TÉCNICO LISBOA

Real-time pulse rate variability via remote photoplethysmography

Pedro Rui Ramalho Constantino

Thesis to obtain the Master of Science Degree in

Biomedical Engineering

Supervisors: Prof. João Miguel Raposo Sanches
Dr. Miguel Constante

Examination Committee

Chairperson: Prof. Rita Homem de Gouveia Costanzo Nunes
Supervisor: Prof. João Miguel Raposo Sanches
Member of the Committee: Prof. Augusto Emanuel Abreu Esteves

September 2021

This work was created using LaTeX typesetting language
in the Overleaf environment (www.overleaf.com).

Acknowledgments

I would like to thank my parents, brother, grandparents and godfather.

I would also like to acknowledge my dissertation supervisors Prof. João Sanches and Dr. Miguel Constante for their insight and support.

Special thanks to Carolina and Lena!

Abstract

Remote medicine is an emerging and important field with the potential to improve patients health at the distance of a teleconsultation. Here, we propose a novel remote photoplethysmography algorithm suited to extract pulse rate variability in real-time from the face of a patient that is being recorded by a consumer-grade webcam. We first test the algorithm with the UBFC dataset, which contains both webcam videos of subjects's faces plus the corresponding blood volume pulse extracted via finger clip photoplethysmography. Average errors between pulse rate variability features extracted from finger photoplethysmography and webcam remote photoplethysmography are as folllows: 20,46 ms for SDNN, 39,21 ms for RMSSD, 16,89% for %LF (nu) and 1,09 for LF/HF ratio. Although this amount of error show that the proposed method is not ready for fine medical diagnostic, we suggest that webcam face videos are a promising source of remote medical data, with potential for patient classification and severity diagnosis. Furthermore we also validate the real-time algorithm with an experiment where both PPG and rPPG are recorded at the same time, from a bluetooth PPG sensor and a computer webcam, respectively, and in which we can see concordance. Because the autonomic nervous system plays a big role in regulating the human heart rate, a remote, real-time pulse rate variability sensor might be of extreme interest for telepsychology and telepsychiatry alike.

Keywords

Facial video; Remote photoplethysmography; Heart rate; Heart rate variability; Pulse rate variability.

Resumo

A medicina remota é um tópico emergente e importante com o potencial de melhorar a saúde de uma pessoa apenas à distância de uma teleconsulta. Neste trabalho, é proposto um novo algoritmo para extrair a varaibilidade do ritmo do pulso a partir de fotopletismografia remota em tempo real, a partir do vídeo de webcam da cara de um paciente. Em primeiro lugar testa-se o algoritmo na base de dados da UBFC, que contém vídeos da cara de sujeitos acompanhados do sinal de pulso obtido a partir de fotopletismografia de dedo. O erro médio entre os parâmetros extraídos a partir do sinal da webcam e do sinal do dedo são os seguintes: 20,46 ms para a SDNN, 39,21 ms para a RMSSD, 16,89% para a %LF (nu) and 1,09 para o rácio LF/HF. Embora estes números demonstrem que o método proposta não será capaz de produzir diagnósticos finos, sugerimos que vídeos de webcam são uma fonte de data biomédica promissora, com potencial para conseguir classificar pacientes e níveis de doença. Para além disso, também validamos o algoritmo em tempo real com uma experiência em que tanto o sinal fotopletismográfico do dedo como o sinal remoto da cara são obtidos ao mesmo tempo, de um sensor bluetooth e de uma webcam de um computar, respetivamente, e no qual se pode ver concordância entre sensores. Como o sistema nervoso autónomo desempenha um papel na regulação do ritmo cardíaco, um algoritmo remoto, que trabalha em tempo-real pode ser de interesse tanto para a telepsicologia, como para a telepsiquiatria.

Palavras Chave

Vídeo facial; Fotopletismografia remota; Ritmo cardíaco; Variabilidade do ritmo cardíaco; Variabilidade do ritmo do pulso.

Contents

List of Figures

List of Tables

Listings

Acronyms

ANS	Autonomic nervous system
AV	Atrioventricular
BVP	Blood volume pulse
ECG	Electrocardiogram
HR	Heart rate
HRV	Heart rate variability
HF	High frequency
LF	Low frquency
PNS	Parasympathetic nervous system
PPG	Photoplethysmography
PR	Pulse rate
PRV	Pulse rate variability
PPT	Pulse transit time
rPPG	remote Photoplethysmography
ROI	Region of interest
SA	Sinoatrial
SNR	Signal-to-Noise Ratio
SNS	Sympathetic nervous system
ULF	Ultra low frequency
VLF	Very low frequency

1

Introduction

Contents

1.1 Background

For several years physicians have monitored heart rhythms through auscultation and have noted that beat-to-beat times shift depending on age, illness and psychological state [1]. From Herophilos and Galen describing the pulse with the naked-eye to Ludwig (smoked drum kymograph) and Einthoven (galvanometer) measuring mechanical and electrical cardiac variables, monitoring the beating of the heart has seen a lot of technological and theoretical advances throughout history [2]. Some of the historical turning points along with introductory concepts follows.

In 1781, Galvani observed "animal electricity" through electric stimulation of frog nerves, setting the basis for biological electrophysiology experimentation [3]. Only many years later, in 1842, Mateucci demonstrated that electrical current accompanies every heart beat, also using a frog. In 1887, Waller publishes the first human Electrocardiogram (ECG). Using electrodes placed on the chest and back of a human, Waller demonstrates that electrical activity preceded ventricular contraction. [4] In 1901, Einthoven invents the string galvanometer, a device that could measure ECG signals accurately and, later in 1906, he proposes the clinical use of ECG patterns to detect arrhythmias [5].

Heart rate (HR) is defined as the number of heartbeats per minute and Heart rate variability (HRV) concerns the fluctuation in the time intervals between adjacent heartbeats [6]. In **Figure 1.1** we can see a short segment of an ECG illustrating five heart beats and the respective time difference between all consecutive pairs of beats. HR variation ensures optimal adaption to environmental challenges [7] and is a reflection of the many physiological factors modulating the normal rhythm of the heart, namely the coordination of autonomic, respiratory, circulatory, endocrine and mechanical influences over time [8].

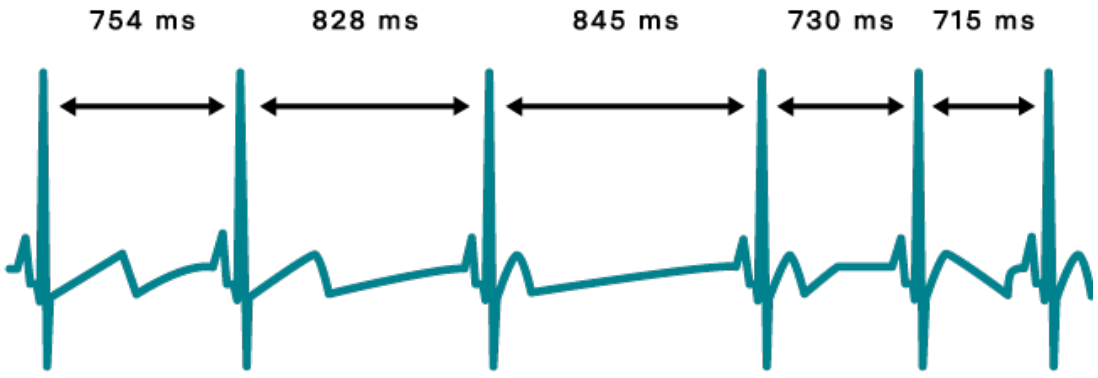

Figure 1.1: Five time differences (in milliseconds) between six heartbeats via electrocardiogram.

Later on, in 1937, Alrick Hertzman introduces Photoplethysmography (PPG), a method to detect the cardio-vascular pulse waves traveling through the body [9]. He proceeds with several publications examining the physiology and potential uses of the PPG waveform [10–13]. The PPG waveform, usually extracted with a pulse oximeter [14], and respective pulse time differences can be seen in **Figure 1.2**.

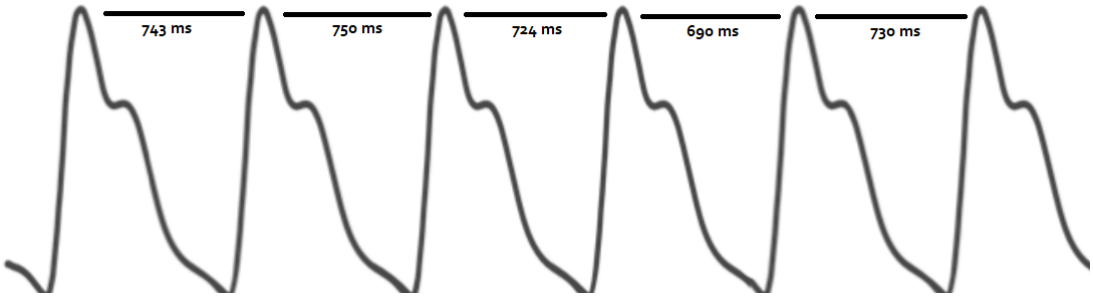

Figure 1.2: Five time differences (in milliseconds) between six pulses via photoplethysmography. This figure is not in scale.

Both ECG and PPG capture the rhythms of the heart, as both can identify the heart beat waveform: computing HR and HRV through ECG R-peaks is similar to calculating Pulse rate (PR) and Pulse rate variability (PRV) through PPG pulse peaks. Although some authors have observed that PRV is an accurate estimate of HRV for young, healthy subjects at rest [15], recent studies strongly suggest that PRV is new biomarker, which can be thought of HRV plus other biomedical information [16].

Although inter-beat time differences have traditionally been extracted from ECG or PPG, other technologies are available, such as smartphone optical pulse sensors, remote (or imaging) photopletysmography (*i.e.* photopletysmography through video camera), ultrasounds, microwave radar, cushion-mounted ballistocardiogram, even toilet seats [8].

In 2008, remote Photoplethysmography (rPPG) research oficially started with reports that a subject's face contains a signal sufficiently rich to measure the PR under ambient light, using only a camera and digital signal processing [17]. It was followed by a research group detecting photoplethysmographic signals using a standard laptop webcam, in 2010 [18]. In 2013, [19] innovated rPPG, introducing a chrominance-based method that used "skin-tone standardization" as a means to extract a more motion-robust signal, arguing that an optimal fixed combination of RGB channel signals can be found based on a ratio of normalized color signals when assuming "standardized" skin, thus eliminating noise derived from specular reflection. Later, in 2016, [20] introduces another chrominance-based method with even superior results.

HRV is popular in psychological science because it allows for lowcost, non-invasive, and accessible autonomic information [21]. Indeed, the ANS activity is seen as a major component of the emotional response in many theories of emotion [22]. Not only it has been reported that a lower HRV may be a predictor of depression [23, 24], but also that HRV is associated with general mental disorder [25], social cognition [26], executive function [27] and emotional regulation [28].

In this thesis, it is considered the case of psychology telemedicine consultations delivered via a video web platform. In this case, the input is video from a subject's face, from which it can be estimated the Blood volume pulse (BVP) signal and consequential PR and PRV measures.

1.2 Novelty

A real-time pipeline for PRV analysis from asynchronous webcam face video is proposed, along with a modification of a recent peak detection algorithm.

1.3 Thesis layout

This thesis is is organized as follows: **Chapter 2** provides an up to date literature review on HRV, PPG, PRV and rPPG principles and methods; **Chapter 3** lays out the digital signal processing pipeline proposed to accomplish real-time pulse rate variability measurements from a consumer-grade webcam; **Chapter 4** describes how the proposed rPPG method was tested both in offline and online contexts and discusses results and, lastly, **Chapter 5** draws final conclusions and proposes future work.

2

Literature review

Contents

This literature review aims at introducing the concepts underlying estimation of PRV from remote photoplethysmographic signals. It is divided in three sections. First, an introduction to HRV (2.1), where we review the physiological beating of the heart and how we can record its time-varying behaviour through the ECG to produce a collection of HRV measures for a range of applications. Next, a small section describing the principles of the photoplethysmography technique (2.2), which appears as an alternative to ECG for obtaining inter beat times. Some examples on how researchers experiment with pulse rate variability within psychology and psychiatry are provided in section (2.3). Finally, an up to date review on the rPPG technique (2.4)..

2.1 Heart rate variability

This section is organized as follows: first, the heart and its beating behaviour are described (2.1.1), followed by a review on the mechanisms behind regulation of heart rate (2.1.2); after that, a brief explanation on how heartbeat times are collected through the ECG (2.1.3); next, RR intervals are introduced as the central piece in HRV analysis (2.1.4), followed by a description of HRV metrics and norms (2.1.5) and finally, applications of HRV to pychological science (2.1.6) and depression (2.1.7) are presented.

2.1.1 Heart and heartbeat

For oxygen, nutrients and waste disposal, cells rely on surrounding interstitial fluid, which is kept stable through continuous exchange between the peripheral tissues and circulating blood. If blood stops moving, supplies of oxygen and nutrients will quickly exhaust and wastes are no longer absorbed. Cardiovascular functions ultimately depend on the heart. [29]

Muscular heart is composed of two atria and two ventricles. See **Figure 2.1**. Ventricles, the lower chambers, pump blood into lungs and arteries, while atria, the upper chambers, receive the returning venous blood. In a single heartbeat, atria and ventricles contract in a coordinated manner so that blood flows in the correct direction, at the correct time. A brief resting phase follows, allowing time for the chambers to relax and prepare for the next heartbeat. These phenomena are associated with the cardiac cycle, which consists of systole (ventricular contraction; blood pressure peaks) and diastole (ventricular relaxation; blood pressure reaches a minimum) [30].

Two types of cardiac muscle cells generate the heartbeat: (1) contractile cells produce strong contractions that cause the heart chamber to shrink and propel blood, and (2) specialized noncontractile muscle cells of the conducting system control modulate contractile cells. Contractile muscle cells, which comprise the majority of cardiac muscle cells, are activated by external action potentials, similarly to skeletal muscle. On the other side, noncontractile muscle cells are less in number and organized as a network made up of two types of cells: nodal cells and conducting cells. Nodal cells are *autorhythmic*,

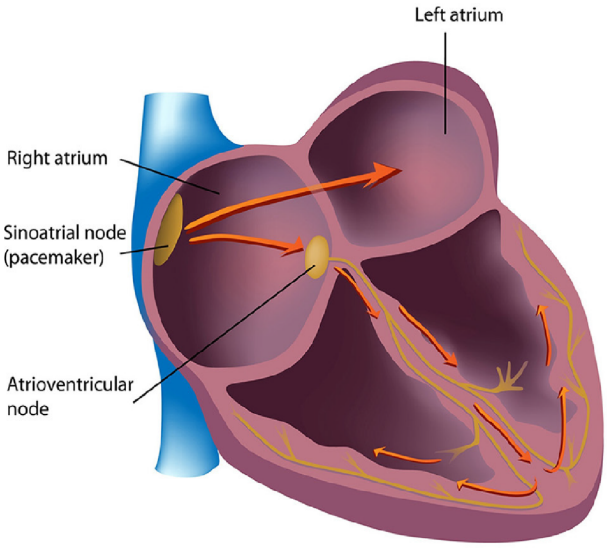

Figure 2.1: Basic muscular anatomy of the heart. *Credit: AlilaSaoMai/Shutterstock.com*

i.e. they contract on their own, without neural or hormonal stimulation, and generate the *pacemaker potentials* responsible for initiating the muscular heartbeat. They are located at the Sinoatrial (SA) and Atrioventricular (AV) nodes. However, nodal cells from the SA node naturally depolarize faster, 70–80 action potentials per minute, than those in the AV node, 40–60 action potentials per minute, being the effective *pacemaker cells* in the heart. In healthy hearts, the SA node generates an electrical impulse that travels through the atria to the AV node in about 0.03 seconds and causes the AV node to fire. Conducting cells interconnect these two nodes and distribute the contractile stimulus throughout the myocardium. Basically, once the stimulus for a contraction is generated at the SA node, it is distributed so that (1) the atria contract together, before the ventricles; and (2) only after, the ventricles contract together. [29, 31–33] See **Figure 2.3 a)**.

2.1.2 Regulation of the heart rate

Although the SA node spontaneously generates the normal heartbeat cardiac rhythm, autonomic motor neurons, circulating hormones and ions can influence the inter-beat interval and magnitude of the myocardial contraction [34]. More specifically, the cardiovascular center, located in the brain stem, integrates sensory information from various bodily receptors and responds through sympathetic and parasympathetic motor neurons (and endocrine systems), adjusting the HR continuously [35]. See **Figure 2.2**.

Cardiac sympathetic nerves target the SA node, AV node, and the bulk of the myocardium and trigger norepinephrine and epinephrine release and binding to beta-adrenergic (b1) receptors located on cardiac muscle fibers, speeding up spontaneous depolarization in the SA and AV nodes (increasing

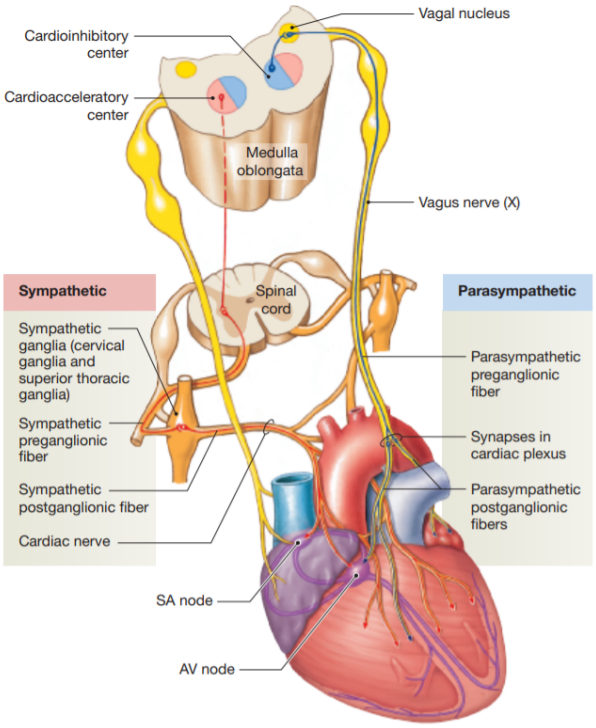

Figure 2.2: Autonomic innervation of the heart. Adapted from (Martini, 2018).

HR) [36]. The parasympathetic vagus (X) nerves also innervate the SA node, AV node, and atrial cardiac muscle and trigger acetylcholine release and binding to muscarinic receptors, decreasing the rate of spontaneous depolarization in the SA and AV nodes (slowing HR) [37]. In the end, there should be a continuous, dynamic balance between Sympathetic nervous system (SNS) and Parasympathetic nervous system (PNS) influences, resulting in some time-varying HR. It is observed that PNS control predominates at rest, resulting on an average of 75 bpm. Furthermore, while sudden changes in HR (up or down) between one beat and the next are parasympathetically mediated, an increase in sympathetic activity is the principal method used to increase HR above the intrinsic level generated by the SA node. PNS activation acts much faster ($< 1s$) on HR than SNS activation ($> 5s$). [38–41] HRV analysis is, thus, a widely accepted method for indirect evaluation of the Autonomic nervous system (ANS) activity [42, 43]

2.1.3 The electrocardiogram

The cardiac electrical phenomena is powerful enough to be detected by electrodes placed on the body surface. The ECG [44] can record the action of the cardiac muscle conduction system, as seen in **Figure 2.3** b). The main features of an ECG are: 1) the P wave, which indicates the depolarization of the atria, 2) the QRS complex, which appears as the ventricles depolarize (here, the electrical signal peaks in amplitude since the mass of the ventricular muscle is much larger than that of the atria) and 3) the T

wave, which accompanies the repolarization of the ventricles. [29]

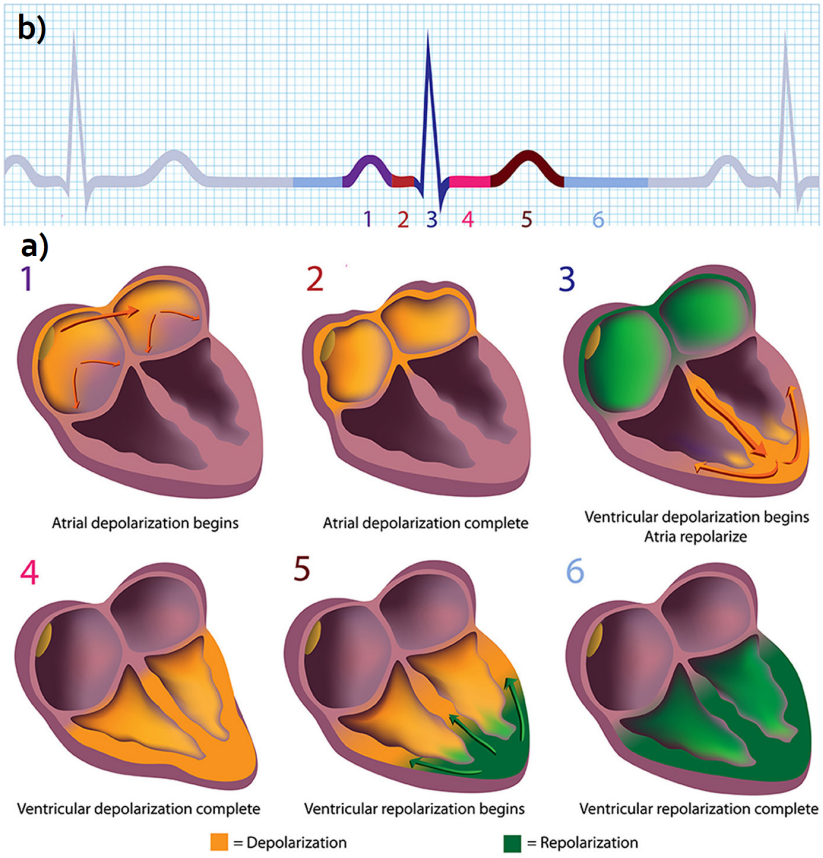

Figure 2.3: a) The depolarization and repolarization of the heart during a beating cycle. b) Associated electrical activity via electrocardiogram. 1 - P wave; 2 - PR segment; 3 - QRS complex; 4 - ST segment and 5 - T wave. *Credit: AlilaSaoMai/Shutterstock.com*

2.1.4 HRV principles

The triad P wave, QRS-complex and T wave and their relationship with heart's contraction can be fully visualized in **Figure 2.3**. HRV is usually measured through QRS-complex distances in the ECG signal [45]. RR intervals refer to the time interval between R-peaks in the QRS-complexes. Normal-to-normal (NN) interval time series contain only RR intervals resulting from sinus node depolarizations, *i.e.*, clean, true intervals [46]. In **Figure 1.1**, we can see five RR intervals computed from the R-peaks of the ECG. From a given sequence of RR intervals extracted for some period of time, it is possible to extract features and statistics that describe the modulation of cardiovascular function, *i.e.* HRV analysis.

2.1.5 HRV metrics and norms

Researchers usually report long-term ($1 - 24h$), short-term ($5min$) and ultra-short-term ($< 5min$) HRV readings, using time-domain, frequency-domain, and non-linear measurements [47]. Short-term and ultra-short-term values are not interchangeable with 24 h values, since longer recording epochs represent processes with slower fluctuations better (*e.g.* circadian rhythms) [48].

Time-domain metrics of HRV quantify the amount of variability in the measurement of RR intervals through statistics like average and standard deviation. Frequency-domain measurements estimate the power spectral density of a given RR signal into four frequency bands: Ultra low frequency (ULF), Very low frequency (VLF), Low frquency (LF), and High frequency (HF) bands. See **Figure 2.4** for a thorough list of HRV metrics organized per analysis domain. Non-linear HRV metrics are not considered in this thesis.

TABLE 1 | HRV time-domain measures.

Parameter	Unit	Description
SDNN	ms	Standard deviation of NN intervals
SDRR	ms	Standard deviation of RR intervals
SDANN	ms	Standard deviation of the average NN intervals for each 5 min segment of a 24 h HRV recording
SDNN index (SDNNI)	ms	Mean of the standard deviations of all the NN intervals for each 5 min segment of a 24 h HRV recording
pNN50	%	Percentage of successive RR intervals that differ by more than 50 ms
HR Max – HR Min	bpm	Average difference between the highest and lowest heart rates during each respiratory cycle
RMSSD	ms	Root mean square of successive RR interval differences
HRV triangular index		Integral of the density of the RR interval histogram divided by its height
TINN	ms	Baseline width of the RR interval histogram

TABLE 2 | HRV frequency-domain measures.

Parameter	Unit	Description
ULF power	ms^2	Absolute power of the ultra-low-frequency band (≤0.003 Hz)
VLF power	ms^2	Absolute power of the very-low-frequency band (0.0033–0.04 Hz)
LF peak	Hz	Peak frequency of the low-frequency band (0.04–0.15 Hz)
LF power	ms^2	Absolute power of the low-frequency band (0.04–0.15 Hz)
LF power	nu	Relative power of the low-frequency band (0.04–0.15 Hz) in normal units
LF power	%	Relative power of the low-frequency band (0.04–0.15 Hz)
HF peak	Hz	Peak frequency of the high-frequency band (0.15–0.4 Hz)
HF power	ms^2	Absolute power of the high-frequency band (0.15–0.4 Hz)
HF power	nu	Relative power of the high-frequency band (0.15–0.4 Hz) in normal units
HF power	%	Relative power of the high-frequency band (0.15–0.4 Hz)
LF/HF	%	Ratio of LF-to-HF power

Figure 2.4: Heart rate variability time-domain and frequency-domain measures. Excluding non-linear measures. Adapted from (Shaffer, 2017).

Regarding time-domain measures, the most common used are standard deviation of NN Intervals (SDNN) and root means successive square difference (rMSSD) [49]. They correspond to short-term HRV changes and are not dependent on day/night variations [45]. These are the time-domain measures considered in this thesis.

Regarding frequency-domain measures. They can be expressed in absolute or relative power. Absolute power is calculated as ms^2 divided by cycles per second (ms^2/Hz). Relative power is estimated as the percentage of total HRV power or in normal units (nu) - it divides the absolute power for a specific frequency band by the summed absolute power of the LF and HF bands. [48] Relative power allows direct comparison between two patients, despite wide variation in specific band power and total power among healthy, age-matched individuals [47]. The ULF band (≤0.003 Hz) indexes fluctuations in RR intervals with a period from 5 min to 24 h and can only be measured using 24 h recordings. In **Fig-**

ure 2.5 we can see how long-term HRV spectral estimates look like. The frequencies in a 5-min long recording can traditionally be divided into other three main bands [50]. The VLF band (0.0033–0.04 Hz) is comprised of rhythms with periods between 25 and 300 s. The LF band (0.04–0.15 Hz) is composed of rhythms with periods between 7 and 25 s and is affected by breathing from 3 to 9 bpm. The HF or respiratory band (0.15–0.40 Hz) is influenced by breathing from 9 to 24 bpm (see respiratory sinus arrhythmia (RSA)). [45,51] It can be said that LF/HF estimates the ratio between the activity of SNS and PNS. Total power is the sum of the energy in the ULF, VLF, LF, and HF bands for 24 h and the VLF, LF, and HF bands for short-term recordings [31].

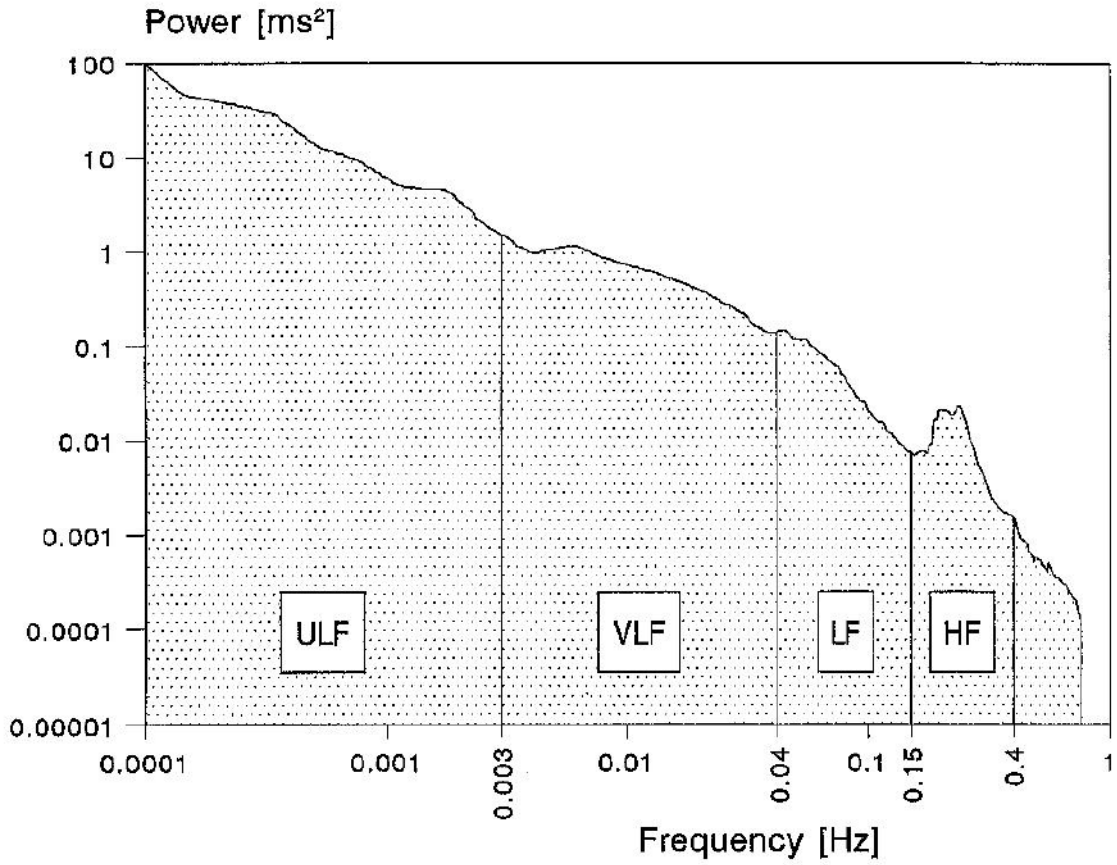

Figure 2.5: Example of an estimate of power spectral density obtained from the entire 24-hour interval of a Holter recording. Adapted from (Task Force, 1997).

Here, we do not account for long-term HRV assessment ($24h$), neither the ULF band in frequency-domain HRV analysis since we are considering telemedicine consultations, which could last a maximum of 1 hour. Furthermore, the case considered in this thesis concerns a patient resting in front of a computer webcam. **Figure 2.6** illustrates a short-term HRV spectrum for a subject at rest. To minimize nonstationarities, 1-min recordings are recommended to assess the HF, while 2-min recordings are recommended to assess the LF [1].

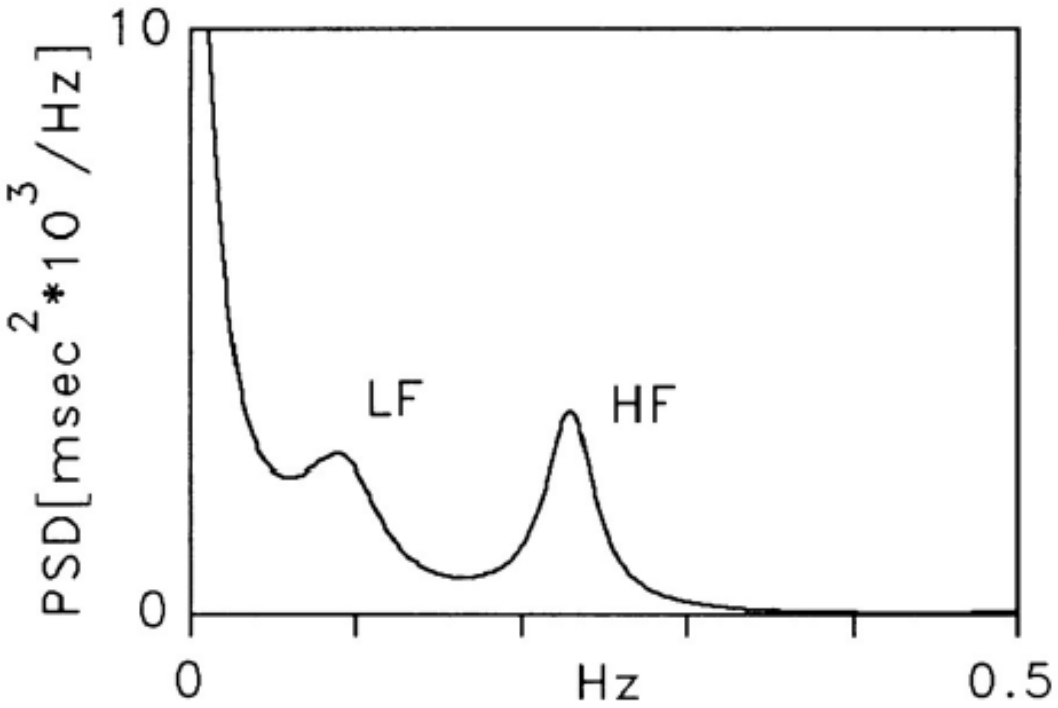

Figure 2.6: Example of a short-term estimate of power spectral density obtained from a patient at rest. Adapted from (Task Force, 1997).

2.1.6 HRV and psychophysiological theory

According to [31], five theories imply HRV in psychophysiological research: the neurovisceral integration model [52], the polyvagal theory [53], the biological behavioral model [54], the resonance frequency model [55], and the psychophysiological coherence model [56]. All five theories focus on vagal tone (parasympathetic modulation). Specifically. In the neurovisceral integration model, it is assumed that the prefrontal cortex is connected to the heart through the central autonomic network and vagus nerve and that higher vagal tone is correlated with better emotional and health regulation and also better executive cognitive performance [52, 57]. In polyvagal theory, the vagus complex functions as a *brake*, as it continuously inhibits the influence of SNS on HR (also reduces the stress response), allowing states of calmness necessary for social behavior. Here, a higher vagal tone is associated to better social functioning. [53], In the biological behavioral model, vagal tone plays a role in regulation of energy exchange by synchronizing respiratory and cardiovascular processes following metabolic and behavioral dynamics. In this model, a higher resting vagal tone means higher adaptability, as this is conceived as a functional energy reserve capacity from which the organism can draw during more active states [54]. In [55], authors suggest that slow paced breathing at the resonance frequency increases vagal tone. Similarly, in the psychophysiological coherence model, a higher vagal tone can be achieved through slow paced breathing [6].

2.1.7 Applications of heart rate variability to depression

In the past, whether HRV could be used as a biomarker for depression was unclear, with research showing mixed findings [58]. While some authors have reported depressed patients to have lower levels of HRV than controls [59, 60], others reported no differences [61–63].

Recent meta-analysis seem to agree that HRV measures correlate with depression to some extent and attribute past mixed results to heterogeneity in relatively small samples, medication confounds and reporting of different HRV measures. In [64], from 13 cross-sectional studies, a small reduction of HRV in depression is reported, though only HF power and RMSSD measures were considered. It should be noted that RMSSD estimates contain a contribution of lower frequency ($< 0.15Hz$), capturing a mix of sympathetic and vagal (parasympathetic) influences [65]. Later, [66] finds compelling evidence to believe that HRV reductions are inversely correlated with depression severity, from a meta-analysis of 18 studies with patients with no cardiovascular complications. Drawing from 6 studies based on children and adolescents, [67] concluded that adolescents who met clinical criteria for depressive disorders exhibited lower resting state HF, which is consistent with research among adults. However, the authors found no no correlation between HF and depressive symptom severity, but were limited to non-clinical samples only. Recently, from 21 studies, [68] comprehensively analyses effect sizes for seven distinct resting-state HRV measures, namely HF, LF, VLF, LF/HF, RMSSD, SDNN and RR intervals. Results suggest that patients with depression are likely to display small reductions in HF, LF, VLF, RMSSD (largest effect size), SDNN and RR intervals and an increase in LF/HF ratio. They conclude that depressive states are not associated with alteration in specific indicators of HRV, but rather with abnormalities in several time- and frequency-domain measures. Moreover, this meta-analysis only considered unmedicated patients, demonstrating that reductions in HRV are prevalent in depressed patients without antidepressants.

In [69], authors found correlation between 15-min ECG-based HRV measures and clinical state of depression. Using 62 depressed patients and 65 non-depressed controls and testing several HRV measures, they have seen a decreased HF%, amongs others, to be highly correlated with a diagnostic of depression. Unfortunately, in this study, RMSSD is not properly acessed.

2.2 Photoplethysmography

This section is divided in two parts: first we introduce the principles behind the PPG technique (2.2.1) and, secondly, we review the relationships and differences between HRV and PRV (2.2.2).

2.2.1 Photoplethysmography principles

For each heart beat, blood is pumped through the circulatory system, from the heart to peripheral tissues, causing a displacement in volume along the way. PPG can be used to detect this pulsatile behaviour [70,71]. Hertzman first introduced the term “photoplethysmography” (”photo” from photon, which is the basis of the optical data analyzed + “plethysmo”, Greek for “enlargement”, referring to the changes in the dermal vasculature) [72]. This technique requires only a light source to illuminate the tissue and a photodetector to measure the sum total of volume changes in any and all blood vessels (e.g., large and small arteries, arterioles, venules, and veins) [73]. The photoplethysmogram is both non-invasive and similar to an arterial blood pressure waveform - which is obtained through an invasive technique. Because of this, the extraction of circulatory data from the photoplethysmogram has been a popular subject of contemporary research. [74]

Although the interaction of light with biological tissue is complex and includes the optical processes of scattering, absorption, reflection, transmission and fluorescence, in general, the greater the blood volume the more the light source is attenuated [75]. Given that, there are two main PPG configurations [76]: 1) in the transmission mode, a tissue sample is placed between the light source and the photo detector, and thus, the measurement site may be limited to tissues where transmitted light can be readily detected (*e.g.* fingertips, earlobes) [77], while 2) in the refection mode, the source and detector are placed side by side, which enables measurements from virtually any point on the skin surface. For a review on optical sensor technology for PPG, see [78]. It matters to add that compared to red/infra red light, green light has greater absorptivity for both oxyhemoglobin and deoxyhaemoglobin, resulting in a better Signal-to-Noise Ratio (SNR) for PPG [79–81].

PPG has many biomedical applications, including: (a) clinical physiological monitoring (blood oxygen saturation, HR, blood pressure, cardiac output and respiration); (b) vascular assessment (arterial disease, arterial compliance and ageing, endothelial function, venous assessment, vasospastic conditions, microvascular blood flowand tissue viability), and (c) autonomic function (vasomotor function and thermoregulation, HRV, orthostatic intolerance and neurology). See [73].

Two limitations associated with PPG are: (a) contact measurement, which means the sensor must be attached firmly to the skin, causing discomfort and impairing movement (*e.g.* sports) and (b) spot measurement, which means that traditional PPG can only extract pulse from one skin site per probe (more probes mean more restriction and discomfort) [76].

2.2.2 HRV vs. PRV

In contrast to the QRS-complex, a pulse wave can be subdivided only into two parts [15]. See **Figure 2.7**. In the anacrotic phase, the pulse rises due to systole-induced blood pressure increase. In the catacrotic phase, which is more prolonged than the anacrotic phase, there is a subsequent decline corresponding to cardiac diastole (it also can contain a secondary peak, called the dicrotic notch, due the closure of the aortic valve [82]). There is some inconclusiveness to the question whether PRV is a substitute for HRV. For each RR interval extracted from the ECG, there is an interval comprising a full pulse cycle length in the PPG, which we call PP interval (see 2.7). Depending on pulse wave velocity and vascular path from the heart to the PPG sensor location, there is a delay between each R peak and the onset of its corresponding pulse wave: this delay is termed Pulse transit time (PPT) and is negatively correlated with blood pressure, arterial stiffness, and age [83]. Deviations between RR intervals and PP intervals can arise in two ways [41]: 1) artifacts and/or noise, or 2) physiological variability in PPT. Furthermore, it has been observed that correlation between PRV and HRV frequency domain variables deteriorates with diminishing sampling rate (noticeably when below 100 Hz). [84].

The difference in name is not simply a matter of signal source (ECG or PPG), but a matter of cardiovascular physiology [85].

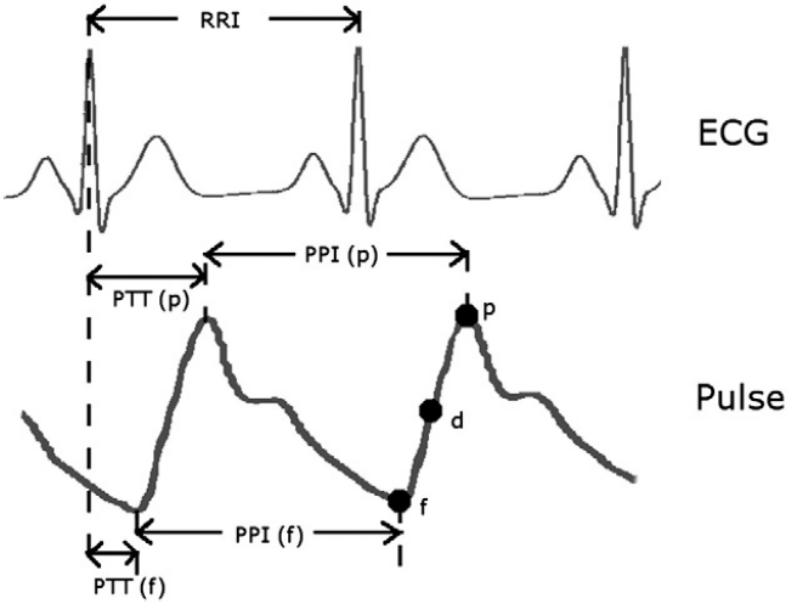

Figure 2.7: Electrocardiogram and a pulse wave signal, illustrating the meaning of RR interval (RRI), pulse transit time (PTT), and pulse interval (PPI). Adapted from (Schäfer, 2013).

Recently, some authors have proposed PRV to be an independent biomarker arguing that, although HRV strongly modulates PRV, PRV is further modulated by many other different sources and factors, which could contain useful biomedical information that is neither error nor noise [16]. See **Figure 2.8**.

Physiological measure	Anatomical location	Conversion of physical property	Modulators	
			Direct	Indirect
ECG R wave ↓	Left ventricular muscle	Electric excitation ↓	Intraventricular conduction, ventricular activation time, electromechanical coupling	Myocardial ischemia, heart diseases
Pre-ejection period ↓	Left ventricle	Muscle force ↓	Preload and after-load, contractility, aortic diastolic pressure	Respiration, blood pressure, body position and exercise, heart failure (alternating pulse)
Aortic pressure elevation ↓	Aorta	Pressure impulse ↓	Stroke volume, aortic dynamic compliance, intrathoracic pressure	Respiration, peripheral resistance
Pulse conduction time ↓	Artery	Pressure wave ↓	Internal radius, wall thickness and elasticity, blood density	Vasomotor sympathetic activity, endothelial function, blood pressure
Tissue volume ↓	Tissue microvasculature	Blood volume ↓	Vascular dynamic compliance, blood flow, venous pressure	Location and body position, body and environmental temperature
PPG pulse wave	Red cell hemoglobin	Light intensity	Absorption, scattering, reflection, and transmission; vascular bed volume	Local red cell count, hemoglobin content, waveform fiducial point

Figure 2.8: Factors affecting the information transmission from electrocardiogram R wave to photoplethysmography pulse wave. Adapted from (Yuda, 2020).

2.3 Pulse rate variability applications in mental health

When applied to psychiatry and psychology, pulse rate variability research targets topics like stress (2.3.1), emotion (2.3.2), among others (2.3.3).

2.3.1 Stress

As early as 2011, [86] suggests that PRV could technically be used as a surrogate of HRV for the purpose of mental stress assessment, although the authors do not test performance with stress levels reported by subjects. More recently, using an elastic net linear regression model, [87] showed correlation between self-reported stress levels and PRV features including SDNN, RMSSD, LF/HF ratio, among others. In this case, the study design proposes to test different groups of people at different experimental stages to arrive to a more robust result. Also in 2018, [88] reported 85.35% accuracy predicting the stress levels reported from subjects during a given moment of the day. For that, it was requested from subjects that they acquire 30-second PPG records three times every day and respective self-reported stress level. These were also learned with a linear regression model.

2.3.2 Emotion

In 2016, [89] reported correlation between video-elicited emotions (happy, fear, sad, angry, surprise, disgust) and both HRV via ECG) and PRV (via finger PPG) features, such as mean RR interval. The

authors suggest, for example, that sad emotion are associated with higher mean of heart rate activity, i.e., smaller mean of RR intervals. In another instance, also using an emotion excitation videos database, [90] accomplishes 60% accuracy classifying three emotions (disgust, anger and happiness) with a support vector machine modelling SDNN, RMSSD, HF and TINN. Very recently, [91] showed that a different PRV pattern appears when a subject gets angry while playing a video game. Specifically, the authors suggest that anger is associated with a higher pulse rate, given a certain baseline.

2.3.3 Other topics

One group reported PRV time-domain features to be particularly low for subjects with a psychosis diagnostic, when compared to depressive and healthy controls [92]. In particular, both SDNN and RMSSD where lower for psychotic patients, irrespective of age, body mass index, smoking, and medication.

Differently from previous applications presented before, which used a traditional PPG sensor, in 2020, [93] uses a smartphone camera at the fingertip of the user in order to predict personality based on PRV features, bringing rPPG and psychology together. The authors suggest that stable, extroverted subjects have higher In this study, personality is measured through a questionnaire-like tool that subjects should complete themselves. Somewhere else [94], the same authors have reported that subjects with higher life satisfaction scores have a higher rMSSD, SDNN, log HF and log LF; that participants who suffer from depression have lower log HF and log LF (this is probably due to medication) and that participants who tend to high-anxiety have lower pNN50 and log HF.

2.4 Remote photoplethysmography

In 2018, rPPG was reportedly [95] the most popular name for a technique that can also be referred to as contactless PPG, camera-based PPG or imaging PPG. Its key component is the camera (e.g. low-cost webcam, mobile phone camera, high-resolution camera), which collects the reflected (or transmitted) photons from the skin [96], resembling the PPG principle reviewed in section (2.2.1).

In this section, it presented the basic model of skin reflection underlying the rPPG algorithm used (2.4.1) and a review of the building blocks of a working rPPG sensor (2.4.2).

2.4.1 Skin reflection model

Although two theories exist to explain the pulsating character of the remote PPG signal, namely blood volume effects [97] and ballistocardiographic effects [98,99], in this thesis we consider only blood volume effects. For that, we introduce a generalized skin reflection model proposed by [20] from which rPPG algorithms can be derived.

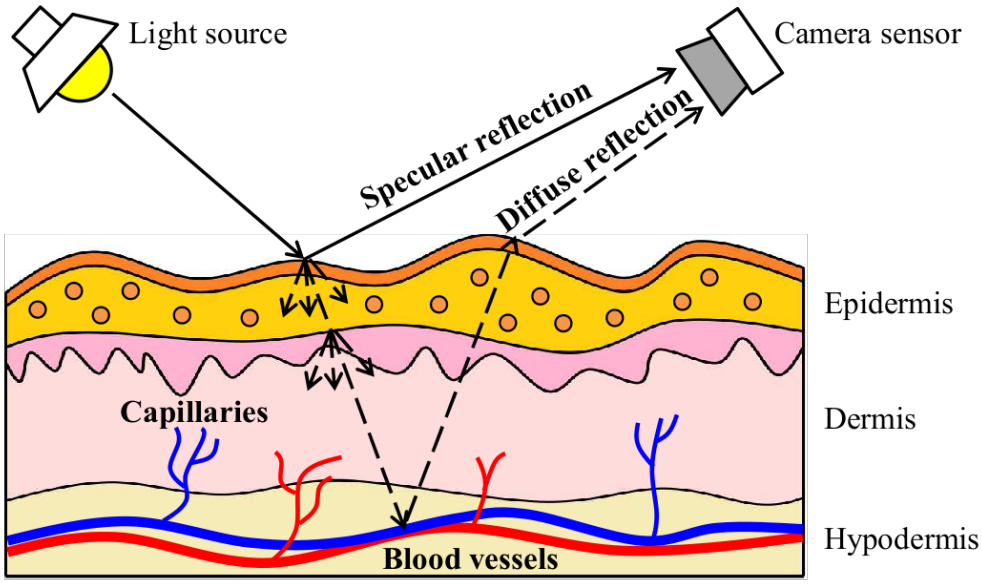

Figure 2.9: Specular and diffuse reflections considered in remote photoplethysmography model. Adapted from (Wang, 2017).

The model considers that there is some light source illuminating a piece of skin tissue, which is being recorded by a remote color camera. See **Figure** 2.9. Furthermore, the light source has a constant spectral composition with varying intensity and the intensity recorded at the camera depends on the distances between light source, skin and camera. According to the diachromatic model, the following time varying function defines the reflection of skin pixel k, at time t:

$$\mathbf{C_k}(t) = I(t) \cdot (\mathbf{v_s}(t) + \mathbf{v_d}(t)) + \mathbf{v_n}(t) \tag{2.1}$$

where $C_k(t)$ denotes the RGB channels (in column); $I(t)$ denotes the luminance intensity level, which absorbs the intensity changes due to the light source as well as to the distance changes between light source, skin tissue and camera; $I(t)$ is modulated by specular reflection $v_s v$ and diffuse reflection $v_d(t)$; the last component $v_n(t)$ denotes the quantization noise of the camera sensor.

Specular reflection is a mirror-like light reflection from the skin surface with no pulsatile information and we can assume that its spectral composition is similar to the light source:

$$\mathbf{v_s}(t) = \mathbf{u_s} \cdot (s_0 + s(t)) \tag{2.2}$$

where u_s denotes the unit color vector of the light spectrum and s_0, $s(t)$ denote the stationary and varying parts of specular reflection.

Diffuse reflection is associated with absorption and scattering of light in skin tissue, where the combination of hemoglobin and melanin lead to a specific chromaticity:

$$\mathbf{v_d}(t) = \mathbf{u_d} \cdot d_0 + \mathbf{u_p} \cdot p(t) \tag{2.3}$$

where u_d denotes the skin's unit color vector, d_0 denotes the stationary reflection strength, u_p denotes the relative pulsatile strengths in RGB channels and $p(t)$ denotes the desired pulse signal.

With (2.2) and (2.3), (2.1) becomes

$$\mathbf{C_k}(t) = I(t) \cdot (\mathbf{u_s} \cdot (s_0 + s(t)) + \mathbf{u_d} \cdot d_0 + \mathbf{u_p} \cdot p(t)) + \mathbf{v_n}(t) \tag{2.4}$$

Finally, if we combine the stationary parts of specular and diffuse reflection into one:

$$\mathbf{u_c} \cdot c_0 = \mathbf{u_s} \cdot s_0 + \mathbf{u_d} \cdot d_0 \tag{2.5}$$

(2.4) becomes:

$$\mathbf{C_k}(t) = I(t) \cdot (1 + i(t)) \cdot (\mathbf{u_c} \cdot c_0 + \mathbf{u_s} \cdot s(t) + \mathbf{u_p} \cdot p(t)) + \mathbf{v_n}(t) \tag{2.6}$$

where $I(t)$ is expressed as the combination of stationary and time varying part and $i(t)$, $s(t)$, $p(t)$ are zero-mean signals.

In terms of (2.6), we can define the problem of rPPG as extracting $p(t)$ from observations $C_k(t)$.

2.4.2 Remote photoplethysmography building blocks

According to [96], the rPPG method can be thought of five steps: recording, image processing, color channel combination, signal processing and analysis.

Recording

The first step involved in remote PPG is to acquire a sequence of video frames capturing human skin tissue (*e.g.* in a webcam call, skin from the human face is available to perform rPPG). See **Figure 2.10**.

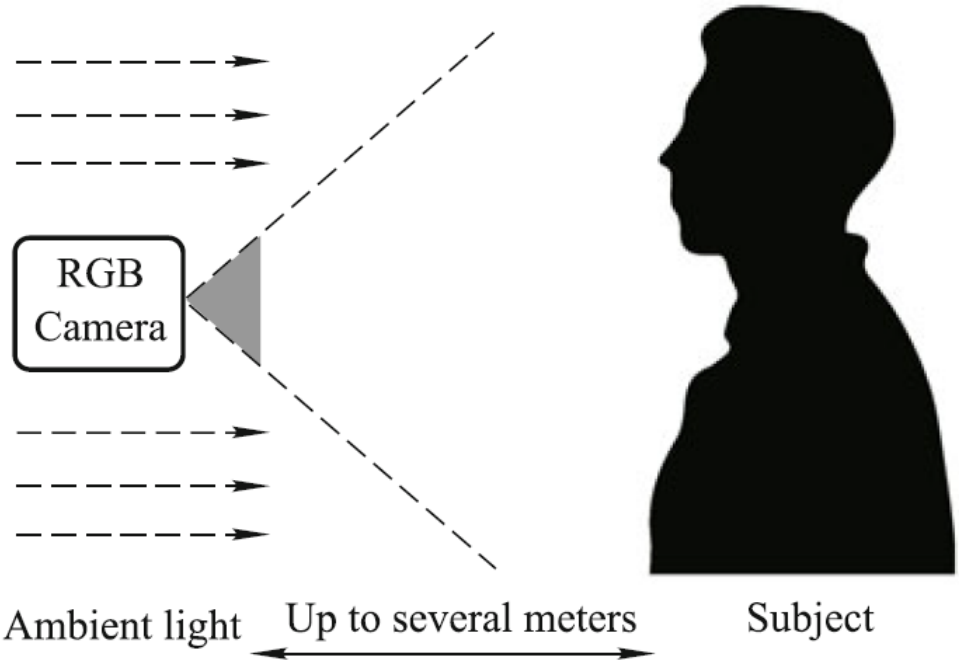

Figure 2.10: Typical setup of a remote photoplethysmography application. Adapted from (Rouast).

Most rPPG research focus on low-cost, consumer-grade, RGB video recorders like webcams, with video resolutions of 640 $\times$ 480 pixels, sampling frequencies of 30 fps or below and handles videos which are stored offline in an uncompressed format [96]. It was suggested that physiological information can be extracted despite compression, though the signal quality is degraded [100]. Recordings are commonly made under laboratory conditions with the subject at rest (or during predefined movements) [95, 96]. Most researchers make use of a single camera, although using multiple cameras can be an advantage in the case of subject motion [101, 102]. Although homogeneous illumination should be the optimal type of illumination for the rPPG technique [103], most works focus on artificial light, ambient light or a combination of both. It is also known that polarized illumination and polarization filters can help to reduce artifacts [104, 105].

Image processing

Once we have a sequence of video frames containing skin, Region of interest (ROI) selection must be performed in order to segment skin pixels from non-skin pixels. In this thesis, only rPPG algorithms based on the human face will be considered, as most works use the subject's face as general ROI. Viola-Jones face detector [106] is a common choice for this. After analysing the spatial distribution of the rPPG signal in the human face, authors showed that the forehead and cheeks provide best results in terms of SNR [107–110]. To go from detected face to segmented skin is a challenge for rPPG. Facial landmarks algorithms are an attractive solution to provide detailed ROI borders and are the solution considered in this thesis. Active appearance model is a statistical model of the human face in which appearance is matched to a given video frame, resulting in a set of coordinates of known facial landmarks [111], which are further exploited to create detailed segmentation of the facial skin. Other approaches similar to this can be found in [108, 112, 113]. An alternative to face detection is applying a skin detection algorithm directly at each video frame [114], though the authors have admitted this method to underperform in cases that there are non-skin regions in the frame which are of similar color to facial skin. A straightforward method to achieve ROI tracking is to simply re-detect the ROI for every frame [95]. Finally, after knowing the coordinates of all skin pixels in some frame, we can extract the raw color signal. This is generally achieved through spatial pooling, which consists in averaging the respective color channel of all pixels contained in the skin ROI of the frame [95]. This approach is known to dilute single-pixel camera noise and is the raw signal extraction method considered in this thesis.

Color channel combination

This step of rPPG aims at transforming the raw color signal extracted previously into a BVP signal that can be further processed to extract measures of interest. This is the second step. Both using the green channel alone or using a combination of the three color signals is common among rPPG work. Here, we consider a combination of color channels. There are two main approaches to combine the three RGB color channels [95, 96]: 1) knowledge-based approaches, which use *a priori* knowledge about how the pulse is pronounced in different color spaces and projection spaces and 2) source separation approaches, which typically use a blind source separation algorithm. In this thesis, we only consider knowledge-based approaches. The chrominance-based method is the most popular knowledge-based approach for BVP extraction. In [19], the authors introduce the first popular chrominance-based method, which uses an empirical weighted combination of R, G, and B channels to yield BVP. The authors of [115, 116] used the u* channel of International Commission on Illumination (CIE) L*u*v* color space as the input signal. Elsewhere it was used chromaticity from the CIE L*a*b* color space [117]. The approach considered in this thesis, called Plane-Orthogonal-to-Skin (POS) method, was proposed by [20]. Simply put, this method first projects RGB channels into the plane orthogonal to skin tone, resulting in

two novel channels, which are then fused by weighting to the desired signal, yielding the one-dimensional BVP signal.

Signal processing

At this stage, the noisy extracted BVP signal has to be filtered. Bandpass filters that cover the expected normal values of heart rate, $0.7 - 4Hz$, are the popular choice for filtering (*e.g.* [113, 118, 119]). Other techniques include: amplitude selective filter [120], continuous wavelet transform filter [115, 121], adaptive filter via CWT [122] or fourier transform [123], emprical mode decomposition [124] and even the Kalman filter [125]. PRV pipelines which aim at beat detection should carefully define the bandpass filter, as PRV analysis suffer from the loss of morphology and temporal information by filtering [126]. In the thesis, we consider the simple bandpass filter.

Analysis

The last step in the rPPG pipeline is to conduct analysis on the extracted BVP signal, which can include estimation of PR, PRV, pulse transit time, pulse wave velocity, oxygen saturation and assessment of vascular state [96]. In this thesis we consider PR and PRV.

Two popular methods for estimating PR from the BVP signal are [95]: 1) frequency analysis and 2) peak detection. Frequency-analysis methods include fast fourier transform [127–129], the most popular, discrete cosine transform (DCT) [130, 131], Welch's method [112, 132], short-time fourier transform (STFT) [133] and autoregressive models [134–136], among others. On the other hand, peak detection methods allow not only for PR estimation (inverting the mean of PP intervals from a certain period yields) but also for PRV analysis. In this thesis we consider peak detection.

3

Methods

Contents

The proposed method takes as input a timestamped stream of video, *i.e.* a sequence of tuples of the form $(frame[i], timestamp[i])$, where i is the counter of video frames captured, $i = 1, 2, 3, ...$, $frame[i]$ is a $(width, height, 3)$ matrix storing each pixel's RGB intensities and $timestamp[i]$ is the time at which the corresponding $frame[i]$ was captured. It is assumed that all frames of the video contain a face. We can see a scheme of the full method in **Figure 3.1**.

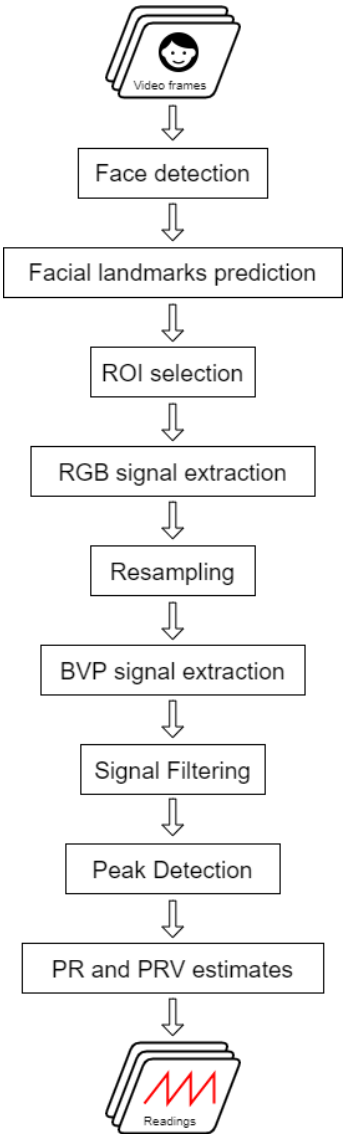

Figure 3.1: Proposed remote photoplethysmography pipeline to extract pulse rate variability in real time.

First, for every video frame, $frame[i]$, an average of the RGB channels over a predicted facial skin ROI, $avg\_rgb[i]$, is produced. This can be accomplished with face detection (3.1), facial landmarks prediction (3.2), ROI selection (3.3) and RGB signal extraction (3.4). Apart from the initialization of the face and landmarks detection algorithms, no additional memory is needed for these steps.

Secondly, we must estimate a BVP signal, $bvp[t]$, from the collected raw RGB signal, $avg\_rgb[i]$. Since following algorithms assume that samples are evenly distributed in time, we must first resample the asynchronous raw RGB signal (3.5). After that, we apply the POS method to transform the skin RGB signal into a BVP signal (3.6). A bandpass filter is further applied to clean the signal for peak detection. Here, array memory is allocated in advance to store the resampled RGB signal, the POS method signal and the filtered BVP signal.

Lastly, PR and its variability are estimated via the filtered BVP signal. Peak detection is firstly applied to the BVP signal in order to detect true heartbeats and, from the peaks, we can compute PP intervals (3.7). From the PP intervals, we can provide estimations of PR and time-domain measures of PRV (3.8). Memory is needed to store detected peaks and PP intervals.

3.1 Face detection

For face detection, *dlib* machine learning library is used [137]. According to *dlib*'s example documentation[1], the face detector is composed of an Histogram of Oriented Gradients, a linear classifier, an image pyramid, and a sliding window detection scheme. This process outputs a rectangle for each detected face in the form $(x_{ul}, y_{ul}, width, hight)$, where the subscript ul refers to the upper left corner of the rectangle enclosing the detected face. See **Figure 3.2**. Here, we assume that only one face is detected and that, for each frame, only one rectangle is output, $rect[i]$. Note that, to process and extract PRV from n faces would require n times the computational power it takes for one (apart from the face detection task which is common to all faces).

Input: $frame[i]$

Output: $rect[i]$

3.2 Facial landmarks prediction

dlib was also used for facial landmarks prediction [137]. It implements an Ensemble of Regression Trees method proposed by [138]. Here, we use a pre-trained model file[2] made available by *dlib*, which was trained on the dataset created by [139]. It receives as input a frame and a rectangle corresponding to the detected face area and outputs $shape[i]$, which, in this case, is a 68-point list containing the pixel coordinates of the 68 facial landmarks modeled by this machine learning instance. In **Figure 3.3** we can visualize the landmark detection and also that some of the fiducial points land outside the face, at non-skin regions.

Input: $frame[i], rect[i]$

[1] http://dlib.net/face_landmark_detection_ex.cpp.html, visited September 21, 2021
[2] http://dlib.net/files/shape_predictor_68_face_landmarks.dat.bz2, visited September 21, 2021

Figure 3.2: Example of face detection.

Output: $shape[i]$

Figure 3.3: Dlib's 68-point facial landmark detection.

3.3 ROI selection

In essence, selecting the ROI consists in, based on the predicted facial landmarks, drawing contours for relevant facial regions and make a decision on which combination of regions should be the target of analysis. Here, the regions comprising the cheeks, the nose and the forehead are considered, while the regions for beard and eyes are removed. Because the 68-point shape model doesn't predict the forehead contour, four forehead points are estimated from the position of the nose and eyebrows. Selection of ROI is ultimately a crucial step for PRV analysis, as noise introduced by sampling non-skin areas strongly

damage the very fine rPPG signal. Since estimated landmarks on the boundary between skin and non-skin areas may land at non-skin areas, even more when the subject moves, a correction is made by modifying x or y coordinates of landmarks. This corrective approach towards *dlib*'s facial prediction has been done elsewhere [140] and can be visualized in **Figure 3.4**. The output, $mask[i]$, consists of a $(width, height)$ binary matrix (*i.e.* it only contains zeros or ones), with ones where there should be skin and zeros where it shouldn't.

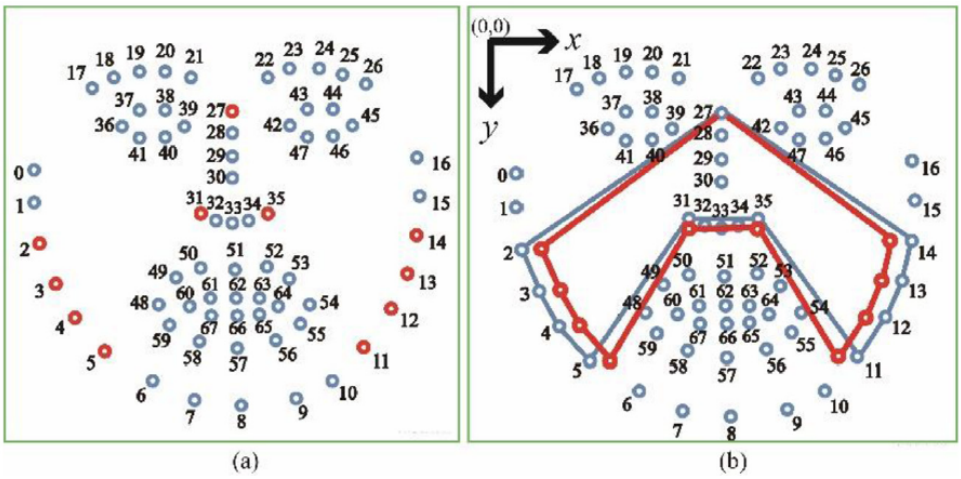

Figure 3.4: Lower head region of interest selection. (a) 68 dlib landmarks. (b) Landmark correction (seen as red). Adapted from (Ryu, 2021).

In the end, after removing the ocular regions and accounting for landmark position correction, we have the skin mask ready, as seen in **Figure 3.5**. Note that ROI tracking is implicitly done so, as ROI positions are predicted anew based on facial landmarks detected for the current frame (not previous frames). In the case a face is not detected, the last ROI position is used until a face is recognized again.

Input: $shape[i]$

Output: $mask[i]$

Figure 3.5: Selected region of interest comprising cheeks, forehead, nose and excluding ocular regions.

3.4 RGB signal extraction

Finally, the mask is applied to the frame via a bitwise operation, creating $masked\_frame[i]$, which is black background plus segmented skin when displayed. $avg\_rgb[i]$ is a $(3, 1)$ vector containing the average red, green and blue intesities over the ROI, $(\bar{r}, \bar{g}, \bar{b})$. The value of $avg\_rgb[i]$ is computed as the sum of RGB intensities of $masked\_frame[i]$ across width and height dividing by the sum of $mask[i]$ (the number of skin pixels). **Figure** 3.6 shows how, for each frame, there are three values being stored: one red average, one green average and one blue average. Displaying the masked frame is useful for debugging the raw signal extraction section.

Input: $frame[i], mask[i]$

Output: $avg\_rgb[i]$

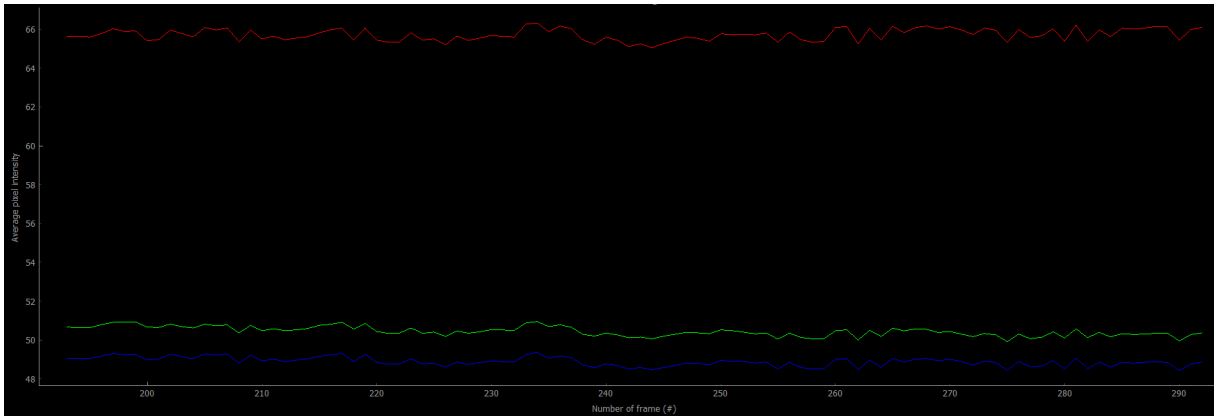

Figure 3.6: Example of average intensity per color channel, across frames.

3.5 Resampling

Within this real-time context (processing live video coming from your own webcam), it can be assumed that the video signal is asynchronous (*i.e.* sampled video frames are not evenly spaced in the time domain). Among many, one reason for this can be that the whole pipeline (including the video capture) is being executed from a single computer thread: for each new video frame, a varying different amount of time is needed to process it (and only after that will the video capture grab the next video sample), creating an irregular sampling pattern in time. In this sense, an online pipeline requires an additional resampling step to produce an evenly spaced signal for next stages, as common signal processing methods assume a fixed sampling rate on their signals. For resampling, linear interpolation is used. Within this scheme, each pair of consecutive frames generates evenly spaced points, according to a working sampling frequency, f_s. Note that the first frame captured (i=1) is seen as the pipeline's beginning of time. Thus, resampling is only performed for $i > 1$.

3.6 BVP signal extraction

To extract a one-dimensional BVP signal from a three-dimensional RGB signal we need a color combination algorithm, reviewed in (2.4.2). Here we use the POS method [20]. The POS method can be found in the *chrominance* category of rPPG algorithms, which means this algorithm integrates skin tone knowledge *a priori*, *i.e.* it requires less knowledge of the BVP signature and is more tolerant to distortion. Specifically, for a single light source (e.g., the fluorescent lamp), the skin pulsatility is the largest in the green-channel, followed by the blue-channel and red-channel. Given that, the authors suggest the first step to reduce dimensionality should be generating two signals, S1 and S2, from multiplying projection matrix **P** with RGB signal. Algebraically, green is accounted more than the blue and the red, and blue is accounted more than the red:

$$S_1(t) = Green(t) - Blue(t) \tag{3.1}$$

$$S_2(t) = Green(t) + Blue(t) - 2 \times Red(t) \tag{3.2}$$

$$P = \begin{bmatrix} 0 & 1 & -1 \\ -2 & 1 & 1 \end{bmatrix} \tag{3.3}$$

After that, alpha-tuning, introduced by [19], is used to arrive at the BVP:

$$h(t) = S_1(t) + \alpha \times S_2(t),\ where\ \alpha = \frac{\sigma(S1)}{\sigma(S2)} \tag{3.4}$$

This step uses knowledge of the BVP to define a rough projection region on the plane orthogonal to the temporally normalized skin-tone direction, and refine an exact projection direction on the plane by real-time tuning. The POS method is a sliding window algorithm. Its window size, $POS_l = f_s \cdot POS_s$, was chosen to correspond to $POS_s = 1.6$ seconds, which encapsulates one cardiac cycle as recommended by the authors. This choice introduces only a 1.6 seconds delay from acquiring the frame to extracting the corresponding facial blood volume, which is short enough for real-time applications.

Furthermore, filtering the raw BVP is essential, since it is easier to find true pulse peaks if the BVP is clean. Filtering is implemented using *SciPy* library [141]. A second-order infinite impulse response (IIR) bandpass butterworth filter, $[0.8, 2.5]$ Hz, was chosen since these frequencies correspond to a human heart rate range of 48 to 150 bpm. See **Figure 3.7** for the BVP signal and respective filtering.

3.7 Peak detection

For peak detection, we use *the mountaineer's method* [142], which was proposed to handle PPG signals regardless of their amplitude. The authors consider that every PPG pulse can be seen as a single

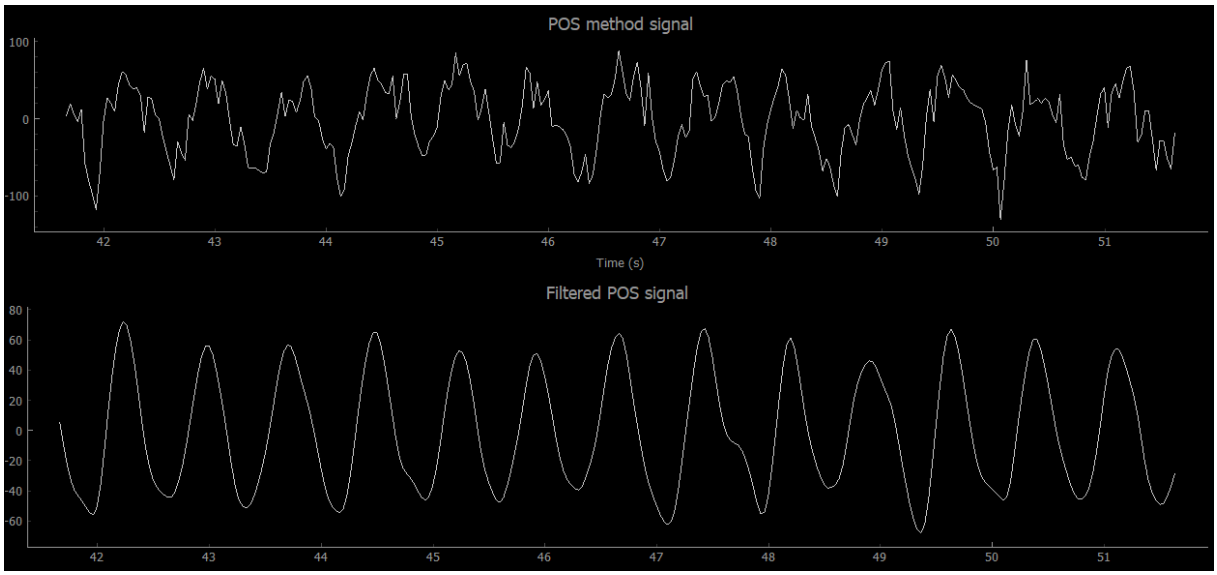

Figure 3.7: Blood volume pulse signal extraction via POS method (upper panel) and respective filtering (lower panel).

mountain, and thus the systolic peak can be thought as the top of the mountain. Their main assumption is that the rising edge preceding the systolic peak is a strictly increasing function and that the peak is reached when the slope changes from positive to negative, the same way a mountaineer would know that he is at the top of the mountain when, after a long time of walking up, he starts walking down. Simply put, the algorithm tracks how many steps it takes to observe changes in the slope of the signal and, like that, is able to detect both peaks and valleys in the mountain-like signal. Furthermore, the method is built so to ignore ripples that do not correspond to neither peaks or valleys, in the sense that if a mountaineer is going down a mountain but has to briefly climb up only to go down again, he can say that he has neither reached a peak or a valley. From the digital signal processing point of view, this is point-by-point, *windowless* algorithm, which is very convenient for a real-time application.

To further select true heartbeats from noise or motion artifact induced peaks, a simple *refereeing* algorithm is proposed after the mountaineer peak detection method. Its purpose is to, based on a previous PR estimate, iteratively predict where the next heartbeat should be, based on the current one. To achieve that, the referee first needs an outside mechanism that provides him with an estimate of PR. In this case, PR is being computed recurrently using non-overlapping 6-seconds portions the BVP signal via autoregressive modelling [143], with model order 12. Based on the current PR estimate and current peak time, he produces a time window where he will go look for the next peak. The window time limits of the referee, RW_{min} and RW_{max}, are set as follows:

$$PP_{estimate} = \frac{1}{\frac{HR_{estimate}}{60}} \tag{3.5}$$

$$next\ peak\ estimate = PP_{estimate} + last\ peak\ time \tag{3.6}$$

$$half\ \ width = \frac{PP_{estimate}}{2} \tag{3.7}$$

$$RW_{min} = next\ peak\ time + half\ \ width \tag{3.8}$$

$$RW_{max} = next\ peak\ time - half\ \ width \tag{3.9}$$

If the mountaineer method only detects one peak within the referee window, that peak is immediately marked as a true peak by the referee. If two or more mountaineer method peaks fall within the referee window, the referee marks as true peak the one closest to the center of the window. If no peak was detected by the mountaineer method within the referee window, the referee adds a beat in the center of the window proceeds from there. In **Figure 3.8** we can see an example of the proposed peak detection scheme.

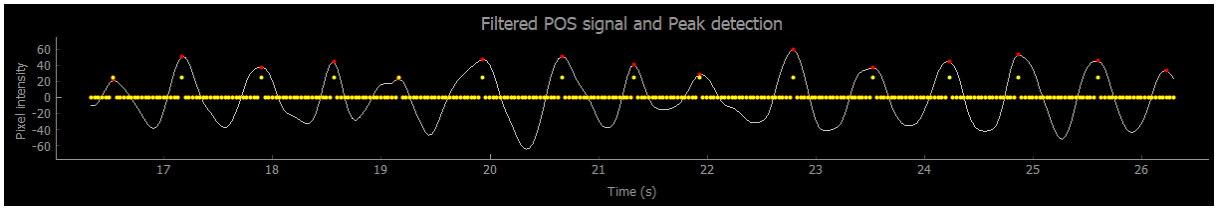

Figure 3.8: Example of proposed real-time peak detection. The red points correspond to peaks detected by the mountaineer method. The yellow points correspond to the peaks marked by the referee. In this case, because there is little distortion to the signal, the mountaineer method does very well and the referee simply marks all mountain peaks detected.

From the detected peaks, we extract PP intervals as the difference between pairs of consecutive peaks.

$$PPinterval_i = IBI_i = PeakTime_i - PeakTime_{i-1} \tag{3.10}$$

In **Figure 3.9** we can see the aspect of a PP signal, where each step represents the time, in milliseconds, between successive pairs of pulses.

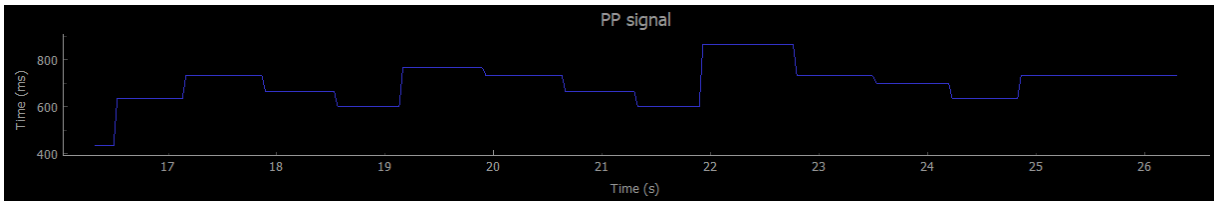

Figure 3.9: The PP signal corresponding to the blood volume pulse signal displayed in **Figure** 3.8.

3.8 HR and PRV estimates

Finally, we produce the biomedical statistics and features that describe a given BVP instance, with a specific collection of PP intervals.

Heart rate is calculated by averaging PP intervals over a time window, and computing the inverse of it:

$$HR_w = \frac{1}{PP_w} \tag{3.11}$$

where PP_w is the mean of all PP intervals that fall within the time window w. Assuming PP data is in seconds, this yields PR in Hertz; multiplying by 60 gives PR in beats-per-minute (bpm). The choice of this time window should reflect the user's requirement (e.g., instantaneous PR, long-term PR).

Two time-domain HRV metrics are considered: SDNN and RMSSD. For a given time window with N PP intervals, the formulas are as follows.

$$SDNN = \sqrt{\frac{1}{N} \times \sum_{1}^{N} (PP_i - \overline{PP})^2} \tag{3.12}$$

$$RMSSD = \sqrt{\frac{1}{N-1} \times \sum_{1}^{N-1} (PP_{i+1} - PP_i)^2} \tag{3.13}$$

We also consider three frequency-domain metrics: normalized LF power, normalized HF power and LF/HF ratio. While Fourier methods [144] and auto-regressive models [143] are popular means of obtaining frequency domain features from the RR signal, they require a PP signal explicitly. According to [145], when dealing with time series with irregular sampling period (i.e. the time between two heart rate samples is dependent of the RR interval duration), it is recommended to use the Lomb-Scargle method. In this case, we don't need to have an explicit PP signal, which makes the method suited for real time applications [146]. In **Figure** 3.10 we can see the Fourier like spectrogram the Lomb-Scargle method yields.

To compute all PRV metrics efficiently, we use the *hrv-analysis* python library [147]. In the end we should have several PR and PRV readings, like we see in **Figure** 3.11

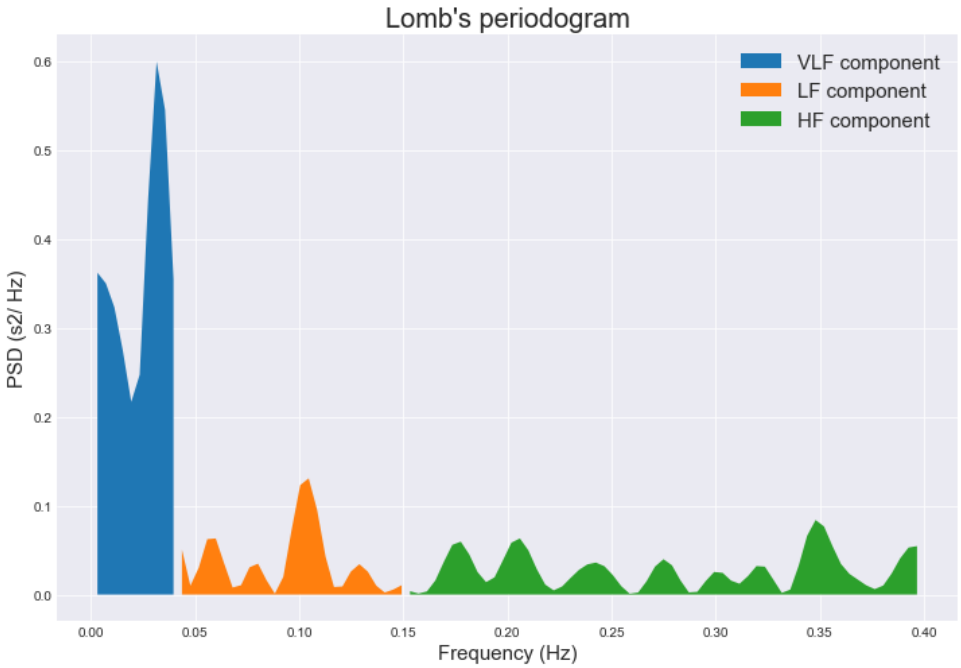

Figure 3.10: An example of a pulse rate variability spectrogram obtained through the Lomb-Scargle method. This spectragrom was created with the python library *hrv-analysis*.

Figure 3.11: The output of the algorithm. In the top left corner, we can see both pulse rate (long-term PR and last-5-PP PR) and pulse rate variability (SDNN, RMSSD, LF, HF and LF/HF) readings.

4

Results and Discussion

Contents

4.1 Offline testing

The proposed method was tested with the *UBFC-RPPG video dataset*, introduced in [148]. This dataset is composed of 40 videos and respective PPG ground truth, gathered from a pulse oximeter finger clip sensor. Each video lasts more than one minute and is recorded with a low cost webcam at 30 fps with a resolution of 640×480 in uncompressed 8-bits RGB format. In these videos, the subjects sit in front of the camera, about one-meter away, and they play a time sensitive mathematical game that supposedly raises the heart rate. This setting shouldn't be too different from that of a medical consultation via webcam, apart from the fact that the videos in this dataset contain very little subject movement and have nice, constant lighting.

Table 4.1 reports the performance of the proposed method for every subject of the UBFC dataset in terms of mean absolute error between estimated and ground truth PR and PRV features. The normalized HF error is not displayed as it is the same as LF.

In videos for which the skin mask is noisy, *i.e.* facial landmarks positions vary drastically even when there is no facial movement, non skin pixels (*e.g.* wall behind the subject) are sampled and fed into the rPPG algorithm introducing distortions in the BVP signal estimation. When distortions are as large as the BVP signal, AR model referee estimation of PR fails and the proposed peak detection scheme collapses. Even though the proposed algorithm shrinks the skin mask to avoid pixels close to the edge of the face, there are extreme cases for which landmark estimates are still too noisy. A very pronounced case of noisy mask is that of subject 41, illustrated in **Figure 4.1**, for which the proposed method performs the poorest. As can bee seen, the landmark prediction fails and the corrections can not exclude the non-skin pixels. This result underlies that selecting skin-only pixels is essential for any rPPG algorithm.

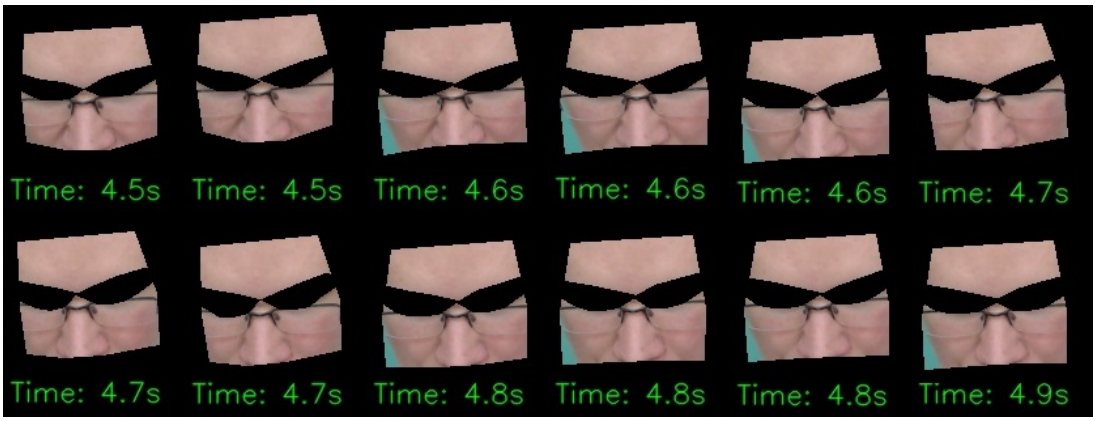

Figure 4.1: Twelve consecutive frames showing skin mask noise, seen here as a blue wall behind the subject. Because the wall pixels can enter and leave the ROI very fast, noisy fluctuations appear in the RGB color signals, which are uncorrelated with blood volume changes in facial vasculature.

Color of skin does not seem to have an impact in the performance of the proposed method, as the average of performance for dark tone is similar to light tone. However, there are only two colored

Table 4.1: Mean absolute error between ground truth and estimated pulse rate and pulse rate variability features, per subject. Columns are marked with asterisks according to the length of the analysis window used: * for 15-second window, ** for 30-second window, *** for full record.

Subject	PR* (bpm)	PR** (bpm)	PR*** (bpm)	SDNN (ms)	RMSSD (ms)	LF (nu)	LF/HF
1	4,1987	3,623055	0,643013	9,133175	14,6971503	9,3313	0,298607
3	2,348539	1,794518	0,553611	8,444861	34,5389541	16,00381	7,642862
4	5,564646	4,251566	6,65146	15,65751	57,393792	4,257184	0,128137
5	1,844587	1,130269	0,921357	15,32402	33,4517674	2,035155	0,032921
8	0,977157	0,838798	0,010891	8,082792	20,3019335	15,21441	0,555628
9	2,969618	1,783416	1,152776	29,3187	45,1347266	32,01188	1,110442
10	0,699623	0,400299	0,567125	6,563537	20,1280719	6,491911	0,287745
11	1,644524	1,036419	0,764924	14,3369	31,0958797	38,89513	2,433866
12	2,683789	2,047038	0,93426	29,25841	44,386202	13,55559	0,26086
13	1,448835	1,327466	0,444564	6,215349	20,8335278	6,970089	0,650175
14	1,205541	1,252071	0,694287	17,47662	63,9381263	20,49469	1,926514
15	6,270456	4,429778	9,782752	1,252001	5,78305055	1,601859	0,034329
16	1,882559	1,639633	0,752636	22,86777	56,247141	15,33861	0,519647
17	1,741078	1,473016	0,241781	15,764	35,6330882	36,36467	1,807558
18	3,357007	2,395363	0,178454	8,955568	16,6111996	17,79268	0,737141
20	1,812351	1,333211	3,854949	26,44222	53,5964425	46,57441	2,712767
22	2,673689	3,1702	0,804041	42,54923	76,2418712	15,89661	0,987123
23	2,908396	3,264732	1,900757	37,88024	60,74649	0,921517	0,037022
24	4,711815	4,720487	1,693969	40,52374	81,9078375	8,998069	0,277155
25	4,653725	3,949907	1,20052	5,323105	18,8029615	22,76256	0,454462
26	1,879532	1,593659	0,223179	21,13367	39,7181849	38,66064	1,556223
27	2,139063	2,20538	2,832174	24,74535	54,8340479	26,53145	3,420723
30	0,688636	0,473665	0,475013	13,37812	35,8015257	29,87593	2,009246
31	0,868949	0,508398	0,365321	9,925167	11,6108863	0,891866	0,020663
32	0,480838	0,523094	0,178176	21,87427	37,6920232	1,58746	0,044037
33	4,790463	5,700498	1,931864	20,28369	28,2869167	15,73327	0,374927
34	1,403921	1,353179	0,521882	13,35422	28,0767238	10,98356	0,553321
35	1,454661	1,714099	0,286558	1,750292	17,8626879	4,271267	0,537495
36	2,175227	2,032103	0,143852	5,62433	14,2635718	13,73998	2,981935
37	1,089043	1,340806	0,381518	2,855368	31,7900006	9,441993	0,362368
38	2,003767	1,542528	2,372209	61,72839	110,799441	41,00793	1,232811
39	0,525357	0,316026	0,027358	11,54208	16,5909121	19,24709	1,743222
40	2,320192	2,181635	0,664368	40,56172	60,0036794	17,40573	0,614505
41	8,659682	8,080893	3,284753	103,5145	70,8345917	0,753694	0,031585
42	0,568467	0,570574	0,14595	3,090669	0,63690635	2,730146	0,347615
43	2,343259	1,293915	0,46283	25,64717	44,2027604	21,53345	0,644212
44	1,390667	1,378657	0,422157	23,13219	46,0356676	26,45901	1,27013
45	1,388663	1,446941	0,229227	6,221288	14,5452721	5,825947	0,75367
46	4,764737	5,337647	0,384027	18,28357	42,1611379	21,59623	0,71285
47	1,490869	1,116672	2,580767	28,39756	71,0768062	35,71173	1,523752

subjects within the dataset, which is a very small statistical sample to make meaningful claims.

Hair in front of the forehead or wearing glasses doesn't seem to lower the performance of the proposed method, even if they are captured by the skin mask. However, these factors can cause abnormalities for the landmark prediction routine (see **Figure** 4.2).

Figure 4.2: Poor landmark prediction due facial occlusion by hair.

The peak detection referee is introduced to deal with signals with distortions and motion artifacts in real time. In **Figure** 4.3 we can see the referee filtering out unwanted, distortion-induced peaks.

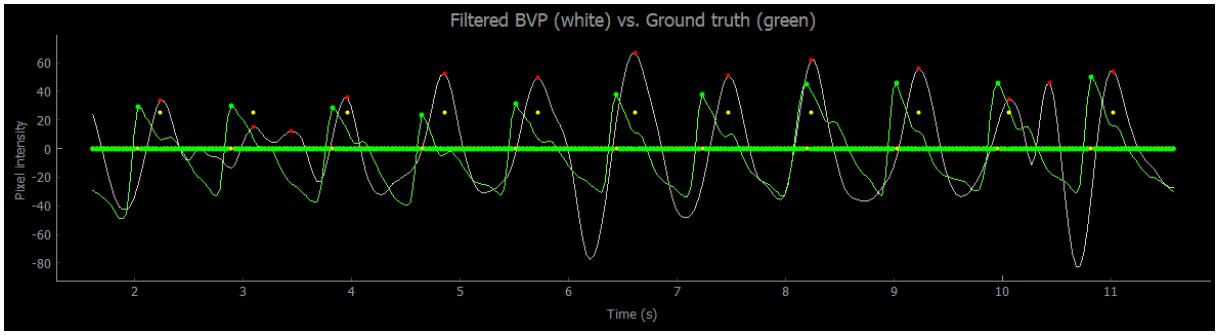

Figure 4.3: An example of good functioning of the peak detection refereeing at around 3.5 seconds and 10.5 seconds. The white line represents the blood volume pulse signal extracted from remote photoplethysmography. Red points are all detected peaks. Yellow points are the peaks produced by the referee. The green signal is the ground truth from a finger photoplethysmography sensor. Green points are ground truth peaks.

When subject motion, illumination changes or mask noise, alone, or together, degrade the signal too much, the peak detection referee starts to fail badly. Besides that, the referee might also fail in situations where a big variation in pulse rate takes place and the trailing auto-regressive model guess becomes useless. In this case, the referee will add abnormal pulses and because it uses the previous beat to

Table 4.2: Average and standard deviation of the estimation error across UBFC subjects, per feature.

	PR* (bpm)	PR** (bpm)	PR*** (bpm)	SDNN (ms)	RMSSD (ms)	%LF (nu)	LF/HF
AVERAGE	2,45	2,16	1,29	20,46	39,21	16,89	1,09
STD	1,75	1,67	1,87	18,6	23,27	12,63	1,35

produce the next one, this might create problems in series. See, for example, **Figure 4.4**.

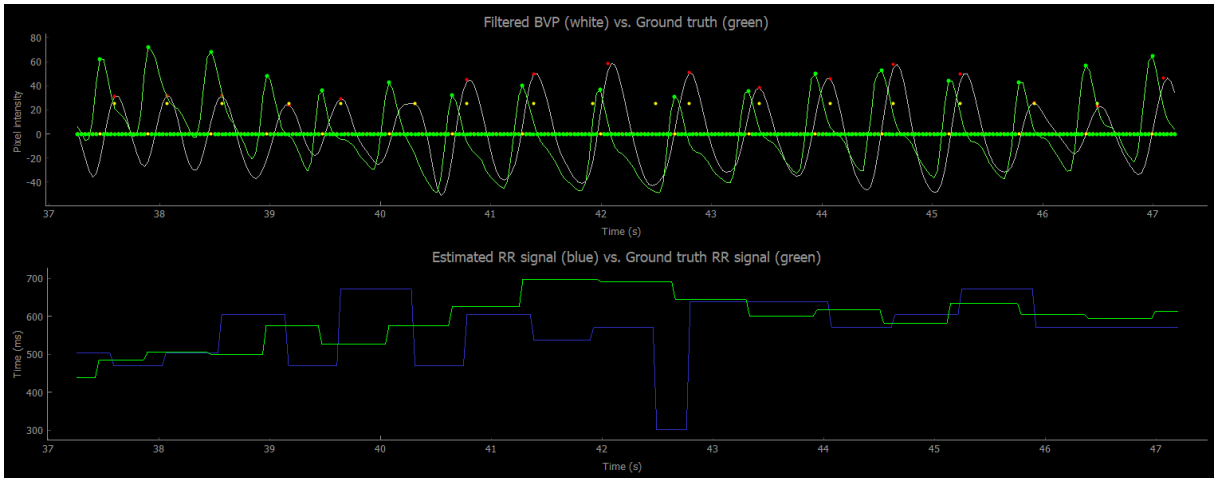

Figure 4.4: An example of bad functioning of the referee. In the upper panel we can see the filtered blood volume pulse (white) against the ground truth photoplethysmography signal (green). In the lower panel we can see the intervals between estimated peaks (blue) and for ground truth peaks (green). In this case, while the mountaineer method (points in red in the upper panel) works properly, the referee (points in yellow) creates problems around 42 seconds. Here, because the subject's pulse rate starts to slow down (visible in the green, the ground truth intervals in the lower panel), it no longer matches the previous pulse rate guess made by the auto-regressive model and the referee fails to include the true peak in its window. Because the referee doesn't detect a peak, it adds one in the center of the window and after that fails again to include the true peak. Finally, it catches up to the true peak detected by the mountaineer method but this creates a pronounced problem in the estimated RR signal (visible in blue in the lower panel, around 42.5 second).

Table 4.2 reports on overall performance of the proposed method when testes with the UBFC dataset, in terms of average error and standard deviation.

Comparing against the state-of-the-art in PR estimation (**Table 4.3**), the proposed method (mean error = 2,16 bpm) performs extremely well, competing with the best previous non-supervised or supervised methods. This result implies that the algorithm performs well for PR measurements in the short-term. Although PR doesn't relate directly to PRV and mental complication, this suggests that the webcam rPPG algorithm should be able to compete with traditional PPG for some applications that need PR measurement. Besides, mobile applications that make PR measurements by placing the smartphone camera on the finger are not alone anymore, as we see that the facial video is just as good.

Regarding PR variability estimations, the method performs good when compared to stat-of-the-art reports (see **Table 4.4**). Regarding estimating SDNN and RMSS, the method performs worse than the one proposed by [144], but better than any other. For %LF (nu) and LF/HF estimation, the proposed

Table 4.3: Performance of state-of-the-art algorithms for pulse rate estimation on the UBFC dataset. Values correspond to the mean absolute error in bpm, either for 30 second windows or full record.

ICA	CHROM	POS	PVM	cPR + fine	MAICA	CK	FaceRPPG	PulseGAN
6,02 [149]	3,7 [149]	4,73 [149]	4,47 [149]	2,1 [150]	3,43 [151]	2,29 [152]	2,37 [144]	2,09 [153]

Table 4.4: State-of-the-art pulse rate variability estimation performance of in terms of mean absolute error for the UBFC dataset.

Method	FaceRPPG [144]	SSF [154]	CHROM	PulseGAN [153]	Proposed method
SDNN (ms)	19	25	38,9 [153]	24,3	20,46
RMSSD (ms)	16	47	93 [154]		39,21
%LF (nu)	20				16,89
LF/HF	1,3				1,09

method outperforms all methods reported.

According to [47], average short-term human HRV features have the following ranges: SDNN ranges from 32-93 ms; RMSSD ranges from 19-75 ms; %LF (nu) ranges from 30-65 and LF/HF ranges from 1.1-11.6. For SDNN, the proposed method has an average error of around 20 ms, which mean we should be able to roughly distinguish three profiles within the SDNN human range. Regarding RMSSD, an average error of 39,21 ms might be too large for biomedical application, since the mean error is almost as big as the allowed range. As for %LF (nu), an average error of 16% should be little enough to distinguish at least two types of autonomic profiles. Finally, the average error for LF/HF ratio estimation is the smallest compared to its respective human range, which suggests that this is the most reliable feature for biomedical application coming out of the proposed algorithm.

4.2 Live demonstration

For the live demonstration, a standard grade laptop (MSI GF63 8RD) runs the whole rPPG pipeline in real-time using the embedded webcam as video input in one thread, and, on another thread, it acquires the traditional PPG signal from a pulse oximeter finger clip sensor, *i.e.* the ground truth signal. The PPG acquisition is mediated through a BITalino board [155], which stores the data at a constant rate and sends it via bluetooth to the laptop. In the end, we can display both signals at the same time and confirm that the rPPG signal follows the PPG signal closely (see **Figure 4.5**), though we can see that the two signals are not perfectly aligned. This delay might be related with: 1) PPG thread starting acquisition first than the webcam thread, or vice versa and 2) the amount of time blood takes to travel from the heart to the face is different from the time it takes travelling from the heart to the finger, and their also target of regulation by the circulatory system control mechanisms. Anyway, we can see that for every PPG-sensor pulse we can count a corresponding delayed rPPG-sensor pulse, confirming the ability of the proposed real-time algorithm to capture pulse rate variability, just like traditional PPG can do.

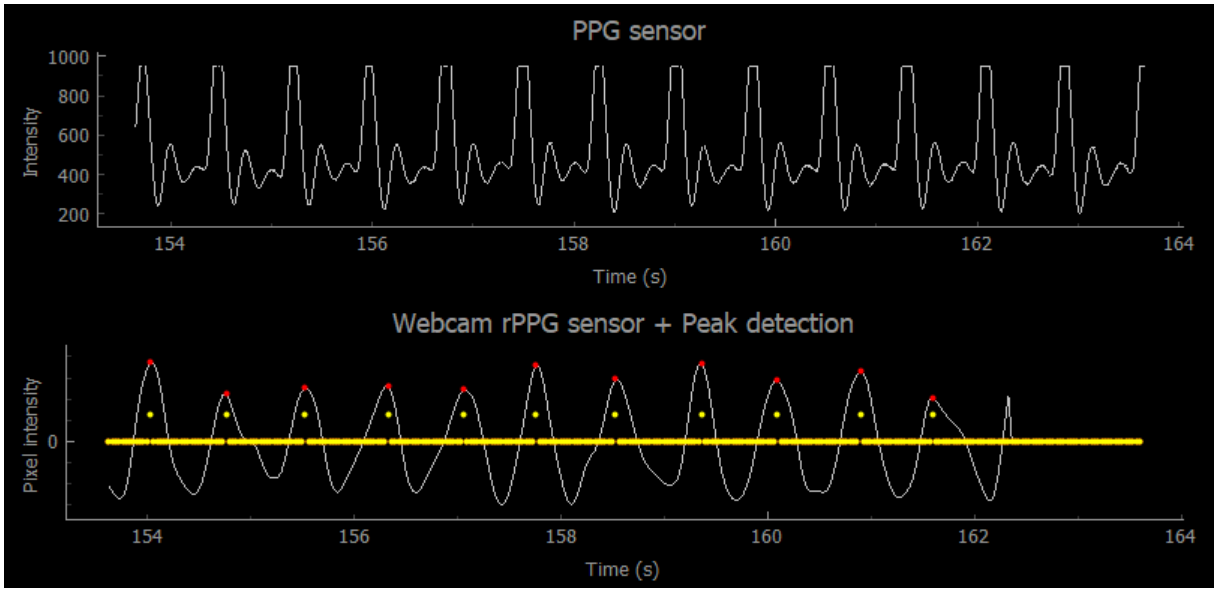

Figure 4.5: Real-time comparison of the photoplethysmography signal obtained through the finger clip sensor (upper panel) and the remote photoplethysmography signal obtained from the webcam video. In this image we can also see the 1.6 second delay of the remote photoplethysmography algorithm.

5

Conclusion

Contents

5.1 Conclusions

This paper has tried to show that not only rPPG is a potential tool for biomedical application to the psychiatric field (and other, like cardiology), but that it can be accomplished in real time to aid doctors in tele consultation.

Virtually all published work on rPPG is based on offline computations [95]. In the offline context (*e.g.* processing a video file from a scientific database), we readily have all the samples at disposal, which allows separation of processes (*e.g.* first detect faces for all frames, then extract signal for all frames, etc). In the real-time context this is not true and each process has to be carefully orchestrated with adjacent processes to make sure each signal processing method is activated only when it has available data to proceed. This is particularly important when there are several windowed processes working in series (*e.g.* the first method operates with a 1.6 second window, thus the second method, which might work with a 2 second window, has to wait for that method, and so on and so forth for following methods). This constraints algorithm choice for the real-time context, which aims for the lowest possible time window between 'acquiring video samples' and 'displaying PRV features'. Thus, simple, low complexity algorithms were preferred over complex ones, in order to get closer to real-time PRV readings.

Accomplishing real-time pulse rate variability means that we can inspect the raw signal and PRV features during recording, which allows to identify artifacts, make sense of the values and overall have control over the precision of the process which is in accordance to recommendations made by [156]. Plus, a real-time rPPG algorithm ensures the doctor has some control over the quality of the measurement. The real-time display not only allows him to search for a sweet spot in terms of patient positioning and ambient lighting, but also to control overall quality of the record.

The most significant limitation of the proposed algorithm is that its automated peak detection can both miss and or add beats, which largely distorts PRV features and overall produces misleading results [157].

According to [158], PRV features extracted from any PPG sensor with a sampling frequency above or equal to 25 Hz should not significantly differ from the HRV features extracted from the ECG, which has much larger sampling frequencies. For a standard laptop, the proposed algorithm runs at 18 Hz, which means improvements have to be made in order for it to run above 25 Hz in regular computers.

This work is relevant because telemedicine is an emerging, cheaper form of providing health services and reliable tools must be developed to support doctors in making decisions within the remote context.

All in all, while the proposed method may be unsuited for very fine diagnostic, i.e be able to separate patients basted on illness severity, it may still be able to distinguish between extremes (*e.g.* healthy patient and severe depression) and this should be promising.

5.2 Future Work

In the future, it should be tested whether PRV features produced by the algorithm correlate with different groups of psychiatric patients. For example, testing whether the proposed rPPG algorithm produces PRV features that are reliable enough to distinguish between healthy patients and patients with severe depression with the help of classification techniques (*e.g.* linear regression).

On the other side, investigating physiological mechanisms that underlie the similarities and differences between finger PPG, rPPG and ECG signals is of importance, as these tools are currently all available and can mean an advancement for several types of medical diagnostic, alone and together.

Bibliography

[1] G. G. Berntson, J. Thomas Bigger Jr, D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone *et al.*, "Heart rate variability: origins, methods, and interpretive caveats," *Psychophysiology*, vol. 34, no. 6, pp. 623–648, 1997.

[2] G. E. Billman, "Heart rate variability–a historical perspective," *Frontiers in physiology*, vol. 2, p. 86, 2011.

[3] M. Rowbottom and C. Susskind, *Electricity and medicine: history of their interaction*. San Francisco Press, 1984.

[4] M. AlGhatrif and J. Lindsay, "A brief review: history to understand fundamentals of electrocardiography," *Journal of community hospital internal medicine perspectives*, vol. 2, no. 1, p. 14383, 2012.

[5] S. S. Barold, "Willem einthoven and the birth of clinical electrocardiography a hundred years ago," *Cardiac electrophysiology review*, vol. 7, no. 1, pp. 99–104, 2003.

[6] R. Mccraty and F. Shaffer, "Heart rate variability: New perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk," *Global Advances in Health and Medicine*, vol. 4, no. 1, pp. 46–61, Jan. 2015. [Online]. Available: https://doi.org/10.7453/gahmj.2014.073

[7] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, "Heart rate variability: a review," *Medical & Biological Engineering & Computing*, vol. 44, no. 12, pp. 1031–1051, Nov. 2006. [Online]. Available: https://doi.org/10.1007/s11517-006-0119-0

[8] D. S. Quintana, G. A. Alvares, and J. A. J. Heathers, "Guidelines for reporting articles on psychiatry and heart rate variability (GRAPH): recommendations to advance research communication," *Translational Psychiatry*, vol. 6, no. 5, pp. e803–e803, May 2016. [Online]. Available: https://doi.org/10.1038/tp.2016.73

[9] A. B. Hertzman, "Observations on the finger volume pulse recorded photoelectrically," *Am. J. Physiol.*, vol. 119, pp. 334–335, 1937.

[10] ——, “The blood supply of various skin areas as estimated by the photoelectric plethysmograph,” *American Journal of Physiology-Legacy Content*, vol. 124, no. 2, pp. 328–340, 1938.

[11] A. B. Hertzman and J. B. Dillon, “Distinction between arterial, venous and flow components in photoelectric plethysmography in man,” *American Journal of Physiology-Legacy Content*, vol. 130, no. 1, pp. 177–185, 1940.

[12] ——, “Applications of photoelectric plethysmography in peripheral vascular disease,” *American Heart Journal*, vol. 20, no. 6, pp. 750–761, 1940.

[13] A. B. Hertzman and L. W. Roth, “The absence of vasoconstrictor reflexes in the forehead circulation. effects of cold,” *American Journal of Physiology-Legacy Content*, vol. 136, no. 4, pp. 692–697, 1942.

[14] T. Aoyagi, “Pulse oximetry: its invention, theory, and future,” *Journal of anesthesia*, vol. 17, no. 4, pp. 259–266, 2003.

[15] A. Schäfer and J. Vagedes, “How accurate is pulse rate variability as an estimate of heart rate variability?: A review on studies comparing photoplethysmographic technology with an electrocardiogram,” *International journal of cardiology*, vol. 166, no. 1, pp. 15–29, 2013.

[16] E. Yuda, M. Shibata, Y. Ogata, N. Ueda, T. Yambe, M. Yoshizawa, and J. Hayano, “Pulse rate variability: a new biomarker, not a surrogate for heart rate variability,” *Journal of physiological anthropology*, vol. 39, no. 1, pp. 1–4, 2020.

[17] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, “Remote plethysmographic imaging using ambient light,” *Optics Express*, vol. 16, no. 26, p. 21434, Dec. 2008. [Online]. Available: https://doi.org/10.1364/oe.16.021434

[18] M.-Z. Poh, D. J. McDuff, and R. W. Picard, “Non-contact, automated cardiac pulse measurements using video imaging and blind source separation,” *Optics Express*, vol. 18, no. 10, p. 10762, May 2010. [Online]. Available: https://doi.org/10.1364/oe.18.010762

[19] G. de Haan and V. Jeanne, “Robust pulse rate from chrominance-based rPPG,” *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, pp. 2878–2886, Oct. 2013. [Online]. Available: https://doi.org/10.1109/tbme.2013.2266196

[20] W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, “Algorithmic principles of remote PPG,” *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1479–1491, Jul. 2017. [Online]. Available: https://doi.org/10.1109/tbme.2016.2609282

[21] J. Heathers and M. Goodwin, “Dead science in live psychology: A case study from heart rate variability (HRV),” *PsyArXiv*, Sep. 2017. [Online]. Available: https://doi.org/10.31234/osf.io/637ym

[22] S. D. Kreibig, “Autonomic nervous system activity in emotion: A review,” *Biological psychology*, vol. 84, no. 3, pp. 394–421, 2010.

[23] A. R. Brunoni, A. H. Kemp, E. M. Dantas, A. C. Goulart, M. A. Nunes, P. S. Boggio, J. G. Mill, P. A. Lotufo, F. Fregni, and I. M. Benseñor, “Heart rate variability is a trait marker of major depressive disorder: evidence from the sertraline vs. electric current therapy to treat depression clinical study,” *International Journal of Neuropsychopharmacology*, vol. 16, no. 9, pp. 1937–1949, 2013.

[24] J. H. Ha, S. Park, D. Yoon, and B. Kim, “Short-term heart rate variability in older patients with newly diagnosed depression,” *Psychiatry research*, vol. 226, no. 2-3, pp. 484–488, 2015.

[25] G. A. Alvares, D. S. Quintana, I. B. Hickie, and A. J. Guastella, “Autonomic nervous system dysfunction in psychiatric disorders and the impact of psychotropic medications: a systematic review and meta-analysis.” *Journal of Psychiatry & Neuroscience*, 2016.

[26] D. S. Quintana, A. J. Guastella, T. Outhred, I. B. Hickie, and A. H. Kemp, “Heart rate variability is associated with emotion recognition: direct evidence for a relationship between the autonomic nervous system and social cognition,” *International journal of psychophysiology*, vol. 86, no. 2, pp. 168–172, 2012.

[27] A. L. Hansen, B. H. Johnsen, J. J. Sollers, K. Stenvik, and J. F. Thayer, “Heart rate variability and its relation to prefrontal cognitive function: the effects of training and detraining,” *European journal of applied physiology*, vol. 93, no. 3, pp. 263–272, 2004.

[28] A. Reinecke, N. Filippini, C. Berna, D. Western, B. Hanson, M. Cooper, P. Taggart, and C. Harmer, “Effective emotion regulation strategies improve fmri and ecg markers of psychopathology in panic disorder: implications for psychological treatment action,” *Translational psychiatry*, vol. 5, no. 11, pp. e673–e673, 2015.

[29] F. H. Martini, J. L. Nath, E. F. Bartholomew, and W. C. Ober, *Fundamentals of anatomy & physiology*. Pearson, 2018.

[30] A. M. Katz, *Physiology of the Heart*. Lippincott Williams & Wilkins, 2010.

[31] F. Shaffer, R. McCraty, and C. L. Zerr, “A healthy heart is not a metronome: an integrative review of the heart’s anatomy and heart rate variability,” *Frontiers in psychology*, vol. 5, p. 1040, 2014.

[32] G. J. Tortora, B. H. Derrickson, G. J. Tortora, and G. J. Tortora, *Tortoras Principles of anatomy & physiology*. Wiley, 2017.

[33] E. N. Marieb and K. Hoehn, *Human anatomy & physiology*. Pearson Education., 2019.

[34] J. Jalife, "Neural control of sinoatrial pacemaker activity," *Vagal control of the heart: Experimental basis and clinical implications*, pp. 173–205, 1994.

[35] B. Fred Shaffer PhD and J. V. MAE, "Heart rate variability anatomy and physiology," *Biofeedback (Online)*, vol. 41, no. 1, p. 13, 2013.

[36] H. Brown, D. DiFrancesco, and S. Noble, "How does adrenaline accelerate the heart?" *Nature*, vol. 280, no. 5719, pp. 235–236, 1979.

[37] B. Sakmann, A. Noma, and W. Trautwein, "Acetylcholine activation of single muscarinic k+ channels in isolated pacemaker cells of the mammalian heart," *Nature*, vol. 303, no. 5914, pp. 250–253, 1983.

[38] J. P. Saul, "Beat-to-beat variations of heart rate reflect modulation of cardiac autonomic outflow," *Physiology*, vol. 5, no. 1, pp. 32–37, 1990.

[39] R. Hainsworth, "The control and physiological importance of heart rate," *Heart rate variability*, pp. 3–9, 1995.

[40] M. L. Ogletree-Hughes, L. B. Stull, W. E. Sweet, N. G. Smedira, P. M. McCarthy, and C. S. Moravec, "Mechanical unloading restores beta-adrenergic responsiveness and reverses receptor downregulation in the failing human heart," *Circulation*, vol. 104, no. 8, pp. 881–886, Aug. 2001. [Online]. Available: https://doi.org/10.1161/hc3301.094911

[41] F. Shaffer and J. Venner, "Heart rate variability anatomy and physiology," *Biofeedback*, vol. 41, no. 1, pp. 13–25, Mar. 2013. [Online]. Available: https://doi.org/10.5298/1081-5937-41.1.05

[42] B. Pomeranz, R. Macaulay, M. A. Caudill, I. Kutz, D. Adam, D. Gordon, K. M. Kilborn, A. C. Barger, D. C. Shannon, and R. J. Cohen, "Assessment of autonomic function in humans by heart rate spectral analysis," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 248, no. 1, pp. H151–H153, 1985.

[43] M. Bootsma, C. A. Swenne, H. H. Van Bolhuis, P. C. Chang, V. M. Cats, and A. Bruschke, "Heart rate and heart rate variability as indexes of sympathovagal balance," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 266, no. 4, pp. H1565–H1571, 1994.

[44] D. B. Geselowitz, "On the theory of the electrocardiogram," *Proceedings of the IEEE*, vol. 77, no. 6, pp. 857–876, 1989.

[45] T. Force, “Standards of measurement, physiological interpretation and clinical use. task force of the european society of cardiology and the north american society of pacing and electrophysiology,” *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.

[46] M. Peltola, “Role of editing of rr intervals in the analysis of heart rate variability,” *Frontiers in physiology*, vol. 3, p. 148, 2012.

[47] F. Shaffer and J. Ginsberg, “An overview of heart rate variability metrics and norms,” *Frontiers in public health*, vol. 5, p. 258, 2017.

[48] T. Kuusela, “Methodological aspects of heart rate variability analysis,” in *Heart Rate Variability (HRV) Signal Analysis*. CRC Press, Oct. 2012, pp. 9–42. [Online]. Available: https://doi.org/10.1201/b12756-4

[49] J. Sztajzel *et al.*, “Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system,” *Swiss medical weekly*, vol. 134, no. 35-36, pp. 514–522, 2004.

[50] T. F. of the European Society Electrophysiology, “Heart rate variability,” *Circulation*, vol. 93, no. 5, pp. 1043–1065, Mar. 1996. [Online]. Available: https://doi.org/10.1161/01.cir.93.5.1043

[51] R. E. Kleiger, P. K. Stein, and J. T. Bigger Jr, “Heart rate variability: measurement and clinical utility,” *Annals of Noninvasive Electrocardiology*, vol. 10, no. 1, pp. 88–101, 2005.

[52] J. F. Thayer, A. L. Hansen, E. Saus-Rose, and B. H. Johnsen, “Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health,” *Annals of Behavioral Medicine*, vol. 37, no. 2, pp. 141–153, 2009.

[53] S. W. Porges, “The polyvagal perspective,” *Biological psychology*, vol. 74, no. 2, pp. 116–143, 2007.

[54] P. Grossman and E. W. Taylor, “Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions,” *Biological psychology*, vol. 74, no. 2, pp. 263–285, 2007.

[55] P. Lehrer, “How does heart rate variability biofeedback work? resonance, the baroreflex, and other mechanisms,” *Biofeedback*, vol. 41, no. 1, pp. 26–31, 2013.

[56] R. McCraty and D. Childre, “Coherence: bridging personal, social, and global health,” *Altern Ther Health Med*, vol. 16, no. 4, pp. 10–24, 2010.

[57] R. Smith, J. F. Thayer, S. S. Khalsa, and R. D. Lane, “The hierarchical basis of neurovisceral integration,” *Neuroscience & biobehavioral reviews*, vol. 75, pp. 274–296, 2017.

[58] K.-J. Bär, W. Greiner, T. Jochum, M. Friedrich, G. Wagner, and H. Sauer, "The influence of major depression and its treatment on heart rate variability and pupillary light reflex parameters," *Journal of affective disorders*, vol. 82, no. 2, pp. 245–252, 2004.

[59] T. Rechlin, M. Weis, A. Spitzer, and W. P. Kaschka, "Are affective disorders associated with alterations of heart rate variability?" *Journal of affective disorders*, vol. 32, no. 4, pp. 271–275, 1994.

[60] G. W. Dalack and S. P. Roose, "Perspectives on the relationship between cardiovascular disease and affective disorder." *The Journal of clinical psychiatry*, 1990.

[61] V. K. Yeragani, R. Pohl, R. Balon, C. Ramesh, D. Glitz, I. Jung, and P. Sherwood, "Heart rate variability in patients with major depression," *Psychiatry research*, vol. 37, no. 1, pp. 35–46, 1991.

[62] M. Lehofer, M. Moser, R. Hoehn-Saric, D. McLeod, P. Liebmann, B. Drnovsek, S. Egner, G. Hildebrandt, and H.-G. Zapotoczky, "Major depression and cardiac autonomic control," *Biological psychiatry*, vol. 42, no. 10, pp. 914–919, 1997.

[63] M.-F. O'Connor, J. J. Allen, and A. W. Kaszniak, "Autonomic and emotion regulation in bereavement and depression," *Journal of psychosomatic research*, vol. 52, no. 4, pp. 183–185, 2002.

[64] J. Rottenberg, "Cardiac vagal control in depression: a critical analysis," *Biological psychology*, vol. 74, no. 2, pp. 200–211, 2007.

[65] G. G. Berntson, D. L. Lozano, and Y.-J. Chen, "Filter properties of root mean square successive difference (rmssd) for heart rate," *Psychophysiology*, vol. 42, no. 2, pp. 246–252, 2005.

[66] A. H. Kemp, D. S. Quintana, M. A. Gray, K. L. Felmingham, K. Brown, and J. M. Gatt, "Impact of depression and antidepressant treatment on heart rate variability: a review and meta-analysis," *Biological psychiatry*, vol. 67, no. 11, pp. 1067–1074, 2010.

[67] J. Koenig, A. H. Kemp, T. P. Beauchaine, J. F. Thayer, and M. Kaess, "Depression and resting state heart rate variability in children and adolescents—a systematic review and meta-analysis," *Clinical psychology review*, vol. 46, pp. 136–150, 2016.

[68] C. Koch, M. Wilhelm, S. Salzmann, W. Rief, and F. Euteneuer, "A meta-analysis of heart rate variability in major depression," *Psychological medicine*, vol. 49, no. 12, pp. 1948–1957, 2019.

[69] R. Hartmann, F. M. Schmidt, C. Sander, and U. Hegerl, "Heart rate variability as indicator of clinical state in depression," *Frontiers in psychiatry*, vol. 9, p. 735, 2019.

[70] C. A. V. J, "Photoelectric plethysmography for estimating cutaneous blood flow," in *Non-Invasive Physiological Measurements vol. I*, P. Rolfe, Ed. London: Academic, 1979, pp. 125–151.

[71] A. Kamal, J. Harness, G. Irving, and A. Mearns, “Skin photoplethysmography — a review,” *Computer Methods and Programs in Biomedicine*, vol. 28, no. 4, pp. 257–269, Apr. 1989. [Online]. Available: https://doi.org/10.1016/0169-2607(89)90159-4

[72] A. B. Hertzman, “THE BLOOD SUPPLY OF VARIOUS SKIN AREAS AS ESTIMATED BY THE PHOTOELECTRIC PLETHYSMOGRAPH,” *American Journal of Physiology-Legacy Content*, vol. 124, no. 2, pp. 328–340, Oct. 1938. [Online]. Available: https://doi.org/10.1152/ajplegacy.1938.124.2.328

[73] J. Allen, “Photoplethysmography and its application in clinical physiological measurement,” *Physiological Measurement*, vol. 28, no. 3, pp. R1–R39, Feb. 2007. [Online]. Available: https://doi.org/10.1088/0967-3334/28/3/r01

[74] A. Reisner, P. A. Shaltis, D. McCombie, H. H. Asada, D. S. Warner, and M. A. Warner, “Utility of the photoplethysmogram in circulatory monitoring,” *Anesthesiology*, vol. 108, no. 5, pp. 950–958, May 2008. [Online]. Available: https://doi.org/10.1097/aln.0b013e31816c89e1

[75] R. R. Anderson and J. A. Parrish, “The optics of human skin,” *Journal of Investigative Dermatology*, vol. 77, no. 1, pp. 13–19, Jul. 1981. [Online]. Available: https://doi.org/10.1111/1523-1747.ep12479191

[76] Y. Sun and N. Thakor, “Photoplethysmography revisited: From contact to noncontact, from point to imaging,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 463–477, Mar. 2016. [Online]. Available: https://doi.org/10.1109/tbme.2015.2476337

[77] E. Tur, M. Tur, H. I. Maibach, and R. H. Guy, “Basal perfusion of the cutaneous microcirculation: Measurements as a function of anatomic position,” *Journal of Investigative Dermatology*, vol. 81, no. 5, pp. 442–446, Nov. 1983. [Online]. Available: https://doi.org/10.1111/1523-1747.ep12522619

[78] J. G. Webster, *Design of pulse oximeters*. Inst. of Physics Publ., 2003.

[79] Y. Maeda, M. Sekine, T. Tamura, A. Moriya, T. Suzuki, and K. Kameyama, “Comparison of reflected green light and infrared photoplethysmography,” in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, Aug. 2008. [Online]. Available: https://doi.org/10.1109/iembs.2008.4649649

[80] J. Lee, K. Matsumura, K. ichi Yamakoshi, P. Rolfe, S. Tanaka, and T. Yamakoshi, “Comparison between red, green and blue light reflection photoplethysmography for heart rate monitoring during motion,” in *2013 35th Annual International Conference of the IEEE*

Engineering in Medicine and Biology Society (EMBC). IEEE, Jul. 2013. [Online]. Available: https://doi.org/10.1109/embc.2013.6609852

[81] E. van Kampen and W. Zijlstra, "Determination of hemoglobin and its derivatives," in *Advances in Clinical Chemistry*. Elsevier, 1966, pp. 141–187. [Online]. Available: https://doi.org/10.1016/s0065-2423(08)60414-x

[82] S. Hoeksel, J. Jansen, J. Blom, and J. J. Schreuder, "Detection of dicrotic notch in arterial pressure signals," *Journal of clinical monitoring*, vol. 13, no. 5, pp. 309–316, 1997.

[83] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological measurement*, vol. 28, no. 3, p. R1, 2007.

[84] N. D. Giardino, P. M. Lehrer, and R. Edelberg, "Comparison of finger plethysmograph to ecg in the measurement of heart rate variability," 2002.

[85] I. CONSTANT, D. LAUDE, I. MURAT, and J.-L. ELGHOZI, "Pulse rate variability is not a surrogate for heart rate variability," *Clinical Science*, vol. 97, no. 4, pp. 391–397, Aug. 1999. [Online]. Available: https://doi.org/10.1042/cs0970391

[86] K. sang Yoo and W. hyung Lee, "Mental stress assessment based on pulse photoplethysmography," in *2011 IEEE 15th International Symposium on Consumer Electronics (ISCE)*. IEEE, Jun. 2011. [Online]. Available: https://doi.org/10.1109/isce.2011.5973841

[87] F. Li, P. Xu, S. Zheng, W. Chen, Y. Yan, S. Lu, and Z. Liu, "Photoplethysmography based psychological stress detection with pulse rate variability feature differences and elastic net," *International Journal of Distributed Sensor Networks*, vol. 14, no. 9, p. 155014771880329, Sep. 2018. [Online]. Available: https://doi.org/10.1177/1550147718803298

[88] J. Park, J. Kim, and S.-P. Kim, "Prediction of daily mental stress levels using a wearable photoplethysmography sensor," in *TENCON 2018 - 2018 IEEE Region 10 Conference*. IEEE, Oct. 2018. [Online]. Available: https://doi.org/10.1109/tencon.2018.8650109

[89] A. D. Wibawa, M. H. Purnomo, A. Marzuki, and L. D. Rumpa, "Physiological pattern of human state emotion based on ecg and pulse sensor." *Journal of Theoretical & Applied Information Technology*, vol. 93, no. 1, 2016.

[90] R. M. Sabour, Y. Benezeth, F. Marzani, K. Nakamura, R. Gomez, and F. Yang, "Emotional state classification using pulse rate variability," in *2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP)*. IEEE, Jul. 2019. [Online]. Available: https://doi.org/10.1109/siprocess.2019.8868781

[91] D. K. Utama, A. D. Wibawa, L. Hakim, and M. H. Purnomo, “Pulse rate pattern for gamers during playing video game in angry emotion,” in *2020 International Electronics Symposium (IES)*. IEEE, 2020, pp. 450–455.

[92] A. Clamor, M. M. Hartmann, U. Köther, C. Otte, S. Moritz, and T. M. Lincoln, “Altered autonomic arousal in psychosis: An analysis of vulnerability and specificity,” *Schizophrenia Research*, vol. 154, no. 1-3, pp. 73–78, Apr. 2014. [Online]. Available: https://doi.org/10.1016/j.schres.2014.02.006

[93] I. Liu, S. Ni, and K. Peng, “Predicting personality with smartphone cameras: A pilot study,” in *2020 IEEE International Conference on Human-Machine Systems (ICHMS)*. IEEE, Sep. 2020. [Online]. Available: https://doi.org/10.1109/ichms49158.2020.9209354

[94] ——, “Happiness at your fingertips: assessing mental health with smartphone photoplethysmogram-based heart rate variability analysis,” *Telemedicine and e-Health*, vol. 26, no. 12, pp. 1483–1491, 2020.

[95] P. V. Rouast, M. T. P. Adam, R. Chiong, D. Cornforth, and E. Lux, “Remote heart rate measurement using low-cost RGB face video: a technical literature review,” *Frontiers of Computer Science*, vol. 12, no. 5, pp. 858–872, Sep. 2018. [Online]. Available: https://doi.org/10.1007/s11704-016-6243-6

[96] S. Zaunseder, A. Trumpp, D. Wedekind, and H. Malberg, “Cardiovascular assessment by imaging photoplethysmography – a review,” *Biomedical Engineering / Biomedizinische Technik*, vol. 63, no. 5, pp. 617–634, Jun. 2018. [Online]. Available: https://doi.org/10.1515/bmt-2017-0119

[97] Z. Marcinkevics, U. Rubins, J. Zaharans, A. Miscuks, E. Urtane, and L. Ozolina-Moll, “Imaging photoplethysmography for clinical assessment of cutaneous microcirculation at two different depths,” *Journal of Biomedical Optics*, vol. 21, no. 3, p. 035005, Mar. 2016. [Online]. Available: https://doi.org/10.1117/1.jbo.21.3.035005

[98] A. V. Moco, S. Stuijk, and G. de Haan, “Ballistocardiographic artifacts in PPG imaging,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 9, pp. 1804–1811, Sep. 2016. [Online]. Available: https://doi.org/10.1109/tbme.2015.2502398

[99] M. J. Butler, J. A. Crowe, B. R. Hayes-Gill, and P. I. Rodmell, “Motion limitations of non-contact photoplethysmography due to the optical and topological properties of skin,” *Physiological Measurement*, vol. 37, no. 5, pp. N27–N37, Apr. 2016. [Online]. Available: https://doi.org/10.1088/0967-3334/37/5/n27

[100] D. J. McDuff, E. B. Blackford, and J. R. Estepp, “The impact of video compression on remote cardiac pulse measurement using imaging photoplethysmography,” in *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*. IEEE, 2017, pp. 63–70.

[101] J. R. Estepp, E. B. Blackford, and C. M. Meier, “Recovering pulse rate during motion artifact with a multi-imager array for non-contact imaging photoplethysmography,” in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2014, pp. 1462–1469.

[102] E. B. Blackford and J. R. Estepp, “Using consumer-grade devices for multi-imager non-contact imaging photoplethysmography,” in *Optical Diagnostics and Sensing XVII: Toward Point-of-Care Diagnostics*, vol. 10072. International Society for Optics and Photonics, 2017, p. 100720P.

[103] A. V. Moco, S. Stuijk, and G. De Haan, “Ballistocardiographic artifacts in ppg imaging,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 9, pp. 1804–1811, 2015.

[104] A. A. Kamshilin, V. Teplov, E. Nippolainen, S. Miridonov, and R. Giniatullin, “Variability of microcirculation detected by blood pulsation imaging,” *PloS one*, vol. 8, no. 2, p. e57117, 2013.

[105] A. V. Moço, S. Stuijk, and G. de Haan, “Motion robust ppg-imaging through color channel mapping,” *Biomedical optics express*, vol. 7, no. 5, pp. 1737–1754, 2016.

[106] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, vol. 1. IEEE, 2001, pp. I–I.

[107] G. Lempe, S. Zaunseder, T. Wirthgen, S. Zipser, and H. Malberg, “Roi selection for remote photoplethysmography,” in *Bildverarbeitung für die Medizin 2013*. Springer, 2013, pp. 99–103.

[108] M. Kumar, A. Veeraraghavan, and A. Sabharwal, “Distanceppg: Robust non-contact vital signs monitoring using a camera,” *Biomedical optics express*, vol. 6, no. 5, pp. 1565–1588, 2015.

[109] S. Kwon, J. Kim, D. Lee, and K. Park, “Roi analysis for remote photoplethysmography on facial video,” in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 4938–4941.

[110] M. Butler, J. A. Crowe, B. R. Hayes-Gill, and P. I. Rodmell, “Motion limitations of non-contact photoplethysmography due to the optical and topological properties of skin,” *Physiological measurement*, vol. 37, no. 5, p. N27, 2016.

[111] T. F. Cootes, G. J. Edwards, and C. J. Taylor, “Active appearance models,” in *European conference on computer vision*. Springer, 1998, pp. 484–498.

[112] X. Li, J. Chen, G. Zhao, and M. Pietikainen, “Remote heart rate measurement from face videos under realistic situations,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 4264–4271.

[113] D. McDuff, S. Gontarek, and R. W. Picard, “Remote detection of photoplethysmographic systolic and diastolic peaks using a digital camera,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 12, pp. 2948–2954, 2014.

[114] K.-Z. Lee, P.-C. Hung, and L.-W. Tsai, “Contact-free heart rate measurement using a camera,” in *2012 Ninth Conference on Computer and Robot Vision*. IEEE, 2012, pp. 147–152.

[115] F. Bousefsaf, C. Maaoui, and A. Pruski, “Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate,” *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 568–574, 2013.

[116] ——, “Automatic selection of webcam photoplethysmographic pixels based on lightness criteria,” *Journal of Medical and Biological Engineering*, vol. 37, no. 3, pp. 374–385, 2017.

[117] Y. Yang, C. Liu, H. Yu, D. Shao, F. Tsow, and N. Tao, “Motion robust remote photoplethysmography in cielab color space,” *Journal of biomedical optics*, vol. 21, no. 11, p. 117001, 2016.

[118] M.-Z. Poh, D. J. McDuff, and R. W. Picard, “Advancements in noncontact, multiparameter physiological measurements using a webcam,” *IEEE transactions on biomedical engineering*, vol. 58, no. 1, pp. 7–11, 2010.

[119] J.-S. Lee, K.-W. Lin, and J.-L. Syue, “Smartphone-based heart-rate measurement using facial images and a spatiotemporal alpha-trimmed mean filter,” *Technology and Health Care*, vol. 24, no. s2, pp. S777–S783, 2016.

[120] W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, “Amplitude-selective filtering for remote-ppg,” *Biomedical optics express*, vol. 8, no. 3, pp. 1965–1980, 2017.

[121] B.-F. Wu, P.-W. Huang, T.-Y. Tsou, T.-M. Lin, and M.-L. Chung, “Camera-based heart rate measurement using continuous wavelet transform,” in *2017 International Conference on System Science and Engineering (ICSSE)*. IEEE, 2017, pp. 7–11.

[122] R.-Y. Huang and L.-R. Dung, “Measurement of heart rate variability using off-the-shelf smart phones,” *Biomedical engineering online*, vol. 15, no. 1, pp. 1–16, 2016.

[123] L. Feng, L.-M. Po, X. Xu, and Y. Li, “Motion artifacts suppression for remote imaging photoplethysmography,” in *2014 19th International Conference on Digital Signal Processing*. IEEE, 2014, pp. 18–23.

[124] X. Sun, P. Yang, Y. Li, Z. Gao, and Y.-T. Zhang, "Robust heart beat detection from photoplethysmography interlaced with motion artifacts based on empirical mode decomposition," in *Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics*. IEEE, 2012, pp. 775–778.

[125] W. J. Jiang, S. C. Gao, P. Wittek, and L. Zhao, "Real-time quantifying heart beat rate from facial video recording on a smart phone using kalman filters," in *2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom)*. IEEE, 2014, pp. 393–396.

[126] J. Allen and A. Murray, "Effects of filtering on multisite photoplethysmography pulse waveform characteristics," in *Computers in Cardiology, 2004*. IEEE, 2004, pp. 485–488.

[127] G. Cennini, J. Arguel, K. Akşit, and A. van Leest, "Heart rate monitoring via remote photoplethysmography with motion artifacts reduction," *Optics Express*, vol. 18, no. 5, p. 4867, Feb. 2010. [Online]. Available: https://doi.org/10.1364/oe.18.004867

[128] S. Kwon, H. Kim, and K. S. Park, "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, Aug. 2012. [Online]. Available: https://doi.org/10.1109/embc.2012.6346392

[129] H. E. Tasli, A. Gudi, and M. den Uyl, "Remote PPG based vital sign measurement using adaptive facial regions," in *2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, Oct. 2014. [Online]. Available: https://doi.org/10.1109/icip.2014.7025282

[130] R. Irani, K. Nasrollahi, and T. B. Moeslund, "Improved pulse detection from head motions using dct," in *2014 international conference on computer vision theory and applications (VISAPP)*, vol. 3. IEEE, 2014, pp. 118–124.

[131] K. Yamamoto, K. Toyoda, and T. Ohtsuki, "Non-contact heartbeat detection by music with discrete cosine transform-based parameter adjustment," in *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2018, pp. 1–6.

[132] M. Fukunishi, D. Mcduff, and N. Tsumura, "Improvements in remote video based estimation of heart rate variability using the welch fft method," *Artificial Life and Robotics*, vol. 23, no. 1, pp. 15–22, 2018.

[133] Y.-P. Yu, B.-H. Kwan, C.-L. Lim, S.-L. Wong, and P. Raveendran, "Video-based heart rate measurement using short-time fourier transform," in *2013 International Symposium on Intelligent Signal Processing and Communication Systems*. IEEE, 2013, pp. 704–707.

[134] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. Clifton, and C. Pugh, “Non-contact video-based vital sign monitoring using ambient light and auto-regressive models,” *Physiological measurement*, vol. 35, no. 5, p. 807, 2014.

[135] P. Nooralishahi, C. K. Loo, and L. W. Shiung, “Robust remote heart rate estimation from multiple asynchronous noisy channels using autoregressive model with kalman filter,” *Biomedical Signal Processing and Control*, vol. 47, pp. 366–379, 2019.

[136] D. J. McDuff, E. B. Blackford, and J. R. Estepp, “Fusing partial camera signals for noncontact pulse rate variability measurement,” *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 8, pp. 1725–1739, 2017.

[137] D. E. King, “Dlib-ml: A machine learning toolkit,” *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.

[138] V. Kazemi and J. Sullivan, “One millisecond face alignment with an ensemble of regression trees,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Jun. 2014. [Online]. Available: https://doi.org/10.1109/cvpr.2014.241

[139] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, “300 faces in-the-wild challenge: database and results,” *Image and Vision Computing*, vol. 47, pp. 3–18, Mar. 2016. [Online]. Available: https://doi.org/10.1016/j.imavis.2016.01.002

[140] J. Ryu, S. Hong, S. Liang, S. Pak, Q. Chen, and S. Yan, “A measurement of illumination variation-resistant noncontact heart rate based on the combination of singular spectrum analysis and sub-band method,” *Computer Methods and Programs in Biomedicine*, vol. 200, p. 105824, Mar. 2021. [Online]. Available: https://doi.org/10.1016/j.cmpb.2020.105824

[141] E. Jones, T. Oliphant, P. Peterson *et al.*, “SciPy: Open source scientific tools for Python,” 2001–. [Online]. Available: http://www.scipy.org/

[142] E. J. Argüello-Prada, “The mountaineer’s method for peak detection in photoplethysmographic signals,” *Revista Facultad de Ingeniería Universidad de Antioquia*, no. 90, pp. 42–50, 2019.

[143] R. Takalo, H. Hytti, and H. Ihalainen, “Tutorial on univariate autoregressive spectral analysis,” *Journal of clinical monitoring and computing*, vol. 19, no. 6, pp. 401–410, 2005.

[144] A. Gudi, M. Bittner, and J. v. van Gemert, “Real-time webcam heart-rate and variability estimation with clean ground truth for evaluation,” *Applied Sciences*, vol. 10, no. 23, p. 8630, 2020.

[145] N. R. Lomb, “Least-squares frequency analysis of unequally spaced data,” *Astrophysics and space science*, vol. 39, no. 2, pp. 447–462, 1976.

[146] G. Ramos, M. Alfaras, and H. Gamboa, "Real-time approach to HRV analysis," in *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies*. SCITEPRESS - Science and Technology Publications, 2018. [Online]. Available: https://doi.org/10.5220/0006641402080215

[147] R. Champseix, "Heart rate variability analysis," https://github.com/Aura-healthcare/hrv-analysis, 2020.

[148] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, and J. Dubois, "Unsupervised skin tissue segmentation for remote photoplethysmography," *Pattern Recognition Letters*, vol. 124, pp. 82–90, Jun. 2019. [Online]. Available: https://doi.org/10.1016/j.patrec.2017.10.017

[149] R. Macwan, S. Bobbia, Y. Benezeth, J. Dubois, and A. Mansouri, "Periodic variance maximization using generalized eigenvalue decomposition applied to remote photoplethysmography estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 1332–1340.

[150] M. Artemyev, M. Churikova, M. Grinenko, and O. Perepelkina, "Robust algorithm for remote photoplethysmography in realistic conditions," *Digital Signal Processing*, vol. 104, p. 102737, 2020.

[151] R. Macwan, Y. Benezeth, and A. Mansouri, "Heart rate estimation using remote photoplethysmography with multi-objective optimization," *Biomedical Signal Processing and Control*, vol. 49, pp. 24–33, 2019.

[152] R. Song, S. Zhang, J. Cheng, C. Li, and X. Chen, "New insights on super-high resolution for video-based heart rate estimation with a semi-blind source separation method," *Computers in biology and medicine*, vol. 116, p. 103535, 2020.

[153] R. Song, H. Chen, J. Cheng, C. Li, Y. Liu, and X. Chen, "Pulsegan: Learning to generate realistic pulse waveforms in remote photoplethysmography," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 5, pp. 1373–1384, 2021.

[154] P. Li, Y. Benezeth, K. Nakamura, R. Gomez, C. Li, and F. Yang, "An improvement for video-based heart rate variability measurement," in *2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP)*. IEEE, 2019, pp. 435–439.

[155] H. P. Da Silva, J. Guerreiro, A. Lourenço, A. L. Fred, and R. Martins, "Bitalino: A novel hardware framework for physiological computing." in *PhyCS*, 2014, pp. 246–253.

[156] E. Peper, F. Shaffer, and I.-M. Lin, "Garbage in; garbage out—identify blood volume pulse (bvp) artifacts before analyzing and interpreting bvp, blood volume pulse amplitude, and heart rate/respiratory sinus arrhythmia data," *Biofeedback*, vol. 38, no. 1, pp. 19–23, 2010.

[157] F. Shaffer and D. C. Combatalade, “Don’t add or miss a beat: A guide to cleaner heart rate variability recordings,” *Biofeedback*, vol. 41, no. 3, pp. 121–130, 2013.

[158] A. Choi and H. Shin, “Photoplethysmography sampling frequency: pilot assessment of how low can we go to analyze pulse rate variability with reliability?” *Physiological Measurement*, vol. 38, no. 3, pp. 586–600, Feb. 2017. [Online]. Available: https://doi.org/10.1088/1361-6579/aa5efa

A

Code of Project

.

Listing A.1: Real time webcam PRV via rPPG python code

```
# Imports
import cv2
import dlib
import numpy as np
import time
import datetime
import bisect
from scipy import signal
from PyQt5 import QtGui
import pyqtgraph as pg
from spectrum import arma2psd, arcovar
```

```
from hrvanalysis import get_time_domain_features,
    get_frequency_domain_features

import auxiliar_functions
import drawing_tools

# Video input
video_capture = cv2.VideoCapture(0)
cap_fps = int(video_capture.get(5))
frame_width = int(video_capture.get(3))
frame_height = int(video_capture.get(4))
size = (frame_width, frame_height)
frame_count = 0

# Memory allocation for incoming real time signals
max_length_of_recording = 60 # min
max_num_of_frames = max_length_of_recording * 60 * cap_fps
resampled_color_signal = np.zeros((max_num_of_frames, 3), dtype=np.float64)
resampled_time = np.arange(0, max_length_of_recording * 60, 1 / cap_fps)
num_resampled_frames = 0
H = np.zeros((max_num_of_frames,), dtype=np.float64)

# Frontal face detector and shape predictor initialization:
# Set p as the path to your .dat file shape predictor
p = "C:\\Users\\pedro constantino\\Desktop\\rPPG_python\\tools\\dlib dat
    files\\shape_predictor_68_face_landmarks.dat"
detector = dlib.get_frontal_face_detector()
predictor = dlib.shape_predictor(p)
mask = np.zeros((frame_height, frame_width, 3), dtype=np.uint8) # Define ROI
    outterbound

# POS method initialization
window_time = 1.6 # 1.6 s
l = int(cap_fps * window_time)
last_frame_POS = 0

# Average color on three channels R, G, B
avg_rgb = np.zeros((max_num_of_frames, 3), dtype=np.float64)
```

```
frame_time_stamps = np.zeros((max_num_of_frames,), dtype=np.float64)
num_of_faces_stamps = np.zeros((max_num_of_frames,), dtype=np.float64)

# Filtration initialization
start_freq_bandpass = 0.8 # Hz
stop_bandpass_freq = 2.3 # Hz
nyquist_freq = cap_fps / 2
Bparam, Aparam = signal.iirfilter(2, [start_freq_bandpass/nyquist_freq,
    stop_bandpass_freq/nyquist_freq], btype='band',
                        analog=False, ftype='butter')
Z, P, K = signal.tf2zpk(Bparam, Aparam)
sos = signal.zpk2sos(Z, P, K)
z = np.zeros((sos.shape[0], 2))
iir_filtered_signal = np.zeros((max_num_of_frames,), dtype=np.float64)

# MMPD (Mountaineer's Method for Peak Detection)
mmpd_frame = 1
mmpd_num_upsteps = 0
mmpd_threshold = 6
mmpd_possible_peak = False
mmpd_possible_valley = False
upstep_possible_peak = 6
mmpd_peaks_signal = np.zeros((max_num_of_frames,), dtype=np.float64)

# Referee intervals
curr_pp_avg = 0
referee_initial_width = 0.3 # seconds
referee_half_width_frames = int((referee_initial_width * cap_fps) / 2)
max_pred_loca = 1000000000
min_pred_loca = 0
referee_boolean = np.zeros((max_num_of_frames,))
referee_p_peaks = []
referee_pp = []
referee_num_peaks = 0
referee_frame = 0

# Raw RR interval signal
raw_pp_signal = np.zeros((max_num_of_frames,), dtype=np.float64)
```

```
last_peak_time = 0
num_of_raw_pp_added = 0

# AR model (periodic)
ar_window_duration = 6 # s
ar_window_length = ar_window_duration * cap_fps
ar_frame = 0

# Estimate heart rate variability
# Time-domain
td_window_duration = 60 # s
td_window_length = int(td_window_duration * cap_fps)

# Plotting variables
qt_plot = QtGui.QApplication([])
win = pg.GraphicsWindow()
pos_plot = win.addPlot(title="POS method data")
pos_plot.setLabel('bottom', 'Time', 's')
pos_plot.setLabel('left', 'Pixel intensity')
pos_curve = pos_plot.plot()
win.nextRow()
filtered_plot = win.addPlot(title="Filtered POS signal and Peak detection")
filtered_plot.setLabel('bottom', 'Time', 's')
filtered_plot.setLabel('left', 'Pixel intensity')
filtered_curve = filtered_plot.plot()
estimated_peaks_curve = filtered_plot.plot()
referee_peaks_curve = filtered_plot.plot()
win.nextRow()
pp_plot = win.addPlot(title="PP signal")
pp_plot.setLabel('bottom', 'Time', 's')
pp_plot.setLabel('left', 'Time', 'ms')
raw_pp_curve = pp_plot.plot()
plot_time_window = 10 # seconds
plt_window_size =  plot_time_window * cap_fps

# FPS estimation variables
frames = 0
cur_fps = 0
```

```
122 last_time = datetime.datetime.now()
123
124 # main loop
125 while frame_count < max_num_of_frames:
126     ret, frame = video_capture.read()
127
128     frame_time_stamps[frame_count] = time.time()
129
130     frames += 1
131
132     # FPS estimation
133     delta_time = datetime.datetime.now() - last_time
134     elapsed_time = delta_time.total_seconds()
135
136     if (elapsed_time != 0):
137         cur_fps = np.around(frames / elapsed_time, 1)
138
139     # Convert frame to gray scale
140     gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
141
142     # Detect faces:
143     rects = detector(gray, 0)
144
145     # For one face
146     if(len(rects) > 0):
147         mask[:,:] = 0
148         shape = predictor(gray, rects[0]) # Get the shape using the predictor
                :
149         shape = auxiliar_functions.shape_to_np(shape) # Convert the shape to
                numpy array
150
151         for region in drawing_tools.roi_regions_of_interest: # Add ROI
                regions
152             cv2.fillConvexPoly(mask, shape[region], (255, 255, 255))
153
154         for region in drawing_tools.ocular_contours: # Remove ocular regions
155             cv2.fillConvexPoly(mask, shape[region], (0,0,0))
156
```

```
157     masked_face = cv2.bitwise_and(frame, mask) # apply the mask
158 
159     number_of_skin_pixels = np.sum(mask>0)
160 
161     # compute mean
162     r = np.sum(masked_face[:,:,2])/number_of_skin_pixels
163     g = np.sum(masked_face[:,:,1])/number_of_skin_pixels
164     b = np.sum(masked_face[:,:,0])/number_of_skin_pixels
165 
166     avg_rgb[frame_count] = np.array([r,g,b]) # if no face detected?
167 
168     frame = cv2.bitwise_and(frame, mask)
169 
170     # Resample color signal
171     if(frame_count > 0):
172         frame_range, interp_values = auxiliar_functions.resample_2
                readings_linear(frame_time_stamps[0], cap_fps, avg_rgb[
                frame_count - 1], frame_time_stamps[frame_count - 1], avg_rgb[
                frame_count], frame_time_stamps[frame_count])
173         resampled_color_signal[frame_range,:] = interp_values
174         num_resampled_frames = num_resampled_frames + frame_range.shape[0]
175 
176     # POS method + IIR filtration
177     if(num_resampled_frames > l):
178 
179         # POS method
180         start = last_frame_POS
181         stop = num_resampled_frames - l
182 
183         for t in range(start, stop):
184             #t = 0
185             # Step 1: Spatial averaging
186             C = resampled_color_signal[t:t+l-1,:].T
187 
188             #Step 2 : Temporal normalization
189             mean_color = np.mean(C, axis=1)
190             diag_mean_color = np.diag(mean_color)
191             diag_mean_color_inv = np.linalg.inv(diag_mean_color)
```

```
192             Cn = np.matmul(diag_mean_color_inv,C)
193
194             #Step 3: Projection to plane orthogonal to skin
195             projection_matrix = np.array([[0,1,-1],[-2,1,1]])
196             S = np.matmul(projection_matrix,Cn)
197
198             #Step 4: 2D signal to 1D signal (tuning)
199             std = np.array([1,np.std(S[0,:])/np.std(S[1,:])])
200             P = np.matmul(std,S)
201
202             #Step 5: Overlap-Adding
203             H[t:t+l-1] = H[t:t+l-1] +  (P-np.mean(P))/np.std(P)
204
205         last_frame_POS = stop
206
207         # IIR filtration
208         iir_filtered_signal[start:stop], z = signal.sosfilt(sos, -H[start:
                stop], zi=z)
209
210     # AR model (periodic)
211     if(last_frame_POS > ar_frame + ar_window_length):
212         ar_stop = ar_frame + ar_window_length
213         ar_values, error = arcovar(iir_filtered_signal[ar_frame:ar_stop], 12)
214         psd = arma2psd(ar_values, sides='centerdc', T=cap_fps)
215         psd = psd[int(len(psd) / 2):]
216         freq_vector = np.linspace(0, cap_fps / 2, len(psd))
217         cov_max_freq_idx = np.argmax(psd)
218         cov_max_freq = freq_vector[cov_max_freq_idx]
219         cov_max_freq_hz = round(cov_max_freq, 2)
220         cov_max_freq_bpm = round(cov_max_freq * 60, 2)
221         curr_pp_avg = 1 / cov_max_freq_hz
222
223         if ar_frame == 0:
224             curr_pp_avg = 1 / cov_max_freq_hz
225         else:
226             if abs(((1 / curr_pp_avg) * 60) - (cov_max_freq_hz * 60)) < 25: #
                    block changes in bpm bigger than 25 bpm
227                 curr_pp_avg = 1 / cov_max_freq_hz
```

```
228
229         ar_frame = ar_stop
230         print('PR estimate last ' + str(ar_window_duration) + 'sec = ' + str(
                cov_max_freq_bpm) + 'bpm')
231
232
233     # MMPPD peak detection
234     if(ar_frame > 0):
235
236         for i in range(mmpd_frame, last_frame_POS):
237
238             if(iir_filtered_signal[i] > iir_filtered_signal[i-1]):
239                 mmpd_num_upsteps = mmpd_num_upsteps + 1
240                 # Looking for the valley
241                 if(mmpd_possible_valley == False):
242                     mmpd_possible_valley = True # Potential valley has been
                            found
243                     value_possible_valley = iir_filtered_signal[i-1]
244                     time_possible_valley = resampled_time[i-1]
245             else:
246                 # Looking for the peak
247                 if(mmpd_num_upsteps > mmpd_threshold):
248                     mmpd_possible_peak = True # Potential peak has been found
249                     value_possible_peak = iir_filtered_signal[i-1]
250                     time_possible_peak = resampled_time[i-1]
251                     upstep_possible_peak = mmpd_num_upsteps
252                 else:
253                     if(mmpd_possible_valley == True):
254                         if(iir_filtered_signal[i] <= value_possible_valley):
255                             value_possible_valley = iir_filtered_signal[i] #
                                    Updating the potential valley
256                             time_possible_valley = resampled_time[i]
257
258                     # A peak has been found
259                     if(mmpd_possible_peak == True): # A peak has been found
260                         if(iir_filtered_signal[i-1] > value_possible_peak):
261                             value_peak = iir_filtered_signal[i-1]
262                             time_peak = resampled_time[i-1]
```

```
263                         else:
264                             value_peak = value_possible_peak
265                             time_peak = time_possible_peak
266 
267                         # Update peak detection variables
268                         time_peak_frame = int(time_peak * cap_fps)
269                         mmpd_peaks_signal[time_peak_frame] = value_peak
270 
271                         # A valley has been found
272                         if(mmpd_possible_valley == True):
273                             value_valley = value_possible_valley
274                             time_valley = time_possible_valley
275                             mmpd_possible_valley = False
276 
277                         mmpd_threshold = 0.37 * upstep_possible_peak #
                                Updating threshold
278                         mmpd_possible_peak = False
279 
280                 mmpd_num_upsteps = 0 # Resetting number of upsteps
281 
282         mmpd_frame = last_frame_POS
283 
284     # Referee
285     if(mmpd_frame > 1):
286         if(referee_num_peaks == 0):
287             # INITIALIZE
288             first_peaks_idx = np.nonzero(mmpd_peaks_signal[:last_frame_POS])[
                    0][0]
289             referee_p_peaks.append(first_peaks_idx / cap_fps)
290             referee_boolean[first_peaks_idx] = 1
291             referee_num_peaks += 1
292             next_peak_pred = int((curr_pp_avg + first_peaks_idx / cap_fps) *
                    cap_fps)
293             referee_half_width_frames = int((curr_pp_avg * cap_fps) / 2)
294             max_pred_loca = next_peak_pred + referee_half_width_frames + 1
295             min_pred_loca = next_peak_pred - referee_half_width_frames
296 
297         while(max_pred_loca < mmpd_frame + 4):
```

```
298
299             peaks_idx = np.nonzero(mmpd_peaks_signal[min_pred_loca:
                    max_pred_loca])[0] + min_pred_loca
300
301             if(peaks_idx.size == 0):
302                 referee_new_peak_index = int(((max_pred_loca + min_pred_loca)
                        / 2))
303             elif(peaks_idx.size < 2):
304                 referee_new_peak_index = peaks_idx[0]
305             else:
306                 peaks_distance_to_pred = peaks_idx - ((max_pred_loca +
                        min_pred_loca) / 2)
307                 peaks_distance_to_pred = np.abs(peaks_distance_to_pred)
308                 referee_new_peak_index = peaks_idx[np.argmin(
                        peaks_distance_to_pred)]
309
310             # Update peak detection variables
311             referee_p_peaks.append(referee_new_peak_index / cap_fps)
312             referee_boolean[referee_new_peak_index] = 1
313             referee_num_peaks += 1
314
315             referee_pp.append(referee_p_peaks[-1] - referee_p_peaks[-2])
316
317             # Make prediction
318             next_peak_pred = referee_new_peak_index + int(curr_pp_avg *
                    cap_fps)
319             referee_half_width_frames = int((curr_pp_avg * cap_fps) / 2)
320             max_pred_loca = next_peak_pred + referee_half_width_frames + 1
321             min_pred_loca = next_peak_pred - referee_half_width_frames
322
323             # Build raw RR signal, filter and downsample
324             if(referee_num_peaks - 1 > num_of_raw_pp_added):
325                 # Build raw RR signal
326                 raw_pp_signal[int(last_peak_time * cap_fps):int(
                        last_peak_time * cap_fps) + 60] = 1000 * referee_pp[-1] #
                         Convert to miliseconds
327                 num_of_raw_pp_added = num_of_raw_pp_added + 1
328                 last_peak_time = referee_p_peaks[-1]
```

```
329
330     # PLOTS
331     if(last_frame_POS > plt_window_size):
332         # Plot POS signal
333         pos_curve.setData(x=resampled_time[last_frame_POS - plt_window_size:
                last_frame_POS], y=-H[last_frame_POS - plt_window_size:
                last_frame_POS])
334
335         # Plot IIR fitration
336         filtered_curve.setData(x=resampled_time[last_frame_POS -
                plt_window_size:last_frame_POS], y=iir_filtered_signal[
                last_frame_POS - plt_window_size:last_frame_POS])
337
338         # Plot peak detection
339         estimated_peaks_curve.setData(x=resampled_time[last_frame_POS -
                plt_window_size:last_frame_POS], y = mmpd_peaks_signal[
                last_frame_POS - plt_window_size:last_frame_POS], pen=None,
                symbol='o', symbolPen=None, symbolSize=4, symbolBrush=('r'))
340         referee_peaks_curve.setData(x=resampled_time[last_frame_POS -
                plt_window_size:last_frame_POS], y = 25 * referee_boolean[
                last_frame_POS - plt_window_size:last_frame_POS], pen=None,
                symbol='o', symbolPen=None, symbolSize=4, symbolBrush=('y'))
341
342         # Plot PP signal
343         raw_pp_curve.setData(x=resampled_time[last_frame_POS -
                plt_window_size:last_frame_POS], y=raw_pp_signal[last_frame_POS -
                 plt_window_size:last_frame_POS], pen=(50,50,200))
344     else:
345         # Plot POS signal
346         pos_curve.setData(x=resampled_time[:last_frame_POS], y=-H[:
                last_frame_POS])
347
348         # Plot IIR fitration
349         filtered_curve.setData(x=resampled_time[:last_frame_POS], y=
                iir_filtered_signal[:last_frame_POS])
350
351         # Plot peak detection
```

```
352         estimated_peaks_curve.setData(x=resampled_time[:last_frame_POS], y =
                mmpd_peaks_signal[:last_frame_POS], pen=None, symbol='o',
                symbolPen=None, symbolSize=4, symbolBrush=('r'))
353         referee_peaks_curve.setData(x=resampled_time[:last_frame_POS], y = 25
                 * referee_boolean[:last_frame_POS], pen=None, symbol='o',
                symbolPen=None, symbolSize=4, symbolBrush=('y'))
354
355         # Plot PP signal
356         raw_pp_curve.setData(x=resampled_time[:last_frame_POS], y=
                raw_pp_signal[:last_frame_POS], pen=(50,50,200))
357     QtGui.QApplication.processEvents()
358
359     # PRINTS
360     # PR
361     if(len(referee_pp) > 1): # Long-term
362         pr_long_term = (1 / np.mean(referee_pp)) * 60
363         cv2.putText(frame, 'Long-term PR: ' + str(round(pr_long_term, 2)) + '
                 bpm', (5, 20), cv2.FONT_HERSHEY_SIMPLEX, 0.7, (0, 100, 255), 1,
                cv2.LINE_AA)
364     if(len(referee_pp) > 5): # Five-rr
365         pr_5 = (1 / np.mean(referee_pp[-5:])) * 60
366         cv2.putText(frame, 'Last-5-PP PR: ' + str(round(pr_5, 2)) + ' bpm', (
                5, 40), cv2.FONT_HERSHEY_SIMPLEX, 0.7, (100, 100, 255), 1, cv2.
                LINE_AA)
367
368     # PRV
369     if(last_frame_POS > td_window_length):
370         min_considered_peak = bisect.bisect_left(referee_p_peaks, (
                last_frame_POS - td_window_length) / cap_fps)
371         considered_intervals = [x * 1000 for x in referee_pp[
                min_considered_peak + 1:]]
372         td_feat = get_time_domain_features(considered_intervals)
373         cv2.putText(frame, 'SDNN_1min: ' + str(round(td_feat['sdnn'], 2)) + '
                 ms', (5, 60), cv2.FONT_HERSHEY_SIMPLEX, 0.7, (100, 100, 200), 1,
                 cv2.LINE_AA)
374         cv2.putText(frame, 'RMSSD_1min: ' + str(round(td_feat['rmssd'], 2)) +
                 ' ms', (5, 80), cv2.FONT_HERSHEY_SIMPLEX, 0.7, (150, 100, 200),
                1, cv2.LINE_AA)
```

```
375             fd_feat = get_frequency_domain_features(considered_intervals, method=
                    "lomb")
376             cv2.putText(frame, 'LF_1min: ' + str(round(fd_feat['lfnu'],3)) + ' nu
                    ', (5, 100), cv2.FONT_HERSHEY_SIMPLEX, 0.7, (255, 153, 204), 1,
                    cv2.LINE_AA)
377             cv2.putText(frame, 'HF_1min: ' + str(round(fd_feat['hfnu'],3)) + ' nu
                    ', (5, 120), cv2.FONT_HERSHEY_SIMPLEX, 0.7, (250, 148, 109), 1,
                    cv2.LINE_AA)
378             cv2.putText(frame, 'LF/HF_1min: ' + str(round(fd_feat['lf_hf_ratio'],
                     2)), (5, 140), cv2.FONT_HERSHEY_SIMPLEX, 0.7, (245, 143, 104), 1
                    , cv2.LINE_AA)
379
380         # Print estimated fps on the image
381         cv2.putText(frame, 'FPS: ' + str(round(cur_fps, 1)), (520, 20), cv2.
                FONT_HERSHEY_SIMPLEX, 0.7, (0, 255, 0), 1, cv2.LINE_AA)
382
383         # Display the resulting frame
384         cv2.imshow("Real-time pulse rate variability", frame)
385
386         frame_count += 1
387
388         # Press 'q' key to exit
389         if cv2.waitKey(1) & 0xFF == ord('q'):
390             break
391
392 # Release everything:
393 video_capture.release()
394 cv2.destroyAllWindows()
395 qt_plot.quit()
396 win.close()
```

Listing A.2: Drawing tools module

```
### Assign IDs to facial landmarks:
JAWLINE_POINTS = list(range(0, 17))
RIGHT_EYEBROW_POINTS = list(range(17, 22))
LEFT_EYEBROW_POINTS = list(range(22, 27))
NOSE_BRIDGE_POINTS = list(range(27, 31))
LOWER_NOSE_POINTS = list(range(31, 36))
RIGHT_EYE_POINTS = list(range(36, 42))
LEFT_EYE_POINTS = list(range(42, 48))
MOUTH_OUTLINE_POINTS = list(range(48, 61))
MOUTH_INNER_POINTS = list(range(61, 68))
ALL_POINTS = list(range(0, 68))
FOREHEAD_POINTS = list(range(68,72)) # predicted manually

### Define adding ROI regions
CENTRAL_FOREHEAD_CUT = [27,68,69]
RIGHT_FOREHEAD_CUT = [27,68,1]
LEFT_FOREHEAD_CUT = [27,69,15]
RIGHT_CHEEK_CUT = [1,3,31,27]
LEFT_CHEEK_CUT = [13,15,27,35]
NOSE_CUT = [27,31,35]
RIGHT_BEARD_CUT = list(range(3,9)) + [33,31]
LEFT_BEARD_CUT = list(range(8,14)) + [35,33]

### Define contours to be removed
RIGHT_OCULAR_CUT_CONTOUR = list(range(17,22)) + [39,40,41,36]
LEFT_OCULAR_CUT_CONTOUR = list(range(22,27)) + [45,46,47,42]

### Define ROI region of interest
ROI_OF_INTEREST = [1,3,31,33,35,13,15,69,68]

### Group ROI regions of interest
roi_regions = [CENTRAL_FOREHEAD_CUT,
               RIGHT_FOREHEAD_CUT,
               LEFT_FOREHEAD_CUT,
               RIGHT_CHEEK_CUT,
               LEFT_CHEEK_CUT,
               NOSE_CUT,
```

```
                RIGHT_BEARD_CUT,
                LEFT_BEARD_CUT]

roi_regions_of_interest = [CENTRAL_FOREHEAD_CUT,
                           RIGHT_FOREHEAD_CUT,
                           LEFT_FOREHEAD_CUT,
                           RIGHT_CHEEK_CUT,
                           LEFT_CHEEK_CUT,
                           NOSE_CUT]


### Group contours to remove
ocular_contours = [RIGHT_OCULAR_CUT_CONTOUR, LEFT_OCULAR_CUT_CONTOUR]
```

Listing A.3: Auxiliar functions module

```
### Imports
import numpy as np
import math

# Correction parameters
cheek_correction_ratio = 0.2
ocular_correction_ratio = 1.4
lower_forhead_ratio = 0.3
upper_forehead_ratio = 1.2

def shape_to_np(dlib_shape, dtype="int"):
    """Converts dlib shape object to numpy array"""

    # Initialize the list of (x,y) coordinates
    coordinates = np.zeros((dlib_shape.num_parts + 4, 2), dtype=dtype)

    # Loop over all facial landmarks and convert them to a tuple with (x,y)
        coordinates:
    for i in range(0, dlib_shape.num_parts):
        coordinates[i] = (dlib_shape.part(i).x, dlib_shape.part(i).y)

    # Add forehead points
    coordinates[dlib_shape.num_parts:] = predict_forehead_box(coordinates)

    # Correct cheek points (2 and 16 are corrected through 28; 4 and 14,
        through 32, 36 respectively)
    coordinates[1] = coordinates[1] + cheek_correction_ratio * (coordinates[2
        7] - coordinates[1]) # point 2
    coordinates[3] = coordinates[3] + cheek_correction_ratio * (coordinates[3
        1] - coordinates[3]) # point 4
    coordinates[13] = coordinates[13] + cheek_correction_ratio * (coordinates
        [35] - coordinates[13]) # point 14
    coordinates[15] = coordinates[15] + cheek_correction_ratio * (coordinates
        [27] - coordinates[15]) # point 16

    # Correct ocular regions points
    # Right ocular region ()
```

```
    right_ocular_region_points = list(range(17,22)) + [39,40,41,36]
    right_ocular_center = np.sum(coordinates[right_ocular_region_points],
        axis=0) / len(right_ocular_region_points)
    for r_ocular_point in right_ocular_region_points:
        coordinates[r_ocular_point] = right_ocular_center +
            ocular_correction_ratio * (coordinates[r_ocular_point] -
            right_ocular_center)

    # Left ocular region ()
    left_ocular_region_points = list(range(22,27)) + [45,46,47,42]
    left_ocular_center = np.sum(coordinates[left_ocular_region_points], axis=
        0) / len(left_ocular_region_points)
    for l_ocular_point in left_ocular_region_points:
        coordinates[l_ocular_point] = left_ocular_center +
            ocular_correction_ratio * (coordinates[l_ocular_point] -
            left_ocular_center)

    # Return the list of (x,y) coordinates:
    return coordinates


def predict_forehead_box(np_shape):
    x_19 = np_shape[18][0]
    y_19 = np_shape[18][1]
    x_26 = np_shape[25][0]
    y_26 = np_shape[25][1]
    x_29 = np_shape[28][0]
    y_29 = np_shape[28][1]
    x_31 = np_shape[30][0]
    y_31 = np_shape[30][1]

    x_nose_vector = x_29 - x_31
    y_nose_vector = y_29 - y_31

    # Mark predicted forehead points
    # FH1
    x_FH1 = int(x_19 + upper_forehead_ratio*x_nose_vector)
    y_FH1 = int(y_19 + upper_forehead_ratio*y_nose_vector)
```

```
64
65      # FH2
66      x_FH2 = int(x_26 + upper_forehead_ratio*x_nose_vector)
67      y_FH2 = int(y_26 + upper_forehead_ratio*y_nose_vector)
68
69      # FH3
70      x_FH3 = int(x_19 + lower_forhead_ratio*x_nose_vector)
71      y_FH3 = int(y_19 + lower_forhead_ratio*y_nose_vector)
72
73      # FH4
74      x_FH4 = int(x_26 + lower_forhead_ratio*x_nose_vector)
75      y_FH4 = int(y_26 + lower_forhead_ratio*y_nose_vector)
76
77      FH_points = np.array([(x_FH1, y_FH1), (x_FH2, y_FH2), (x_FH4, y_FH4), (
            x_FH3, y_FH3)], dtype=np.int32)
78
79      return FH_points
80
81  def resample_2readings_linear(ts0, resampling_rate, xi, tsi, xj, tsj):
82      ti = tsi - ts0
83      tj = tsj - ts0
84      slope = (xj - xi) / (tj - ti)
85      starting_frame = math.ceil(ti * resampling_rate)
86      end_frame = math.floor(tj * resampling_rate)
87      frame_range = np.arange(starting_frame, end_frame + 1)
88      values = np.array([xi + slope * ( x / resampling_rate - ti) for x in
            frame_range], dtype=np.float64)
89      return frame_range, values
```