	offline	No	Yes	Yes
Proposed method	2020	Pulse Qualification	online	No	Yes	Yes

Table 3.1: **State of the art methods using smartphone PPG sensing.**

Chapter 4

System Overview

This chapter describes the developed system. Firstly, a slight introduction is required to elaborate on the design choices and system structure in Section 4.1. Then in Section 4.2 a general overview of the entire system is given. Hereafter, each module of the proposed system is discussed shortly, whereas the extensive analysis is reserved for Chapters 5, 6, 7 and 8. The chapter concludes with a couple of illustrations to showcase the developed smartphone application.

4.1 System Design

In order to understand the design decisions, a revision of the thesis problem is required. The goal of this thesis is to design a system that operates in real-time and obtains high quality PPG signals. As the related studies have shown in Section 3.2, two important challenges need to be overcome:

1. Camera sensor configuration
2. Human related artifacts

The former consists of all factors related to the light source, photo-receiver and the sensing of the blood volume change via light intensity change. Since the smartphone hardware will not be changed as was done in prior SoA studies, the solutions proposed in this thesis are software-based and related to the settings of the smartphone camera. The idea is that *there exists an optimal camera sensor configuration* for obtaining clean PPG signals. The human related artifacts affect the quality of the PPG signal by introducing distortions. Examples of these influences are body motions, finger movement, finger placement, finger pressure and respiration. A repeating pattern in these influences is that they are all related to motion or displacement of the contact between sensor and finger. Some of these *motion artifacts* can be prevented by instructing the user before measurements. Finger pressure however, is an artifact that is unconsciously applied by the user and is hard to self-regulate. As such, mechanism need to be in place to detect finger pressure and *inform* the user to increase or decrease this pressure.

Up to this point, the importance of clean PPG signals is clear but how the degree of quality can be measured is not discussed. As was stated before, clean

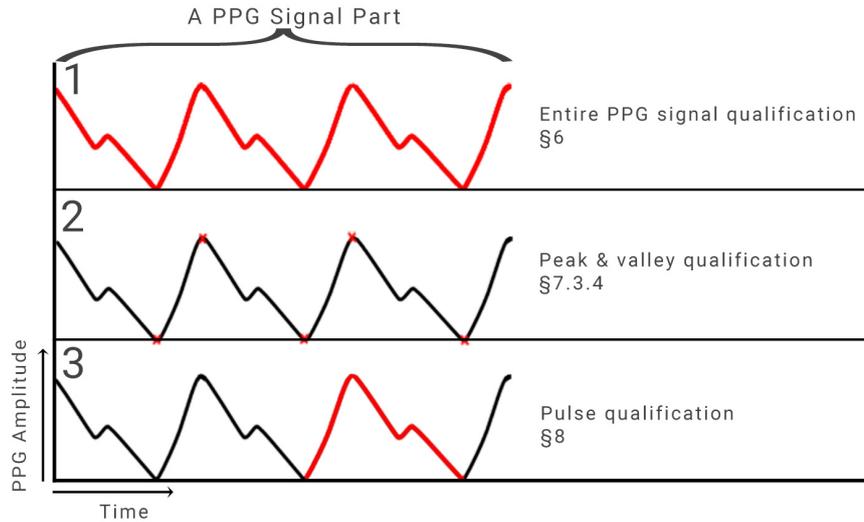


Figure 4.1: **The three qualification stages proposed in the system. From top to bottom: 1) Finger pressure detection with frequency analysis of the entire PPG signal part. 2) Qualification using the peaks & valleys after pulse segmentation. 3) Qualification of the single pulses.**

PPG signals are a means to an end and only improve the performance of applications. A framework is required to quantitatively measure the PPG signal quality. The PPG signal consists of a repetition of cardiac pulses as was shown in 2.2 and the information provided by these pulses is used. However, to obtain these pulses from the PPG signal, a segmentation algorithm is required. There exist several algorithms for pulse segmentation and at the time of writing, research on this topic continues. Moreover, these methods need to be able to handle and detect corruption of the PPG signal.

Another important design choice is related to time and computing power as they are a precious resources that should be conserved. As such, the testing of the signal quality is spread out over the several modules in the system. Via this way, corrupt signals can be detected at an earlier stage, preventing the continuation of processing a corrupt signal. This spread of signal qualification over multiple modules has a structure to it, illustrated by figure 4.1. After a couple seconds of PPG signal generation, first the finger pressure is checked by evaluating the generated PPG part, shown by **stage 1**. Then if a good finger pressure is detected, the PPG signal can be segmented which results in a couple PPG pulses. A PPG pulse consists of a start, end and systolic peak. The start and end of a pulse are always the lowest values and will be called ‘valleys’. Qualification can be done using only these three features of all the identified pulses in the PPG part shown by **stage 2**. The entire PPG part is discarded if this stage is not passed. The last step, **stage 3**, is the actual pulse qualification. Here, all data of the pulse is used to check whether the pulse is clean or corrupt. This is in contrast with the previous qualification stage which used only three features of the pulse and discards all pulses if the PPG part is corrupt.

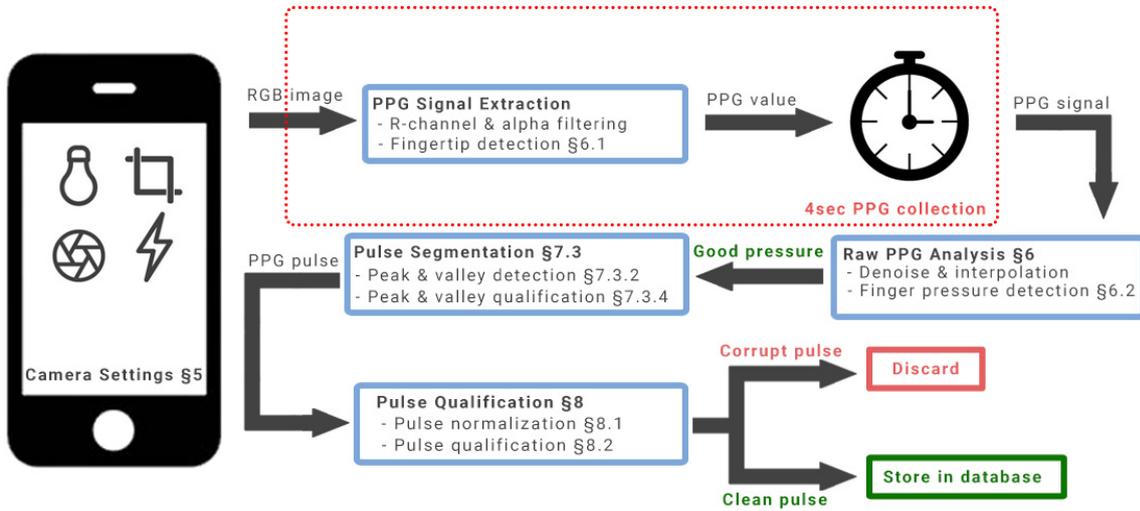


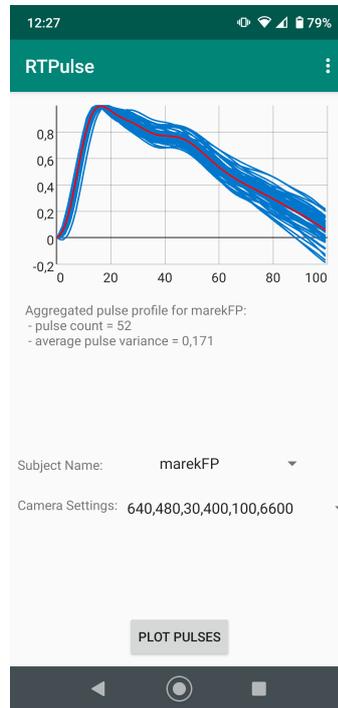
Figure 4.2: The subsequent stages that encompass the PPG processing system inside the smartphone.

4.2 System Diagram

With the different design choices in mind, an illustration of the proposed system is shown in figure 4.2. First, the camera settings related to image capture are configured. Examples of these settings which will be extensively discussed in Chapter 5. After an image is captured, a PPG value needs to be constructed based on all the pixel intensity values in the image, which is done by the **PPG Signal Extraction** module. Only the red color channel of a pixel is used for the construction of the PPG value. If no fingertip is present the system will simply ignore the signal and wait for the next image. After 4 seconds of signal capture and PPG value extraction, the resulting PPG signal is then passed on to the **Raw PPG Analysis** stage (Chapter 6). Here, the PPG signal is filtered, interpolated and the signal is centered around the horizontal axis. Then, a frequency analysis takes place to determine the finger pressure. Only if the finger pressure is qualified as ‘good’, the PPG signal is passed to the **Pulse Segmentation** stage (Chapter 7). Here, the systolic peaks are identified, where each peak corresponds to one PPG pulse. Then, based on the peaks, the start and end of a pulse are identified where the latter two are called ‘valleys’. Based on the peaks and valleys in the 4 second PPG signal, some additional qualification can be done by evaluating the amplitude and time distances of the peaks and valleys. The amplitude distance between a valley and peak represents the pulse amplitude whereas the distance between subsequent valleys or peaks the pulse period. This qualification and segmentation method uses thresholds that are adaptively changed over time to account for variations in e.g. heart-rate. The entire PPG signal is discarded if it fails to pass this qualification stage. In case of a pass, the PPG pulses segmented in the previous stage are then passed 1-by-1 to the **Pulse Qualification** stage (Chapter 8). Here, the pulses are normalized on a



(a) Interface during PPG signal extraction.



(b) Plot of all collected and qualified pulses for a user.

Figure 4.3: Screenshots of the smartphone application.

time and amplitude scale such that they can be compared independent of heart rate and amplitude. This is followed by the actual pulse qualification which consists of two parts:

1. Improve the PPG pulse quality
2. Detect corrupt pulses

Finally, the pulses labelled as ‘corrupt’ are discarded whereas the ‘clean’ pulses are stored in a database.

The system is implemented on an Android smartphone. Figure 4.3 shows two screenshots of the smartphone application. Figure 4.3(a) shows a preview of the camera, real-time statistics, finger pressure feedback and a real-time plot of the extracted PPG signal. Figure 4.3(b) shows the aggregated normalized pulses for an individual with specific camera settings.

Chapter 5

Camera Settings

Section 5.1 introduces PPG metrics that will be monitored to track the effects of the camera settings. Then, the different camera settings available in modern smartphones to configure the camera sensor will be discussed in Section 5.2. Thereafter, the relationship between the different settings and the PPG signal is evaluated in Section 5.3. As to why only these settings have been considered is discussed in appendix A.

5.1 PPG Evaluation Metrics

The first metric is the **PPG amplitude** of the PPG waveform as it is captured over time. The importance of this property is related to Signal-to-Noise-Ratios (SNRs). Recall that the pixel intensities in a frame are averaged, which produces a value between [0-255]. As the blood volume and the captured red color change over time, the values fluctuate between a maximum and a minimum pixel intensity value as can be seen in figure 5.1. The distance between upper and lower envelope in figure 5.1 is referred to as the PPG amplitude.

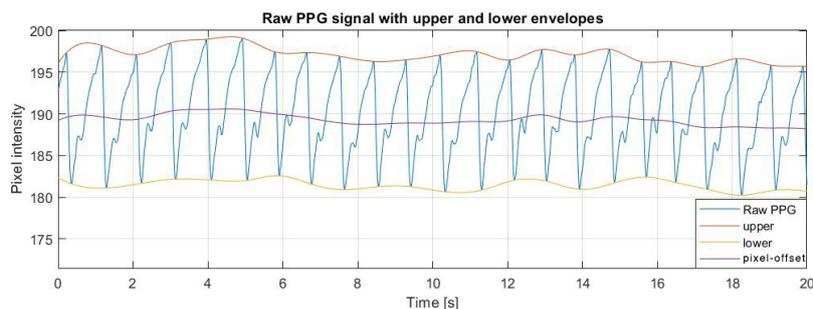


Figure 5.1: **A raw PPG signal with upper and lower envelopes. The line that propagates through the middle of the signal represents the pixel offset.**

The second metric is the **pixel offset**, which defines the level along which the PPG waveform propagates in time and can be approximated with the lower and upper envelope. The physical meaning behind the value itself is related to

how bright or dark the captured images are. Section 5.2 will discuss multiple camera settings that influence the brightness of the captured frames, which can be tracked via this metric.

The third and last property is related to the PPG waveform morphology. Since the morphology of a cardiac pulse is used to identify or authenticate individuals, it is crucial to establish whether the camera settings have any influences. However, at this stage a quantitative evaluation is very complex. The quantitative evaluation of the individual pulses from the PPG takes place in chapter 9, as it needs a segmentation algorithm which will be covered in chapter 7.

5.2 Smartphone Camera Settings

With the introduction of the smartphone, large scale improvements have taken place with regards to photography and video recording. Part of these improvements is related to the customizability of the camera settings, which tries to mimic the features of the analog cameras. The influences of these settings on the PPG signal need to be explored. The following sections discuss the relevant camera settings related to recording, light exposure and image-processing techniques.

5.2.1 Video Recording

Recording encompasses the settings related to the actual recording of the video but unrelated to the color of the pixels. Two main factors are the number of frames collected for the video and the number of pixels in a single frame.

Frame Rate

The PPG signal contains important features that need to be captured, such as the start and end of cycles, systolic and diastolic peaks, and dicrotic notches. A high frame rate is desired, but smartphones come with a maximum supported frame rate. A study by Fujita et. al. [13] showed that a frame rate between [30-60Hz] resulted in comparable PPG signals as for frame rates above 60Hz. As such, a frame rate equal or above 30fps is desired. Another important aspect here, is that the frame rate is an *approximation*. In reality, the time between each frame suffers from a slight variability, called *frame rate jitter*. Since this is mainly device and application dependent, an evaluation is only performed for the used smartphone, a Motorola Moto G7 Plus, in section 5.3.

Frame Resolution

Before frame resolution can be discussed, first an understanding of the camera sensor is required. A digital smartphone camera consists of arrays of light-sensitive areas which are called ‘photosites’. These photosites are often incorrectly called pixels, which causes confusion when used for expressing image resolution. The number and arrangement of these photosites is *fixed* and so is the sensor resolution since they are physical elements. These photosites capture light entering the camera lens and quantifies a value for the light intensity. The information from these photosite can then be mapped to pixels to create an image. The frame resolution is *variable* as it determines how many photosites

are used for the construction of a single pixel value. The sensor resolution varies on a per device-basis and is uncontrollable by the user, whereas the frame resolution is standardized and *can be changed*. Although the details of each PPG extraction method vary, the main component remains the same: the averaging of pixel intensities in a Region-of-Interest (RoI) within the frame, which will be discussed in section 5.2.4. Furthermore, an analysis is required for the benefit of large resolutions, which come at the cost of added computational weight (section 5.3).

5.2.2 Light Exposure

This section covers the different settings affecting pixel intensity values *during frame generation*. Depending on the type of smartphone, different settings related to light exposure can be adjusted manually. The settings related to light exposure are usually introduced as the ‘exposure triangle’ due to their interdependent interaction with light.

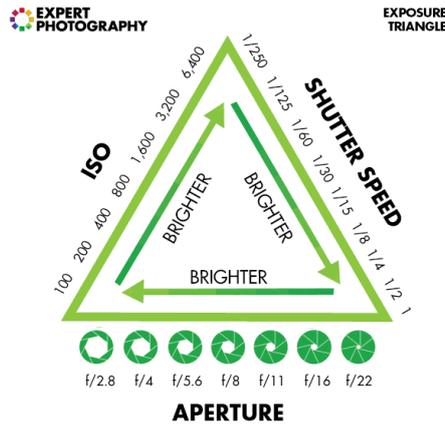


Figure 5.2: Exposure triangle demonstrating the relationship between ISO, shutter speed and aperture aspects [20].

As is shown with figure 5.2, the exposure triangle consists of ISO, shutter speed and aperture, where each parameter will be explained in more detail in the following sections. The combination of these settings affect the PPG amplitude and pixel offset and will be evaluated in section 5.3.

Flash Light LED

Most smartphones are equipped with a flashlight LED, but because of the wide range of smartphones available, the used flash lights differ greatly. The wavelength and intensity of the emitted light play an important role in capturing the PPG signal as was shown in chapter 2. Unfortunately, almost all Android devices are hardware-limited with respect to the flash light usage and as such its influences cannot be controlled and have to be taken into consideration during development.

ISO

ISO stands for ‘International Organization of Standardization’ and expresses the digital camera’s *sensitivity to light*. ISO numbers generally range from [100-6400] (unit-less) and represent low to high sensitivity to light. This parameter allows video recording in darker areas by using a higher ISO value, but comes with a trade-off that the video would contain more grain and digital noise, degrading the quality of the video or image.

Shutter Speed

The shutter speed describes the amount of *time that the shutter is open*, during which the photosites are exposed to light, expressed as a fraction of seconds, e.g. 1/200 s. The longer the shutter is open, the more light is exposed to the photosites and the brighter the image will be.

Aperture

The Aperture is *the size of the opening* of the camera and is used to create depth of field. The smaller the hole, the less light is exposed to the photosites and vice versa for larger holes. However, most smartphone cameras have a fixed aperture setting that cannot be changed and thus needs to be considered a constant. The specific value is model dependent.

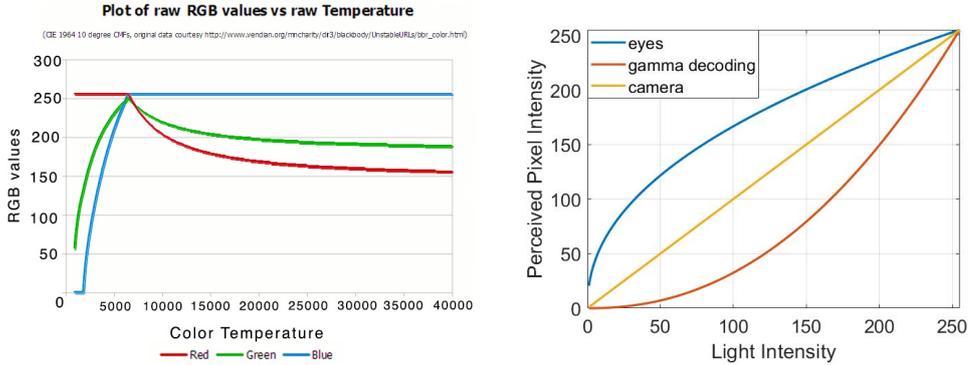
5.2.3 Image-Processing

This section covers the different settings affecting pixel intensity values *after frame generation*. In comparison with the previous settings, the influence of these on the PPG signal is harder to grasp.

White Balance

In order to understand white balance, an introduction to ‘color temperature’ is required. Color temperature, measured in Kelvins [K], is the warmth or coolness of white light as perceived by the camera. Different light sources have different color temperatures also known as color casts. Although named ‘white light’, the distribution of the visible colors in the light spectrum varies with color temperature [18]. An analogy can be made with a gas-flame. At lower temperatures, the flame has a red or yellowish color, but as the temperature increases the flame turns more white or even blueish. In photography, low color temperatures (3000K) produce images primarily dominated by red colors whereas high color temperatures are dominated by blue colors (9000K). The human eyes automatically adjust for the differences in temperatures such that e.g. white color will always be perceived as white, whereas cameras cannot. The mathematical relationship between color temperature and the specific RGB values has been studied by T. Helland [15]. Figure 5.3(a) shows this relation. The white color point occurs at 6500K-6600K, for which the RGB channels are all equal to a pixel intensity value of 255. This is the *maximum brightness*.

Since only the red color channel is relevant for constructing the PPG signal, it is important to keep the maximum range of [0-255] PI, which is [1000K-6600K] for color temperatures.



(a) Relationship between color temperature and maximum RGB values (Modified from [15]).

(b) In order to mimic the human eyes, a gamma correction is applied to the linear camera curve. However, linearity is desired for PPG so gamma decoding is applied.

Figure 5.3: Non-linear transformations on the pixels need to be accounted for.

Gamma Correction

An important but often forgotten aspect related to digital cameras is the gamma correction. This technique comes from the fact that human eyes and digital cameras perceive light differently. The intensity or brightness of light can be quantified with ‘luminance’. The link between luminance and pixel intensity is defined by ‘gamma’ (γ) and the differences in perception of light between the human eyes and smartphone camera is bridged with gamma correction. Figure 5.3(b) shows the linear response between the light present and the light observed for the camera. This is because if twice as many photons (thus light) hit the camera sensors, the resulting response is twice as large. In the case of human eyes, the relation shows a non-linear response, which stems from a biological advantage, that enables human eyes to work over a wider range of luminance, because more dark tones can be distinguished. In order to go from what our eyes see to a linear response, a gamma decoding or correction is required, illustrated by the orange curve in figure 5.3(b). Fortunately, the mapping between actual and perceived light can be set manually. The gamma curve can be approximated with a mathematical formula:

$$Y = 255 \times \left(\frac{X}{255} \right)^\gamma \quad [\text{PI}] \quad (5.1)$$

Here, X is the original scene luminance (PI), γ a constant and Y the output pixel intensity recorded (PI). A γ value of $\frac{1}{2.2}$ is commonly used in digital cameras, but for the PPG extraction a linear mapping is desired to preserve the PPG waveform. As such, γ is set to 1.0 to maintain the linear relation between observed light intensity and pixel intensity value.

Paper	Author	Date	Description
[49]	E.J. Wang	2016	RoI chosen as a square around the center of the frame with dimensions: $\frac{1}{2} \times$ width and $\frac{1}{2} \times$ height of the frame. Average pixel intensity to obtain the PPG signal.
[4]	N. Bui	2017	K-means clustering of all pixels in clusters of good/bad pixels with enough variation. Compare time variation with heart rate obtained with FFT. Average remaining pixel intensities.
[22]	E. Jonathan	2010	Arbitrary 10x10 pixel RoI selection based on empirical observations.
[42]	C.G. Scully	2012	Frames with resolution 720x480. RoI size of 50x50 pixels but location in the frame unspecified. Average pixel intensity to obtain the PPG signal.
[46]	F. Tabei	2019	Average all pixels in the entire frame.
[30]	J. Liu	2019	Average of all pixels in the entire frame, for multiple frames, to obtain PPG information about the cardiac cycles. Determine the frames corresponding to the min-max PPG amplitude values. In the min-max frames, determine the pixels that satisfy the ‘distance condition’ to obtain a mask. Apply the mask to all frames of cycle to improve the PPG signal.
[21]	W.J. Jiang	2014	Average all pixels in the entire frame.
[50]	G. Zhang	2019	Pixel selection on a per frame basis with empirically determined thresholds. Sum all pixel intensities to obtain the PPG signal.
[28]	F. Lamonaca	2012	Adaptive RoI circle

Table 5.1: **State of the art RoI selection methods.**

5.2.4 Region of Interest Selection

The last parameter to be discussed is not a camera settings, but related to the PPG signal extraction process. Region of Interest (RoI) is defined as the pixel region from a frame used to construct the PPG signal. Regular PPG sensors usually consist of a single photo-diode, whereas the camera sensor has a large array of photosites. As a result, instead of one light intensity value as for the traditional case, multiple pixels are to be considered. Not all pixels within a frame carry valuable information. Some are distorted by noise, others by motion of the finger on the camera lens. Some studies do not take RoI selection into account and use all pixels in the entire frame to construct the PPG signal. Other studies do consider RoI selection and offer solutions that vary in complexity and computational intensity. Table 5.1 shows an overview of the state of the art solutions.

As can be seen, a couple of studies ([46],[21]) average all pixels in the entire frame. Before discussing the other methods, it is important to explain why averaging the entire frame is detrimental for proper PPG signal extraction. Recall that a camera tries to capture the changes in blood volume over time via changes in light intensity, which is subsequently mapped to pixel intensity values in frames.

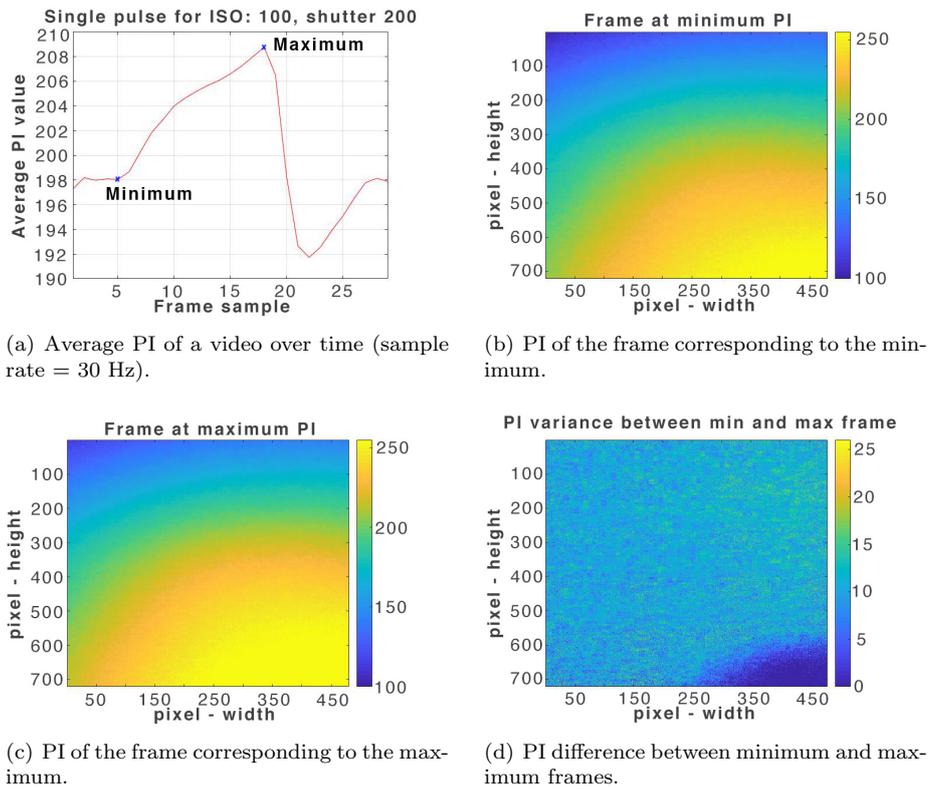


Figure 5.4: Pixel intensity change over time in a single PPG cycle.

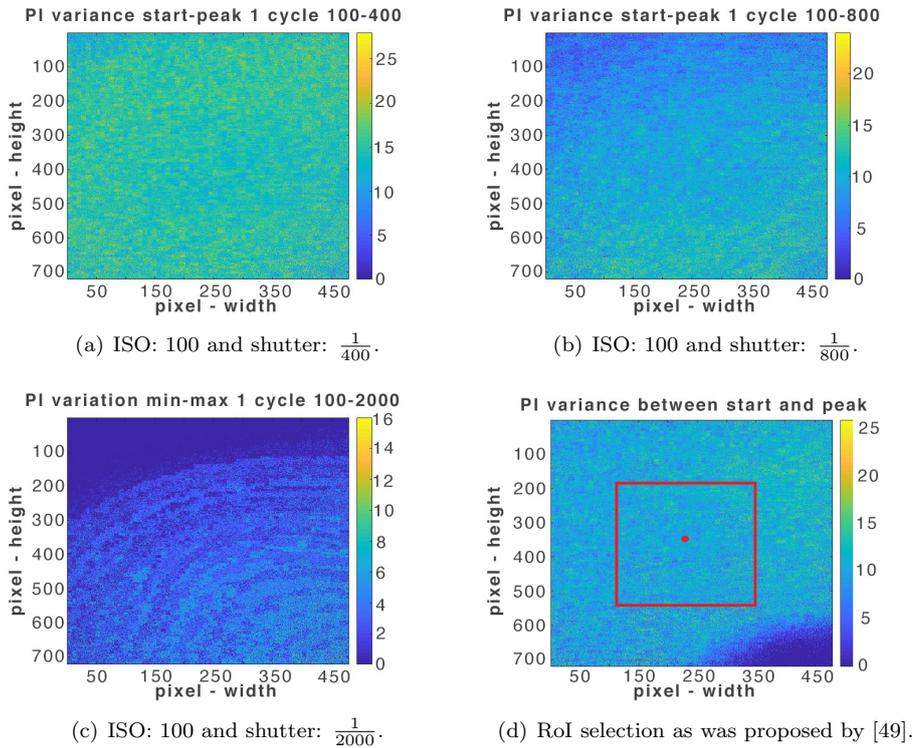


Figure 5.5: Influence of shutter speed on PI variations over time

Figures 5.4(b) and 5.4(c) show the frames at the time instances when the red color intensity of the averaged PPG signal is locally minimal and maximal in figure 5.4(a), corresponding to a diastolic peak and start valley of cycle respectively¹. The flash light is located in the bottom-right with respect to the finger and the closer a pixel is to the flash, the higher the pixel intensity (PI). By calculating the differences in PI for each pixel in the min. and max. frame, as shown with figure 5.4(d), it can be seen that the pixels near the flash light are saturated with a maximum value of 255. The light intensity as perceived in reality is much stronger, but cannot be quantified due to the arbitrary PI range. Fortunately, this can be solved by changing the camera settings, such as reducing the sensitivity to light or reducing the exposure time. Figure 5.5(c) shows the opposite case, where the pixels furthest from the flash are not reached by light. No saturation occurs in figure 5.5(a) and only slightly in figure 5.5(b).

The solution proposed by [49] is illustrated with figure 5.5(d). By selecting only the center part of the frame, the previous problems are eliminated, no matter the orientation of the flash with respect to the camera, making it suitable for different smartphones. Furthermore, the outer regions of the frame are more subject to ambient light if the finger is moved slightly. An additional advantage is that previously all PI values in the frame were averaged, whereas now only a fraction ($\frac{1}{4}$ th) of the frame is considered, speeding up the computational process significantly. *This RoI will be used for the system in this thesis.*

Methods proposed by [4], [30] and [50] do not select a continuous region, but individual pixels based on how they change over time, which gives rise to two different problems. The first problem is related to computational intensity. If a 720x480 pixel frame is considered, 345,600 pixels need to be evaluated *individually* and given the fact that a frame rate of 30fps (or more) is not uncommon, 10,368,000 pixels need to be evaluated *in the span of single second* which is not feasible for a real-time system. The second problem is that the authors assumes that the location where the blood volume changes occur is stationary, but in reality the resulting changes in light intensity are not necessarily mapped to the same pixel over time, because of e.g. finger movement. The influences of minor movements (a couple of pixels) can be prevented by considering larger continuous pixel regions, as was proposed by [49], instead of individual pixels.

Now that the RoI is defined, the pixel values in a frame need to be combined into a single value. Some methods simply average the pixel values in the frame, but this might lead to incorrect values if the RoI is not chosen carefully. A solution could be to use the median value of the pixel values but then the granularity of the signal becomes an integer which is very large. Fortunately, there exists a middle ground between the mean and median methods: α -trimmed mean filtering [9]. With this method, the outliers of the pixel intensity distribution are removed and the remainder is averaged. By carefully choosing α , saturated values can be excluded. An α -value of 0 corresponds to the mean and a value of 0.5 to the median. An α -value of 0.10 removes the values till the first 10th percentile and the values higher than the 90th percentile of the sorted pixel intensity values and will be used in the system.

¹Recall that light intensity and blood volume are inversely related (Section 2.1)

5.2.5 Summary

Due to the large amount of possible parameters, evaluating each one without even considering the possible inter-relations is a tedious and unnecessary task. Several parameters can be fixed or their ranges of values can be limited, based on what the system desires. Starting with the frame rate, 30 fps is the maximum frame rate the Motorola Moto G7 Plus offers. Secondly, large resolutions are unfeasible due to real-time constraints with all the processing overhead. Thirdly, any color temperature in the range of 1000K-6600K would suffice, but information from the blue and green channel can be valuable for e.g. finger detection which will be discussed in Section 6.1. As such, the color temperature for white balance is fixed to 6600K, since the red channel needs to take benefit of the entire pixel range [0-255]. Since aperture is fixed, only ISO and shutter speed are to be considered for the evaluation regarding light exposure. The range of values ISO and shutter speed can take is reduced substantially after some preliminary experiments. High ISO settings (ISO>800) would result in over-saturated images. What this means is that the camera sensor is too light-sensitive and would remain at its maximum value of 255 over time, irrespective of blood volume change. Shutter speed on the other hand, would result in images that are mostly black (near 0 pixel intensity) for very fast speeds ($< 50\mu s$) and as such no blood volume change could be measured. For slow speeds ($> 1ms$), the camera sensor would be over-exposed and the pixel values would be saturated with brightness. Summarised in table format:

Camera Setting	Value
Frame Rate	30 fps
White Balance	6600K
Gamma (γ)	1.0
Aperture	f/1.7
Resolution	320x240, 640x480, 800x600, 1024x768
ISO	100, 200, 400, 800
Shutter Speed	1/200, 1/400, 1/800, 1/2000 [s]

Table 5.2: Camera settings for evaluation with the used smartphone. The settings colored in **red** are variable and will be studied in section 5.3, whereas the others are *fixed*.

5.3 Evaluation

5.3.1 Frame Resolution

Recall the difference between sensor resolution and frame resolution as was discussed in Section 5.2. The sensor resolution is fixed, and expresses the number of photosites in the camera sensor. The frame resolution indicates how many photosites are combined to produce a single pixel value. The construction of the PPG signal is similar: multiple pixel values are combined (averaged) in the RoI to produce single PPG value. Decreasing or increasing the number of pixels would only affect the granularity of the PPG signal. What is meant with granularity is the size of the change in PPG value due to a change in one pixel with magnitude 1. As an example, the theoretical granularity of the PPG signal can be calculated for a frame resolution of 320x240 pixels. The size of RoI was defined as a square in the center of the frame and amounts to $\frac{1}{4}th$ of the original area. The pixels in the RoI are then averaged via α -trimming and with an $\alpha = 0.10$ only 80% of the pixels in this frame are retained. The granularity is then calculated as:

$$Granularity = 80\% \times \frac{4}{Width * Height} = \frac{1}{15360} = 0.000065 \quad (5.2)$$

The interpretation of this value is that if a single pixel increases in intensity with magnitude 1, this results in a change of 0.000065 in amplitude value of the averaged PPG value, which should not be the limiting factor for the PPG signal quality. The decision on which frame resolution should be used depends on two factors:

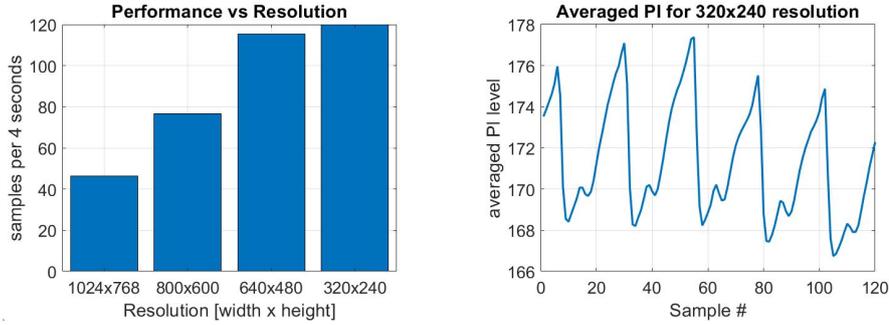
- Computational weight
- PPG signal quality

The first metric can be measured by monitoring the time it takes to process a frame. Ideally, this time should be low enough such that the desired frame rate of 30fps can be achieved. There are several ways to measure the processing time but one needs to be careful by not introducing unnecessary overhead that could affect the measurements. The method used performs measurements of 60 seconds with a timer that triggers each 4 seconds reporting the number of frames processed and the corresponding PPG values. Ideally for a frame rate of 30fps this will be 120 frames for each segment.

The second metric is more difficult to quantitatively measure. It is important that key PPG pulse properties remain identifiable, such as the dicrotic notch or diastolic peaks and that the granularity is high enough. The evaluation of this metric will be discussed in chapter 9.

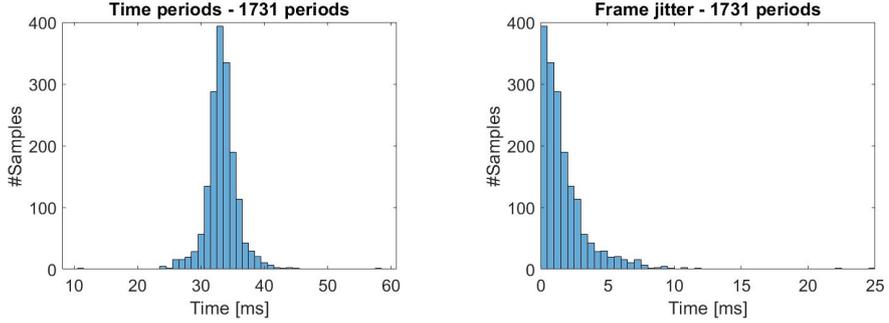
Figure 5.6(a) shows the influence of large images on the system performance. The larger resolutions significantly affect the actual sample rate, which is not alleviated until a resolution of 320x240 is used. It can be concluded that frame resolutions above 320x240 are detrimental for the system application.

Figure 5.6(b) shows a PPG signal obtained with resolution of 320x240. It can be observed that PPG features such as the dicrotic notch and diastolic peak are clearly distinguishable.



(a) Average number of samples processed in 4 second segments. (b) A 4 second PPG segment for 320x240 frame resolution.

Figure 5.6: Impact of resolution on application performance and PPG



(a) Histogram of the calculated time periods. (b) Histogram of the frame jitter, which is the absolute difference between period and mean.

Figure 5.7: Evaluation of the time periods and frame jitter in the smartphone application.

5.3.2 Frame Jitter

Now that the frame resolution is determined, the second frame-related aspect can be evaluated which is jitter. The periodic task of processing the frame consists of the following sub-tasks:

1. Acquire the image from the camera
2. Obtain a single PPG value by evaluating the pixels in the image
3. Perform fingertip detection (which will be discussed in section 6.1)

Since the frame-rate is 30fps, it means that a delay of $\approx 33.33\text{ms}$ between consecutive frames is desired. Any deviation from this expected period is jitter. To evaluate this, a 60 second long measurement at 30fps is performed to obtain 'periods'. A period is defined as the time passed between 2 timestamped tasks. For this experiment, 1731 periods with a resolution of 1 ms were collected. Figure 5.7(a) shows a histogram distribution of the periods. The mean period is equal to 33.33ms, as was expected. To get a better grasp of the jitter variance, the mean is subtracted from the time period, and the result is shown by figure 5.7(b). The standard deviation of the time periods, $\sigma(T)$, is equal to 2.83ms which for a bell-shaped curve is the 68th percentile. The 95th percentile = 5.67ms which also happens to be $\approx 2\sigma(T)$. Another important observation is the severity of the outliers. Albeit their occurrences are unlikely, 25ms of jitter is considerable when compared to the mean period. However, the overhead introduced by tracking these periods can also play a role in the variability of the periods. Given the jitter variance, the frames will not be time-stamped.

5.3.3 Light Exposure

The range of ISO and shutter speed values that will be evaluated are displayed in table 5.3.

Camera Setting	Value Range
ISO	100, 200, 400, 800
Shutter Speed	1/200, 1/400, 1/800, 1/2000 [s]

Table 5.3: ISO and shutter speed values that will be evaluated.

With 2 variables and 4 possible values, 16 experiments need to be done to explore all combinations. A single experiment consists of 120 seconds of PPG measurements on the right index finger. The metrics used to compare the ISO and shutter speed combinations are the PPG amplitude and pixel offset. The measurements are videos recorded with the different camera settings and evaluated *offline* in MATLAB[®] ver. R2018b. The results are displayed in tables 5.4 and 5.5.

	Exposure Time			
ISO	1/200	1/400	1/800	1/2000
100	170.7348	98.5701	47.5453	11.4754
200	222.6345	164.9828	112.3258	39.3939
400	247.9105	228.9164	179.2379	81.7558
800	254.9863	249.6753	212.4715	150.9384

Table 5.4: Pixel offset along which the PPG signal propagates over time. The intensity values range from 0 (black) to 255 (white).

Table 5.4 shows the pixel offset at which the PPG signal fluctuates. Although the pixel offset is not constant over time and also depends on finger related factors, it does give a good approximation whether the PPG signal is created with saturated pixels or not. Large values close to the maximum pixel intensity of 255 indicated saturation due to extreme exposure to light. On the other hand, low values close to the minimum pixel intensity value of 0 suffer from under-exposure, because very little light is able to reach the camera. This idea is confirmed by the findings of table 5.5, where the saturated PPG signals have very small PPG amplitudes.

	Exposure Time			
ISO	1/200	1/400	1/800	1/2000
100	12.1565	13.0804	6.7899	3.3339
200	5.8817	6.9425	12.8817	4.4928
400	2.9864	5.4884	11.6682	10.324
800	0.0261	3.6198	8.9487	9.0787

Table 5.5: The mean PPG amplitude expressed in pixel intensity (PI) for the different camera settings, where lower (red) values indicate worse performance.

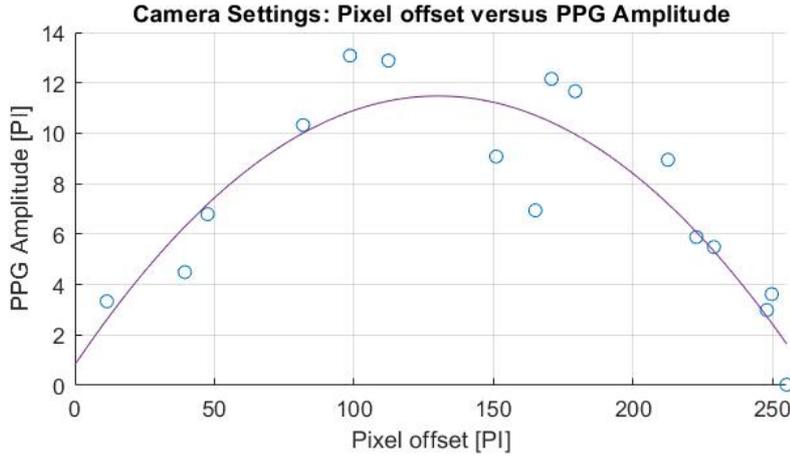


Figure 5.8: **Pixel offset plotted versus mean PPG amplitude. The curve is a cubic approximation of the data-points.**

Combining the values from tables 5.4 and 5.5 results in the curve shown in figure 5.8. It can be seen that the PPG amplitude is maximized for values located in the middle of the pixel intensity range [80-179], centered around ≈ 123 . However, the pixel offset can easily shift by motion artifacts and a single optimal configuration is hard to obtain. For instance, ambient light can increase the pixel offset of the signal. The skin color, structure and underlying tissue in combination with finger thickness also play an important role. As such, any setting that does not result in pixel-intensity saturation as a result of over- or under-exposure can be used.

5.3.4 PPG Signal Generation

Recall from section 5.2.4 that only a part of the frame (the RoI) is used and the PPG signal is generated via α -filtering. SoA studies that operate in real-time, average the pixel values in the entire frame and do not consider the camera settings. To compare the two methods, first, an experiment is conducted where a 60 second video is recorded using a standard smartphone camera application with automatic camera control. This video is then processed using the SoA method to generate a PPG signal. For the second experiment, the camera is manually controlled with an ISO = 100, shutter speed = $\frac{1}{400}$, and fixed parameters from table 5.2, whereafter the RoI and α -filter are applied. Table 5.6 shows that for the automatic parameters, the pixels are saturated, demonstrated by the high pixel offset and the small PPG amplitude.

Camera Control	PPG Amplitude	Pixel Offset
Manual + Average Frame	7.2	134.0
Manual + RoI & α -filter	7.2	144.9
Automatic	2.9	250.8

Table 5.6: **Automatic camera control leads to saturated pixels, which negatively affect the PPG amplitude and waveform.**

Furthermore, there is no difference in PPG amplitude between averaging the frame or using an RoI with the α -filter. However, the latter is faster and more efficient as it only needs to evaluate $\frac{1}{4}th$ of the frame.

5.3.5 Conclusions

Summarizing this chapter, there are some general conclusions that can be drawn. Automatic camera control is dangerous as it leads to pixel saturation and incorrect PPG signal generation, thus manually choosing the camera parameters is desired. Table 5.7 shows the settings that should hold for any device. Because the aperture is different per device, the *ISO and shutter speed should be chosen in such a way that the saturation of the pixel intensity values is avoided*. A general approach for determining the right ISO and shutter speed settings would be as follows. First, set the ISO to a low value, e.g. 100. Then, increase the shutter speed until the image is no longer saturated by light and the pixel offset lies within the [80-179] PI range.

Camera Setting	Value
Frame Rate	≥ 30 fps
White Balance	6600K
Gamma (γ)	1.0
Resolution	320x240

Table 5.7: **Camera settings that apply to all smartphones.**

Furthermore, ROI selection of the center of a frame and α -trimming is introduced to decrease the influences of pixel saturation. Analysis of the frame rate jitter showed that 95% of the occurrences has a jitter of ≤ 5.67 ms with only a few large outliers. Because this jitter is relatively low and several other factors play a role in the frame capture, tracking the time-stamp of a frame is deemed unnecessary.

Chapter 6

Raw PPG Signal Analysis

The previous chapter covered the first steps to extracting the PPG signal from the smartphone camera. However, the quality of this reconstructed PPG signal is subject to several sources of disturbances and noise.

In this chapter, a method is introduced in Section 6.1 that automatically detects a fingertip to start measurements. Lastly, the finger pressure on the camera lens impact the PPG signal and needs to be tackled, which will be covered in Section 6.2.

6.1 Fingertip Detection

6.1.1 Methods

With the intention of using the system as a smartphone app, an automatic finger detection algorithm would alleviate the need for users to start/stop the testing and measuring. The placement and position of the fingertip play an important role. Two similar methods to handle this have been implemented in literature. The first method by Liu et. al. [30] evaluates the dominance of the red channel over the green and blue channel for each pixel. Equation 6.1 shows the calculation of the dominance of the red channel for frame t :

$$T \leq Pr(x, y) = \frac{r_{(x,y)}(t)}{r_{(x,y)}(t) + g_{(x,y)}(t) + b_{(x,y)}(t)} \quad (6.1)$$

where $T(=0.85)$ is the threshold and $r(x, y)$, $g(x, y)$, $b(x, y)$ denote the light intensity in the red, green and blue channel at pixel location (x, y) , respectively. If Pr is higher than the threshold T , the pixel is said to be dominated by the red channel. If 95% of the pixels are dominated by the red channel, a fingertip is detected. The dominance depends on the fact that the RGB values have equal ranges (0-255) which is dictated by the white balance as was demonstrated in Section 5.2. Liu et. al. only mention that white balance is locked, but not at what temperature and thus RGB ranges are unknown.

The second method was proposed by Zhang et. al. [50]. Here they empirically derived thresholds for the averaged RGB channels and also the standard

deviation of the red channel (σR) within a frame. Table 6.1 shows the used settings where the values represent pixel intensity values in a range of [0-255].

Threshold Condition
$Ave(R) \geq 240$
$\sigma R \leq 20$
$Ave(G) < 1$
$Ave(B) < 75$

Table 6.1: **Conditions for evaluating Video Usability [50].**

Observe the requirements for the red channel R . The average intensity $Ave(R)$ needs to be 240 but the standard deviation σR is required to be smaller than 20. These values are alarming because the maximum pixel intensity is 255 which is only 15 more than the average-threshold, while σR is allowed to exceed this. This means that for these thresholds, several pixel values will be saturated resulting in incorrect PPG pulses. Furthermore, the values for these thresholds are extremely dependent on the used camera settings as was described in Section 5.2, which are not considered.

6.1.2 Evaluation

Before the method proposed by Liu et. al. could be implemented, a modification was required. The threshold for dominance, T , had to be lowered from 0.85 to 0.70 because no fingertip would be detected at all. During evaluation, the fingertip was sometimes placed improperly yet the system ‘detects’ a fingertip. Examples of this are shown in figures 6.1(a) and 6.1(b) on the next page, where the camera is not entirely covered by the fingertip, yet the black areas are dominated by the red channel. There is a reason why the finger detection algorithm fails. Consider a black pixel with a RGB value of $R = 1$, $G = 0$ and $B = 0$, equation 6.1 would produce a domination value of 1.0 (100%), *yet no finger is present*. Since the shutter speed is fast and ISO is low, images are very dark if no finger is applied to the camera-lens, making the previously described scenario likely.

6.1.3 Solution

A new algorithm is proposed that can properly detect the finger, illustrated with figure 6.1(c). An 8x8 raster is applied on top of the frame, where each cell is of equal size. The average red pixel intensity is calculated for each cell and compared to a threshold, $minR = 30$. If one of the cells fails to satisfy the threshold, the finger is not placed correctly and further processing is stopped while informing the user to adjust their finger. In figure 6.1(c) the top left cells would fail the threshold, indicating incorrect finger placement. The threshold value is based on figure 5.8 from chapter 5, because values below 30 would result in signals with too small of a PPG amplitude and this way the threshold can accommodate multiple camera setting configurations.

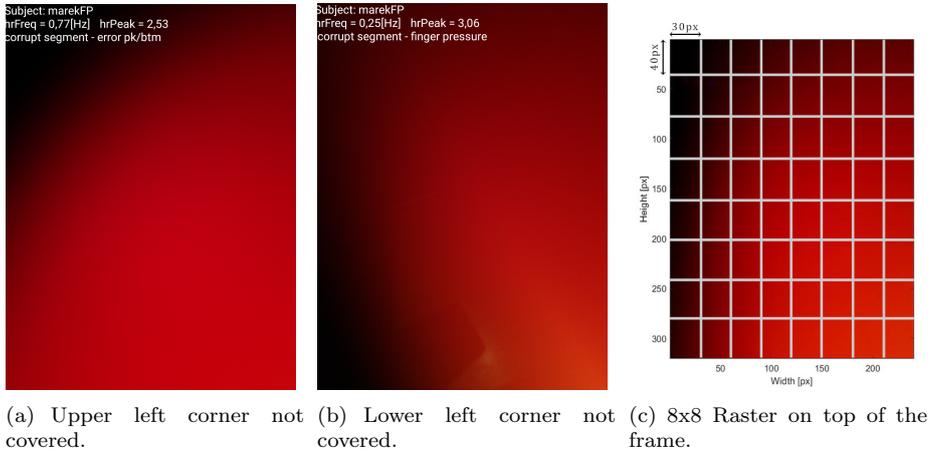


Figure 6.1: a), b) The original algorithm does not account for fingertip placement as it ‘detects’ a fingertip in the situations depicted above. c) Proposed method to evaluate cells in the raster for correct placement detection.

6.2 Finger Pressure Detection

Although some disturbances on the PPG signal are exclusive to smartphone sensing, others have been around since the birth of PPG sensing in 1937 [16]. A.B. Hertzman, who coined the term PPG, stated in his early works: “*The most important source of error and the one most difficult to control is movement of the skin with respect to the plethysmograph*” [17]. Furthermore, Hertzman also found that the contact pressure between the skin and plethysmograph is crucial for PPG. Most of these artifacts can be limited as they are very noticeable but the finger pressure is more subtle as it is difficult for the user to control. When the finger is pressed against the smartphone camera lens, a force is applied on the blood vessels under the skin surface. As a result, the blood occupying these regions is pushed out which ‘flattens’ the skin slightly at the fingertip as could be seen in figure 2.4. While the heart keeps pumping, blood cannot easily flow into these capillaries due to the obstructions.

6.2.1 Methods

The effect of finger pressure is application dependent. Trivial physiological features such as heart rate estimations [37] do not tackle finger pressure, yet obtain good results. As more research was put in PPG sensing and the smartphone was introduced, medical applications extended to more complex medical features and security that require a higher PPG signal quality than before.

A recent study by Chandrasekhar et. al. [8] underlines the fact that neglecting finger pressure in scenarios without finger-cuff severely affects features used for determining medical conditions, which is also supported by [14] regarding arterial stiffness. Figure 6.2 highlights the influence of contact pressure on the PPG pulse morphology. As the finger pressure increases, the dicrotic notch and diastolic peak decrease with respect to the systolic peak.

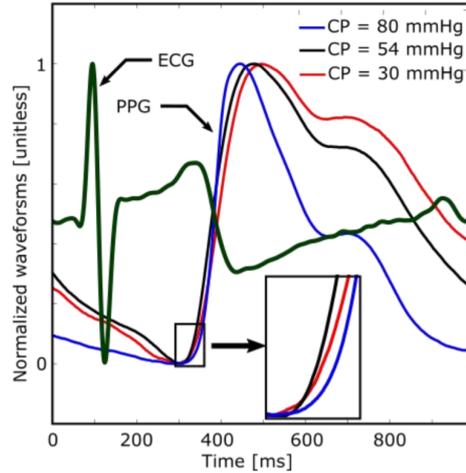


Figure 6.2: **Impact of contact pressure (CP) on PPG pulse waveform [8]. The dicotic notch and diastolic peak vary in amplitude depending on the CP.**

Medically-certified PPG sensors such as pulse oximeters use finger-cuffs to regulate the pressure. There are multiple studies that build devices which can be attached to a smartphone for PPG signal extraction, often including a finger-cuff [49] to regulate finger pressure, or to measure the contact pressure separately, as demonstrated by [45] and [7]. This approach is similar to what was done for movement artifacts. The downside of these methods, is the requirement of additional hardware. More recently, there have been studies for PPG measurements methods without finger-cuff. Tabei et. al. [46] use the PPG amplitude and compare it to an average of the entire frame, but this only works in an *offline* setting since it requires knowledge of multiple minutes of data. Bui et. al. [4] apply spectral analysis in real-time to detect finger pressure. Table 6.2 summarizes the finger pressure detection SoA works, highlighting the most relevant method in green.

Author	Date	Description
Wang [49]	2016	(Hardware) HemaApp: finger-cuff
Sim [45]	2018	(Hardware) Thermo-pneumatic regulator to measure and regulate contact pressure at the wrist site.
Chandrasekhar [7]	2018	(Hardware) Measure with separate force transducer and obtain optimum via user feedback.
Tabei [46]	2018	(Software) Compares the Average Amplitude Value of the <i>entire</i> signal with segments of the signal in an <i>offline</i> setting.
Bui [4]	2017	(Software) PhO2: Fast Fourier Transform of PPG signal and evaluated amplitude of HR component in an <i>online</i> setting.

Table 6.2: **State of the art techniques to combat finger pressure using either hardware or software solutions.**



Figure 6.3: **Finger above smart-
phone [4].**



Figure 6.4: **Finger below smart-
phone.**

6.2.2 Method Description

Bui et. al. [4] are the only authors that tackles finger pressure using only information from the PPG signal in a *real-time* smartphone design. In their study, the authors distinguished between three types of pressure: weak pressure, appropriate pressure and strong pressure. Weak pressure occurs when the fingertip barely makes contact with the camera lens and only a faint pulse can be detected, whereas strong pressure completely occludes the capillaries and arteries. The definition of appropriate pressure is rather obscure, since it is defined as ‘*the pulse is clearly observed*’. The authors state that in clean PPG signals with appropriate pressure, a large fraction of the signal power is concentrated in the heart rate frequency. The less noise and distortions in a short segment of 4 seconds, the stronger the amplitude corresponding to the heart rate component. The Fast Fourier Transform (FFT) is applied to a PPG segment and the largest peak correspond to the heart rate. The amplitude is then evaluated to determine the finger pressure.

6.2.3 Limitation

Recall that under normal conditions, the heart rate ranges from [30-240bpm] which correspond to a frequency range of [0.5-4Hz]. The spectrum is assumed to be clean in this range, but noise and distortions within this range cannot be filtered. Furthermore, the frequency resolution is very low (0.25Hz) due to the short segment time (4 seconds). Another important factor for finger pressure is how the finger is applied to the smartphone camera-lens. To elaborate on this matter, two possible scenarios for holding the smartphone are shown in figures 6.3 and 6.4. In the first scenario (figure 6.3), the user is entirely responsible for the applied finger pressure, whereas in the second scenario (figure 6.4), the smartphone also applies a pressure to the finger due to gravitational forces. This force on the finger depends on the angle at which the smartphone is held, further complicating the problem. The first scenario is desired in a controlled research environment because the only pressure applied is from the user, however it is unrealistic for commercial usage. This is because the PPG signal extraction is used in combination with an app that provides visual feedback to the user, which is only possible if the screen faces the user. Although some smartphones are equipped with a back and front facing camera, the flash is normally located on the back and as such, the second scenario, as depicted in figure 6.4, is more realistic.

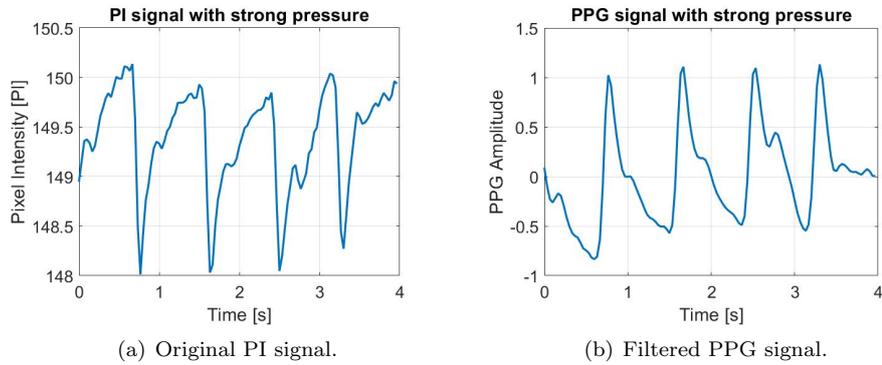


Figure 6.5: Strong finger pressure results in small PI amplitude of 2 and changes the waveform morphology.

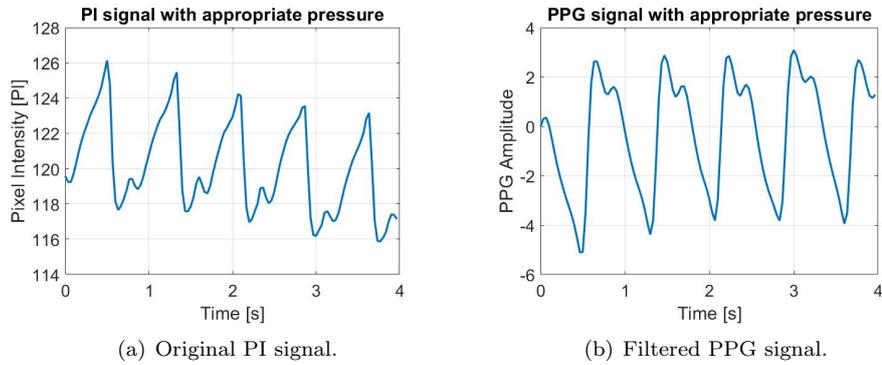


Figure 6.6: Appropriate finger pressure with large PI amplitude of 8.

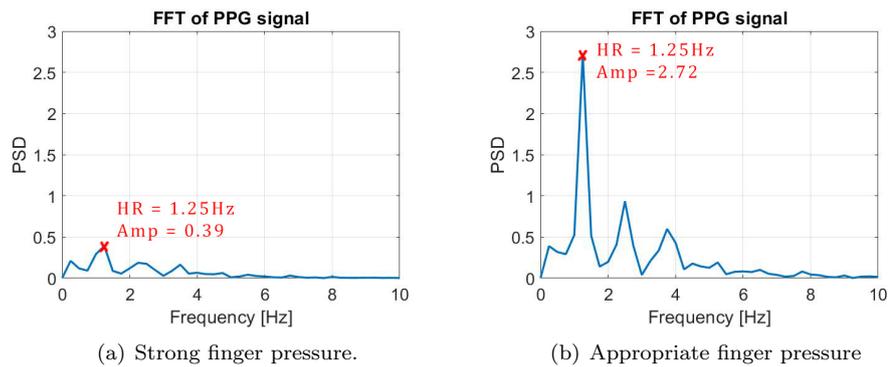


Figure 6.7: Power Spectral Density Distribution after the Fast Fourier Transform (FFT). Strong finger pressure results in small amplitudes for the heart rate frequency.

6.2.4 Results

Before the FFT can be applied, the signal needs to be filtered and the pixel offset needs to be accounted for. If the latter is not done, an extremely large DC-component is included in the spectral-analysis, because the signal is not centered around the horizontal axis. The filtering is discussed in detail in Section 7.2.1. Figures 6.5 and 6.6 show that different finger pressures result in PPG signals with different PI amplitude and waveform in the original and filtered segments. There is a large amplitude drop going from the systolic peak to the dicrotic notch and peak for strong finger pressure (figure 6.5). This amplitude drop is much smaller for appropriate finger pressure (figure 6.6) and the overall amplitude is much larger. These observations are in correspondence with the measurements done by [8] in figure 6.2. Here, the dicrotic notch and diastolic peak become more distant from the systolic peak as the contact pressure (CP) increases. Figure 6.7 shows the frequency distribution for both strong and appropriate pressure. The heart rate frequency corresponds to the largest peak. It can be seen that for appropriate pressure, a much larger peak amplitude ($\times 7$) is detected than for strong finger pressure. Since strong pressure results in undesired distortions in the PPG waveform, it is important to detect it. Because the peak amplitudes for the heart rate component differ so greatly for the two cases, they can be distinguished by using a threshold. However, the threshold should hold for multiple individuals.

The peak amplitude should be larger than 0.5 and the heart rate frequency should lie within the heart rate range of $0.5 \leq HR \leq 4$ [Hz]. This threshold was empirically determined. If these conditions are satisfied, the finger pressure is deemed correct. Otherwise, the user will be informed to adjust the pressure.

Another important observation during measuring was the impact of finger temperature. This was already discovered by Hertzman [17] in 1938, where under cold conditions, the blood circulation rapidly diminished in the finger and thus a low-amplitude PPG signal was observed. Although the heat generated by the flash light can increase the finger temperature, it can also lead to discomfort or even damage the skin for measurements in the order of minutes.

6.3 Conclusions

In this chapter, a new method was introduced to detect the presence of the fingertip by rasterizing frames and evaluating the resulting cells. Moreover, it was shown that finger pressure is a subtle artifact that has a large impact on the PPG signal amplitude and waveform. It is important to track the finger pressure and provide feedback to the user to obtain an appropriate amount. Despite the impact, finger pressure was neglected during the early periods of PPG sensing with the smartphone (2010). Only recently it was addressed and a method from the literature is implemented to track finger pressure (Bui et. al. [4]). This method evaluates the frequency domain characteristics of the PPG signal and can differentiate between strong and appropriate finger pressure. In the former case, the user is informed to adjust the pressure. Finger temperature also plays an important role, as the blood circulation in the finger is highly susceptible to heat. Before measuring, the finger should be warm to obtain good PPG signals. All in all, the first stage of qualifying the PPG signal is done and the signal can now be segmented, covered in the next chapter.

Chapter 7

Pulse Segmentation

The goal of this thesis is to develop a system that provides PPG signals of high quality. In order to obtain PPG pulses, the original PPG signal needs to be segmented. This is a complex process and several studies have been done on parts of the segmentation progress.

Section 7.1 covers the methods developed in literature that aid in this segmentation process and a set of metrics is used to evaluate them. Hereafter, Section 7.2 introduces a pulse segmentation algorithm proposed by this thesis, which uses a few elements of a state of the art (SoA) method. The chapter concludes with a summary in Section 7.3.

7.1 State of the Art

Although the details vary, the segmentation fundamentally depends on the detection of the systolic peaks. There exist several studies that explore this *systolic peak detection*, but not many studies actually segment the PPG signal. In order to segment the actual pulses, the start and end point need to be identified, which will be referred to as *valley detection*. However, it is still worth to consider these systolic peak detection studies, since elements of the different methods can be combined. Since PPG signals are not obtained from perfect scenarios but the real world, *motion and noise artifacts* (MNAs) [46] are present. MNAs are disturbances on the PPG signal due to noise or user movement. Additionally, the sensor type plays an important role. PPG signals from dedicated medical equipment provide higher quality than from a smartphone camera. Furthermore, the method needs to be able to segment pulses for all possible waveform types, since each individual has a unique one.

In order to develop and test a method, PPG data from several subjects is needed, which can be hard to collect due to time restrictions and requirement of authorization. Fortunately, many public *PPG databases* are available, but there are drawbacks because PPG signals are *too clean*:

- PPG signals are collected with high quality medical equipment.
- PPG signals have been processed with unknown methods.
- PPG signals are collected under ideal conditions with as few as possible MNAs, from e.g. the intensive care and from subjects under anesthesia.

These databases can be used to develop algorithms, but only as complement for datasets with both clean and corrupt signals.

Furthermore, data used for training the method should be different than the data used for testing. Otherwise the parameters would be too tailored for a specific dataset. Another important factor is whether the method is intended for real-time pulse segmentation. For instance, a segmentation method in an *offline* environment has complete knowledge of e.g. 120 seconds of PPG data of a subject, whereas in a *real-time* environment only a fraction of this data is available. Another important metric that is related to the real-time constraint is the *complexity*. A method could obtain near perfect results with e.g. a neural network (NN), but the training of this NN takes egregious amounts of time. Furthermore, the *verification* of the method is important. For credibility, the method should be compared with SoA methods in the paper.

Implementing and evaluating every possible method is a tedious and time consuming task, but fortunately several methods can be eliminated based on the information provided in their papers. Table 7.1 shows the methods and how they compare with each other based on the previously described metrics. *All* methods were developed with PPG signals from medical equipment or online databases and not with PPG signals from a smartphone camera. The methods are in chronological order and it can be seen that the early methods did not account for MNAs as they obtained PPG signals from MNA-free environments. From the methods in table 7.1, only two methods score positively on almost all metrics, where MNA elimination played an important role. These methods are:

- [10] Event-Related Moving Average by Elgendi et. al. (2013)
- [12] Moving Average + thresholds by Fischer et. al. (2017)

The first method only detects systolic peaks and has a couple other limitations that can be overcome, which will be discussed in section 7.2. The second method does more and actually segments pulses and uses a large list with thresholds conditions. The problem with these threshold values is that they are completely tailored for the databases with which the algorithm was developed. Implementing this method would require re-calibrating all thresholds which are device dependent. This is in conflict with the goal of having a generalized application. However, certain fundamental aspects are used to determine the PPG signal quality, which will be discussed in section 7.2.4.

Author	Year	Description	MNA elimination	Real-time	Complexity	Dataset	Verification
[1] Aboy et. al.	2005	Peak detection: Multi Stage bpf and HR estimation	No, HR estimation not robust	Iterative stages	Medium	IC, fingerclip, anesthesia	Acceptance intervals of 8.0, 16.0, 24.0 and 48 ms
[11] Farooq et. al.	2010	Pulse segmentation: Zero-crossing of first derivative + thresholds	Fingerclip and cover to reduce ambient light	Yes	Low	Fingerclip	x
[27] Karlen et. al.	2011	Peak detection: Adaptive Frequency Estimators	Unable to detect large and rapid changes in HR	Yes	Medium	Intensive Care, finger clip, anesthesia	Comparison with 2 old and outdated algorithms
[3] E. Billauer MATLAB®	2012	Peak detection: Local Min/Max + threshold between potential peak and surrounding values.	No.	Yes	Low	-	-
[26] Karlen et. al.	2012	Pulse segmentation: Gaussian Filters + HR filtering	Yes, but high recovery cost after artifact detection.	Yes	High	Intensive Care, finger clip, anesthesia	Long recalibration → good pulses missed
[25] Karlen et. al.	2012	Pulse segmentation: Incremental Line Segmentation (ILS)	Yes, but bad performance	Yes	Low	Intensive Care, finger clip, anesthesia	Artifacts are 50% classified incorrectly as pulse
[10] Elgendi et. al.	2013	Peak detection: Event-Related Moving Average	Yes, but uses static thresholds	Possible	High	Dataset with a lot of motion artifacts etc.	High stress situation: heat, physical exercises etc.
[19] Dae-geun et. al.	2014	Peak detection: Slope Sum Function (SSF)	Only peak qualification, initialization requires clean PPG	Yes	Low	Strict peak location deviation → 5 ms	Separation of training and testing sets
[12] Fischer et. al.	2017	Pulse segmentation: Moving average filter + abs min/max + thresholds.	Uses a checklist for artifact detection, but many arbitrary thresholds	Yes	High	No pre-processed datasets, diverse data	Good analysis of datasets used by others (that are bad)
[6] Campbell et. al.	2018	Peak detection: Wavelet decomposition	No	Yes	Low	Only tested on a <i>single</i> person	No comparison with other contemporary methods

Table 7.1: State of the art pulse segmentation and peak detection methods.

7.2 Proposed Method

This section introduces an algorithm for the extraction of cardiac pulses from the captured PPG signal. Inspiration was given by Elgendi et. al. [10], whose peak detection method for systolic peaks in PPG signals will be discussed in section 7.2.2. During development, PPG signals from a private data set containing 24 people with 4 minute smartphone recordings each was used. Additionally, public datasets with regular PPG sensor data were used [32],[33].

7.2.1 De-Noising

The goal of this stage is to reduce the noise artifacts in the PPG signal while preserving crucial information. The most conventional method is a band-pass filter. The existence of the dicrotic notch should be preserved and thus a bandwidth of [0.5Hz - 12.5Hz] is chosen [30]. A 2nd order Butterworth meets the requirements as it has an almost flat frequency response in the pass-band. The PPG signal is ‘spline-interpolated’ from 30Hz to 60Hz and mirrored in the horizontal axis to obtain the known PPG waveform (Discussed in section 2.2). From this stage on, the specific amplitude of the signal is irrelevant and only the waveform, pulse duration and proportions matter. A 4 second segment before and after filtering is shown in figures 7.1(a) and 7.1(b).

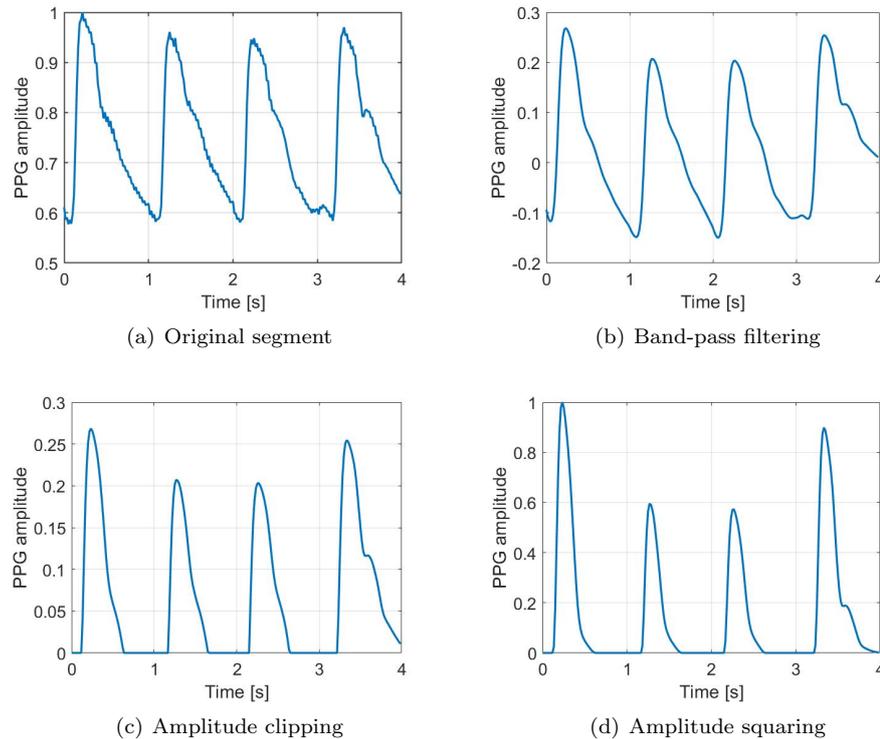


Figure 7.1: A 4 second PPG segment that undergoes different processing techniques.

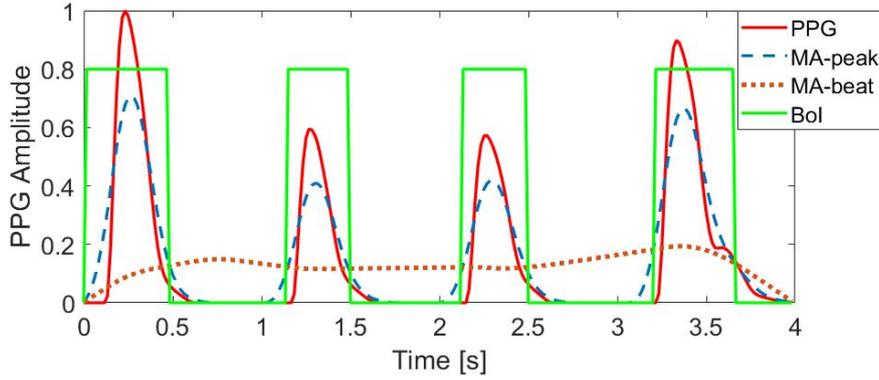


Figure 7.2: **A 4 second segment after BoI generation.**

7.2.2 Systolic Peak Detection

For peak detection, a copy of the filtered signal is used, whereas the original is left *unmodified*. First, the potential peak locations are amplified whereas the other parts of the PPG signal are attenuated. This is accomplished by first ‘clipping’ the filtered ppg signal amplitude below zero. Then, the clipped signal is squared to amplify the stronger components, shown in figures 7.1(c) and 7.1(d).

It should now be possible to detect the systolic peak area via a method called ‘Block of Interest’ (BoI) generation. Two moving average filters `MA_peak` and `MA_beat` are required with different window sizes, $W1$ and $W2$. As the names already indicate, `MA_peak` is used to amplify the systolic peak area whereas `MA_beat` represents the average of a cardiac pulse. The corresponding window sizes $W1$ and $W2$ are static in the method of [10] with values of $W1 = 111ms$ and $W2 = 667ms$, where the latter corresponds to a heart rate of 90bpm. A threshold is generated by adding an offset, α , to `MA_beat` to account for slight baseline drift in the PPG signal. The threshold value is given by equation 7.1:

$$TH = MA_beat[n] + \alpha \quad [\text{a.u.}] \quad (7.1)$$

A BoI is generated any time `MA_peak` *exceeds* the TH up until `MA_peak` falls below the threshold. This is repeated for the entire segment. Figure 7.2 shows how blocks are generated. Each BoI is tested whether it actually captures a systolic peak or not. This is done by comparing the time-length of a block with $W1$, the window size of the `MA_peak`. Since the systolic peaks have been identified in a heavily processed PPG signal, they need to be mapped back to the original filtered segment. The maximum value within in window of 50ms left and right from the detected peak position is labelled as the systolic peak.

Limitations

The previously described method has several limitations that need to be overcome, before it can be used in a smartphone applications. Firstly, the method was developed and tested in an entirely offline environment. This has already been addressed by only providing 4 seconds segments to the algorithm. Secondly, the window sizes $W1$ and $W2$ for the moving average algorithms are *static values*. These values were obtained by brute-forcing the optimal values from a large PPG data-set. Due to these static thresholds, the method fails for cases where the heart rate is much higher or lower than 90 bpm, since systolic peaks are either missed or over-detected because diastolic peaks are incorrectly identified as systolic peaks. Thirdly, the same data used for development was also used for testing. Lastly, only peak detection was performed and for proper pulse

segmentation the detection of valleys is required. In short, the following aspects need to be addressed:

1. Offline → Real-time environment
2. Static → Dynamic window sizes
3. One dataset → Different training and testing datasets
4. Systolic peak detection → Pulse segmentation

Improvements

A PPG smartphone application needs to be robust against any external factors. One of these factors is the heart rate of the user. As was mentioned in section 7.2.1, the heart rate of healthy people lies in the range of [25bpm - 240bpm], which corresponds to pulses of [2400ms - 250ms] respectively. Elgendi et. al. [10] only considers pulse widths of $W2 = 667ms$, thus an estimate of the pulse width is needed, prior to the construction of the moving average filters. One way to solve this is to analyze the segment in the frequency domain. Fortunately, this was already done for the finger pressure detection method discussed in section 6.2 and the heart rate obtained in that step can be used here. In the original method the window sizes were chosen as $W1 = 111ms$ and $W2 = 667ms$, which are assumed to be related. Dynamic window sizes can be derived from the estimated heart rate: $W1 = \frac{W2}{6}$ and $W2 = \frac{fs}{f_{hr}}$ where f_{hr} is the heart rate in [Hz] and fs the sample frequency in [Hz]. After the window sizes $W1$ and $W2$, have been set, the algorithm continues as described at the start of Section 7.2.2 till the peaks have been derived from the BoIs. Section 7.2.4 will go more in depth on the details of the heart rate estimation.

7.2.3 Valley Detection

While systolic peak detection is a relatively complex process, the valley detection of a cardiac pulse is simple. As was extensively discussed in Section 2.2, a cardiac pulse starts with a rapid rise in amplitude until the systolic peak is reached, whereafter it gradually declines. *This steep rise is common to all cardiac pulses.* The systolic peak is always preceded by the start of the pulse, which can easily be identified as the local minimum in a window *before* the systolic peak. This window, $W3$, is a fraction of the original cardiac pulse and can be derived from $W2$ as follows: $W3 = \frac{W2}{3}$. If the minimum value in this window is the furthest away from the systolic peak, it means that the PPG signal is still descending and the window size $W3$ is expanded to $W4 = 1.5 \cdot W3$. If no local minimum can be found, the peak is discarded. The peaks and valleys are displayed in figure 7.3(a) on the next page.

7.2.4 Peak & Valley Decision Logic

Due to MNAs or over-detection due to diastolic peaks, some of the peaks and valleys derived in the previous sections can be incorrectly identified. Figure 7.4 shows two diastolic peaks incorrectly being identified as a systolic peak. To combat this incorrect classification, some qualification indicators are put in place: only the peaks and valleys will be used for qualification.

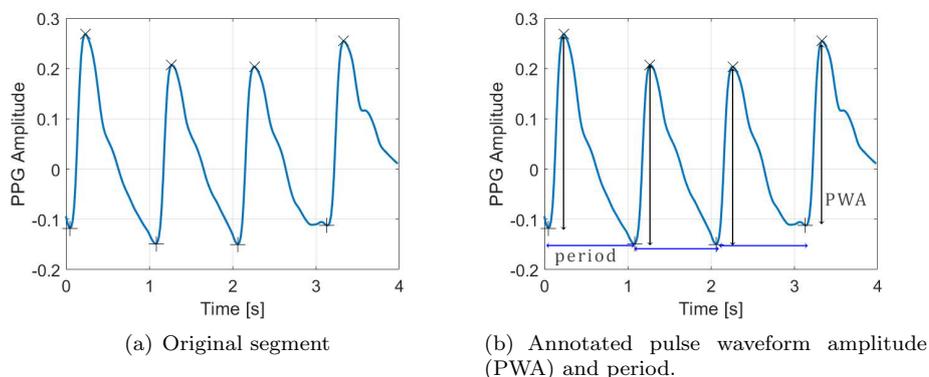


Figure 7.3: **The pulse width amplitude (PWA) is calculated with the first valley and systolic peak of a pulse. The period is the distance between 2 subsequent valleys.**

There should be at least 2 valleys in the segment, as otherwise no pulse can be detected. Then, there are two ways in which the peaks and valleys can be evaluated by using the physiological properties inherent to cardiac cycles. The first method is related to the ‘vertical’ distance between a peak-valley pair. This is referred to as ‘Vertical Distancing’ and the distance is called the Pulse Waveform Amplitude (PWA). The second method uses the duration of pulses, the period, and looks at the time between subsequent valleys or peaks. An illustration of both properties within a segment is given in figure 7.3(b).

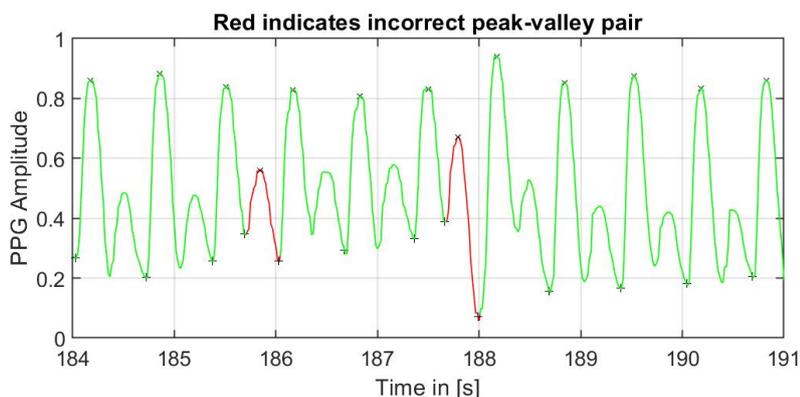


Figure 7.4: **A 4 second segment with correct peak-valley pairs in green and an incorrect peak-valley pair in red.**

Vertical Distancing

Inspiration was drawn from a PWA method proposed by Fischer et. al. [12]. The general idea is that the PWA of subsequent pulses does not change rapidly. This can be quantified by comparing the calculated PWA of a peak-valley

pair to a threshold value. This threshold can be either static or dynamic. Static thresholds have the downside of becoming less robust over time whereas dynamic thresholds need to be initialized and derived. Fischer et. al. [12] compare the PWA values of the previous correct pulse and the current pulse under observation. If the PWAs lie within a range of [25% - 400%], the current PWA is also deemed correct. But determining the right range has a trade-off. If the bounds are too loose, diastolic peaks as in figure 7.4 were classified as correct, but making them too strict did not cope with the natural variance in PWA of subsequent pulses. As such, the following solution is proposed. The three most recent ‘correct’ PWA values are stored in a FIFO (First In First Out) buffer. Each new PWA value is then compared with the mean of the buffer and if the values lie within a range of [50% - 200%] of each other the new PWA value is deemed correct [40]. These thresholds were empirically determined. The buffer is then updated with this value and the process continues. In case of initialization or re-calibration, the buffer needs to be filled with new values.

Pulse Duration

The pulse duration can be derived by calculating the time passed between subsequent systolic peaks. Previously the heart rate was estimated via the FFT with a resolution of 0.25Hz, which gave a coarse estimate. By calculating the average pulse duration in the segment, a new estimate of the HR can be made. As was shown in section 5.2, the camera frame rate is 30fps which results in a sample resolution of $\frac{1}{30}$ Hz. This new resolution is a significant improvement from 0.25Hz. If the old and new estimates differ less than T Hz, the segment is deemed clean and the old estimated heart rate can be updated. T is defined as:

$$T = \pm \frac{2fs}{n} \quad [Hz] \quad (7.2)$$

Here fs is the sample frequency in [Hz] and n the number of data samples. The moving average window sizes $W1$, $W2$ and $W3$ are updated with the estimated heart rate from the pulse duration. In case the estimated heart rates from the pulse duration and FFT are in disagreement, the peaks and valleys are labeled as corrupt and a re-initialization is requested. This means that for the next segment, the moving average window sizes $W1$, $W2$ and $W3$ are reinitialized with the heart rate estimate of the FFT. The PWA buffer is also emptied and reinitialized as described in Section 7.2.4.

7.3 Conclusions

The proposed pulse segmentation method uses elements of the SoA. However, the peak detection method by Elgendi et. al. [10] had several limitations. Firstly, it used static thresholds and was evaluated with offline data-sets, which severely limits the applicability to a real-time domain. Secondly, only peak detection was performed, whereas for proper segmentation the start and end of a pulse are required. As such, improvements to the algorithm are proposed in order to make it adaptive, real-time and let it segment pulses. Finally, the pulse waveform amplitude (PWA) and periods are used to determine whether the PPG signal is of good quality.

Chapter 8

Pulse Qualification

Before the segmented ‘candidate pulses’ can be used, some form of qualification is required. However, many challenges need to be overcome regarding the question: “What is a correct PPG pulse?”. The uniqueness of the PPG signal on a person-by-person basis makes it challenging to create a reference signal that applies to *all* cases. As such, the following sections cover methods that mainly tackle large anomalies. The first section explains the normalization process. Section 8.2 covers a method for separating *two* pulses that have been incorrectly classified as *one* pulse. Lastly, section 8.3.2 discusses several methods related to only a fraction of the signal: the rising slope from the start of the pulse till the systolic peak.

8.1 Pulse Normalization

In the previous chapter, the systolic peaks and pulse valleys have been identified within the segment, such that the pulses can be extracted. During this process, the amplitude of the PPG signal was altered due to e.g. filtering, but the morphology remains largely intact. Furthermore, the heart rate of the user at the time of extraction influences the pulse duration. In order to compare different pulse periods with each other, the amplitude is normalized, and the waveform is interpolated to 100 samples.

8.2 Double Pulse Problem

On some occasions, the segmentation algorithm introduced in the previous chapter misses a peak, and as a result 2 pulses are classified as one. Fortunately, these double pulses can be detected and more importantly separated, such that they can be processed individually. This way, the PPG signal is *corrected* instead of *discarded*. The detection of a double pulse is relatively easy and the used method is demonstrated in section 8.2.1. Then, the pulses need to be separated, which will be discussed in section 8.2.2.

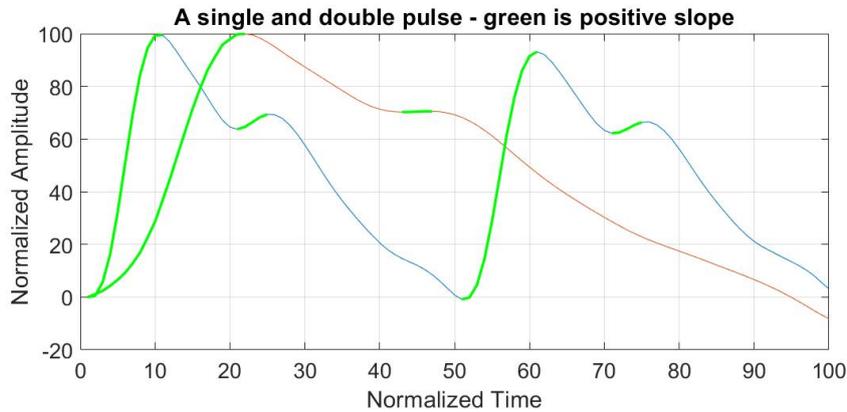


Figure 8.1: **Green segments indicate a positive slope. Each pulse starts with a positive slope segment with large amplitude.**

8.2.1 Detection

The detection of double pulses can be illustrated with figure 8.1, which shows a normal single pulse and a double pulse. If we only consider the normal pulse, it *always* starts with a positive slope segment till the systolic peak. Then, the pulse can continue in two possible ways:

- A negative slope till the end of the pulse in absence of the dicrotic notch.
- A negative slope till the dicrotic notch, positive till the diastolic peak and negative till the end of the pulse.

The focus here lies on the positive segments, marked with green in figure 8.1. No matter what waveform, the first positive slope part is substantially larger in amplitude than *any other* positive slope segment and is *dominant*. To conclude, a **single pulse** consists of either:

- A single large positive slope segment.
- Two positive slope segments, where the first dominates the latter.

A **double pulse** is simply the combination of two pulses, and the properties from above hold in twofold, which means:

- There are *always* two large positive slope segments.

These two large segments are comparable in size and do not dominate each other, making it clearly distinguishable from the single pulse.

8.2.2 Separation

Now that the double pulses are detected, they need to be separated, such that they can be processed individually. For this, the end of the first and the start of the second pulse need to be detected, which is located somewhere in the middle, within in a range of [40-60] in normalized time as demonstrated by figure 8.1. The local minimum in this range indicates the separation point, after which the whole pulse normalization and qualification can be repeated.

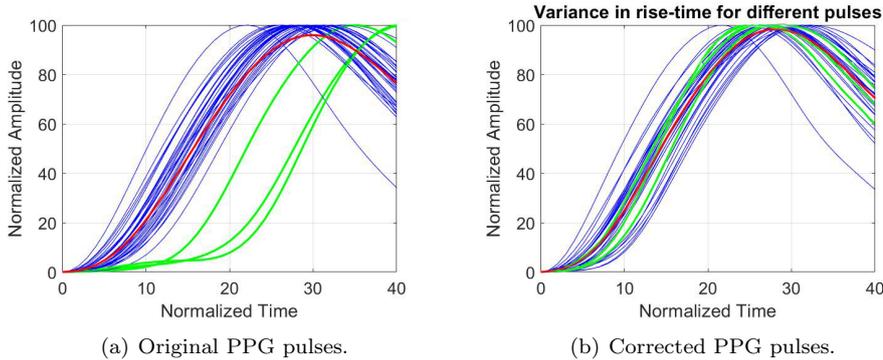


Figure 8.2: Delayed starts of the pulses result in shifted copies of similar pulse waveforms (green), which can be corrected.

8.3 Systolic Pulse Part

First, a method is proposed for detecting and *correcting* pulses where the rise-time is delayed, which is discussed in Section 8.3.1. The second method is a simple threshold check for the time when the systolic peak occurs. This method *discards* incorrect pulses.

8.3.1 Rise-time Correction

Whilst the systolic peak has a strong definition and its occurrence can be easily detected, this is not the case for the start of the pulse. The start of a pulse is the tipping point whereafter the PPG signal monotonically increases till the systolic peak. However, the precise location of this point is difficult to determine, which can be illustrated with figure 8.2(a). The figure shows 3 delayed pulses, which can be corrected.

Solution

To correct the delayed start, a solution is proposed in figure 8.3 on the next page. In figure 8.3(a), a pulse from the set visible in figure 8.2 is displayed with an obvious delay at the beginning. This delay can be detected via a method using the euclidean distance. First, a straight line is drawn from point (0,0) to the occurrence of the systolic peak, $(t_{peak}, 100)$. This line holds the same number of samples as the waveform in this specific range. Then, the smallest euclidean distance from each point to the straight line is calculated and the largest euclidean distance of all these points is used to determine the start.

In the case of correct pulses, the pulse will be very close to the straight line, whereas for delayed pulses a large value is measured. The threshold is determined as $TH_{dist} = 25$. The point at which the euclidean distance is the largest (black cross), corresponds to the rough area of the start of the pulse. However, this point is slightly too far up the slope and an offset to the left is required. The offset was empirically determined as 5 samples, shown with a red cross in figure 8.3(b). Because the pulse no longer consists of 100 samples, it needs to be normalized again. The resulting corrected pulse is illustrated in figure 8.3(c). Applying this method on all pulses from figure 8.2(a) results in the pulses illustrated in figure 8.2(b), which are more compact. However, care

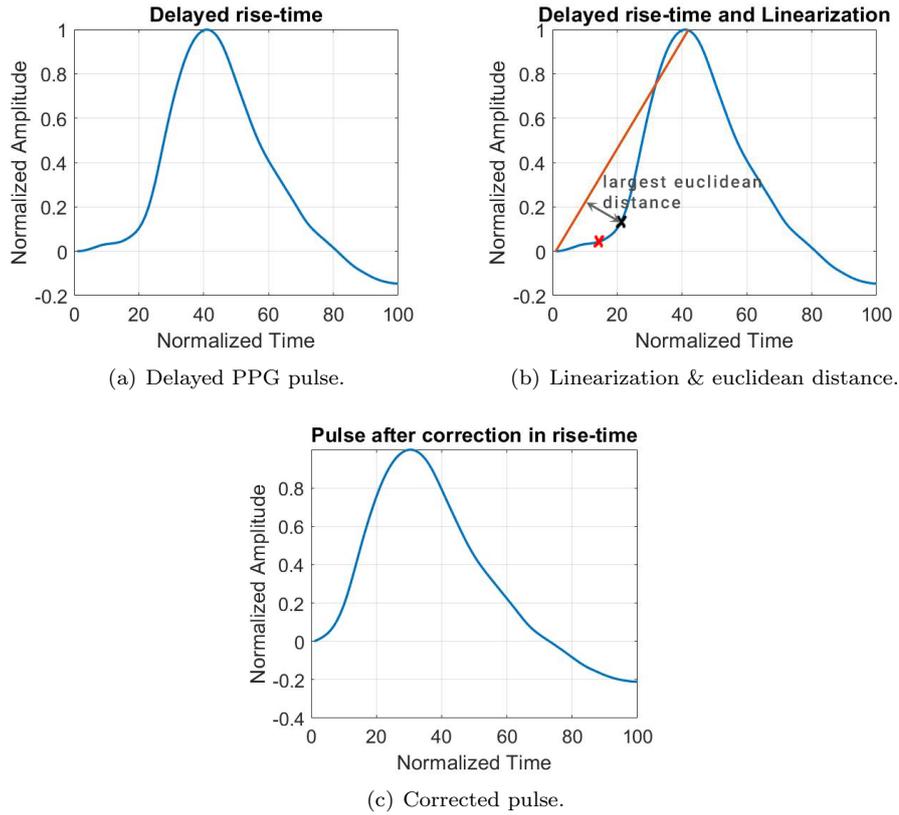


Figure 8.3: **Rise-time correction. Red cross is the new pulse start.**

needs to be taken with this method, because it could remove a large part of the pulse due to severe corruption. In those cases, the pulse needs to be discarded. This can be easily detected by comparing the amplitude of the start and end of the pulse, which should not deviate more than half the peak amplitude, A_{peak} .

8.3.2 Discarding Incorrect Pulses

There is one aspect of the PPG pulse that is common to *all* pulses: the systolic peak. This fundamental property can be used to discard incorrect pulses. The peak *always* occurs in the first half of the waveform. In normalized time [0-100], it always occurs in the first 40% of the pulse. As such, a simple threshold can be used to discard incorrect pulses:

$$t_{peak} \leq 40 \quad (8.1)$$

Where t_{peak} is the occurrence in time of the systolic peak.

Chapter 9

Evaluation

This chapter covers the evaluation of the developed smartphone application, the individual contribution of each stage in the system and how it fares versus a state of the art (SoA) PPG sensor. Two metrics are used to compare the methods, one focuses on the quality of the extracted PPG pulses and the other that measures the efficiency at which they are obtained, the acceptance rate. These performance metrics will be discussed in section 9.1 and 9.2 respectively. Section 9.3 covers the contribution of each system component and the improvement achieved. Finally, section 9.4 compares the smartphone application with the dedicated PPG sensor.

9.1 Measure of Pulse Variation

There are many ways to express the variation between pulses in a dataset. The most obvious way is to calculate the variance Var or standard deviation σ of each pulse, but there is a downside illustrated with figure 9.1. The orange pulse is *identical* to the blue mean-pulse, but delayed by 3 samples. Features extracted from these pulses would produce the exact same values, which should be reflected in the quality metric. However, the variance is calculated between points occurring at the *exact same time* (horizontal arrow in the figure). The large distance between the points in figure 9.1 will result in a large variance, which would mean the quality of the orange pulse is graded worse than is the case in reality.

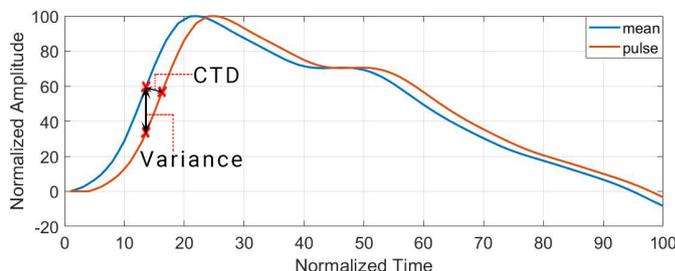


Figure 9.1: The blue line represents the mean of all pulses collected in one experiment and the orange line is a pulse from this set.

The cross track error (CTE), derived from the cross track distance (CTD), does not have this problem [39]. The cross track distance is obtained by calculating the smallest euclidean distance from the point on the blue pulse to the orange pulse, as shown in figure 9.1. A pulse consists of 100 points and the 100 resulting CTDs are averaged to obtain a single CTD representing that pulse. Finally, averaging the CTDs for all pulses in the dataset results in the cross track error (CTE), which is the performance metric for the quality of PPG pulses. Large CTE values indicate large variance and thus a low quality dataset, whereas small CTE values mean that the dataset is of good quality.

9.2 Acceptance Rate

As the PPG signals are obtained from the smartphone camera, there is a long process with stages of qualification involved which either keep or discard the PPG pulses. If the qualification is strict, it can take a very long time to obtain enough pulses for the desired application. On the other hand, if the qualification is too loose, many corrupt pulses are graded as clean. The quality aspect is measured by the cross track distance and for determining the efficiency at which pulses are collected, the acceptance rate, r_a , can be used:

$$r_a = \frac{n_{collected}}{n_{hr}} \quad (9.1)$$

Where $n_{collected}$ is the number of pulses extracted and n_{hr} the total number of pulses possible, derived from the heart rate in the experiment. The heart rate is calculated via the FFT over the entire data for the PPG sensor case. For the smartphone application, only the pulses are stored and the heart rate is obtained by averaging the duration of the pulses collected. However, the PPG signal is collected over a period of 4 seconds to offer real-time feedback, whereafter it is processed to extract the pulses. Sometimes, a pulse is only partially captured in this window and the other part will be collected in the next segment, shown with figure 9.2.

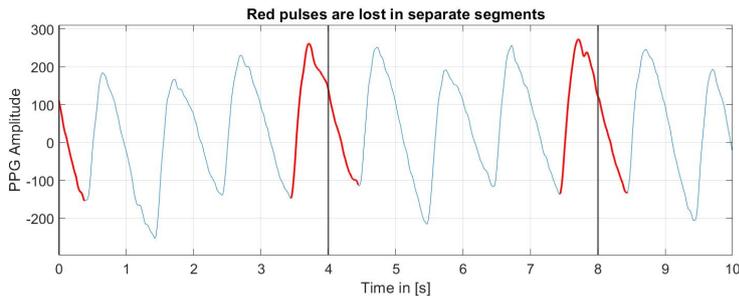


Figure 9.2: **The red pulses are lost as they fall between 2 segments.**

Due to processing overhead of the application between the 4 second periods, these parts cannot simply be stitched together to create a pulse, because no samples are collected during that downtime. Furthermore, the lower the heart rate the greater the impact of losing one pulse due to overlap, as there are less pulses in the segment than with a high heart rate. With the current system, only

70% of the pulses in the PPG signal are collected at best. Since the PPG sensor and smartphone operate with the same design limitations, the acceptance rate needs to be adjusted. In the worst case, there is one pulse lost in the overlap per segment and equation 9.1 is adjusted to:

$$r_a = \frac{n_{collected}}{n_{hr} - n_{segments}} \quad (9.2)$$

where $n_{segments}$ is the number of segments in the experiment. This way, r_a takes not only the design implementation into account, but also the heart rate. r_a as defined in equation 9.2 will be used for the evaluation in section 9.4.

9.3 Impact of Individual System Stages

The system proposed in this thesis consists of several stages and the influence of each stage needs to be evaluated. However, in order to obtain the CTE a pulse segmentation (PS) algorithm is needed at minimum. The following aspects will be evaluated:

1. The peak detection by Elgendi et. al [10] only detect peaks, not pulses and as such, the valley detection from section 7.2.3 is added. This method will be referred to as ‘Original PS’.
2. The proposed pulse segmentation (New PS) including all algorithms proposed in chapter 7.
3. Finger pressure detection (FPD) from chapter 6 is added to the New PS stage.
4. Pulse qualification (PQ), which checks the quality of individual pulses from chapter 8. With the addition of this stage, the entire pipeline consists of New PS, FPD and PQ.

9.3.1 Experiment Setup

For the evaluation, 4 individuals participated in the experiments. Measurements were obtained from the right index finger. It is difficult to evaluate each stage within the smartphone application and instead a video is recorded which will then be processed offline in MATLAB[®] ver. R2018b. The camera settings are manually set according to the recommendations in section 5.3.5. The real-time environment is simulated by providing the algorithms with non-overlapping 4 second PPG segments one at a time. As was mentioned in section 6.2, continuous measurements can only be done for a relatively short time due to the heat of the flash light, which is why a period of 60 seconds was chosen. Subjects were asked to first apply 30 seconds of appropriate pressure, whereafter they would increase the pressure to a level of ‘strong pressure’ and hold that pressure for 30 seconds. This way, the obtained PPG signal consists of a ‘clean’ and ‘corrupt’ portion.

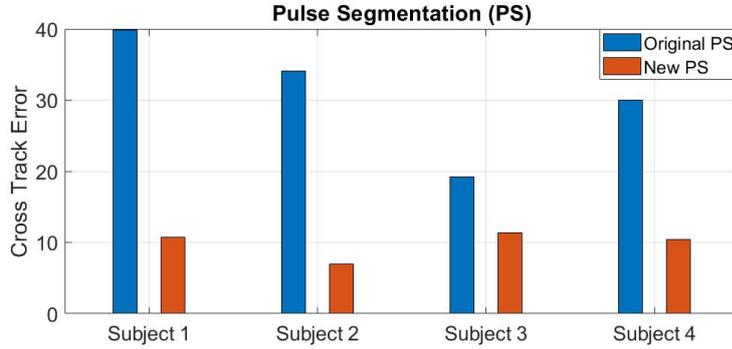


Figure 9.3: **The improvements on Elgendi’s [10] method reduce the CTE by 67.94% on average.**

Firstly, a comparison is made between the pulse segmentation algorithms. Figure 9.3 shows that the original PS algorithm by Elgendi et. al. is unable to handle user-related distortions for all subjects to varying degrees. The peak & valley decision logic from the new PS plays an important role in discarding corrupt segments. Although the average reduction in CTE of 67.94% is substantial, The average CTE = 9.87 is still too large. Further processing is required by adding FPD and PQ to New PS and their effects are displayed in figure 9.4.

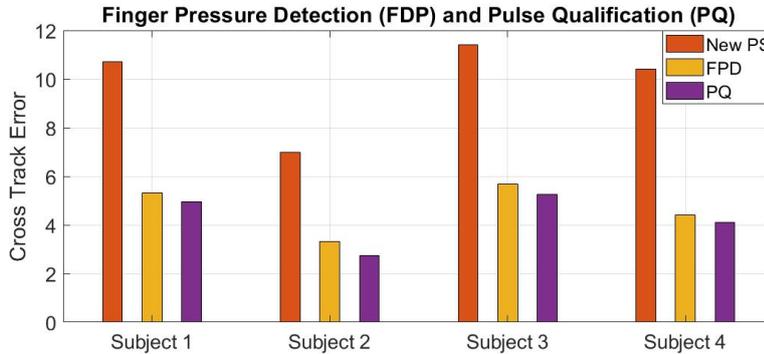


Figure 9.4: **Adding FPD and PQ to the system reduces the CTE by 50% on average, compared to only using the new PS.**

The orange bar in figure 9.4 refers to the pulse segmentation as proposed in this thesis, which was evaluated above and corresponds to the same orange bar from figure 9.3. The other bars introduce the addition of finger pressure detection (FPD) and pulse qualification (PQ). This means that ‘PQ’ in figure 9.4 is the complete combination of the new PS, FPD and PQ. Figure 9.4 shows that the finger pressure detection reduces the CTE with 52.57% on average, which is to be expected. Pulse qualification only reduces the CTE with 8.89% for all subjects, as it eliminates only a few corrupt pulses.

9.3.2 Summary of Improvements

Figure 9.5 shows the average CTE reduction per stage. The CTE with the original pulse segmentation method from figure 9.3 serves as the baseline.

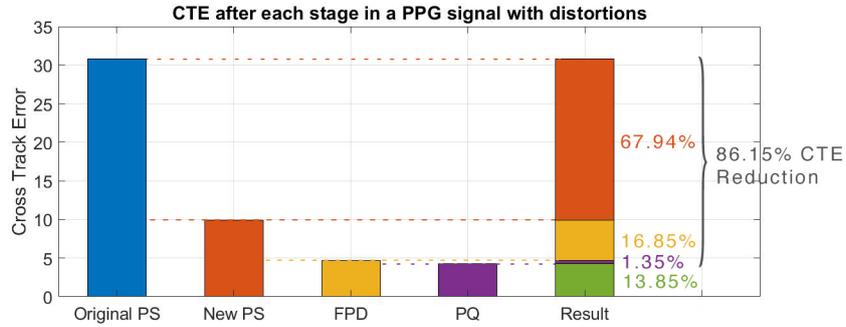


Figure 9.5: **The addition of each stage reduces the CTE with 86.15%, where the CTE with the original pulse segmentation serves as the 100% baseline.**

Because each of the three stages checks the signal quality, severe distortions are easily detected and thus removed at an earlier stage. As such, the first stage, the new pulse segmentation (PS), is responsible for the largest CTE reduction (67.94%). A large part of the PPG signals in the experiment were so distorted due to the strong finger pressure, that even without finger pressure detection (FPD) several corrupt PPG signal parts could be eliminated. The addition of FPD stage detects the remaining distorted pulses. As the pulses get further in the qualification pipeline, there are less distortions in the signal and those present are more subtle. This is why only a very small improvement is observed for the pulse qualification (PQ) stage. Finally, the combination of all stages leads to a reduction in CTE of 86.15% when compared with the bare minimum, the SoA pulse segmentation.

9.4 PPG Sensor versus Smartphone App

In order to evaluate the smartphone application and the influence of the camera settings, a comparison with an actual PPG sensor needs to be made. Both will use the same pulse segmentation and pulse qualification methods to aggregated pulses. The only difference is how they perceive light intensity changes and how this is converted to a PPG signal. Furthermore, the finger pressure is handled differently. Two experiments of 60 seconds with 2 weeks between them are held to also account for the time-invariance of PPG signals.

9.4.1 Experiment Setup

The PPG sensor is called SDPPG by APMKorea [2] and is depicted in figures 9.6(a).

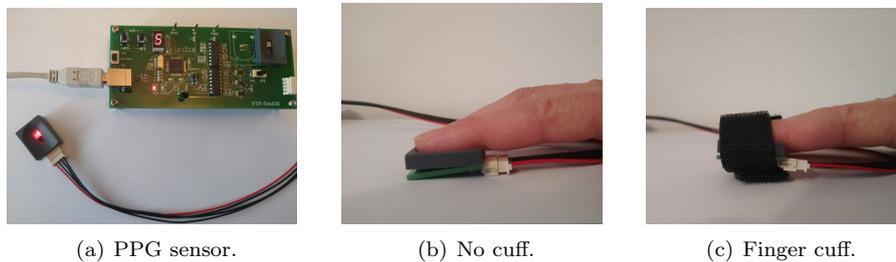


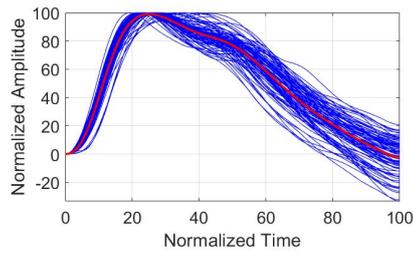
Figure 9.6: **PPG sensor modification to make it similar to medical devices.**

The sensor unit consists of a red LED and photodiode. Furthermore, a finger mould helps the user in correctly placing the finger. However, 2 modifications are required for the sensor to be comparable with SoA medical equipment. The first modification is related to the finger pressure. With the setup illustrated in figure 9.6(b), the user is entirely responsible for the degree of pressure. Fortunately this is easily solved with a custom-made finger-cuff of ‘velcro’, as depicted in figure 9.6(c). The finger pressure is now constant and no longer controlled by the user. The second modification is related to ambient light. Medical PPG sensors try to eliminate disturbances of ambient light via a cover over the finger. Here this is accomplished by covering the finger with a cloth, such that only the light of the LED is received by the photodiode. The device is provided with software that allows for logging the data with a sample rate of $f_s = 2000\text{Hz}$. This data is then provided as 4 second segments to the algorithms to simulate the real-time environment.

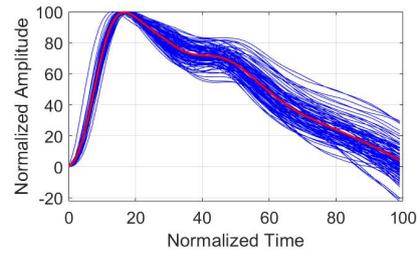
The smartphone requires two hands to operate which means that first the PPG sensor is tested and not long thereafter the smartphone. The pulses are stored in a SQLite database on the smartphone and the database can either be moved from the smartphone to PC, or simply evaluated in the application itself.

9.4.2 Results

Figures 9.7, 9.8, 9.9 and 9.10 show the pulses collected in the two experiments of 60 seconds each for 4 different people.

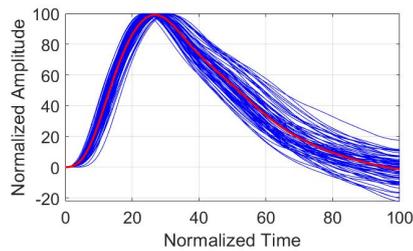


(a) PPG Sensor: $r_a = 0.86$, CTE = 3.44

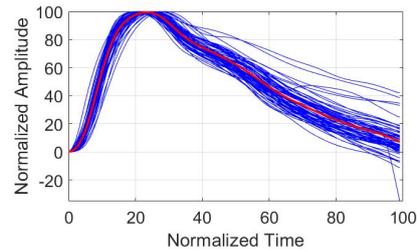


(b) Smartphone: $r_a = 0.92$, CTE = 3.43

Figure 9.7: **Subject 1: Aggregated plot of pulses collected within two experiments of 60 seconds.**

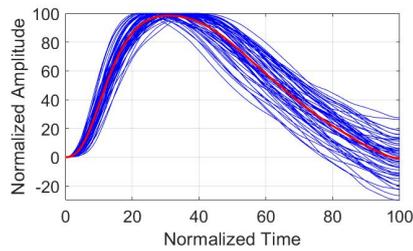


(a) PPG Sensor: $r_a = 0.86$, CTE = 3.33

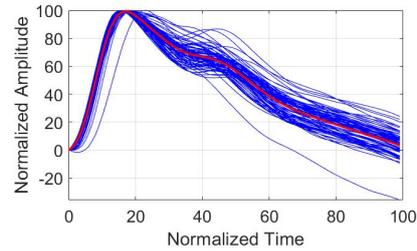


(b) Smartphone: $r_a = 0.72$, CTE = 3.23

Figure 9.8: **Subject 2: Aggregated plot of pulses collected within two experiments of 60 seconds.**

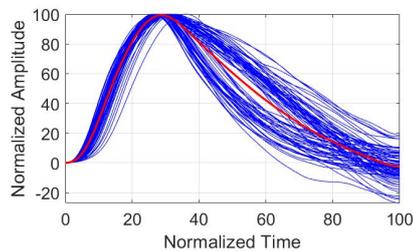


(a) PPG Sensor: $r_a = 0.74$, CTE = 3.91

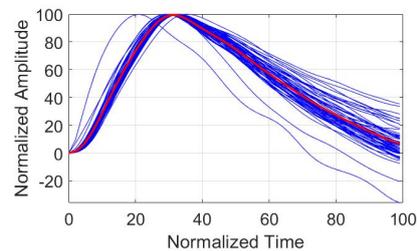


(b) Smartphone: $r_a = 0.94$, CTE = 3.54

Figure 9.9: **Subject 3: Aggregated plot of pulses collected within two experiments of 60 seconds.**



(a) PPG Sensor: $r_a = 0.96$, CTE = 4.50



(b) Smartphone: $r_a = 0.70$, CTE = 2.80

Figure 9.10: **Subject 4: Aggregated plot of pulses collected within two experiments of 60 seconds.**

The heart-rate of the subjects was in the range of [50-73] BPM, which means that in theory, 50 to 73 pulses are present in a 60 second window, which should result in at least 100 pulses for 2 experiments. However, the number of pulses actually collected is much lower as was explained in section 9.2. Furthermore, the $r_a=0.96$ with the PPG sensor for subject 4 indicates that, from the pulses available, almost all were good enough to keep.

Surprisingly, all subjects obtain a lower CTE for the pulses collected with the smartphone compared with the regular PPG sensor. This is only at the cost of 5,7 and 24 pulses for subjects 1, 2 and 4 respectively. Subject 3 on the other hand saw an increase of 19 in the number of pulses collected. The reduced number of pulses is due to the finger pressure detection. Users are responsible for exerting the correct pressure via feedback from the smartphone, which means that segments will be discarded if the pressure is incorrect. Users need to continuously monitor the finger pressure themselves. This is not the case for the PPG sensor as the finger pressure is fixed by the finger-cuff.

Subjects 2, 3 and 4 show a couple pulses with large deviation from the other pulses in the smartphone case. During the real-time pulse collection, these pulses passed all requirements for being ‘clean’. They can only be removed *after* the PPG pulse aggregation if desired. For the PPG sensor, two different waveforms can be distinguished for subject 4, as shown by figure 9.11.

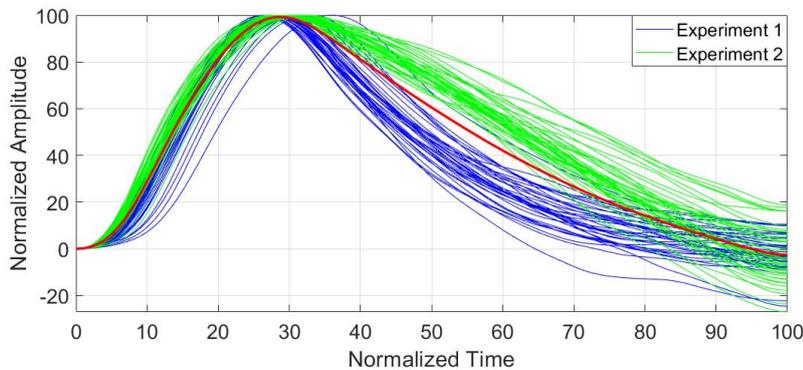


Figure 9.11: **Two different pulse waveforms are obtained per experiment with CTEs of 2.97 and 2.55 for experiments 1 and 2 respectively.**

These ‘waveform classes’ are obtained during the two separate experiments and the change in shape is due to the contact pressure between sensor and finger. In hindsight, for experiment 1 the finger-cuff was too tight, which gives rise to a waveform corresponding to a strong contact finger pressure as was shown in section 6.2. Although no dicrotic notch or diastolic peak are present, the signal initially rapidly declines from the systolic peak towards the end. For experiment 2, the finger-cuff was less tight, and the decline is more gradual.

9.4.3 Summary of Improvements

Figure 9.12 shows a summary of the CTE for all tests where figure 9.12(a) shows the comparison in CTE per sensor device and figure 9.12(b) shows the acceptance rate r_a . For the CTE, an average reduction of 12.63% is obtained by going from the PPG sensor to the smartphone. The average r_a only drops with 3.53% going from the PPG sensor to the smartphone.

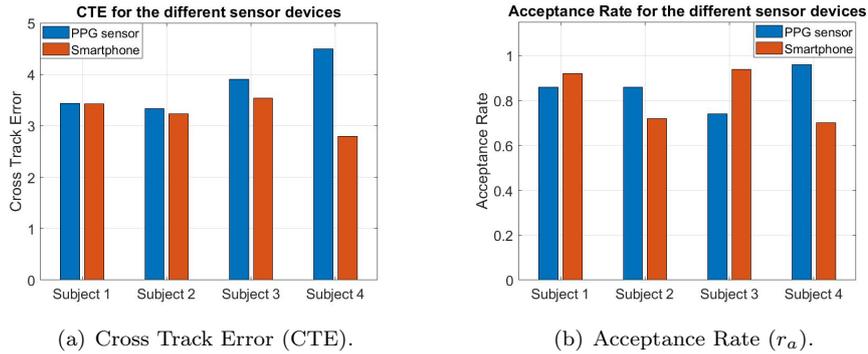


Figure 9.12: **The CTE of the aggregated pulses is lower for the smartphone application. The average reduction in CTE by using the smartphone is 12.63%.**

Although the CTE for subject 4 with the PPG sensor is worse due to the finger pressure difference, it is shown that overall the smartphone application can perform equally or even better.

Chapter 10

Conclusions

10.1 Conclusions

In the recent years, a surge in interest for mobile health monitoring is seen. Several smartphone applications are developed to estimate e.g. heart rate, blood pressure or blood oxygen level, but the correctness of these estimates has become more concerning. More specifically, photoplethysmography (PPG) via the smartphone camera and flash light proves to be challenging, as different sources of noise and motion by the user can distort the PPG signal. For instance, the pressure applied by the finger on the sensor severely impacts signal quality. This was originally solved with a finger-clip or cuff, but a method without additional hardware is desired. Furthermore, the different parameters of the smartphone camera are not properly explored and as such there is a risk that sub-optimal settings are used. Moreover, many methods involving PPG signals are evaluated in an offline setting. Lastly, PPG signals from a single individual are similar, but different from other individuals. This uniqueness of the PPG signal complicates the robustness of algorithms, as they need be as general as possible.

In this thesis, a more extensive study of the smartphone camera settings is provided than was previously done in literature, highlighting their importance for correct PPG signal extraction. A smartphone application was developed that not only extracts the PPG signal but also verifies the quality. PPG signals obtained are 86.15% more compact than with a SoA method. A novel real-time pulse segmentation algorithm is proposed to extract pulses from the PPG signal. As pulses are aggregated over time, they can be viewed in the application and a quality indicator, the cross track error, is calculated to show the variation between the captured pulses.

The smartphone application was compared with a dedicated PPG sensor for evaluation purposes. 4 Subjects were asked to participate in the experiments and on average the signals obtained from the smartphone were 12.63% more compact, indicating pulses of higher quality. This came only at the cost of a 3.53% lower acceptance rate. As a result, the developed system performs better or in the worst case, similarly as the regular PPG sensor.

Lastly, this thesis introduced a real-time system that provides high quality pulses, which is a means to an end. Other related studies can take advantage of the contributions and recommendations done here to improve their meth-

ods. Furthermore, the evaluation of the smartphone camera settings revealed an unexplored rabbit hole, covering the image processing pipeline present in smartphones (appendix A). These unexplored research areas keep the future uncertain but promising.

10.2 Future Work

Although the initial goals of this thesis were achieved, the system cannot be considered as finished. First of all, the task scheduling and PPG signal extraction can be done more efficiently, such that more pulses can be collected in a period. Right now a pulse that is split between 2 measurement windows is lost, which can be prevented by buffering the frames. However, a complex scheduler is required that manages the tasks, their priority and duration.

Secondly, only the tip of the veil was lifted concerning the camera settings and processing techniques in the smartphone. Many transformations are applied on an image captured by the camera before it finally is displayed. These transformations could be non-linear, which affect the pulse shape in an unknown way. In chapter 9 it could be seen that using a different sensor device already affects the shape of the pulses collected. Appendix A goes into more detail regarding the image processing pipeline and the related future work.

Thirdly, the application was only tested with a single smartphone, the Motorola G7 Plus. Attempts were made to test with the Redmi 5A phones available at the TU Delft, but they require several modifications of the firmware such that the camera settings can be manually adjusted. This was not possible to do in time for this thesis. The application will behave differently on another device, as the camera, flashlight and image processing pipeline are different. Furthermore, older smartphones are unable to use the application due to the need of manual camera access, which was only released in 2014.

Last but not least, in order to make the system more robust for different individuals, a large and varied data set is required. This way, edge cases can be discovered and a better understanding of the features shared among all pulses is obtained. Furthermore, physical factors such as skin color and thickness need to be properly explored, as the skin tone influences the penetration and reflection of light.

Appendix A

Image Processing Pipeline

Before addressing the main concern provided with the information here, first a revisit of the history of PPG and its fundamental is required.

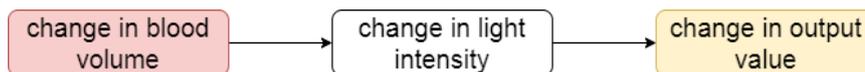


Figure A.1: **Signal transformations required before blood volume change can be monitored.**

PPG can be briefly described as follows: A change in blood volume can be observed by measuring the change in light intensity at that specific site, given a light emitter and receiver. This relationship has been proven over and over again and will be referred to as the *qualitative relation*. This property applies to all methods developed in literature with regards to PPG monitoring. The second property is called the *quantitative relation* which mathematically describes the relationship between blood volume change and light intensity change. Blood volume changes cannot be directly observed and as such a sensor is required that introduces an additional transformation to e.g. electric current, as is illustrated with figure A.1. In contrast with the previous property, this transformation is completely dependent on the used method.

Before the invention of smartphones, PPG monitoring in medical applications was done with sensors consisting of an emitting light with a known wavelength from the visible light spectrum and a photodiode as receiver. A photodiode is a device that converts light to an electric current with a known quantitative relation, since these devices have been studied broadly. Important to note here is that based on the generated current, a mapping back to light intensity can be made since that relation is known.

The invention of the smartphone would revolutionize the world as was known before then, providing endless possibilities including for the measurement of biometric signals. Smartphones were equipped with cameras and flashlights and it took not very long for researchers to find a link with contemporary PPG monitoring [37]. The flashlight would act as the light emitter and the smartphone

camera as the receiver. The general idea is that as the blood volume changes and results in a change of light intensity, this change can be tracked by monitoring the pixel intensity levels of images shot by the camera over time. More specifically, the RGB channels of the pixels in an image. However, the flashlight consists of a wide range of wavelengths instead of one and information about these wavelengths is often kept private by the manufacturers. Furthermore, the wide range of different commercial smartphones with different flashlights complicates the matter since specific hardware information is also kept secret. As such, the relation between light intensity levels and generated pixel intensity levels is unknown and has not been thoroughly studied, in stark contrast with the contemporary PPG monitoring methods. Initially, many studies assumed that the mapping of light intensity to observed output is the same as for the regular state of the art sensors, without considering the aforementioned factors.

Fortunately, some studies recognized the importance of considering the smartphone camera properties for proper PPG monitoring. For instance, there have been many studies on the RGB channel selection for the pixels in an image; which channel provides the ‘best’ information [23]. Algorithms have been developed to construct PPG signals from these channels with mathematical formulas or even select individual pixels in an image as was depicted in table 5.1 of Section 5.2. In 2019, Liu et. al. [30] considered multiple camera settings such as sensor sensitivity (ISO), flashlight intensity, focus, shutter speed and white balance for their PPG application. However, they only cover superficial parameters that can be changed in a smartphone, which happen to be mainly photography related parameters, whereas PPG monitoring is completely different. For some studies, this could be attributed to the limited access to the camera hardware.

“The ability to edit its image processing pipeline has been held exclusively by the camera manufacturers. Google is trying to change the game with the release of the Camera 2 API for Android¹.”

The case of limited access to the camera only applies to the studies before the release of the Camera 2 API in 2014. Liu et. al. [30] consider only a minuscule fraction of the settings that interact with the Camera 2 API. Figure A.2 shows a generalized image processing pipeline present in smartphones, with the several stages required to go from the raw sensor values to an image that can be displayed. Only stage 6, White-balancing and stage 8, Exposure curve, have been covered in literature with the addition of stage 12, Gamma curve by this thesis. This is only $\frac{3}{12}$ th of the entire pipeline. Recall that the goal of controlling these settings is to obtain a quantitative relation between the changing light intensity and pixel intensity. Each of the stages in figure A.2 introduces a transformation or modification on the raw values as perceived by the smartphone camera sensor, altering the original information it carried. At the time of writing, no research has been done on the contributions of each stage to the final PPG signal extracted from the RGB pixel intensity levels.

There is no clear reason as to why nobody has investigated the image processing pipeline, but there are a few possible explanations. *Disclaimer:* the

¹Min Jae Kim - May 15 2019 - <https://medium.com/@kimminjae/android-camera2-api-pipeline-manipulation-dc3becd59a36>

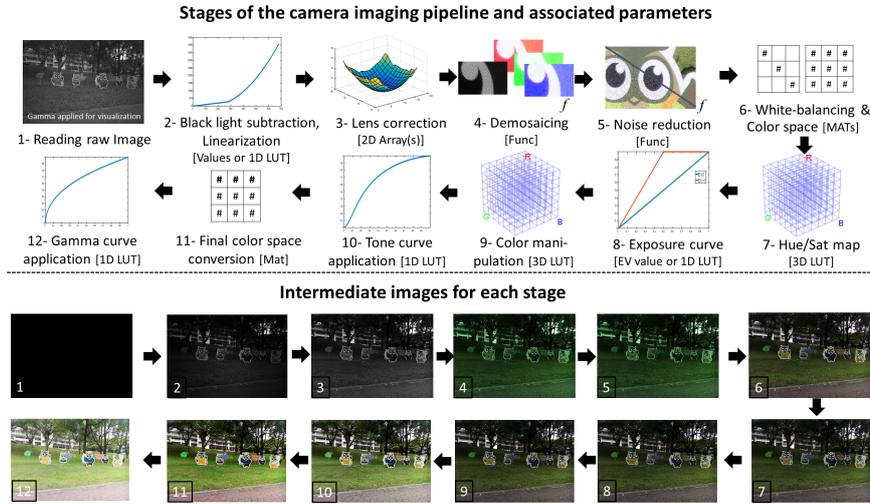


Figure A.2: A generalized image processing pipeline of a smartphone [24].

following is based on what I have researched the past year and the knowledge that I have gained about the topic and as such is mostly opinion driven.

The *first* reason is that several papers covering smartphone PPG monitoring verify their methods *offline*. What this means is that first, a video is recorded with a smartphone and then on a PC the method is applied to obtain results. Developing a custom app to record videos is seen as unnecessary since many apps on the AppStore or PlayStore are available. However, these apps are intended for ‘normal’ photography and allow users to tune camera parameters on a rather basic and limited level even with Camera 2 API capabilities.

The *second* reason is that the image processing pipeline is a complicated process that even varies per device. Investigating each stage would be a very tough and tedious task. Smartphones do allow for raw sensor output captures, but the interpretation is device dependent which introduces the portability issue of apps developed on one device and then used on another.

Thirdly, a large part of smartphone PPG monitoring has been focused on obtaining the heart-rate, which is rather trivial. This is because only the peak of the PPG pulse is required. The quality of the remainder of the pulse is irrelevant, which means that such methods are much more robust versus MNAs and image processing transformations. The moment features of the pulse waveform are considered, these influences become important.

Based on these findings, we are left with many questions. Can we omit the entire pipeline and directly convert the raw sensor data to a PPG signal? What is the influence of each stage on the PPG signal? Can we use the pipeline to increase the PPG signal quality? In order to answer these questions, research on the image processing pipeline is required, which is well beyond the scope of this thesis and could be a thesis on its own. What I have tried to do here, is bring attention to a research area yearning for exposure as it has been neglected, despite its possibly crucial impact on the PPG signal.

Bibliography

- [1] M. Aboy, J. McNames, Tran Thong, D. Tsunami, M.S. Ellenby, and B. Goldstein. An automatic beat detection algorithm for pressure signals. *IEEE Transactions on Biomedical Engineering*, 52(10), September 19 2005.
- [2] APMKorea. Sdppg_v2.0 user guide. [http://apmkr.com/bio-device/SDPPG%20V2.0%20user%20guide\(En\).pdf](http://apmkr.com/bio-device/SDPPG%20V2.0%20user%20guide(En).pdf), 2020. Last accessed: Sep. 3, 2020.
- [3] Eli Billauer. peakdet: Peak detection using matlab. <http://billauer.co.il/peakdet.html>, 2012. Last accessed: Sep. 4, 2020.
- [4] Nam Bui, Anh Nguyen, Phuc Nguyen, Hoang Truong, Ashwin Ashok, Thang Dinh, Robin Deterding, and Tam Vu. Pho2: Smartphone based blood oxygen level measurement systems using near-ir and red wave-guided light. *SenSys '17*, November 6, 2017.
- [5] Nam Bui, Jianliang Xiao, Robin Deterding, Thang Dinh, Tam Vu, Nhat Pham, Jessica Barnitz, Zhanan Zou, Phuc Nguyen, Hoang Truong, Taehong Kim, Nicholas Farrow, and Anh Nguyen. ebp: A wearable system for frequent and comfortable blood pressure monitoring from user's ear. In *MOBICOM'19*, pages 1–17, 05 2019.
- [6] Jake Campbell, James Chase, and Phillip Bones. Near-real-time detection of pulse oximeter ppg peaks using wavelet decomposition. *IFAC-PapersOnLine*, 51:146–151, January 2018.
- [7] Anand Chandrasekhar, Chang-Sei Kim, Mohammad Yavarimanesh, Keerthana Natarajan, Jin-Oh Hahn, and Ramakrishna Mukkamala. Smartphone-based blood pressure monitoring via the oscillometric finger-pressing method. *Science Translational Medicine*, 10(431), March 7, 2018.
- [8] Anand Chandrasekhar, Mohammad Yavarimanesh, Keerthana Natarajan, Jin-Oh Hahn, and Ramakrishna Mukkamala. Ppg sensor contact pressure should be taken into account for cuff-less blood pressure measurement. In *IEEE Transactions on Biomedical Engineering (Early Access)*, February 28, 2020.
- [9] Sergey Chernenko. Alpha-trimmed mean filter. <http://www.librow.com/articles/article-7>. Last accessed: Sep. 19, 2020.
- [10] Mohamed Elgendi, Ian Norton, Matt Bearley, Derek Abbott, and Dale Schuurmans. Systolic peak detection in acceleration photoplethysmograms

- measured from emergency responders in tropical conditions. *PLOS ONE*, 8(10), October 22, 2013.
- [11] Umar Farooq, Dae-Geun Jang, Jang-Ho Park, and Seung-Hun Park. Ppg delineator for real-time ubiquitous applications. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2010.
 - [12] Christoph Fischer, Benno Dömer, Thomas Wibmer, and Thomas Penzel. An algorithm for real-time pulse waveform segmentation and artifact detection in photoplethysmograms. *IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS*, 21(2), March 2017.
 - [13] Daisuke Fujita and Arata Suzuki. Evaluation of the possible use of ppg waveform features measured at low sampling rate. *IEEE Access*, 7, May 15, 2019.
 - [14] Andris Grabovskis, Zbignevs Marcinkevics, Uldis Rubins, and Edgars Kvišis-Kipge. Effect of probe contact pressure on the photoplethysmographic assessment of conduit artery stiffness. *J. of Biomedical Optics*, 18(2), February 1, 2013.
 - [15] Tanner Helland. How to convert temperature (k) to rgb: Algorithm and sample code. <https://tannerhelland.com/2012/09/18/convert-temperature-rgb-algorithm-code.html>, 2012. Last accessed: Jun. 12, 2020.
 - [16] Alrick B. Hertzman. Photoelectric plethysmography of the fingers and toes in man. In *Proceedings of the Society of Experimental Biology and Medicine*, 1937.
 - [17] Alrick B. Hertzman. The blood supply of various skin areas as estimated by the photoelectric plethysmograph. In *Proceedings of the Society of Experimental Biology and Medicine*, July 18, 1938.
 - [18] Cambridge in Color. Tutorials: White balance. <https://www.cambridgeincolour.com/tutorials/white-balance.htm>, 2005. Last accessed: Jun. 11, 2020.
 - [19] Dae-Geun Jang, Sangjun Park, Minsoo Hahn, and Seung-Hun Park. A real-time pulse peak detection algorithm for the photoplethysmogram. *International Journal of Electronics and Electrical Engineering*, pages 45–49, January 2014.
 - [20] Jenn Mishra. How to easily understand the exposure triangle: Aperture, shutter speed & iso. <https://expertphotography.com/exposure-triangle/>, 2020. Last accessed: Jun. 11, 2020.
 - [21] Wen Jun Jiang, Peter Wittek, Li Zhao, and Shi Chao Gao. Adaptive thresholding with inverted triangular area for real-time detection of the heart rate from photoplethysmogram traces on a smartphone. In *Conf Proc IEEE Eng Med Biol Soc*. PubMed, August 26–30, 2014.
 - [22] E. Jonathan and Martin Leahy. Investigating a smartphone imaging unit for photoplethysmography. *Physiological Measurement*, 31(11), September 24, 2010.

- [23] Alexei A. Kamshilin and Nikita B. Margaryants. Origin of photoplethysmographic waveform at green light. *Physics Procedia*, 86(2017), February 20, 2017.
- [24] Hakki Can Karaimer and Michael S. Brown. A software platform for manipulating the camera imaging pipeline. In *European Conference on Computer Vision (ECCV)*, 2016.
- [25] Walter Karlen, Mark Ansermino, and Guy Dumont. Adaptive pulse segmentation and artifact detection in photoplethysmography for mobile applications. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 8 1, 2012.
- [26] Walter Karlen, Mark Ansermino, and Guy Dumont. Photoplethysmogram signal quality estimation using repeated gaussian filters and cross-correlation. *Physiological Measurement*, 33(10):1617–1629, September 2012.
- [27] Walter Karlen, Chris Petersen, Jennifer Gow, Mark Ansermino, and Guy Dumont. Photoplethysmogram processing using an adaptive single frequency phase vocoder algorithm. In *Biomedical Engineering Systems and Technologies (BIOSTEC)*, January 26, 2011.
- [28] Francesco Lamonaca, Yuriy Kurylyak, Domenico Grimaldi, and Vitaliano Spagnuolo. Reliable pulse rate evaluation by smartphone. In *2012 IEEE International Symposium on Medical Measurements and Applications Proceedings*. IEEE, May 18–19, 2012.
- [29] Tom Lister, Philip A. Wright, and Paul H. Chappell. Optical properties of human skin. *Journal of Biomedical Optics*, 17(9), September 24, 2012.
- [30] Jian Liu, Cong Shi, Yingying Chen, Hongbo Liu, and Marco Grutese. Cardiocam: Leveraging camera on mobile devices to verify users while their heart is pumping. *MobiSys '19*, June 18, 2019.
- [31] Jing Liu, Bryan Ping-Yen Yan, Wen-Xuan Dai, Xiao-Rong Ding, Yuan-Ting Zhang, and Ni Zhao. Multi-wavelength photoplethysmography method for skin arterial pulse extraction. *Biomed Opt Express*, 7(10), October 1, 2016.
- [32] M. A. F. Pimentel and A. E. W. Johnson and P. H. Charlton and D. Birrenkott and P. J. Watkinson and L. Tarassenko and D. A. Clifton. Bidmc ppg and respiration dataset. <https://www.physionet.org/content/bidmc/1.0.0/>, 2016. Last accessed: Sep. 4, 2020.
- [33] M. Kachuee and M. M. Kiani and H. Mohammadzade and M. Shabany. Cuff-less blood pressure estimation data set. <https://archive.ics.uci.edu/ml/datasets/Cuff-Less+Blood+Pressure+Estimation>, 2015. Last accessed: Sep. 4, 2020.
- [34] Yuka Maeda, Masaki Sekine, and Toshiyo Tamura. The advantages of wearable green reflected photoplethysmography. *J Med Syst*, 35(5), May 18, 2010.

- [35] Kenta Matsumura, Peter Rolfe, Jihyoung Lee, and Takehiro Yamakoshi. iphone 4s photoplethysmography: Which light color yields the most accurate heart rate and normalized pulse volume using the iphysiometer application in the presence of motion artifact? *PLOS ONE*, 9(3), March 11, 2014.
- [36] K. Matthes. Untersuchungen über die sauerstoffsättigung des menschlichen arterienblutes” [studies on the oxygen saturation of arterial human blood]. *Naunyn-Schmiedeberg’s Archives of Pharmacology (in German)*, 179(6), 1935.
- [37] P. Pelegris, K. Banitsas, T. Orbach, and K. Marias. A novel method to detect heart beat rate using a mobile phone. In *Engineering in Medicine and Biology Society (EMBC)*. IEEE, August31 2010.
- [38] Ming-Zher Poh, Daniel McDuff, and Rosalind Picard. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics express*, 18:10762–74, 05 2010.
- [39] J. Rounsaville, J. Dvorak, and Tim Stombaugh. Methods for calculating relative cross-track error for asabe/iso standard 12188-2 from discrete measurements. *Transactions of the ASABE*, 59:1609–1616, 12 2016.
- [40] Alan Rozanski, Ehtasham Qureshi, Mara Bauman, George Reed, Giora Pillar, and George A. Diamond. Peripheral arterial responses to treadmill exercise among healthy subjects and atherosclerotic patients. *Circulation*, 103(16), April 2001.
- [41] Christopher Scully, Jinseok Lee, Joseph Meyer, Alexander Gorbach, Domhnall Granquist-Fraser, Yitzhak Mendelson, and Kaye Chon. Physiological parameter monitoring from optical recordings with a mobile phone. *IEEE transactions on bio-medical engineering*, 59:303–6, 07 2011.
- [42] Christopher G. Scully, Jinseok Lee, Joseph Meyer, Alexander M Gorbach, Domhnall Granquist-Fraser, Yitzhak Mendelson, and Ki H Chon. Physiological parameter monitoring from optical recordings with a mobile phone. *IEEE Trans Biomed Eng.*, 59(2), October 19, 2012.
- [43] J.W. Severinghaus. Takuo aoyagi: discovery of pulse oximetry. *Anesth Analg*, 105(6), December 2007.
- [44] Jai Sim, Bongyoung Ahn, and Il Doh. A contact-force regulated photoplethysmography (ppg) platform. *AIP Advances*, 8:045210, 04 2018.
- [45] Jai Kyoung Sim, Bongyoung Ahn, and Il Doh. A contact-force regulated photoplethysmography (ppg) platform. *AIP Advances*, 8(4), April 6, 2018.
- [46] Fatemehsadat Tabei, Rajnish Kumar, Tra Nguyen Phan, David D. McManus, and Jo Woon Chong. A novel personalized motion and noise artifact (mna) detection method for smartphone photoplethysmograph (ppg) signals. *IEEE Access*, 6, October 16, 2018.
- [47] Toshiyo Tamura, Yuka Maeda, Masaki Sekine, and Masaki Yoshida. Wearable photoplethysmographic sensors—past and present. *Electronics*, 3(2):282–302, 2014.

- [48] Toshiyo Tamura, Yuka Maeda, Masaki Sekine, and Masaki Yoshida. Wearable photoplethysmographic sensors—past and present. *Electronics*, 3(2), April 23, 2014.
- [49] Edward Jay Wang, William Li, Doug Hawkins, Terry Gernsheimer, Colette Norby-Slycord, and Shwetak N. Patel. Hemaapp: Noninvasive blood screening of hemoglobin using smartphone cameras. *UBICOMP '16*, September 12, 2016.
- [50] Gaobo Zhang, Zhen Mei, Yuan Zhang, Xuesheng Ma, Benny Lo, Dongyi Chen, and Yuanting Zhang. A non-invasive blood glucose monitoring system based on smartphone ppg signal processing and machine learning. In *IEEE Transactions on Industrial Informatics*. IEEE, February 20, 2020.