

A Statistical Framework for Non-Contact Heart Rate Estimation via Photoplethysmogram Imaging

by

Brendan Chwyl

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2016

© Brendan Chwyl 2016

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contribution

One accepted conference paper and one published journal paper comprise the majority of this thesis. Brendan Chwyl (B.C.) was the major contributor and the first author on both publications. The parts of the thesis that have been co-authored, as well as co-author contributions, are detailed in the following paragraphs.

The first paper is titled *Time-frequency domain analysis via pulselets for non-contact heart rate estimation from remotely acquired photoplethysmograms* and was published in Volume 1 of Vision Letters. Sections 3.2, 3.3, 3.4 were copied verbatim and modified to add necessary detail and improve the flow of this thesis. In this paper, B.C. designed and implemented the algorithm. B.C., Audrey G. Chung (A.G.C.), Robert Amelard (R.A.), Jason Deglint (J.D.), and Alexander Wong (A.W.) performed data collection, and B.C., A.G.C., A.W., and David A. Clausi (D.A.C.) contributed to writing and editing the paper.

The second paper is titled *Remote heart rate measurement through broadband video via stochastic bayesian estimation* and was accepted to the 13th Conference on Computer and Robot Vision. Sections 3.5 and 4.1 were copied and modified to include added detail and improve overall flow. B.C., and A.W., designed and implemented the algorithm. B.C., A.G.C., R.A., J.D., and A.W. were responsible for data collection, and B.C., A.G.C., A.W., and D.A.C., contributed to writing and editing the paper.

Abstract

Although medical progress and increased health awareness over the last 60 years have reduced death rates from cardiovascular disease by more than 75%, cardiovascular disease remains one of the leading causes of death, hospitalization, and cause of prescription drug use. Resting heart rate can act as an independent risk factor in cardiovascular mortality, while more detailed blood volume waveforms can offer insight on blood pressure, blood oxygenation, respiration rate, and cognitive stress.

Electrocardiograms (ECGs) are widely used in the clinical setting due to their accurate measurement of heart rate and detailed capture of heart muscle depolarization, making them useful in diagnosis of specific cardiovascular conditions. However, the discomfort caused by the required adhesive patches, as well as the relatively high cost of ECG machines, introduces the need for an alternative system when only the resting heart rate is required. Photoplethysmography (PPG), the optical acquisition of blood volume pulse over time, offers one such solution. The pulse oximeter, a device which clips onto a thin extremity and measures the amount of transmitted light over time, is widely used in a clinical setting for heart rate and oxygen saturation measurements in cases where ECG is unnecessary or unavailable. Recently, a technique has been demonstrated to construct a blood volume pulse signal without the need for contact, offering a more sanitary and comfortable alternative to pulse oximetry. This technique relies on camera systems and is known as PPG imaging (PPGI). However, the accuracy of PPGI methods suffers in realistic environments with error incurred by motion, illumination variation, and natural fluctuation of the heart rate. For this reason, a statistical framework which aims to offer higher accuracy in realistic scenarios is proposed.

The initial step in the framework is to construct a PPG waveform, a time series correlated to hemoglobin concentration. Here, an importance-weighted Monte Carlo sampling strategy is used to construct a PPG waveform from many time series observations. Once the PPG waveform is established, a continuous wavelet transform is applied, using the so-called *pulselet* as the mother wavelet, to create a response map in the time-frequency domain. The average of frequencies corresponding to the maximum response over time is used as the heart rate estimation.

To verify the efficacy of the proposed framework, tests were run on two data sets; the first consists of broadband red-green-blue (RGB) colour channel video data and the second contains single channel near infrared video data. In the first case, improvements over state-of-the-art methods were shown, however; in the second case, no statistically significant improvement was observed.

Acknowledgements

I would like to thank my parents, Ed and Mary Chwyl, for their constant encouragement, my supervisors Alexander Wong and David Clausi, for their unwavering support and guidance, and the members of the Vision and Image Processing Group for adding companionship to the past few years.

Table of Contents

List of Tables	viii
List of Figures	ix
Nomenclature	xi
1 Introduction	1
1.1 Thesis Scope	2
1.2 Objectives	2
1.3 Contributions	2
2 Background	4
2.1 Heart Rate Analysis	4
2.2 Photoplethysmography Theory	5
2.3 Photoplethysmography Imaging	8
2.4 Photoplethysmography Imaging Waveform Construction	9
2.5 Photoplethysmography Waveform Analysis	12
2.6 Wavelet Theory	13
3 Methodology	19
3.1 Algorithm Overview	19
3.2 Sample Region Registration and Point Tracking	20

3.3	Skin Erythema Model	21
3.4	Photoplethysmogram Waveform Estimation	22
3.4.1	Bayesian Minimization	22
3.4.2	Posterior Probability Estimation	23
3.5	Photoplethysmogram Waveform Analysis	25
4	Experiments	27
4.1	Broadband Red-Green-Blue Channel Data Set	27
4.1.1	Experimental Setup	27
4.1.2	Results and Discussion	29
4.2	Near-Infrared Data Set	32
4.2.1	Experimental Setup	32
4.2.2	Results and Discussion	33
5	Conclusions	36
5.1	Key Observations	36
5.2	Concluding Remarks	36
5.3	Future Research and Recommendations	37
	References	39

List of Tables

4.1	Mean error, mean absolute error, and root mean squared error for the stationary and gradual motion categories of the RGB experiment.	29
4.2	Mean error, mean absolute error, and root mean squared error for all videos in the RGB experiment.	31
4.3	Mean error, mean absolute error, and root mean squared error for the NIR experiment.	33

List of Figures

1.1	A general flow chart for PPGI algorithms with the main contributions of this work identified in red.	3
2.1	Elements of the ECG waveform [8].	6
2.2	Example of a contact photoplethysmography system (a finger pulse oximeter).	7
2.3	Absorption spectra of oxygenated hemoglobin and deoxygenated hemoglobin in the visible and near-infrared wavelengths generated with data provided by Prahl [49]	8
2.4	Differences in PPG waveform shape between younger subjects (a) and older subjects (b) [33].	9
2.5	a) and b) show two different signals containing the same frequency content occurring at different time intervals. The magnitude of the Fourier transform for both signals is computed and plotted in c) and d) , respectively, showing identical results for each of the different signals.	15
2.6	A window with relatively poor time localization is shown in a) , while a window with relatively good time localization is shown in b) . The resulting frequency response of the Fourier transform performed on the windows indicated in a) and b) is shown in c)	16
2.7	A comparison of response maps generated for signal a) with the y-axis representing frequency and the x-axis representing time. b) is generated by the STFT using a small window, c) , is generated by the STFT using a large window, and c) is generated with the WT.	18

3.1	Flow chart of the proposed algorithm. Many skin erythema time series, $\psi(t)$, are used to produce an estimate of the PPG waveform, $\hat{\phi}(t)$. The contributions of each signal are weighted based on $\gamma(\Psi(u))$, which considers the Fourier transform of $\psi(t)$, $\Psi(u)$. The continuous wavelet transform is then applied to produce $ P(u, t) $, the magnitude of the wavelet response in the time-frequency domain.	20
3.2	Example of the final cheek region, outlined in green, from which the PPG waveform is constructed.	21
3.3	The pulselet in the time domain scaled to frequencies of a) 42 beats per minute, b) 140 beats per minute, c) 240 beats per minute, as well as d) , an example of a single period of a PPG waveform acquired through finger pulse oximetry [33].	26
4.1	Frame capture from a video in the tested data set. Ambient lighting conditions were used while a red LED was positioned behind the participant as a means of acquiring ground truth.	28
4.2	Pulselet response maps for two videos from the stationary category, a) and b) , and two videos from the gradual motion category, b) and c) , for a single participant. Response values are represented as a heat map, with the greatest responses indicated in bright yellow and the weakest responses indicated in dark blue.	30
4.3	The magnitude of the frequency response of a PPG waveform resulting from the FFT is shown in a) , while the response map of the same PPG waveform resulting from the CWT is shown in b) . The maximum peak of a) is 66.27 bpm, the average of maximum frequencies over time of b) is 75.36 bpm, and the ground truth heart rate is 76.50 bpm.	32
4.4	Geometric configuration of experiment. A participant is imaged via a coded hemodynamic imaging system with an applied tungsten-halogen illumination source. Ground truth is simultaneously collected using a finger pulse oximeter [5].	34

Nomenclature

\aleph	Spatial window size
α	Acceptance probability scaling parameter
γ	Acceptance probability
λ	Wavelength
μ	Pulselet center frequency
$\phi(t)$	Photoplethysmogram waveform
$\psi(t)$	Skin erythema transform (time domain)
$\Psi(u)$	Skin erythema transform (frequency domain)
σ	Time-frequency localization parameter
$\xi(t)$	Mother wavelet (time domain)
$\Xi(u)$	Mother wavelet (frequency domain)
ζ	Noise floor
a	Wavelet frequency scale parameter
b	Wavelet time shift parameter
L	Noise floor calculation window size
N	Number of observations
T	Total duration of constructed PPG waveform

t	Time
u	Frequency
X	Set of accepted observations
x	Sample point

Chapter 1

Introduction

Heart rate is the number of complete pulsations of the heart per unit time, most commonly measured in beats per minute. Resting heart rate has been shown to be an independent risk factor for cardiovascular mortality [26] and changes in resting heart rate over time may be indicative of a heart condition [1]. Despite significant medical advancements and fitness awareness efforts, cardiovascular disease remains the leading cause of death worldwide and is expected to account for as many as 23.6 million deaths per year by the year 2030 [1]. Fortunately, heart rate offers a simple metric to predict and monitor heart conditions. Currently, electrocardiograms (ECG) and pulse oximeters are used for clinical measurement of heart rate. ECG relies on the electrical measurement of heart depolarization and can be uncomfortable due to the required physical contact. In addition, the required equipment is relatively costly. Pulse oximetry offers a less expensive alternative and works by measuring the amount of light transmitted through an extremity, however the need for physical contact persists. A more sanitary, efficient, and comfortable alternative is remote heart rate measurement via imaging. This works in a similar manner as pulse oximetry, except rather than measuring the transmitted light passing through an extremity, the light reflected from the skin region is measured over time. This technique is known as photoplethysmography imaging (PPGI). While theoretically superior to contact measurements, PPGI systems are not yet widely adopted in clinical health care due to their relative novelty and lack of proven reliability.

1.1 Thesis Scope

This work is focused on the remote construction of clean measurements of blood volume over time, henceforth referred to as photoplethysmography (PPG) waveforms, in uncontrolled environments. This work also explores how PPG waveforms can be used to accurately estimate heart rate. As such, a framework is proposed with two main contributions:

- The PPG waveform is constructed via Bayesian minimization, with the required posterior probability estimated via an importance-weighted Monte Carlo method, and is presented in Chapter 3.4 [18]. This serves to increase robustness to gradual motion and illumination variation to maintain consistent measurements and improve accuracy.
- A continuous wavelet transform, performed using the so-called *pulselet* as the mother wavelet, is used to analyze the PPG waveform. This technique is presented in Chapter 3.5 [17] and allows for PPG waveform analysis in the time-frequency domain, mitigating error caused by natural heart rate variation which is experienced by typical Fourier domain analysis methods.

1.2 Objectives

There are three primary objectives surrounding the design and implementation of a framework for improved PPGI. The first objective is to improve the accuracy using PPGI in uncontrolled environments e.g., in the presence of motion or lighting artifacts. The second objective is to ensure a low cost solution by using commercially available, off-the-shelf digital cameras. The third objective of this framework is to demonstrate that it can be extended to narrow band methods (e.g., single channel infrared) to accommodate existing hardware set-ups.

1.3 Contributions

While the specifics of each step differ across all PPGI methods, every PPGI method follows the same general outline. First, the sample region, typically an area of exposed skin, is located within the initial video. This region is either tracked throughout the video sequence or assumed to be static. At each new frame, an additional point of the PPG waveform is

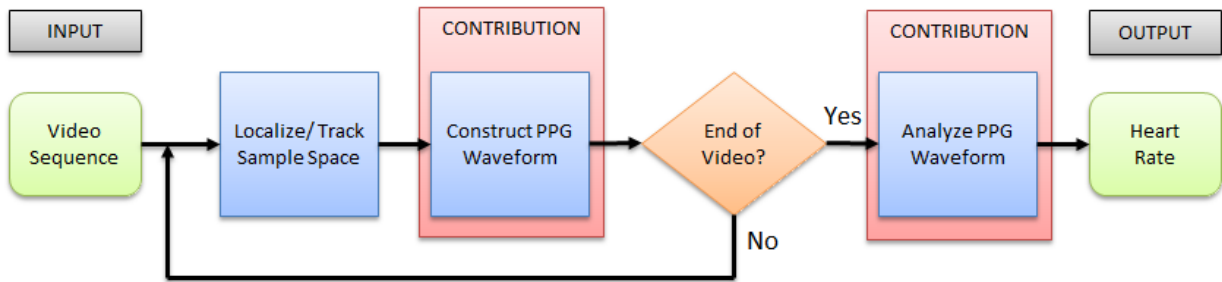


Figure 1.1: A general flow chart for PPGI algorithms with the main contributions of this work identified in red.

calculated and appended. Once the video sequence ends or the amount of data is deemed sufficient, the waveform is analyzed to produce a heart rate estimate.

In this work, a novel method for PPG waveform construction is introduced in which Bayesian minimization is used to produce a clean signal from a set of skin erythema observations, with the required posterior probability estimated with an importance-weighted Monte Carlo method. A novel PPG waveform analysis method is also introduced in which the continuous wavelet transform is used to produce a response map in the time-frequency domain. A general flow chart can be seen in Fig 1.1, with particular contributions in red.

Chapter 2

Background

In this chapter, current state-of-the-art for heart rate measurement and publications relating to the two main thesis contributions are described. The theory behind photoplethysmography is detailed in Section 2.2. Papers relating to PPG waveform construction are described in Section 2.4 while papers relating to PPG waveform analysis are detailed in Section 2.5. Finally, the theory and intuition behind the wavelet transform is explained in Section 2.6.

2.1 Heart Rate Analysis

The current gold standard for heart rate measurement in clinical applications is the electrocardiogram (ECG) [35]. ECGs rely on electrodes, which are placed on the subject's body with adhesive, and measure the depolarization in the heart muscle accompanying each heart beat. The resulting waveform contains components representative of the P-wave (atrial depolarization), QRS-wave (ventricles depolarization), and the T-wave and the U-wave (ventricular repolarization) [10]. An example of such a waveform can be seen in Fig. 2.1. The multiple phases of depolarization are captured in great detail, helping with accurate diagnosis, monitoring, and prediction of a multitude of cardiovascular conditions, including the diagnosis of congenital heart disease in adults [36] and the prediction of ventricular arrhythmias in patients undergoing implantable cardioverter-defibrillators therapy [11]. While invaluable for clinical scenarios due to its high accuracy and detail, the ECG relies on contact measurements, making it uncomfortable and impractical for extended monitoring. In addition, ECG machines are relatively expensive, averaging a cost

of roughly \$2,200 [12]¹. As such, ECG machines are largely inaccessible for personal use and impractical when only a measurement of the heart rate is required.

Photoplethysmography [32] (PPG) offers a non-invasive and cost effective means of measuring the change in blood volume over time. PPG methods rely on the absorption properties of hemoglobin, a protein found in red blood cells that is indicative of blood volume, to optically measure the change in blood volume over time, resulting in a PPG waveform. The current gold standard in PPG heart rate measurements are Pulse oximeters [7], devices which shine a light through a thin extremity (e.g., finger, ear lobe) and detect the intensity of transmitted light. The intensity of transmitted light is changed when blood is present in the extremity, and a time series correlated to blood pulse is constructed. Because of their relatively low cost, averaging \$179.99 [50]², pulse oximeters are widely used for personal health care or medical screening procedures. However, the need for physical contact makes these devices unsuitable for monitoring subjects for lengthy periods of time, or unsuitable for patients with wounds or burns.

Recently, photoplethysmography imaging (PPGI) has become a popular area of research. PPGI employs camera based systems to allow for the non-contact measurement of heart rate and offers a more sanitary, more efficient, and less obtrusive alternative to ECG and PPG methods. The fundamental theory behind PPG methods is explored in detail in Section 2.2.

2.2 Photoplethysmography Theory

Photoplethysmography is the optical construction of a plethysmogram, a volumetric measure of an organ as it varies through time. In the context of this thesis, plethysmograms refer to the volumetric measure of blood volume over time, such that heart rate may be inferred. Photoplethysmography was first introduced by Challoner and Ramsay [14] when they sought to design a device capable of measuring cutaneous blood flow non-invasively at any skin region in real time. The most basic PPG systems consist of only two main components: a light source to illuminate the skin region and a photodetector to measure the intensity of transmitted or reflected light as well as emitted light in the case of infrared systems. Skin tissue and bone, as well as the presence of blood, attenuates a proportion of the applied light due to multiple scatterings, reflection, and absorption [6]. If measured at an appropriate wavelength, the attenuation caused by skin tissue and bone is small or

¹Indicated price is averaged across 50 different ECG machine models.

²Indicated price is averaged across 10 different pulse oximeter models

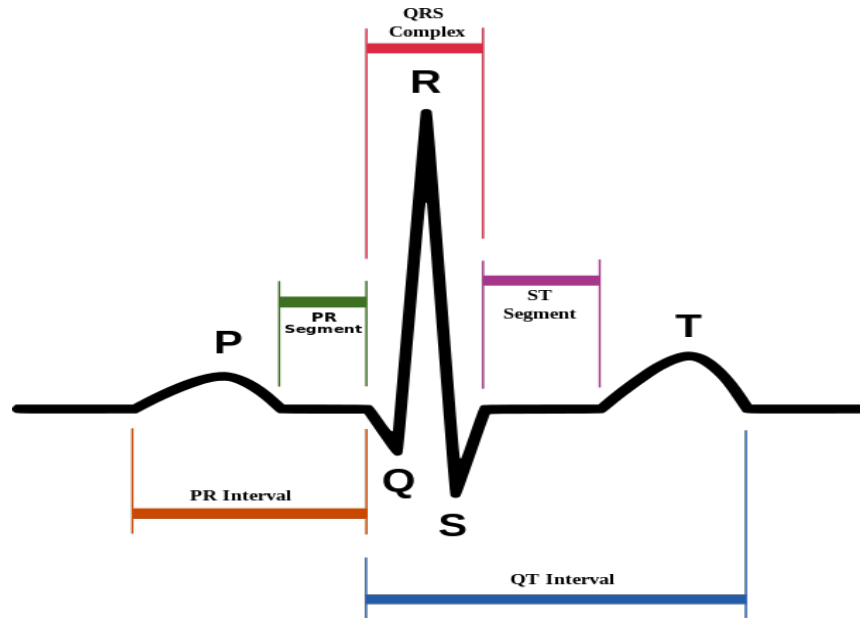


Figure 2.1: Elements of the ECG waveform [8].

constant, allowing the measurement of blood content by measuring the amount of attenuated reflected or transmitted light. An example of a contact PPG system can be seen in Fig. 2.2.

The selection of wavelength to be measured is an important factor for accurate estimation of heart rate via PPG. Hemoglobin, a protein responsible for transporting oxygen in blood, plays an important role in selecting an appropriate wavelength for heart rate estimation. More specifically, choosing wavelengths which correspond to high levels of absorption results in larger attenuation of the measured light, subsequently increasing the magnitude of blood volume measurements. However, hemoglobin absorption varies with oxygenation, introducing the possibility that attenuation may occur from two sources: the concentration of hemoglobin and the oxygenation of hemoglobin. A figure displaying the absorption spectra for both oxygenated and deoxygenated hemoglobin can be seen in Fig. 2.3. Many past works [14, 7] measure light at a wavelength of 805nm, an isosbestic point for hemoglobin. At this wavelength, oxygenated and deoxygenated hemoglobin have the same absorption coefficient, thus removing the impact of blood oxygenation on hemoglobin absorption. This subsequently leaves hemoglobin concentration as the only variable altering the level of light attenuation. While many methods measure a wavelength of 805nm, isosbestic points also exist at shorter wavelengths [52], allowing for PPG systems operating in the

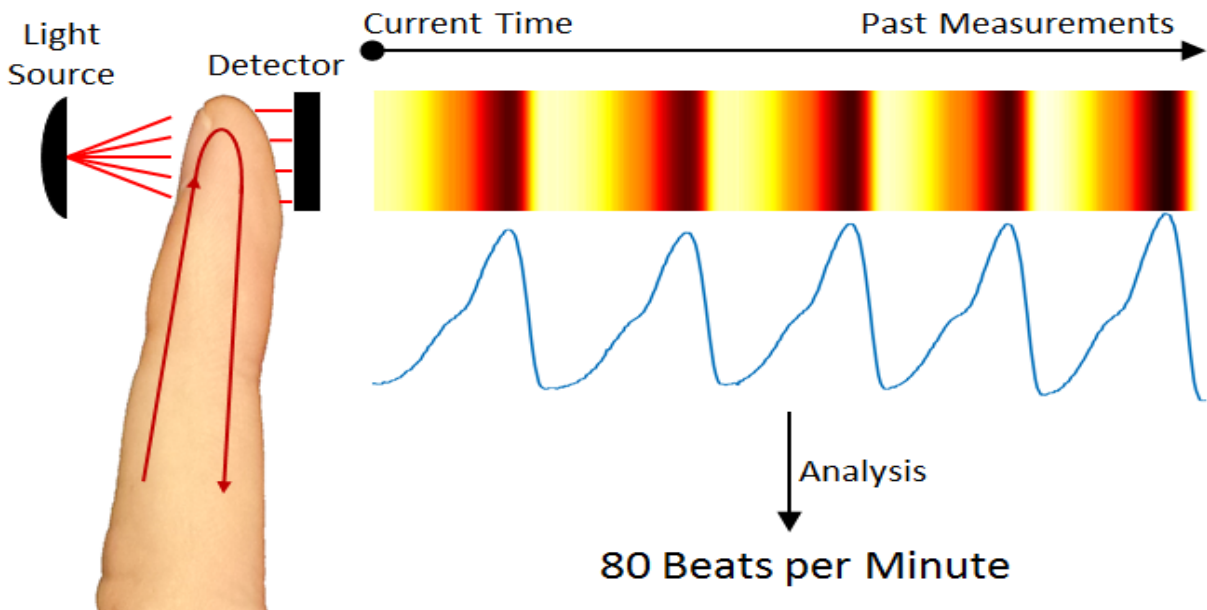


Figure 2.2: Example of a contact photoplethysmography system (a finger pulse oximeter).

visible spectrum.

The waveform captured by a PPG system consists of two main components: a pulsatile, or 'AC', component, and a quasi-DC component [4]. The pulsatile component has a fundamental frequency which corresponds to the average heart rate and is superimposed onto the quasi-DC component, a slowly varying signal caused by illumination changes, motion artefacts, respiratory activity, vasoconstrictor waves, Traube Hering Mayer (THM) waves, and thermoregulation.

In younger subjects (aged 20 to 40), PPG waveforms typically consist of a prominent initial peak, a valley, and a second weaker peak. Conversely, in older subjects (aged 40 to 60), PPG waveforms typically lack a distinct second peak and instead show a more gradual slope in its place [33]. An example of each category is shown in Fig. 2.4. Further still, PPG waveforms can vary based on the location at which it was measured. For example, a PPG waveform acquired from a subject's finger will differ from a PPG waveform acquired from a subject's ear lobe [3]. While PPG waveforms lack the detail of ECG waveform, numerous clinical applications of PPG exist [4] (e.g., blood oxygen saturation, blood pressure, respiration rate).

Of particular interest in this work is the analysis of a PPG waveform through time in order to infer heart rate. In such applications, the fundamental frequency of the PPG

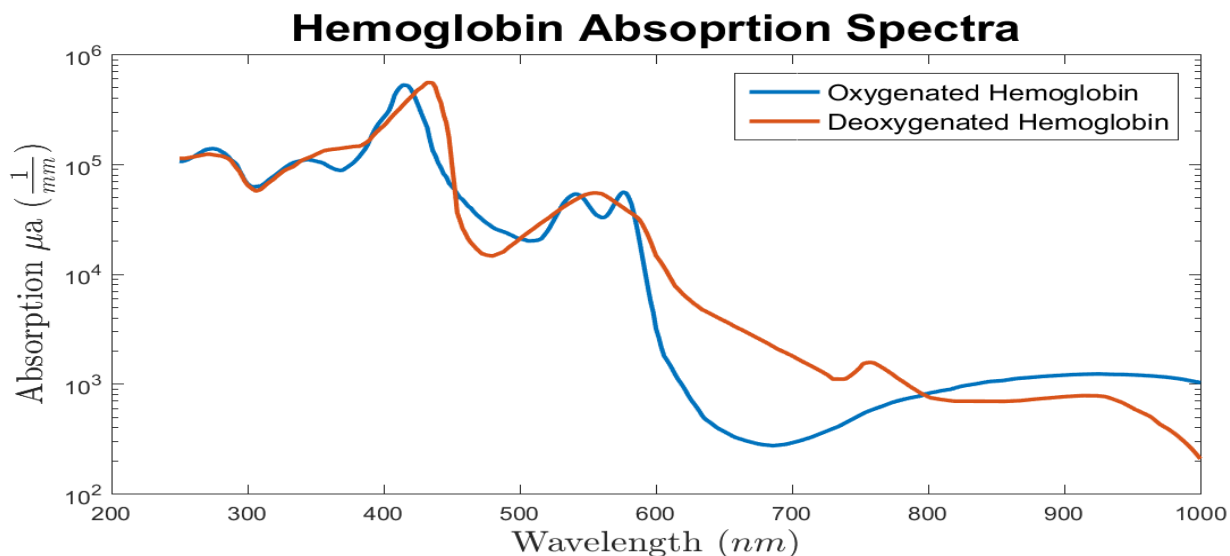


Figure 2.3: Absorption spectra of oxygenated hemoglobin and deoxygenated hemoglobin in the visible and near-infrared wavelengths generated with data provided by Prahl [49]

waveform must be estimated through the analysis of the time domain, frequency domain, or time-frequency domain. A more in depth explanation of the various PPG waveform analysis techniques is provided in Section 2.5.

2.3 Photoplethysmography Imaging

Photoplethysmography Imaging (PPGI) has recently gained traction as a non-contact means of acquiring a PPG waveform by measuring the light reflected off of a subject's skin throughout a sequence of images. PPGI systems offer a more sanitary and less obtrusive means of constructing a PPG waveform than contact PPG methods. In addition, PPGI facilitates novel applications such as telemedicine, affective computing, and monitoring subjects in motion. However, PPGI introduces new difficulties into the process of constructing and analyzing PPG waveforms.

The first challenge that is introduced by PPGI is the need to determine the regions which contain relevant data. Unlike the case of contact PPG, the entirety of each acquired sensor reading, in this case an image, is not necessarily useful for determining blood volume. In fact, only the skin regions within each image are useful and must be determined.

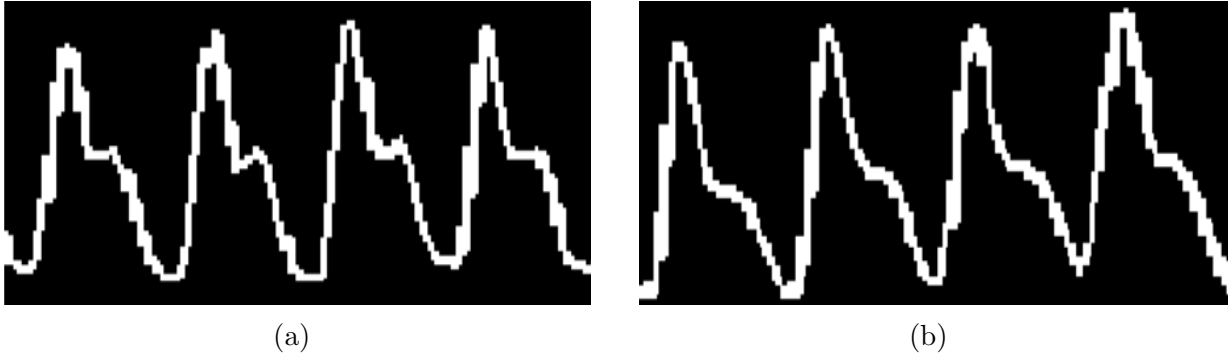


Figure 2.4: Differences in PPG waveform shape between younger subjects (a) and older subjects (b) [33].

Secondly, motion artefacts, while also capable of corrupting contact PPG waveforms, pose a more complex problem in PPGI system. This problem is introduced by the fact that subjects are always in motion, whether it is caused by a spontaneous reaction (e.g., smiling or laughing) or by a deliberate action (e.g., moving to another location). Even when consciously attempting to sit still motion is introduced via blood flow or respiration. In the construction of a PPG waveform, it is important to take measurements at a consistent location to accurately observe the fluctuation in blood volume over time. However, once the relevant regions are identified, it is not a trivial task to accurately track them over time.

Lastly, because PPGI systems measure the magnitude of reflected light, illumination levels and illumination variation greatly affects the quality of a PPG waveform. The strength of a reflected signal is dependant on the magnitude of illumination, and if the illumination level is too low, the acquired signal can be dominated by inherent camera noise. In addition, illumination variation can introduce false peaks and troughs in collected data, leading to invalid heart rate measurements.

2.4 Photoplethysmography Imaging Waveform Construction

Many PPGI methods rely on active illumination and optical filters to construct a PPG waveform. These methods aim to leverage the absorption properties of hemoglobin by applying and filtering light at certain wavelengths, as well as increasing the magnitude of

reflected light in an effort to lessen the effects of camera noise.

Cennini *et al.* [13] proposed a system in which two photodiodes are housed in a cylindrical tube, each behind a different optical filter, having cut-off wavelengths of $\lambda = 480\text{nm}$ and $\lambda = 700\text{nm}$, respectively. This device was used to collect reflected light, generated by two types of LEDs ($\lambda = 970\text{nm}$ and $\lambda = 450\text{nm}$), from the palm of a subject’s hand. Wieringa *et al.* [60] have also proposed a system which constructs a PPG waveform from a subject’s palm. It relies on a 3-wavelength LED-ringlight ($\lambda = 660\text{nm}$, $\lambda = 810\text{nm}$, $\lambda = 940\text{nm}$) to actively illuminate the scene while a modified monochrome CMOS camera is used to capture the signal. Sun *et al.* [53] proposed another system which constructs a PPG waveform from measured reflectance of the palm. This system actively illuminated the subject’s palm with two infrared light sources (both with $\lambda = 880\text{nm}$) and used a monochrome CMOS camera with a spectral range of $\lambda = [500 - 1000]\text{nm}$ to measure reflectance and encoded results with 10 bits in order to capture minute fluctuations in reflected light.

Van Gastel *et al.* [56] explore two camera configurations in an effort to create a motion robust PPGI system. The first setup consists of three monochrome cameras, each using a separate optical filter with center wavelengths at $\lambda_1 = 675\text{nm}$, $\lambda_2 = 800\text{nm}$, and $\lambda_3 = 842\text{nm}$. The PPG waveform is then constructed as a linear combination of the average pixel values within a manually defined face region for each camera. The second setup uses a single red-green-blue (RGB) camera in which the near-infrared (NIR) filter was replaced with a filter designed to occlude the visible light spectrum ($\lambda = [400 - 700]\text{nm}$). The PPG waveform was created as a linear combination of the average NIR response of the red, green, and blue sensors within a manually determined face region.

While these methods are capable of producing relatively clean PPG waveforms closely correlated to those acquired via pulse oximeters, such systems require custom hardware and involved configurations. Cheaper, more convenient, and more widely accessible alternatives have been explored by creating the PPG waveform from commercially available broadband RGB cameras under ambient illumination, alleviating the need for active illumination.

Poh *et al.* [46] proposed a method for remote heart rate estimation from broadband RGB video in which traces of the average red, green, and blue channels within an automatically determined face bounding box were created over time. Independent component analysis (ICA) was performed to decompose the raw RGB traces into three independent source signals. For the sake of automation, the second source signal was always used as the PPG waveform. However, the PPG waveform produced by this method is subject to the random order of which the ICA algorithm returns the separated source signals, which can potentially lead to selecting a waveform with no relevant data. In addition, by using the

entire face region, spontaneous motion such as smiling and blinking risk corrupting the signal.

Poh *et al.* [47] extend upon their previous work to improve the analyzed PPG waveform in two main ways. First, moving average filters were applied to each of the raw RGB traces and detrending on the independent source signals was performed. In addition, the final PPG waveform selection was better automated by selecting the independent source signal with the highest magnitude frequency response. While the issue of mistakenly selecting the wrong separated source signal is largely addressed, the encapsulation of irrelevant data in the waveform remains a potential source of error. In addition, the introduction of smoothing and detrending filters risks suppressing relevant signal information.

Li *et al.* [42] aim to further improve upon these works by more precisely defining the skin region, better accounting for fluctuations in ambient lighting conditions, and removing motion artefacts. Initially, a skin region mask is registered from automatically detected facial land marks. Within the skin region, the average green channel is averaged at each time step to create an initial PPG waveform. To compensate for ambient illumination changes, another average green channel time series is created from the background region, determined via a distance regularized level set evolution (DRLSE) algorithm [41], and subtracted from the initial PPG waveform. This illumination compensated waveform is further compensated for motion via non-rigid motion elimination, a moving average filter, and detrending techniques. While this method compensates for most sources of PPG waveform noise, the detrending and smoothing filters risk removing or suppressing frequency data necessary for accurate heart estimation. In addition, this method relies on a signal generated from the scene’s background to compensate for illumination, however, dynamic backgrounds or incorrect segmentation can introduce errors in this signal and propagate to the final PPG waveform.

Rather than averaging across entire skin regions, Chung *et al.* [16] proposed a method which creates a PPG waveform from few selected points. Each point is registered automatically and tracked throughout the video sequence. The PPG waveform itself is made based on the average skin erythema, a measure of skin redness which is closely correlated to hemoglobin concentration, at each point. While a clean PPG waveform is typically created, the contribution from each point is evenly weighted. Due to the small number of points (typically 10 were used), a single point with noisy data can impact the final waveform.

While most heart rate estimation methods explicitly use the principals of PPG to construct a PPG waveform, some methods rely on the motion incurred by the head as heart pulses to the brain. Balakrishnan *et al.* [9] proposed such a method. The vertical motion of face regions are tracked throughout the video sequence and the variations in

y -coordinate of each region are used to create a PPG waveform³. Because this method does not rely on imaging exposed skin, this method is robust to situations in which the face is covered or not visible, however, this method relies on vertical motion, it is very susceptible to motion noise.

2.5 Photoplethysmography Waveform Analysis

Once a PPG waveform has been constructed, one of many different techniques is used to infer heart rate. The benefits and drawbacks of each of these techniques is detailed in the following section.

Time domain analysis of a PPG waveform is typically performed by measuring the time between consecutive pulses, a quantity known as the inter-beat interval (IBI). While possible to achieve a heart rate estimate via counting the number of peaks over time, this technique is typically avoided as the error is heavily influenced by the window size and number of beats. The inverse of IBI between any two peaks is roughly the heart rate frequency, however, this relation experiences variability due to natural heart rate fluctuations and waveform noise. Poh *et al.* [47] apply a proprietary peak detection method to their PPG waveform and use an average of all IBIs to estimate heart rate. Balakrishnan *et al.* [9] first detect peaks by segmenting the PPG waveform into equally sized windows and classify the maximum value in each of the resulting windows as a peak. The average IBI for each peak is then calculated, from which a heart rate estimate is produced. While these techniques have been shown to produce good results, is robust to heart rate variation throughout time, and can be easily extended to estimate heart rate variability (HRV), they can incur large error from incorrectly detected and undetected peaks. This makes them very susceptible to any noise, caused by motion artefacts and illumination variation, that is present in the processed waveform.

Analyzing the frequency content of the PPG waveform is another common method for estimating heart rate. Poh *et al.* [46], applies a 128-point Hamming bandpass filter ([0.7 – 4.0] Hz) to their PPG waveform and then applies the fast Fourier transform (FFT). The frequency corresponding to the maximum response is selected as the heart rate frequency. Chung *et al.* [16] also transform the PPG waveform into the frequency domain via the FFT. This method selects the frequency corresponding to the maximum peak within an operating interval of [40 – 100] beats per minute as the heart rate frequency. Li *et al.* [42] apply a Hamming bandpass filter to the PPG waveform and analyze its power spectral

³Not constructed via PPG, but referred to as a PPG waveform as it is used in heart rate estimation.

density distribution using Welch’s method [59], an estimate of power as a function of frequency. The frequency with the maximum power response is then selected as the heart rate frequency. While trials on subjects in a constrained environment produce good results, frequency analysis methods can incur error in two main ways. First, if filter parameters are not carefully selected, it is possible to suppress frequency content relevant to the estimation of heart rate. Second, the PPG waveform is not strictly periodic as it changes with physical activity, respiratory rate, and cognitive stress. Subsequently, the frequency domain may not contain a single distinct peak, but rather the response may be spread of across multiple frequency bins, causing the inherently high magnitude low frequency content to be more prominent than the actual heart heart frequency.

Techniques have been explored to compensate for heart rate variation over time. De Haana and ven Leest [22] iteratively perform the Fourier transform within a sliding window and create a spectrogram of frequency response over time, a technique known as the short time Fourier transform (STFT). This allows for better analysis of quasi-periodic waveforms than the widely used Fourier transform by offering insight into how the frequency of a signal varies through time. However, the window size is an important consideration in this method and must be carefully selected; if it is too small, then high frequency components may be missed, whereas if it is too large, the time localization is poor. Many methods make use of the wavelet transform (WT), which offers a similar effect as the STFT without the need for a precisely defined window size. Both the STFT and WT are described further in Section 2.6. Lee *et al.* [40] and Fu *et al.* [27] both use wavelets to filter the acquired PPG waveforms prior to analysis in the time domain. Lee *et al.* [40] reduced motion artefacts captured by a finger pulse oximeter, leading to more accurate heart rate estimates. Fu *et al.* [27] also performed wavelet filtering on waveforms acquired from a finger pulse oximeter and showed an improved accuracy over using a moving-average filter approach. The multi-resolution capabilities of wavelet filtering versus the single window size of the moving-average filter was noted as a main contributor to the improved results.

2.6 Wavelet Theory

While the Fourier transform allows a signal to be decomposed into a weighted sum of sinusoidal functions of varying frequency and phase, time localization is not possible. Consider a signal comprised of two sinusoids at adjacent intervals: $y_1(t) = \sin(\omega_1 t)|_{t=a}^{t=b} + \sin(\omega_2 t)|_{t=c}^{t=d}$. The magnitude of the Fourier transform of $y_1(t)$, $\mathfrak{F}(y_1(t))$, results in two distinct peaks representative of ω_1 and ω_2 . Now consider a second signal in which the order of the two sinusoids is reversed: $y_2(t) = \sin(\omega_2 t)|_{t=a}^{t=b} + \sin(\omega_1 t)|_{t=c}^{t=d}$. The magnitude of the Fourier

transform of y_2 , $\mathfrak{F}(y_2(t))$, will produce the same magnitude as the Fourier transform of y_1 . An example of such a situation is visible in Fig 2.5.

Time localization can be achieved by applying the Fourier transform within a fixed-sized sliding window, a technique known as the short time Fourier transform (STFT). By decreasing the size of the windowed Fourier transform, more accurate time localization can be achieved; however, there exists a fundamental trade off between frequency resolution and time resolution. As the time resolution is increased the frequency resolution is decreased and visa versa. An example of this trade off is visible in Fig.2.6. The window size in the STFT is therefore an important consideration as this parameter is fixed throughout the transform, limiting the STFT to a single resolution.

Another limitation of the Fourier transform is that it is restricted to a sinusoidal basis function, restricting its ability to analyze non-stationary signals. However, in the context of biomedical signal processing, this is insufficient as many waveforms, including the PPG waveform, are non-stationary. In addition, biomedical signals typically contain transient patterns that carry important information [2].

The wavelet transform (WT) offers an alternative to the STFT and is capable of time-frequency analysis at multiple resolutions while not depending on a sinusoidal basis function [20]. This allows for more suitable analysis of non-stationary signals with better resolution at high frequencies and better time resolution at low frequencies. The wavelet transform relies on a family of functions, defined as

$$\xi_{a,b}(t) = \frac{1}{\sqrt{a}}\xi\left(\frac{t-b}{a}\right), a > 0, b \in \mathbb{R}, \quad (2.1)$$

where ξ represents the *mother wavelet*, a fixed function which is scaled by the scale parameter, a , and shifted through time by the translation parameter, b [19]. The mother wavelet can be customized to suit specific applications, however, it must be measurable and meet the criterion that it is both absolutely integrable and square integrable, defined mathematically as

$$\int_{-\infty}^{\infty} |\xi(t)| dt < \infty \quad \text{and} \quad \int_{-\infty}^{\infty} |\xi(t)|^2 dt < \infty. \quad (2.2)$$

These requirements are necessary such that the mean of each wavelet in the wavelet family is zero and each wavelet is contained within the Hilbert space [19]. The WT allows the waveform to be represented by a family of wavelets, each multiplied by a coefficient and derived from the mother wavelet by altering scale and shifting through time. With the

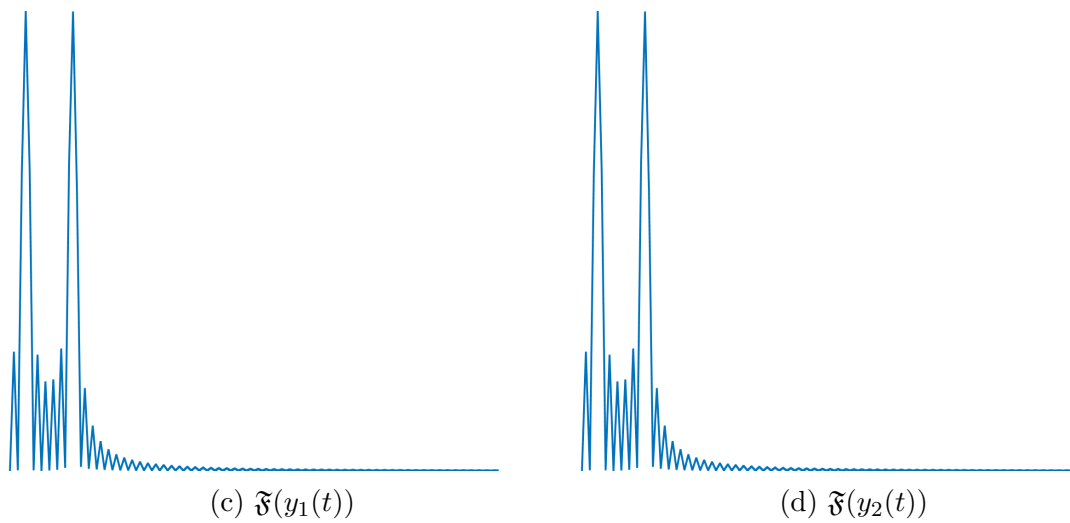
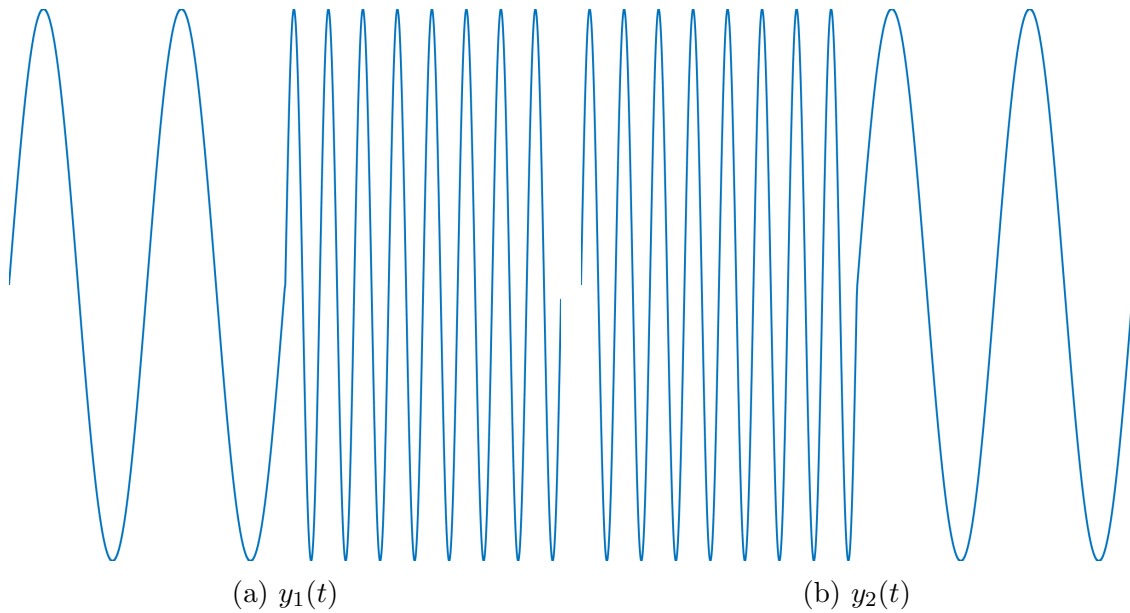
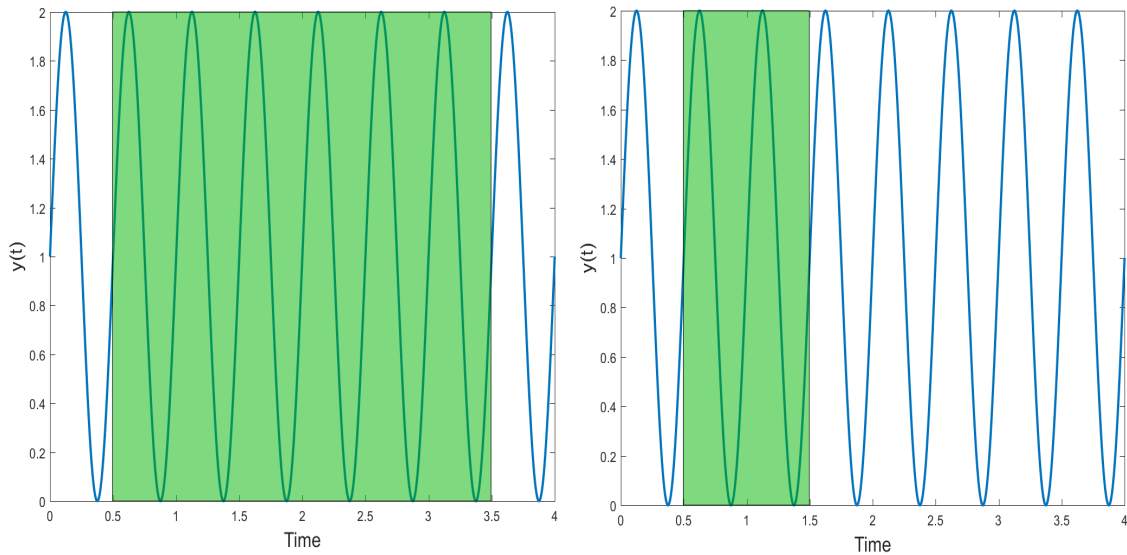
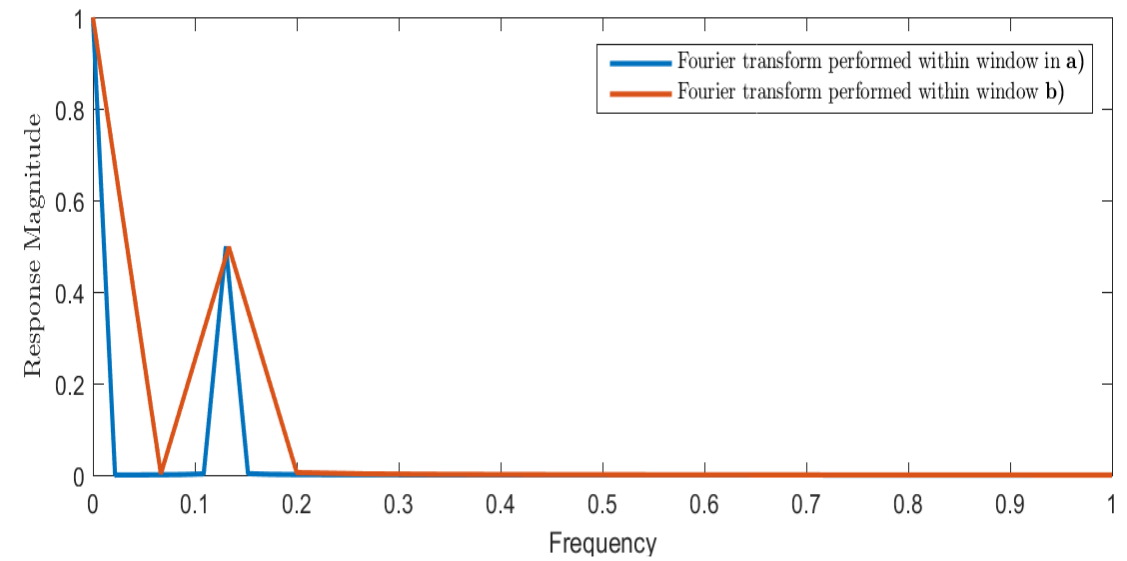


Figure 2.5: **a)** and **b)** show two different signals containing the same frequency content occurring at different time intervals. The magnitude of the Fourier transform for both signals is computed and plotted in **c)** and **d)**, respectively, showing identical results for each of the different signals.



(a) Window with relatively poor time localization. (b) Window with relatively good time localization.

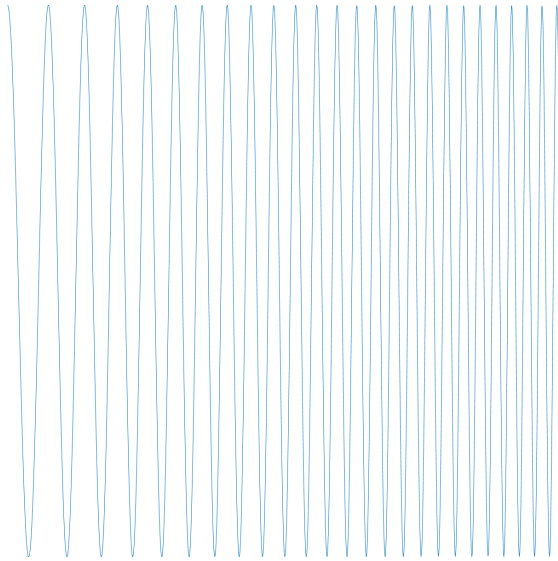


(c) Resulting frequency response of the Fourier transform performed in each of the two windows.

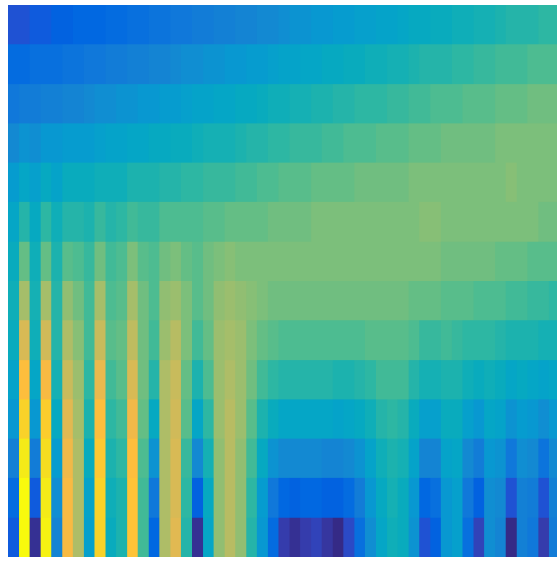
Figure 2.6: A window with relatively poor time localization is shown in **a)**, while a window with relatively good time localization is shown in **b)**. The resulting frequency response of the Fourier transform performed on the windows indicated in **a)** and **b)** is shown in **c)**.

knowledge of each wavelet's scale and translation, a time-frequency response map can be created. A comparison of a wavelet response map to a short time Fourier transform response map is visible in Fig. 2.7. The trade off between time resolution and frequency resolution is apparent when comparing the relatively short windowed STFT and the relatively large windowed STFT. In addition, the smaller windowed STFT does not adequately capture one full period of the analyzed signal and artefacts are seen early in time when the analyzed signal is at low frequency. The mother wavelet for the WT performed in this example was scaled by a factor of 2 and as such, was interpolated such that the frequencies were linearly spaced and consistent with the STFT examples. It can be seen that the time resolution of the WT is very fine while the frequency resolution was not sacrificed.

Two types of wavelet transform exist: the continuous wavelet transform (CWT) and the discrete wavelet transform (DWT). The main difference with these two transforms lies in the way the mother wavelet is scaled. In the case of the DWT, the scale parameter is discretized to integer powers of 2 (i.e., 2^3). In the case of CWT, the scale parameter is more finely discretized to fractional powers of 2 (i.e., $2^{3/2}$). The primary trade off between these two methods is that the CWT is more computationally complex, but allows for finer scaling of the mother wavelet. Conversely, the DWT is computationally more simple, but limits the granularity of the scaling parameter. In this work, the continuous wavelet transform was used, though the implementation was discretized such that it could be implemented on a computer.



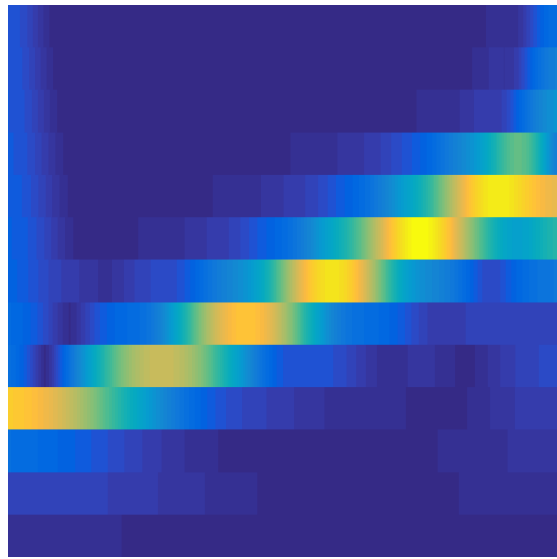
(a) Linear chirp function.



(b) Response map generated with the STFT with a relatively small window.



(c) Response map generated with the STFT with a relatively large window.



(d) Response map generated with the wavelet transform (logarithmic scale).

Figure 2.7: A comparison of response maps generated for signal **a**) with the y-axis representing frequency and the x-axis representing time. **b**) is generated by the STFT using a small window, **c**), is generated by the STFT using a large window, and **d**) is generated with the WT.

Chapter 3

Methodology

This chapter provides a general algorithm overview, a description of the overall methodology, and details relating to the main contributions of this work. Section 3.1 provides a brief description of the overall algorithm as well as an algorithm flow chart. Sample region registration and point tracking is outlined in Section 3.2 while the skin erythema model is covered in Section 3.3. The two main contributions of PPG waveform construction and PPG waveform analysis are detailed in Section 3.4 and Section 3.5, respectively.

3.1 Algorithm Overview

To produce an accurate heart rate estimate remotely using video, an accurate PPG waveform must first be constructed. Chwyl *et al.* [18] demonstrated that a reliable PPG waveform estimate can be obtained through Bayesian minimization. This Bayesian minimization uses a set of skin erythema time series as observations, established from points on the cheek region, with the required posterior probability estimated via an importance-weighted Monte Carlo approach. Once constructed, the continuous wavelet transform (CWT) is applied, using the pulselet as the mother wavelet, to allow for analysis in the time-frequency domain. The scales used in the CWT are selected to correspond to frequencies encapsulating realistic heart rate values ($[f_l, f_h]$ Hz). The average of frequencies corresponding to the maximum response of the CWT at each time step is obtained and multiplied by 60 to produce HR_{bpm} , a heart rate estimate in the standard unit of measurement, beats per minute (bpm). An algorithm flow diagram can be seen in Fig. 3.1.

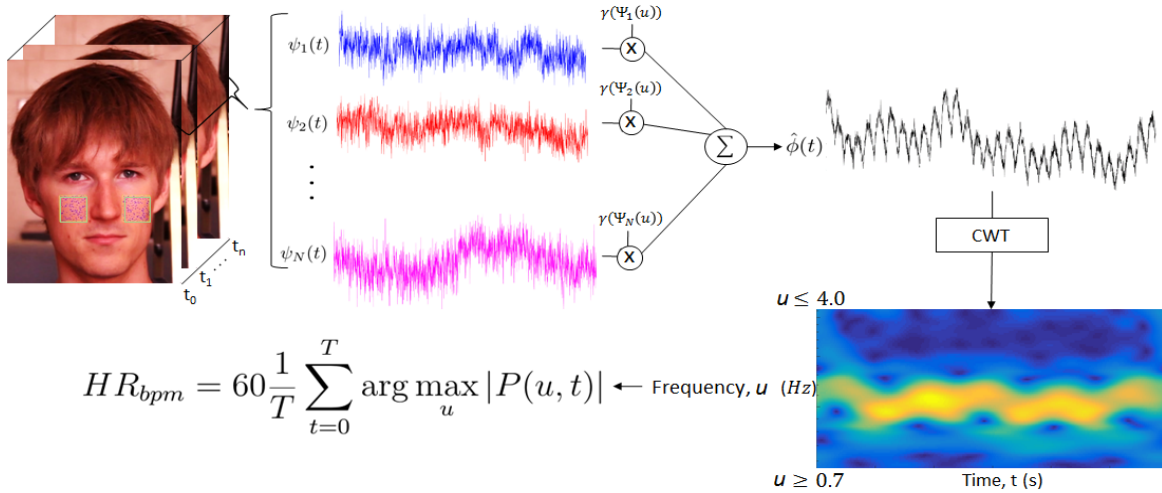


Figure 3.1: Flow chart of the proposed algorithm. Many skin erythema time series, $\psi(t)$, are used to produce an estimate of the PPG waveform, $\hat{\phi}(t)$. The contributions of each signal are weighted based on $\gamma(\Psi(u))$, which considers the Fourier transform of $\psi(t)$, $\Psi(u)$. The continuous wavelet transform is then applied to produce $|P(u, t)|$, the magnitude of the wavelet response in the time-frequency domain.

3.2 Sample Region Registration and Point Tracking

To obtain measurements relevant to heart rate estimation, the subject’s skin region must be localized. The cheek region’s low facial skin thickness [30] makes it an ideal location for obtaining measurements. Because there is less substance for light to penetrate, reflected light from the cheek region undergoes less scattering and is more likely to provide a strong signal.

The localization of the cheek region relies heavily on a cascade object detection method proposed by Viola and Jones [58]. This method iteratively applies classifiers of increasing discernment to quickly discard areas of unimportance, enabling it to perform classification tasks in real time. To register the cheek region, the subject’s face is first identified using the cascade object detection method trained specifically to identify human faces. After obtaining the face region, the subject’s eye pair is identified by applying a version of the object detection algorithm trained on human eye pairs within the identified face region. Once acquired, the eye pair bounding box, EP , is translated and scaled to produce the cheek bounding box C , as



Figure 3.2: Example of the final cheek region, outlined in green, from which the PPG waveform is constructed.

$$\begin{aligned}
 C_x &= EP_x & C_{width} &= EP_{width} \\
 C_y &= EP_y + 1.4 * EP_{height} & C_{height} &= 1.2 * EP_{height},
 \end{aligned}
 \tag{3.1}$$

where the subscripts x and y represent the coordinates of the top left corner of each bounding box, and the subscripts $width$ and $height$ represent the width and height of each bounding box, respectively. The resulting bounding box is split into two sections with the middle fifth of C omitted to avoid sampling from the nose. An example of the final cheek region is visible in Figure 3.2.

Once the cheek region is localized, N points are stochastically sampled with uniform probability. To compensate for motion, each point, x , is updated through time, t , as

$$x_t = f(x_{t-1}), \tag{3.2}$$

where $f(\cdot)$ is the tracking function proposed by Tomasi *et al.* [55].

3.3 Skin Erythema Model

While narrowband methods inherently focus on wavelengths closely correlated with hemoglobin concentration, a suitable model which correlates to hemoglobin concentration must also

be applied in the case of broadband red-green-blue (RGB) measurements. Skin erythema, the measure of redness of the skin, offers a biologically motivated means of heart rate estimation [16] due to its high correspondance to hemoglobin concentration [29]. While numerous skin erythema models exist [24, 23, 21], the skin erythema model proposed by Yamamoto *et al.* [62] was used due to the broadband nature of our captured signal. The skin erythema transform, $\psi(t)$, is formulated for a single point, x , as

$$\psi(t) = \log_{10} \left(\frac{1}{\bar{x}_g(t)} \right) - \log_{10} \left(\frac{1}{\bar{x}_r(t)} \right) \quad (3.3)$$

where $\bar{x}_g(t)$ and $\bar{x}_r(t)$ represent the average values of the green and red channels, respectively, within a window of size $\aleph \times \aleph$ surrounding x . The spatial average is taken in a window surrounding x in order to reduce the effects of tracking error and point noise.

3.4 Photoplethysmogram Waveform Estimation

To accurately estimate heart rate, to construct a clean PPG waveform is imperative. By incorporating prior knowledge of the expected PPG signal into a statistical framework, a set of observations can produce a relatively clean PPG waveform. In this approach, a skin erythema time series is constructed for each of N sample points, stochastically sampled from within the cheek region, to establish a set observations. Bayesian minimization is then applied, where the posterior probability is inferred using an importance-weighted Monte Carlo sampling approach [31].

3.4.1 Bayesian Minimization

By considering numerous skin erythema time series observations, a statistical framework can be used to reliably produce an estimate of the PPG waveform, $\hat{\phi}(t)$. Each skin erythema time series is acquired by applying the skin erythema transform, $\psi(t)$, to a certain sample point, x , as it is tracked throughout the video sequence. Bayesian minimization is then used to construct $\hat{\phi}(t)$ and is formulated as

$$\hat{\phi}(t) = \arg \min_{\hat{\phi}(t)} E(((\hat{\phi}(t) - \phi(t))^2) | X), \quad (3.4)$$

where $E(\cdot)$ represents the expectation and X represents the set of skin erythema time series observations. Substituting the formal definition of $E(\cdot)$ into Eq. 3.4 yields

$$\hat{\phi}(t) = \arg \min_{\hat{\phi}(t)} \left(\int (\phi(t) - \hat{\phi}(t))^2 p(\phi(t)|X) d\phi(t) \right), \quad (3.5)$$

where $p(\phi(t)|X)$ is the posterior probability. To solve the minimization, the derivative of each side is taken [43] to produce

$$\frac{\delta}{\delta d\hat{\phi}(t)} \left(\int (\phi(t) - \hat{\phi}(t))^2 p(\phi(t)|X) d\phi(t) \right) = \int -2(\phi(t) - \hat{\phi}(t)) p(\phi(t)|X) d\phi(t). \quad (3.6)$$

Setting Eq. 3.6 to zero and simplifying produces

$$\begin{aligned} 0 &= \int -2(\phi(t) - \hat{\phi}(t)) p(\phi(t)|X) d\phi(t) \\ 0 &= -2 \int (\phi(t) p(\phi(t)|X) d\phi(t) - \hat{\phi}(t) p(\phi(t)|X) d\phi(t)) d\phi(t) \\ \int \hat{\phi}(t) p(\phi(t)|X) d\phi(t) &= \int \phi(t) p(\phi(t)|X) d\phi(t) \\ \hat{\phi}(t) \int p(\phi(t)|X) d\phi(t) &= \int \phi(t) p(\phi(t)|X) d\phi(t) \\ \hat{\phi}(t) &= \int \phi(t) p(\phi(t)|X) d\phi(t). \end{aligned} \quad (3.7)$$

The posterior probability, $p(\phi(t)|X)$, is required to solve this equation; however, it is difficult to solve for analytically. Therefore, we use an importance-weighted Monte Carlo sampling approach to estimate $p(\phi(t)|X)$ [15, 61, 43].

3.4.2 Posterior Probability Estimation

To estimate the posterior probability via an importance-weighted Monte Carlo sampling approach, a set of reliable observations, X , must first be established. The acceptance probability, $\gamma(\Psi_i(u))$, defining the likelihood that the i^{th} skin erythema time series, $\psi_i(t)$, is accepted into the set X , is calculated as

$$\gamma(\Psi_i(u)) = \exp\left(-\alpha \frac{\max(|\Psi_i(u)|)}{\zeta}\right) \quad (3.8)$$

subject to $u_l \leq u \leq u_h$,

where $\Psi_i(u)$ is the Fourier transform of $\psi_i(t)$, ζ is the noise floor of $\Psi_i(u)$, and α is an empirically determined scaling parameter. Because it is expected that clean observations will have a prominent response within the range of reasonable heart rates, this metric was designed to omit observations with weak frequency responses in this frequency band. The noise floor, ζ , is obtained in two steps. First, a window of width L containing the smallest windowed variance within $|\Psi_i(u)|$ is identified. The centre of the identified window is denoted as u_c . Second, ζ is calculated as the average of frequency response values in a window of width L centred at u_c . Locating the window centre, u_c , is formulated as

$$u_c = \arg \min_z Var(|\Phi(u)|) \quad (3.9)$$

subject to $z - \frac{L}{2} \leq u \leq z + \frac{L}{2}$,

where $Var(|\Phi(u)|)$ is the windowed variance of $|\Phi(u)|$ within the range $z - \frac{L}{2} \leq u \leq z + \frac{L}{2}$. The noise floor, ζ , can then be calculated as the average magnitude of frequency responses within a window of width L centred at u_c and is formulated as

$$\zeta = \sum_{j=u_c - \frac{L}{2}}^{u_c + \frac{L}{2}} \Psi_i(j). \quad (3.10)$$

Once X is established, an estimate, $\hat{p}(\phi(t)|X)$, of the posterior probability can be produced from a weighted histogram:

$$\hat{p}(\phi(t)|X) = \frac{\sum_{i=0}^N \gamma(\Psi_i(u)) \delta(\psi_i(t) - \phi(t))}{\sum_{i=0}^N \gamma(\Psi_i(u))}, \quad (3.11)$$

where $\delta(\cdot)$ is the Dirac function, N is the number of observations and $\gamma(\Psi_i(u))$ is the acceptance probability. Because the acceptance probability is designed to accept clean signals into the set, it is reused as a weighting function in the posterior probability estimation. Once acquired, $\hat{p}(\phi(t)|X)$ can be used in Eq. 3.7 to calculate $\hat{\phi}(t)$.

3.5 Photoplethysmogram Waveform Analysis

The continuous wavelet transform (CWT) offers a technique for analyzing the frequency of a non-stationary signal as it changes through time, avoiding the spread of frequency content incurred by analyzing a PPG waveform across its entire duration. The analyzing wavelet used in the CWT in our method was chosen due to its similar characteristics to a PPG waveform constructed with a finger pulse oximeter, and is referred to in this thesis as the *pulselet*. Similarities between the pulselet and the PPG waveform produced by a finger pulse oximeter can be seen in Fig. 3.3, specifically with the initial systolic peak followed by a lesser diastolic peak. The pulselet, $\Xi(au)$, is based on the bump wavelet [45] and is formulated in the frequency domain as

$$\Xi(au) = \begin{cases} \exp\left(1 - \frac{1}{1 - \frac{(au - \mu)^2}{\sigma^2}}\right) & \frac{(\mu - \sigma)}{a} \leq u \leq \frac{(\mu + \sigma)}{a}, \\ 0 & \text{otherwise.} \end{cases} \quad (3.12)$$

where μ represents the pulselet's center frequency in $\frac{rad}{sample}$, σ controls the balance between frequency localization and time localization, a is the scale parameter, and u is the frequency.

Using this pulselet, the CWT was performed, resulting in a map of frequency responses in the time-frequency domain. The heart rate in beats per minute, HR_{bpm} , is estimated as

$$HR_{bpm} = \frac{60}{T} \sum_{t=0}^T \arg \max_u |P(u, t)| \quad (3.13)$$

where $|P(u, t)|$ represents the magnitude of the frequency response to the pulselet centred at frequency u and time t , and T represents the total duration of $\phi(t)$.

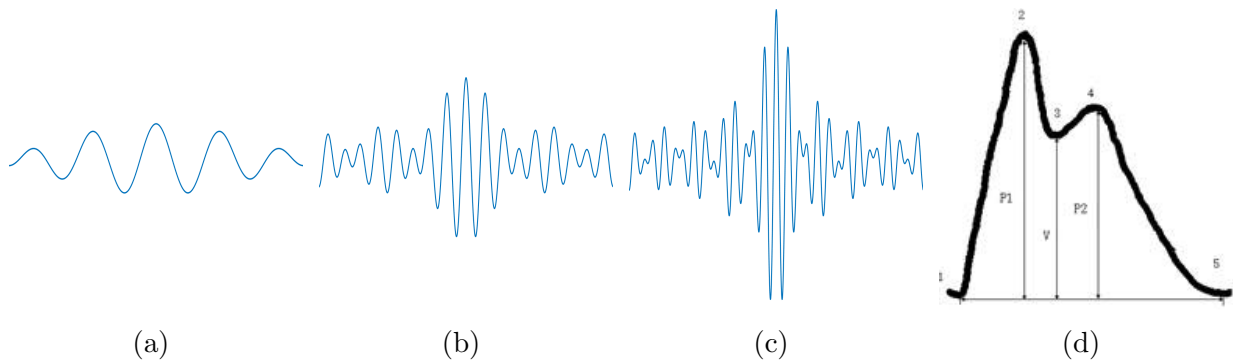


Figure 3.3: The pulselet in the time domain scaled to frequencies of **a)** 42 beats per minute, **b)** 140 beats per minute, **c)** 240 beats per minute, as well as **d)**, an example of a single period of a PPG waveform acquired through finger pulse oximetry [33].

Chapter 4

Experiments

Two experiments were performed to verify the efficacy of the proposed method. The experimental setup of a broadband red-green-blue channel experiment, as well as the subsequent results, are described in Section 4.1. Similarly, the experimental setup and subsequent results of an experiment performed on a near-infrared data set collected by Amelard *et al.* [5] are outlined Section 4.2.

4.1 Broadband Red-Green-Blue Channel Data Set

4.1.1 Experimental Setup

To evaluate the effectiveness of the proposed method, a data set consisting of 30 videos was recorded. The data set contains five different participants with six videos per participant. The videos are each 30 seconds in length and were captured at 80 frames per second with a Chameleon3 camera produced by Point Grey Research, Inc. [48]. The recorded videos are divided into two categories; stationary, and gradual motion. For each participant, four stationary videos and two gradual motion videos exist. The stationary category consists of videos where the participant was asked to sit still throughout the recording, whereas the gradual motion category contains videos where the participant was encouraged to gradually move and exhibit natural motion, such as speaking and facial expressions.

Ground truth was obtained by including a red LED in each frame in the video sequence. The LED was wired to an Easy Pulse 1.1 finger pulse oximeter (manufactured by Embedded Lab) [51] and programmed to blink with each detected pulse. The location of the red



Figure 4.1: Frame capture from a video in the tested data set. Ambient lighting conditions were used while a red LED was positioned behind the participant as a means of acquiring ground truth.

LED was manually determined in each video sequence and a red channel time series was created. The Fourier transform was applied to the red channel time series and the frequency corresponding to the maximum magnitude in the frequency domain was used as the ground truth heart rate frequency. To ensure the light emitted from the red LED did not interfere with the ambient reflected light, the red LED was positioned behind the participant as shown in Fig. 4.1.

The values of u_l and u_h were chosen as 0.7 Hz and 4.0 Hz, respectively, to encapsulate a wide range of heart rate values while representing average heart rate values [1]. The experiment was conducted with parameter values of $\alpha = 2.0$, $\aleph = 5$, $L = 35$, $\sigma = 0.7$, $\mu = 4.0$, and a sampling rate of 1% when initially sampling from the cheek region, as these values were empirically determined to produce strong results.

To compare against state of the art methods, Poh *et al.* [46] (2010), Poh *et al.* [47] (2011), Li *et al.* [42], and Chwyl *et al.* [18] were implemented. For the method proposed by Poh *et al.* [47] (2011), a novel peak detection algorithm is used; however, this particular peak detection algorithm is left unspecified in the publication. Therefore, the Fourier transform was applied to the final PPG waveform and the frequency corresponding to the maximum peak was selected as the heart rate frequency.

Table 4.1: Results of state-of-the-art comparison subdivided into two categories: stationary and gradual motion. The mean error (M_e), standard deviation (σ_e), mean absolute error ($|M_e|$), and root mean squared error ($RMSE$) are tabulated with the best results for $|M_e|$ and $RMSE$ indicated in boldface. All results are shown in terms of beats per minute (bpm).

	Stationary			Gradual Motion		
	M_e (σ_e)	$ M_e $	$RMSE$	M_e (σ_e)	$ M_e $	$RMSE$
Pulselet Method	-0.16 (3.67)	2.14	3.58	-3.33 (7.48)	5.49	7.84
Chwyl 2015 [18]	-3.23 (12.24)	5.25	12.36	-11.00 (18.45)	11.56	20.68
Li 2014 [42]	20.25 (21.12)	21.16	29.26	15.19 (18.24)	18.35	23.74
Poh 2011 [47]	29.88 (24.04)	30.77	38.35	31.74 (24.39)	32.07	40.03
Poh 2010 [46]	-12.71 (17.19)	14.14	21.38	-10.65 (28.48)	22.81	30.41

4.1.2 Results and Discussion

From our data set, the error was calculated as the difference between the heart rate estimates and the ground truth heart rate measurements for each video. The mean error (M_e), standard deviation of the mean error (σ_e), mean absolute error ($|M_e|$), and root mean squared error ($RMSE$) were calculated in beats per minute (bpm) for each method and are tabulated in Table 4.1 with the best results indicated in bold face. It can be seen that for all conditions, our pulselet method has lower $RMSE$ and $|M_e|$ values, and the M_e values are closer to zero than other state of the art methods with smaller σ_e . This implies that the estimates are not biased (i.e., consistently high) and are consistently close to ground truth. Pulselet response maps corresponding to videos from each the stationary and gradual motion categories can be seen in Fig. 4.2, showing the heart rate variation throughout the video. Natural heart rate variations are visible for both cases, allowing these fluctuations to be included in the average heart rate estimates. Though out of the scope of this work, the sinusoidal nature of these variations are of note, and likely provide insight into respiratory rate [4].

Heart rate estimates from the set of videos with gradual motion typically resulted in higher error than heart rate estimates from the videos with no motion; however, the methods which use detrending and filtering algorithms ([47] and [42]) are less affected. This is expected as the presence of motion introduces the possibility of changing illumination levels, varying face orientation, and tracking error, all of which introduce noise to the PPG waveform. In the methods proposed by Poh *et al.* [47] (2011) and Li *et al.* [42], the low frequency illumination variation is suppressed with the detrending methods while

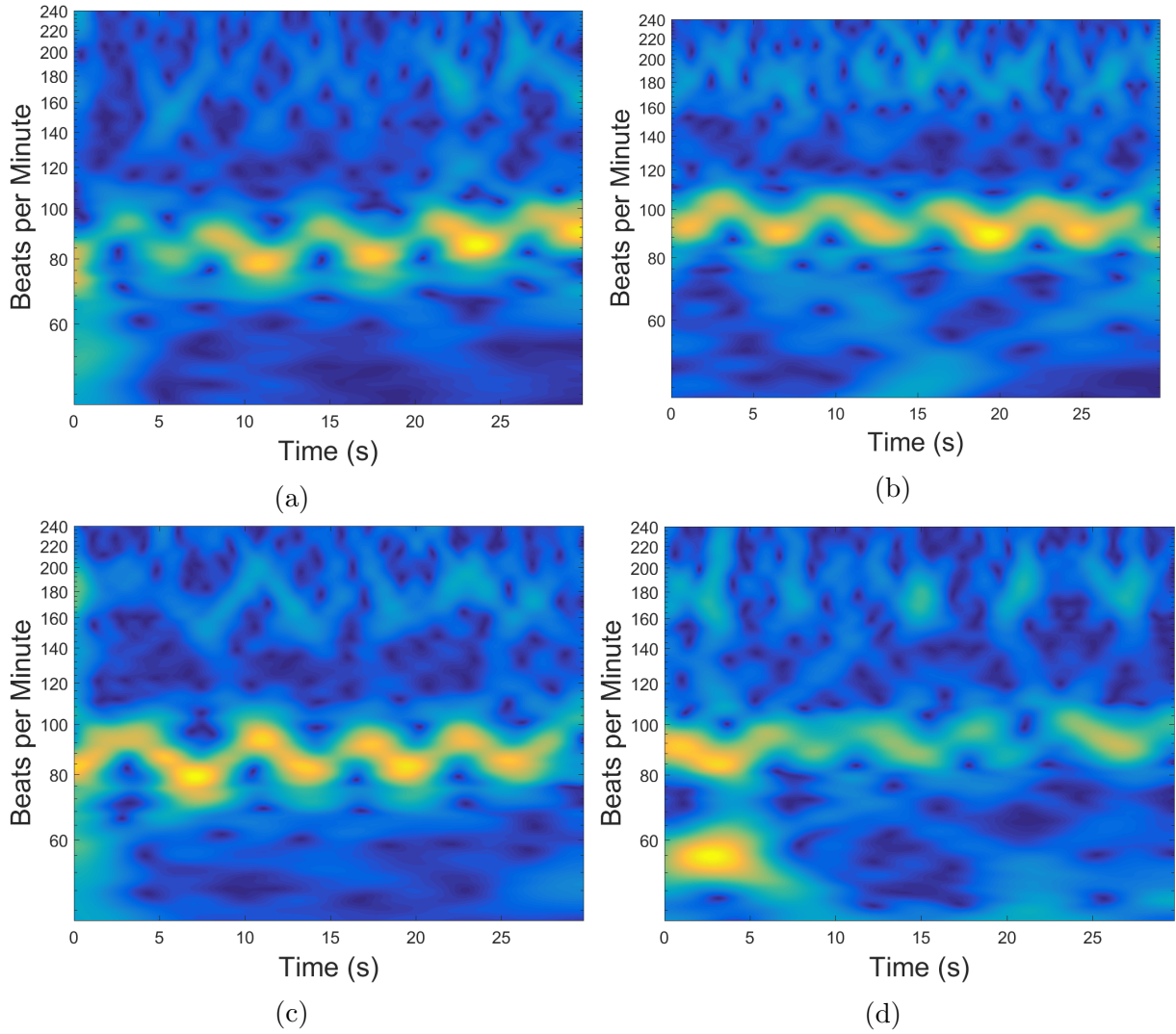


Figure 4.2: Pulselet response maps for two videos from the stationary category, **a)** and **b)**, and two videos from the gradual motion category, **b)** and **c)**, for a single participant. Response values are represented as a heat map, with the greatest responses indicated in bright yellow and the weakest responses indicated in dark blue.

Table 4.2: Results of state-of-the-art comparison for both stationary conditions and gradual motion. The mean error (M_e), standard deviation (σ_e), mean absolute error ($|M_e|$), and root mean squared error ($RMSE$) are tabulated with the best results for $|M_e|$ and $RMSE$ indicated in boldface. All results are shown in terms of beats per minute (bpm). T-tests performed between the absolute errors of the pulselet method and the absolute errors of each other method are also tabulated.

	All Motion Conditions			
	M_e (σ_e)	$ M_e $	$RMSE$	t -test
Pulselet Method	-1.22 (5.34)	3.26	5.39	n/a
Chwyl 2015 [18]	-5.82 (14.76)	7.35	15.63	$t(29) = 1.97, p > 0.001$
Li 2014 [42]	18.57 (20.35)	20.22	27.54	$t(29) = 4.37, p < 0.001$
Poh 2011 [47]	30.50 (24.17)	31.20	38.92	$t(29) = 6.08, p < 0.001$
Poh 2010 [46]	-12.02 (21.64)	17.03	24.76	$t(29) = 4.49, p < 0.001$

the applied filters remove high frequency noise, thus compensating for noise introduced through motion.

The methods proposed by Li *et al.* [42] and Poh *et al.* [47] (2011) have positive mean errors, suggesting these methods produce high heart rate estimates. This may be explained by the detrending algorithms and filtering used in both methods. Such methods aim to remove low frequency noise such as ambient illumination changes; however, relevant frequency content may be erroneously suppressed in the process.

Conversely, the methods proposed by Poh *et al.* [46] (2010) and Chwyl *et al.* [18] produce negative mean errors, caused by heart rate estimates lower than ground truth. In these methods, it is likely that natural heart rate fluctuation caused the frequency content to be spread over multiple frequency bins. In addition, detrending algorithms and filtering are not applied to these methods, increasing the likelihood of noise and causing less prominent frequency responses at the heart rate frequencies. An example of such a situation can be seen in Fig. 4.3. It can be seen that the FFT response is spread across multiple frequency bins, resulting in two prominent peaks on either side of the ground truth heart rate (76.50 bpm). In cases where the heart rate frequency is not prominent, the maximum magnitude of the frequency response within the range $[u_l, u_h]$ Hz will likely be located near u_l due to the inherently dominant low frequency content.

Though the pulselet method also produces a negative average error, this is primarily attributed to noise in the constructed PPG signal. Unlike the other aforementioned methods, signal analysis is performed in the time-frequency domain, allowing heart rate fluctuations to be captured and included in the average heart rate estimate. However, the

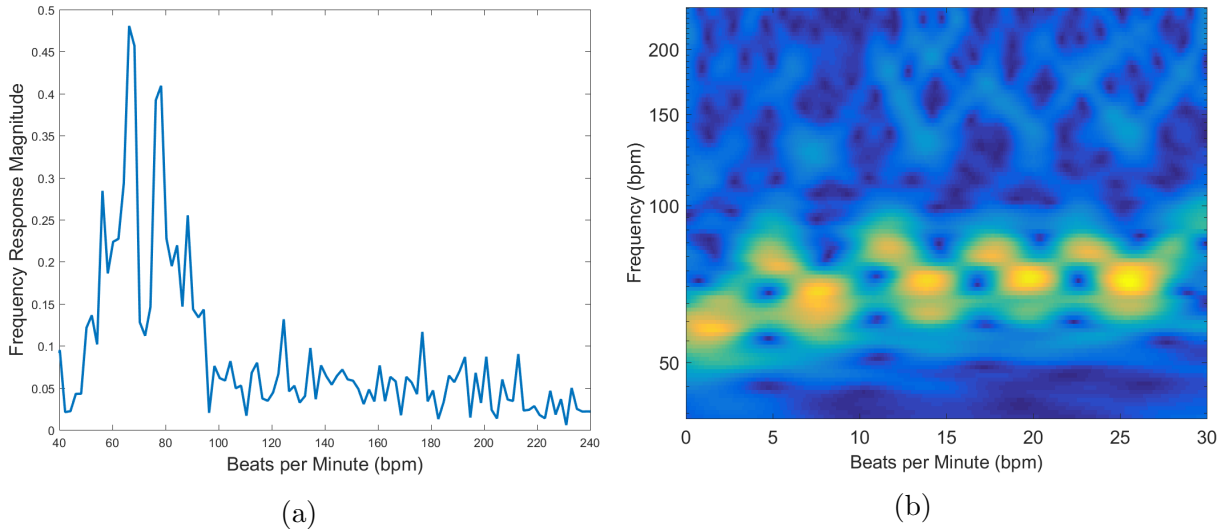


Figure 4.3: The magnitude of the frequency response of a PPG waveform resulting from the FFT is shown in **a)**, while the response map of the same PPG waveform resulting from the CWT is shown in **b)**. The maximum peak of **a)** is 66.27 bpm, the average of maximum frequencies over time of **b)** is 75.36 bpm, and the ground truth heart rate is 76.50 bpm.

lack of filtering or detrending remains, offering an explanation as to why the heart rate estimates are lower than the ground truth heart rate on average.

A t-test performed on the absolute errors of each method show significant improvement over Poh *et al.* [46] (2010), Poh *et al.* [47] (2011), and Li *et al.* [18] with t-test results of $t(29) = 4.37, p < 0.001$, $t(29) = 6.08, p < 0.001$, and $t(29) = 4.49, p < 0.001$, respectively. Though a statistically significant improvement was not shown against Chwyl *et al.* [18] ($t(29) = 1.97, p > 0.001$), results indicate a 10.24 bpm reduction in *RMSE* across all videos.

4.2 Near-Infrared Data Set

4.2.1 Experimental Setup

To verify the robustness of the proposed framework we tested the system on a narrow-band near-infrared (NIR) data set created by Amelard *et al.* [5]. Videos of 24 different

participants (age ($\mu \pm \sigma$) = 28.7 ± 12.4) were collected using a coded hemodynamic imaging (CHI) system consisting of a PointGrey GS3-U3-41C6NIR-C camera with an optical bandpass filter ($\lambda = 850 - 1000nm$) at a distance of $1.5m$ from the participant. Ground truth was simultaneously collected via the Easy Pulse 1.1 finger pulse oximeter [51] and the scene was illuminated with a $250W$ tungsten-halogen lamp, also placed at a distance of $1.5m$ from the participant. An example of the experimental configuration can be seen in Fig. 4.4. This study was approved by a University of Waterloo Research Ethics committee.

To remain consistent with the broadband light experiment, the values of u_l and u_h were chosen as 0.7 Hz and 4.0 Hz, respectively, and parameter values of $\alpha = 2.0$, $\aleph = 5$, $L = 35$, $\sigma = 0.7$, $\mu = 4.0$, and a sampling rate of 1% when initially sampling from the cheek region, were used. For all videos in this data set the cheek region was manually defined. To offer a comparison to state of the art, the algorithms proposed by Li *et al.* [42] and Chwyl *et al.* [18] were implemented and modified to use the near-infrared channel in place of the green channel waveform and skin erythema waveform, respectively. Comparison against narrowband state of the art methods was not performed, as the primary novelty in these methods is accomplished in the data capture stage by using unique imaging devices.

4.2.2 Results and Discussion

Table 4.3: Average of mean error (M_e), mean absolute error ($|M_e|$), and root mean squared error ($RMSE$) across the entire near-infrared data set calculated in beats per minute for the algorithms proposed by Li *et al.* [42], Chwyl *et al.* [18], as well as the pulselet method. The standard deviations of M_e is also tabulated (σ_e). The best values for $|M_e|$ and $RMSE$ are indicated in boldface. T-tests performed between the absolute errors of the pulselet method and the absolute errors of each other method are also tabulated.

	M_e (σ_e)	$ M_e $	$RMSE$	t-test
Pulselet Method	0.45 (1.80)	1.39	1.90	n/a
Chwyl 2015 [18]	-1.80 (6.40)	2.38	6.50	$t(20) = 0.73, p > 0.001$
Li 2014 [42]	0.17 (1.17)	0.86	1.16	$t(20) = 1.79, p > 0.001$

The results of the near-infrared experiment are summarized in Table 4.3. It can be seen that although the method proposed by Li *et al.* [42] achieves both the lowest mean absolute error and root mean squared error, the pulselet method is comparable. T-tests, performed between the absolute errors of the pulselet method and the absolute errors of Li *et al.*'s [42] method, yielded a result of $t(20) = 1.79, p > 0.001$, indicating the improvement is statistically insignificant. The same t-test was performed between the pulselet

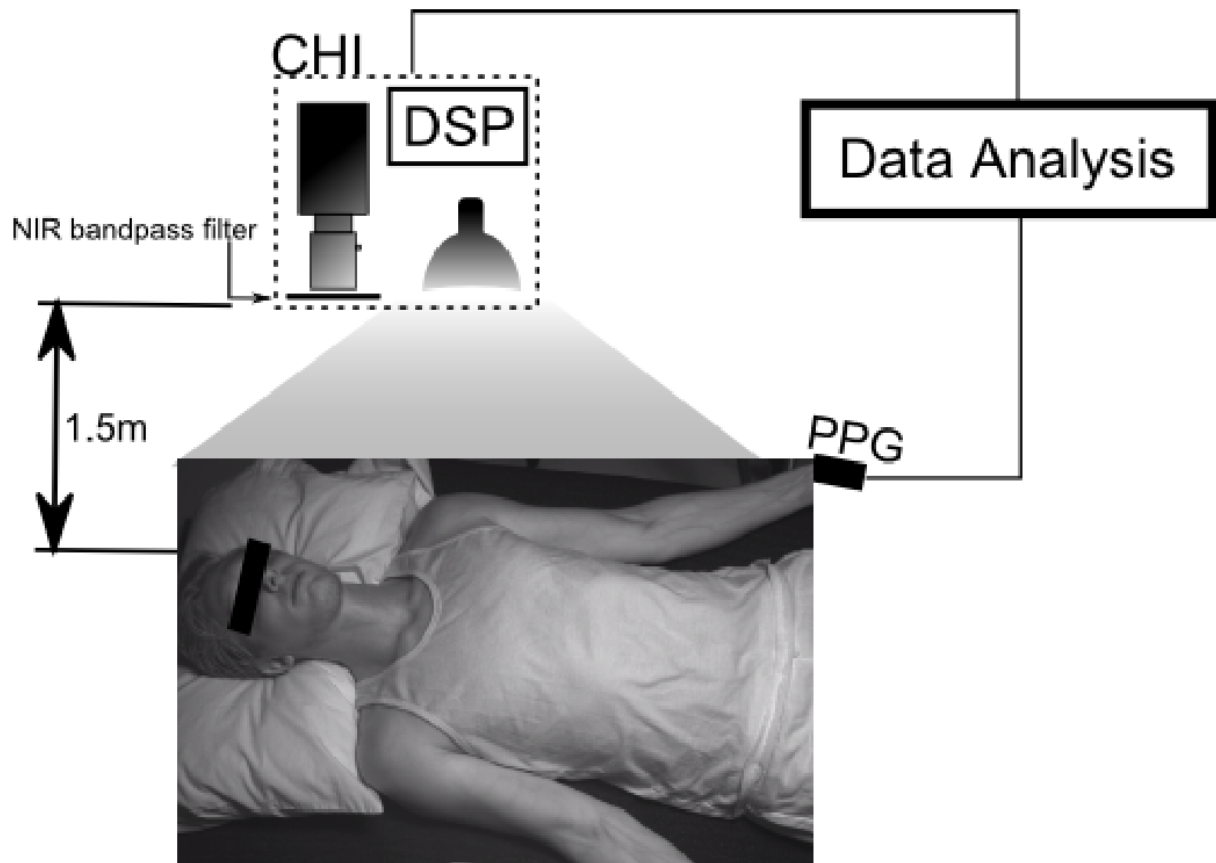


Figure 4.4: Geometric configuration of experiment. A participant is imaged via a coded hemodynamic imaging system with an applied tungsten-halogen illumination source. Ground truth is simultaneously collected using a finger pulse oximeter [5].

method and Chwyl *et al.*'s [18] method and produced a result of $t(20) = 0.73, p > 0.001$, once again indicating statistical insignificance.

It is thought that the lack of motion, the supine position of individuals throughout the experiment, and the narrowband device for PPG waveform acquisition played a role in the similarity of results. With the PPG waveform precisely acquired, the main contribution in PPG waveform construction is not prominent, including the added robustness to motion and non-ideal illumination conditions. The contribution of PPG waveform analysis in the time-frequency domain is also less prominent, as participants in a supine position typically have less natural fluctuation in heart rate, and the fluctuation occurs at lower frequencies [63]. This largely reduced the need for analysis of a quasi-periodic signal, meaning any benefits of analyzing the waveform in the time-frequency domain are less apparent. In addition, the detrending methods used by Li *et al.* [42] remove low frequency content and are better suited for removing this type of slow moving variation from the waveform in preprocessing, allowing a greater level of accuracy in heart rate estimation.

Chapter 5

Conclusions

5.1 Key Observations

When applied to broadband RGB data, the method detailed in this work outperforms state of the art methods for two primary reasons. First, it constructs the PPG waveform from many separate observations, rejecting those which are contaminated with significant noise, attenuating those which suffer minor artefacts, and amplifying those with the cleanest waveforms. Because observations are taken stochastically over the cheek region, this allows for a degree of robustness against unevenly applied illumination, as well as illumination variation throughout the video. Second, by analyzing the constructed PPG waveform in the time-frequency domain via the wavelet transform, the quasi-periodicity of the PPG waveform is well handled. By averaging the maximum frequency over time, results obtained by this method outperform those obtained by other state of the art methods. In the near-infrared experiment, natural heart rate fluctuations were minimized due to the supine position of all participants, partially contributing to statistically insignificant results when compared to state-of-the-art methods.

5.2 Concluding Remarks

A method for non-contact heart rate estimation via Bayesian minimization and time-frequency domain analysis was outlined. Results indicate a statistically significant improvement over most current state of the art methods for broadband systems and a comparable performance to state-of-the-art methods when applied to narrowband systems.

This method offers a cost effective and computationally efficient means of estimating heart rate. Cost is saved since this method is capable of working with a multitude of camera configurations, as demonstrated by the near-infrared experiment, and does not require active illumination. The versatility of this framework is beneficial because, unlike many state-of-the-art methods, this method is independent of hardware. In addition, this method lends itself well to parallelization, making it capable of high computational performance and fast calculations. A cost effective system for non-contact monitoring of heart rate in the presence of gradual motion and ambient illumination conditions would allow for convenient and hygienic screening of patients, as well as allow for novel applications such as telemedicine and affective computing. While results are encouraging, it is recommended that further testing be conducted and certain aspects of this method improved.

5.3 Future Research and Recommendations

Though results indicate an improvement over state of the art systems, further improvements can be made. Items to consider for future research are listed below.

- Comprehensive testing
- Improve motion compensation
- Increase accuracy of heart rate estimate from time-frequency response map
- Explore time domain analysis
- Assess additional biometrics

Comprehensive testing can be performed to further identify benefits and limitations of this work. The lack of statistical significance in the results acquired in Section 4.2 suggests that the proposed framework’s benefit over state-of-the-art methods are lessened in stationary conditions. Furthermore, the framework’s ability to provide consistent results in the presence of illumination variation was not fully explored. Further experiments could help to fully define the benefits and drawbacks of the proposed framework.

Improve motion compensation by enhancing the way sample points are tracked. It was noted that the sample points were not accurately tracked on a consistent basis due to the lack of unique local features within the cheek region. By using a separate set of tracking points and updating the sample points via a geometric transform, better motion

compensation may be achieved. Since the tracking points are not directly used to construct the PPG waveform, it is possible to locate them anywhere on the participant containing unique local features. Another possible improvement is to update the sample points based on the dense optical flow between consecutive frames. While the point tracking function uses local spatial information surrounding each point, dense optical flow considers the entire image. While potentially more robust, this method is more computationally inefficient.

Increase accuracy of heart rate estimate from time-frequency response map
While reasonable results were obtained, the final component of the outlined framework, which produces a final heart rate estimation from a time-frequency response map, is simple and could likely be improved with a more sophisticated algorithm. Currently, an average of frequencies corresponding to the maximum response over time is considered as the heart rate frequency, however, this does not consider outliers or discontinuities. By enforcing a certain level of smoothness between consecutive time steps, noise affecting brief durations of the PPG waveform can be suppressed and results can likely be improved.

Explore time domain analysis by discarding or attenuating wavelet coefficients which fall below a selected magnitude threshold in the time-frequency domain. This allows for a de-noised waveform to be reconstructed with only the important frequencies present. It is then possible to calculate the average inter-beat interval via peak detection. This will reduce any error introduced at the edges of the waveform when analyzing in the time-frequency domain.

Assess additional biometrics such as breath rate and cognitive stress by analyzing the way the PPG waveform varies through time. As mentioned, respiratory rate can cause fluctuations in heart rate, allowing breath rate to be inferred by analyzing how heart rate varies through time. Similarly, cognitive stress [44] has been linked to the ratio of high and low frequency responses of a PPG waveform.

It is recommended that further exploration on this method be conducted as current results are encouraging. Future works could further define limitations, improve accuracy, robustness, and give insight into a wider variety of biomedical information.

References

- [1] Target heart rates. <http://www.heart.org>. Accessed: 2016-01-25.
- [2] Metin Akay and Claudia Mello. Wavelets for biomedical signal processing. In *Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*, volume 6, pages 2688–2691. IEEE, 1997.
- [3] J Allen and A Murray. Variability of photoplethysmography peripheral pulse measurements at the ears, thumbs and toes. In *Science, Measurement and Technology, IEE Proceedings-*, volume 147, pages 403–407. IET, 2000.
- [4] John Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological measurement*, 28(3):R1, 2007.
- [5] Robert Amelard, L Richard Hughson, Danielle K Greaves, Kaylen J Pfisterer, Jason Leung, David A Clausi, and Alexander Wong. Non-contact hemodynamic imaging reveal the jugular venous pulse waveform. Preprint; arXiv:1604.05213v2.
- [6] R Rox Anderson and John A Parrish. The optics of human skin. *Journal of Investigative Dermatology*, 77(1):13–19, 1981.
- [7] Takuo Aoyagi. Pulse oximetry: its invention, theory, and future. *Journal of anesthesia*, 17(4):259–266, 2003.
- [8] Agateller (Anthony Atkielski). Normal cardiac complex from electrocardiogram, 2009. [Online; accessed August 23, 2016].
- [9] Ganesh Balakrishnan, Frederic Durand, and John Guttag. Detecting pulse from head motions in video. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 3430–3437. IEEE, 2013.

- [10] Lena Biel, Ola Pettersson, Lennart Philipson, and Peter Wide. Ecg analysis: a new approach in human identification. *Instrumentation and Measurement, IEEE Transactions on*, 50(3):808–812, 2001.
- [11] C Jan Willem Borleffs, Roderick WC Scherptong, Sum-Che Man, Guido H van Welsenes, Jeroen J Bax, Lieselot van Erven, Cees A Swenne, and Martin J SchaliJ. Predicting ventricular arrhythmias in patients with ischemic heart disease clinical application of the ecg-derived qrs-t angle. *Circulation: Arrhythmia and Electrophysiology*, 2(5):548–554, 2009.
- [12] Discount Cardiology. Ekg machines. [ONLINE]. Available: <http://discountcardiology.com/EKG-Machines/>, 2016.
- [13] Giovanni Cennini, Jeremie Arguel, Kaan Akşit, and Arno van Leest. Heart rate monitoring via remote photoplethysmography with motion artifacts reduction. *Optics express*, 18(5):4867–4875, 2010.
- [14] A V J Challoner and C A Ramsay. A photoelectric plethysmograph for the measurement of cutaneous blood flow. *Physics in Medicine and Biology*, 19(3):317–328, 1974.
- [15] Ming-Hui Chen. Importance-weighted marginal bayesian posterior density estimation. *Journal of the American Statistical Association*, 89(427):818–824, 1994.
- [16] Audrey Chung, Xiao Yu Wang, Robert Amelard, Christian Scharfenberger, Joanne Leong, Jan Kulinski, Alexander Wong, and David A Clausi. High-resolution motion-compensated imaging photoplethysmography for remote heart rate monitoring. In *SPIE BiOS*, pages 93160A–93160A. International Society for Optics and Photonics, 2015.
- [17] Brendan Chwyl, Audrey G Chung, Robert Amelard, Jason Deglint, Alexander Wong, and David Clausi. Time-frequency domain analysis via pulselets for non-contact heart rate estimation from remotely acquired photoplethysmograms. In *Computer and Robot Vision*, 2016.
- [18] Brendan Chwyl, Audrey G Chung, Jason Deglint, Alexander Wong, and David Clausi. Remote heart rate measurement through broadband video via stochastic bayesian estimation. *Vision Letters*, 1(1), 2015.
- [19] Albert Cohen and Jelena Kovacevic. Wavelets: The mathematical background. In *Proc. IEEE*. Citeseer, 1996.

- [20] Ingrid Daubechies et al. *Ten lectures on wavelets*, volume 61. SIAM, 1992.
- [21] JB Dawson, DJ Barker, DJ Ellis, JA Cotterill, E Grassam, GW Fisher, and JW Feather. A theoretical and experimental study of light absorption and scattering by in vivo skin. *Physics in medicine and biology*, 25(4):695, 1980.
- [22] G de Haan and A Van Leest. Improved motion robustness of remote-ppg by using the blood volume pulse signature. *Physiological measurement*, 35(9):1913, 2014.
- [23] BL Diffey, RJ Oliver, and PM Farr. A portable instrument for quantifying erythema induced by ultraviolet radiation. *British Journal of Dermatology*, 111(6):663–672, 1984.
- [24] JW Feather, DJ Ellis, and G Leslie. A portable reflectometer for the rapid quantification of cutaneous haemoglobin and melanin. *Physics in medicine and biology*, 33(6):711, 1988.
- [25] Jin Fei and Ioannis Pavlidis. Thermistor at a distance: unobtrusive measurement of breathing. *Biomedical Engineering, IEEE Transactions on*, 57(4):988–998, 2010.
- [26] Kim Fox, Jeffrey S Borer, A John Camm, Nicolas Danchin, Roberto Ferrari, Jose L Lopez Sendon, Philippe Gabriel Steg, Jean-Claude Tardif, Luigi Tavazzi, and Michal Tendera. Resting heart rate in cardiovascular disease. *Journal of the American College of Cardiology*, 50(9):823–830, 2007.
- [27] Tsu-Hsun Fu, Shing-Hong Liu, and Kuo-Tai Tang. Heart rate extraction from photoplethysmogram waveform using wavelet multi-resolution analysis. *Journal of medical and biological engineering*, 28(4):229–232, 2008.
- [28] Marc Garbey, Nanfei Sun, Arcangelo Merla, and Ioannis Pavlidis. Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. *Biomedical Engineering, IEEE Transactions on*, 54(8):1418–1426, 2007.
- [29] Hao Gong and Michel Desvignes. Hemoglobin and melanin quantification on skin images. In *Image Analysis and Recognition*, pages 198–205. Springer, 2012.
- [30] Richard Y Ha, Kimihiro Nojima, William P Adams Jr, and Spencer A Brown. Analysis of facial skin thickness: defining the relative thickness index. *Plastic and reconstructive surgery*, 115(6):1769–1773, 2005.
- [31] W Keith Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

- [32] Alrick B Hertzman. Photoelectric plethysmography of the fingers and toes in man. *Experimental Biology and Medicine*, 37(3):529–534, 1937.
- [33] Yanchun Hu, Yanxia Hu, and Shenju Song. Photoplethysmographic waveforms recorded from patients with untreated hypertension. *Clinical Medicine Research*, 4(5):163–167, 2015.
- [34] Markus Huelsbusch and Vladimir Blazek. Contactless mapping of rhythmical phenomena in tissue perfusion using ppgi. In *Medical Imaging 2002*, pages 110–117. International Society for Optics and Photonics, 2002.
- [35] Louis Nelson Katz and Alfred Pick. *Clinical electrocardiography*. Lea & Febiger, 1956.
- [36] Paul Khairy and Ariane J Marelli. Clinical use of electrocardiography in adults with congenital heart disease. *Circulation*, 116(23):2734–2746, 2007.
- [37] Samer Khudairi. Morpholio presents photoplethysmography technology transfer. <http://www.designboom.com/design/morpholio-photoplethysmography-technology-transfer-28-04-2014/>. (28 April 2014).
- [38] Lingqin Kong, Yuejin Zhao, Liquan Dong, Yiyun Jian, Xiaoli Jin, Bing Li, Yun Feng, Ming Liu, Xiaohua Liu, and Hong Wu. Non-contact detection of oxygen saturation based on visible light imaging device using ambient light. *Optics express*, 21(15):17464–17471, 2013.
- [39] Sungjun Kwon, Hyunseok Kim, and Kwang Suk Park. Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 2174–2177. IEEE, 2012.
- [40] CM Lee and YT Zhang. Reduction of motion artifacts from photoplethysmographic recordings using a wavelet denoising approach. In *Biomedical Engineering, 2003. IEEE EMBS Asian-Pacific Conference on*, pages 194–195. IEEE, 2003.
- [41] Chunming Li, Chenyang Xu, Changfeng Gui, and Martin D Fox. Distance regularized level set evolution and its application to image segmentation. *Image Processing, IEEE Transactions on*, 19(12):3243–3254, 2010.
- [42] Xiaobai Li, Jie Chen, Guoying Zhao, and Matti Pietikainen. Remote heart rate measurement from face videos under realistic situations. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 4264–4271. IEEE, 2014.

- [43] Dorothy Lui, Amen Modhafar, Jeffrey Glaister, Alexander Wong, and Masoom A Haider. Monte carlo bias field correction in endorectal diffusion imaging. *Biomedical Engineering, IEEE Transactions on*, 61(2):368–380, 2014.
- [44] Daniel McDuff, Sarah Gontarek, and Rosalind Picard. Remote measurement of cognitive stress via heart rate variability. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 2957–2960. IEEE, 2014.
- [45] Michel Misiti, Yves Misiti, Georges Oppenheim, and Jean-Michel Poggi. Matlab wavelet toolbox user’s guide. version 3. 2004.
- [46] Ming-Zher Poh, Daniel J McDuff, and Rosalind W Picard. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics express*, 18(10):10762–10774, 2010.
- [47] Ming-Zher Poh, Daniel J McDuff, and Rosalind W Picard. Advancements in non-contact, multiparameter physiological measurements using a webcam. *Biomedical Engineering, IEEE Transactions on*, 58(1):7–11, 2011.
- [48] Inc. Point Grey Research. Chameleon3. <https://www.ptgrey.com/chameleon3-usb3-vision-cameras>. (13 June 2016).
- [49] S. Prahl. *Optical absorption of hemoglobin*. Oregon Medical Laser Center, Portland, Oregon, USA, 1999.
- [50] PulseOximeter.org. Products. <http://pulseoximeter.org>. (14 June 2016).
- [51] R-B. Easy pulse (version 1.1) sensor overview (part 1). <http://embedded-lab.com/blog/?p=7336>. (20 April 2013).
- [52] Andr Roggan, Moritz Friebel, Klaus Drschel, Andreas Hahn, and Gerhard Mller. Optical properties of circulating human blood in the wavelength range 400–2500 nm. *Journal of biomedical optics*, 4(1):36–46, 1999.
- [53] Yu Sun, Sijung Hu, Vicente Azorin-Peris, Stephen Greenwald, Jonathon Chambers, and Yisheng Zhu. Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise. *Journal of biomedical optics*, 16(7):077010–077010, 2011.

- [54] Mika P Tarvainen, Perttu O Ranta-Aho, Pasi A Karjalainen, et al. An advanced detrending method with application to hrv analysis. *IEEE Transactions on Biomedical Engineering*, 49(2):172–175, 2002.
- [55] Carlo Tomasi and Takeo Kanade. *Detection and tracking of point features*. School of Computer Science, Carnegie Mellon Univ. Pittsburgh, 1991.
- [56] Mark van Gastel, Sander Stuijk, and Gerard de Haan. Motion robust remote-ppg in infrared. *Biomedical Engineering, IEEE Transactions on*, 62(5):1425–1433, 2015.
- [57] Wim Verkruysse, Lars O Svaasand, and J Stuart Nelson. Remote plethysmographic imaging using ambient light. *Optics express*, 16(26):21434–21445, 2008.
- [58] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–511. IEEE, 2001.
- [59] Peter D Welch. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, 15(2):70–73, 1967.
- [60] FP Wieringa, Frits Mastik, and AFW Van der Steen. Contactless multiple wavelength photoplethysmographic imaging: a first step toward spo2 camera technology. *Annals of biomedical engineering*, 33(8):1034–1041, 2005.
- [61] Alexander Wong, Akshaya Mishra, Wen Zhang, Paul Fieguth, and David A Clausi. Stochastic image denoising based on markov-chain monte carlo sampling. *Signal Processing*, 91(8):2112–2120, 2011.
- [62] Tadamasa Yamamoto, Hirotsugu Takiwaki, Seiji Arase, and Hiroshi Ohshima. Derivation and clinical application of special imaging by means of digital cameras and image j freeware for quantification of erythema and pigmentation. *Skin Research and Technology*, 14(1):26–34, 2008.
- [63] VK Yeragani, R Pohl, R Berger, R Balon, and K Srinivasan. Relationship between age and heart rate variability in supine and standing postures: a study of spectral analysis of heart rate. *Pediatric cardiology*, 15(1):14–20, 1994.