Video camera monitoring to detect changes in haemodynamics

DPhil Thesis

Jonathan Daly
St John's College

Submitted in Partial Fulfilment of the Requirements
for the Degree Doctor of Philosophy

Supervisor:
Prof. Lionel Tarassenko

Institute of Biomedical Engineering
Department of Engineering Science
University of Oxford

Michaelmas Term 2016

Abstract

Video camera monitoring to detect changes in haemodynamics

Patients in hospital can be prone to sudden, life-threatening changes in their cardiovascular state. Haemodynamic parameters such as blood pressure, pulse transit time (PTT) and perfusion can be monitored in clinical situations to identify these changes as early as possible. Continuous blood pressure is usually monitored using a catheter placed into a major artery, but this is invasive and involves risk to the patient. In the last decade, the field of non-contact vital sign monitoring has emerged, with growing evidence that the remote photoplethysmogram (rPPG) signal can be used to estimate vital signs using video cameras. If the analysis of the rPPG signal can be expanded to include the estimation of haemodynamic parameters, it could result in methods for the continuous, non-contact monitoring of a subject's haemodynamic state.

In a physiology study, a series of video recordings were made of 43 healthy volunteers. The subjects sat in a purpose-built chamber, and the composition of the air was carefully adjusted to cause the subjects to experience large, controlled changes in blood oxygen levels. To validate the video camera algorithms, reference data were also collected. Along with the volunteer study, a clinical study was performed to acquire data in a challenging clinical environment. Data were collected from patients on haemodialysis in the Renal Unit, a population likely to experience sudden changes in haemodynamics.

The reference data from the Renal Unit study were analysed to determine the extent to which PTT and mean arterial pressure (MAP) are related. The correlation coefficients and linear fits were found on a global and a per-subject basis. In addition, the video recordings from the Physiology study were processed to derive rPPG signals, and these signals were analysed to obtain estimates for PTT. Local rPPG signals were also derived for different regions of interest, and the waveforms were analysed using a novel application of the technique of signal averaging to produce spatial maps of perfusion and blood flow.

The correlation between conventionally measured PTT and MAP was found to be weaker in the haemodialysis population than has been shown elsewhere in the literature, except for a sub-set of patients. The results of the video analysis showed that PTT could be estimated robustly and consistently, although direct validation of these estimates was not possible because of the different method used to calculate the reference PTT. For most subjects, the spatial mapping methods produced robust maps that were consistent over time. These results suggest that it is possible to detect changes in haemodynamics using a video camera, and that this could have applications in healthcare, providing that challenges such as subject movement and clinical validation can be overcome.

Abstract (simple English)

Patients in hospital sometimes have sudden changes in their blood pressure and blood flow, which can be very dangerous. It is important to keep track of these changes very carefully in some patients, but this often requires a needle to be inserted into the patient's arm. In the last few years, researchers have shown that video cameras can be used to measure people's vital signs, such as their heart rate. If video cameras are also able to keep track of certain other measurements, it might be possible to monitor someone's blood pressure and blood flow without attaching sensors to them.

In an experiment, we made video recordings of 43 healthy volunteers. They sat in a room while being recorded, and the gases in the air were adjusted to change the amount of oxygen dissolved in their blood. Some medical measurements were also collected using a probe so that the video camera measurements could be checked later. We also collected video and medical measurements from hospital patients having dialysis treatment for kidney failure.

We analysed the data from the dialysis patients to see if there was a link between their blood pressure and a measurement called pulse transit time (PTT), which tells doctors how fast a patient's pulse travels along their blood vessels. We also analysed the video recordings from the volunteer experiment, and used them to estimate each volunteer's PTT. We then used the video recordings to make some maps of each volunteer's skin, showing how much blood was flowing and where it was travelling.

We found that there was a link between the dialysis patients' blood pressure and their PTT, but that in most patients the link was not as strong as researchers have suggested in other scientific papers. We also showed that it is possible to estimate PTT using only a video camera, but we were not able to check whether these estimates were accurate. We were able to produce maps of blood flow for most of the volunteers, which did not change much over time. This research suggests that it is possible to keep track of a patient's blood pressure and blood flow using only a video camera. This could eventually be used in hospitals in the future.

Acknowledgements

This thesis would not have been possible without the help and support of my supervisor, Prof. Lionel Tarassenko, whose guidance has been invaluable. I owe similar thanks to my colleagues in the research group—Mauricio Villarroel, Alessandro Guazzi, João Jorge, Hamish Tomlinson and Sitthichok Chaichulee—for their enthusiasm, knowledge and advice. I am also grateful for the expert assistance of our clinical team, who worked tirelessly to ensure our many studies were a success: Sheera Sutherland, Dr Clare MacEwen, Prof. Chris Pugh, Sara Davis, Dr Gabrielle Green, Dr Kenny McCormick and Dr Matthew Frise.

I have had the joy and privilege of being surrounded by the best friends, family and teammates I can imagine, and I could never thank them enough for their continued love and support. For them, I am immensely grateful.

This work was supported by the Engineering and Physical Sciences Research Council, the Wellcome Trust and the NIHR Biomedical Research Centre Programme, Oxford.

Declaration

I declare that this thesis is entirely my own work and, except where stated, describes my own research.

J. Daly
St John's College

Contents

List of Figures

List of Tables

List of Abbreviations

AAMI . . . . Association for the Advancement of Medical Instrumentation
AOI . . . . . Area Of Interest
ADC . . . . . Analogue-to-Digital Converter
BBI . . . . . Beat-to-Beat Interval
BHS . . . . . British Hypertension Society
BP . . . . . Blood Pressure
BR . . . . . Breathing Rate
BSS . . . . . Blind Source Separation
CCD . . . . . Charge-Coupled Device
CMOS . . . . Complementary Metal-Oxide-Semiconductor
cPTT . . . . . conventional Pulse Transit Time (from heart to finger)
dPTT . . . . differential Pulse Transit Time (between any two sites)
chk . . . . . cheek
CT . . . . . Computed Tomography
CVD . . . . . Cardiovascular Disease
DBP . . . . . Diastolic Blood Pressure
ECG . . . . . Electrocardiogram
ESRD . . . . End-Stage Renal Disease
FFT . . . . . Fast Fourier Transform
FIR . . . . . Finite Impulse Response
fhd . . . . . forehead
fps . . . . . frames per second
Hb . . . . . (Deoxygenated) Haemoglobin
HbO_2 . . . . . Oxygenated Haemoglobin
HR . . . . . Heart Rate
ID . . . . . Identifier
IP . . . . . Inductance Pneumography

LED . . . . . Light-Emitting Diode
MAE . . . . . Mean Absolute Error
MAP . . . . . Mean Arterial Pressure
mmHg . . . . millimetres of mercury
MRI . . . . . Magnetic Resonance Imaging
nck . . . . . neck
PPG . . . . . Photoplethysmogram
PTT . . . . . Pulse Transit Time
PWV . . . . . Pulse Wave Velocity
QRS . . . . . QRS (feature of ECG signal)
REC . . . . . Research Ethics Committee
RGB . . . . . Red Green Blue
RMSE . . . . Root Mean Squared Error
ROI . . . . . Region Of Interest
ref . . . . . reference
rPPG . . . . . remote photoplethysmogram
rPTT . . . . . remote Pulse Transit Time (from heart to any site)
SBP . . . . . Systolic Blood Pressure
SNR . . . . . Signal-to-Noise Ratio
SpO_2 . . . . . Peripheral Oxygen Saturation

Chapter 1

Introduction

1.1 Clinical motivation

The evolution of modern medicine has been driven heavily by technology. Early medical imaging techniques such as X-ray photography have largely been replaced with magnetic resonance imaging (MRI) and computed tomography (CT) scanning; open surgery can now be done instead using minimally-invasive laparoscopic techniques; cancer can now be treated with precisely aimed beams of radiation or targeted therapy drugs. Modern equipment and techniques have changed the way patients are diagnosed and treated.

Despite these advances in technology, patients in hospital still have their vital signs monitored in much the same way as fifty years ago. Heart rate (HR) is usually monitored using the electrocardiogram (ECG), which has been used clinically for over a century [26]. Nurses commonly estimate breathing rate (BR) by manually counting a patient's chest wall movements for a pre-determined period of time [67], measure temperature using a tympanic thermometer [103], and measure blood pressure (BP) using an inflatable arm cuff [10]. The last major change in routine patient monitoring was the introduction of pulse oximetry in the 1970s, which uses the photoplethysmogram (PPG) to measure the changes in the absorption of light by haemoglobin in the red blood cells during the

cardiac cycle. This allows the level of oxygen in the blood to be quantified in the form of peripheral oxygen saturation (SpO_2).

These conventional methods of vital sign monitoring are extremely well established, but they have disadvantages. For example, ECG leads and pulse oximeter probes are uncomfortable and restrictive to patient movement, and raise the risk of a hospital-acquired infection [170, 83]. In addition, some monitoring methods (such as BP measurement using an arm cuff) produce only intermittent measurements; continuous methods of monitoring are seen as important tools for the early identification of deteriorating patients [128]. There is therefore a clinical need for alternative ways of monitoring patients, and particularly for methods that are continuous and do not involve physical contact with the patient.

1.2 Non-contact monitoring

Within the last decade, a new field of research in biomedical engineering has emerged that seeks to address the problems of conventional monitoring. Non-contact vital sign monitoring (also called remote vital sign monitoring) uses video cameras to estimate the values of vital signs without the need for additional equipment. Inexpensive video cameras are embedded in laptops, mobile phones and webcams, hence non-contact monitoring could have applications for monitoring patients in their own homes. In addition, the idea of turning any consumer-grade video camera into a medical device is an attractive possibility for telemedicine and mobile health, particularly in regions with limited medical or financial resources.

While the field of non-contact monitoring is growing rapidly, early efforts were focused only on the estimation of HR and BR. Studies have shown that this can be done reliably and accurately in a clinical setting [153], and so the focus has now shifted to the other physiological parameters video cameras may be able to measure. One possibility is the

estimation of SpO_2 [44, 160, 1]; another is blood pressure (BP), which may be possible to estimate using other camera-derived parameters such as pulse transit time (PTT) [106, 139, 61]. PTT is the time taken for a blood pressure wave to travel between two points, and is introduced in Chapter 2.

1.3 Thesis overview

In this thesis, we investigate whether it is possible to detect changes in haemodynamic parameters using only a video camera. These haemodynamic parameters include BP, perfusion and direction of blood flow. To test this hypothesis, three main areas of investigation are considered:

1. **An investigation into the clinical utility of conventionally-measured PTT as a surrogate for BP in patients on haemodialysis.** While the relationship between PTT and BP is well-studied, this investigation concerns a specific population of hospital patients. Patients on haemodialysis often experience sudden drops in BP, but do not typically have their BP monitored continuously. The investigation in Chapter 3 will consider whether PTT could provide any useful information to clinicians as an alternative to continuous BP monitoring.

2. **A comparison between PTT measured conventionally and estimated from video of a subject's face.** To date, PTT has been measured from video in several studies, but each of these studies has used short video recordings, typically less than a minute long, and camera-derived estimates were not compared with reference values. In contrast, Chapter 5 uses extended video recordings and compares PTT estimates against a reference, providing a means of partial validation (although this is limited as the two measurements are not directly comparable).

3. **A novel method of mapping perfusion and blood flow in areas of exposed skin.** Chapter 6 proposes spatial mapping of both the amplitude and phase of the

local cardiac signal, and links them to the local level of perfusion and the direction of local blood flow respectively. Similar mapping has been demonstrated in several recent papers, but to our knowledge this method is the first to attempt continuous monitoring using such maps, which could have uses in the field of intensive care medicine. This could lead to a method of monitoring perfusion clinically using visible light, which has the potential to become a new modality of medical imaging.

The physiological background to PTT and its relationship with BP is provided in Chapter 2, and this relationship is then investigated in Chapter 3 in a population of patients on haemodialysis. The background of non-contact vital sign monitoring is explored in Chapter 4. Chapter 5 reports on a comprehensive investigation into the estimation of PTT using video cameras. Chapter 6 goes on to describe the methods and results of the perfusion mapping investigation. The thesis is concluded in Chapter 7, which includes suggestions for further work to continue the research.

1.4 Data

In order to develop and assess the performance of camera-based methods of detecting changes in haemodynamics, video recordings and reference data were collected as part a clinical study undertaken within the Oxford University Hospitals NHS Trust (Oxford, UK). The study was given approval by an NHS research ethics committee (REC). In addition, video and reference data were collected as part of a volunteer study in the University of Oxford's Department of Physiology, Anatomy and Genetics (Oxford, UK), which was given approval by the Central University REC.

1.4.1 Renal Unit study

Patients suffering from end-stage renal disease usually require haemodialysis treatment [96], in which blood is continuously removed, filtered of waste and returned to the body.

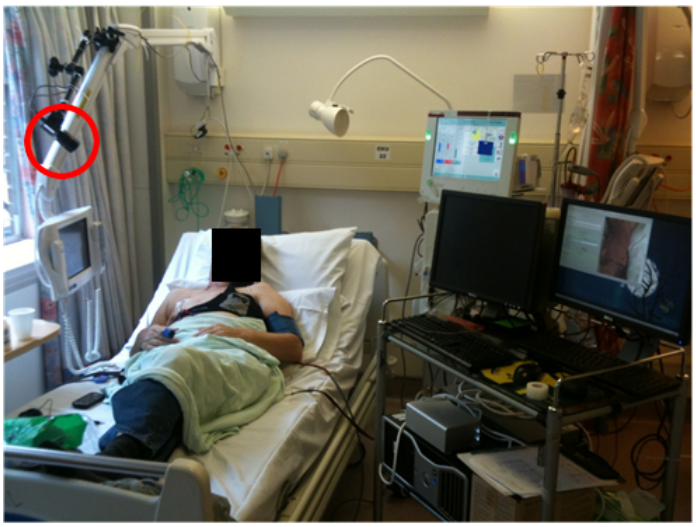


Figure 1.1: Patient being monitored during a haemodialysis session as part of the Renal Unit study. The video camera is circled in red.

Because of the changes in blood volume during haemodialysis, a patient's vital signs can undergo large changes during a haemodialysis session. This patient population is therefore an ideal group in which to investigate haemodynamic relationships.

In a clinical study[1], data were collected from 60 patients undergoing haemodialysis in the Renal Unit of Oxford's Churchill Hospital, with data from up to 12 separate haemodialysis sessions being recorded per subject[2]. Subjects were given an identifier (ID) of DFXX, where XX is a sequential numerical identifier. 21 patients did not consent to video recording of their sessions, and so only reference data were collected for these subjects. A total of 198 sessions from the remaining 39 subjects were recorded using the video camera. An example of the equipment setup is shown in figure 1.1.

Reference data

The reference data recorded included the following:

1. ECG, finger PPG and PTT recorded using a Black Shadow device (Stowood Scientific Instruments Ltd., Oxford, UK)

[1] NHS REC reference number 13/SC/0397
[2] One subject, DF05, gave consent to be recorded for 25 sessions.

   (a) ECG: sampling rate 256 Hz, 1 µV amplitude resolution

   (b) PPG: sampling rate 256 Hz, 16-bit amplitude resolution

   (c) PTT: sampling rate 5 Hz, 1 ms amplitude resolution

2. Continuous BP waveform, systolic blood pressure (SBP), mean arterial pressure (MAP) and diastolic blood pressure (DBP) recorded using a Finometer device (Finapres Medical Systems, Amsterdam, Netherlands), which uses the technique of vascular unloading (see section 2.3.3)

   (a) BP waveform: sampling rate 200 Hz, 0.1 mmHg amplitude resolution

   (b) SBP, MAP, DBP: one measurement per cardiac cycle, 1 mmHg amplitude resolution

3. Intermittent SBP, MAP and DBP (1 mmHg amplitude resolution) recorded by oscillometry using a Braun BP monitor (Braun GmbH, Kronberg, Germany)

Video camera data

The video camera used in this study was a Flea3 model (FL3-U3-13Y3M-C) manufactured by Point Grey (Point Grey Research Inc., Richmond, BC, Canada), recording raw uncompressed 12-bit monochrome video using a CCD sensor (see section 4.1.3) at a frame rate of 40 frames per second (fps) and a resolution of 1280×1024 pixels. Because each pixel required two bytes to store 10 bits of data, this meant that a four-hour recording of a haemodialysis session required around 1.5 terabytes of storage. A custom C/C++ application running on a dedicated computer (using the Linux-based Fedora operating system) was used to record the video data, which were synchronised with the reference data using the Network Time Protocol.

The high data transfer rates meant that the equipment ran at the limits of existing technology. Although the camera hardware was capable of higher frame-rates of up to 150 fps, the limitations of the other computing equipment meant that these higher frame

rates could not be sustained. 40 fps was therefore chosen as the highest frame rate that could be supported for robust data collection.

The camera was placed approximately 1 m away from the subject at the start of the session, and was manually focused on the subject's face. Camera parameters including exposure, gain, frame rate and shutter speed were set manually to a constant value. This prevented the camera firmware from adjusting the parameters automatically during the session. Through feedback with the clinical team involved in the study, the protocol was refined so as to make recordings as consistent as possible across all sessions and subjects. All recordings took place in the same hospital bed, and lighting was provided using a combination of natural light and the artificial fluorescent lighting in the ward.

Subject demographics

The demographics of the study subjects are given in table 1.1.

Total subjects		**60**
Sex (% male)		76.7
Age (years)		61.9 ± 15.9
Height (cm)		169.6 ± 9.1
Weight (kg)		81.6 ± 21.0
BMI ($kg\,m^{-2}$)		27.9 ± 6.4
Time in ESRD (months)		42.9 ± 76.0
Time on haemodialysis (months)		31.1 ± 37.8
Diabetes (%)		38.3
Ethnicity (%)	Asian	6.7
	Black	6.7
	White	86.7
Renal diagnosis (%)	Diabetic Nephropathy	21.7
	Hypertensive/Renovascular Disease	16.7
	Glomerulonephritis	31.7
	Polycystic Kidney Disease	3.3
	Other	26.7
12-month outcome (%)	Haemodialysis	50.0
	Kidney transplant	20.0
	Deceased	28.3
	Withdrew	1.7

Table 1.1: Demographics of all subjects in the Renal Unit study (mean ± standard deviation). BMI: body mass index, ESRD: end-stage renal disease, 12-month outcome: status of subject 12 months after enrolment in the study.

1.4.2 Physiology study

The Renal Unit study described in the previous section involved collection of real clinical data. Recording in a clinical environment, however, had a number of disadvantages. Firstly, the environment was largely uncontrolled, resulting in large changes in ambient lighting throughout recordings. Secondly, the movement of real hospital patients is not constrained; subjects in the Renal Unit study were able to move within the confines of their hospital bed. This tends to produce large movement artefacts in the camera-derived signals. Thirdly, the physiology of hospital patients tends to be highly non-uniform, making it difficult to optimise the algorithms for a healthy population.

Figure 1.2: A subject being monitored as part of the Physiology study.

To address these issues, a volunteer study[3] was conducted in the Department of Physiology, Anatomy and Genetics at the University of Oxford. In this study, volunteers were placed in a purpose-built chamber within which the inspired concentrations of nitrogen, oxygen and carbon dioxide could be carefully controlled to create an environment able to induce hypoxia (low blood oxygen) or hypercapnia (high blood carbon dioxide) [52]. The levels of these gases were adjusted by a qualified medical doctor during each 40-minute session such that the volunteers had SpO_2 values that followed a reproducible pattern. The lighting was also carefully controlled. The protocol for the study is described in full in a previous publication from our research group [44]. A total of 86 video recordings from 43 participants were made (43 hypoxia sessions, and 43 hypercapnia sessions). Subjects were given an ID of HP3-XX, where XX is a sequential numerical identifier. An example of the equipment setup is shown in figure 1.2.

[3]Central University REC reference number MSD-IDREC-C3-2014-003

Reference data

Reference data were recorded from the participants using two Black Shadow devices (see section 1.4.1). Participants had a pulse oximeter probe on the index finger of each hand, a three-lead ECG, and abdominal and thoracic elastic belts to measure BR using inductance pneumography (IP). The recorded signals consisted of the following:

1. ECG: sampling rate 256 Hz, 1 µV amplitude resolution
2. PPG: sampling rate 256 Hz, 16-bit amplitude resolution
3. PTT: sampling rate 5 Hz, 1 ms amplitude resolution
4. SpO2: sampling rate 32 Hz, 0.1% amplitude resolution
5. IP: sampling rate 64 Hz, 16-bit amplitude resolution

Video camera data

The camera used in the study was a 3CCD progressive scan AT-200CL model manufactured by JAI (JAI A/S, Copenhagen, Denmark), recording raw uncompressed video at a frame rate of 16 fps and a resolution of 1628×1236 pixels. The video is in 3-channel RGB format, with 12 bits per channel for a total of 6 bytes per pixel. This meant that a 40-minute recording session required around 550 gigabytes of storage space.

The software and equipment setup is similar to that of the Renal Unit study. The camera was positioned approximately 1.5 m from the subject. All windows were covered so as to minimise ambient lighting from outside the chamber, and subjects were illuminated using light emitting diode (LED) arrays. The interior of the chamber was lined with sheets of Tyvek 1443R (Dupont, Wilmington, DE, USA) to ensure controlled reflection from the light sources, two 30 cm 576-LED Mosaic lighting panels (Bowens International Ltd., Colchester, UK).

Study volunteers

All volunteers were healthy males between the age of 18 and 40 who gave their informed consent to participate in the study. Subjects were chosen such that there was a representative spectrum of skin colour in the cohort; each of the six skin types on the Fitzpatrick scale, a classification system for skin colour [38], were represented in the subject population. The following two exclusion criteria applied:

- Contraindication to exposure to hypoxia or hypercapnia, as judged by a clinician (e.g. epilepsy, respiratory disease)
- Claustrophobia preventing the subject entering the test chamber

Volunteers were compensated for their time, and were free to withdraw from the study at any time and for any or no reason.

1.4.3 Computing requirements

In both of the studies described in this chapter, raw uncompressed video data had to be collected and stored for future analysis. Raw video was required because video compression algorithms tend to remove small variations in colour, and this variation is crucial in performing physiological analysis of the data. The total amount of data recorded during the studies was in the order of one petabyte (one million gigabytes). Because these data had to be stored safely (with the appropriate levels of redundancy for backup purposes), the video files were compressed using the open source bzip2 lossless compression algorithm (`http://bzip.org/`), which allowed a reduction in file size of approximately 50%.

A significant proportion of the work required for this thesis went into creating and maintaining custom software for the purposes of data collection, storage and analysis[4].

[4]The codebase for this software was contributed to by several members of the research group, including M. Villarroel, A. Guazzi, J. Jorge and the author.

Chapter 2

Pulse Transit Time

Pulse transit time (PTT) is a haemodynamic parameter that can be measured to learn information about a patient's cardiovascular state. It is easily measured in a non-invasive way, and is often used to assess arterial stiffness[1]. PTT has also been proposed as a surrogate measurement for BP, and could potentially be used for continuous BP monitoring [40, 79, 121].

PTT is defined as the time taken for a blood pressure wave to travel between two points along an arterial path. PTT is related to the pulse wave velocity (PWV), the speed at which the blood pressure wave propagates along the vasculature, by the equation

$$PWV = \frac{L_{AB}}{PTT_{AB}} \tag{2.1}$$

where L_{AB} is the length of an arbitrary arterial path $\overrightarrow{AB}$, and PTT_{AB} is the pulse transit time along that path (see figure 2.1).

It is important to understand that the PWV is the velocity of the pressure wave travelling through the blood, and not the velocity of the blood itself. Using air as an analogy, the PWV corresponds to the speed of sound, whereas the velocity of the blood corresponds to wind speed. Conceptually, the PWV represents the group velocity of a

[1] Arterial stiffness is a measurement of how elastic the arteries are, and therefore how hard the heart must work to pump blood through them.

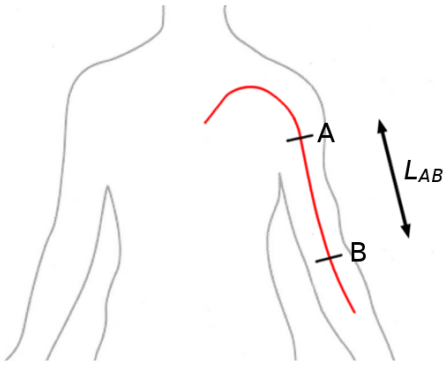


Figure 2.1: Illustration of PTT along an arterial path. PTT_{AB} is the time taken for the blood pressure wave to propagate from point A to point B along distance L_{AB}.

blood pressure wave, and the blood flow velocity represents the wave's phase velocity. To illustrate this difference, although it typically takes the blood approximately 20 seconds to arrive at the finger from the left ventricle of the heart [68, 179], PTT measured between the heart and finger is typically around 250 ms [141, 69].

This chapter explores the physiological background of PTT, and its relationship with BP. Section 2.1 contains an introduction to BP measurement; sections 2.2 and 2.3 introduce the ECG and PPG, two physiological signals which are often used to calculate PTT; and section 2.4 describes the conventional measurement of PTT using these signals. Finally, in section 2.5 the relationship between PTT and BP is explored.

2.1 Blood pressure

Blood pressure (BP) is an important indicator of cardiovascular health, and is routinely measured by clinicians. BP can provide information about heart function, blood vessel health and fluid balance, and abnormal BP is associated with a large number of medical conditions. Chronic hypertension (high BP) plays a major role in cardiovascular disease, and is the single greatest modifiable risk factor for ischaemic stroke [133]. Hypotension (low BP) can indicate more acute problems such as internal bleeding, particularly after traumatic injury or surgery. It is therefore important to measure BP accurately in the

context of acute medicine in the hospital, and important also to monitor certain patients over the course of months or years.

BP must be maintained by the heart in order for blood to be consistently delivered to the body's organs. If BP drops too low, organs such as the brain will not receive enough oxygen and nutrients; an excessively high BP will damage the vasculature. BP is therefore closely monitored and controlled by the body's sympathetic and parasympathetic nervous systems[2], which maintain a constant BP in a healthy individual, even in the presence of environmental changes (such as temperature swings or changes in posture). BP is maintained by constantly adjusting the HR, the stroke volume (the amount of blood pumped by the heart per contraction), the total volume of blood in the vasculature, and the degree of vasodilation or vasoconstriction in the blood vessels [24].

In most cases, a patient's BP is measured intermittently. The frequency of BP measurement depends on the situation; daily measurements are sufficient to diagnose and monitor long-term health problems such as chronic hypertension. In the case of acute hypotension, BP changes can occur much more rapidly; patients at risk of hypotension must therefore have their BP measured more frequently. One such group of patients is those in intensive care, where hypotension can indicate problems such as hypovolaemia (low blood volume, which can be caused by internal bleeding) or septic shock [33]. If not treated promptly and effectively, acute hypotension can lead to organ failure and death [101]. BP is usually monitored continuously in intensive care units using an invasive arterial line [82], which means that any decrease in BP can be detected as it occurs; this allows intervention to occur much earlier than intermittent BP monitoring—even on an hourly basis—would allow.

Although continuous BP monitoring has clear advantages, conventional methods of continuous monitoring are invasive and have associated risks (such as infection [136]).

[2]The sympathetic and parasympathetic nervous systems are the two branches of the autonomic nervous system, which is responsible for the unconscious regulation of internal organs and bodily functions. Loosely, the sympathetic nervous system is responsible for the body's "fight-or-flight" response, whereas the parasympathetic nervous system stimulates "rest-and-digest" activities.

Because of these risks, many patients who might otherwise be monitored continuously are not. Neonates in high-dependency units, for example, usually do not have an arterial line inserted despite the fact hypotension occurs in 20% of very low birthweight infants [28]. Similarly, patients on haemodialysis experience symptomatic hypotension during 20-30% of haemodialysis sessions [114], yet do not have their BP monitored continuously. For these groups, a non-invasive method of continuous BP monitoring would be of real benefit.

2.1.1 Conventional measurement

The heart is unlike most mechanical pumps, in that its beating action causes a pulsatile pressure over each cardiac cycle. BP is therefore recorded as a maximum or systolic blood pressure (SBP) and a minimum or diastolic blood pressure (DBP). SBP is the peak pressure reached during contraction of the heart's ventricles, and DBP is the lowest pressure during ventricular relaxation. Clinicians also measure mean arterial pressure (MAP), which is defined as the mean pressure over the whole cardiac cycle. MAP is usually estimated based on SBP and DBP using the formula

$$MAP \simeq \frac{2}{3}(DBP) + \frac{1}{3}(SBP) \tag{2.2}$$

where the ratio 2:1 comes from the relative duration of diastole and systole as the heart beats [102]. BP measurements are often presented in the form 120/80 (93), where the first two numbers indicate a typical SBP and DBP respectively [134], and the number in parentheses indicates the MAP estimated using equation 2.2. BP values are almost always given in units of millimetres of mercury (mmHg) since mercury manometers were the first accurate pressure gauges.

The values of SBP and DBP depend on where they are measured; as the pressure pulse travels distally along the vasculature, the SBP rises and the DBP falls. This is

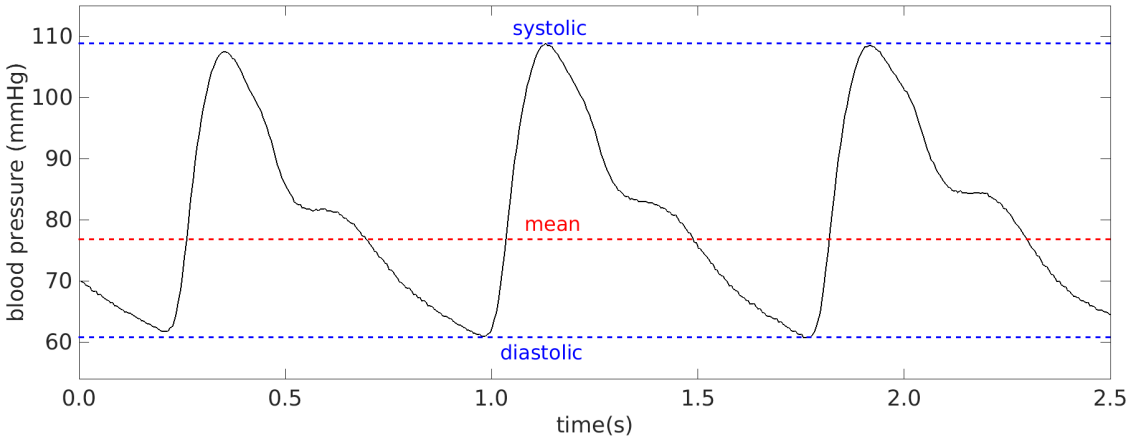


Figure 2.2: Example of an arterial blood pressure waveform, with SBP and DBP indicated in blue, and MAP indicated in red. The feature at approximately 1.4 s is the dicrotic notch, caused by the closure of the heart's aortic valve.

caused by reflection of the pressure pulse at vessel bifurcations, and by a reduction in blood vessel compliance moving downstream from the aorta [95].

BP values are not absolute, but are instead zero-referenced against atmospheric pressure. BP is usually measured non-invasively using a sphygmomanometer [10], but can also be measured invasively using an arterial line [143].

Invasive measurement

The conventional method of continuous BP measurement requires the placement of an arterial line. An arterial line consists of a catheter coupled with a pressure transducer, which is inserted into a major artery (commonly the radial or femoral artery) in order to measure the pressure directly. An arterial line is considered to be the most accurate method of BP measurement [165], but it also has a number of disadvantages: the process of placing the catheter is painful and invasive, and is associated with possible complications such as infection and bleeding [136].

Figure 2.2 shows an example of a blood pressure waveform recorded using an arterial line.

Non-invasive measurement

BP is usually measured non-invasively using a sphygmomanometer [10], consisting of an inflatable cuff, a pressure gauge and a pump. The cuff is put in position (normally around the upper arm), inflated to above SBP to restrict blood flow and then slowly allowed to deflate until the pressure drops below DBP and blood flow returns to normal. The pressure gauge measures the air pressure inside the cuff during this process, and the BP is determined either manually using the auscultatory method, or automatically using the oscillometric method. In the auscultatory method, a trained clinician uses a stethoscope to listen for characteristic noises (Korotkoff sounds); these noises are caused by the partially occluded blood flow through the arteries, and can be used to estimate SBP and DBP. The oscillometric method instead uses the oscillations in the pressure recorded by the gauge to determine when MAP is reached; MAP is determined by the pressure at which the oscillations have the largest amplitude, and values of SBP and DBP are extrapolated from the MAP. The oscillometric method is typically done automatically using electronic equipment to control cuff inflation and to record the BP measurements.

BP measurement using a sphygmomanometer is non-invasive, but is time consuming and can only be performed at intermittent intervals. The methods of determining BP from the gauge also have flaws. The auscultatory method relies on the skill of the listener to determine when salient pressures have been reached, which means that the accuracy of the method depends on the individual performing the measurements [6]. Modern measurement of BP using the oscillometric method is automatic and therefore independent of the user's skill level, but can sometimes produce inaccurate BP readings (for example when there is a mismatch between the size of the cuff and the patient's upper arm diameter [17]). Oscillometric devices also require calibration, and so their accuracy depends on the method used to calibrate them [9]. Studies have shown that the BP obtained from the auscultatory and different oscillometric methods can differ significantly; in a study of 381 children, Weaver *et al.* [166] found that auscultatory SBP

was 6.4 ± 6.5 mmHg lower and DBP was 8.7 ± 8.7 mmHg higher than oscillometric SBP and DBP respectively. Landgraf *et al.* [78] found a mean discrepancy over 337 patients of 1.95 ± 5 mmHg for SBP and 1.3 ± 4 mmHg for DBP between the two methods, with the auscultatory method resulting in consistently greater BP values than oscillometric devices. Using a different oscillometric device, Park *et al.* [119] found that it gave significantly higher readings than the auscultatory method. They compared the two methods in 7,208 children, finding a mean difference of 10.2 ± 7.0 mmHg for SBP and 4.7 ± 9.3 mmHg for DBP.

Validation

New methods or equipment for measuring BP must be validated against an existing method to ensure they are accurate. They are typically validated against invasive measurements obtained using an arterial line, which is considered to be the gold standard [165]. There are two widely used standards for evaluation of BP accuracy: the British Hypertension Society (BHS) protocol and the US Association for the Advancement of Medical Instrumentation (AAMI) standard [112]. The BHS protocol grades devices according to what proportion of their measurements are within a certain range of the gold standard, whereas the AAMI standard requires a mean error of no more than 5 ± 8 mmHg (mean ± standard deviation).

2.1.2 New methods of measurement

There is no established method of BP measurement that is both continuous and non-invasive, but a number of methods have been proposed. In 2003, a review of non-invasive continuous methods of BP monitoring by Parati *et al.* [118] highlighted two main fields of development: the vascular unloading technique, which is discussed in section 2.3.3, and the method of arterial tonometry. This latter technique involves measurement of the physical motion of major arteries as the blood flow pulses, accounting

for the non-linear relationship between vessel volume and pressure due to the elastic vessel walls. Matthys and Verdonck [93] reviewed the field in 2002, concluding that despite problems with motion artefacts and calibration, tonometry showed promise and could potentially be improved through combination with other physiological signals. Another study by Cheng *et al.* [21] found that although tonometry could be used to estimate central BP, the reliability of the method was poor when cuff BP was used to calibrate the peripheral waveforms recorded by the tonometry device. Devices including the T-line Tensymeter (Tensys Medical Inc., San Diego, CA, USA) [135] and the Vasotrac (Medwave Inc., Arden Hills, MN, USA) [11, 60] have been trialled clinically.

A more recent review in 2013 by Chung *et al.* [23] covered the techniques of vascular unloading and tonometry, but also included the use of PTT as a surrogate for BP, which is the main focus of this chapter. Chung's review also highlighted the expense associated with alternative methods of BP measurement compared to conventional ones.

A 2016 review of new techniques of haemodynamic monitoring by Teboul *et al.* [154] concluded that methods have "evolved considerably" from invasive equipment (like the arterial line) towards less invasive options, and from intermittent measurement to continuous monitoring. The authors noted that non-invasive monitoring is increasingly being used in critically ill patients, where they are considered reliable enough to displace conventional invasive monitoring. The methods assessed in the paper included waveform analysis of various physiological signals, oesophageal Doppler ultrasound, vascular unloading, arterial tonometry, and correlation with PTT (which is the focus of this chapter).

2.2 Electrocardiography

Electrocardiography measures the change in electrical potential across the heart over time. It is used to monitor HR, but is also useful in the diagnosis of conditions such as

arrhythmia (irregular heartbeat) or cardiovascular disease [30]. The electrocardiogram (ECG) can be used in combination with the photoplethysmogram (see section 2.3) to measure PTT, the subject of this chapter. This process is detailed in section 2.4.

ECG is measured using electrodes placed on a patient's skin. These electrodes detect the electrical activity of the heart as it beats, producing a periodic signal which can be displayed to clinicians for analysis. An ECG has a characteristic shape, which varies depending on the exact placement of the electrodes. The waveform has three main features per heartbeat: the P-wave, which is produced by the depolarisation of the atria; the QRS complex, which is due to ventricular depolarisation, indicating the start of ventricular contraction; and the T-wave, which is caused by the repolarisation of the ventricles. A diagram of an ECG waveform is shown in figure 2.3.

Ventricular depolarisation is sudden, and is the largest and most prominent feature in the ECG. The QRS complex is therefore used as a salient point when calculating HR; QRS detection algorithms such as the Pan-Tompkins detector (see section 5.7.3) are used to identify QRS complexes in an ECG waveform.

Abnormalities in the ECG waveform can be interpreted by trained clinicians and used in the diagnosis of cardiovascular conditions. For example, a raised segment in the ECG between the QRS complex and the T-wave can indicate acute myocardial infarction (a heart attack), which must be treated immediately.

Because the ECG is produced by the electrical activity of the heart, it provides information about the timing of the blood pressure pulse at the pulse's origin. This allows the timing of the BP pulse to be determined at its most proximal[3] point, without the pulse having travelled through any of the distal vasculature.

[3]In this context, proximal means close to the heart, and distal means far from the heart.

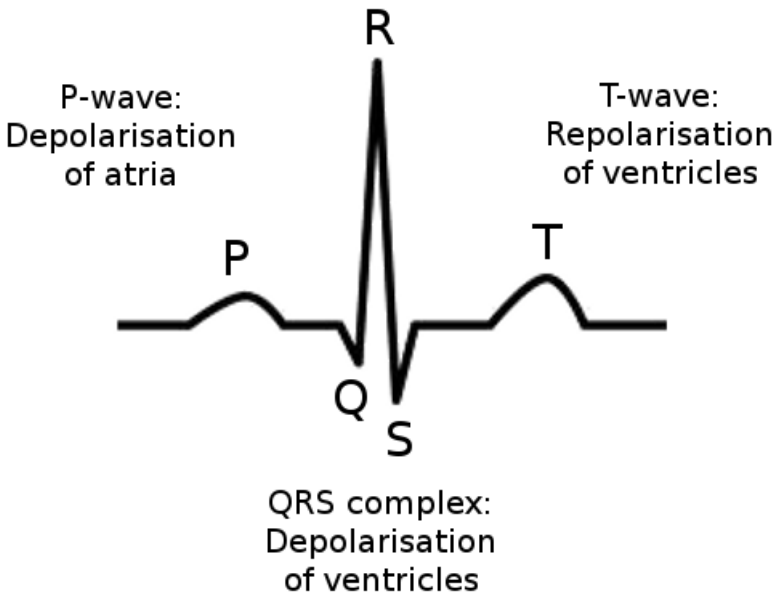


Figure 2.3: Annotated ECG waveform.

2.3 Photoplethysmography

Photoplethysmography uses light to measure the change in blood volume within blood vessels over time. The photoplethysmogram (PPG) can be used to derive important physiological parameters such as SpO_2, HR and BR [3]. Several methods of estimating BP using the PPG signal have also been proposed, which are discussed in section 2.3.3. PPG can also be used in combination with ECG to measure PTT, the subject of this chapter. This process is detailed in section 2.4.

2.3.1 Measurement

A PPG signal is typically obtained from a pulse oximeter probe attached to the finger or ear, which shines light into the skin using a light source (usually a pair of LEDs). A photodiode then measures the amount of light transmitted through the tissue (transmission PPG) or, less commonly, the amount reflected from it (reflectance PPG). This measurement is converted into a voltage which is then sampled to produce a

PPG time-series. By measuring the attenuation of the light after passing through the tissue, changes in the tissue's composition can be tracked over time. The PPG from tissue containing blood vessels has a DC component due to the absorption of light by bone and skin tissue, and an AC component corresponding to the pulsatile blood flow (all of which is assumed to be arterial) [3]. Figure 2.4 shows the origin of the PPG waveform.

The pulsatile PPG signal is similar in appearance to a BP waveform (compare figures 2.2 and 2.4, for example) because the pressure and volume within the blood vessel are closely related. At higher pressures, the vessel wall stretches to hold more blood (causing more light to be reflected from the tissue), and at lower pressures it relaxes to hold less blood (causing more light to be transmitted through the tissue). In the case of reflectance PPG, the waveform peaks during systolic pressure (when the blood vessels hold the most blood); in transmission PPG, however, the waveform peaks during diastolic pressure (when more light passes through the emptier blood vessels). Transmission PPG waveforms are usually inverted before being displayed so that they peak during systolic pressure[4].

[4]Clinicians are used to seeing PPG waveforms peak during systolic pressure, and the inversion of transmission PPG waveforms prevents them from becoming confused.

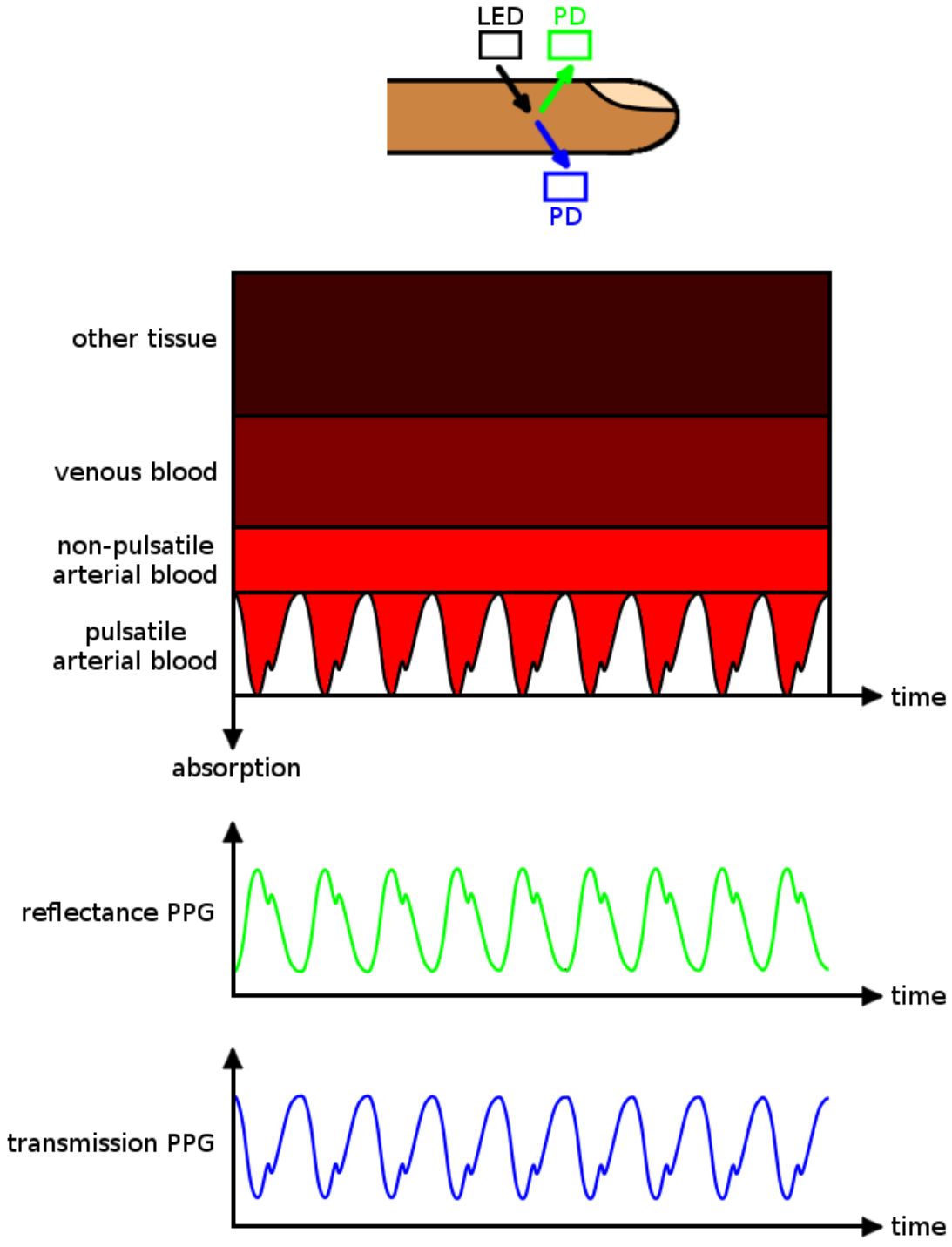


Figure 2.4: Origin of the PPG waveform. Top: reflectance PPG (green) and transmission PPG (blue) measured at the finger using LEDs and photodiodes. Middle: absorption of light over time by different tissue components. Bottom: resulting PPG waveforms.

2.3.2 Oxygen saturation

Blood oxygen saturation can be measured using a pulse oximeter by comparing the relative absorption of light at different wavelengths. Haemoglobin, which is a protein responsible for carrying oxygen in the blood, absorbs these wavelengths differently depending on whether oxygen is bound to it (see figure 2.5). Well oxygenated blood contains mostly oxyhaemoglobin, which absorbs relatively more infrared light than red light; poorly oxygenated blood contains more deoxygenated haemoglobin, which absorbs relatively more red light than infrared light. Pulse oximeters typically use LEDs of two characteristic wavelengths, usually 660 nm (red) and 910 nm (near infra-red), which are strobed alternately at high frequency to compare the blood's relative absorption.

Peripheral oxygen saturation (SpO_2) measures the percentage of haemoglobin in the peripheral vasculature that is saturated with oxygen. SpO_2 typically lies between 95% and 99%. It is defined as

$$SpO_2 = 100\% \times \frac{[HbO_2]}{[HbO_2] + [Hb]} \tag{2.3}$$

where $[HbO_2]$ is the concentration of oxyhaemoglobin and [Hb] is the concentration of deoxygenated haemoglobin in the peripheral blood. This concentration can be determined by measuring the attenuation of light as it passes through the tissue. SpO_2 then becomes a function of the ratio R of the PPG signal's DC and AC components at each wavelength. R is expressed mathematically as

$$R = \frac{\log_{10}\left(\frac{I_{AC}}{I_{DC}}\right)_{\lambda 1}}{\log_{10}\left(\frac{I_{AC}}{I_{DC}}\right)_{\lambda 2}} \tag{2.4}$$

where I_{AC} is the amplitude of the pulsatile signal, I_{DC} is the DC amplitude, and $\lambda 1$ and $\lambda 2$ are the two characteristic wavelengths. Appendix A describes the derivation of this equation.

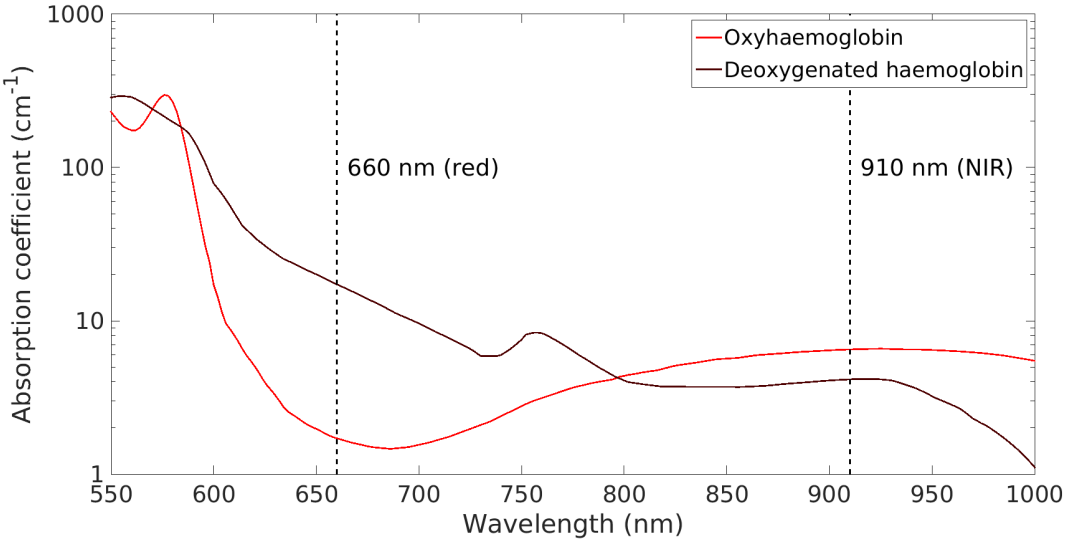


Figure 2.5: Absorption spectrum of haemoglobin in its oxygenated and deoxygenated forms. NIR: near infra-red. Data compiled by S. Prahl [127], publicly available at http://omlc.org/spectra/hemoglobin/summary.html.

The relationship between R and SpO_2 is complicated by the light-scattering behaviour of red blood cells, and so in practice a lookup table is used to compute SpO_2 from R. The accuracy of SpO_2 measured using pulse oximetry is typically within 2-3% for values above 80% [47] (it is considered unethical to test lower values, and so the accuracy decreases below an SpO_2 of 80%).

2.3.3 Estimation of BP from PPG

The AC component of the PPG signal is caused by the pulsatile nature of BP, and the BP and PPG waveforms are therefore closely related. This has enabled the oscillometric method (see section 2.1.1) to be adapted by some research groups to use the PPG signal instead of the output from a pressure sensor; the PPG probe is placed downstream of the arm cuff, and the PPG waveform is used to determine when MAP (and, by extrapolation, SBP and DPB) is reached in the cuff pressure [81, 111].

Because of the relationship between the BP and PPG waveforms, many studies have proposed methods of estimating BP from PPG, which would allow continuous (and

non-invasive) BP monitoring. While PPG amplitude is correlated with change in blood volume, this does not translate directly into change in pressure due to the elastic nature of the arterial wall, which makes the relationship between pressure and blood volume non-linear.

Vascular unloading

The vascular unloading technique seeks to overcome this non-linearity by measuring the pressure required to keep the blood volume constant. It was first proposed in 1973 by Peňáz [123], who combined a pressure cuff with a pulse oximeter to be placed over the finger. The PPG signal produced by the oximeter was fed into a feedback control loop, and the output was used to control the pressure in the cuff via an electro-pneumatic transducer. By constantly adjusting the cuff pressure so that the blood volume in the finger remained constant (indicated by a constant finger PPG), Peňáz was able to estimate BP continuously. However, the system was not able to adjust to changes in vessel tone (vasodilation and vasoconstriction), which could cause a change in the PPG amplitude independent of the BP; the system therefore produced inaccurate estimates of BP after such changes in vessel tone.

A commercial device exploiting the principle of vascular unloading was developed by Finapres Medical Systems (Amsterdam, Netherlands), which entered the market with Finapres in 1986. Their technology was validated in a number of studies in the late 1980s: Imholz *et al.* [59] investigated the reliability of Finapres in subjects undergoing procedures designed to induce rapid changes in BP, while Parati *et al.* [117] compared the device to an arterial line in patients both at rest and during laboratory testing. In the last decade the same company has produced the Finometer, which incorporates automatic calibration using the oscillometric signal to account for changes in vessel tone and to mitigate other confounding factors in BP measurement. The newer device also allows the reconstruction of brachial arterial pressure from the pressure measured at the

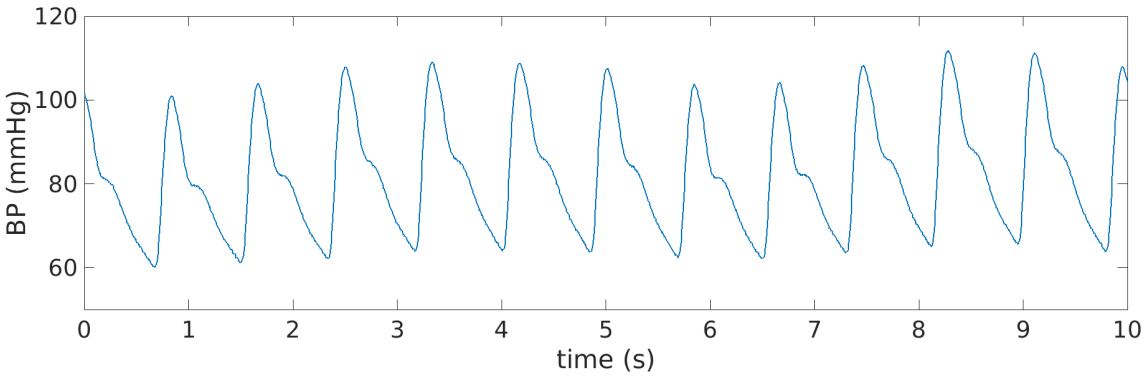


Figure 2.6: Reconstructed brachial BP waveform recorded by the Finometer device. Data from subject DF05 of the Renal Unit clinical study.

finger [45]. Many studies have favourably compared the device's measurements to those obtained from conventional measurements [45, 122, 137], although some have questioned whether its accuracy is sufficient for use in critical care medicine [37, 91]. Figure 2.6 shows an example of the BP waveform recorded by the Finometer device.

Waveform shape

Other studies have attempted to estimate BP values using parameters derived from the shape of the PPG waveform. In 2008, Suzuki and Oguri [150] used multiple regression analysis to predict SBP in 96 healthy subjects. They used parameters derived from the waveform and the acceleration pulse wave (the twice-differentiated PPG signal), reporting a standard deviation of 8.2 mmHg in the error between their predicted SBP and reference BP measured using the auscultatory method. Suzuki and Oguri proposed another method in 2009 [151] using features obtained from the shape of the PPG signal as the inputs to a series of weak binary classifiers, which were then aggregated into a multi-class strong classifier. The authors trained the classifier on 195 subjects, and found an error of -1.2 $\pm$ 11.7 mmHg between estimated and measured SBP on a 21-subject test set. Both of these studies had greater variance in their measurement errors than the 8 mmHg of standard deviation allowed by the AAMI [112], and the limited population of the two studies makes it difficult to assess the potential utility of the approach.

Monte-Moreno [100] proposed an alternative approach in 2011. A variety of machine learning techniques (including ridge linear regression, a multilayer perceptron neural network, support vector machines and random forests) were tested, and the random forest technique was found to be the most accurate for BP estimation. PPG-based features were used as the inputs to the random forest classifier, 10-fold cross-validation was used to train the algorithms on a training/validation set of 328 individuals, and the classifier was tested on 82 test subjects. Monte-Moreno found that 91.0% of the SBP measurements and 93.9% of the DBP measurements were within 15 mmHg of the reference, meeting the BHS criteria for a grade B device [112] but not for a grade A device. In addition, the BP produced by Monte-Moreno's techniques was not continuous.

In 2013, Ruiz-Rodríguez *et al.* [132] performed a study on 572 critically ill adult patients in which BP calculated from the PPG waveform was compared to invasive arterial line monitoring. The authors used a set of parameters derived using waveform analysis as inputs to a regression model, and found that the error between SBP measured using the two methods was -2.98 $\pm$ 19.35 mmHg. Although the mean error for the proposed method was less than the AAMI requirement of 5 mmHg, the standard deviation was much greater than the allowed 8 mmHg [112].

2.4 Conventional measurement of PTT

Although PTT can be measured along any length of the arterial vasculature, it is conventionally measured between the heart and the finger, as the arterial path that runs between these two sites is one of the longest in the body. The heart is almost always used as the proximal site, but other options for the distal site include other extremities such as the toes and ears [110, 171].

PTT is defined in terms of the blood pressure wave's arrival, but this is usually difficult to measure directly. Instead, PTT is conventionally measured in a clinical setting

using the ECG (see section 2.2) and the finger PPG (see section 2.3).

The ECG's most prominent feature is the QRS complex (see section 2.2), which indicates the start of ventricular contraction. The R-wave of the ECG is therefore used as a proximal reference point for the departure of the blood pressure pulse. The distal reference point is the systolic rise in the PPG, which indicates the arrival of the blood pressure pulse. The exact point to be used is a subject of some debate in the literature, but a common choice is the point in the PPG that lies halfway up the systolic rise [141]. PTT measured in this way will be referred to in this thesis as cPTT (where the c denotes 'conventional'), and can be broken down into two components: the pre-ejection period between the manifestation of the R-wave and the onset of mechanical contraction of the ventricles, and the period in which the blood pressure wave propagates along the vasculature [84].

Although cPTT is the conventional method of measurement, PTT can be measured between any two points in the arterial vasculature using any cardiac signal recorded at those points. Figure 2.7 describes the calculation of cPTT using ECG and PPG, and of any other PTT using PPG signals recorded at two sites. Note that when measuring PTT between two PPG signals, any reference point can be chosen as long as it is consistent between the signals; the peak of the signal is usually used as it is simple to locate.

An example of calculating cPTT using real waveforms (from the dataset described in section 1.4.1) is shown in figure 2.8.

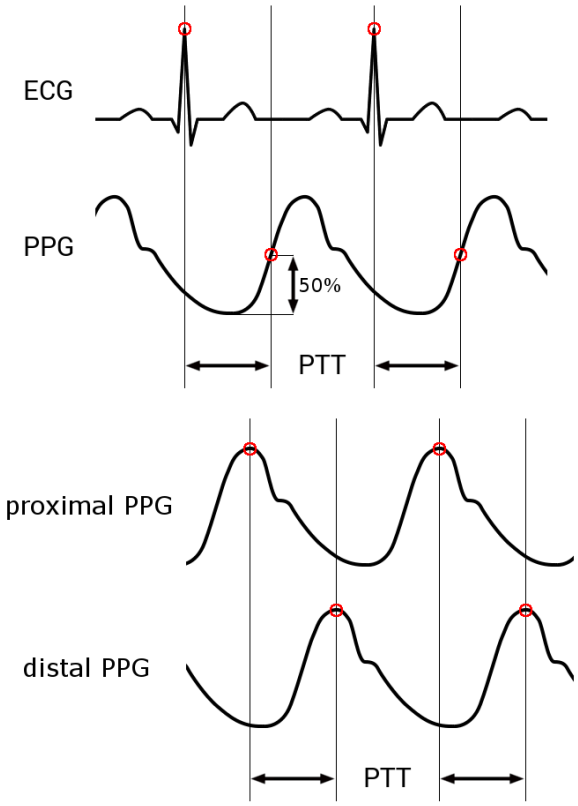


Figure 2.7: Methods of calculating PTT. Left: cPTT from ECG and PPG recorded distally at the finger. Right: PTT from two PPG signals, one proximal and one distal.

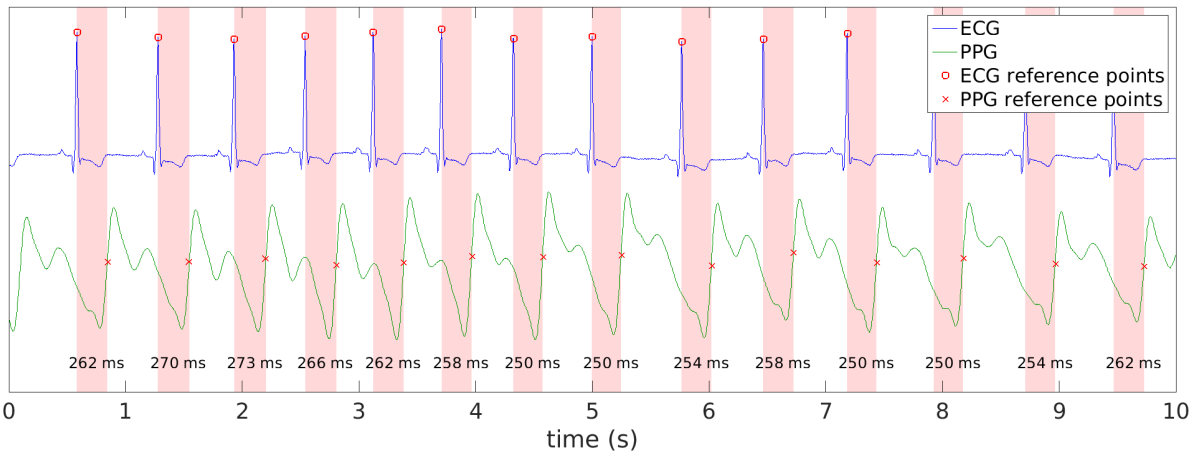


Figure 2.8: Example showing calculation of cPTT from real synchronised ECG and PPG. For each pulse, the red shaded area represents the cPTT.

2.5 Relationship between PTT and BP

The link between PTT and BP was first investigated in 1955 by Thomas [156]. Section 2.5.1 explains the theory behind the relationship, and section 2.5.2 discusses some of the studies that have investigated its clinical significance.

2.5.1 Theory

When a substance is disturbed by a change in pressure, a pressure wave propagates through it with a velocity c. That velocity (often referred to as the speed of sound) depends on the substance's material properties, and can be expressed as

$$c = \sqrt{\frac{\partial p}{\partial \rho}} \tag{2.5}$$

where p is the pressure and ρ is the mass density of the medium. Alternatively, Hooke's law can be used to express the speed of sound as

$$c = \sqrt{\frac{E}{\rho}} \tag{2.6}$$

where E is the bulk modulus (or modulus of bulk elasticity in the case of gases). From this equation it can be seen why c is usually much larger in solids and liquids, which tend to be much denser than gases. For example, the speed of sound in air at sea level is approximately $340\,\mathrm{m\,s^{-1}}$, whereas the speed of sound in water at $25\,^{\circ}\mathrm{C}$ is approximately $1497\,\mathrm{m\,s^{-1}}$.

While calculating the speed of sound in air is relatively simple, calculating the speed of a blood pressure pulse is more complicated. This is because the blood is flowing through a blood vessel with elastic walls that are capable of stretching and distorting. The relationship between PTT and BP is therefore affected by the stiffness and distensibility of arteries, which have thick walls composed of muscle and elastic fibres.

An equation for this situation was derived independently in 1878 by Moens [98, 99] and Korteweg [75], who coupled Newton's second law of motion with some simplifying assumptions: that the vessel wall is isotropic, and that there is insignificant change in vessel area and wall thickness during the pulse. This produced the Moens-Korteweg equation, which expresses the PWV in an elastic vessel as

$$c = \sqrt{\frac{E_{inc}h}{2r\rho}} \tag{2.7}$$

where c is the PWV, E_{inc} is the incremental elastic modulus of the tube wall, h is the wall thickness, r is the tube radius and ρ is the mass density of the blood.

The Moens-Korteweg was adapted by Bramwell and Hill [15] in 1922 to use easily measurable haemodynamic parameters. The Bramwell-Hill equation expresses the PWV as

$$c = \sqrt{\frac{V}{\rho}\frac{\partial P}{\partial V}} \tag{2.8}$$

where V is the blood volume and P is the BP.

From the Bramwell-Hill equation, the link between PWV and BP is clear. As BP increases, the elastic blood vessel stretches (causing V to increase) and becomes less compliant (causing the $\partial P/\partial V$ term to increase). An increase in BP therefore causes an increase in PWV (or equivalently, a decrease in PTT).

Heard *et al.* [50] further describe the derivation of a relationship between PTT and BP by assuming that blood is incompressible (i.e. ρ is constant), a common assumption for a water-like fluid. This allows the change in pressure during the pulse to be simplified to

$$\Delta P \propto c^2 \left(\frac{\Delta V}{V}\right) \tag{2.9}$$

where ΔP is the change in pressure and $\Delta V/V$ is the fractional change in vessel

volume as the pulse propagates within the vessel. Recalling from equation 2.1 that PTT is inversely related to the PWV, it follows that an inverse relationship exists between pulse pressure ΔP and the PTT. This relationship is complicated by the presence of the $\Delta V/V$ term, which depends on the vascular tone (the degree to which a blood vessel is constricted or dilated). In addition, the use of PTT to estimate absolute values such as SBP, DBP and MAP requires calibration as only the pulse pressure ΔP (the difference between SBP and DBP) is related to PTT as described in equation 2.9.

Other expressions of the relationship between PTT and BP have been proposed. In a recent review by Mukkamala *et al.* (2015) [105], the authors note three popular models. The first is derived from the Moens-Korteweg equation by expressing E_{inc} as

$$E_{inc} = E_0 e^{\alpha P} \tag{2.10}$$

where E_0 and α are subject-specific parameters. Substituting this into equation 2.7 and rearranging produces

$$P = -\frac{2}{\alpha} \ln (PTT) + \frac{1}{\alpha} \ln \frac{2r\rho L^2}{hE_0} \tag{2.11}$$

which can be simplified to

$$P = K_1 \ln (PTT) + K_2 \tag{2.12}$$

where K_1 and K_2 are subject-specific parameters which must be determined by regression of paired PTT and BP values.

Mukkamala *et al.* also discuss the model

$$P = \frac{K_1}{PTT} + K_2 \tag{2.13}$$

which is derived by considering vessel compliance. As the authors note, this model is

supported by a number of studies over a wide range of BP values [40, 178, 104].

Finally, a simpler model is used throughout the literature: studies assume PTT is linearly related to BP, describing the relationship simply as a gradient and intercept. Thus the relationship has the form

$$P = K_1(PTT) + K_2 \tag{2.14}$$

where K_1 and K_2 are the patient-specific gradient and intercept respectively.

These simple models are not asymptotically correct as PTT approaches 0 or ∞, and others have sought to address this using more complicated non-linear models [167], but Mukkamala *et al.* point out the advantages of the simpler models, which are accurate within physiologically plausible ranges of BP. Because of this limited physiological range, and taking into account the many confounding factors of PTT, simple linear regression will be used in this thesis to quantify the relationship between PTT and BP. Chapter 3 includes a comparison of the models presented in this section.

2.5.2 Empirical relationship in the literature

Several studies have attempted to assess the accuracy of BP estimated using PTT as a surrogate. Early studies in the field of psychophysiology found relationships between PTT and SBP, but disagreed about the relationship between PTT and DBP or MAP [144, 40, 79]. In 2000, Chen *et al.* [20] continuously estimated SBP by combining higher-frequency PTT information with lower-frequency BP measured intermittently using an arm cuff. This process essentially involved repeated calibration using the arm cuff, and adjustment of the continuous SBP estimates from changes in the value of the PTT. The authors showed that it was possible to track changes in SBP in this way, but recognised the need for repeated calibration to obtain estimates that were accurate over longer periods of time.

As part of a clinical evaluation of the DxTek monitor (DxTek Inc., Chestnut Hill, MA, USA), which produces continuous BP estimates using PTT measurements, Heard *et al.* [50] found an accuracy comparable to that of oscillometric devices. The DxTek monitor calculates BP values based on the cPTT measured from ECG and PPG, the instantaneous HR, and calibration using an initial external measurement of BP. The authors found that when calibrated using an arterial line, the error in SBP between monitor and arterial line was 0.5 ± 10.1 mmHg, compared to -4.0 ± 9.6 mmHg between oscillometry and arterial line measurements. A larger study on 85 subjects by Poon and Zhang [126] in 2005 found that after patient-specific calibration, cPTT could be used to estimate SBP with an error of 0.6 ± 9.8 mmHg, and DBP with an error of 0.9 ± 5.6 mmHg. They found that this was comparable to the error between BP measured by a clinician using the auscultatory method and the oscillometric method.

Others have questioned the use of PTT as a surrogate for BP. A study of 12 subjects by Payne *et al.* [121] in 2006 found that although there was a correlation between cPTT and SBP, it was not reliable enough to be used a surrogate measurement. The authors acknowledged, however, that cPTT may be useful as a surrogate for change in BP. In 2002, Nitzan *et al.* [110] explored the relationship between PTT and age, finding that PTT decreased as a function of the subject's age and SBP; since age is a further confounding factor in the PTT-BP relationship, calibration would therefore be required on an individual basis. This further reinforces the hypothesis that PTT may be better used as an indicator for changes in BP rather than to determine absolute values.

A summary of studies investigating the relationship between PTT and BP is given in table 2.1.

Study	N	Outcome
Steptoe *et al.* (1976) [144]	5	$r = -0.96$ (cPTT $\rightarrow$ MAP)
Geddes *et al.* (1981) [40]	6	$r = -0.82$ (cPTT $\rightarrow$ DBP in dogs)
Lane *et al.* (1983) [79]	4	$r = -0.62$ (cPTT $\rightarrow$ SBP) $r = -0.48$ (cPTT $\rightarrow$ MAP) $r = -0.32$ (cPTT $\rightarrow$ DBP)
Chen *et al.* (2000) [20]	20	$e = +0.06 \pm 0.68$ % (SBP estimate $\rightarrow$ reference)
Poon & Zhang (2005) [126]	85	$e = +0.6 \pm 9.8$ mmHg (SBP estimate $\rightarrow$ reference) $e = +0.9 \pm 5.6$ mmHg (DBP estimate $\rightarrow$ reference)
Ahlstrom *et al.* (2005) [2]	10	$r = +0.80$ (SBP estimate $\rightarrow$ reference)
Payne *et al.* (2006) [121]	12	$r = -0.62$ (cPTT $\rightarrow$ SBP) $r = -0.28$ (cPTT $\rightarrow$ MAP) $r = -0.14$ (cPTT $\rightarrow$ DBP)
Douniama *et al.* (2009) [35]	22	$r = +0.43$, $e = 6.7$ mmHg (SBP estimate $\rightarrow$ reference) $r = +0.42$, $e = 5.6$ mmHg (MAP estimate $\rightarrow$ reference) $r = +0.37$, $e = 5.3$ mmHg (DBP estimate $\rightarrow$ reference)
McCarthy *et al.* (2011) [94]	6	$e = -0.8 \pm 5.3$ mmHg (SBP estimate $\rightarrow$ reference)
Gesche *et al.* (2012) [41]	50	$r = +0.83$ (SBP estimate $\rightarrow$ reference)
Patzak *et al.* (2015) [120]	12	$r = +0.95$ (SBP estimate $\rightarrow$ reference) $r = +0.42$ (DBP estimate $\rightarrow$ reference)

Table 2.1: Summary of studies investigating the relationship between PTT and BP. N: number of subjects, r: correlation coefficient, e: error (mean $\pm$ standard deviation).

2.5.3 Confounding factors

In reality, there are a large number of confounding factors when considering the correlation between PTT and SBP. These confounding factors include age, gender, BR and the presence of cardiovascular conditions such as hypertension, diabetes and hypercholesterolaemia [80, 176, 97, 39]. In practice, therefore, PTT may be useful only in detecting changes in BP (as discussed in section 2.5.2) rather than in measuring absolute values of SBP, DBP and MAP, even with calibration for individual patients [41]. Examples of cases where the PTT-MAP relationship is atypical can be seen in Chapter 3.

Chapter 3

PTT in patients on haemodialysis

This chapter will investigate the relationship between pulse transit time (PTT) and mean arterial pressure (MAP) in a population of patients undergoing haemodialysis treatment. This population is unusual in that it is very heterogeneous, containing a range of ages, demographics and comorbidities. Since most studies into the relationship between PTT and MAP have focused on healthy subjects and more homogeneous groups (see section 2.5), it is likely that the strength and nature of the relationship will vary throughout the haemodialysis population.

Section 3.1 introduces the physiology of haemodialysis. Section 3.2 quantifies the relationship between PTT and MAP, globally and on a per-subject and per-session basis. Section 3.3 goes on to examine the data longitudinally, assessing the extent to which data between sessions is reproducible, before the chapter is concluded in section 3.4, which discusses the results.

3.1 Physiology of haemodialysis

Haemodialysis is a process in which blood is continuously removed, filtered of waste and returned to the body. This task is performed by the kidneys in healthy individuals, but must be done artificially in the case where a patient's kidneys are failing to work

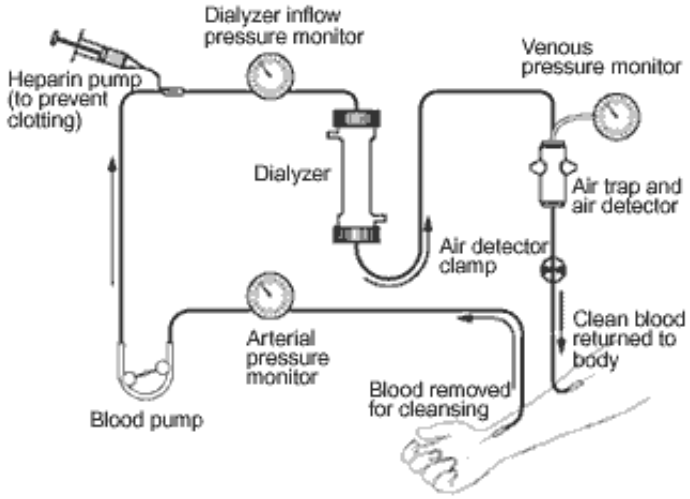


Figure 3.1: Diagram of a simple haemodialysis circuit (Source: National Institute of Diabetes and Digestive and Kidney Diseases, USA [113]).

normally. Haemodialysis allows the removal of waste products such as urea from the blood, which would otherwise build up to dangerous levels.

Patients on haemodialysis have their blood removed via a catheter, which is placed into a major blood vessel (often an arteriovenous fistula in the arm, a surgically created link between an artery and vein). The blood is then pumped into a dialyser, where it flows in counter current exchange with dialysate. The waste products in the blood diffuse into the dialysate across a semi-permeable membrane, but the composition of the dialysate is such that diffusion of other components (such as salts) is minimised to prevent loss. The fluid composition of the blood is also monitored and adjusted during this process. Finally, the cleaned blood is filtered to remove dangerous air bubbles, and returned to the body via another catheter. A diagram of the process of haemodialysis is shown in figure 3.1.

Haemodialysis is typically required two to three times a week, and sessions last around four hours [27]. The vital signs of patients on haemodialysis can change dramatically throughout a session due to the changes in fluid balance; removing too much fluid causes symptoms such as cramps, nausea, headaches and dizziness. Some patients are particularly prone to intra-dialytic hypotension, in which BP drops suddenly and

intervention is required [115, 29]. These sudden changes in BP make these patients a particularly interesting population to study when investigating the link between PTT and BP.

3.1.1 End-stage renal disease

End-stage renal disease (ESRD) is the most severe stage of chronic kidney disease. According to the National Kidney Foundation in the USA, ESRD is reached either when a patient requires haemodialysis, or when their glomerular filtration rate drops below 15 ml/min/1.73m^2 [71]. Patients with ESRD require regular haemodialysis sessions [96], and ideally a kidney transplant. Common comorbidities include increasing age, diabetes, cardiovascular disease (see section 3.1.2), and poor nutrition [129]. As well as sudden hypotension, patients on haemodialysis are also prone to long-term hypertension which can influence renal outcome [130].

3.1.2 Cardiovascular disease

Cardiovascular disease (CVD) is a common comorbidity of ESRD; over 50% of the deaths of ESRD patients are caused by CVD [25], a far greater proportion than in the general population. The overwhelming majority of ESRD patients are hypertensive, possibly due to inadequate haemodialysis; there is evidence that longer haemodialysis sessions are associated with improved control of BP [88]. Hypertension is a major risk factor for CVD [142], and therefore certainly contributes to this higher rate of CVD in ESRD patients.

3.2 Correlation between PTT and BP

In order to investigate the correlation between PTT and BP in the haemodialysis population, reference data were collected as part of the Renal Unit study described in

section 1.4.1. Continuous BP was recorded using the Finometer device, and PTT was recorded using the Black Shadow device. Continuous BP was chosen for this analysis rather than intermittent cuff BP, as each session produced hours of continuous BP data but only a few cuff measurements.

In the literature, conventional heart-to-finger PTT (cPTT, see section 2.4) has been variously compared with SBP, DBP and MAP. In this chapter, MAP has been used because it was shown by MacEwen (2016) [89]—based on an analysis of the same data set—to correlate more strongly with cPTT.

3.2.1 Global

The global correlation between cPTT and MAP was first found across all subjects. The raw data were noisy, with calibration artefacts present throughout the Finometer waveforms and non-physiological high-frequency variation present in both MAP and cPTT signals. This noise was dealt with by aggregating paired cPTT MAP measurements for each haemodialysis session, with one measurement per 60-second time period (see figure 3.2 for an example). For each haemodialysis session, the measurements were found as follows:

1. The raw time-series of both cPTT and MAP during the haemodialysis session were

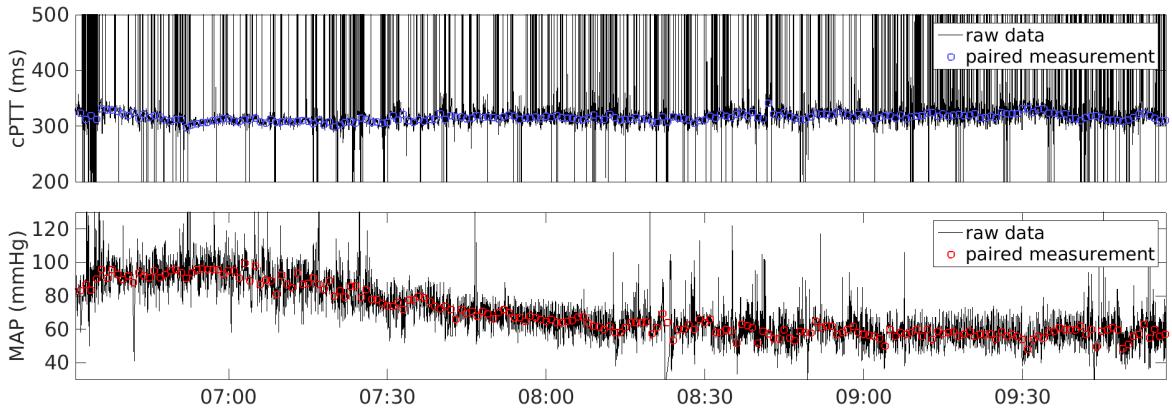


Figure 3.2: Example of raw cPTT and MAP data during a haemodialysis session recorded from subject DF05.

loaded

2. The time-series were processed to flag measurements far outside the possible physiological range. The criteria for flagging were extremely conservative:
   (a) cPTT outside a range[1] of 50-500 ms
   (b) MAP outside a range[2] of 30-300 mmHg
3. Additionally, time periods were flagged where any of the following criteria applied:
   (a) cPTT was outside the physiological range for more than 60 seconds
   (b) MAP was outside the physiological range for more than 60 seconds
   (c) No measurements of cPTT were acquired for more than 60 seconds
   (d) No measurements of MAP were acquired for more than 60 seconds
4. The flagged measurements from step 2 were removed
5. The time-series were median-filtered using a window size corresponding to a time period of roughly 60 seconds. The window sizes chosen were:
   (a) 300 for cPTT, since the cPTT sampling rate was 5 Hz, or 300 measurements per minute
   (b) 60 for MAP, since the MAP time series contained one measurement per heartbeat, or approximately 60 measurements per minute for a typical HR of 60 bpm
6. The time-series were linearly interpolated onto a common set of timestamps, with one timestamp every 60 seconds, starting at the first point of overlap between the cPTT and MAP time-series and ending at the final point of overlap

[1] The final paired cPTT measurements had a 5th, 50th and 95th percentile of 235 ms, 291 ms and 391 ms respectively, which is is similar to the typical values of cPTT reported in the literature [141, 69]

[2] The final paired MAP measurements had a 5th, 50th and 95th percentile of 51 mmHg, 90 mmHg and 128 mmHg respectively.

7. Sections of the interpolated cPTT and MAP time series lying within the time periods identified in step 3 were removed

This process produced a set of 68,478 paired measurements of cPTT and MAP, each representing a 60-second section of a haemodialysis session. The measurements were aggregated and are shown in figure 3.3. The correlation between cPTT and MAP was found using Pearson's correlation coefficient, which is defined as

$$r(A,B) = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{A_i-\mu_A}{\sigma_A}\right)\left(\frac{B_i-\mu_B}{\sigma_B}\right) \tag{3.1}$$

where μ_A and σ_A are the mean and standard deviation of dataset A, and μ_B and σ_B are the mean and standard deviation of dataset B respectively. The significance of the correlation was found by testing the null hypothesis that the correlation coefficient is zero, using a Student's t-test with a test statistic of

$$t = r\sqrt{\frac{N-2}{1-r^2}} \tag{3.2}$$

where N is the number of paired measurements.

In total there were 68,403 paired measurements across 358 sessions recorded from 37 subjects. The paired cPTT and MAP measurements had a global correlation coefficient of $r = -0.233$, indicating significant ($p < 0.00001$) weak negative correlation between PTT and mean arterial pressure across all subjects. Also shown on the figure are the least squares fits to each of the models in section 2.5.1: the logarithmic model in equation 2.12, the hyperbolic model in equation 2.13 and the linear model in equation 2.14. Each model has a similar global performance, with an RMSE of around 23 mmHg, and each model has a similar fit within the physiological range of MAP. For simplicity, and due to the similar performance of the models, only the linear model will be considered in further investigation.

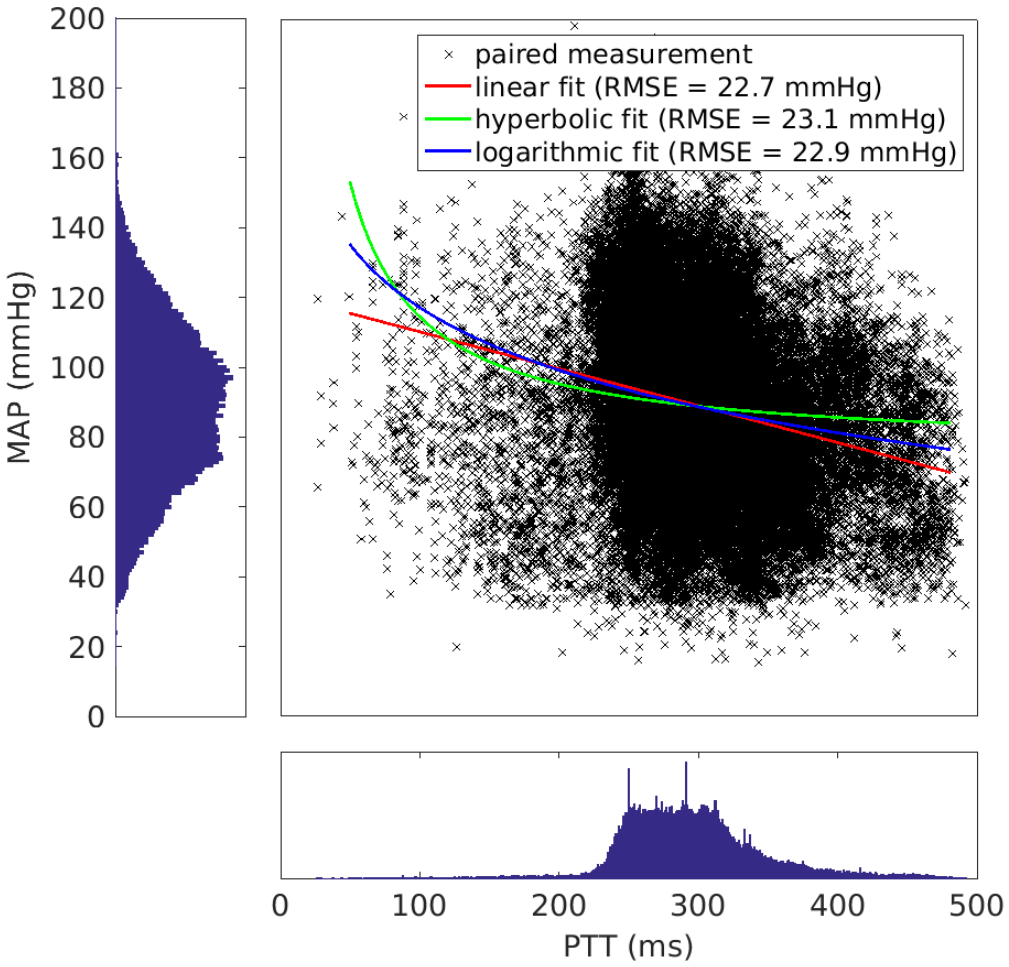


Figure 3.3: Global relationship between paired cPTT and MAP measurements across all haemodialysis sessions, with least squares fits from the three models described in section 2.5.1, and RMSE (in mmHg) shown in the legend. The histograms show the physiological range of MAP and cPTT in the population.

3.2.2 Per-subject correlations

The heterogenerity of the population of patients on haemodialysis means that it is difficult to perform global analysis. Rather, it may be more appropriate either to perform the analysis on a per-subject or per-session basis.

To investigate the relationship between cPTT and MAP on a per-subject level, the correlation coefficient between cPTT and MAP was calculated for each subject by considering only sessions recorded from that patient. The mean per-subject correlation coefficient was -0.124, and the standard deviation was 0.203. After fitting each subject's

data to the linear model (equation 2.14), the mean per-subject RMSE was 15.4 mmHg and the standard deviation was 3.1 mmHg.

Most patients show significant but weak negative correlation between cPTT and MAP, but the correlation depends on the subject. The relationship for each subject is shown in table 3.1, which includes the correlation coefficient between cPTT and MAP, and quantifies the linear fit for that subject using RMSE. The linear fit is also described using the gradient and intercept of the line, which are equivalent to the parameters K_1 and K_2 respectively in the linear model from equation 2.14. Figure 3.4 shows histograms of correlation coefficient, RMSE, gradient and intercept for all subjects.

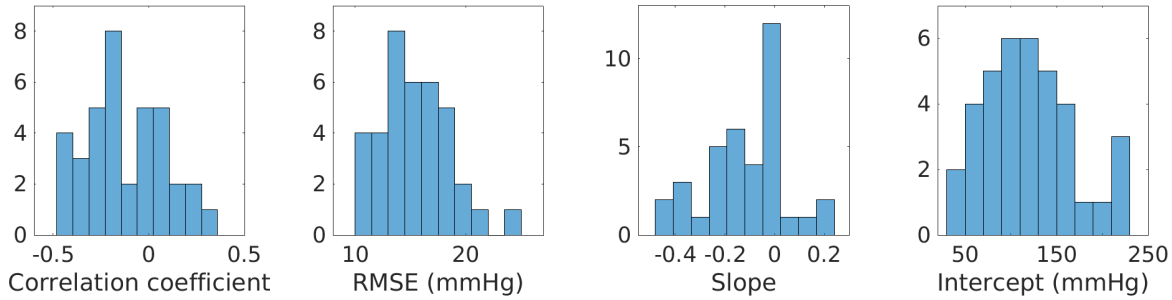


Figure 3.4: Histograms showing the per-subject parameters of the relationships between cPTT and MAP.

#	Subject	Count	r	RMSE	gradient	intercept
1	DF48	2572	-0.458	18.4	-0.303	173.6
2	DF40	260	-0.430	14.4	-0.425	213.7
3	DF05	4986	-0.423	15.4	-0.426	206.7
4	DF17	2831	-0.404	16.7	-0.211	164.3
5	DF11	2172	-0.389	18.1	-0.192	121.0
6	DF01	1221	-0.379	19.9	-0.368	221.5
7	DF55	1016	-0.336	11.3	-0.348	153.4
8	DF43	2560	-0.308	14.4	-0.197	137.7
9	DF37	2615	-0.280	18.5	-0.178	139.9
10	DF34	1976	-0.256	17.0	-0.369	212.1
11	DF21	1741	-0.256	17.1	-0.206	161.0
12	DF30	1696	-0.229	16.1	-0.133	118.4
13	DF54	1845	-0.211	15.2	-0.196	141.9
14	DF02	210	-0.205	10.9	-0.095	110.2
15	DF20	1992	-0.176	13.5	-0.126	106.2
16	DF13	1961	-0.167	11.0	-0.076	113.3
17	DF04	237	-0.166	10.0	-0.086	102.4
18	DF07	1063	-0.165	12.5	-0.153	147.0
19	DF45	2120	-0.163	14.3	-0.149	157.7
20	DF08	1873	-0.148	15.0	-0.131	134.4
21	DF60	1653	-0.084	13.0	-0.045	120.7
22	DF18	2651	-0.060	20.8	-0.054	122.2
23	DF26	2034	-0.052	15.0	-0.021	85.5
24	DF12	2556	-0.047	13.1	-0.036	97.2
25	DF23	2391	-0.043	15.0	-0.019	79.8
26	DF27	1478	-0.029	18.9	-0.009	79.5
27	DF53	1352	-0.011	13.6	-0.006	78.5
28	DF59	1090	+0.024	24.4	+0.008	65.1
29	DF22	1733	+0.027	11.6	+0.013	96.3
30	DF14	649	+0.027	16.9	+0.010	101.8
31	DF41	1682	+0.043	15.3	+0.007	50.3
32	DF33	1338	+0.067	19.4	+0.022	105.7
33	DF24	2604	+0.119	17.9	+0.021	79.3
34	DF44	2307	+0.169	14.2	+0.056	62.0
35	DF49	2183	+0.201	13.9	+0.231	39.6
36	DF35	3253	+0.254	16.1	+0.223	38.2
37	DF19	502	+0.355	12.6	+0.097	59.5
-	**Total**	**68403**	**-**	**-**	**-**	**-**

Table 3.1: Strength of the cPTT-MAP linear fit shown for each subject of the Renal Unit study, ranked by correlation coefficient. Count: number of paired measurements of cPTT and MAP, r: correlation coefficient, RMSE: root mean squared error in mmHg. Subjects without any paired measurements have been excluded.

3.2.3 Per-session correlations

In addition to the global and per-subject correlations, the cPTT-MAP relationship was also investigated on a per-session basis. These per-session relationships are shown in appendix B. Across all subjects and sessions, the mean per-session correlation coefficient was -0.102, and the standard deviation was 0.309. The mean RMSE of the linear model was 12.3 mmHg, and the standard deviation was 4.3 mmHg.

3.3 Longitudinal data

In order to visualise the relationship between cPTT and MAP during a subject's haemodialysis sessions, longitudinal data are presented for three subjects with a relatively strong negative correlation between cPTT and MAP. The data are presented in figure 3.5, which shows a clear relationship between cPTT and MAP during most sessions for these subjects. Subject DF05, for example, exhibits a rise in MAP and a fall in cPTT during almost all sessions.

As table 3.1 shows, some subjects had a slightly positive correlation (or no correlation) between cPTT and MAP. Longitudinal data from three such subjects are shown in figure 3.6. Unlike figure 3.5, which showed clear negative correlation between cPTT and MAP, figure 3.6 instead shows cPTT and MAP with slight positive correlation.

Figure 3.7 shows paired measurements from two sessions: one with negative correlation, and one with positive correlation. This figure demonstrates that the use of correlation to assess the relationship between cPTT and MAP has some flaws. The session with negative correlation displays two clustered groups of measurements, neither of which has any obvious negative correlation. When combined, however, the clusters produce a relatively strong negative correlation. This is an example of Simpson's Paradox, in which a trend in grouped data is not present in the individual groups. In this case, the BP is markedly lower in the second half of the session compared to the first, and the PTT is

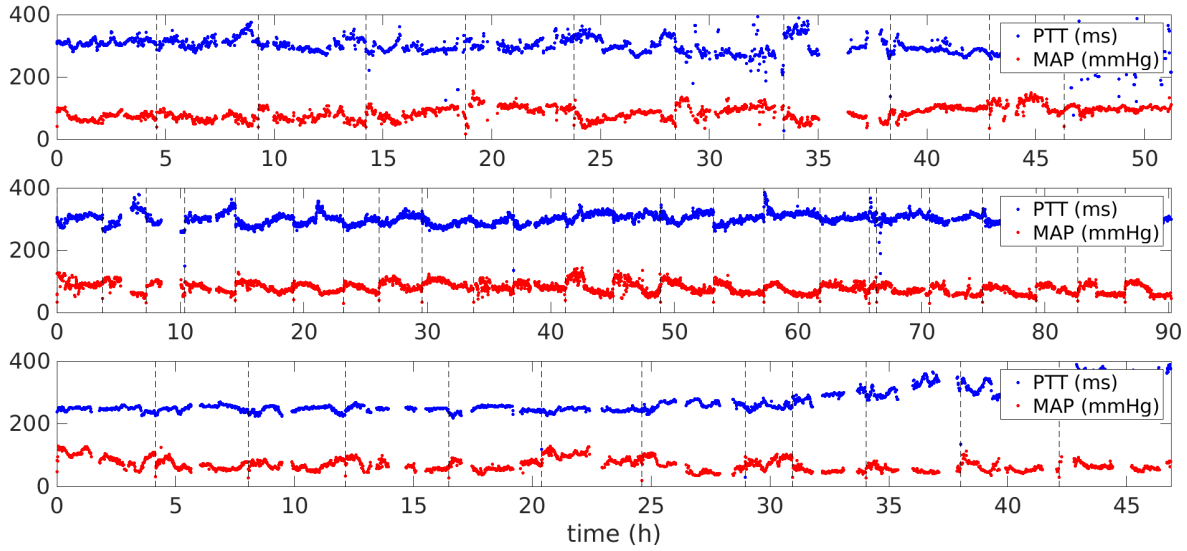


Figure 3.5: Longitudinal data from subjects with a relatively strong negative correlation between paired cPTT and MAP measurements. Data from subjects DF48, DF05 and DF11 (top to bottom). Sessions are displayed back-to-back, separated by black dashed lines.

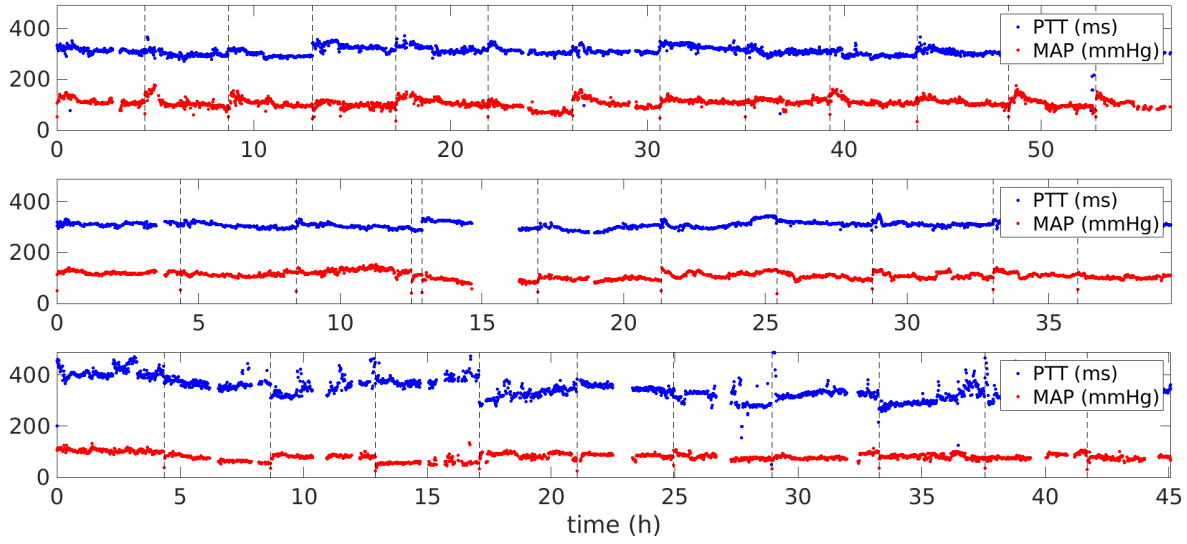


Figure 3.6: Longitudinal data from subjects with slightly positive or no correlation between paired cPTT and MAP measurements. Data from subjects DF35, DF49 and DF44 (top to bottom). Sessions are displayed back-to-back, separated by black dashed lines.

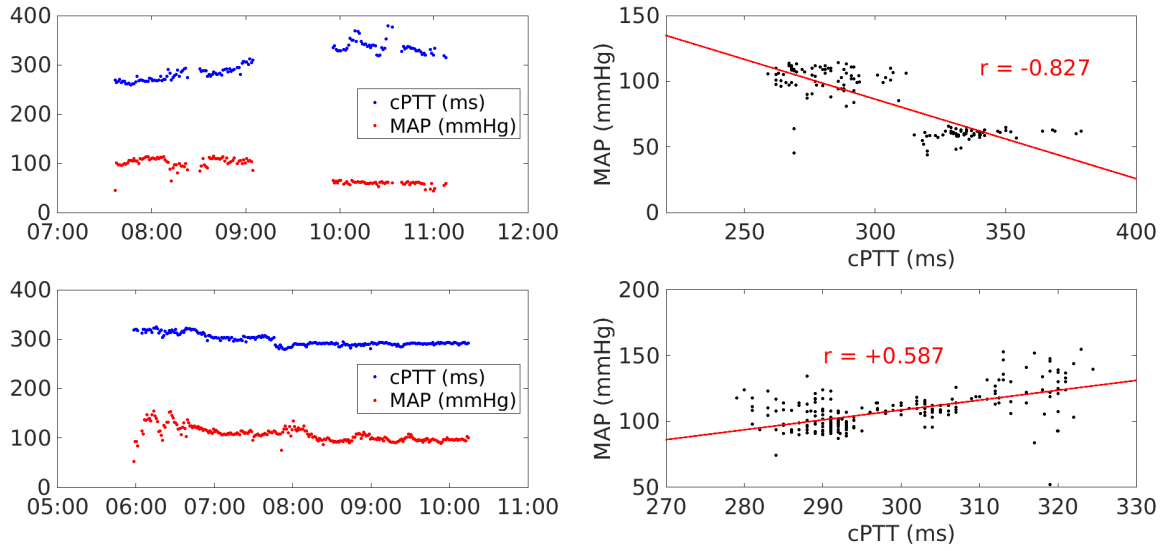


Figure 3.7: Paired measurements from two haemodialysis sessions: one from subject DF05 (top) showing negative correlation, and one from subject DF35 (bottom) showing positive correlation.

correspondingly higher; within these two parts of the session, however, the correlation is not apparent.

In the case of the second session, figure 3.7 shows a relatively high positive correlation coefficient. The longitudinal data, however, exhibits some opposing movement in cPTT and MAP, notably at approximately 08:00 where there is a sudden increase in MAP and a decrease in PTT. This seems to have very little effect on the overall correlation coefficient.

3.4 Discussion

Overall, the correlation between cPTT and MAP was weaker than has been reported elsewhere [92, 105] (see also section 2.5.2). Most other studies, however, have analysed data from healthy volunteers, who comprise a more predictable, homogeneous population. The subjects in this study, by contrast, are much more heterogeneous, and so the correlation is unlikely to be as strong. Factors such as CVD disease can themselves

affect the PTT-MAP relationship (see section 2.5.3), and other factors such as HR and fluid levels change dynamically throughout haemodialysis, potentially confounding the relationship. Lower correlations between PTT and MAP could also be explained by patients with compromised cardiovascular function; for example, a decrease in the ability to autoregulate blood flow in the organs via changes in vessel tone (vasoconstriction and vasodilation). In such patients, changes in MAP might not manifest as a change in PTT because of the unpredictable changes in blood vessel volume.

The heterogeneity of the patient population also explains the huge variation between subjects, which is much larger than is often reported in the literature. Another possible cause of the weaker correlations is the length of time for which subjects were monitored; other studies tended to monitor for minutes, rather than hours, resulting in fewer measurements. Generally, it is easier to observe higher correlations in a small number of measurements, and so the large amount of data analysed in this chapter puts this investigation at a potential disadvantage as far as assessing correlation is concerned.

In a similar analysis of the relationship between cPTT and MAP, and using the same data set as that of this chapter, MacEwen [89] showed that HR had a weak effect on the strength of the correlation, and found that the presence of peripheral vascular disease was significantly predictive of a less negative correlation. The study also noted that positive correlation could be the result of a session with limited change in both cPTT and MAP. These factors could explain the subjects and sessions in which no negative correlation is observed.

The continuous BP measurements were acquired using a Finometer device, which uses technique of vascular unloading to estimate brachial BP (see section 2.3.3). Other studies have used continuous invasive BP measurement using an arterial line (section 2.1.1); in this clinical study this method of measurement was not available. The use of the Finometer, which estimates BP, introduces a level of error to the models since PTT is not being compared to directly-measured BP. In addition, MAP was used as a

comparator in this chapter, whereas many others have used SBP or DBP; this is because Finometer MAP was previously shown by MacEwen (2016) [89] to have a smaller error when compared to MAP than the corresponding error for SBP.

Simpson's Paradox, which is discussed in section 3.3 could explain why the global correlation (-0.233) is slightly stronger than the mean per-subject correlation (-0.124), and in turn why the mean per-subject correlation is stronger than the mean per-session correlation (-0.102). The global correlation may be more prominent simply because there is a larger spread of measurements. This is supported by the increased standard deviation of correlation coefficient in the per-session data (0.309) than in the per-subject data (0.203), which shows that the variation in correlation between sessions can be averaged out by aggregating the sessions.

For the linear model, the error decreases from a global RMSE of 22.7 mmHg to a mean per-subject RMSE of 15.4 mmHg and a mean per-session RMSE of 12.3 mmHg. This suggests that the model should be personalised to each patient for greater accuracy, and ideally to each session. In practice, fitting a model for each session would not be feasible as the process requires a large number of measurements which are not available until late in the session.

It is clear from the results in this chapter that the relationship between PTT and MAP varies greatly between subjects, and even between haemodialysis sessions for the same subject. The relationship is so weak in some individuals that it is not of practical clinical use, but is strong enough in others that PTT has the potential to be used as a minimally-invasive continuously acquired surrogate for MAP. Identifying the subjects in which the PTT-MAP relationship has a strong correlation is therefore of great importance; further investigation of patient sub-populations could help to identify groups of patients for which this type of monitoring may be clinically useful. Regardless, in any real clinical scenario, calibration would be required frequently if cPTT were to be used as a surrogate for MAP [105].

Chapter 4

Non-contact monitoring

Within the last decade, the field of non-contact vital sign monitoring has advanced rapidly. Wieringa *et al.* in 2005 [169], Takano and Ohta in 2007 [152], and Verkruysse *et al.* in 2008 [161] were among the first to demonstrate that visible light video cameras can be used to capture a remote reflectance photoplethysmography (rPPG) waveform. The principle of rPPG is a simple one: as the heart beats, the apparent colour of the skin changes due to the blood's pulsation. There are various factors contributing to this colour change, which are described in section 4.1.1. Although this colour pulsation is too small to be seen with the naked eye, it can be detected and measured using a video camera positioned up to several metres away.

Non-contact monitoring allows physiological signals to be recorded without the need for sensors to be in physical contact with the patient. Patients in hospital often have to wear several sensors for vital sign monitoring, and these sensors quickly become cumbersome and irritating, particularly if they have to be worn for a long period of time.

There is also risk associated with attaching sensors to patients. Probes such as ECG leads [83] and finger probes for pulse oximeters [170] can harbour bacteria and are therefore an infection risk; this is particularly true for reusable equipment. Many studies describe burn injuries associated with MRI which are caused by electromagnetic

induction in sensor cables [62, 34, 107]. In addition, some patients have skin that is sensitive and likely to be damaged by such sensors (such as neonates with delicate skin, or patients with extensive burn injuries) [16]. Non-contact monitoring completely eliminates these risks, and also reduces expenditure on consumables such as disposable sensors.

4.1 Remote photoplethysmography

Although it cannot be seen with the naked eye, the skin constantly fluctuates in colour as the heart beats. This colour change was demonstrated by Wu *et al.* in 2012, who developed a framework for the magnification of small motions and colour changes in video[1]. One of the examples the authors showed was the colour change seen in a human subject's face over time (figure 4.1).

Section 2.3 introduced the concept of photoplethysmography (PPG), which uses light to measure the change in volume of blood vessels. The measurement of PPG has been part of standard clinical care since the invention of pulse oximetry in the 1970s [3], and the technology has changed little since. The measurement of remote plethysmography (rPPG) using a video camera involves similar principles to pulse oximetry, but table 4.1 shows several key differences.

This section discusses the physiological origin of the rPPG signal (section 4.1.1), the

[1]The software is available at `http://people.csail.mit.edu/mrub/vidmag/`, where it can be downloaded for non-commercial research purposes.

Figure 4.1: Example of the Eulerian Video Magnification framework by Wu *et al.*, showing four frames from a subject's face with colour changes amplified. Figure modified from [173].

	Conventional PPG	Remote PPG
Light source	Monochromatic LEDs	Any, often ambient light
Sensor	Photodiodes	Camera image sensor
Distance	Requires physical contact	Up to several metres
Site	Typically finger or ear	Any visible skin region

Table 4.1: Key differences between conventional PPG measurement using a pulse oximeter, and measurement of rPPG using a video camera.

technology involved in the recording process (section 4.1.3), and the method of deriving the rPPG (section 4.1.2). In addition, section 4.1.4 includes a review of the various names used for the signal throughout the literature.

4.1.1 Origin

Colour is subjective; it is not an inherent property of light, but rather is perceived as a result of the biological signals sent from the eye to the brain. The eye is sensitive to light in the visible range of the electromagnetic spectrum, but also has the ability to differentiate between wavelengths within this range; the brain interprets this as colour.

Vision and colour

The human retina contains two types of photoreceptor cells: rod cells, which are responsible mainly for vision in dim light and peripheral vision, and cone cells, which work better in brighter conditions and are responsible for the perception of colour. Humans have three[2] types of cone cell, each of which is sensitive to a particular range of wavelengths (see figure 4.2). The perception of colour arises from the difference in the signals received from the different types of cone cell, which the brain is able to interpret. For example, light stimulating the long-wavelength cone cell much more than the medium-

[2]Colour-blindless is a fault in one of the types of cone cell, reducing the ability to differentiate between colours. Conversely, it has been reported that some humans have four types of cone cell, resulting in an increased ability to differentiate between colours. In comparison, most mammals only have two types of cone cell, and certain animals have many more (the mantis shrimp has 16 cone cells, giving it the ability to see into the ultraviolet range).

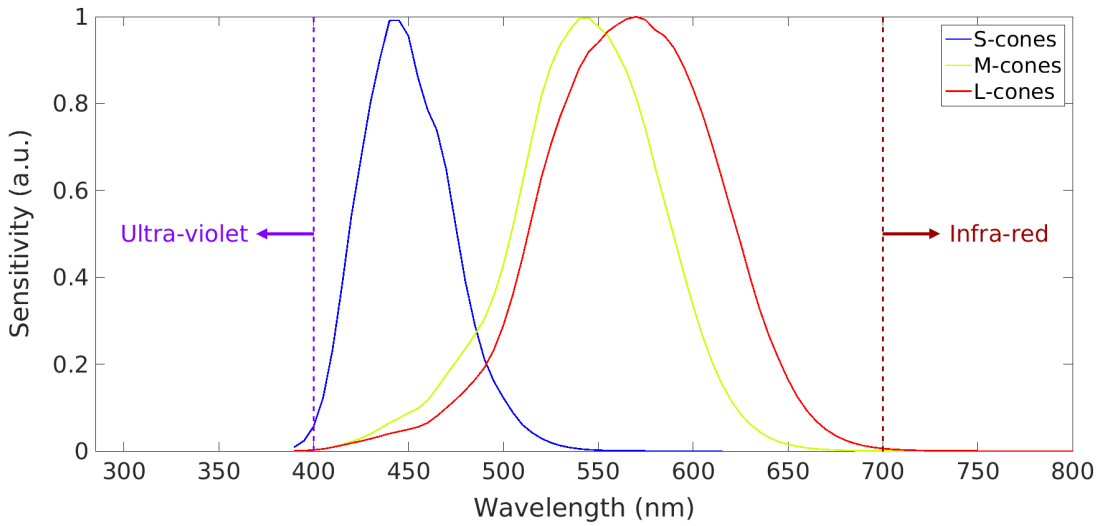


Figure 4.2: Normalised sensitivity of human cone cells to light of different wavelengths. Raw data from Stockman and Sharpe (2000) [145], used with permission.

or short-wavelength cells would be interpreted by the brain as red[3].

The colour of an object is determined by which wavelengths of light it variously absorbs, reflects and emits. When white light is shone on skin, for example, the shorter (bluer) wavelengths are absorbed more than the longer (redder) wavelengths, which are reflected causing the eye to receive light of longer wavelength. This makes the skin appear more red than blue[4].

Light-skin interaction

The origin of the colour change seen in figure 4.1 is not straightforward as it is complicated by the interaction between skin and light. For example, visible light only penetrates 2-3 mm into the skin [13]; beyond this distance the light is attenuated too much to be reflected back to the eye. For this reason, the colour change can only be caused by blood

[3]White light, which is broad-spectrum and contains many different wavelengths, can be approximated by combining frequencies that stimulate each type of cone cell equally. The primary colours are commonly red, green and blue as these colours maximise the difference between the responses of the various cone cells.

[4]It is a common misconception that veins are blue. In fact, they merely appear blue when seen through several millimetres of skin; subcutaneous fat allows mainly blue light to penetrate it, and therefore only blue light can be reflected back.

vessels close to the skin's surface; this means that major arteries and veins are unlikely be the source of the colour change as they are too deep under the skin's surface. The colour change must therefore be caused by minor blood vessels such as arterioles, capillaries and venules. Since the colour change is pulsatile, and vessels of the venous system experience little or no pulsatile blood flow, it is likely that arterioles are the main source of the colour change [90].

It has also been shown that, in the case of RGB cameras, the green channel contains the pulsatile signal with the greatest amplitude [161, 124, 125]. It is hypothesised that this is because the green channel contains wavelengths which are more likely to be absorbed by blood (which would reflect longer red wavelengths) but which penetrate further into the skin than shorter blue wavelengths.

Components of the rPPG signal

The apparent colour changes in the skin that give rise to the rPPG signal are complex as they are caused by multiple factors. These factors are a subject of debate in the literature, but the following components are generally believed to contribute to the rPPG signal.

1. **Blood vessel volume.** The beating of the heart causes a pulsatile blood pressure, which cycles between systolic and diastolic pressure (see section 2.1.1). Blood vessels stretch as the pressure surges within them, increasing the volume of blood and changing the intensity of the light reflected from the vessels. This has the effect of changing the perceived colour of the skin. In addition, longer-term changes in vessel volume are a result of vasoconstriction and vasodilation, the contraction and relaxation of the muscular blood vessel walls. This process allows the vessels to control the pressure of the blood passing through them, which is necessary to maintain a constant flow rate to organs that require a steady supply of oxygen. Vasoconstriction also prevents excessive heat loss in cold environments; similarly,

vasodilation allows heat to be radiated out of the body. These changes in blood vessel diameter can cause obvious skin colour changes, such as pale skin in cold weather or flushed skin in a patient with a fever.

2. **Red blood cell alignment.** Surprisingly, a pulsatile colour change is observed when blood is pumped through a rigid vessel (such as a glass tube). As the pressure wave propagates through the blood, red blood cells change their alignment. This alignment affects the light-scattering properties of the blood [109] and causes a periodic change in the intensity of reflected light, which manifests as a periodic change in apparent skin colour.

3. **Oxygen saturation.** SpO_2 (see section 2.3.2) has a distinct effect on skin colour; low levels of oxygen in peripheral tissue (such as the fingers or lips) causes a condition called cyanosis, in which the skin appears blue or purple. This colour effect is due to the differing absorption spectra of deoxygenated haemoglobin and reduced haemoglobin (as shown in figure 2.5). SpO_2 has an effect on both the DC and AC components of the rPPG signal [43].

4. **Cardiac-synchronous movement.** The beating of the heart causes surges of pressure in the arteries, and the resulting forces on the blood vessels cause subtle changes in the position of the head [8]. This periodic movement contributes to the rPPG waveform since the area of skin being observed by the video camera fluctuates with the movement.

These components should all be considered when examining the information contained within an rPPG signal (this information will be investigated in section 4.3).

4.1.2 Derivation

The rPPG signal is derived from the raw video in order to condense the information contained within the video frames to a one-dimensional time-series. Each colour channel

of the video can be thought of as an $M \times N \times T$ matrix, where M is the number of pixels in the vertical direction, N is the number of pixels in the horizontal direction, and T is the number of frames. Each value in the matrix represents the light intensity recorded by a particular pixel at a particular time. This 3-dimensional matrix is reduced to the rPPG signal in the form of a 1-dimensional $T \times 1$ matrix.

The process consists of three steps:

1. Selection of a region of interest (ROI). This ROI should cover the area of skin over which the rPPG signal is to be measured.

2. Spatial averaging of the pixel intensities within the ROI. This produces one value for each video frame.

3. Combination of each frame's averaged value to produce an rPPG time-series.

The selection of an appropriate ROI is discussed in section 4.4.3. The spatial averaging is typically done using the mean, which is representative of the distribution of pixel intensities within the ROI as long as the distribution of pixel values within the ROI is approximately Gaussian. Figure 4.3 shows that this is indeed the case for the examples shown.

The steps above are represented graphically in figure 4.4.

In mathematical terms, the process of deriving the rPPG signal is

$$rPPG_n = \frac{1}{|R|} \sum_{i \in R} \gamma_n^i \tag{4.1}$$

where $rPPG_n$ is the value of rPPG at frame n, R is the set of pixels in the ROI (and $|R|$ is the size of that set), and γ_n^i is the intensity of pixel i in frame n.

The resulting rPPG time-series is usually processed further; time-series from different colour channels can be combined, for example, or the rPPG signal can be filtered to isolate the cardiac-synchronous component.

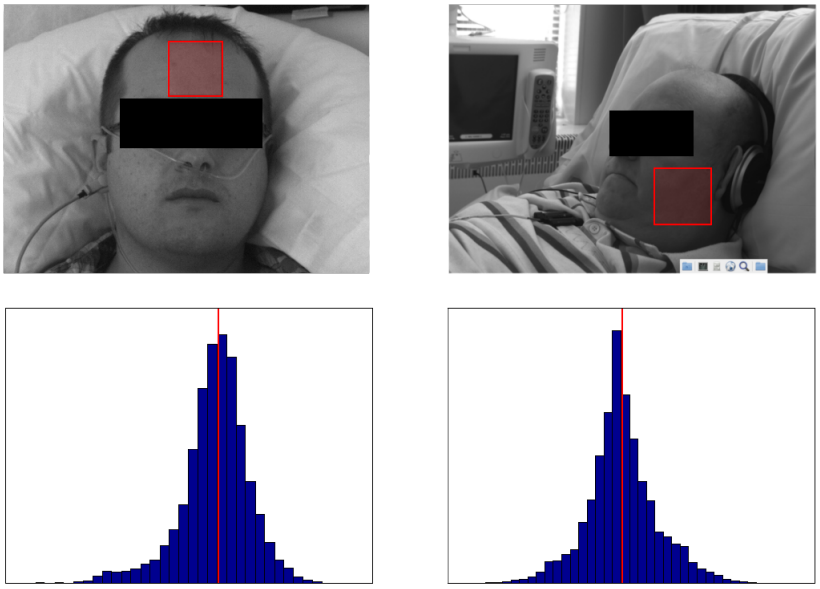
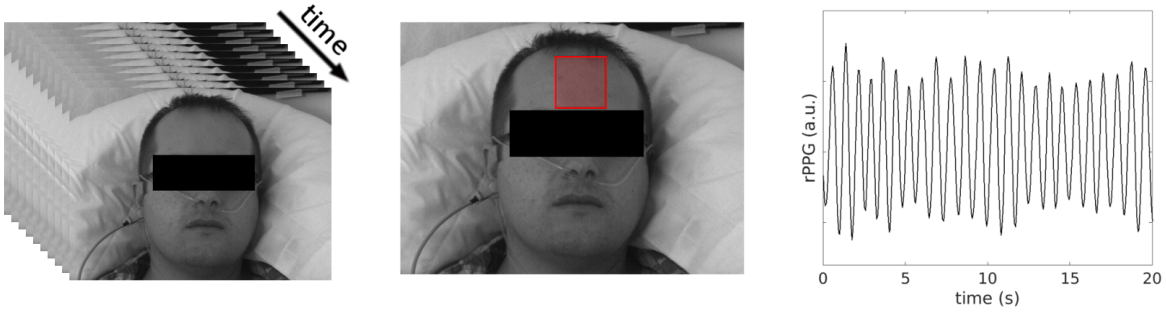


Figure 4.3: Examples of ROI selection (top) and pixel intensity distribution (bottom) for subject HP3-05 of the Physiology study and subject DF27 of the Renal Unit study. The distribution in each case is approximately Gaussian, with the mean (vertical red line) therefore being an appropriate measure of average pixel intensity.



(a) Frames from source video **(b)** Selected ROI (red box) **(c)** Raw rPPG signal

Figure 4.4: Derivation of rPPG signal from an ROI. Data from subject HP3-05 of the Physiology study.

4.1.3 Video recordings

Studies have shown that rPPG signals can be derived using video cameras ranging from high-end custom-built sensors [169] to commercially available digital cameras [161] and webcams [124, 125]. However, the quality of the video recordings determines the quality of the derived rPPG signals. The recording process must therefore be thought through very carefully in order to extract as much information as possible from the rPPG.

Camera sensors

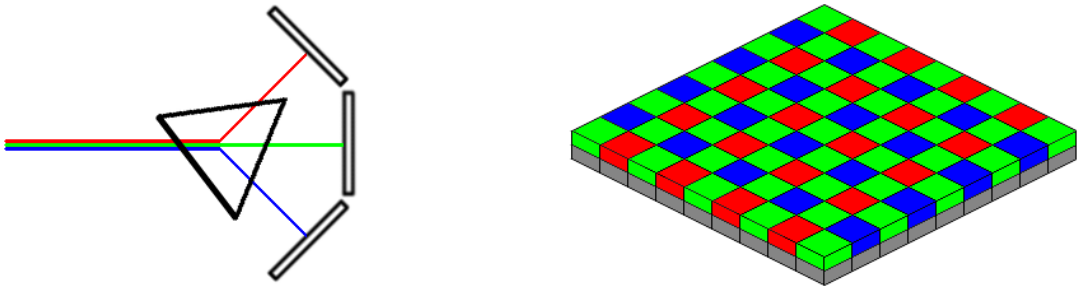
Video cameras use many different types of image sensors. The sensor is responsible for converting light to an electrical current, and this analogue signal is then converted to a digital output and sampled. Each pixel on the image sensor produces its own value of light intensity, and these are combined to form an image.

The two main types of image sensors are charge-coupled device (CCD) sensors and complementary metal-oxide-semiconductor (CMOS) sensors. CCD sensors have one amplifier in the corner of the sensor which reads and amplifies the charge generated by each pixel in the sensor. These charges are read sequentially, pixel-by-pixel. CMOS sensors, conversely, have individual amplifiers for each pixel. This allows the charge from every pixel to be read and processed in parallel.

Cameras with CCD sensors generally produce higher-quality images, but are more expensive and consume much more power. By comparison, CMOS cameras are cheaper and consume less power, but the electronics required by each pixel result in a lower sensitivity to light. CCD technology is older and well established, but CMOS devices are now considered to have comparable performance for both commercial [18] and specialised [73] purposes.

Colour image sensors are more complicated than monochrome sensors as they have to distinguish between light of different wavelengths. 3CCD image sensors (figure 4.5a) use a prism to separate incoming light into ranges of wavelengths corresponding to red,

green and blue light. Each of these ranges is directed onto its own CCD sensor, and the output from the three CCD sensors are combined into a colour image. 3CCD sensors produce high-quality images, but are more expensive since three different CCD sensors are required. Cheaper options include single CCD sensors with a Bayer filter, which is a mosaic pattern of optical filters that allows only red, green or blue light to pass through to each pixel (figure 4.5b). The values from neighbouring single-colour pixels are then interpolated to obtain RGB values for individual pixels. This interpolation results in a lower-quality image.



(a) 3CCD sensor **(b)** Bayer filter sensor

Figure 4.5: Two types of colour image sensor. The 3CCD sensor combines information from three CCD sensors to create an image. The Bayer filter interpolates between adjacent pixels to create an image.

Although colour images are most commonly represented using red, green and blue (RGB) values, other channels can be used. In RGB colour space, each channel represents an orthogonal dimension. Alternative colour spaces include CMYK—cyan, magenta, yellow and key (black)—which is commonly employed in the printing industry; HSV, which encodes hue, saturation and lightness[5] and is often used in computer graphics or image editing; and YCbCr, which encodes luminance and chrominance[6] and is used in video and digital photography. Colour transformations can be used to map between

[5]The HSV system uses cylindrical coordinates to locate a colour in the RGB colour model, in an attempt to be more intuitive than the standard Cartesian coordinates of the RGB system.

[6]Luminance encodes the brightness of an image, and chrominance encodes the variation in colour from the mean.

colour spaces.

The image sensors and colour spaces used in our studies are described in section 1.4.

Video camera parameters

Video camera parameters such as frame rate, resolution and bit depth will affect the quality of the rPPG signals derived from a video. Each parameter must therefore be chosen with a compromise between video quality (which will affect the rPPG quality) and the resources required to record and store the video data.

Frame rate, for example, will determine the sampling frequency of the rPPG as each frame produces one sample of rPPG. A high frame rate is therefore desirable, but an excessive frame rate will produce very large video files with limited benefits.

The resolution of the image depends on the number of pixels in the image sensor and can affect the amount of noise in the rPPG signal; for a given area of skin, a very low resolution video will produce a rPPG signal using values from only a few pixels, which are each prone to electrical sensor noise. In higher resolution images, this noise tends to be cancelled out through the rPPG derivation process (specifically the averaging of pixel values); this averaging is less effective at removing sensor noise for a small number of pixels. Again, very high resolution video cameras produce large video files and so a compromise must be made between quality and size.

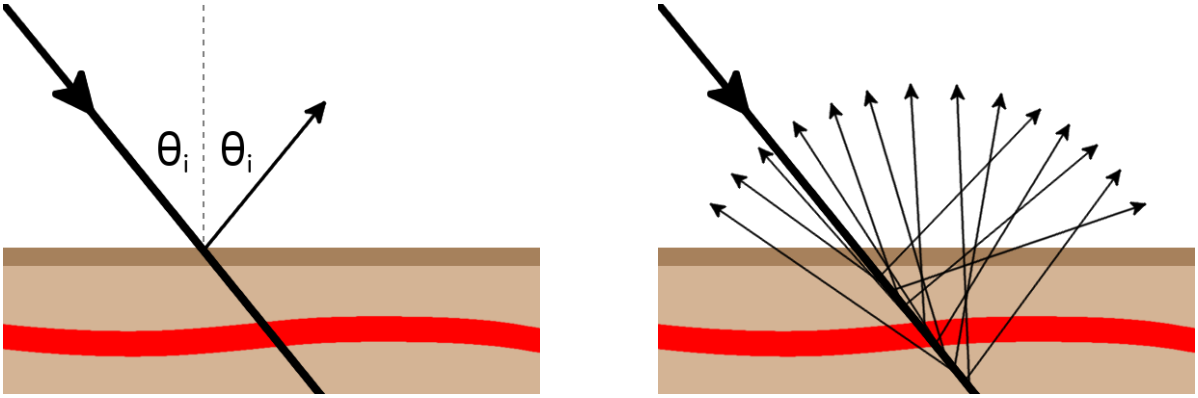
The bit depth of the image is determined by the image sensor's analogue-to-digital converter (ADC). An 8-bit ADC will produce digital intensity values of between 0 and 255. A 12-bit ADC will produce readings of between 0 and 4096, giving 16 times the colour resolution, but requires more expensive image sensors and produces larger video files.

The camera parameters in our studies were chosen so as to allow the highest possible quality rPPG waveform to be produced (within the constraints of the available recording equipment). These parameters are discussed in section 1.4.

Subject lighting

The lighting of the subject is another important consideration. This lighting should be chosen to minimise the amount of specular reflection in the image. This is because specular reflection consists mainly of light that has not penetrated the skin, and therefore contains minimal physiological information. Diffuse reflection, by contrast, is more likely to be the result of interaction with the skin and therefore contains more physiological information. Figure 4.6 shows the difference between specular and diffuse reflection.

Direct lighting of the subject should be avoided; a concentrated source of light will bounce off the subject's skin via specular reflection. More diffuse lighting, in which the source of light is less concentrated (such as ambient daylight), is preferable to maximise levels of diffuse reflection. Specular reflection can present as an overexposed (highlighted) area on the subject's skin.



(a) Specular reflection **(b)** Diffuse reflection

Figure 4.6: Comparison of specular and diffuse reflection in the context of skin. Specular reflection is less likely than diffuse reflection to have penetrated the skin, and is therefore less likely to contain useful physiological information.

Lighting powered by mains electricity is also likely to produce low-quality rPPG signals. Such types of lighting often flicker at the mains frequency (50 or 60 Hz), which can produce aliased constant-frequency rPPG noise. Preferable sources of lighting include LED lamps, which do not flicker, and natural daylight.

The studies in this thesis used different types of lighting. The Renal Unit study used

ambient lighting due to restrictions on the environment around the patient, whereas the Physiology study used carefully controlled LED illumination (see section 1.4).

4.1.4 Signal name

The rPPG signal described in this section has a variety of names in the literature, including:

- imaging PPG (iPPG/IPPG) [149]
- PPG imaging (PPGI) [174]
- remote PPG (rPPG) [32]
- DistancePPG [77]
- camera based PPG (cbPPG) [158]
- non-contact PPG (NCPPG) [53]

Because a consensus on a name has not been reached in the field, it is important to take note of the various names so that researchers can keep track of relevant publications (regardless of the terminology used). Throughout this thesis, the signal is referred to as rPPG.

4.2 Vital sign monitoring

Just as conventional PPG can be used to estimate vital signs (see section 2.3), it is possible to estimate vital signs from rPPG. In fact, rPPG signals are potentially more useful than conventional PPG because multiple rPPG signals can be derived using different ROIs, corresponding to different sites on the skin. This means that each signal has spatial information associated with it, which allows the possibility of spatial and

temporal mapping of cardiovascular and haemodynamic information (explored in Chapter 6).

Originally, rPPG was used to derive HR and BR, but non-contact monitoring using video cameras has been proposed for the measurement of parameters such as SpO_2 [44, 160, 1], PTT [106, 139, 61] and even pain or stress [70, 14]. In this section, some of the non-contact monitoring literature is presented to provide context for the work described in this thesis.

4.2.1 Heart rate and breathing rate

HR and BR were the first physiological parameters to be measured using video cameras. In 2005, Wieringa *et al.* [169] used near-infra-red illumination and a monochrome CMOS camera to isolate a reflectance rPPG signal in the skin of 7 subjects, finding a weak pulsatile variation due to HR superimposed on a clear BR frequency. When the experiment was repeated for one subject at an increased frame rate, the authors found peaks in the frequency spectrum that matched the BR frequency, the HR frequency and its second harmonic. They highlighted problems with low signal-to-noise ratio (SNR) and motion artefacts. In 2007, Takano and Ohta [152] used spectral analysis of video recordings from 14 healthy subjects captured using a CCD camera to identify peaks corresponding to HR and BR.

Other authors used consumer-grade video cameras in the same way, with one study finding that high-performance cameras and webcams performed similarly in HR measurement [148]. Verkruysse *et al.* [161] used a consumer digital camera in 2008 to find HR and BR from subjects' faces, using only ambient light as illumination. The authors also produced maps of the amplitude and phase of the rPPG in the face. In 2010, Poh *et al.* [124] derived HR from webcam videos of the faces of 12 subjects, using blind source separation (BSS) to combine information from the webcam's colour channels. The aim of BSS is to isolate a set of source signals (in this case, a cardiac signal and noise

sources) using a set of mixed signals (the three colour channels), and is performed using techniques such as principal component analysis or independent component analysis. The authors included automatic face detection as a method of selecting an ROI in the video. They compared the estimated HR to a reference, finding a mean error of -0.05 ± 2.29 bpm in still subjects and an error of 0.64 ± 4.59 bpm in subjects allowed to move slightly. In a subsequent paper [125], the same authors went on to compute HR as well as further parameters such as BR and HR variability for 12 subjects, finding a mean error of 0.95 ± 0.83 bpm for HR, and 0.12 ± 1.33 bpm for BR. It should be noted that these videos were recorded in very controlled circumstances, with subjects at rest, and for only one minute.

More recently, further refinements have been made to improve the estimation algorithms in less controlled circumstances. A study by de Haan and Jeanne (2013) [31] used chrominance-based rPPG instead of BSS to attempt to account for motion artefact. By creating two orthogonal chrominance signals from three colour channels, the authors aimed to eliminate specular reflection. They showed that the use of chrominance instead of BSS improved the accuracy of HR estimation, particularly for moving subjects. In 2014, our research group published a study [153] which used auto-regressive modelling to estimate vital signs including HR and BR, finding that the error between estimated HR and the reference values from a pulse oximeter was comparable to the error between simultaneous pulse oximeter measurements at the ear and finger.

A paper reviewing the field of non-contact HR measurement was published in 2014 by Kranjec *et al.* [76]. The authors compared new methods of non-contact measurement, and highlighted RGB imaging as one of the most promising; they determined the main advantages to be the low cost and the potential to monitor multiple subjects simultaneously. The complexity of the processing methods, as well as the low temporal resolution, were identified as disadvantages.

Another review by Sun and Thakor [149], published in 2015, focused more generally

Camera type	High-end specialised camera Consumer-level digital camera Phone camera
ROI selection	Manual Automatic using face detection Automatic using feature selection + tracking
rPPG derivation	Spatial averaging Concatenation of independent high-SNR subsignals
Channel selection	Single channel (usually green) Linear combination of channels using BSS techniques
Noise reduction	Moving-average filter Band-pass filter Wavelet processing
Vital sign estimation	Time-domain (peak detection, autoregressive modelling) Frequency domain (Fourier transform, wavelet transform) Learning-based (support vector regression, hidden Markov model)

Table 4.2: Different options for deriving and processing rPPG signals, as identified by Sun and Thakor [149].

on the field of measuring PPG signals using video cameras. The authors presented the different options for each step of processing (table 4.2), and identified some recent advances in the field. These included new methods of reducing motion artefact, the use of signal processing techniques other than Fourier-based methods (which assume stationarity of the rPPG signal), and the expansion of rPPG analysis to estimate multiple physiological parameters rather than just HR and BR.

4.2.2 Oxygen saturation

Although early research focused mainly on HR and BR, the field of non-contact monitoring is expanding to include measurement of other physiological information. Scully *et al.* [138] showed in 2012 that trends in SpO_2 (see section 2.3.2) could be tracked by comparing the ratio of colour intensities in video recorded using a mobile phone. Although their methods required physical contact between the video camera and the subject's fingertip, the principle for measuring SpO_2 from rPPG ought to be the same, although the contact

method does not suffer from problems caused by shadows and lighting changes.

The first studies to investigate the measurement of SpO_2 using rPPG were conducted by Humphreys *et al.* [56, 57] and by Wieringa *et al.* [168] between 2005 and 2007. The authors used CMOS cameras with optical filters at specific wavelengths to estimate SpO_2, but found that their methods were vulnerable to changes in ambient lighting. Little progress was made in the field until 2013, when Kong *et al.* [74] used ambient light coupled with narrow-band filters to estimate SpO_2, using the whole face as a ROI.

Prior to the work described in this thesis, our research group at the University of Oxford also made a contribution to the field of remote SpO_2 estimation. Tarassenko *et al.* (2014) [153] showed that the ratio of rPPG amplitudes in the red and blue channel (from an RGB camera) was correlated with SpO_2 in adult subjects. No optical filters were used, and the only source of illumination was ambient light. In 2015, Villarroel *et al.* [162] showed that the same method could be used in a clinical setting to estimate SpO_2 in neonates.

Further correlation between camera-derived estimates and reference SpO_2 have been demonstrated by Bal (2015) [7] and by Liu *et al.* (2015) [86]. In a paper published by our research group, Guazzi *et al.* (2015) [44] showed, using data from the Physiology study described in section 1.4.2, that SpO_2 could be estimated in a robust way by considering a large number of ROIs across the face, weighting each ROI according to the SNR of the rPPG signal within it.

In 2016, Verkruysse *et al.* [160] demonstrated that the error between camera-derived estimates of SpO_2 and the reference values were small enough to make camera-based pulse oximetry feasible. Again, subject motion was found to be the biggest challenge, and would have to be addressed further for camera-based SpO_2 estimation to be considered a clinical possibility.

Most recently, novel methods have been proposed by Addison (2016) [1], who estimated SpO_2 using wavelet analysis of the video signal, and by van Gastel *et al.* (2016) [159], who

made use of the rPPG signal's characteristic pulse vector. This pulse vector—originally described by the same group in a previous study [32]—is a representation of the rPPG pulsation, describing the direction and magnitude of the skin colour variation within RGB colour space. The authors compared the pulse vector from different video recordings, showing that the pulse vector changes with SpO_2, and that (for a given light source and camera) each value of SpO_2 has an expected pulse vector.

These novel methods could have advantages over conventional ratio-based methods in terms of robustness to subject motion.

4.2.3 PTT and spatial mapping

In the same way that cPTT is derived from the ECG and finger PPG signals (see section 2.4), the PTT between any two locations in the vasculature can be calculated from the PPG recorded at those sites (see figure 2.7). It follows that by deriving rPPG from two separate regions of video of a subject, the PTT between those sites could be found. In this way, by detecting changes in the remotely-measured PTT, it may be possible to detect changes in haemodynamic parameters using only a video camera.

A disadvantage of this approach is that the conventional cPTT—measured from heart to finger—cannot be measured without electrodes attached to the skin. Also, the heart is located much too deep inside the chest for a representative rPPG signal to be derived; rPPG measured at the chest would instead be measuring changes in superficial skin perfusion in the area. PTT derived using video would therefore have to be measured between two sites at which skin is exposed to the camera, such as the neck, face and limbs. The relationship between BP and PTT measured between such sites has not been investigated in the literature to the same extent as the relationship between BP and cPTT.

Other publications have reported methods that include spatial mapping of certain parameters within a video. The original work of Verkruysse *et al.* [161] included maps

of the power contained within localised rPPG signals, as well as maps showing phase differences in the rPPG signals derived from different parts of the face. Although the authors did not use this terminology, this phase difference is equivalent to local PTT between two nearby regions.

A more detailed review of the literature is included in later chapters. Section 5.1 presents the recent papers that have estimated PTT using video cameras, and section 6.1 discusses publications that have attempted to use cameras to perform spatial mapping of various haemodynamic parameters.

4.3 rPPG frequency content

Despite the growth of the field of non-contact monitoring, most studies have used the rPPG signal only to derive HR or BR; relatively few have thoroughly investigated the origin of the rPPG signal to see what additional information the waveform contains.

In their review of the rPPG literature, Sun and Thakor (2015) [149] suggested that the opto-physiological model from which the rPPG signal arises is poorly understood. They cite the complexity of the light-skin interaction, combined with the physiology of blood pulsation in elastic blood vessels, as reasons for this lack of understanding. The authors suggest that the model could be refined by interpreting new experimental results, which might lead to a better understanding of the origin of the rPPG signal's various components.

As an example of the uncertainty over the existing rPPG model, Kamshilin *et al.* (2015) [64] reported some experimental observations suggesting that the light intensity changes captured by the rPPG signal are caused mainly by elastic deformation of the skin, rather than by pulsation of arterial blood. They argued that the pulsation of arterial pressure leads to periodic deformation of the capillary bed, which modulates the light-scattering properties of the skin and leads to a change in reflected light intensity.

The authors pointed to the presence of different phases in the rPPG signal as evidence for this claim.

In order to understand the complexities of camera-derived PTT (which is the focus of Chapter 5), this section will investigate the frequency components of the rPPG signal, comparing them to the components of conventional PPG signals.

4.3.1 Frequency-domain techniques

The frequency content of an rPPG signal can be determined using the Fast Fourier Transform (FFT), which converts the original signal from the time domain to a new representation in the frequency domain. A 30-second section of rPPG data from subject HP3-33 of the Physiology study was analysed in this way. The raw rPPG section produced from each colour channel (sampling rate 16 Hz) was detrended, multiplied by a Hamming window and processed using a 512-point FFT.

Figure 4.7 shows the rPPG signals and their associated frequency spectra. Each spectrum plots the magnitude of the FFT against frequency. The figure shows that each channel contains a prominent low-frequency component and cardiac component (the latter at approximately 1.2 Hz, or 72 bpm). The signals and spectra both have a common scale across the channels, showing that the green channel has the largest-amplitude rPPG and the strongest cardiac component. This agrees with the results reported in previous studies (see section 4.1.1).

For comparison, figure 4.8 shows the corresponding section of PPG acquired using a conventional finger pulse oximeter, as well as its frequency spectrum. The PPG signal has a high SNR (shown by the clear harmonics of the cardiac component in the frequency spectrum at 2.4 and 3.6 Hz). Note that this spectrum is not directly comparable with those of the rPPG signals because the PPG waveform output by the Black Shadow device has already been processed by proprietary algorithms, which include band-pass filtering.

The FFT can be used to calculate the frequency spectrum for a short window, typically

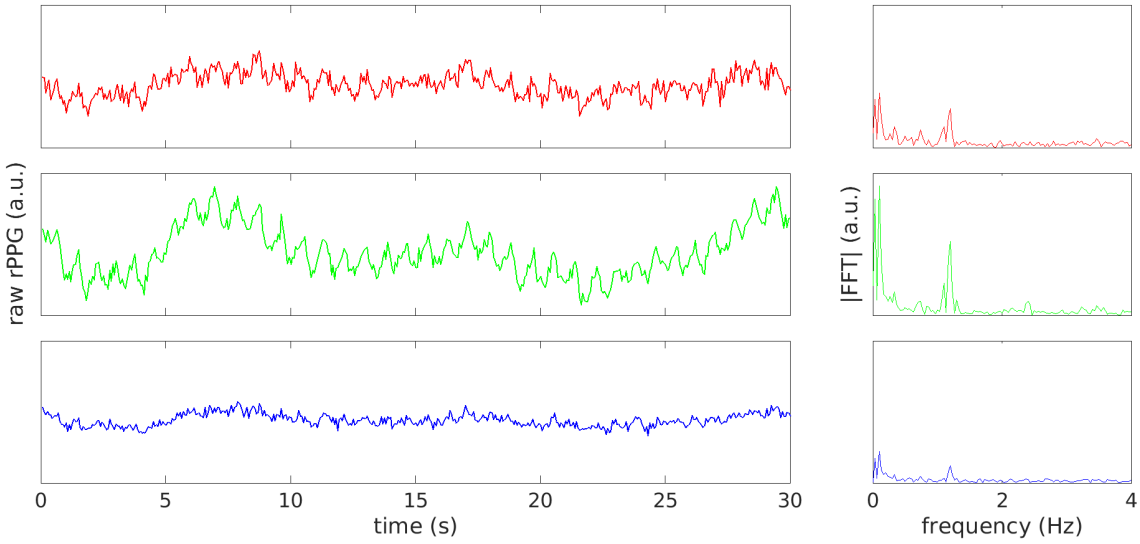


Figure 4.7: Raw rPPG signal (left) and FFT frequency spectra (right) for red, green and blue colour channels.

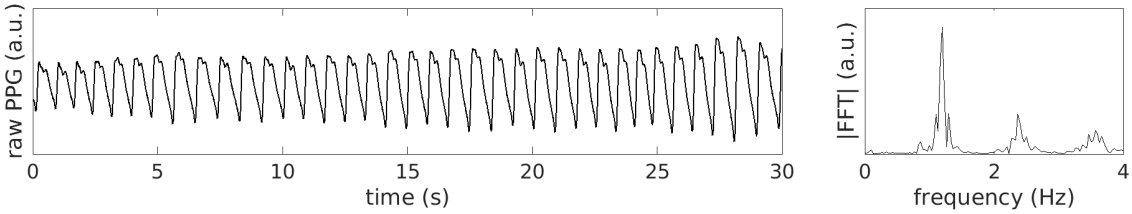


Figure 4.8: Raw finger PPG signal (left) and FFT frequency spectrum (right).

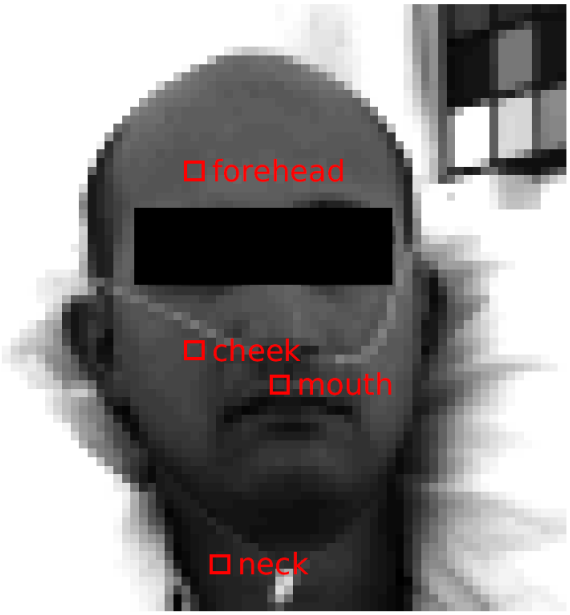


Figure 4.9: Locations of ROIs over which rPPG is derived.

30 seconds or less. Over longer periods of time, slight variability of HR and BR produces a spectrum with heavily smeared frequency components and with no temporal information. An appropriate tool to visualise the frequency components is the spectrogram, which is a representation of the frequency spectrum over time.

Some examples of rPPG spectrograms are shown in the following figures. The spectrograms are calculated using the raw rPPG signals (green channel) derived from four different sites on the subject's face, with each signal first being individually detrended to remove the DC component. The colour intensity of the spectrogram represents the magnitude of the FFT at a given frequency and time (with blue corresponding to a smaller FFT magnitude, and yellow corresponding to a larger magnitude). The spectrograms are all found by calculating the FFT over windows of 30 s, with successive windows overlapping by 25 s. Figure 4.9 shows the locations of the 24×24 pixel ROIs from which the rPPG signals are derived. Additional spectrograms, which show the effect of ROI size on the frequency components, are presented in appendix C.

4.3.2 Breathing components

Figure 4.10 shows spectrograms for each rPPG signal (with a common colour axis) over a range of possible breathing frequencies. The equivalent spectrogram for the finger PPG signal is shown in figure 4.11. The reference BR, found using the inductance pneumogram (IP) signal, is superimposed on each spectrogram in red.

Like those of the finger PPG signal, the spectrograms from the subject's forehead, cheek and neck show very little breathing-synchronous information, whereas the spectrogram of the rPPG measured at the mouth has a strong breathing component. This is because the mouth ROI contains high-gradient edges in the video frames, and the breathing-synchronous movement of the subject causes these edges to move periodically; this creates a corresponding breathing-synchronous component in the mouth rPPG signal. By comparison, the other ROIs contain few edges in the video frames, and the breathing component is thus much weaker.

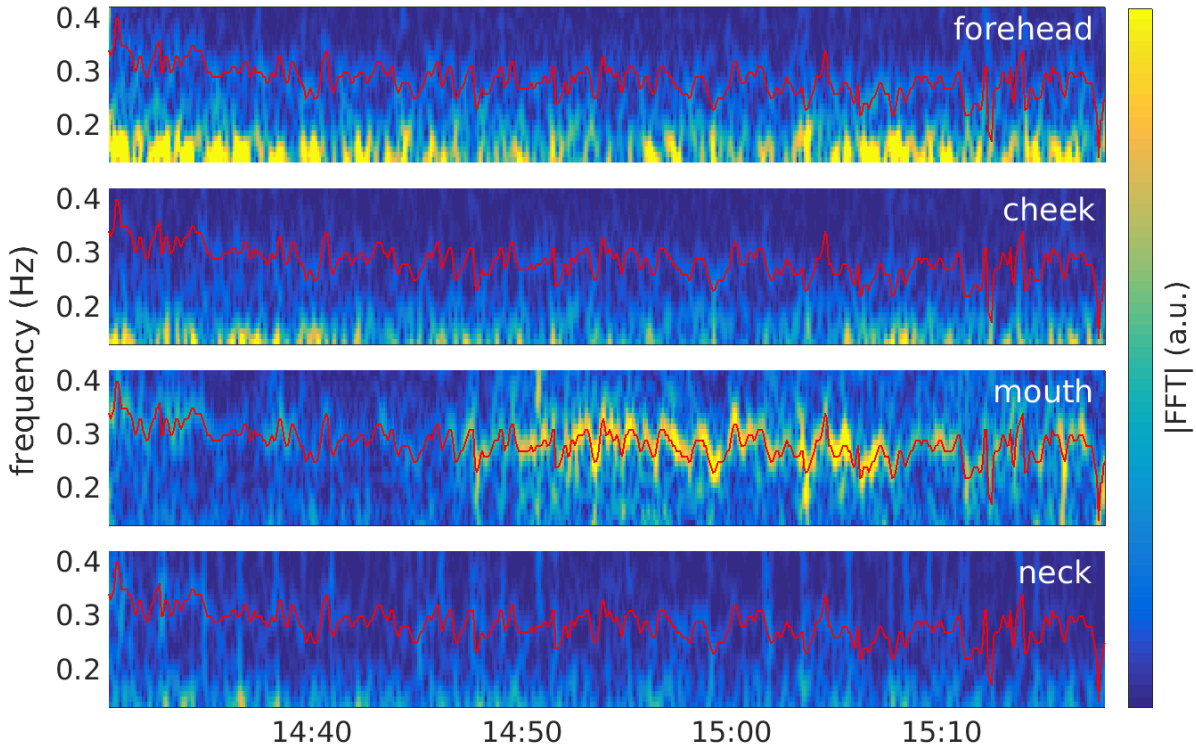


Figure 4.10: Breathing-frequency rPPG spectrograms from four different ROIs (locations in white text). The reference BR from the IP is superimposed in red.

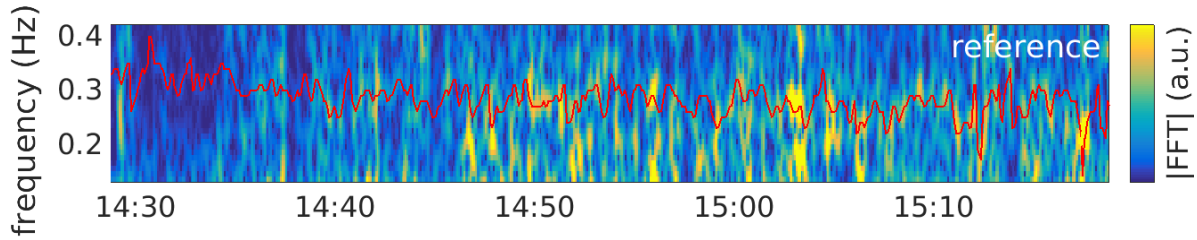


Figure 4.11: Breathing-frequency spectrogram for the conventional finger PPG signal. The reference BR from the IP is superimposed in red.

4.3.3 Cardiac components

Figure 4.12 shows spectrograms for each detrended raw rPPG signal, this time over a range of possible cardiac frequencies. These spectrograms were created in the same way as those in the previous section; the only difference is the frequency range. Again, the colour scale is consistent between sites. Figure 4.13 shows the corresponding spectrogram for the finger PPG signal. The reference HR, found using the ECG, is superimposed on each spectrogram in red.

Unlike the breathing-frequency spectrograms, the forehead and cheek ROIs exhibit the strongest cardiac-synchronous component, with the mouth and neck ROIs being affected more by transient noise (broad-spectrum power presenting as bright vertical bars in the spectrogram). The forehead and cheek areas are least prone to movement artefact as they are relatively isolated from high-gradient edges in the video frames. They are also relatively well perfused compared to the neck and mouth and therefore produce a higher-amplitude rPPG signal.

In the finger PPG signal there is again a clear HR throughout the spectrogram. The cardiac component appears to have comparable strength to the cardiac components of the forehead and cheek rPPG signals.

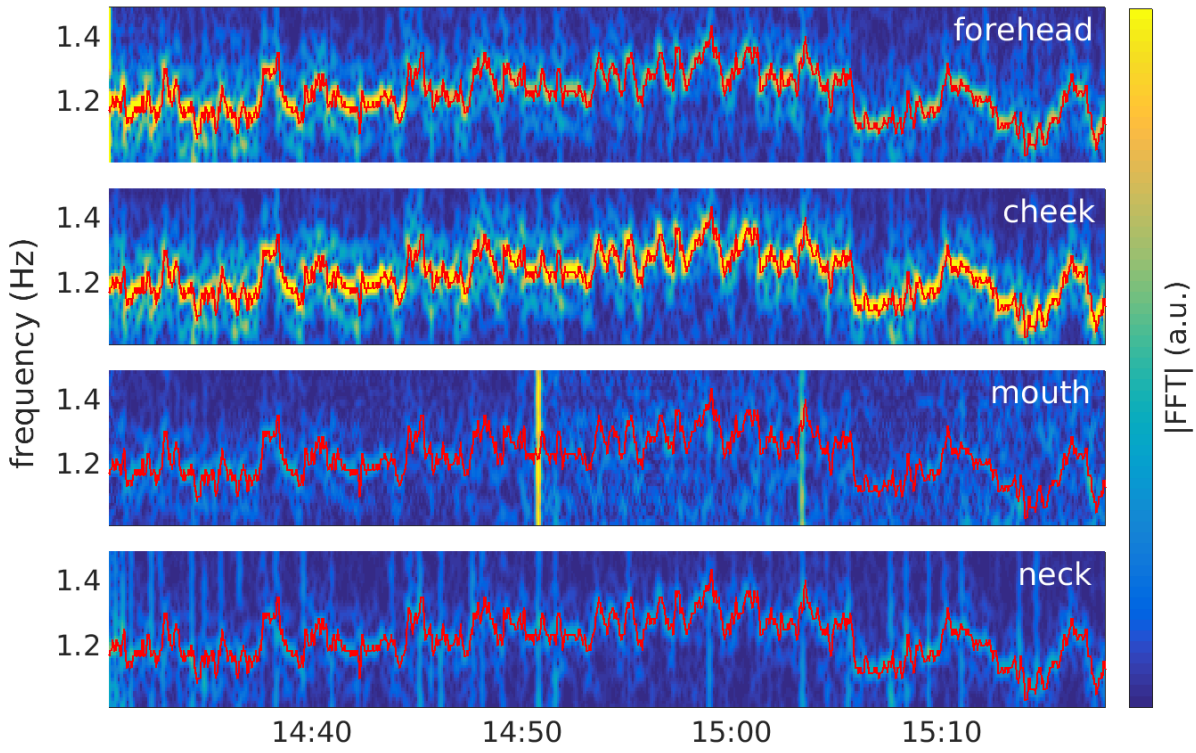


Figure 4.12: Cardiac-frequency rPPG spectrograms from four different ROIs (locations in white text). The reference HR from the ECG is superimposed in red.

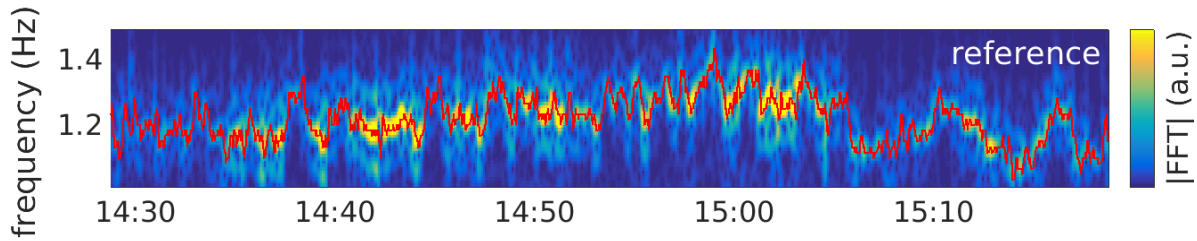


Figure 4.13: Cardiac-frequency spectrogram for the conventional finger PPG signal. The reference HR from the ECG is superimposed in red.

4.4 Challenges of non-contact monitoring

Although non-contact vital sign monitoring has a number of advantages over conventional monitoring, it also presents several challenges. These include dealing with a subject's movement, noise in rPPG signals that does not affect conventional PPG signals, the issue of ROI selection, and the difficulty of validation. These issues are discussed in this section.

4.4.1 Subject movement

Conventional PPG signals are acquired using pulse oximeters (see section 2.3), which are usually attached to a patient's finger. Pulse oximeters are prone to movement artefact [49], but this is usually due to sudden, obvious movement that is transient in nature.

Remotely acquired rPPG signals, by contrast, are derived by choosing an ROI in the video frame which corresponds to an area of the subject's skin. If the subject moves relative to the camera, a static ROI will correspond to a different part of the frame; in other words, a subject's motion is equivalent to movement of the ROI itself. This is illustrated in figure 4.14, which shows four frames from a session of the renal unit study (subject DF43). A static ROI is set up initially to cover the subject's cheek, but the subject's motion causes it to migrate to locations such as the hand and mouth[7].

Many different strategies have been suggested to address these issues, but large-scale movement of the subject causes such fundamental disruption to the rPPG signal that it is rendered useless for vital sign estimation. It is usually sufficient simply to detect such large scale movement, and stop monitoring until the motion ceases. This is reasonable in a clinical setting because if a patient is moving, they are very likely to be healthy and monitoring is less of a concern; the loss of physiological data for a few seconds is a regular occurrence, even with conventional monitoring.

[7]The final frame of figure 4.14 shows a nurse blocking the camera, another drawback which does not have to be considered when performing conventional monitoring.

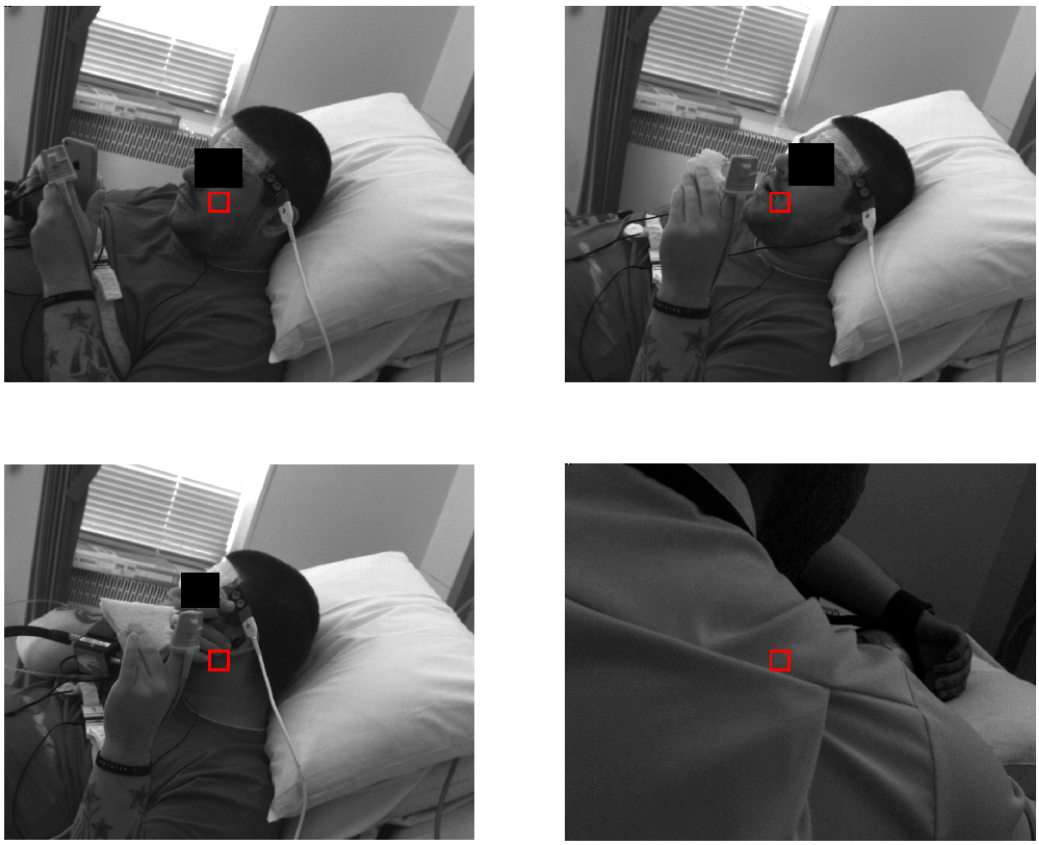


Figure 4.14: Example of subject movement leading to inconsistent ROI location relative to the subject. Images from subject DF43 of the Renal Unit study.

Motion that is not transient, however, must be dealt with in a more intelligent way so as not to impede continuous monitoring. Such motion includes the act of breathing, which can produce a dominant component in the rPPG signal (see section 4.3).

A common technique for addressing subject motion involves image tracking such that the ROI moves with the subject. ROIs can be initialised either manually [161, 162], or using an automated face detection algorithm such as the Viola-Jones detector [124, 22]. Features within the ROI can then be tracked across frames using algorithms such as the Kanade-Lucas-Tomasi feature tracker [87, 157]. Face detection and feature tracking are fields of research within their own right, and are therefore not addressed in this thesis.

4.4.2 Noise and artefacts

Physiological signals are commonly subject to noise. In this context, noise is defined as any unwanted component in the rPPG signal. Whether a component is unwanted can be determined by the purpose for which the rPPG signal is used; the breathing-synchronous component, for example, could be considered noise when trying to estimate HR. Similarly, the cardiac component in the signal might be considered noise for the purpose of determining BR.

Conventional PPG signals are subject to sources of non-physiological noise including movement artefact (usually low-frequency), electrical noise from mains-powered devices (at 50 or 60 Hz), and broad-spectrum electronic noise from the pulse oximetry circuitry. rPPG signals are also subject to these types of noise, but are affected by other noise sources as well. Each pixel in the camera's image sensor is affected by image noise, random variation in brightness caused by the discrete nature of small-scale electric charge (this can include shot noise). With mains-powered lighting, there is often a constant-frequency noise in the rPPG signal (see section 4.1.3).

High-frequency noise can be reduced in conventional PPG by low-pass filtering, but this is more difficult in the case of rPPG. rPPG signals are likely to be captured at lower

sampling frequencies than conventional PPG, which causes high-frequency noise to be aliased down to lower frequencies. These lower frequencies are more difficult to remove with low-pass filtering as they lie closer to the desired rPPG signal.

Our 2014 study [153] attempted to remove background noise using autoregressive modelling. rPPG signals were derived from two ROIs: one genuine signal at the desired location on the subject's skin, and one from a background portion of the image (such as the wall behind the subject). Autoregressive models were fitted to each rPPG signal, and poles and zeros were identified for each model. The poles in the background model corresponded to noise rather than cardiac or breathing frequencies, which allowed similar poles in the genuine model to be dismissed, leaving only the poles containing physiological information. We were able to use this method of pole cancellation to remove aliased frequency components due to artificial light flicker.

4.4.3 ROI size

To achieve a high SNR in the rPPG signal, the output from multiple pixels is averaged within an ROI (see section 4.1.2). Because the random noise from each pixel tends to have zero mean, the noise decreases as the ROI size increases (assuming the noise in each pixel is independent).

Choosing large ROIs minimises noise, but causes other problems. The rPPG signal derived from a large ROI is, physiologically speaking, the average cardiac signal in that region; this means that spatial information about the local phase and amplitude of the rPPG signal within that ROI is lost. This is undesirable in certain situations, for example when investigating PTT (see Chapter 5) or when attempting to map the local rPPG waveform (see Chapter 6). In addition, having a large ROI can leave the rPPG signal more prone to movement artefact if the ROI contains edges or other features.

The size of the ROI also affects the relative size of the cardiac and breathing components within the rPPG signal [153]. rPPG signals derived from small ROIs tend to

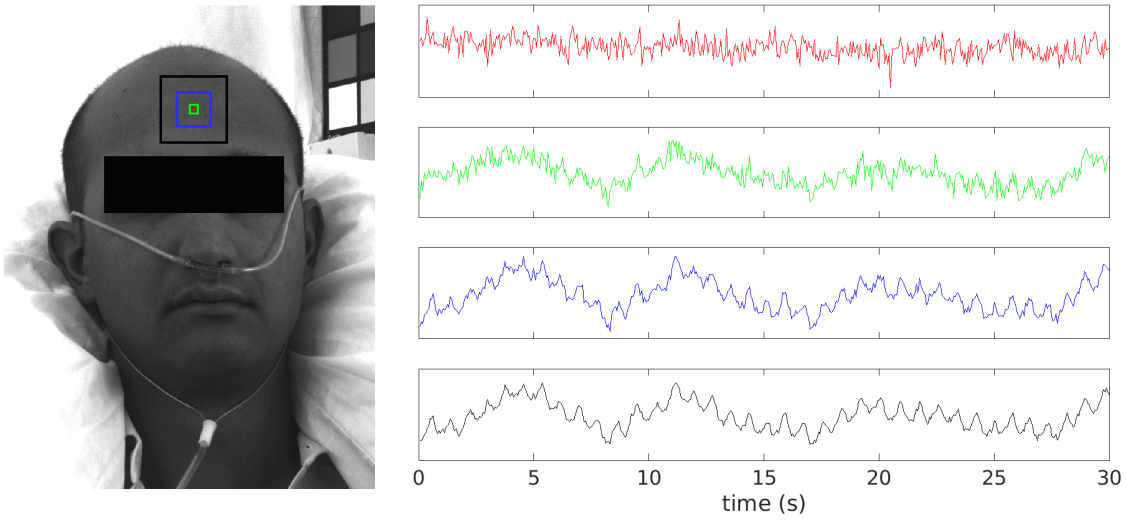


Figure 4.15: Effect of ROI size on the rPPG signal. rPPG signals are shown for ROIs of size 1×1 pixel (red), 12×12 pixels (green), 60×60 pixels (blue), and 120×120 pixels (black).

contain a relatively large breathing component because the subject's breathing-synchronous movement dominates; features such as edges and corners are large relative to the ROI and their movement greatly affects the rPPG. In large ROIs, the cardiac component of the rPPG signal is larger relative to the breathing component because the features within the ROI are relatively small.

Figure 4.15 shows the effect of ROI size on the rPPG signal. The increase in SNR with ROI size is clearly visible, and the prominence of the cardiac component also increases (relative to the breathing component) with ROI size.

4.4.4 Validation

As with any emerging medical technique, the vital signs estimated through non-contact monitoring must be validated against the existing standard before being accepted for clinical use by the medical community. This validation process typically involves comparison of the remotely-acquired estimates against a gold standard, which is considered to be the ground truth. For example, camera-derived HR estimates can be validated

against the reference HR found using ECG by comparing the error between the values produced by the two techniques.

Validation typically involves quantitative assessment of the error between the estimate and reference values. The mean absolute error (MAE) or the root mean squared error (RMSE) between paired measurements is often reported, as is mean ± standard deviation.

Some physiological estimates are more difficult to validate as they cannot be compared directly with a reference. For example, a video camera might be used to estimate the PTT between a subject's cheek and forehead (see Chapter 5), whereas the reference cPTT is measured between the subject's heart and finger (see section 2.4). In these cases, the reporting of direct error between paired estimates is not appropriate as the two values are unlikely to be the same. In these situations, other options are available. The correlation between the two sets of data (estimate and reference) can be quantified using Pearson's correlation coefficient (see section 3.2). Alternatively, regression can be used to produce a model for the two sets of data (for example, a linear fit between reference and estimate), and the model's error can then be quantified.

Validation can also be complicated by the use of proprietary algorithms in conventional clinical equipment. Taking HR as an example, clinical monitors use proprietary signal processing techniques to window and filter the HR values before they are reported to the user. In order to protect their intellectual property, medical technology companies do not usually make these algorithms publicly known; direct comparison of estimates and reference data can therefore be misleading.

Another difficulty with validation arises from the timestamping of data. Often, separate systems are used to collect different physiological data, with each measurement being timestamped by the relevant piece of equipment. If the internal clocks of the two systems are not synchronised, validation can be extremely difficult as it may be impossible to pair estimates with the correct reference value. In these cases, an attempt must be made to synchronise the two sets of timestamps retrospectively.

Chapter 5

PTT from video

Chapter 4 explored the background of non-contact vital sign monitoring using video cameras. This chapter will extend some of these ideas to consider whether PTT can be estimated using non-contact monitoring, as was briefly introduced in section 4.2.3.

PTT must be measured using two cardiac-synchronous signals; the delay is measured between the proximal signal and the distal signal. The two signals conventionally used are the ECG and the finger PPG, producing what will be denoted cPTT (see section 2.4). This chapter's hypothesis is that the rPPG signal derived from video recordings (see section 4.1) may be used to replace one or both of the ECG and finger PPG; that is, a version of PTT can be calculated by measuring the time delay between ECG and rPPG waveforms, and another version by measuring between rPPG signals from different sites on the skin. The former will be denoted remote PTT (rPTT), and the latter denoted differential PTT (dPTT) as it can be thought of as the difference between the rPTT measured at the two sites. The different types of PTT are laid out in figure 5.1 and table 5.1.

Section 5.1 explores the existing work on camera-derived PTT that has recently been published. Section 5.2 introduces the vasculature of the face, and discusses the complexity of calculating PTT in this region.

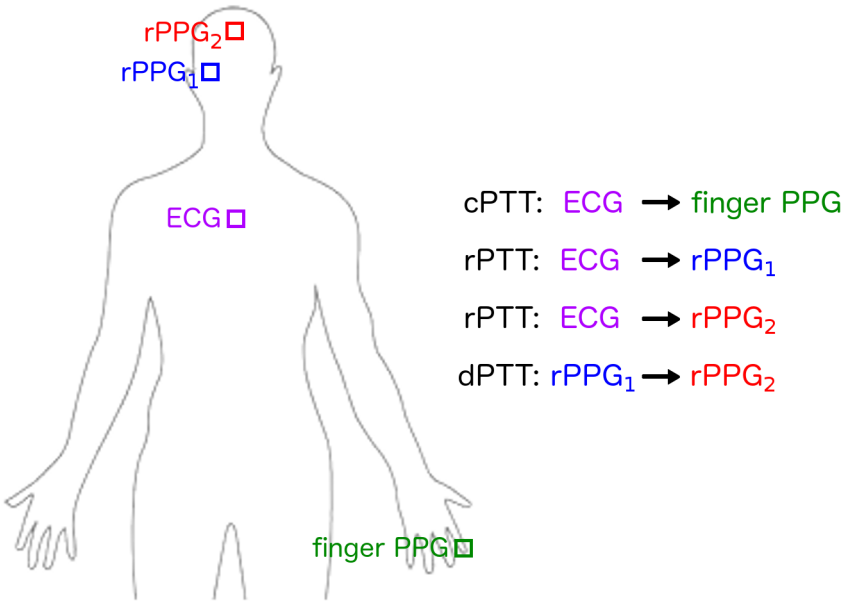


Figure 5.1: Example of the different types of PTT that can be measured using conventional equipment and a video camera.

Name	Notation	Proximal signal	Distal signal
Conventional PTT	cPTT	ECG	finger PPG
Remote PTT	rPTT	ECG	$rPPG_1$ or $rPPG_2$
Differential PTT	dPTT	$rPPG_1$	$rPPG_2$

Table 5.1: The different types of PTT and the abbreviations used to refer to them.

The steps for obtaining an estimate of rPTT and dPTT from video include preprocessing of the video (section 5.3), selection of an ROI (section 5.4), derivation of an rPPG signal (section 5.5), processing of the rPPG signal (section 5.6) and finally calculation of the rPTT(section 5.7) and dPTT (section 5.8). The results (section 5.9) are then discussed in section 5.10.

5.1 Previous work

Before the work for this thesis was started, there was no real attempt in the field to calculate PTT using video cameras. As was mentioned in section 4.2.3, Verkruysse *et al.* [161] came close to measuring PTT by considering the phase difference between rPPG signals, but the authors did not extend their work to investigate PTT.

Since 2014, however, several papers that measure camera-derived PTT have been published. Shao *et al.* (2014) [140] measured dPTT in 10 subjects using rPPG signals derived from the mouth and palm areas, finding average values of 30-40 ms between face and hand. The authors made no attempt to assess the validity of their dPTT measurements.

Murakami *et al.* (2015) [106] derived rPPG signals from the wrist and ankle, and estimated dPTT from the time difference of the pulse peaks. Using video recordings from 10 male subjects, the authors found a correlation coefficient between camera-derived dPTT and SBP of -0.879, indicating strong negative correlation. However, only 10 patients were monitored—each for just 30 seconds—producing one value of dPTT and SBP per subject. This could account for the apparent strong correlation reported in this study, which is comparable with the correlation between BP and cPTT in the literature (see section 2.5.2).

In another study, Kaur *et al.* (2015) [70] created rPPG signals from the face and hand, and estimated dPTT by cross-correlating the two signals. dPTT was estimated

from 15 volunteers in both "normal" and "stressed" states, and the authors concluded that dPTT was a marker for stress (since the average dPTT value was higher when the subjects were in the stressed state). Again, only short video recordings were used, with subjects remaining in each state for one minute. The authors claim a statistically significant difference between dPTT measured in the stressed and normal states, but the small cohort and short duration of recording cast some doubt on the validity of this claim.

In 2016, a paper by Balasingham *et al.* [139] was published in which the authors estimated rPTT and dPTT using a combination of ECG and rPPG signals derived from a subject's forehead and palm. Three subjects were recorded at rest and after exercise, and BP was measured using an invasive arterial line. Across nine 3-minute recordings, the authors found a mean correlation coefficient of -0.716 between SBP and rPTT, and of -0.605 between SBP and dPTT. Again, these results should be considered in the context of the small number of subjects and limited length of the recordings.

Another recent study by Jeong and Finkelstein (2016) [61] estimated dPTT from a subject's head and palm. Nine ten-second recordings were obtained from each of 7 volunteers, who were double monitored for both cPTT and SBP. The authors found an average correlation of 0.85 between cPTT and dPTT, and of -0.80 between SBP and dPTT. In the latter case, it was noted that although each subject had a relatively high correlation, the gradient of the fit was different, indicating that individualised calibration would be required for the accurate estimation of SBP in each patient.

Volynsky *et al.* (2016) [164] took a slightly different approach, asking volunteers to press their palm against a glass surface and recording from the other side of the glass for 30 seconds. Illumination was controlled using LEDs emitting light at a specific wavelength, and a filter was used to allow only this wavelength to enter the camera, with the resulting video recordings used along with ECG to estimate rPTT. The authors did not attempt to validate the rPTT estimates, but noted that comparable values of rPTT

were obtained from the left and right hands of each subject.

Each of the studies mentioned in this section attempted to estimate PTT using video cameras, but used only short recordings from a limited number of subjects. In this chapter, longer recordings from over 40 subjects will be used to assess the robustness and consistency of PTT estimation.

5.2 Vasculature of the face

The relationship between BP and conventional heart-finger cPTT has been thoroughly investigated in the literature (see section 2.4). By comparison, the relationship between BP and non-conventional PTT (such as the rPTT and dPTT measured in this chapter) is less well understood. While the path of blood flowing from the heart to the finger is straightforward (with no arterial branching), the path of blood elsewhere in the vasculature is more complex.

The parts of skin visible to a video camera are usually limited to the face, neck and hands depending on the clothing being worn by the subject. Throughout this chapter, ROIs will be selected from the parts of the head consistently visible to the camera, specifically the neck, cheek and forehead.

The blood supply to the face is complicated by the presence of different blood supplies. Blood does not flow directly from the cheek to the forehead, for example; the two areas of skin are part of different branches of the complex vasculature, so it is not immediately obvious where the pulse should arrive first. Most of the blood to the forehead and cheek is supplied from the external carotid via the temporal (forehead) and facial (cheek) arteries, but the blood to the area of the eyes is supplied from the supraorbital artery via the internal carotid.

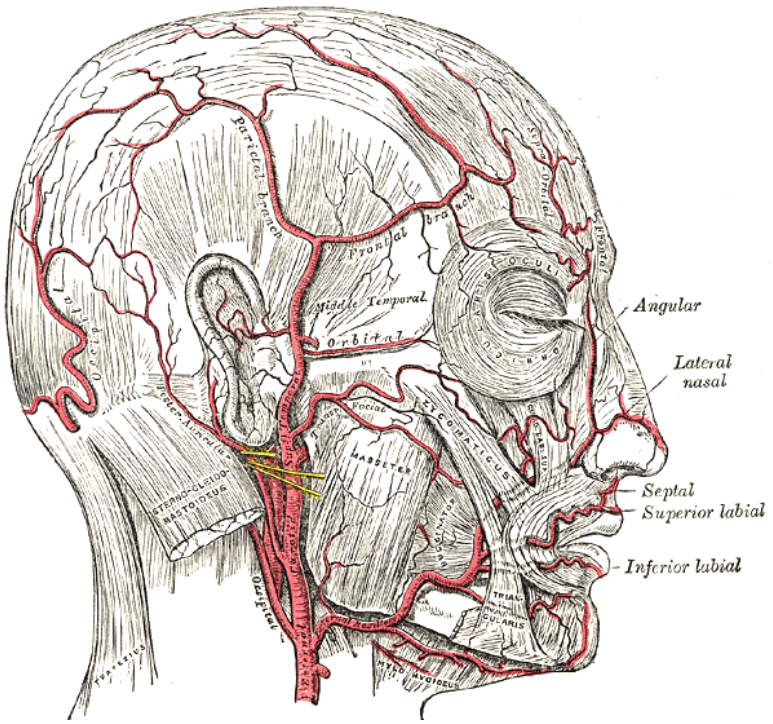
This branching complicates the measurement of dPTT; if local blood vessels in two sites are fed by different branches of the arterial tree, then measuring dPTT between

those sites is likely to be misleading (since the pulse does not travel directly from one site to the other). For example, measuring dPTT between the left cheek and the right cheek is likely to produce a dPTT that is close to zero, since the pulse arrives in each cheek at approximately the same time (fed by different branches of the vasculature). When measuring dPTT between two sites, therefore, the assumption of a direct path between those sites must be assessed carefully.

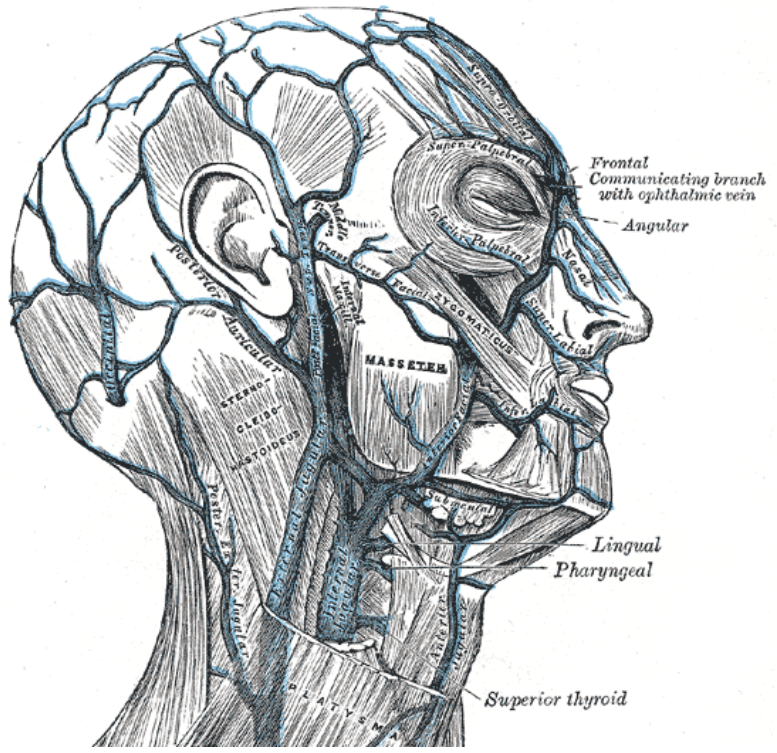
Figure 5.2 shows the vasculature of the face; note that the largest blood vessels run up the neck, which also has less vessel branching than other areas. The neck may therefore be a good site to investigate camera-derived PTT.

In addition, camera-derived PTT is based on rPPG signals, which are affected more by superficial blood flow in the skin than conventional PPG signals [90]. This could affect the extent to which PTT measured in the face is affected by vasoconstriction and vasodilation.

Despite these complications, it is likely that a sudden change in BP which changes the cPTT would also affect facial rPTT and dPTT in some way. Since the focus is on detecting *changes* in BP (rather than measuring absolute values), looking for a change in camera-derived PTT should therefore be a valid strategy.



(a) Arteries of the face



(b) Veins of the face

Figure 5.2: Anatomy of the facial vasculature. Adapted from the online edition of Gray's Anatomy [42], which is available at http://www.bartleby.com/107/ (public domain).

5.3 Preprocessing of video recordings

This chapter involves analysis of data produced in the Physiology study, which is described in detail in section 1.4.2. The data set consists of 40-minute video recordings from 43 subjects, each of whom were subjected to a hypoxic environment in order to produce changes in their vital signs. Data from this study were used because subjects were volunteers (rather than hospital patients) and were therefore more likely to remain still and produce robust estimates of camera-derived PTT. In addition, the video recording environment was more carefully controlled than in the Renal Unit study, producing rPPG signals with a relatively higher SNR.

As described in section 1.4, the amount of data produced by the video camera in the three studies is extremely large; the binary video files are saved in a compressed format so that they can be stored long-term. This means that analysis of the video files is a slow process as the already large files need to be decompressed before the data are accessible. In addition, conventional decompression is not feasible due to the file size. Rather than writing the uncompressed data to file in an intermediate step, decompression therefore has to be performed on-the-fly within primary computer memory. This allows the processing to be performed an order of magnitude faster than would otherwise be possible.

The video files from the Physiology study were therefore subjected to a preprocessing stage[1], with the goal of cropping and downsampling each video and saving it in a simpler format that would be easier to load and access using MATLAB (Mathworks, MA, USA), a suite of mathematical computing software. This allowed the video data to be analysed more quickly, and prevented unnecessary repetition of the decompression process.

For each recording session, the video files were preprocessed as follows:

1. A custom C++ application was used to load the files into memory, while performing

[1]The preprocessing was carried out using software from a codebase contributed to by several members of our research group (M. Villarroel, A. Guazzi, J. Jorge and the author), and the preprocessed data has previously been used in other publications [44, 43].

decompression on-the-fly using the Boost C++ libraries (http://www.boost.org).

2. The raw uncompressed video data were loaded from the C++ application into MATLAB using a MEX function. Within MATLAB the video existed in the form of a $M \times N \times 3 \times T$ matrix, representing T frames of a $M \times N$ pixel video with 3 colour channels (red, green and blue).

3. The subject's face was located in the video's first frame using a Viola-Jones frontal face detector, a robust method of object detection [163] which is commonly used in the field of computer vision to solve the problem of face detection. The bounding box of the face (plus a margin) was defined as the area of interest (AOI) of size $M_a \times N_a$ pixels.

4. For each colour channel and for each frame in the video:

   (a) The Viola-Jones detector was again used to locate the face. If the centroid of the face had moved more than a certain threshold from the previous AOI's centroid, the AOI was re-initialised about the new centroid.

   (b) The AOI was split into a grid, with each element of the grid containing a 12×12 pixel area of the frame.

   (c) The pixels within each element were summed, downsampling the original $M_a \times N_a$ AOI to a $M_a/12 \times N_a/12$ image.

5. The downsampled frames from each colour channel were recombined to form a $M_a/12 \times N_a/12 \times 3 \times T$ matrix.

6. This matrix of downsampled video data was then saved in the form of a MATLAB .mat file.

This process (also shown in figure 5.3) reduced each set of binary compressed video files (approximately 500 GB of data) to a .mat file with size of the order of 2 GB. Only

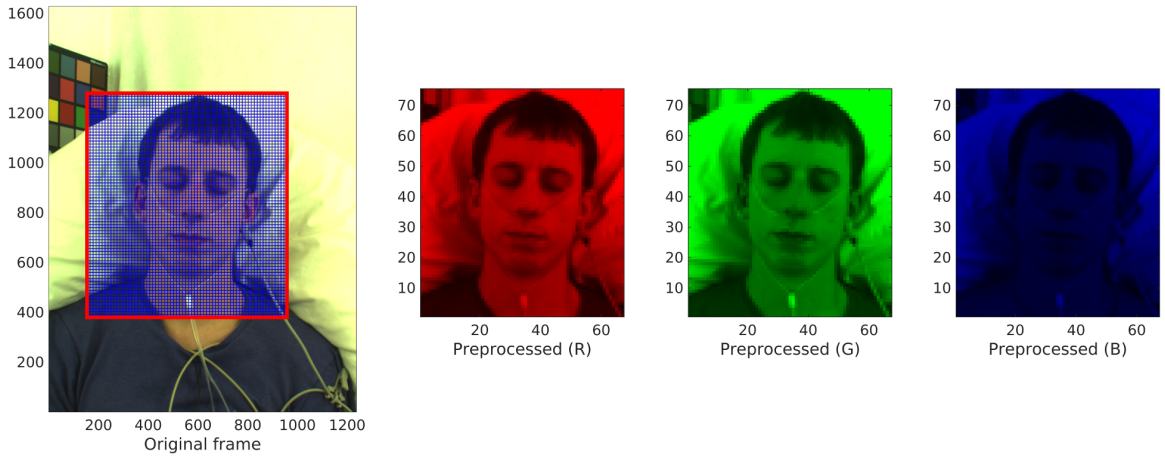


Figure 5.3: Preprocessing of a video frame. Left: original frame, overlaid with AOI (red) and processing grid (blue). Right: red, green and blue channels of preprocessed image. Note the lower spatial resolution caused by the downsampling.

spatial information below a resolution of 12×12 pixels was lost; in effect, each video was reduced to a set of rPPG signals, each found using a element of size 12×12 pixels. This element size was chosen as a compromise; smaller grid elements would produce rPPG signals with a low SNR, and larger elements would reduce the spatial resolution of any further analysis (see section 4.4.3).

5.4 ROI selection

In order to assess the optimal site for the estimation of rPTT, and to allow the calculation of dPTT between ROIs, multiple ROIs were identified for each subject of the Physiology study. ROIs on the forehead, neck and cheek were manually chosen using reference frames from the video recordings. These areas were chosen as they contain relatively smooth exposed skin, are consistently visible to the camera, and are hypothesised to be sufficiently far apart that a non-zero dPTT can be calculated between them.

2×2 element ROIs were chosen, which is equivalent to 24×24 pixels in the original video (since each element in the preprocessed video contains data from a 12×12 pixel area). This size was chosen as it was small enough to contain local information, but large

enough for a reasonable SNR (see section 4.4.3).

The process of choosing the ROIs for a given subject was as follows:

1. The preprocessed frame (see section 5.3) at the 20-minute mark was displayed. This frame was chosen as it occurred approximately halfway through the session, when the subject was more likely to be still.

2. The ROIs were manually selected by determining the coordinates of:

   (a) a point on the forehead above the subject's eye, two centimetres above the subject's eyebrow

   (b) a point on the cheek below the subject's eye, level with the subject's nostrils

   (c) a point on the neck below the subject's eye, level with the subject's chin.

3. A 2×2 element ROI was selected centred on each point to produce a forehead, cheek and neck ROI.

4. Where possible, the left eye (and therefore the left side of the face) was used for consistency.

5. Where the subject's pose or clothing complicated the selection of an ROI, appropriate adjustments to the ROI were made by visual inspection, with the criteria above used as a starting point.

In addition to the ROIs above, further ROIs were chosen to investigate the local variation in rPTT and dPTT around the forehead and neck regions. A particular grid element was chosen as a central point, and additional ROIs were chosen by selecting the elements above, below, and to either side (leaving a gap of one element). Together, these smaller ROIs create a cross formation. These ROIs were chosen to be 1×1 elements (equivalent to 12×12 pixels in the original video), as the smaller size allowed for greater localisation of the rPTT information (at the expense of decreasing the SNR). This

Figure 5.4: Example of ROI selection for subject HP3-33 of the Physiology study. Left: forehead (red), cheek (green) and neck (blue) ROIs. Right: cross formation of ROIs (white), centred on the forehead and neck.

increased localisation was required because each ROI in the cross formation is very close to the others.

The cross formations were constructed around the forehead location chosen in step 2a, and around the neck location chosen in step 2c.

An example of all of the selected ROIs for one subject is given in figure 5.4.

As described in section 4.4.1, image tracking can be used in an attempt to ensure that the ROI moves with the subject. However, such methods of image tracking are outside the scope of this thesis, and so static ROIs are used instead; subjects in the Physiology study were instructed to remain still as much as possible so as to minimise the problems associated with movement.

5.5 Derivation of the rPPG signal

The rPPG signal was derived from each ROI using the method described in section 4.1.2, which consists of spatial averaging of the grid elements within the ROI. By producing a

single rPPG value per frame, an rPPG time series was created with a sampling rate of 16 Hz, equal to the video frame-rate. The green colour channel was used as this channel tends to have the largest SNR (see section 4.1.1).

5.6 rPPG processing

Before being analysed further, the rPPG waveforms had to be processed to remove noise using filtering, and resampled to a higher sampling frequency to increase the resolution of the PTT estimates.

5.6.1 Filtering

Digital filtering is a common way of isolating the cardiac-synchronous component of the rPPG signal, used ubiquitously throughout the rPPG literature (see section 4.4.2). However, filtering must be performed with particular care for the purposes of calculating PTT, because it has the effect of shifting peak location. Zero-phase filtering is used throughout this thesis so as to avoid the introduction of delays, but even zero-phase filtering can affect peak location. This is because the removal of high-frequency information tends to cause steep gradients to be flattened. Figure 5.5 shows the effect of filtering on a finger PPG signal from the Physiology study. The filter is a band-pass zero-phase finite impulse response (FIR) filter tuned to cardiac frequencies. Note that as the filter order increases, the amplitude of the pulse and the location of the peak both change.

Although filtering does affect the peak location, it does not affect the PTT because the shift is consistent; the peak-to-peak distance in figure 5.5 is the same regardless of the filter used.

It is important to note that, generally speaking, absolute values of PTT are of less interest to clinicians than relative changes. Different methods of calculating PTT can

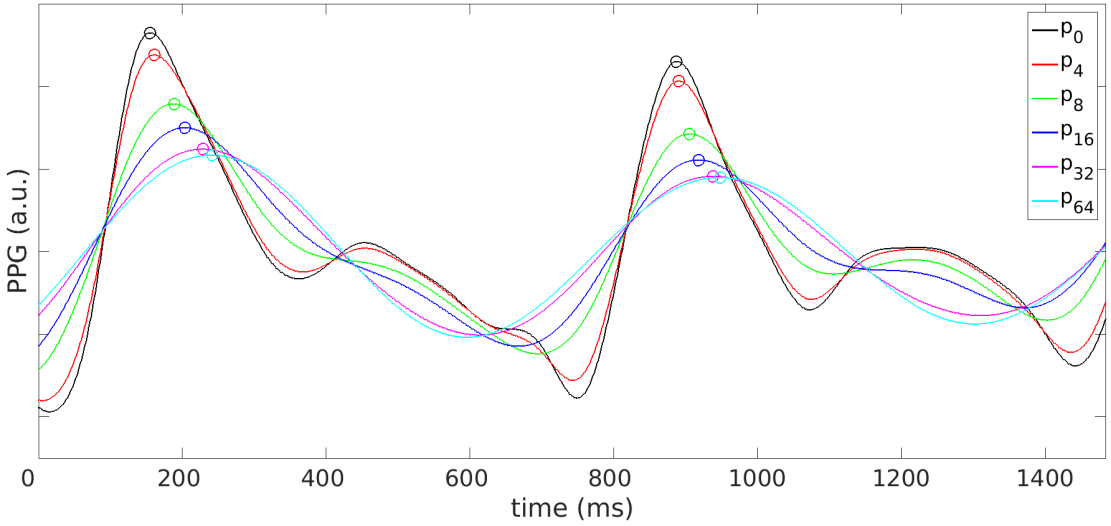


Figure 5.5: Effect of filtering a finger PPG signal using a band-pass zero-phase FIR filter.

produce markedly different measurements, but a sudden change in PTT (as long as it is measured via a consistent method) will always be of interest as it suggests the possibility of a change in BP (the relationship between BP and PTT has been explored in Chapter 2). A small shift in peak location caused by filtering is acceptable because the shift is consistent; relative changes in PTT will be unaffected. Figure 5.6 demonstrates this consistency, showing the beat-to-beat interval (BBI) for filtered PPG signals compared to unfiltered signals. The areas in which there is poor agreement correspond to noisy sections in the underlying PPG waveform.

Chosen filter

The filter used was a FIR band-pass filter designed to retain cardiac frequencies at a range of possible resting HR. The coefficients of the filter were chosen using MATLAB's digital filter design tool. The cut-off frequencies were chosen based on the 1st and 99th percentiles of HR recorded from adult subjects of the Physiology study and Renal Unit study (see figure 5.7), producing a filter with an attenuation of approximately 5.6 dB at 0.77 Hz (46 bpm) and 2.10 Hz (126 bpm) below and above the pass band respectively. The

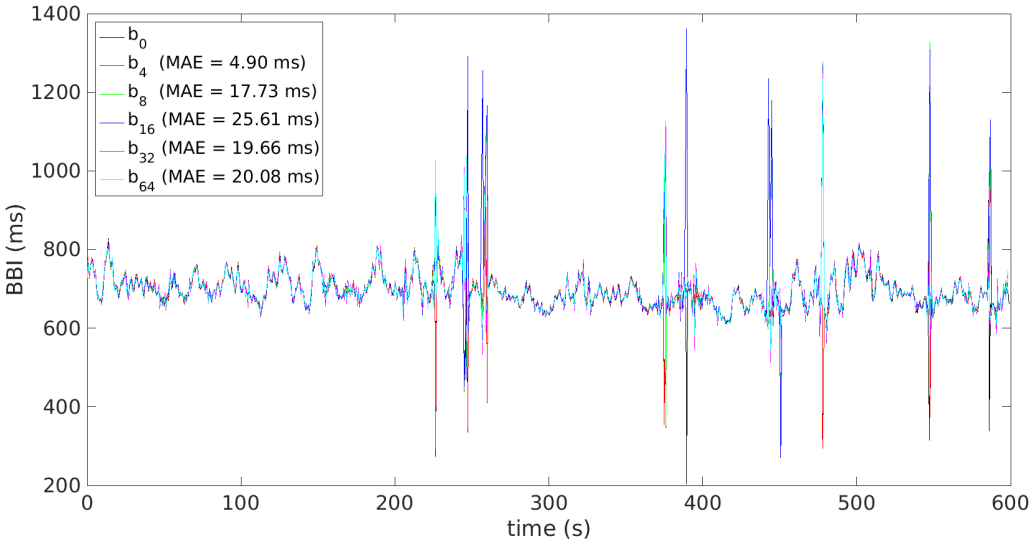


Figure 5.6: Effect of filtering on BBI measurement. The filtered signals are labelled b_n, where n is the order of the filter (the unfiltered signal is labelled b_0). The MAE of the peak locations is also shown.

attenuation at the cut-off frequency is typically 3dB, but is higher in this case because the filter order is relatively large and the cut-off frequencies are small compared to the sample rate. Zero-phase filtering was achieved using MATLAB's `filtfilt` function, which filters the signal in both the forward and reverse direction. In practice, therefore, the cut-off frequencies were attenuated by approximately 11.2 dB, and the effective filter order was twice that of the underlying band-pass filter.

An effective filter order of 64 was chosen to obtain strong attenuation outside the pass-band.

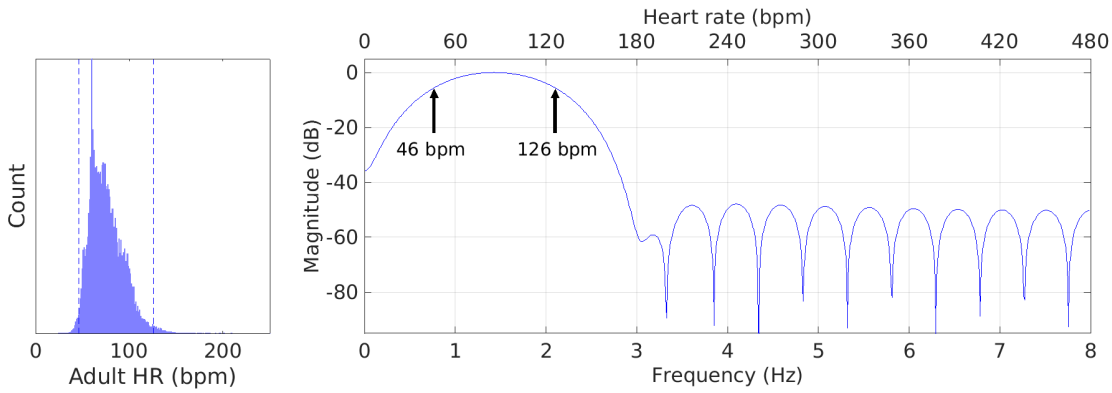


Figure 5.7: Histogram of HR values recorded in the Physiology and Renal Unit studies (left) and frequency response of the filters described in section 5.6.1 (right). 1st and 99th percentiles are marked with a dotted line. Note the peak at 60bpm in the histogram, caused by subjects with a pacemaker.

5.6.2 Resampling

In many cases, the peak locations within the rPPG signal need to be known as precisely as possible. When calculating dPTT between two rPPG signals, the resolution of the dPTT estimate will depend on the frame rate of the video. This frame rate (16 Hz in the case of the Physiology study) is relatively low compared to that of conventional PPG, which is commonly sampled at 256 Hz. Since peak detection can only locate peaks to the nearest sample, this means that the time difference between true peak location and detected location can be up to half the sampling period in the worst-case scenario. Such a difference can cause a significant error in the calculation of PTT if the sampling rate is low.

To minimise this error in peak detection, the rPPG time series can be resampled to 1kHz using piecewise cubic spline interpolation, taking into account the fact that the rPPG waveform is a near-sinusoid. This process has the effect of reconstructing peaks poorly located as the result of using a low sampling frequency (see figure 5.8), allowing their peaks to be estimated more accurately. Note that resampling does not create information that was absent in the original rPPG signal; rather, resampling is used to estimate peak location more accurately using the prior information that the

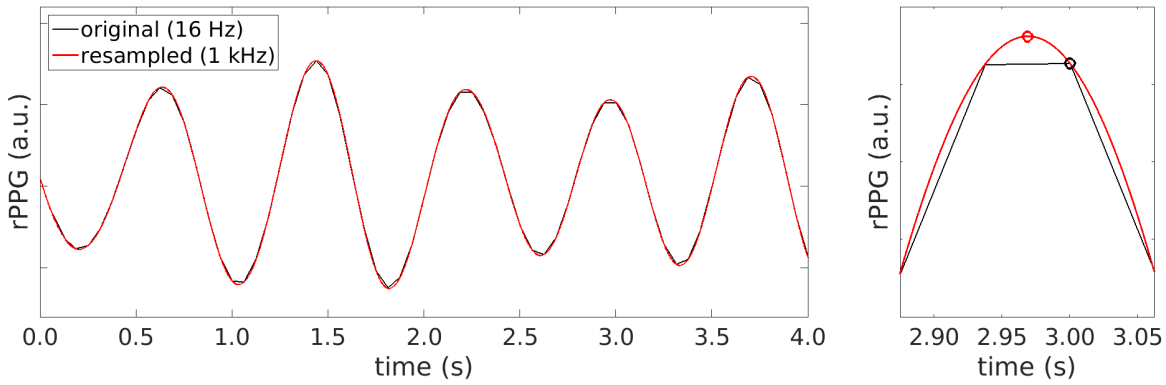


Figure 5.8: Reconstruction of an rPPG waveform using cubic spline interpolation. Left: original and resampled waveforms, with the original showing sampling artefact around the peaks. Right: a peak detected using the original waveform (black circle) and the resampled waveform (red circle), with the peak location differing by 31 ms, equivalent to approximately half the sampling period of the original signal.

underlying PPG waveform varies smoothly between any two samples.

The choice of cubic splines as a method of interpolation is justified in appendix D.

Peak reconstruction error

In this section, the peak reconstruction properties of cubic spline interpolation are investigated.

Three signals were created with a sampling rate of 1024 Hz: a sinusoid, a triangle waveform and a pseudo-PPG created by superimposing four sinusoids. Peak detection was performed on each signal to find a ground truth for the location of the peaks in the signals.

Each signal was downsampled to a test frequency f_t, and then resampled back to the original 1024 Hz using cubic spline interpolation. Peak detection was again performed on the reconstructed signal, and the average error was then found between the locations of the reconstructed peaks and the ground-truth peaks.

The mean and median errors are shown for a range of f_t in figure 5.9. The triangle waveform has the largest error (since it contains very high frequencies which are poorly represented by a downsampled signal), but in the case of the sinusoid and the pseudo-PPG

signals the error rapidly decreases below 1 ms above a f_t of 5 Hz.

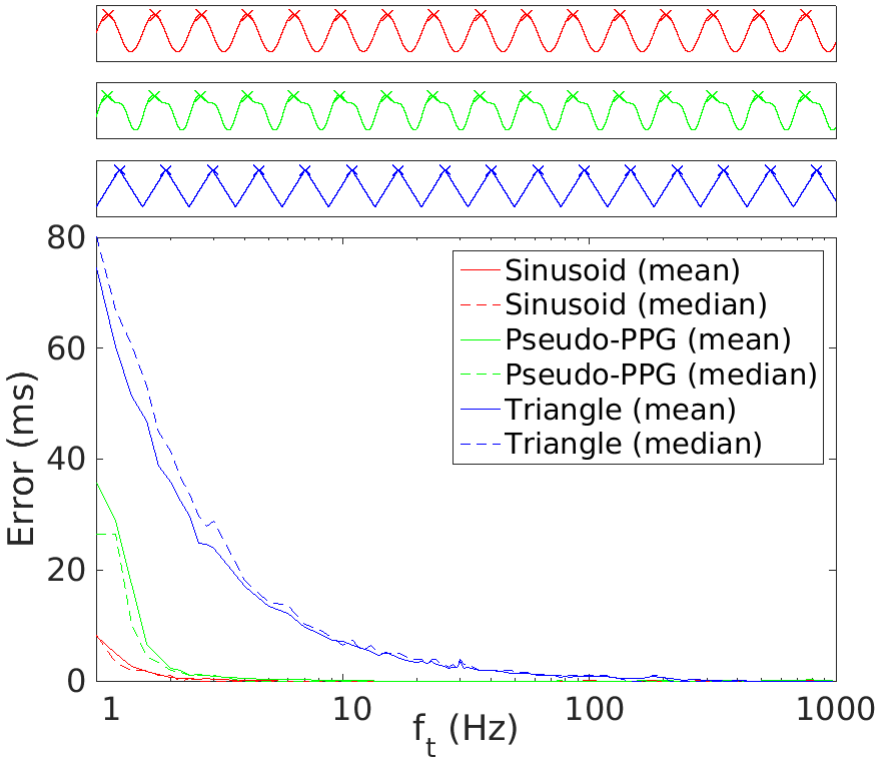


Figure 5.9: Average error between reconstructed peaks and ground-truth peaks for three different artificial signals.

The peak reconstruction was also tested on real signals recorded from subject HP3-33 of the Physiology study: a PPG signal recorded conventionally at the finger, and an rPPG signal recorded from the subject's forehead. In this case, the ground-truth was considered to be the locations of the peaks in the original signal. The error is shown in figure 5.10, which shows that the error drops below 1 ms above a f_t of 10 Hz. This indicates that the use of cubic spline interpolation is sufficiently accurate to locate peaks in rPPG signals with sampling rates of over 10 Hz.

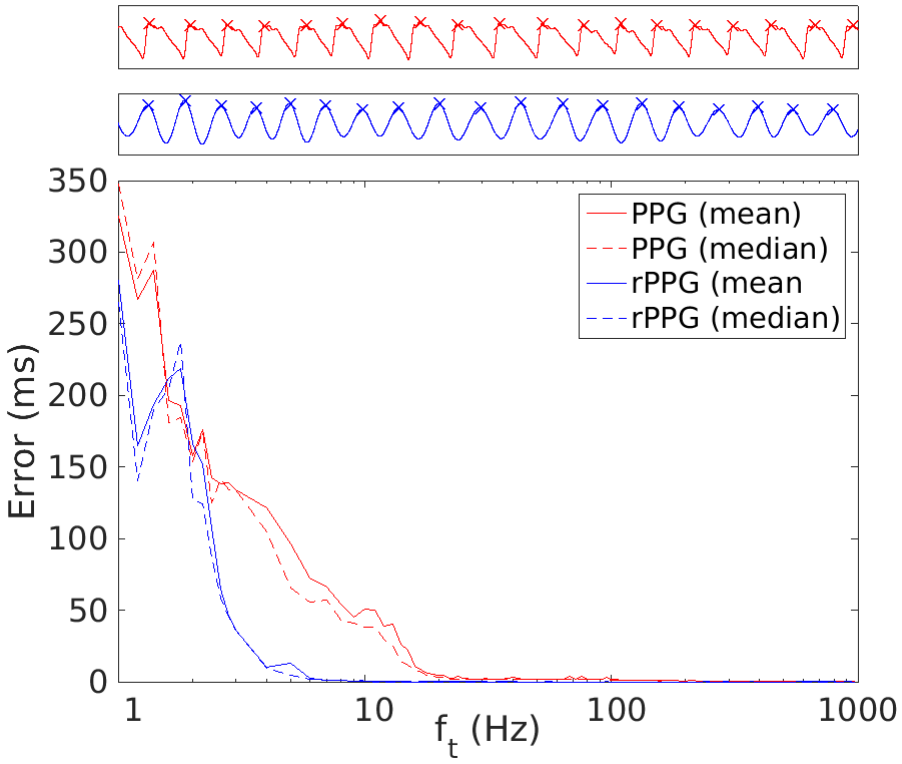


Figure 5.10: Average error between reconstructed peaks and ground-truth peaks for a real finger PPG signal and a real rPPG signal recorded from a subject's forehead.

5.7 rPTT estimation

The method for estimating rPTT consisted of four stages, which are described in more detail in the sections below. The stages included synchronisation of the ECG and rPPG signals (section 5.7.1), interpolation onto common timestamps (section 5.7.2), calculation of raw rPTT (section 5.7.3), and finally processing of the rPTT data (section 5.7.4).

5.7.1 Synchronisation

In the Physiology study, the reference data were collected by the Stowood Black Shadow device (see section 1.4.2). This equipment was separate from the computer controlling the video camera, resulting in asynchronous data collection. The reference and camera data therefore had inconsistent timestamps caused by clock drift between the two computers,

even though the data were recorded simultaneously.

We calculated (and compensated for) the lag between the reference and camera data streams by using the physiology of the subject to synchronise the two systems. This was achieved by cross-correlating first the finger PPG and the camera rPPG signal, and then the reference HR and the HR estimated from the rPPG signal.

Cross-correlation measures the similarity of two signals as one is shifted in time relative to the other. The output cross-correlation function between signals f and g is therefore a function of the time lag τ between the two systems (assumed to be constant), defined as

$$f \star g = \int_{-\infty}^{\infty} f^*(\tau)\, g(t+\tau)\, \mathrm{d}\tau \tag{5.1}$$

where f^* is the complex conjugate of f, and $\star$ denotes the cross-correlation operation.

Since the PPG and rPPG are both being generated by the same source (the beating of the heart), they should be approximately in phase throughout the session, with only a small phase difference due to the different measuring sites[2]. Similarly, the reference HR estimated by the Stowood equipment should align with the HR estimated from the rPPG. The location of the peak in the cross-correlation function therefore determines the most likely lag between the two systems.

For each 40-minute session, the steps of calculating the lag between the PPG data streams were as follows:

1. The finger PPG and camera rPPG signals were each filtered using the filter described in section 5.6.1, and then linearly interpolated onto a set of common timestamps with a sampling frequency equal to the frame-rate of the video camera (16 Hz).

2. The two signals were normalised by subtracting their respective mean and dividing by their respective standard deviation.

[2]This phase difference gives rise to the differential pulse transit time between the reference PPG's distal site (the finger) and the camera rPPG's distal site (e.g. the cheek).

3. The cross-correlation function between the two PPG signals was found.

4. The maximum in the cross-correlation was found along with its corresponding phase shift (in samples).

5. The shift was divided by the common sampling frequency to produce a phase lag estimate.

Steps 3-4 can be expressed mathematically as

$$\phi = \arg\max_{t} \int_{-\infty}^{\infty} f^*(\tau) g(t+\tau)\,\mathrm{d}\tau \tag{5.2}$$

where ϕ is the lag, f^* is the complex conjugate of the finger PPG signal, g is the camera rPPG signal and τ is the time by which g is shifted with respect to f^*.

Similarly, the steps of calculating the lag between the HR data streams were as follows:

1. The HR derived from the rPPG signal (by performing beat detection during sliding 30-second windows) and the reference HR were linearly interpolated onto a set of common timestamps with a sampling frequency of 1 Hz.

2. The cross-correlation function between the two HR signals was found.

3. The maximum in the cross-correlation was found along with its corresponding phase shift (in samples).

4. The shift was divided by the common sampling frequency to produce a phase lag estimate.

An example of this synchronisation is provided in figure 5.11, which shows the camera and reference HR data before and after synchronisation, and the cross-correlation function which determines the correct lag of -74 s.

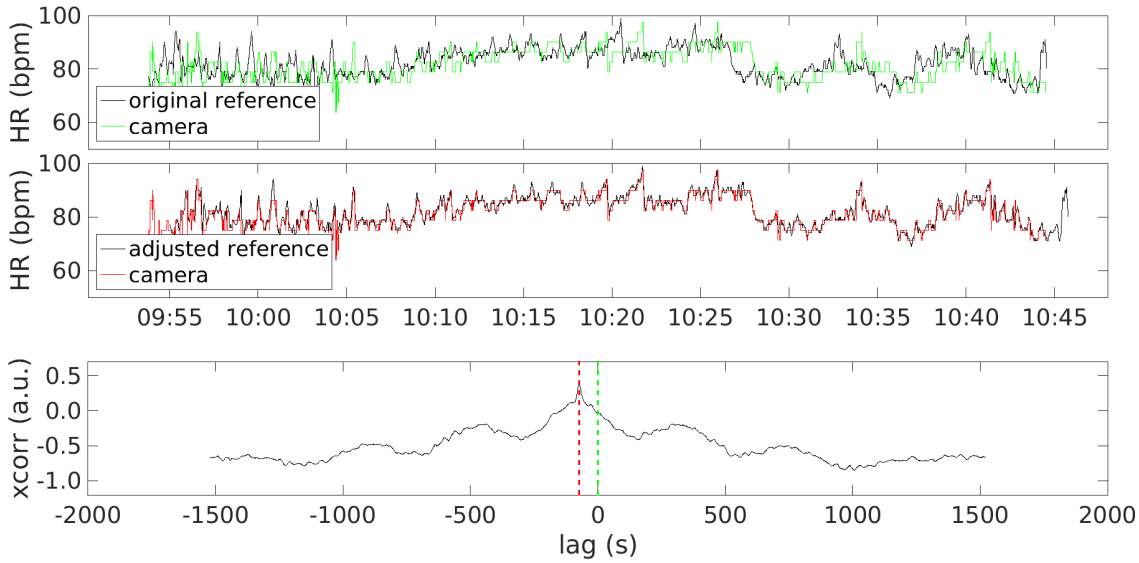


Figure 5.11: Example of lag estimation using cross correlation of the two HR data streams. Top: data before synchronisation. Middle: data after synchronisation. Bottom: cross correlation function showing the zero-lag point (green dashed line) and the point of maximum cross-correlation (red dashed line).

For 34 of the 43 sessions in the Physiology study, the PPG-derived lag and the HR-derived lag agreed closely (within 3 seconds of each other), and the more accurate PPG-derived lag was therefore chosen. Of the remaining nine sessions, the correct lag was chosen manually in seven sessions by visual inspection of the waveforms, and two sessions were discarded because a lag leading to alignment of the two signals could not be found.

Once the lag was determined, the timestamps of the reference data for each session were adjusted accordingly such that the camera and reference data became synchronised. This adjustment artificially forced the phase difference between the PPG and rPPG signals to be zero; the real (non-zero) phase difference between the signals was not recoverable. The synchronisation could therefore only be considered accurate to within the period of one heart beat (approximately $\pm$ 500 ms).

The distribution of the lags across all sessions is shown in figure 5.12. Every session had a negative lag, indicating that the reference timestamps are early relative to the

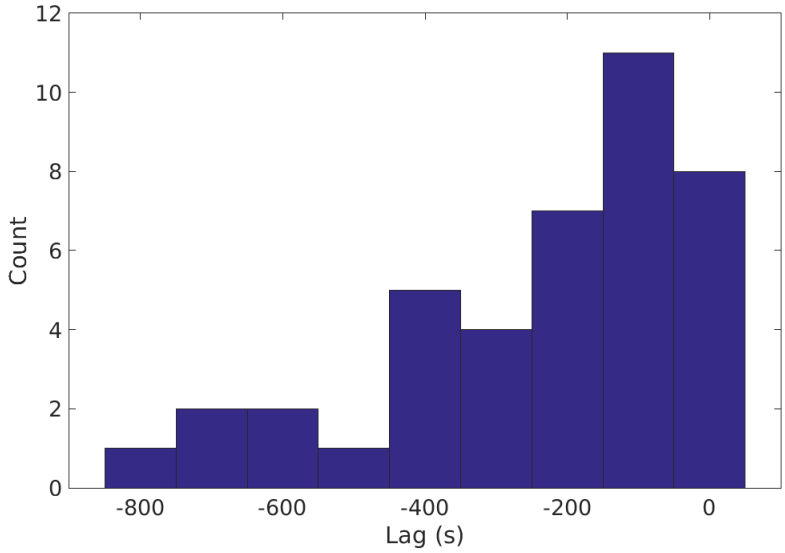


Figure 5.12: Distribution of the lag calculated between reference and camera data across all sessions of the Physiology study. A negative lag indicates that the reference timestamps are early relative to the camera timestamps.

camera timestamps. The lag had a mean value of −231.9 s and a standard deviation of 212.0 s. The biggest lag was −771.3 s, and the smallest was −8.5 s. The large order of magnitude of the lag suggests that the Stowood equipment's internal clock was faulty; the timestamps of the computer controlling the video camera should be considered reliable as the computer time was synchronised at the start of each recording session using standard Network Time Protocol.

5.7.2 Interpolation

The raw rPPG signal has a sampling rate of 16 Hz, determined by the frame rate of the video camera; by contrast, the reference ECG has a sampling rate of 256 Hz. To simplify the process of calculating rPTT between the reference ECG and the camera rPPG signals, each signal was interpolated onto a common set of timestamps.

The resolution of the rPTT depends on the sampling rate of the input signals. This is because the reference points required to calculate rPTT can only be located to the nearest sample. In order to achieve a resolution of 1 ms, the common timestamps were chosen

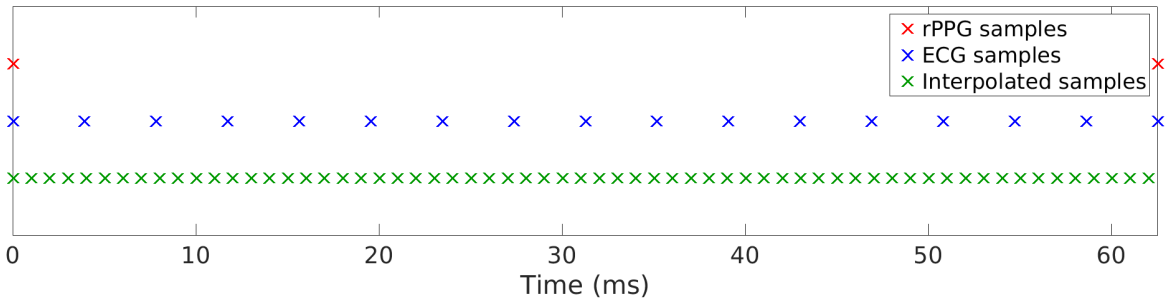


Figure 5.13: Illustration of the interpolation process in terms of timestamps. rPPG samples (red, originally at 16 Hz) and ECG samples (blue, originally at 256 Hz) are interpolated onto a set of common timestamps (green, at 1 kHz).

such that the signals would have a sampling rate of 1 kHz. Cubic spline interpolation was used so that the peak location could be estimated more accurately (see section 5.6.2).

The interpolation process resulted in two signals, each with a sampling rate of 1 kHz and a set of common timestamps. An illustration of the interpolation process is shown in figure 5.13.

5.7.3 Calculation

The rPTT calculation sequence was used to estimate raw rPTT from the synchronised and interpolated ECG and rPPG signals. The sequence first required R-wave detection to be performed on the ECG signal, and peak/trough detection to be performed on the rPPG signal.

ECG R-peak detection

The R-peaks in the QRS complexes were detected using the algorithm developed by Pan and Tompkins (1985) [116]. This algorithm involves filtering, differentiating and squaring of the ECG signal, followed by adaptive thresholding to discriminate between QRS complexes and T-waves or artefacts (see appendix E for more detail).

rPPG peak/trough detection

Peak/trough detection was performed using a decaying threshold detector, which used a decaying threshold in order to adjust to sudden changes in amplitude. For extra reliability it was run both forwards and backwards on the signal to check for missed peaks and troughs, particularly during the initialisation period of the detector.

Listing 5.1 provides pseudocode for the peak/trough detection algorithm. Only the pseudocode for the peak detection portion is shown; the trough detection code is identical except for minor changes (e.g. reversed inequalities, `vMax` replaced with `vMin`). The parameters used were 250 samples for `DECAY_THRESHOLD` (corresponding to 250 ms for an rPPG waveform resampled to 1 kHz), 0.1 for `DELTA_FRACTION` and 0.5 for `DECAY_FACTOR`.

[htb]

```
// Initialisation period.
vMax =: max(initialisationWindow) // Value of local maximum.
vMin =: min(initialisationWindow) // Value of local minimum.
delta =: (vMax - vMin) * DELTA_FRACTION
limit =: vMax - delta // Limit below which mode switches.
decayCounter =: 0 // Decay delta when this reaches a threshold.

for value in inputSignal
  if MODE == PEAK
    // In PEAK search mode.
    if value > limit
      // Above limit. Remain in PEAK search mode.
      if value > vMax
        // Just found a new local maximum.
        vMax =: value, iMax =: indexof(value) // Store maximum.
        decayCounter =:0 // Reset decay counter.
        limit =: vMax - delta // Update limit.
      else
        // Haven't found new local maximum.
        decayCounter++
        if decayCounter > DECAY_THRESHOLD
          // We have been searching for a peak for a while.
          delta =: delta * DECAY_FACTOR // Decay the delta value.
          decayCounter =: 0 // Reset decay counter.
          limit =: vMax - delta
        end
      end
    else
      // Below limit. Switch to TROUGH search mode.
      mode =: TROUGH
      nPeaks++, nTroughs++
      // Record most recent peak and trough.
      iPeaks(nPeaks) =: iMax, iTroughs(nTroughs) =: iMin
      delta =: (vMax - vMin) * DELTA_FRACTION // Update delta.
      decayCounter =: 0 // Reset decay counter.
      vMin =: value // Initialise current value as the minimum.
      limit =: vMin + delta // Establish new limit.
    end
  else
    // In TROUGH search mode.
    ...
  end
end
```

Listing 5.1: Pseudocode for peak/trough detection of an input signal.

Example outputs from the R-peak detection and peak/trough detection algorithms are shown in figure 5.14.

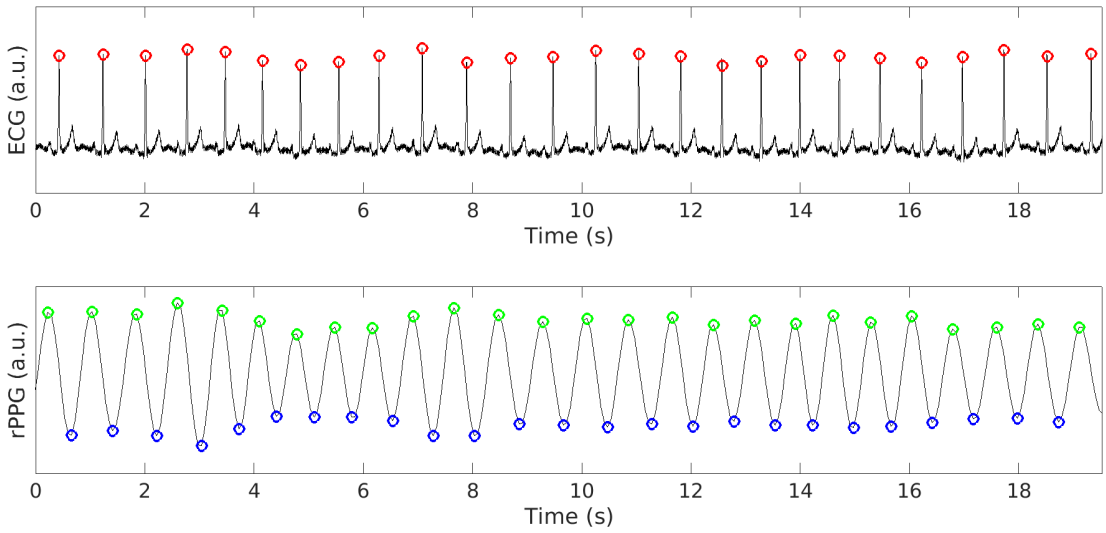
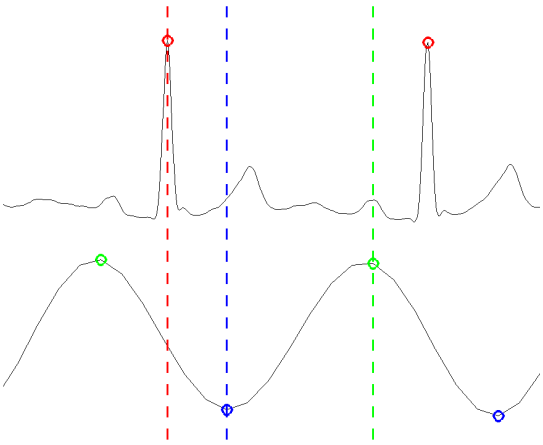


Figure 5.14: Example of R-peak detection on an ECG signal (top), and peak/trough detection on an rPPG signal (bottom). Data from subject HP3-33 of the Physiology study.

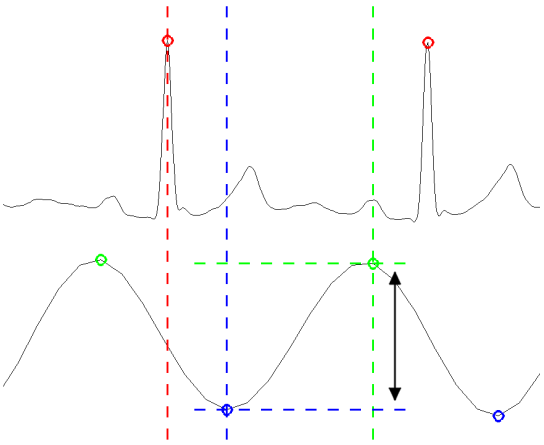
Calculation sequence

Once the R-peaks have been identified in the ECG, and once the peaks and troughs of the rPPG signal have been located, the rPTT can be calculated. For each R-peak, which acts as the proximal reference point for the rPTT, the calculation sequence is as follows:

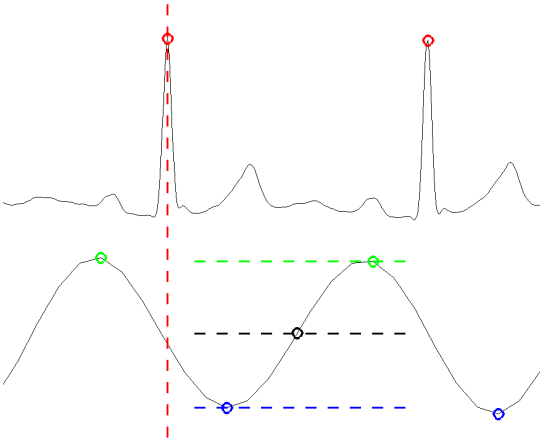
1. Find the first trough occurring in the rPPG after the R-peak
2. Find the first peak occurring after this trough
3. Calculate the amplitude value halfway between the peak and the trough amplitudes
4. Moving from the trough to the peak, find the sample from the rPPG time series whose value is closest to this amplitude
5. Set the index of this rPPG sample as the distal reference point for the rPTT
6. Find the number of samples separating the proximal reference point (i.e. the R-peak) and the distal reference point
7. Convert this number to a time delay by dividing by the sampling rate



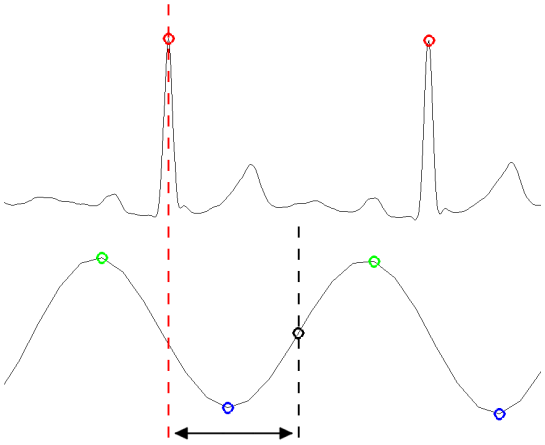
(a) ECG and rPPG with R-peak (red), trough (blue) and peak (green) marked.



(b) Trough-peak amplitude indicated (arrow).



(c) Mid-amplitude point identified (black circle).



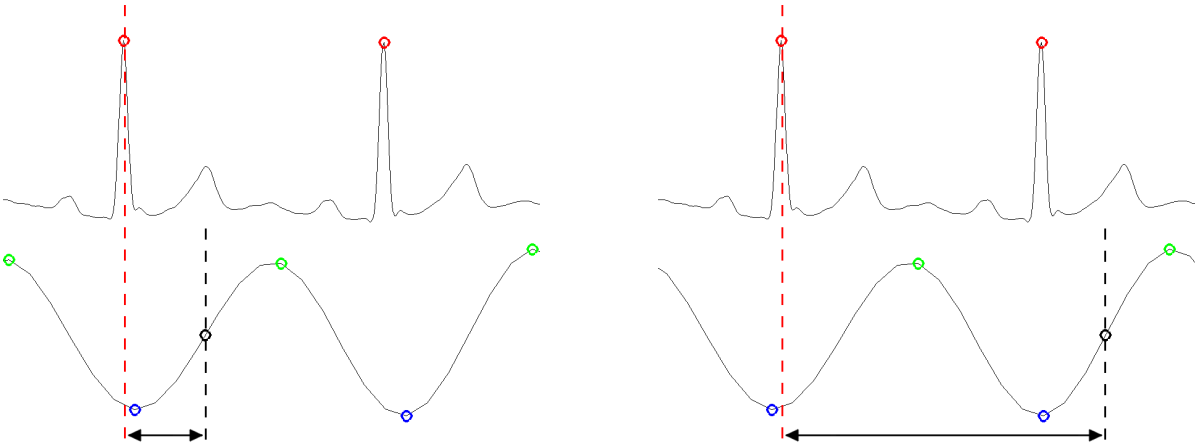
(d) Resulting rPTT shown (arrow).

Figure 5.15: Sequence for rPTT calculation between an ECG and rPPG signal.

This sequence is shown graphically in figure 5.15.

5.7.4 Processing

The raw rPTT estimates were produced on a beat-to-beat basis, and exhibited noise caused by spurious calculation sequences. The estimates were therefore processed by first unwrapping the phase of the estimates, and then windowing.



(a) Trough occurs just after R-peak. **(b)** Trough occurs just before R-peak.

Figure 5.16: Example of PTT phase wrapping: a small change in trough location causes a jump in rPTT. This can be corrected by subtracting the R-R interval from the second rPTT value.

Unwrapping

Physiologically, there must be a delay between the R-peak and the systolic rise in the rPPG signal because the systolic pressure takes time to propagate along the vasculature. However, this was not always the case in the synchronised ECG and rPPG signals because the synchronisation was only accurate to within one heartbeat (see section 5.7.1). In cases where the rPPG troughs occurred very close to the R-peaks, the rPTT calculation sequence was prone to phase wrapping of the rPTT values (since the calculation sequence looks for the first trough after each R-peak). A relative change of only one sample in rPPG trough location could therefore cause two rPTT estimates to differ by a period equal to the time between R-peaks (the R-R interval). An example of this is shown in figure 5.16.

This ambiguity in phase was corrected by performing phase unwrapping on the rPTT estimates. This unwrapping process relied upon the assumption that rPTT does not change rapidly, but rather over a timescale in the order of minutes. Any rPTT estimate differing from the median of the previous 120 (equivalent to approximately two minutes' worth of data) by more than the R-R interval was corrected in the appropriate direction

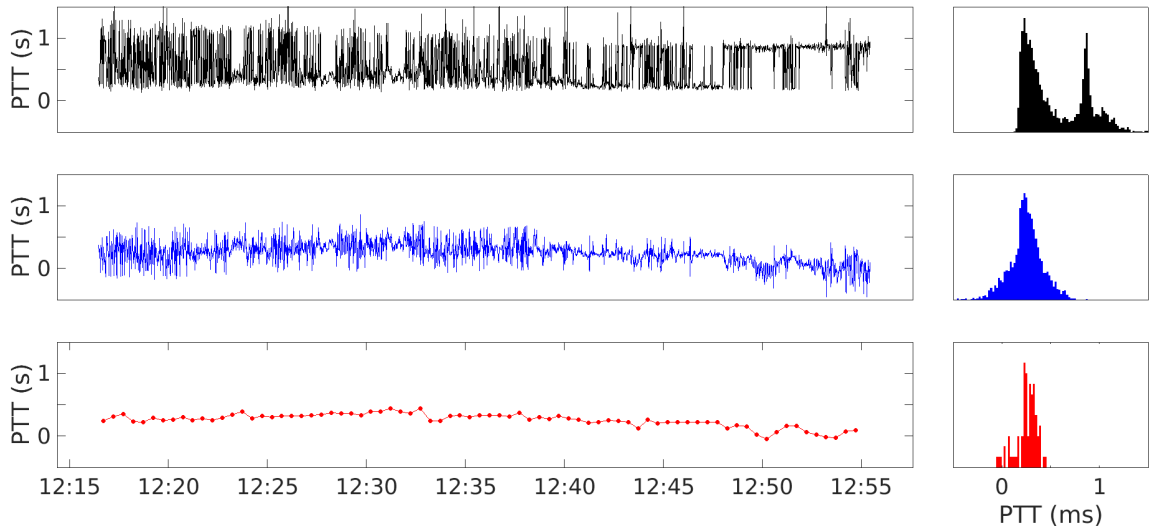


Figure 5.17: Steps of rPTT processing: raw (top), unwrapped (middle) and windowed (bottom). Note the bimodal histogram of raw rPTT values, which is reduced to one mode by unwrapping. Data from subject HP3-32 of the Physiology study.

by one R-R interval.

Windowing

In healthy volunteers (who remain seated throughout the recording session), rPTT is unlikely to change over timescales shorter than a minute. In order to smooth the time-series and remove noise, the rPTT signal is windowed into 30-second segments, and the median value of rPTT is found. This produces one value of rPTT per 30 seconds of recording. Windows in which more than half the reference cPTT measurements were invalid (not physiologically plausible) were discarded, with no value of rPTT being estimated for that window.

An example of rPTT processing is shown in figure 5.17.

5.7.5 Validation

The rPTT calculation method estimates PTT between ECG and rPPG signals, and as such is similar to the Stowood algorithm which is used to calculate the reference cPTT in

the Renal Unit study (see section 1.4.1). The Stowood algorithm is proprietary and not publicly available, but uses the same principle as the method described in this section (though conventional finger PPG is used instead of camera-derived rPPG).

The rPTT estimation algorithm was therefore validated against the Stowood algorithm using the data from the Renal Unit study. cPTT was calculated between the ECG and finger PPG signals (using the rPTT estimation algorithm) to produce estimates, and these were compared with Stowood's cPTT data as a reference. The reference cPTT calculated by Stowood has a low SNR, with frequent signal dropout. Therefore, physiologically implausible values (of more than 500 ms or less than 50 ms) were removed from the reference data, with the cPTT estimates treated in the same way. Paired cPTT values were produced by calculating the median of both the estimate and reference cPTT during 60-second windows. This length of window was chosen because PTT in healthy volunteers (who remain seated) is physiologically unlikely to change within a shorter timescale. Windows with less than 45 seconds of acceptable PTT values (in either the estimate or reference data) were ignored. Data from the first or last 5 minutes of a haemodialysis session were also ignored; these time periods typically showed artefacts caused by setup or removal of nursing equipment.

Out of 381 haemodialysis sessions with reference cPTT data, 5 were excluded due to insufficient reference data (less than five minutes' worth); 68 were excluded because more than half of the reference cPTT values were physiologically implausible; and 12 were excluded due to excessively noisy reference data (cPTT with a standard deviation of more than 100 ms after removal of physiologically implausible data). For the remaining 67,902 paired measurements from 296 sessions, the RMSE between reference and estimate was 12.5 ms and the MAE was 5.6 ms. The mean and standard deviation of the error was 0.2 ± 12.5 ms. The paired measurements are shown along with a Bland-Altman plot in figure 5.18.

These errors provide a lower bound of the possible accuracy of camera-derived PTT

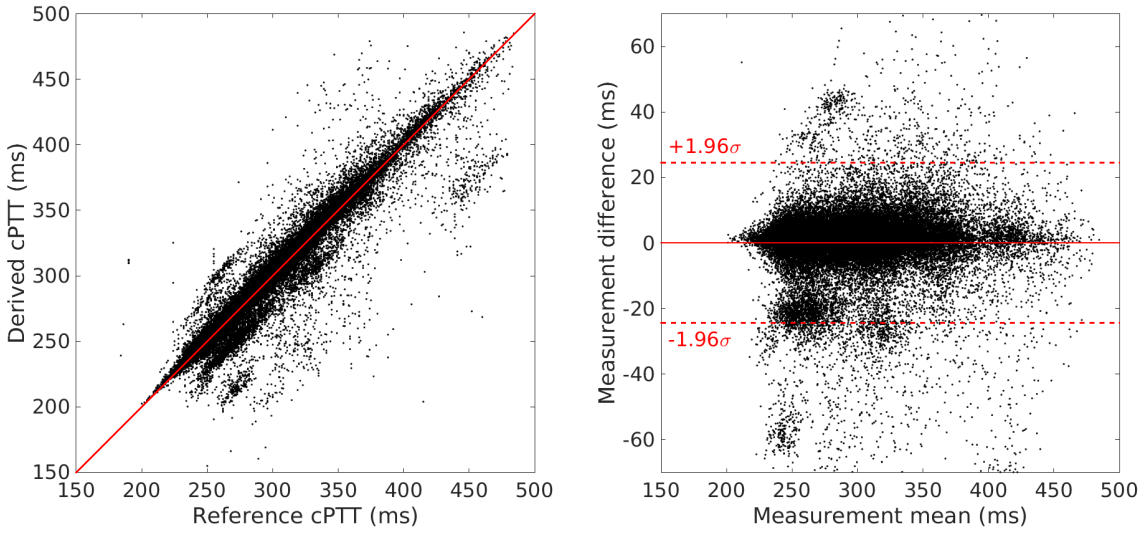


Figure 5.18: Paired measurements (left) and Bland-Altman plot (right) of reference and estimated cPTT using data from the Renal Unit study.

data when compared to the reference cPTT data reported by the Stowood algorithm. The rPTT calculation algorithm can be expected to result in a minimum RMSE of 12.5 ms, since that error represents the difference between the two algorithms when run over the same data set.

5.7.6 Frequency content

In normal circumstances, BP falls slightly during inhalation and rises during exhalation[3]. During inhalation, the pressure within the pleural cavity decreases, causing air to enter the lungs. This lowering of pressure is transmitted to the aorta, reducing the blood pressure [72, 46]. Figure 5.19 shows an example of this effect in a continuous BP waveform obtained using the Finometer device (see section 2.3.3).

As discussed in Chapter 2 and Chapter 3, there is usually an inverse relationship between PTT and BP. It would therefore be reasonable to expect to see a breathing component in a PTT signal, where the PTT increases slightly during inhalation and

[3]When this change in pressure is large (greater than 10 mmHg), it is referred to as pulsus paradoxus. This is considered pathological, and may be a sign of an underlying problem.

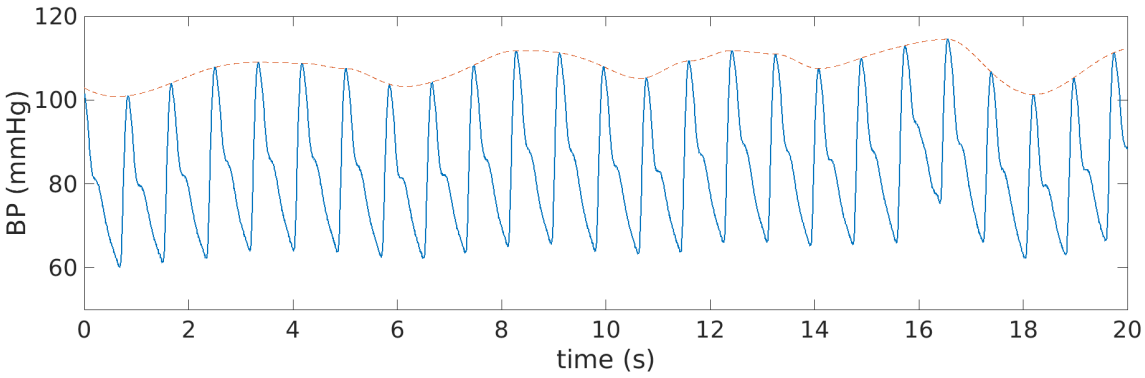


Figure 5.19: BP waveform (blue) showing a breathing component (red dashed line). Data from subject DF05 of the Renal Unit study.

decreases again during exhalation.

To investigate whether this is the case, the frequency content of the rPTT signal can be investigated in the same way that the content of the rPPG signals were investigated in section 4.3. Spectrograms were created (see section 4.3.1) for the camera-derived rPTT signal and reference cPTT signals estimated from subject HP3-33 of the Physiology study, using windows of 30 s overlapping by 25 s. ROIs were selected from the same sites of skin shown in figure 4.9.

Figure 5.20 shows spectrograms for each rPTT signal over a range of possible breathing frequencies. The colour scale is consistent between sites. The equivalent spectrogram for the reference cPTT signal is shown in figure 5.21. The reference BR, found using the inductance pneumogram (IP) signal, is superimposed on each spectrogram in red.

The spectrogram of the reference cPTT shows a very clear breathing component that matches the BR measured by the IP signal; the subject's cPTT is strongly modulated by their breathing. In the case of the camera-derived rPTT, however, the breathing component is not so clear. There is perhaps a subtle breathing component to the rPTT signal estimated at the forehead and cheek, but the mouth- and neck-derived rPTT signals are too dominated by broad-spectrum noise to see a breathing component.

It is possible that there is a physiological basis for the lack of breathing component in the camera-derived rPPT signal. When measuring cPTT, the blood travels from the

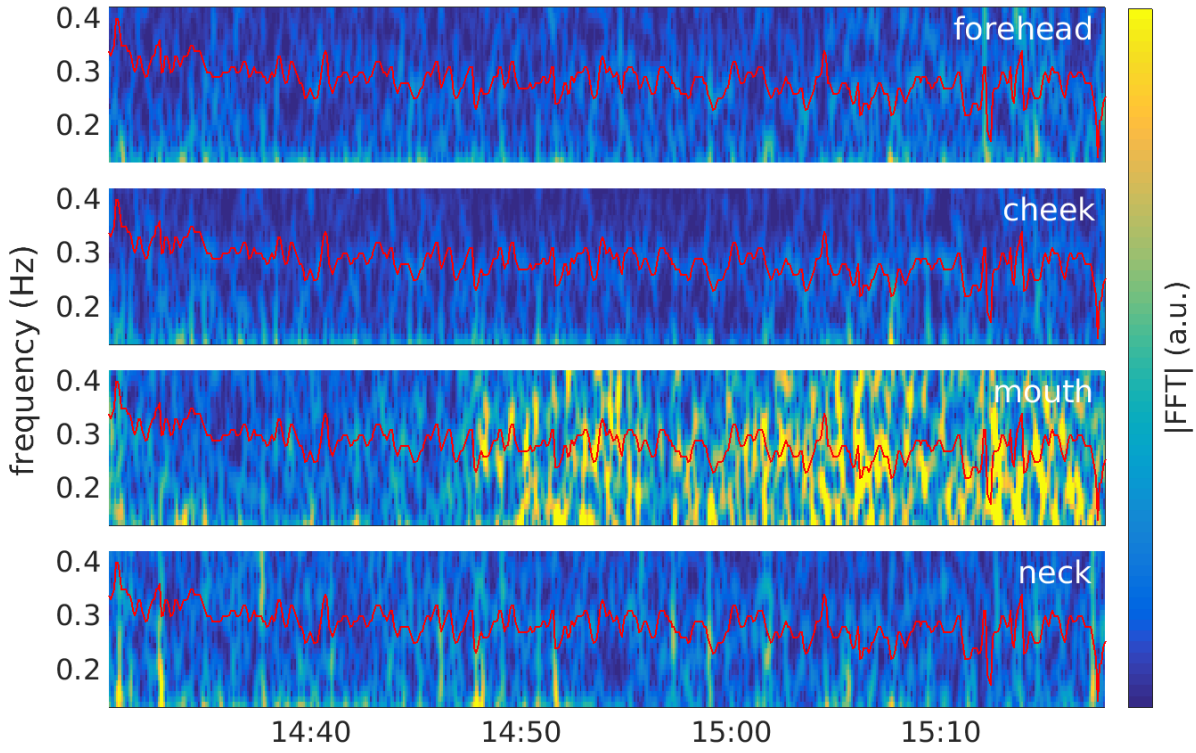


Figure 5.20: Breathing-frequency rPTT spectrograms from four different ROIs (locations in white text). The reference BR from the IP is superimposed in red.

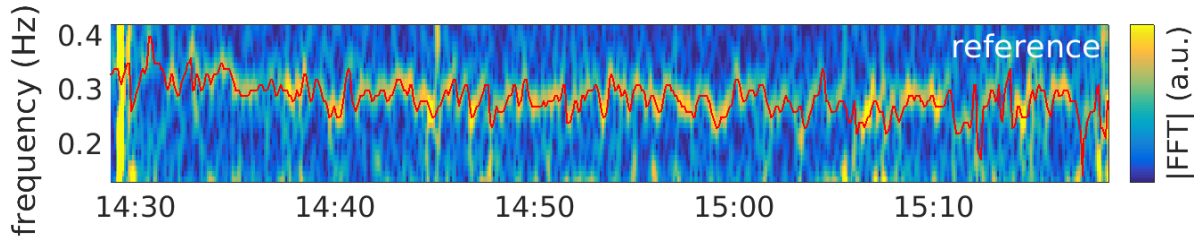


Figure 5.21: Breathing-frequency spectrogram for the reference cPTT signal. The reference BR from the IP is superimposed in red.

heart to the finger almost entirely via major blood vessels, and the route is very direct. In contrast, the skin of the face is perfused via a complex vasculature that involves more branching and smaller, superficial blood vessels. With this difference in mind, the small variations in BP (and, correspondingly, PTT) may become diminished, leading to a loss of breathing frequencies in the power spectrum of rPTT measured at the face.

5.8 dPTT estimation

The beginning of this chapter describes the calculation of differential pulse transit time (dPTT) using rPPG signals derived from two ROIs in the video: one proximal (i.e. closer to the heart) and one distal (i.e. further from the heart).

The method for estimating dPTT consisted of three stages: resampling (section 5.8.1), calculation of raw dPTT (section 5.8.2), and finally processing of the dPTT data (section 5.8.3).

5.8.1 Resampling

Since the two rPPG signals were derived from the same video recording, they are synchronous in time; calculation of dPTT therefore does not require the synchronisation step that rPTT calculation does. However, in order to improve the precision of peak detection (which can only locate peaks to the nearest sample), the two signals were resampled from their original frequency of 16 Hz to a new frequency of 1 kHz. This increased the resolution of the peak location algorithm (and therefore the eventual dPTT resolution) from 62.5 ms to 1 ms. The resampling was performed by cubic spline interpolation using the method described in section 5.6.2. It is worth stating again that information about the peak location is not being created by the resampling process; rather, assumptions about the underlying rPPG waveform are used to estimate the peak location more accurately.

5.8.2 Calculation sequence

The dPTT calculation sequence is simpler than that for rPTT because two rPPG signals are being compared; no ECG is involved. This means that any reference point in the rPPG signals can be used for dPTT calculation, as long as the same reference point is used in both signals. The need for location of the trough/peak midpoint is therefore eliminated, and the reference point problem becomes a simple case of peak detection. The beat-to-beat dPTT can then be calculated using the time difference between corresponding peaks of the rPPG signals.

The peak detection algorithm in section 5.7.3 was used to locate the rPPG peaks. For each peak p_i in the proximal rPPG signal, dPTT was calculated as follows:

1. The closest peak in the distal rPPG signal to p_i was identified as the corresponding peak q_j
2. The number of samples separating p_i and q_j was found
3. The sample difference was converted to a time delay by dividing by the sampling rate

This method assumes that the peak detector identified all peaks in both signals correctly. The process of identifying the index j of the closest peak to p_i can be expressed mathematically as

$$j = \arg\min_{n}(|p_i - q_n|) \tag{5.3}$$

and the dPTT value (in seconds) can then be calculated as

$$dPTT_i = \frac{p_i - q_j}{f_s} \tag{5.4}$$

where f_s is the sample rate in Hz, and $dPTT_i$ is the ith value of the dPTT time-series.

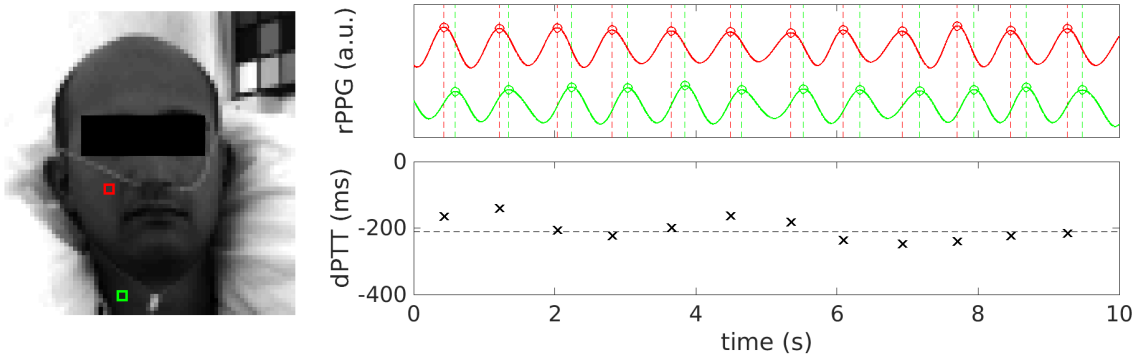


Figure 5.22: Example of calculating dPTT using rPPG signals from subject HP3-33 of the Physiology study. Top: filtered proximal and distal rPPG signals. Bottom: beat-by-beat interval between corresponding peaks (black crosses) and median (black dashed line).

.

It is not always obvious which of the two rPPG signals is recorded at a proximal site, and which is recorded at a distal site. In these cases, the sign of the dPTT value indicates at which site the pulse arrives first; the method does not rely on the correct signal being identified as proximal or distal.

An example of this method of calculating dPTT is shown in figure 5.22. In this example, the pulse apparently arrives at the cheek before the neck; this may not necessarily be the case, however, as there is a phase ambiguity in the dPTT values (see section 5.8.3). Note also the possible breathing-synchronous changes in the beat-by-beat intervals, which was explored in section 5.7.6.

5.8.3 Processing

The dPTT calculation sequence was performed on a beat-by-beat basis, as it was for rPTT. The resulting dPTT time-series exhibit noise from spurious sequences, and are therefore processed to address this problem.

Phase ambiguity

The dPTT calculation method relies on finding the time difference between corresponding rPPG peaks at each site. However, given the periodicity of the rPPG waveforms it is impossible to distinguish positive differences from negative ones without relying on assumptions about the direction of blood flow.

For example, in the case where a subject has a steady HR of 60 bpm, a dPTT of +100 ms from neck to cheek is indistinguishable from a dPTT of -900 ms because the peaks would have the same relative difference. This ambiguity in phase can be addressed by including assumptions about the magnitude of the dPTT (in the example given, a dPTT of +100 ms is more likely because the neck and cheek are relatively close), or about the direction of blood flow (if the pulse arrives at the neck before the cheek, then a dPTT of +100 ms must be the correct phase).

Making these assumptions can be dangerous, however, since the vasculature of the face is complex (see section 5.2); the proximity of two sites on the face may not guarantee a dPTT that is small in magnitude since they may be fed by different branches of the vasculature. In addition, while arterial blood flows from neck towards the cheek, venous blood flows in the opposite. The observed sign of the dPTT can therefore not be assumed carelessly; either phase may be correct.

Windowing

As was the case for rPTT, the dPTT values from healthy volunteers are unlikely to change over short timescales. In order to improve the SNR of the dPTT time-series, the values were windowed and the median value found over each window. This process was identical to the windowing method used in section 5.7.4 to process the dPTT.

5.9 Results

The methods described in this chapter were used to estimate rPTT and dPTT using video recordings from the Physiology study described in section 1.4.2. The results of the analysis are presented in this section.

In the figures of this section, "reference", "forehead", "cheek" and "neck" are abbreviated "ref", "fhd", "chk" and "nck" respectively.

5.9.1 rPTT estimation

Values of rPTT were estimated using the methods described in section 5.5, and windowed and paired with the appropriate reference values (section 5.7.4). The reference cPTT data were measured by Stowood's Black Shadow device (section 1.4.2).

Data were analysed from 41 sessions of the Physiology study. Two sessions were excluded as the reference and camera data could not be synchronised (see section 5.7.1). The change in rPTT and cPTT compared to the start of the session was also found (denoted ΔrPTT and ΔcPTT respectively).

At the global level, there was no correlation between camera-derived rPTT and reference cPTT, or between ΔrPTT and ΔcPTT. Correlation coefficients were effectively zero regardless of the site at which rPTT was measured. At the session level, however, some correlation was observed between ΔrPTT and ΔcPTT.

Figure 5.23 shows three sessions exhibiting positive correlation between reference and combined estimates, and figure 5.24 shows three examples of sessions exhibiting no such correlation.

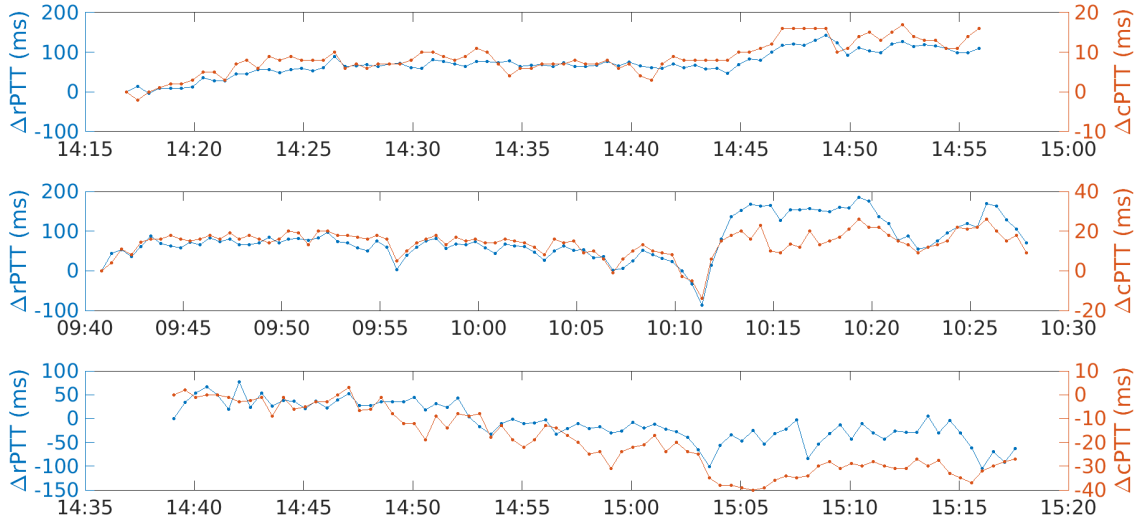


Figure 5.23: Examples of sessions with positive correlation between reference ΔcPTT and combined ΔrPTT estimates. Sessions are shown for subjects HP3-47 (top), HP3-37 (middle) and HP3-30 (bottom).

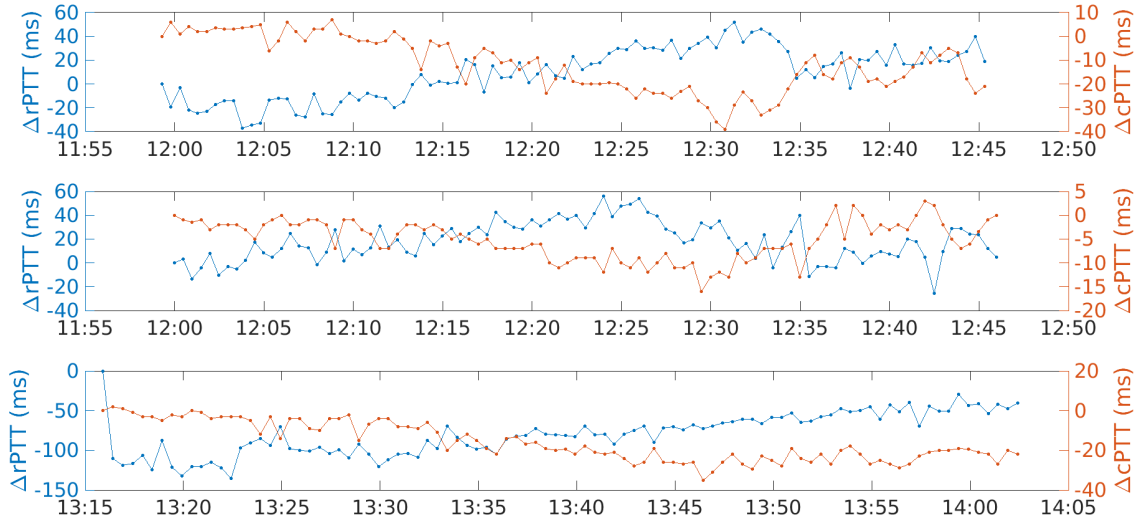


Figure 5.24: Examples of sessions with no positive correlation between reference ΔcPTT and combined ΔrPTT estimates. Sessions are shown for subjects HP3-24 (top), HP3-13 (middle) and HP3-22 (bottom).

5.9.2 dPTT estimation

Values of rPTT were estimated using the methods described in section 5.7, and windowed and paired with the appropriate reference values (section 5.8.3). Again, data were analysed from 41 sessions of the Physiology study with two sessions excluded as in section 5.9.1. The change in dPTT compared to the start of the session is denoted ΔdPTT.

Consistency with rPTT

To assess the consistency between dPTT and rPTT measurements, dPTT measured between two sites was compared to the difference between the rPTT estimates at each site. For example, cheek-to-forehead dPTT was compared to the difference between cheek rPTT and forehead rPTT; the two should be the same, as is suggested by the example in figure 5.25. These relationships are shown in figure 5.26, and again in figure 5.27 with the scale adjusted to ignore outliers, which are generally caused by a failure of the peak detection algorithm due to low SNR in the rPPG signals.

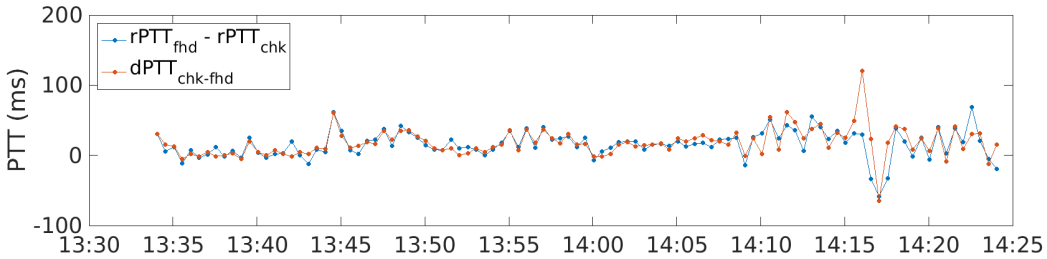


Figure 5.25: Example of the relationship between dPTT and the difference between rPTT measured at the equivalent sites. Data from subject HP3-07 of the Physiology study.

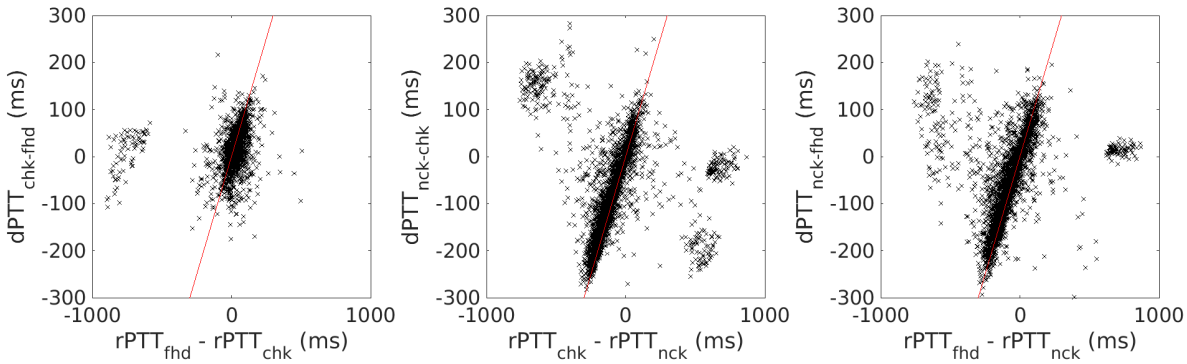


Figure 5.26: Global relationships between dPTT and the difference between rPTT measured at the equivalent sites, showing all points. Red lines show the expected $y = x$ relationship.

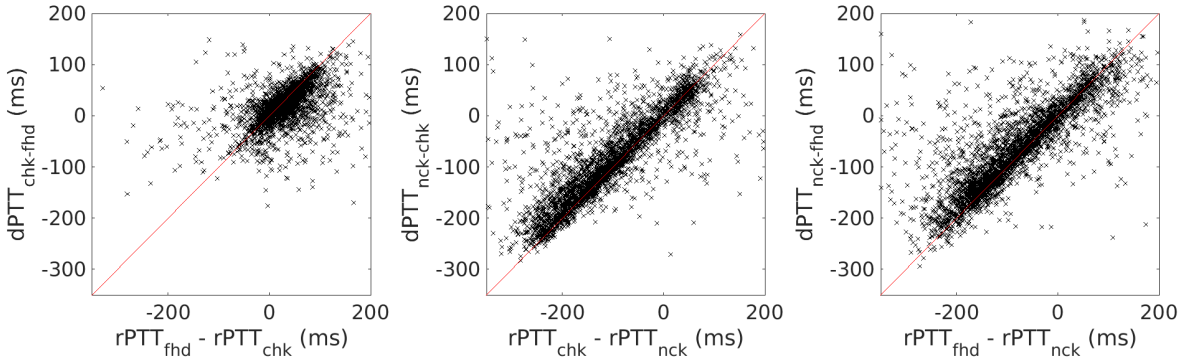


Figure 5.27: Global relationships between dPTT and the difference between rPTT measured at the equivalent sites, with scale adjusted to ignore outliers. Red lines show the expected $y = x$ relationship.

5.9.3 Local variation in PTT

To investigate local spatial variation, rPTT and dPTT was found for some subjects of the study using the cross formation of ROIs described in section 5.4. These cross formations were used to assess spatial variation at the neck and forehead. rPTT was estimated using the ECG and the rPPG derived at each ROI; dPTT was estimated by looking at the phase delay between the rPPG signal measured at the central ROI and each neighbouring ROI. An example of this is shown in figure 5.28.

Forehead area

For each of the sessions with camera-derived PTT data, rPTT and dPTT were estimated at each ROI in the cross formation centred on the subject's forehead. The median values over the whole session were then found. Five sessions were excluded as the forehead ROI was on the right hand side, and four were excluded as the subject's hair obscured the forehead. Box plots of the median dPTT values are shown in figure 5.29; each box plot shows the distribution of the median values (one from each session) at a particular ROI in the cross formation.

Neck area

The process was repeated using the cross formation centred over the subject's neck. Five sessions were excluded as the neck ROI was on the right hand side; two were excluded as the neck was not visible; and two were excluded due to excessive noise in the camera-derived PTT data. Box plots of the median values of dPTT are shown in figure 5.30.

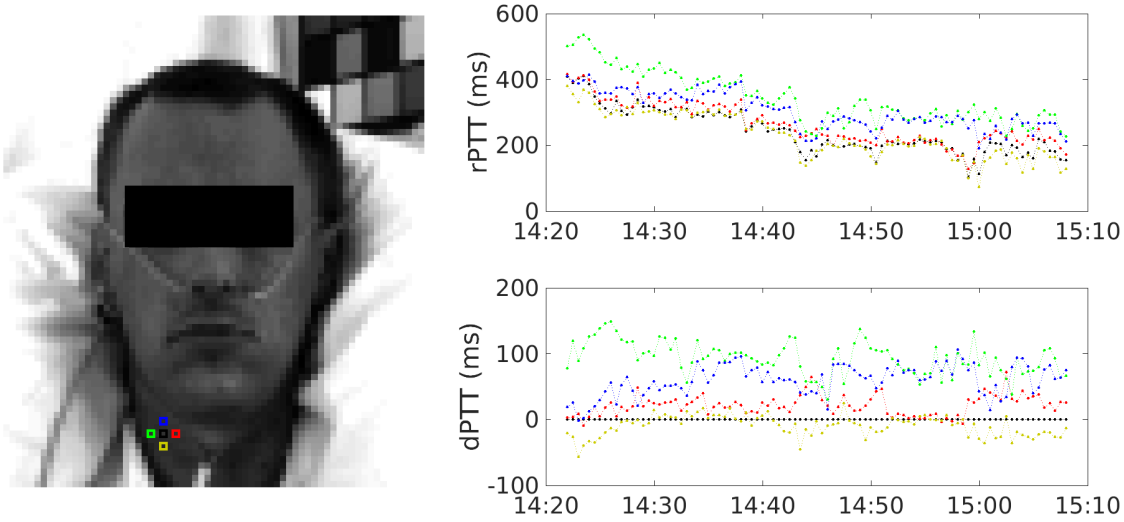


Figure 5.28: Example of camera-derived PTT estimated at multiple sites around a central location on the neck. Each PTT trace is coloured according to the ROI at which it was derived. Data from subject HP3-14 of the Physiology study.

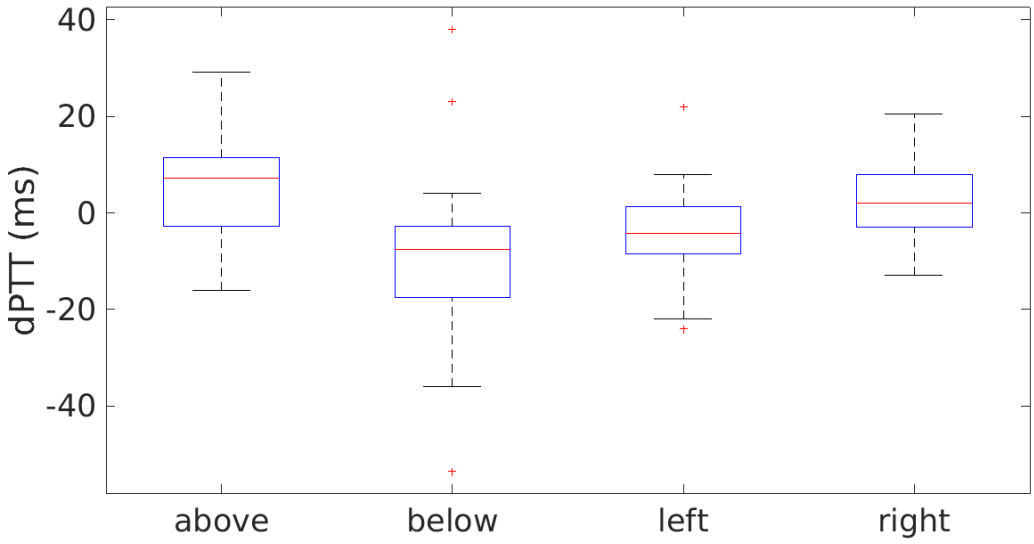


Figure 5.29: Box plots showing the dPTT measured at each ROI in the **forehead** cross formation. Each box plot describes the distribution of the median values (one from each session) at one of the ROIs, showing the minimum and maximum (black dotted line), interquartile range (blue box), median (red line) and outliers (red crosses) of the distribution from that location.

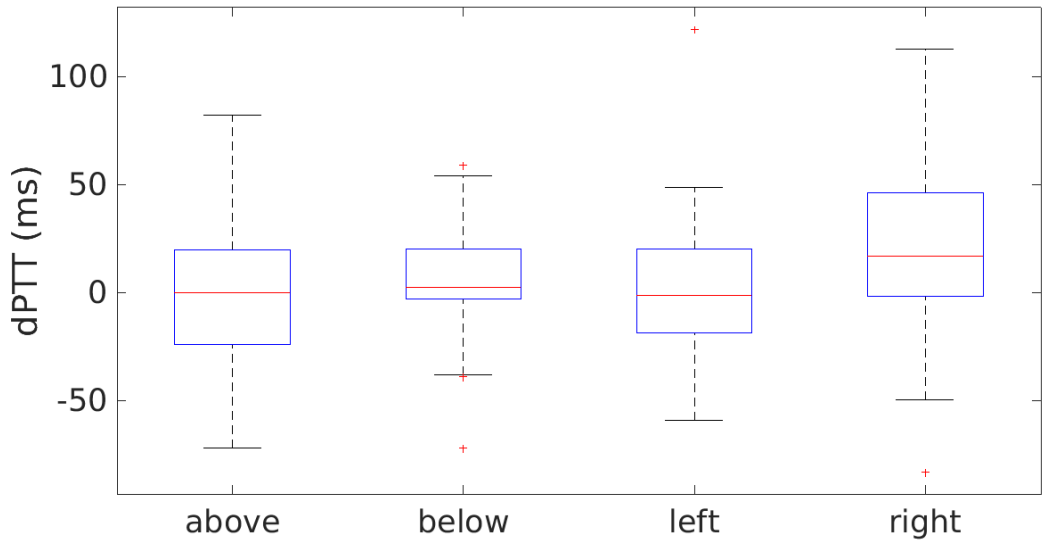


Figure 5.30: Box plots showing the dPTT measured at each ROI in the **neck** cross formation. Each box plot describes the distribution of the median values (one from each session) at one of the ROIs, showing the minimum and maximum (black dotted line), interquartile range (blue box), median (red line) and outliers (red crosses) of the distribution from that location.

5.10 Discussion

The results presented in section 5.9 suggest that it is possible to estimate PTT using a video camera. Longitudinal sections of rPTT and dPTT are generally consistent over time, indicating that the methods proposed in this chapter are reasonably robust. The extent to which these camera-derived PTT estimates are clinically useful, however, is not clear.

There is a high level of agreement between the dPTT measured between two sites compared to the difference between the rPTT estimates at each site. This agreement is exemplified in figure 5.25, and can be seen globally in figure 5.26. This agreement is further evidence of the robustness of both the rPTT and dPTT estimates.

On a global level, there was effectively no correlation between cPTT and rPTT or dPTT. This lack of global correlation at any of the sites suggests that neither rPTT or dPTT can be used as a robust surrogate for cPTT, as there is no consistent relationship between the variables across subjects of the study.

On a session-based level, however, some relationships were observed between reference and estimates. Strong positive correlation is particularly visible in the sessions shown in figure 5.23, in which ΔcPTT and ΔrPTT clearly follow the same trends. In the case of subjects HP3-37 and HP3-47, the same small variations are seen in both reference and estimate values. This suggests that, in the case of these two sessions, camera-derived PTT could have been used as a surrogate for cPTT for those subjects of the study.

Other sessions, however, exhibit no such positive correlation (figure 5.24). This could be partly due to the natural variation between individuals that was observed in Chapter 3, although the heterogeneity of that population probably results in a higher variation than should be seen in the healthy volunteers seen in this chapter. The large range in the degree of correlation could also be due to the difficulty in selecting consistent ROIs between subjects of the study. Many subjects had large areas of skin clearly visible to the camera, but this was not always the case. Sources of inconsistency in this area included:

- Hair, which frequently caused the forehead to be partially or fully obscured
- Clothing, which fully covered the neck of many subjects
- Pose of the subject, which affected the level of visibility of the neck
- Differences in facial structure, which resulted in more or less skin being visible to the camera

In addition, the level of similarity in vascular anatomy across subjects is unknown. The same section of skin from two subjects might have slight differences in the way they are supplied with blood; these differences might have an effect on the measured PTT because they are closer to or further from other parts of the vasculature.

In most subjects, the rPTT measured at the neck was the largest in magnitude, and the rPTT measured at the cheek was the smallest. This suggests that the pulse tends to arrive at the cheek before the forehead, and at the forehead before the neck. The

apparent lag between arrival at the cheek and neck is not intuitive, since the blood in the larger arterial vessels travels up the neck towards the top of the head; it could be explained, however, by the phase ambiguity which exists in the measurement of PTT (see sections 5.7.4 and 5.8.3). It is possible that in fact, the blood arrives at the neck before the rest of the face, and that the PTT at the cheek and forehead is measured with respect to the previous pulse, giving rise to an apparently smaller PTT. Alternatively, it is possible that the rPPG signals include components caused by venous return; blood flows in the facial veins from the head back towards the heart, and the pressure pulse in the veins would therefore arrive in the neck after the rest of the face. It is usually held, however, that the rPPG signal arises mainly from blood flow in the arteries and arterioles rather than the veins [90], but this has been disputed by some [64].

Looking at the longitudinal data in figures 5.23 and 5.24, the range of ΔrPTT was much higher than ΔcPTT. Typically, rPTT changed by more than 100 ms during a session, whereas cPTT usually changed by less than 50 ms. This could be because rPPG is produced by more superficial blood vessels than conventional finger PPG (see section 5.2), and camera-derived PTT therefore depends more on blood vessel tone.

Looking at the cross formation of ROIs gives some further information about the direction of pulse propagation in the face. Figure 5.29 shows a more negative dPTT below and to the left of the central point, and a more positive dPTT above and to the right of the central point. Recalling that the dPTT values are estimated between each surrounding point and the centre point, these box plots suggest that the pulse tends to propagate from left to right (lateral to medial, since the ROIs cover the left side of the forehead) and from below to above. These results seem to agree with the direction of blood flow suggested in figure 5.2, which shows arterial blood vessels travelling from the temple towards the centre and top of the forehead.

Figure 5.30 shows that in most subjects, the dPTT between the central and left (more lateral) ROI is less than the dPTT between the central and right (more medial) ROI,

suggesting that the pulse propagates in the medial direction, from the side of the neck towards the centre. The vertical propagation is less clear, however; the dPTT measured at the ROI above and below the central ROI has approximately the same mean value. This suggests that there is no clear difference in pulse arrival time between the two ROIs.

A major limitation of camera-derived PTT estimation is the adverse effect that subject movement would have on the estimates. A sudden movement of the subject's head would cause artefact in the rPPG signals derived from the face, and would also leave the ROIs covering a slightly different part of the skin. This could have an effect on the estimated PTT if the new site is supplied by a different branch of the arterial vasculature. However, this adverse effect of movement could be reduced with the use of ROI tracking techniques, which are not used in this thesis.

A limitation specific to dPTT estimates is the assumption that the pulse travels directly from one site to the other. This may not necessarily be the case, especially in areas with arterial branching; the two sites may instead be fed by different branches of the vasculature. dPTT should therefore be considered reliable only when the pulse travels directly from one site to the other. Alternatively, dPTT can be measured between sites that are very close to each other, since nearby sites are likely to be fed by the same branch of the vasculature. This is the approach taken in Chapter 6, in which local dPTT is spatially mapped.

Chapter 6

Spatial haemodynamic mapping

Chapter 5 introduced the calculation of PTT using a video camera, focusing on specific ROIs which were chosen manually. ROIs corresponding to the subject's forehead, cheek and neck were identified, and the spatial variation of camera-derived PTT was investigated in section 5.9.3 by considering small ROIs around a central point. There is, however, no reason to limit ourselves to several ROIs; PTT can be calculated using any site visible to the camera. In this chapter, many small ROIs are combined to form a grid which covers the subject's entire face. This introduces the possibility for near real-time spatial mapping of PTT, but also of other haemodynamic parameters. It is hypothesised that perfusion is one of the parameters that can be monitored in this way; since this would be useful in the context of critical care medicine for early detection of problems such as shock [85, 5, 19], the possibility of real-time measurement of perfusion using a completely non-invasive method would be a useful development.

Section 6.1 summarises the existing research that has attempted spatial mapping using video cameras. Section 6.2 describes the process of deriving local[1] rPPG waveforms, and section 6.3 introduces the technique of signal averaging, which is used to increase the SNR of the local signals. Section 6.4 goes on to explain the methods used to derive local

[1]Throughout this chapter, the term "local" will be used to refer to parameters measured at a specific site on a subject's skin. This is in contrast to "global" signals or parameters, which are measured over large areas (such as the whole face).

average rPPG waveforms, from which local haemodynamic parameters are then derived in section 6.5. Sections 6.6 and 6.7 then describe how these parameters are combined into spatial maps. The results of these mapping methods are presented in section 6.8, and these results are discussed in section 6.9.

6.1 Previous work

Since the field of non-contact vital sign monitoring first emerged, the focus has moved from simple vital sign estimation to more complex analysis of the information available in video data. Even in the early stages of the field's development, however, research suggested that haemodynamic parameters could be spatially mapped. Such *et al.* showed in two papers in 1996 [146] and 1997 [147] that the DC output from CCD sensors could be used to provide information about parameters such as venous refill time. The same research group showed in a study by Wu *et al.* (2000) [175] that rPPG could be derived at several locations on the hand to obtain spatial information about arterial blood pulsation.

In another early paper to touch upon the idea of spatial mapping, Wu (2003) [174] used near-infra-red light to derive rPPG signals at several sites on the skin, and proposed the idea of spatial mapping of perfusion levels by finding the amplitude of localised rPPG signals. Two other early studies on rPPG, by Wieringa *et al.* (2005) [169] and Verkruysse *et al.* (2008) [161], both created images showing a spatial map of relative rPPG waveform power. A PhD thesis completed in 2008 by Hülsbusch [54] also showed that perfusion could be measured using the amplitude of the local rPPG signal. This thesis included techniques to perform motion artefact reduction in the rPPG signals.

Starting in the 2010s, multiple studies attempted true spatial mapping of more advanced physiological parameters. Kamshilin *et al.* [63] showed in 2011 that by mapping the amplitude and phase of the local rPPG signal in the hands of subjects, changes in perfusion after scratching of the skin's surface could be detected. The authors used

monochromatic illumination and also recorded the reflectance signals using a monochrome video camera, avoiding the problems associated with the use of ambient light. This study was perhaps the first to attempt true mapping of multiple haemodynamic parameters over a large area (the entire hand), rather than simply measuring parameters at multiple distinct sites, or mapping over very limited regions.

A study by Amelard *et al.* (2016) [4] found evidence that some areas of the neck (in proximity to the carotid artery) produced an arterial pulse waveform, whereas more lateral areas (lying over the jugular vein) produced what appeared to be a venous pulse waveform. The authors pointed to a large phase difference between the rPPG signals in neighbouring regions (which were often in antiphase), and noted that these regions lay over the path of the major blood vessels in the neck, which were located via ultrasound.

Most recently, a study by Kamshilin *et al.* (2016) [65] mapped the local rPPG phase in the lower face. The authors recorded short videos (32 seconds) of 36 subjects with their chin resting on a support to keep their head still, and used a grid of small ROIs to estimate local rPTT. The focus of this study was on accurate rPTT measurement rather than spatial mapping, however; the authors claimed that certain ROIs produced noisy rPPG signals with "stochastically enhanced" dicrotic notches, and marked these ROIs as unreliable.

Table 6.1 lists studies in which spatial mapping of haemodynamic parameters has been attempted. None of these studies have attempted to validate their results, presumably because of the lack of a gold standard against which to compare.

Although the methods used in these studies are similar to the mapping techniques presented in this thesis, the authors all gathered their data under very controlled circumstances. Close-up short video recordings were made of the hands or faces of subjects with minimal movement, and the subjects were illuminated using controlled lighting. By contrast, our video recordings are much longer (which allows for the possibility of spatial mapping over time), and many include realistic levels of movement.

Study	Year	Site	Parameter
Such [146]	1996	leg	venous refill time
Such *et al.* [147]	1997	leg	venous refill time
Wu *et al.* [175]	2000	limbs	venous refill time and rPPG
Hülsbusch *et al.* [55]	2002	limbs	rPPG power
Wieringa *et al.* [169]	2005	wrist	rPPG power
Verkruysse *et al.* [161]	2008	face	rPPG power
Hülsbusch [54]	2008	various	rPPG amplitude
Kamshilin *et al.* [63]	2011	palm	rPPG amplitude/phase
Rubins *et al.* [131]	2011	palm	rPPG amplitude
Kamshilin *et al.* [66]	2013	palm	rPPG amplitude/phase
Teplov *et al.* [155]	2014	wrist	rPPG amplitude/phase
Nakano *et al.* [108]	2014	hand	venous compliance
Yang *et al.* [177]	2015	face	blood flow velocity
Amelard *et al.* [4]	2016	neck	correlation with finger PPG
Iakolev *et al.* [58]	2016	hand	rPPG amplitude/phase
Kamshilin *et al.* [65]	2016	face	rPPG phase

Table 6.1: Studies performing physiological mapping of a haemodynamic parameter.

In addition, the method of mapping used in this thesis involves signal averaging (see section 6.3), which is a novel approach that allows higher temporal and spatial resolution within the maps.

6.2 Derivation of local rPPG

In order to spatially map haemodynamic parameters, local information must be obtained from the video camera data. For each subject of the Physiology study (dataset described in section 1.4.2), a local rPPG signal was derived for each element in a grid covering the face. Since such a grid was already created in section 5.3, this step consisted only of concatenating the rPPG values across all frames into a time-series, with one local rPPG time-series created for each element in the grid.

Each element of the grid corresponds to a 12×12 pixel area of the original video, which covers a facial area of approximately 3 mm $\times$ 3 mm. As discussed in section 4.4.3, an ROI of this size is likely to produce a rPPG signal with low SNR. This low SNR causes problems with rPPG processing (peak detection performs poorly with a noisy signal, for example) and the local signal therefore has to be processed before it can be analysed further.

The local rPPG signals from each element in the grid were filtered using the band-pass FIR filter described in section 5.6.1, and each signal was then resampled to 1 kHz using cubic spline interpolation. This was done to improve the resolution of subsequent processing methods, as explained in section 5.6.2.

Note that a local rPPG signal will have a comparable amplitude to a global rPPG signal derived over the entire face. This is because the rPPG signals are derived by calculating the average pixel intensity (rather than the summed intensities) within the relevant ROI (see section 4.1.2).

6.3 Signal averaging

Signal averaging exploits the periodicity of an underlying signal in a noisy waveform to increase SNR [48]. Summing sections of the waveform serves to build up the coherent signal, while the uncorrelated noise is cancelled out. Signal averaging works on the assumption that there is no correlation between the noise and the underlying signal, and that the noise has zero mean and constant variance.

6.3.1 Principles

An observed waveform $\mathbf{p}$ can be considered to have a periodic signal component $\mathbf{q}$ and a noise component $\mathbf{r}$. The nth observation of the waveform can be written as

$$\mathbf{p_n} = \mathbf{q} + \mathbf{r_n} \tag{6.1}$$

where $\mathbf{q}$ is the underlying signal waveform, which is assumed to be identical for all observations due to its periodicity[2], and $\mathbf{r_n}$ is the noise, which is different for each observation.

The power SNR γ_1 of $\mathbf{p_n}$ can be expressed as

$$\gamma_1 = \frac{E[|\mathbf{q}|^2]}{E[|\mathbf{r_n}|^2} \tag{6.2}$$

where E is the expectation operator, defined as

$$E[x] = \frac{1}{N}\sum_{i=1}^{N} x_i \,. \tag{6.3}$$

Averaging N observations of $\mathbf{p_n}$ produces

$$\bar{\mathbf{p}} = \frac{1}{N}\sum_{n=1}^{N}(\mathbf{q} + \mathbf{r_n}) = \mathbf{q} + \frac{1}{N}\sum_{n=1}^{N}\mathbf{r_n} \tag{6.4}$$

since $\mathbf{q}$ is constant. The SNR γ_2 of the averaged signal then becomes

$$\gamma_2 = \frac{E[|\mathbf{q}|^2]}{E\left[\left|\frac{1}{N}\sum_{n=1}^{N}\mathbf{r_n}\right|^2\right]} = N^2\frac{E[|\mathbf{q}|^2]}{E\left[\left|\sum_{n=1}^{N}\mathbf{r_n}\right|^2\right]} \,. \tag{6.5}$$

Assuming that the noise observations are uncorrelated with each other, the denominator in equation 6.5 can be simplified to

$$E\left[\left|\sum_{n=1}^{N}\mathbf{r_n}\right|^2\right] = \sum_{n=1}^{N} E\left[|\mathbf{r_n}|^2\right] = NE\left[|\mathbf{r_n}|^2\right] \tag{6.6}$$

which can be substituted back into equation 6.5 to produce

[2] In a perfectly periodic signal, the same waveform sample (i.e. the sample observed at the same time relative to the start of the waveform) will have a constant value.

$$\gamma_2 = N\frac{E[|\mathbf{q}|^2]}{E\left[|\mathbf{r_n}|^2\right]}\,. \tag{6.7}$$

Equations 6.7 and 6.2 can be combined to show that

$$\gamma_2 = N\gamma_1 \tag{6.8}$$

which indicates that averaging N observations increases the power SNR by a factor of N (or, equivalently, the amplitude SNR by a factor of $\sqrt{N}$). This factor is reduced if the underlying signal $\mathbf{q}$ is not exactly periodic, or if the noise $\mathbf{r}$ has a periodic component or a degree of correlation with $\mathbf{q}$.

6.3.2 Reference points

In order for periodic sections of a signal to be summed, these sections must first be identified. Each section must start at the same point relative to the signal's period. For a constant-frequency signal, this is straightforward; the signal can be split into sections determined by the period, with each section being the same length and following directly on from the previous one.

In the case of a physiological signal such as the PPG or rPPG, the frequency is only approximately constant as it changes with the HR. Reference points in the signal must be identified and used to select sections, which can then be averaged. This reference point is usually derived from another independent signal of the same frequency; for example, the R-wave of the ECG could be used to select sections of the rPPG.

6.4 Local average waveforms

Signal averaging was performed on the local rPPG signals described in section 6.2 with the goal of producing average rPPG pulse waveforms for each grid element. Reference

points in the signal were chosen with respect to a global rPPG waveform, derived over the subject's entire face. This is a novel concept, since reference points are usually found using external signals such as the ECG; in this case, however, the global rPPG signal is largely independent of the local rPPG signal and can therefore be used to select reference points. The global signal was filtered and resampled in the same way as the local signals. For each element in the grid, the process was as follows:

1. A 30 s window was selected from both the global and local rPPG signals, since the subjects' haemodynamic parameters are unlikely to change over such timescales (given that they are healthy and sitting still)

2. Signal averaging was performed on the windowed signals to construct global and local rPPG waveforms:

   (a) The peak detector described in section 5.7.3 was used to choose reference points in the global signal

   (b) A section of the global and local waveforms surrounding each reference point (the peak of the global waveform) was selected. Each section had a width equal to the heart beat interval, estimated by taking the median difference between peaks in the global waveform.

   (c) The sections were summed to produce an average waveform for both the global and local signals

3. The window was shifted forwards by 5 s, and the process repeated from step 1.

This produced one local rPPG waveform for each element of the grid, as well as one global waveform, for each 30-second window. An example of this, where the local rPPG signal is derived from part of the forehead, is shown in figure 6.1. Figure 6.2 shows an equivalent example, but where the ECG signal was used to select reference points. The

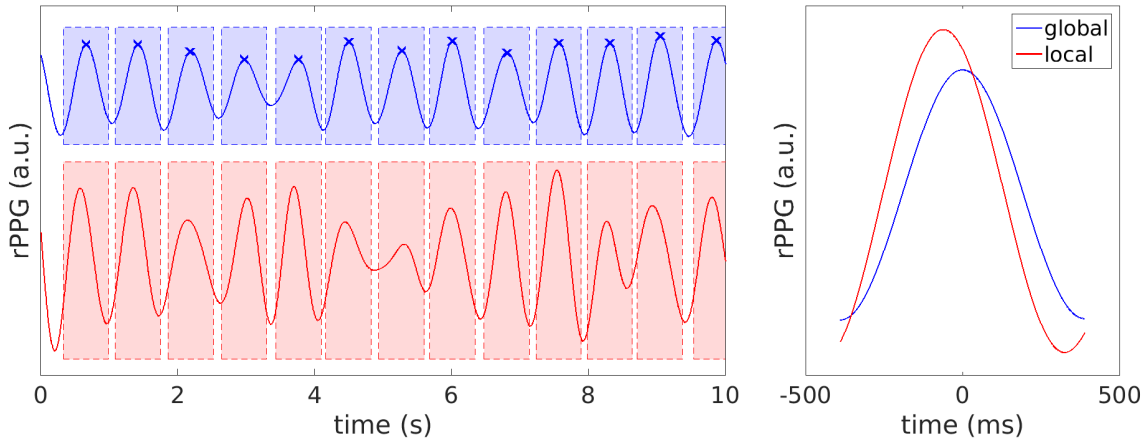


Figure 6.1: Example of signal averaging using the global rPPG waveform to obtain reference points. Left: global rPPG signal (blue) and local PPG signal derived at the forehead (red), with global reference points at the peaks (blue crosses) and sections of the waveform around each reference point (boxes). Right: The averaged waveforms, obtained by summing the raw signal within the boxes.

averaged local waveform has a different phase to the local waveform in figure 6.1, but is otherwise very similar in shape.

In figure 6.1, the local rPPG signal has a phase lead relative to the global reference, and the amplitude is larger than the global reference. This suggests that, in this specific example, the pulse arrives early at the forehead compared to the average over the face, and that the pulse is stronger compared to the average over the face.

Figure 6.3 shows an example of the local waveforms for each element across a subject's entire face. For display purposes, the waveforms are scaled such that their amplitudes are equal.

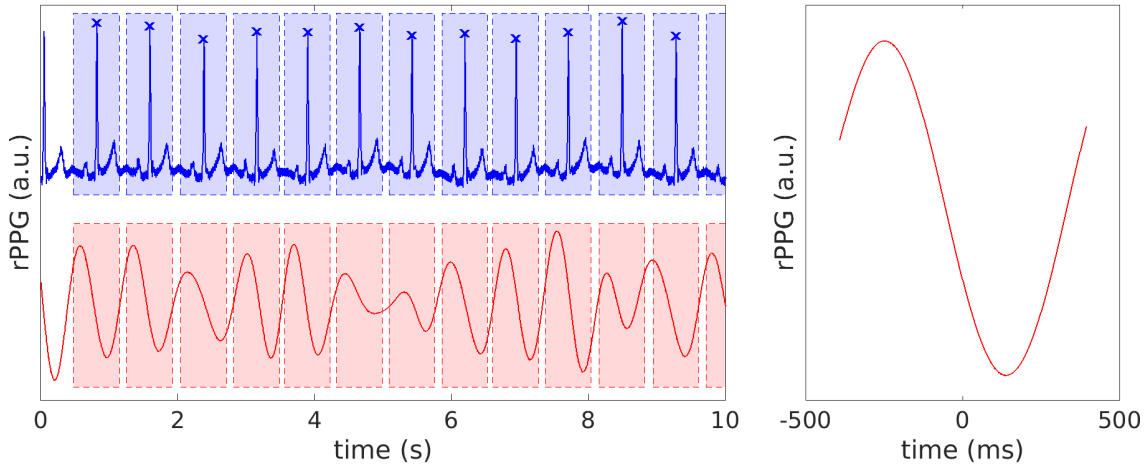


Figure 6.2: Example of signal averaging using the ECG to obtain reference points. Left: ECG signal (blue) and local PPG signal derived at the forehead (red), with ECG reference points at the R-wave of the QRS complexes (blue crosses) and sections of the waveform around each reference point (boxes). Right: The average local waveform, obtained by summing the raw signal within the boxes.

Figure 6.3: Local rPPG waveforms over a subject's entire face. Blue: global reference waveform. Red: local waveform.

6.5 Haemodynamic parameters

The local waveforms can be compared with the global waveform to estimate certain parameters. For example, the amplitude of the waveform can be used to estimate local perfusion levels, and the relative phase of the waveform can be used to estimate the direction of local blood flow.

6.5.1 Perfusion

Areas of skin that are well perfused have a good blood supply, and are likely to produce a local rPPG waveform with a large amplitude. Similarly, poorly perfused areas are likely to produce a local rPPG waveform with a smaller amplitude. By assessing the rPPG amplitude of each element in the grid, perfusion levels can be compared spatially as well as temporally.

The amplitude of each local rPPG waveform is found by subtracting the minimum value from the maximum. The amplitude is found independently for each averaged waveform (and therefore corresponds to 30 seconds of video).

6.5.2 Blood flow

By comparing the relative phase difference between the global waveform and the local waveform, the relative time of arrival of the pulse can be determined (just as it is when calculating dPTT in section 5.8). If the average time of arrival of the pulse is known at each location in the grid, the direction of local blood flow can be determined.

The relative phase of each local rPPG waveform is found using cross-correlation, which is described in section 5.7.1. The phase was found by calculating the cross-correlation function between the local and global waveforms (equation 5.1), and determining the location of the peak in this function. The corresponding lag was then converted to a time difference by dividing by the sampling frequency of 1 kHz.

This method was preferred to simple comparison of the peaks in the signal averaged waveforms because it was more robust to noisy signals, which often resulted in unexpected waveform shapes.

Optical flow

Optical flow can be used to assess relative motion in a series of images by determining the apparent velocity of objects. It could therefore be used as a tool to visualise the apparent velocity of the rPPG pulse as it moves across the face.

Figure 6.4 shows an example of optical flow being used in this way. Using data from subject HP3-33 of the Physiology study, the local signal-averaged rPPG waveforms from one 30 second window were combined into a series of images. Each image corresponds to a point in time (i.e. a point along the x-axis of the average waveforms in figure 6.1), and the intensity at a location corresponds to the local amplitude at that time. MATLAB's Computer Vision System Toolbox was then used to estimate the horizontal and vertical components of velocity in each image, found using the Horn-Schunck method [51]. The figure shows the mean estimated velocity over the duration of the pulse. An animation showing the optical flow over the course of the pulse can be found at https://www.dropbox.com/s/x39s8ck69ivkh9e.

The figure suggests that the pulse propagates upwards and laterally from the neck, but more uniformly over the skin of the face.

Figure 6.4: Example of optical flow measured over the course of a rPPG pulse.

6.6 Spatial mapping

Because each element in the grid has a value associated with it, spatial maps can be created for haemodynamic parameters such as local perfusion and blood flow.

6.6.1 Creation of maps

The maps are created by assigning each element a colour based on the value of a local parameter. The value of the parameter determines the intensity of the colour at that location. An example of a map of local amplitude is shown in figure 6.5. The map indicates that blood is pulsing most strongly in the middle of the forehead, which implies that the forehead is the best-perfused area of skin that is visible to the video camera. These results agree with the spectrograms in section 4.3.3, which indicated that the forehead and cheek rPPG signals contained a stronger cardiac component than those derived at the mouth or neck, and can be coupled with figure 6.4 to suggest that blood flow in the forehead is strong, but uniform (with no obvious direction of flow in the optical flow image).

6.6.2 Spatial filtering

Because each pixel of a map corresponds to a small area of skin, the SNR is relatively low even after signal averaging. It is unlikely that neighbouring pixels will have a large difference in local amplitude or phase, and spatial filtering of the map is therefore performed in order to smooth large discontinuities between adjacent pixels. The maps were filtered using a 2-dimensional Gaussian filter with size 3×3 pixels ($\sigma = 0.5$).

6.6.3 Skin masking

Areas of the image that do not contain skin have no underlying rPPG waveform, but may still contain cardiac-synchronous information; a subject's clothing, for example, may

exhibit periodic movement caused by the pulsation of blood (see section 4.1.1). The local waveform produced by signal averaging is usually physiologically meaningless in such areas, but it still has a measurable amplitude (usually small due to the lack of coherence) and phase, neither of which are likely to be physiologically meaningful.

A method of skin detection is therefore required to determine the relevant regions of a given image. Once areas of skin are identified, a binary mask can be produced and then applied to the maps in order to present only physiologically meaningful information.

There are many different approaches to skin detection, which is a field of research in itself. The skin detection algorithm used in this thesis has been used previously within our research group, and is described fully by Guazzi (2016) [43]. The algorithm involves the creation of a Gaussian mixture model, which is a probabilistic model created by assuming that observed data points are generated from a mixture of Gaussian distributions. By determining the most likely parameters of these distributions, the model forms a probability density function, and can be used for classification.

The steps of the algorithm are as follows:

1. The bounding box of the subject's face is found using the Viola-Jones frontal face detector [163]

2. A bimodal Gaussian mixture model in colour space is fitted to the pixel values inside the bounding box

3. The Gaussian in the model describing the largest number of pixels is taken as the likelihood of a pixel containing skin, and the other Gaussian is taken as the likelihood of a pixel containing non-skin

4. Each pixel is classified as either skin or non-skin depending on which likelihood function is greater for that pixel

5. A binary skin mask is created from the pixel classifications

6. The mask is morphologically dilated (the neighbours of any pixel classified as skin are also classified as skin)

7. All contiguous skin regions except the largest are removed (classified as non-skin)

8. The remaining mask is multiplied by the original skin mask to reverse the dilation

An example of a skin mask can be seen in figure 6.5.

Figure 6.5: Example of mapping local rPPG amplitude in subject HP3-33 of the Physiology study. Top left: processed frame showing subject's position. Top right: raw amplitude map. Bottom left: mask showing areas identified as skin. Bottom right: processed amplitude map (after filtering and masking).

6.7 Spatio-temporal mapping

Since the method described in section 6.4 produced one set of local waveforms per window, a map can be produced corresponding to each window. By combining successive maps into a video, this allows for physiological mapping across both space and time. Spatio-temporal mapping is useful because it allows changes in perfusion and blood flow to be monitored over time.

In addition, comparing successive maps allows the robustness of the method to be assessed. If the maps contain a large amount of incoherent noise, successive maps are unlikely to appear similar. If, however, the SNR in the maps is relatively high, then successive maps are likely to look similar.

In order to make successive maps visually comparable, the colour axis was scaled to global extremes. This scale was determined by plotting a histogram of local rPPG parameters from all of the processed maps (i.e. after filtering and masking) from a subject's recording session, and selecting a set of percentile values as the minimum and maximum colour (see figure 6.6). Values outside these percentiles were saturated to the nearest colour. An example of this is shown in figure 6.7, which shows the original maps exhibiting a scale that is not robust to outliers (such as the bright pixels near the subject's left eye at 14:34:20), and the same maps scaled using the 1st–99th and 5th–95th percentiles.

The 1st and 99th percentiles were chosen as the scaling values because they resulted in consistent maps (unlike the raw values) without excessive saturation at the extremes (unlike the maps scaled to the 5th and 99th percentile values).

Combining maps across time also allows for the possibility of temporal filtering. The local amplitude or phase of the local rPPG signal in each element of the grid is unlikely to change rapidly over time, and so changes to an element's local amplitude or phase could be filtered in time to smooth sudden changes. Temporal filtering was not considered necessary in this chapter since the SNR in the maps was acceptable without it.

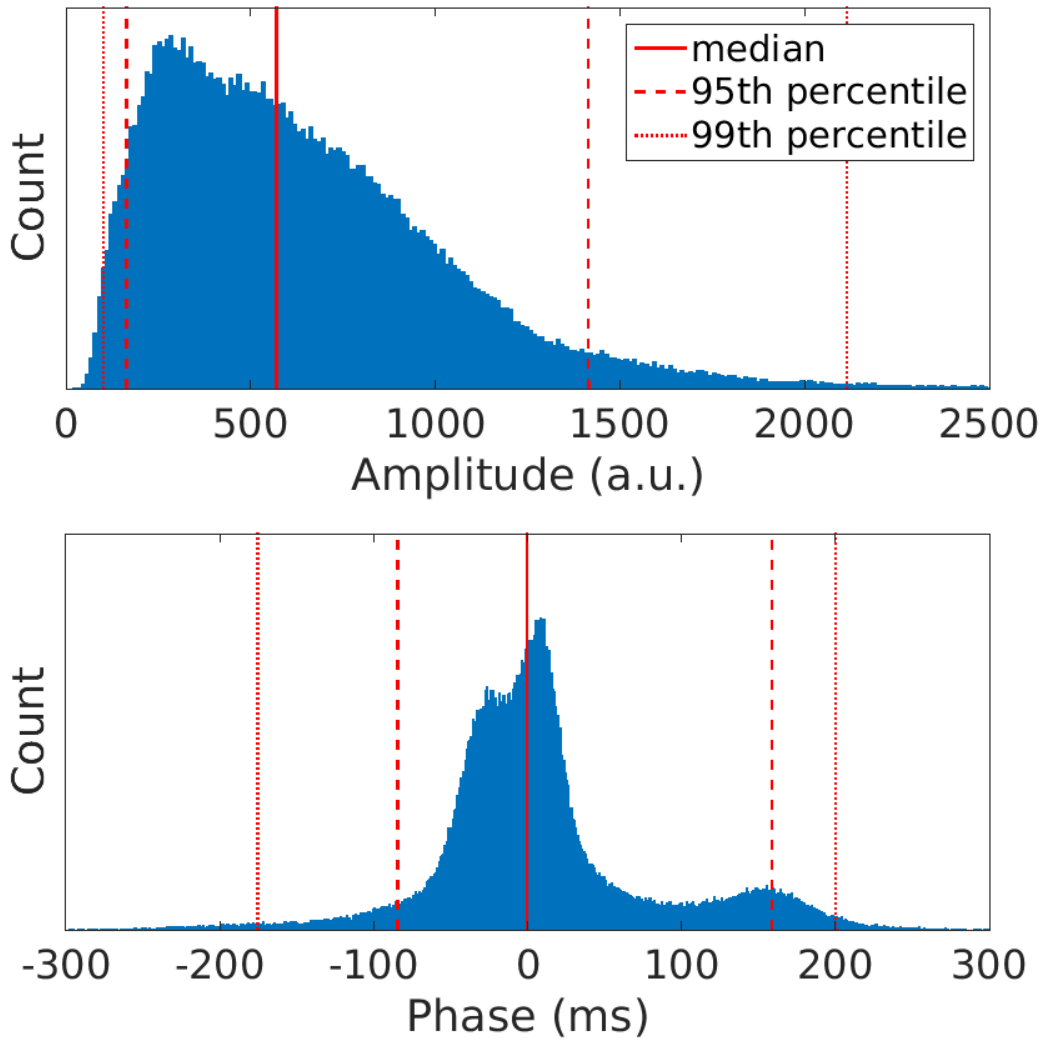


Figure 6.6: Histogram of the values of local rPPG amplitude and phase across all processed maps for subject HP3-33.

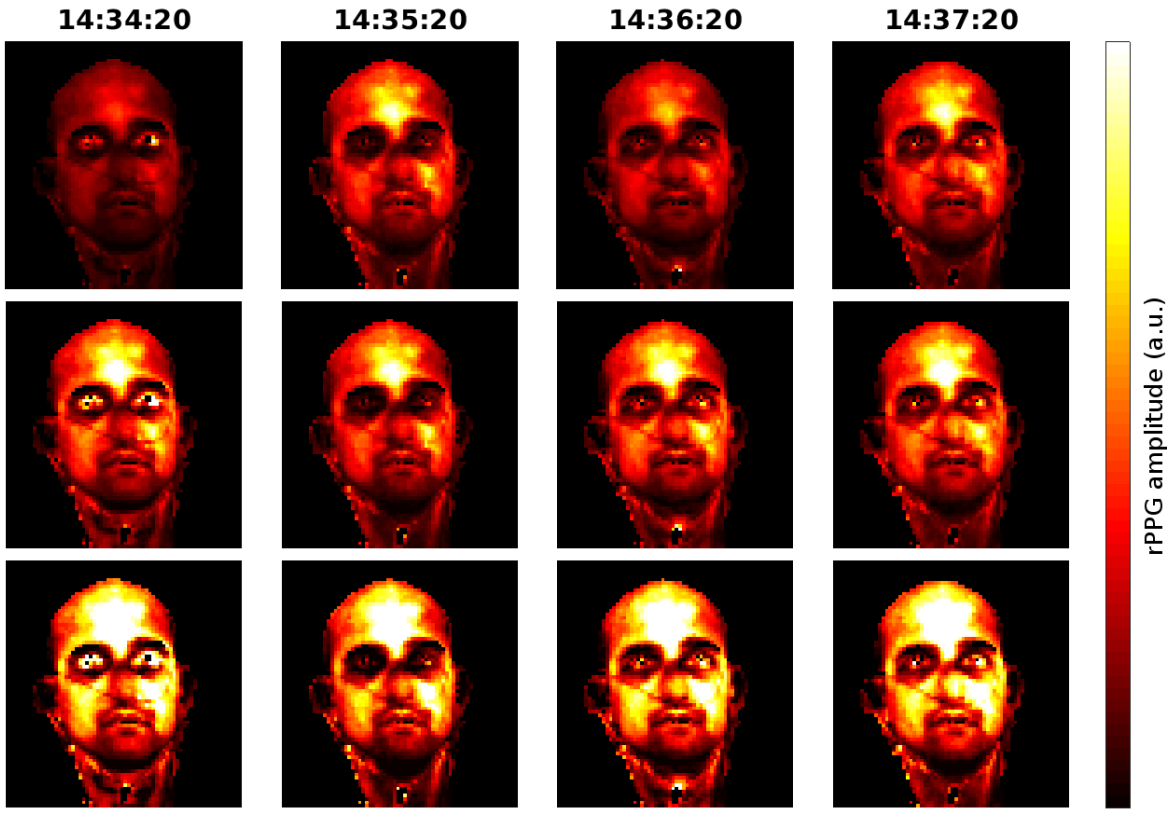


Figure 6.7: Four maps of local rPPG amplitude, each covering a 30-second window, each separated by one minute. Top row: original map after filtering and masking. Middle row: original map scaled to 1st–99th percentile values. Bottom row: original map scaled to 5th–95th percentile values.

6.8 Results

The methods described in section 6.6 were used to create spatial maps of local rPPG amplitude and phase for each of the 43 subjects of the Physiology study. Maps from 12 of these subjects are shown in figure 6.8 (showing local rPPG amplitude) and figure 6.9 (showing local rPPG phase).

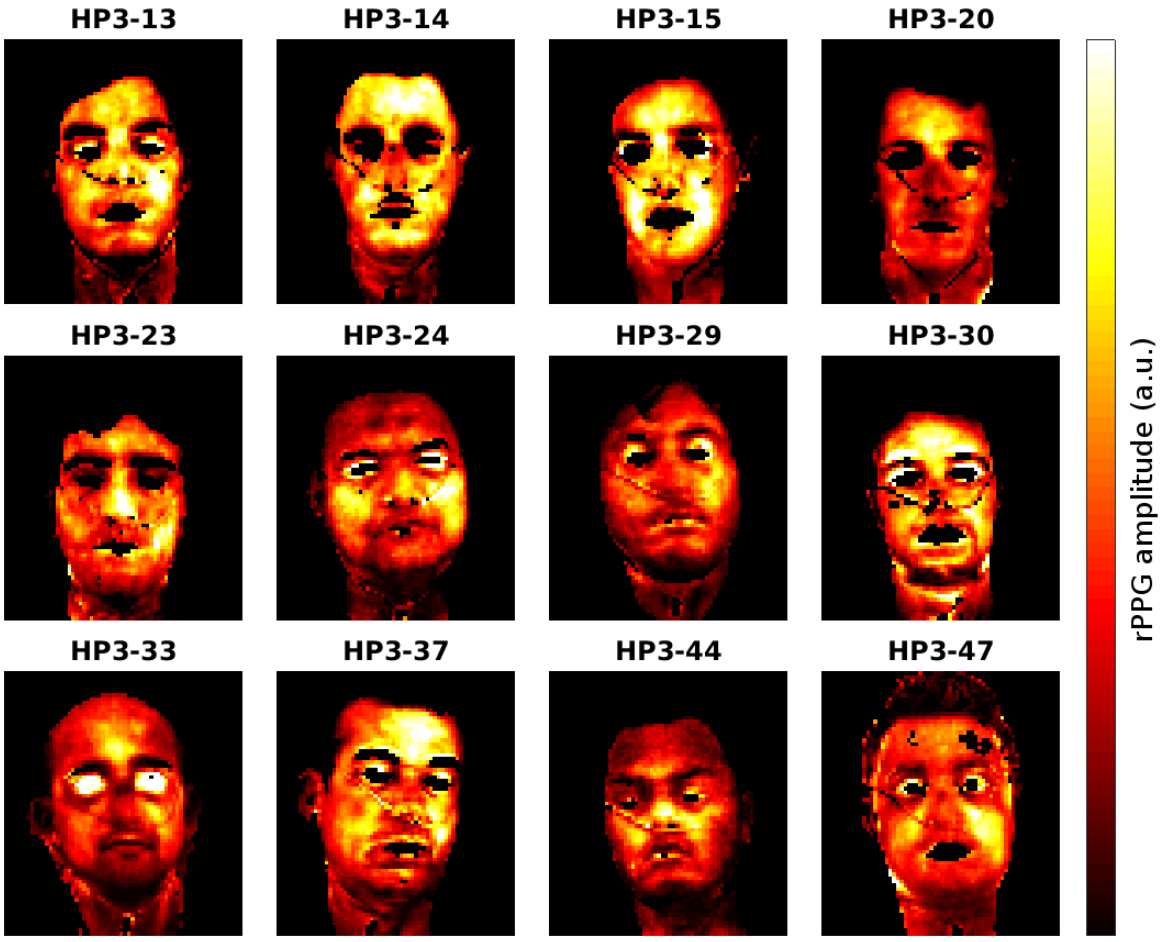


Figure 6.8: Spatial maps of local rPPG amplitude from 12 subjects of the Physiology study.

The method of spatio-temporal mapping in section 6.7 was also used to create animations showing how the maps changed over time. Figure 6.10 shows a summary of the spatio-temporal maps over the course of a session for three subjects, and videos of spatio-temporal mapping for several subjects of the Physiology study can be found at https://www.dropbox.com/sh/lyss0p8tvydzokc/AABJOuVSfTxNIJu7AnvACo1Oa.

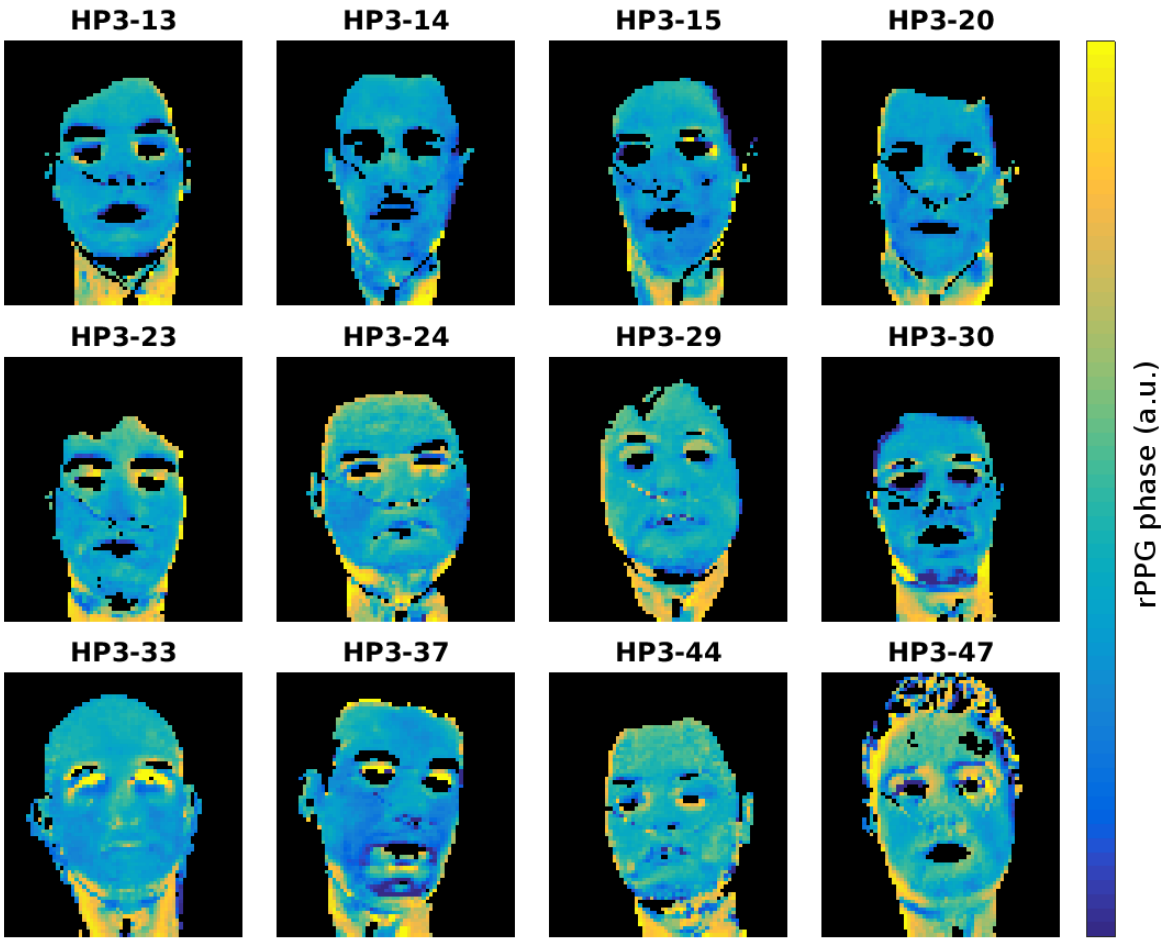


Figure 6.9: Spatial maps of local rPPG phase from 12 subjects of the Physiology study.

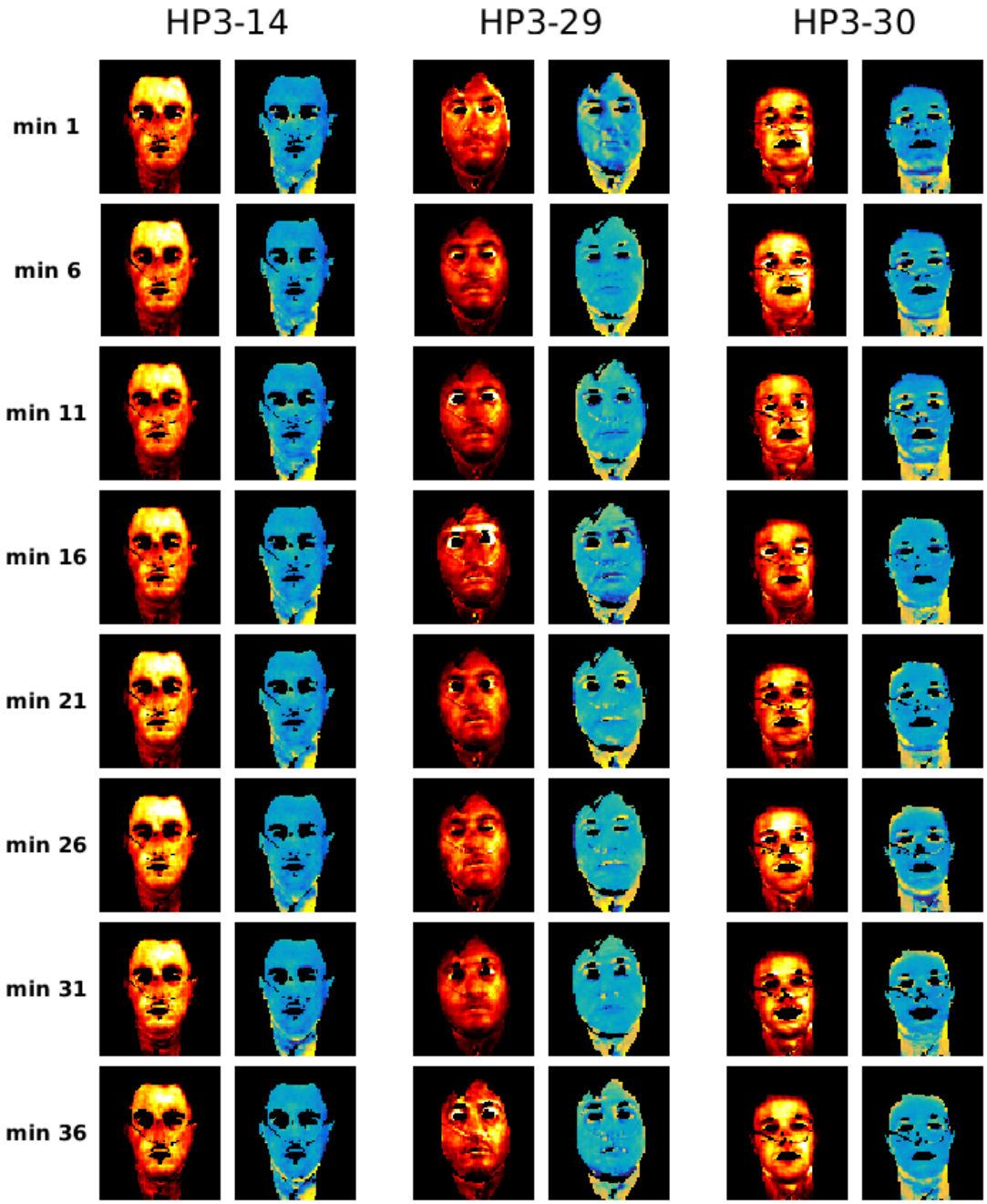


Figure 6.10: Maps of local rPPG amplitude (left columns) and phase (right columns) over the course of three different sessions.

6.9 Discussion

The results in section 6.8 show the spatial maps of haemodynamic parameters that can be produced using only video cameras. In this section, these results are discussed to assess their validity, robustness, and possible clinical utility.

Figure 6.8 shows a map of local rPPG amplitude for 12 subjects of the Physiology study. The maps show that the local rPPG amplitude is largest in the forehead and cheek regions, with areas near the neck, chin and temples showing a less prominent cardiac signal. The cheek and forehead are known to be well supplied with blood (see figure 5.2); these mapping results therefore appear to support the hypothesis that better perfused areas produce a stronger local rPPG signal. It could therefore be reasonable to think of a map of local rPPG amplitude as a map of local perfusion.

In some of the amplitude maps, the strongest signal appears to be associated with edges in the image. For example, see the amplitude map derived from subject HP3-45, in which a bright line can be seen corresponding to the subject's nasal cannula. Similarly, the raw map of local amplitude in figure 6.5 contains a bright line corresponding to the top of the subject's head. These apparent edge-detecting properties are due to the component of the rPPG signal that is caused by physical movement of the subject (see section 4.1.1); an edge that moves into and out of a small ROI will therefore produce an rPPG signal with a large amplitude.

Figure 6.9 shows the corresponding maps of local rPPG phase for the same 12 subjects. In all 12 subjects, there is an apparent phase lag between the rPPG signal in the face and the neck, with the pulse appearing to arrive at the neck after the rest of the face. This could be a manifestation of the phase ambiguity discussed in section 5.8.3, since it is not possible to differentiate large phase lags from large phase leads in a periodic signal, or could be a result of the rPPG signal being produced by venous blood rather than arterial blood (as was suggested in section 5.10).

The maps also suggest that the pulse arrives at the middle of the neck, before moving

laterally out towards both sides (indicated by the yellow sides and bluer middle in most subjects). This feature is also seen in the optical flow image shown in figure 6.4, which also shows a "velocity" in the medial-to-lateral direction.

More generally, the local rPPG phase across the face appears to be broadly similar; there is no prominent variation in phase in any particular direction. There are, however, subtle differences between different regions; most subjects have slightly bluer regions in the cheek area than the forehead area, suggesting that the pulse arrives at the cheek slightly before the forehead. Some subjects also appear to have a small change in phase across the forehead, indicating that the pulse arrives near the temples before propagating to the centre and top of the forehead.

Again, the maps show prominent changes in phase around edges in the original frames; most of the images in figure 6.9 show large phase differences near the edges of the subjects' faces. The orientation of the edge appears to affect these differences. Subject HP3-44, for example, has yellow edges on one side of the forehead, and dark blue edges on the other. Figure 6.11 shows a probable cause of this effect: if a subject periodically moves their head due to their pulse, static ROIs placed over the edge of the head will produce rPPG signals with opposing phases depending on the orientation of the edge. In the figure, if the subject's head moves slightly towards the left of the image (against a white background), the average pixel intensity in the red ROI will decrease, while the intensity in the blue ROI will increase. If the green ROI contains no edges, the rPPG phase measured here would be closer to the true phase since movement effects will not dominate the ROI.

The results of the spatio-temporal mapping (see figure 6.10) are also encouraging, as they show that both the phase and amplitude maps remain consistent over time. Most of the differences between these maps are close to prominent edges in the original images (such as around the eyes and mouth), and the artefacts caused by these edges are different from map to map. In clear areas of skin, however, the maps stay largely unchanged;

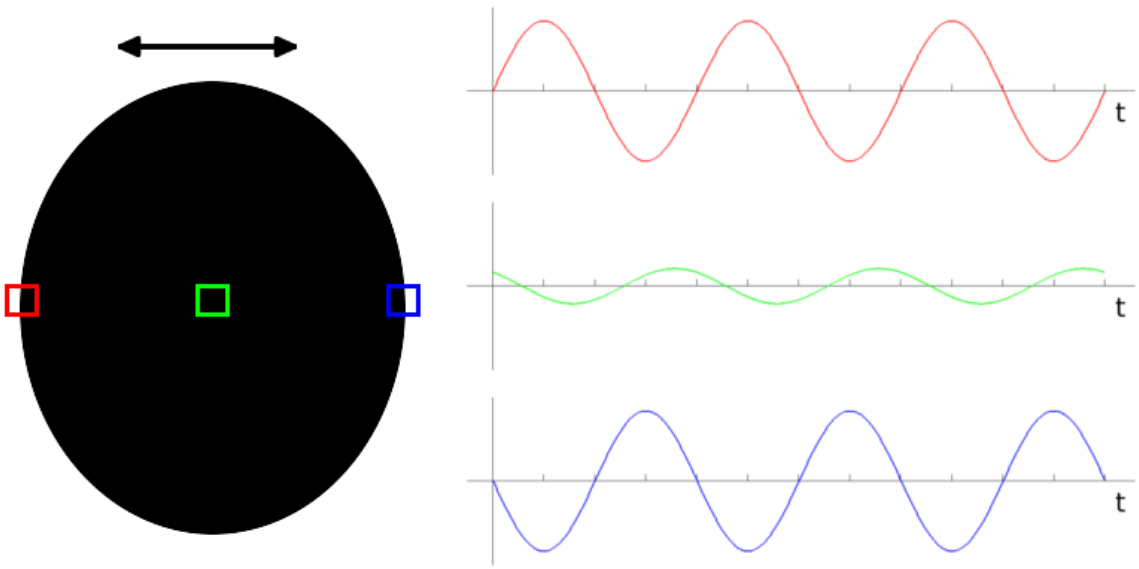


Figure 6.11: Effect of oscillatory motion on phase. The same oscillation (black arrow) will give rise to different rPPG phases if it is measured at edges with different orientations (in terms of intensity gradients).

this is expected, since physiologically meaningful changes in perfusion or blood flow are unlikely in a population of healthy volunteers who remain seated.

It was not possible to produce robust maps for several of the subjects of the Physiology study. Of the 43 subjects, 12 produced poor or failed maps. Figure 6.12 shows some examples of these failed maps, which had a number of causes. These causes, along with some corresponding examples in figure 6.12, included:

- Failure of the skin detection algorithm (due to inability to distinguish between the skin and background modes in the model), resulting in an invalid skin mask (e.g. HP3-11)
- Failure to derive an rPPG signal with a high SNR, particularly in subjects with darker skin (e.g. HP3-42)
- Excessive movement of the subject during the 30-second window (e.g. HP3-38)
- Domination of the map by edges, due either to a low rPPG SNR or unusually large cardiac-synchronous subject movement (e.g. HP3-45)

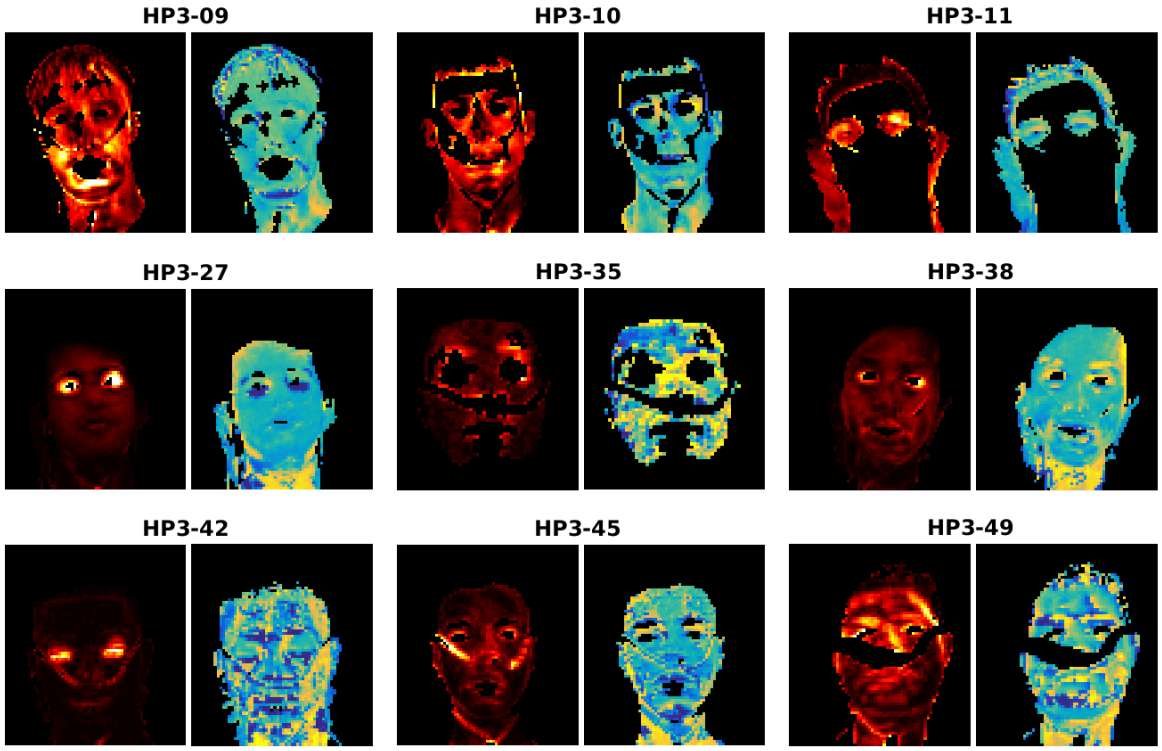


Figure 6.12: Examples of poor or failed spatial maps.

The results suggest that it may be possible to perform real-time mapping of perfusion and blood flow using only a video camera. This could have implications in the fields of intensive care (where sudden changes in haemodynamics could be the first sign of patient deterioration) and perfusion imaging. In the latter case, this spatial mapping has the potential to become an alternative to current methods of perfusion measurement, which include techniques such as Doppler ultrasound [36], MRI with arterial spin labelling [172], and PET/SPECT (two modalities of nuclear functional imaging) [12].

Chapter 7

Conclusions and future work

This chapter concludes the thesis, and begins with a review of the main results in section 7.1. The limitations of the research are discussed in section 7.2, and section 7.3 gives some suggestions for the directions in which future research could be focused. Finally, some concluding remarks can be found in section 7.4.

7.1 Thesis review

Chapter 1 introduced the clinical motivation for this thesis, arguing that conventional methods of patient monitoring have inherent disadvantages compared to new methods of non-contact continuous monitoring. Three areas of investigation were identified. The chapter also contained a detailed description of the video recordings and reference data used in subsequent chapters.

In Chapter 2, the background of PTT was explored. Three physiological signals (along with the vital signs that can be derived from them) were introduced: the continuous blood pressure waveform, from which SBP, MAP and DBP can be derived; the ECG, which is used to calculated HR; and the PPG, commonly used to estimate both HR and SpO_2. The conventional measurement of cPTT from the ECG and PPG was described, and a literature review was presented in which the relationship between cPTT and BP

was investigated.

Chapter 3 was an investigation into the relationship between cPTT and BP in a population of patients on haemodialysis. The physiology of haemodialysis was introduced, and the correlation between the two variables was then explored in depth. Firstly the global relationship was quantified over the whole patient population (section 3.2.1), before the individual relationships were assessed first on a per-subject basis (section 3.2.2) and then on a per-session basis (section 3.2.3). Next, longitudinal examples of data were presented showing examples of subjects for whom there was a clear negative correlation between cPTT and MAP (figure 3.5), and examples of subjects for whom there was no obvious relationship (figure 3.6). In the discussion (section 3.4), it was noted that the observed relationships were weaker than had been reported in the literature, probably because of the heterogeneity of the patient population in the Renal Unit study. It was also suggested that the use the Finometer device to estimate BP (rather than a conventional arterial line) could have made it more difficult to determine the relationship between the two parameters. Although the per-session correlations were found to be slightly stronger than the per-subject and global correlations, the relationship is too variable to be of practical clinical use except for a sub-set of patients; regular calibration would be required for these individuals, however. Further study of the PTT-MAP relationship in patient sub-populations (or in patients with different primary diseases) may help identify groups of patients for whom continuous monitoring of PTT may be clinically useful.

From Chapter 4 onwards, the focus switched from conventional monitoring to non-contact monitoring using video cameras. The principles behind the derivation of the rPPG signal were explained, and section 4.2 was a literature review of the field of non-contact vital sign monitoring using video cameras. An investigation into the frequency content of rPPG signals followed, and the chapter finished with a discussion of the challenges associated with non-contact monitoring.

In Chapter 5, an investigation was carried out to compare conventional cPTT to

camera-derived PTT. The chapter began with a short review of the literature related to camera-derived PTT estimates, and the physiology of the facial vasculature was then briefly discussed. The next sections described the methods used to estimate rPTT (measured between the ECG and the rPPG signals) and dPTT (measured between two rPPG signals), which required video preprocessing, ROI selection, and derivation of the rPPG signal. Some results were then presented and discussed in sections 5.9 and 5.10. These demonstrated that PTT could be estimated robustly and consistently using only a video camera, although direct validation of these estimates was not possible because of the different sites used to measure cPTT.

The chapter produced some interesting results, notably those shown in figures 5.23 and figure 5.29. The former figure suggested that rPTT and cPTT have very similar trends in some cases, which could be of clinical use (since a sudden rise in rPTT might be associated with a decrease in BP, for example). In Figure 5.29, the results suggested that the pulse in the forehead tends to propagate from the temple towards the centre of the forehead, which is in keeping with the known physiology of the facial arteries. These results showed that camera-derived PTT estimates have physiological relevance, and can even provide spatial information.

This spatial information was explored in depth in Chapter 6, which investigated whether local haemodynamic parameters (such as rPPG amplitude and phase) could be spatially mapped. A short review of the literature was first presented, including papers attempting such spatial mapping using video cameras. Methods were then presented to create representative local rPPG waveforms for a given area of skin; these methods included a novel application of the technique of signal averaging to increase the SNR. Links were also proposed between local rPPG amplitude and perfusion levels, and between local rPPG phase and blood flow. The methods for spatial mapping were then described, before results of the mapping process were presented in section 6.8 and discussed in section 6.9. The results suggested that it is possible to perform spatial

mapping of haemodynamic parameters using only video cameras; for most subjects, robust maps were produced over time showing that the forehead and cheeks were the most perfused areas (see figure 6.8), which is consistent with the known physiology of the facial vasculature. The local phase information was less useful as it tended to be fairly uniform across the face, but with some unexpected variation; most of the maps in figure 6.9 show the blood seemingly arriving at the neck after the rest of the face. Some possible causes were suggested: firstly, that the camera is in fact partially mapping venous return rather than arterial outflow; and secondly, that the mapping methods give rise to phase ambiguities that are being incorrectly resolved. It was also noted that the mapping methods failed for several of the subjects, for reasons such as failure of the skin detection algorithm, noise and artefact in the rPPG signals (particularly in subjects with darker skin), excessive subject movement, and edge effects.

7.2 Limitations

Compared to the most of the previous studies that have been published in the field of non-contact vital sign monitoring, the analysis performed in this thesis involves longer recordings and less precise control over the conditions under which subjects were monitored. The analysis has a number of limitations that should be considered when assessing the results.

Some of these limitations come in the form of the study protocols. These include the following:

- Continuous BP was not measured using an arterial line (which is the gold standard) in either of the studies. Continuous BP was acquired in the Renal Unit study, but this was done using a Finometer device which only provides an estimate of BP. The relationships between cPTT and MAP found in Chapter 3 should be considered with this in mind, since they would be affected by any error in the estimation of

continuous BP. In the Renal Unit study, where continuous BP was not available, presumed changes in BP cannot be verified due to the lack of a reference.

- In the Physiology study, healthy male volunteers were used instead of hospital patients, and video was recorded in much more controlled circumstances. In any hospital environment outside of intensive care (as was the case in the Renal Unit study, for example), subjects tend be much more active, leading to frequent occurrence of movement artefact in derived rPPG signals. Light levels also fluctuate with activity in the patient's environment, and direct line of sight with the patient is often disturbed. In a clinical situation, camera-derived PTT estimates will therefore be less robust than for the results presented in Chapters 5 and 6.

- In the course of the video recordings, reference data were acquired simultaneously by different monitoring equipment. The data were timestamped by each monitor according to their internal clocks, and are therefore subject to error and clock drift. An attempt was made to keep the different monitors synchronised during data acquisition, but with limited success. In some cases, the video data had to be synchronised with the reference data using the subject's physiology (see section 5.7.1, for example), which is only a partial solution. In the case of the rPTT estimates in Chapter 5, this meant that the absolute rPTT values were physiologically meaningless, as a result of the unknown time offset between the ECG and rPPG time-series. Only the changes in rPTT were meaningful.

As well as the limitations arising from the study protocols, further limitations affect the results presented in Chapters 3, 5 and 6. These are as follows:

- The relationships between cPTT and MAP found in Chapter 3 varied between subjects, and even between sessions recorded from the same subject. Because no global relationship was observed, it is unlikely that cPTT will ever be used clinically as a surrogate for BP. Instead, patients for whom there is a useful relationship

between cPTT and MAP would have to be identified, and the parameters of this relationship would have to be used to calibrate for the individual. This would require a great deal of further research into specific patient cohorts, with the goal of identifying the patient characteristics which are likely to lead to meaningful cPTT estimates (such as a healthy vascular physiology).

- Although Chapter 5 suggested that rPTT and dPTT could be robustly estimated from video recordings, direct validation of these estimates was not possible with the available reference data. Because cPTT is measured from heart to finger, and rPTT was measured from heart to face, the two values are not directly comparable. Validation against a gold standard is a fundamental requirement for the introduction of any new technology into a clinical setting, and the analysis in this thesis allows only a qualitative assessment of whether cPTT, rPTT and dPTT can be compared in a useful way.

- In both Chapters 5 and 6, it was observed that robust results could not be generated for every subject of the Physiology study. In particular, it was noted that the rPPG signals derived from subjects with darker skin had a lower SNR; this is a widespread limitation of video camera monitoring methods [160, 16, 149], and future research needs to be directed towards vital sign estimation in subjects with darker skin tones. In addition, the skin detection algorithm in section 6.6.3 failed to perform correctly on a number of subjects, and spatial mapping of haemodynamic parameters was not possible in these cases.

7.3 Future work

Some of the limitations of the thesis can be addressed with further research. The field of non-contact vital sign monitoring is relatively new, and expanding quickly; there are many directions in which this research can be taken, but three specific areas directly

follow on from this thesis.

- **Validation of camera-derived PTT using a comparable reference, and comparison with the BP gold standard.** In order to be comparable, the reference and camera PTT must use the same sites in the vasculature; for example, video of a subject's hand can be used along with ECG to derive heart-to-finger rPTT, which can then be compared directly with reference cPTT. Alternatively, rPTT estimated at the face could be compared to PTT derived between ECG and PPG measured using an ear probe. Ideally, these PTT estimates should then be compared with BP measured using an arterial line (the gold standard for continuous BP monitoring) to assess whether changes in PTT correspond to opposing changes in BP. A study protocol involving an arterial line is unlikely to be given ethical approval unless the arterial line would be used regardless; for this reason, patients in intensive care are a good choice of study population as they usually have an arterial line in place. There are two further advantages to choosing patients in intensive care: firstly, they are less likely to move as they tend to be sedated; and secondly, they tend to be at risk of sudden changes in blood pressure, which are useful for assessing the relationship between PTT and BP.

- **Spatial mapping of perfusion and blood flow in at-risk patients, with a focus on vascular physiology.** An appropriate clinical study could be helpful to determine the extent to which spatial haemodynamic mapping is useful to clinicians. This study could focus on a site in which the vasculature is simpler and better understood, such as the hand and arm (where there is little arterial branching, and the direction of blood flow is clearer). Alternatively, a study could investigate patients that have undergone a procedure involving free tissue transfer, where a flap of skin tissue is grafted from one location to another. Haemodynamic mapping of a graft location should reveal the perfusion levels within the graft, and might also reveal a local rPPG phase difference between the flap and the surrounding

tissue. A study of this sort could be helpful in assessing the performance of the mapping, and the technique could eventually be used to provide some indication to clinicians as to how well the graft is perfused in the days that follow the procedure.

- **Continuous monitoring of hospital patients, including 24-hour recording.** This monitoring could include established methods including HR and BR estimation, as well as some of the newer methods of haemodynamic monitoring proposed in this thesis. Ideally, any clinical study would attempt 24-hour monitoring, which would require infra-red illumination so as not to disturb the patients at night; the monitoring algorithms would therefore have to be adapted to work with rPPG signals derived from infra-red illumination. An appropriate patient population might be those in high-dependency care, such as patients recovering during the 48 hours following major surgery.

7.4 Concluding remarks

The results presented in this thesis suggest that aspects of haemodynamic monitoring can be performed using only video cameras. Specifically, it appears that PTT can be estimated robustly from video data. In addition, the results indicate that spatial mapping of haemodynamic parameters is possible, and that these maps are likely to contain physiological information that is clinically useful. This opens up the possibility that these mapping methods could form the basis of a fast and inexpensive imaging modality using visible light (or possibly infra-red illumination for night-time monitoring), which could have applications in critical care medicine.

If technology and algorithms continue to advance at their current rate, the advantages offered by non-contact monitoring could result in video cameras becoming a fundamental part of hospital care in the future.

Appendix A

Pulse oximetry

In pulse oximetry (see section 2.3.2), the proportion of haemoglobin that is carrying oxygen is determined by measuring the intensity of light passing through peripheral tissue. The peripheral oxygen saturation SpO_2 is defined as

$$\mathrm{SpO}_2 = 100\% \times \frac{c_o}{c_o + c_r} \tag{A.1}$$

where c_o is the concentration of oxyhaemoglobin and c_r is the concentration of deoxygenated haemoglobin (see equation 2.3).

The attenuation of light as it passes through a medium is related to the concentration according to Beer's law, which states that the intensity of transmitted light decreases exponentially as the concentration increases. The attenuation also depends on the distance travelled through the tissue according to Lambert's law, which states that the intensity of the transmitted light decreases exponentially as the distance increases. The two laws can be combined to form the Beer-Lambert law, which states that

$$A = a(\lambda)cl \tag{A.2}$$

where A is the absorbance, a is an absorption coefficient that depends on wavelength,

c is the concentration and l is the path length. The absorbance can be related to the initial light intensity I_i and the transmitted intensity I_t using the formula

$$A = \log_{10}\left(\frac{I_i}{I_t}\right) \tag{A.3}$$

which can be substituted into equation A.2 and rearranged to produce

$$I_t = I_i 10^{-a(\lambda)cl}\,. \tag{A.4}$$

Accounting for both oxyhaemoglobin and deoxygenated haemoglobin—and extending to two separate wavelengths—produces

$$\begin{aligned} I_{t1} &= I_{i1} 10^{-(a_{o1}c_o + a_{r1}c_r)l} \\ I_{t2} &= I_{i2} 10^{-(a_{o2}c_o + a_{r2}c_r)l} \end{aligned} \tag{A.5}$$

where a_{on} is the absorption coefficient of HbO_2 at wavelength n, a_{rn} is the absorption coefficient of Hb at wavelength n, c_o is the concentration of HbO_2 and c_r is the concentration of Hb. These equations can be rearranged to

$$\begin{aligned} \frac{1}{l}\log_{10}\left(\frac{I_{i1}}{I_{t1}}\right) &= a_{o1}c_o + a_{r1}c_r \\ \frac{1}{l}\log_{10}\left(\frac{I_{i2}}{I_{t2}}\right) &= a_{o2}c_o + a_{r2}c_r \end{aligned} \tag{A.6}$$

and combined to produce

$$\frac{\log_{10}\left(\frac{I_{i1}}{I_{t1}}\right)}{\log_{10}\left(\frac{I_{i2}}{I_{t2}}\right)} = \frac{a_{o1}c_o + a_{r1}c_r}{a_{o2}c_o + a_{r2}c_r}\,. \tag{A.7}$$

Using the substitution

$$R = \frac{\log_{10}\left(\frac{I_{i1}}{I_{t1}}\right)}{\log_{10}\left(\frac{I_{i2}}{I_{t2}}\right)} \tag{A.8}$$

allows c_o to be expressed as

$$c_o = \frac{a_{r1} - a_{r2}R}{a_{o2}R - a_{o1}} c_r \tag{A.9}$$

which can be substituted into equation A.1 to produce the general oximetry equation

$$\frac{c_o}{c_o + c_r} = \frac{a_{r1} - a_{r2}R}{(a_{r1} - a_{o1}) + (a_{o2} - a_{r2})R} \tag{A.10}$$

which expresses the SpO_2 as a function of R as the absorption coefficients are constant.

This general oximetry equation can be extended to *pulse* oximetry by considering only the absorption due to arterial blood. It is assumed that the AC component of the PPG signal is caused only by pulsatile arterial blood (see section 2.3.1). The SpO_2 can then be determined by subtracting the DC absorption from the total absorption, considering only the absorption due to arterial blood.

In this situation, the initial intensity is the PPG signal's DC baseline (denoted I_{DC}), and the transmitted intensity is sum of the DC baseline and the AC amplitude (denoted I_{AC}). Equation A.10 still holds, but R is now modified to

$$R = \frac{\log_{10}\left(\frac{I_{AC}}{I_{DC}}\right)_{\lambda 1}}{\log_{10}\left(\frac{I_{AC}}{I_{DC}}\right)_{\lambda 2}} \tag{A.11}$$

and is often referred to as the ratio of ratios, the ratio of the relative amplitude of the PPG's AC and DC components at each wavelength. The SpO_2 can then be written

$$SpO_2 = 100\% \times \frac{a_{r1} - a_{r2}R}{(a_{r1} - a_{o1}) + (a_{o2} - a_{r2})R}\,. \tag{A.12}$$

Appendix B

Additional cPTT-MAP relationships

In section 3.2, the relationship between cPTT and MAP was investigated using data from the Renal Unit study. The relationship was quantified for each subject in table 3.1, which pooled the data from each subject's various haemodialysis sessions before assessing the linear relationship for each of the subjects.

The relationship between cPTT and MAP within a given session was discussed in section 3.2.3. Figure B.1 shows histograms of correlation coefficient, RMSE, gradient and intercept across all sessions of the study.

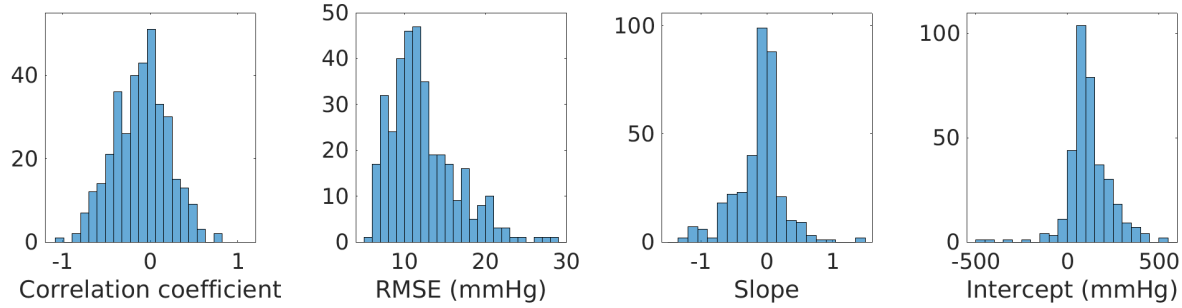


Figure B.1: Histograms showing the per-session parameters of the relationships between cPTT and MAP.

Table B.1 describes each subject's average per-session relationship, giving the mean and standard deviation for each session's correlation coefficient, and each session's RMSE, gradient and intercept for the best linear fit (a separate fit was found for each session).

#	Subject	N	r	RMSE	gradient	intercept
1	DF48	11	-0.286 ± 0.223	13.5 ± 3.0	-0.267 ± 0.318	161.5 ± 93.0
2	DF40	3	-0.340 ± 0.468	10.8 ± 5.0	-0.531 ± 0.573	241.5 ± 166.1
3	DF05	25	-0.523 ± 0.303	11.2 ± 4.2	-1.966 ± 6.841	667.8 ± 2048.1
4	DF17	12	-0.024 ± 0.153	11.2 ± 2.8	-0.035 ± 0.132	106.3 ± 38.8
5	DF11	12	-0.294 ± 0.279	12.5 ± 3.9	-0.352 ± 0.412	159.2 ± 102.6
6	DF01	11	-0.081 ± 0.257	16.7 ± 5.2	-0.043 ± 0.361	134.0 ± 95.7
7	DF55	5	-0.283 ± 0.271	10.3 ± 2.1	-0.343 ± 0.321	152.7 ± 91.4
8	DF43	12	+0.046 ± 0.302	11.6 ± 3.6	+0.108 ± 0.392	57.0 ± 103.9
9	DF37	12	-0.151 ± 0.192	15.0 ± 5.3	-0.184 ± 0.385	139.3 ± 125.6
10	DF34	11	-0.343 ± 0.179	12.7 ± 2.7	-0.513 ± 0.316	249.8 ± 76.9
11	DF21	8	-0.029 ± 0.289	12.5 ± 4.4	-0.067 ± 0.430	118.7 ± 131.4
12	DF30	9	-0.166 ± 0.326	14.1 ± 4.4	-0.019 ± 0.886	91.4 ± 235.4
13	DF54	10	-0.055 ± 0.291	13.7 ± 3.0	-0.011 ± 0.562	91.4 ± 152.7
14	DF02	2	-0.270 ± 0.018	11.6 ± 1.2	-0.113 ± 0.094	114.1 ± 27.2
15	DF20	10	-0.208 ± 0.278	10.8 ± 3.2	-0.194 ± 0.240	129.0 ± 77.8
16	DF13	10	-0.164 ± 0.169	10.0 ± 2.4	-0.141 ± 0.189	132.8 ± 59.0
17	DF04	1	-0.166 ± 0.000	10.0 ± 0.0	-0.086 ± 0.000	102.4 ± 0.0
18	DF07	8	-0.164 ± 0.152	10.3 ± 3.1	-0.229 ± 0.249	169.0 ± 69.5
19	DF45	10	-0.063 ± 0.185	10.7 ± 1.8	-0.060 ± 0.168	133.0 ± 50.4
20	DF08	10	-0.146 ± 0.171	11.9 ± 4.4	-0.166 ± 0.175	143.6 ± 47.4
21	DF60	8	-0.245 ± 0.315	9.5 ± 1.6	-0.395 ± 0.471	209.5 ± 126.1
22	DF18	12	-0.103 ± 0.273	15.1 ± 4.4	-0.191 ± 0.392	157.4 ± 103.8
23	DF26	10	+0.021 ± 0.115	9.9 ± 2.5	+0.006 ± 0.057	76.3 ± 20.9
24	DF12	12	-0.025 ± 0.294	9.1 ± 2.3	-0.032 ± 0.284	94.5 ± 71.0
25	DF23	11	-0.121 ± 0.186	11.3 ± 5.5	-0.085 ± 0.159	96.8 ± 44.2
26	DF27	11	-0.111 ± 0.371	13.5 ± 5.7	-0.093 ± 0.160	108.3 ± 54.6
27	DF53	8	-0.001 ± 0.306	9.3 ± 1.7	-0.077 ± 0.220	97.4 ± 63.4
28	DF59	9	+0.089 ± 0.404	18.6 ± 6.9	+0.024 ± 0.140	61.9 ± 39.4
29	DF22	8	-0.042 ± 0.216	10.1 ± 1.7	-0.049 ± 0.231	111.4 ± 56.5
30	DF14	10	+0.063 ± 0.234	15.1 ± 4.0	+0.131 ± 0.504	60.2 ± 177.9
31	DF41	10	+0.009 ± 0.295	12.7 ± 4.1	-0.007 ± 0.056	54.6 ± 19.5
32	DF33	7	-0.111 ± 0.235	17.0 ± 3.6	-0.075 ± 0.283	128.6 ± 68.7
33	DF24	12	-0.002 ± 0.247	14.7 ± 3.5	-0.016 ± 0.160	93.2 ± 67.1
34	DF44	11	+0.054 ± 0.167	9.0 ± 1.8	+0.043 ± 0.078	65.5 ± 34.8
35	DF49	11	+0.141 ± 0.268	9.4 ± 2.4	-0.068 ± 0.810	125.4 ± 237.6
36	DF35	13	+0.357 ± 0.253	12.8 ± 3.3	+0.421 ± 0.332	-22.1 ± 105.2
37	DF19	3	+0.176 ± 0.408	11.5 ± 3.6	+0.151 ± 0.287	43.2 ± 86.5
-	**Total**	**358**	**-**	**-**	**-**	**-**

Table B.1: Average strength of the per-session cPTT-MAP relationship shown for each subject of the Renal Unit study. The subjects are ranked as in table 3.1. The average per-session relationship is described using mean ± standard deviation. N: number of haemodialysis sessions, r: correlation coefficient, RMSE: root mean squared error (in mmHg) for best linear fit. Subjects without any paired measurements have been excluded.

Appendix C

Additional rPPG spectrograms

In section 4.3, spectrograms were created showing the frequency content of rPPG signals derived at the forehead, cheek, mouth and neck. The spectrograms were produced using data from subject HP3-33 of the Physiology study, and showed the magnitude of the FFT measured during 30-second windows of time. Additional spectrograms are shown in this appendix to demonstrate the effect ROI size has on the frequency content of the rPPG signals.

Firstly, spectrograms are presented over a range of possible breathing frequencies. The original result, figure 4.10, was created using ROIs of size 24 × 24 pixels. Equivalent spectrograms are shown here for rPPG signals derived from 12 × 12 pixel ROIs (figure C.1), and from 48 × 48 pixel ROIs (figure C.2). These larger and smaller ROIs are centred around the same locations on the face as the original ROIs in figure 4.9.

As expected (see section 4.4.3), the breathing component from the 12 × 12 pixel ROIs is slightly stronger than that from the original 24 × 24 pixel ROIs, and is slightly weaker in the case of the 48 × 48 pixel ROI. This is particularly apparent when looking at the spectrograms from the mouth ROI, which contains more bright yellow in the case of the smaller ROI.

The same spectrograms are also shown over a range of possible cardiac frequencies.

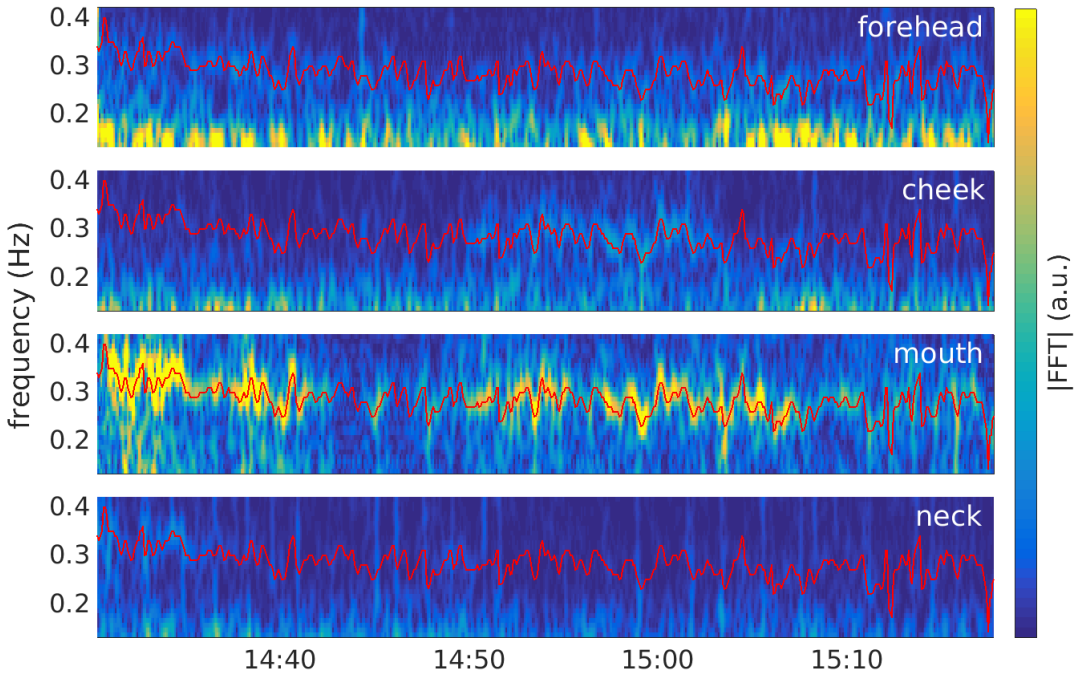


Figure C.1: Breathing-frequency rPPG spectrograms from four different 12 × 12 pixel ROIs (locations in white text). The reference BR from the IP is superimposed in red.

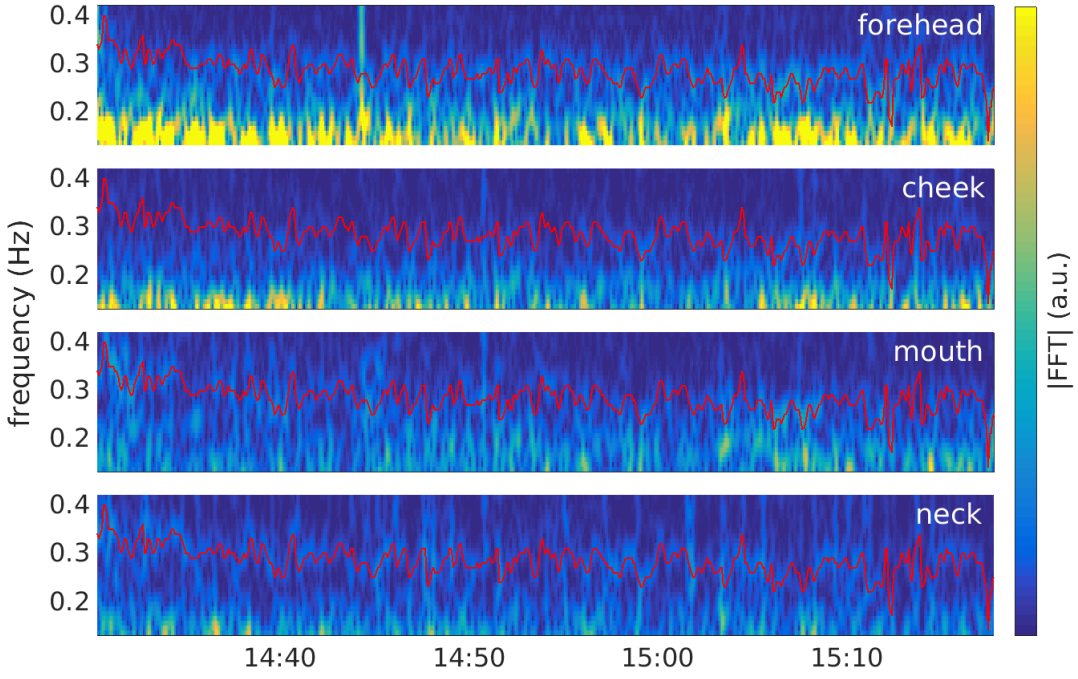


Figure C.2: Breathing-frequency rPPG spectrograms from four different 48 × 48 pixel ROIs (locations in white text). The reference BR from the IP is superimposed in red.

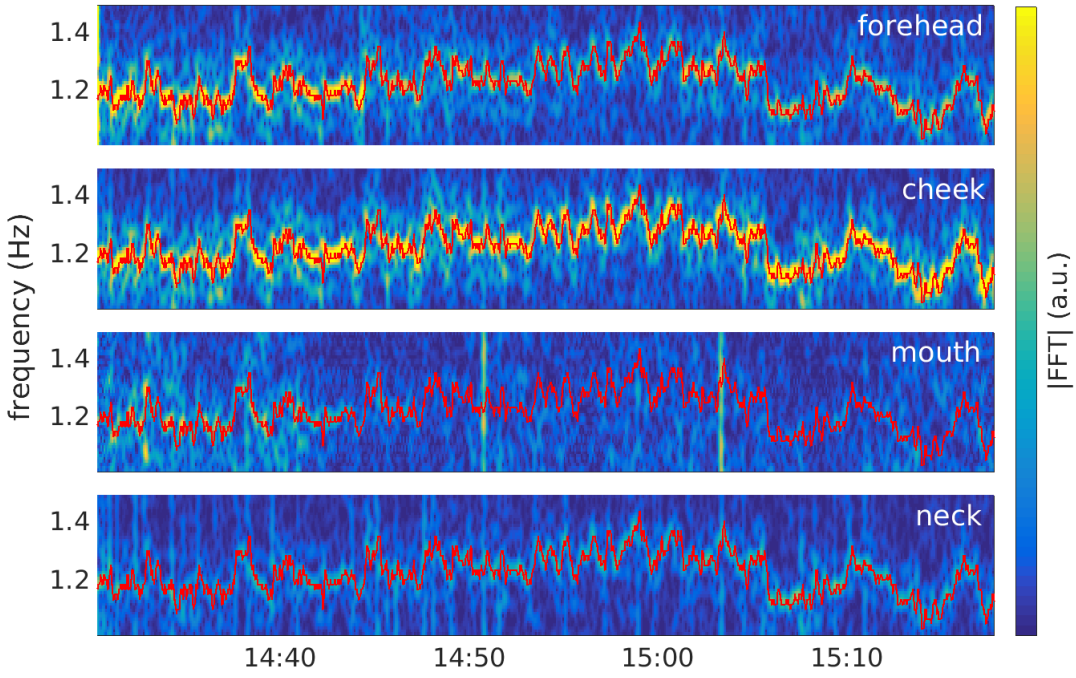


Figure C.3: Cardiac-frequency rPPG spectrograms from four different 12 × 12 pixel ROIs (locations in white text). The reference HR from the ECG is superimposed in red.

Figure C.3 shows spectrograms created from 12 × 12 pixel ROIs, and figure C.4 shows spectrograms created from 48 × 48 pixel ROIs. These figures can be compared with the original spectrogram in figure 4.12, which was created using 24 × 24 pixel ROIs.

It appears that ROI size has minimal effect on the spectrograms, which appear to show similar cardiac components in each case.

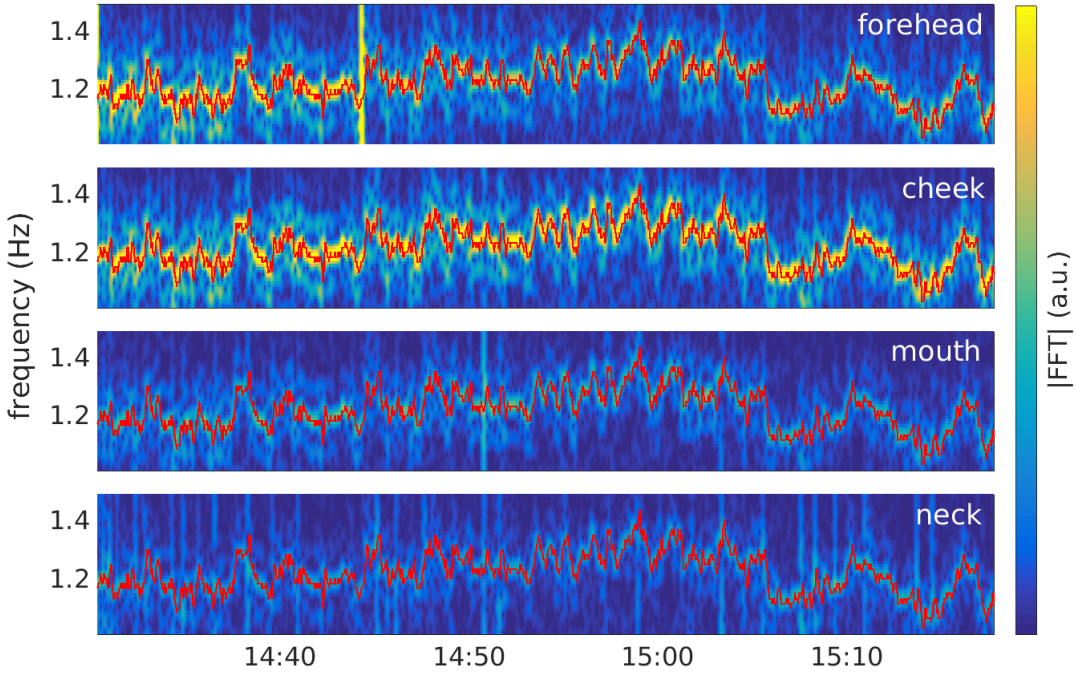


Figure C.4: Cardiac-frequency rPPG spectrograms from four different 48 × 48 pixel ROIs (locations in white text). The reference HR from the ECG is superimposed in red.

Appendix D

Piecewise cubic spline interpolation

This appendix justifies the choice of piecewise cubic spline interpolation for resampling of rPPG signals to higher sampling frequencies.

D.1 Piecewise interpolation

In piecewise interpolation, an interpolant function is created from many sub-functions, with each sub-function occupying an interval between two consecutive samples of the original signal. This is in contrast to a global interpolant function, which must be fitted to the samples across their entire range.

For longer input functions with many oscillations, interpolation using a global function requires either a high-order polynomial or a complex periodic function composed of periodic basis functions such as sinusoids. The parameterisation of these global functions is computationally expensive, and the functions are prone to overfitting and oscillation.

Piecewise interpolation produces a function of the form

$$F(x) = \begin{cases} f_1(x), & \text{if } x_0 < x < x_1 \\ f_2(x), & \text{if } x_1 < x < x_2 \\ & \vdots \\ f_N(x), & \text{if } x_{N-1} < x < x_N \end{cases} \tag{D.1}$$

where N sub-functions are combined to produce a piecewise interpolant function $F(x)$ from $N+1$ samples. A piecewise interpolation function is preferable to a global function as it allows small errors without requiring a very high-order polynomial fit. A piecewise interpolant can be forced to fit the input samples exactly by ensuring that

$$\begin{aligned} f_1(x_0) &= y_0 \\ f_1(x_1) = f_2(x_1) &= y_1 \\ f_2(x_2) = f_3(x_2) &= y_2 \\ &\vdots \\ f_{N-1}(x_{N-1}) = f_N(x_{N-1}) &= y_{N-1} \\ f_N(x_N) &= y_N \end{aligned} \tag{D.2}$$

where (x_n, y_n) is the nth raw sample and there are N sub-functions in total.

Additional equations can also be constructed to force the piecewise function to be smooth.

D.2 Interpolant function order

Piecewise interpolation requires the parameterisation of N sub-functions, each of which exists between two neighbouring input samples (of which there must be $N+1$). Low-complexity sub-functions are therefore preferred so as to keep computational requirements to

a minimum. Using polynomial sub-functions, the complexity is minimised by using low-order polynomials.

The simplest possible function is a zeroth-order polynomial of the form $f_n(x) = a_n$, which corresponds to nearest-neighbour interpolation. Similarly, one can use first-order polynomials of the form $f_n(x) = a_n + b_n x$, which corresponds to linear interpolation. Neither of these methods allow the peaks of the input signal to be located more accurately; the interpolated signal will always have its peaks at the original locations. They are therefore unsuitable for the purposes of peak location.

Increasing the complexity, quadratic functions of the form $f_n(x) = a_n + b_n x + c_n x^2$ can be used. The use of higher-order polynomials allows more complex constraints on the sub-functions; for example, it is desirable that the interpolant function is smooth so that its derivative has no discontinuities, and so the equations

$$\begin{aligned} f_1'(x_1) &= f_2'(x_1) \\ f_2'(x_2) &= f_3'(x_2) \\ &\vdots \\ f_{N-1}'(x_{N-1}) &= f_N'(x_{N-1}) \end{aligned} \tag{D.3}$$

can be enforced along with those in equation D.2. These equations impose constraints on the sub-functions, and produce a system of simultaneous equations that can be solved to find the coefficients (see section D.3 for an example). Piecewise quadratic interpolation allows more accurate estimation of peak location, and produces an output which is smooth through to its first derivative, but is prone to oscillation as there are few degrees of freedom in the sub-functions.

Cubic sub-functions (also referred to as splines) produce an interpolated function that is continuous through to its second derivative. The functions are of the form

$f_n(x) = a_n + b_n x + c_n x^2 + d_n x^3$, which allows the constraints

$$\begin{aligned} f_1''(x_1) &= f_2''(x_1) \\ f_2''(x_2) &= f_3''(x_2) \\ &\vdots \\ f_{N-1}''(x_{N-1}) &= f_N''(x_{N-1}) \end{aligned} \tag{D.4}$$

to be enforced. When combined with equations D.2 and D.3, this again produces a system of simultaneous equations that can be solved to find the coefficients of each sub-function. Cubic interpolation does not suffer from oscillation to the extent that quadratic interpolation does.

Higher order polynomials (4+) can be used to create interpolated functions with continuous higher derivatives, but this gives diminishing returns for the increase in computational expense. Cubic spline interpolation is considered sufficient for most purposes, and in the case of PPG signals produces an interpolated signal that appears completely smooth.

Figure D.1 shows a visual comparison of different methods of interpolation. The figure shows that although each of the methods produce a continuous interpolant, the linear interpolant is not smooth (its first derivative is discontinuous, exhibiting sudden steps). The quadratic interpolant is continuous though to its first derivative, and the cubic interpolant is continuous through to its second derivative. Cubic interpolation therefore produces the smoothest function of the three methods.

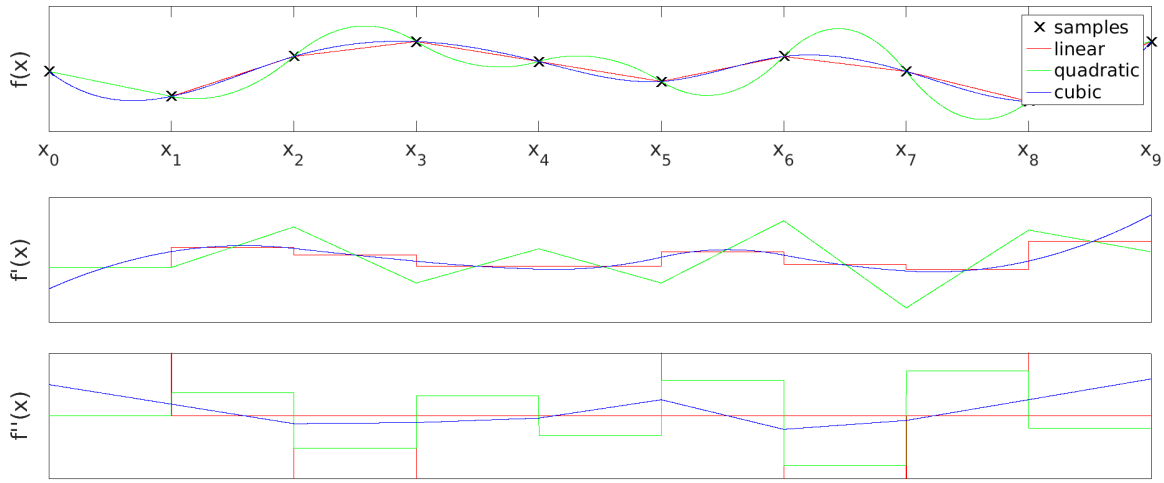


Figure D.1: Comparison of piecewise linear, quadratic and cubic interpolation. Top: raw samples and interpolants. Middle: first derivative of interpolants. Bottom: second derivative of interpolants.

D.3 Quadratic interpolation example

In piecewise quadratic interpolation (see section 5.6.2), quadratic sub-functions are used as interpolants between a set of samples to create a piecewise interpolated function (see figure D.2). This process involves the solution of a set of simultaneous equations. Each sub-function is of the form $f_n(x) = a_n + b_n x + c_n x^2$.

A first set of equations arises from the requirement that each sub-function passes through the samples either side of it, producing

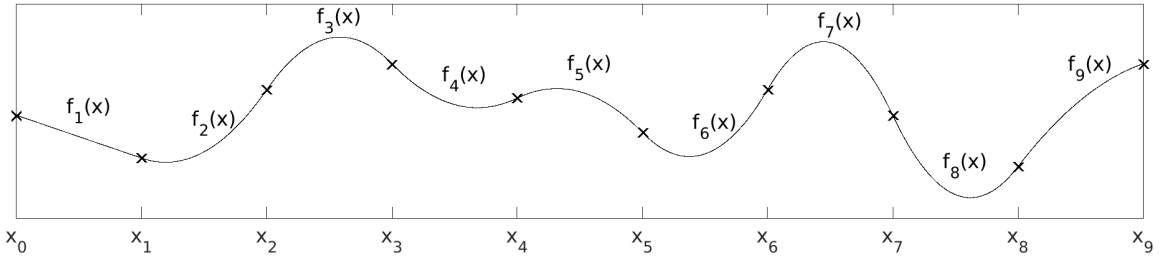


Figure D.2: Diagram of piecewise quadratic interpolation. Black crosses: original samples. Black line: piecewise quadratic interpolant.

$$
\begin{aligned}
a_1 &+ b_1 x_0 &+ c_1 x_0^2 &= y_0 \\
a_1 &+ b_1 x_1 &+ c_1 x_1^2 &= y_1 \\
a_2 &+ b_2 x_1 &+ c_2 x_1^2 &= y_1 \\
a_2 &+ b_2 x_2 &+ c_2 x_2^2 &= y_2 \\
& & \vdots & \\
a_{N-1} &+ b_{N-1} x_{N-2} &+ c_{N-1} x_{N-2}^2 &= y_{N-2} \\
a_{N-1} &+ b_{N-1} x_{N-1} &+ c_{N-1} x_{N-1}^2 &= y_{N-1} \\
a_N &+ b_N x_{N-1} &+ c_N x_{N-1}^2 &= y_{N-1} \\
a_N &+ b_N x_N &+ c_N x_N^2 &= y_N
\end{aligned}
\tag{D.5}
$$

where (x_n, y_n) is the nth sample and there are N sub-functions (covering $N+1$ samples) overall. Note that there are $2N$ equations in total.

A second set of equations arises from the desire to have a continuous first derivative (in order to achieve a smooth piecewise function). A continuous first derivative requires that $f'_{n-1}(x_{n-1}) = f'_n(x_{n-1})$. Since $f'_n(x) = b_n + 2c_n x$, this produces

$$
\begin{aligned}
b_1 &+ 2c_1 x_1 &- b_2 &- 2c_2 x_1 &= 0 \\
b_2 &+ 2c_2 x_2 &- b_3 &- 2c_3 x_2 &= 0 \\
& & \vdots & & \\
b_{N-2} &+ 2c_{N-2} x_{N-2} &- b_{N-1} &- 2c_{N-1} x_{N-2} &= 0 \\
b_{N-1} &+ 2c_{N-1} x_{N-1} &- b_N &- 2c_N x_{N-1} &= 0
\end{aligned}
\tag{D.6}
$$

which contributes another $N-1$ equations.

Combining the equations in D.5 and D.6, the system has $3N-1$ equations in total. There are $3N$ unknowns in total (three coefficients in each of the N sub-functions), and the system is therefore underdetermined; there is one more unknown than there are equations, meaning in this case that there are infinitely many solutions. Another equation must therefore be created using a boundary condition. A simple boundary condition is

$$c_1 = 0 \tag{D.7}$$

which constrains f_1 to be linear. Adding this equation to the system allows for a unique solution.

The equations in D.5, D.6 and D.7 can be combined into matrices to produce an equation of the form $\mathbf{Ax} = \mathbf{b}$. The equation is

$$\begin{bmatrix}
1 & x_0 & x_0^2 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & x_1 & x_1^2 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & x_1 & x_1^2 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & x_2 & x_2^2 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & x_{N-2} & x_{N-2}^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \dots & 1 & x_{N-1} & x_{N-1}^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & x_{N-1} & x_{N-1}^2 \\
0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 1 & x_N & x_N^2 \\
0 & 1 & 2x_1 & 0 & -1 & -2x_1 & \dots & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 2x_{N-1} & 0 & -1 & -2x_{N-1} \\
0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
a_1 \\ b_1 \\ c_1 \\ a_2 \\ b_2 \\ c_2 \\ \vdots \\ a_{N-1} \\ b_{N-1} \\ c_{N-1} \\ a_N \\ b_N \\ c_N
\end{bmatrix}
=
\begin{bmatrix}
y_0 \\ y_1 \\ y_1 \\ y_2 \\ \vdots \\ y_{N-2} \\ y_{N-1} \\ y_{N-1} \\ y_N \\ 0 \\ 0 \\ \vdots \\ 0
\end{bmatrix} \tag{D.8}$$

where the sub-function coefficients form the unknown matrix $\mathbf{x}$. The equation can be solved for $\mathbf{x}$ by rearranging to $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$; given the sparsity of the matrix $\mathbf{A}$, numerical inversion is straightforward even with a large number of sub-functions.

Once $\mathbf{x}$ has been found, the sub-functions can be created and combined to form a

global piecewise interpolant.

Note that a similar system can be set up using cubic sub-functions, but the matrices involved are much larger.

Appendix E

QRS detection algorithm

The method used in section 5.7.3 to detect ECG R-peaks is based on the QRS detection algorithm proposed by Pan and Tompkins (1985) [116]. The method uses the following steps:

1. Band-pass filtering of the ECG (with a pass-band of 5-15 Hz)
2. Differentiator filtering to highlight the QRS complex
3. Squaring of the resulting signal to produce a power signal
4. Averaging with a moving window to reduce noise
5. Adaptive thresholding of the resulting pulse waveform to discriminate QRS complexes from noise artefact or large T-waves. This process includes:
   (a) Location of peaks in the pulse waveform
   (b) Thresholding of peaks using continuously adapting threshold values (one to identify QRS complexes, another to identify noise), which are adjusted depending the size of the peaks
   (c) Searchback for missed QRS complexes if no peak occurs within a reasonable period following the last identified QRS complex

(d) Elimination of multiple QRS detections within a refractory period following the last identified QRS complex

(e) Discrimination of T-waves based on slope and time since last identified QRS complex

Bibliography

[1] P. S. Addison. Modular continuous wavelet processing of biosignals: extracting heart rate and oxygen saturation from a video signal. *Healthcare Technology Letters*, 3(2):111–115, Jun 2016.

[2] C. Ahlstrom, A. Johansson, F. Uhlin, T. Lanne, and P. Ask. Noninvasive investigation of blood pressure changes using the pulse wave transit time: a novel approach in the monitoring of hemodialysis patients. *Journal of Artificial Organs*, 8(3):192–197, 2005.

[3] J. Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, 28(3):1–39, Mar 2007.

[4] R. Amelard, R. L. Hughson, D. K. Greaves, K. J. Pfisterer, J. Leung, D. A. Clausi, and A. Wong. Non-contact hemodynamic imaging reveals the jugular venous pulse waveform. *Scientific Reports*, 7:40150, 2016.

[5] M. Antonelli, M. Levy, P. J. Andrews, J. Chastre, L. D. Hudson, C. Manthous, G. U. Meduri, R. P. Moreno, C. Putensen, T. Stewart, and A. Torres. Hemodynamic monitoring in shock and implications for management. International Consensus Conference, Paris, France, 27-28 April 2006. *Intensive Care Medicine*, 33(4):575–590, Apr 2007.

[6] G. G. Arabidze, V. Petrov, and J. A. Staessen. Blood pressure by Korotkoff's auscultatory method: end of an era or bright future? *Blood Pressure Monitoring*, 1(4):321–327, Aug 1996.

[7] U. Bal. Non-contact estimation of heart rate and oxygen saturation using ambient light. *Biomedical Optics Express*, 6(1):86, Jan 2015.

[8] G. Balakrishnan, F. Durand, and J. Guttag. Detecting Pulse from Head Motions in Video. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 3430–3437, Jun 2013.

[9] E. Balestrieri and S. Rapuano. Instruments and Methods for Calibration of Oscillometric Blood Pressure Measurement Devices. *IEEE Transactions on Instrumentation and Measurement*, 59(9):2391–2404, Sep 2010.

[10] G. Beevers, G. Y. Lip, and E. O'Brien. ABC of hypertension: Blood pressure measurement. Part II-conventional sphygmomanometry: technique of auscultatory blood pressure measurement. *The BMJ*, 322(7293):1043–1047, Apr 2001.

[11] K. G. Belani, J. J. Buckley, and M. O. Poliac. Accuracy of radial artery blood pressure determination with the Vasotrac. *Canadian Journal of Anesthesia*, 46(5 Pt 1):488–496, May 1999.

[12] G. A. Beller and S. R. Bergmann. Myocardial perfusion imaging agents: SPECT and PET. *Journal of Nuclear Cardiology*, 11(1):71–86, 2004.

[13] T. Binzoni, A. Vogel, A. H. Gandjbakhche, and R. Marchesini. Detection limits of multi-spectral optical imaging under the skin surface. *Physics in Medicine and Biology*, 53(3):617–636, Feb 2008.

[14] V. Blažek, N. Blanik, C. R. Blažek, M. Paul, C. Pereira, M. Koeny, B. Venema, and S. Leonhardt. Active and Passive Optical Imaging Modality for Unobtrusive Cardiorespiratory Monitoring and Facial Expression Assessment. *Anesthesia & Analgesia*, 124(1):104–119, Jan 2017.

[15] J. C. Bramwell and A.V. Hill. The velocity of the pulse wave in man. *Proceedings of the Royal Society of London B: Biological Sciences*, 93:298–306, 1922.

[16] C. Brüser, C. H. Antink, T. Wartzek, M. Walter, and S. Leonhardt. Ambient and Unobtrusive Cardiorespiratory Monitoring Techniques. *IEEE Reviews in Biomedical Engineering*, 8:30–43, 2015.

[17] A. Bur, H. Herkner, M. Vlcek, C. Woisetschlager, U. Derhaschnig, G. Delle Karth, A. N. Laggner, and M. M. Hirschl. Factors influencing the accuracy of oscillometric blood pressure measurement in critically ill patients. *Critical Care Medicine*, 31(3):793–799, Mar 2003.

[18] B. S. Carlson. Comparison of modern CCD and CMOS image sensor technologies and systems for low resolution imaging. In *Sensors, 2002. Proceedings of IEEE*, volume 1, pages 171–176 vol.1, 2002.

[19] M. Cecconi, D. De Backer, M. Antonelli, R. Beale, J. Bakker, C. Hofer, R. Jaeschke, A. Mebazaa, M. R. Pinsky, J. L. Teboul, J. L. Vincent, and A. Rhodes. Consensus on circulatory shock and hemodynamic monitoring. Task force of the European Society of Intensive Care Medicine. *Intensive Care Medicine*, 40(12):1795–1815, Dec 2014.

[20] W. Chen, T. Kobayashi, S. Ichikawa, Y. Takeuchi, and T. Togawa. Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent calibration. *Medical & Biological Engineering & Computing*, 38(5):569–574, 2000.

[21] H.-M. Cheng, D. Lang, C. Tufanaru, and A. Pearson. Measurement accuracy of non-invasively obtained central blood pressure by applanation tonometry:

A systematic review and meta-analysis. *International Journal of Cardiology*, 167(5):1867–1876, 2013.

[22] A. Chung, X. Y. Wang, R. Amelard, C. Scharfenberger, J. Leong, J. Kulinski, A. Wong, and D. A. Clausi. High-resolution motion-compensated imaging photoplethysmography for remote heart rate monitoring. In *Proceedings of SPIE*, volume 9316, pages 93160A–93160A–5, 2015.

[23] E. Chung, G. Chen, B. Alexander, and M. Cannesson. Non-invasive continuous blood pressure monitoring: a review of current applications. *Frontiers in Medicine*, 7(1):91–101, Mar 2013.

[24] T. M. Coffman. Under pressure: the search for the essential mechanisms of hypertension. *Nature Medicine*, 17(11):1402–1409, 2011.

[25] A. J. Collins. Cardiovascular mortality in end-stage renal disease. *The American Journal of the Medical Sciences*, 325(4):163–167, Apr 2003.

[26] J. K. Cooper. Electrocardiography 100 years ago. Origins, pioneers, and contributors. *New England Journal of Medicine*, 315(7):461–464, Aug 1986.

[27] C. Couchoud, J. Kooman, P. Finne, T. Leivestad, O. Stojceva-Taneva, J. B. Ponikvar, F. Collart, R. Kramar, A. de Francisco, and K. J. Jager. From registry data collection to international comparisons: examples of haemodialysis duration and frequency. *Nephrology Dialysis Transplantation*, 24(1):217–224, Jan 2009.

[28] S. J. Dasgupta and A. B. Gill. Hypotension in the very low birthweight infant: the old, the new, and the uncertain. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, 88(6):F450–454, Nov 2003.

[29] A. Davenport. Can advances in hemodialysis machine technology prevent intradialytic hypotension? *Seminars in Dialysis*, 22(3):231–236, 2009.

[30] D. De Bacquer, G. De Backer, M. Kornitzer, and H. Blackburn. Prognostic value of ECG findings for total, cardiovascular disease, and coronary heart disease death in men and women. *Heart*, 80(6):570–577, 1998.

[31] G. de Haan and V. Jeanne. Robust pulse rate from chrominance-based rPPG. *IEEE Transactions on Biomedical Engineering*, 60(10):2878–2886, Oct 2013.

[32] G. de Haan and A. van Leest. Improved motion robustness of remote-PPG by using the blood volume pulse signature. *Physiological Measurement*, 35(9):1913–1926, Aug 2014.

[33] R. P. Dellinger, M. M. Levy, J. M. Carlet, J. Bion, M. M. Parker, R. Jaeschke, K. Reinhart, D. C. Angus, C. Brun-Buisson, R. Beale, T. Calandra, J.-F. Dhainaut, H. Gerlach, M. Harvey, J. J. Marini, J. Marshall, M. Ranieri, G. Ramsay, J. Sevransky, B. T. Thompson, S. Townsend, J. S. Vender, J. L. Zimmerman, and J.-L. Vincent. Surviving sepsis campaign: International guidelines for management

of severe sepsis and septic shock: 2008. *Critical Care Medicine*, 36(1):296–327, 2008.

[34] M. F. Dempsey and B. Condon. Thermal injuries associated with MRI. *Clinical Radiology*, 56(6):457–465, Jun 2001.

[35] C. Douniama, C. U. Sauter, and R. Couronne. Blood pressure tracking capabilities of pulse transit times in different arterial segments: A clinical evaluation. *Computing in Cardiology*, pages 201–204, Sep 2009.

[36] S. O. Dymling, H. W. Persson, and C. H. Hertz. Measurement of blood perfusion in tissue using doppler ultrasound. *Ultrasound in Medicine & Biology*, 17(5):433–444, 1991.

[37] A. Elvan-Taşpinar, L.A.M.S. Uiterkamp, J.M. Sikkema, M.L. Bots, H.A. Koomans, H.W. Bruinse, and A. Franx. Validation and use of the Finometer for blood pressure measurement in normal, hypertensive and pre-eclamptic pregnancy. *Journal of Hypertension*, 21(11):2053–2060, 2003.

[38] T. B. Fitzpatrick. Soleil et peau. *Journal de Médecine Esthétique*, 2(7):33–34, 1975.

[39] J. Y. Foo and C. S. Lim. Pulse transit time as an indirect marker for variations in cardiovascular related reactivity. *Technology and Health Care*, 14(2):97–108, 2006.

[40] L. A. Geddes, M. H. Voelz, C. F. Babbs, J. D. Bourland, and W. A. Tacker. Pulse transit time as an indicator of arterial blood pressure. *Psychophysiology*, 18(1):71–74, Jan 1981.

[41] H. Gesche, D. Grosskurth, G. Kuchler, and A. Patzak. Continuous blood pressure measurement by using the pulse transit time: comparison to a cuff-based method. *European Journal of Applied Physiology*, 112(1):309–315, Jan 2012.

[42] H. Gray and W. H. Lewis. *Anatomy of the human body.* Lea and Febiger, Philadelphia, 20th edition, 1918. Online edition published in 2000 by Bartleby.com.

[43] A. Guazzi. *Revealing Physiological Changes Through Colour.* PhD thesis, University of Oxford, 2016.

[44] A. Guazzi, M. Villarroel, J. Jorge, J. Daly, M. C. Frise, P. A. Robbins, and L. Tarassenko. Non-contact measurement of oxygen saturation with an RGB camera. *Biomedical Optics Express*, 6(9):3320–3338, Sep 2015.

[45] I. Guelen, B. E. Westerhof, G. L. Van Der Sar, G. A. Van Montfrans, F. Kiemeneij, K. H. Wesseling, and W. J. Bos. Finometer, finger pressure measurements with the possibility to reconstruct brachial pressure. *Blood Pressure Monitoring*, 8(1):27–30, Feb 2003.

[46] O. Hamzaoui, X. Monnet, and J. L. Teboul. Pulsus paradoxus. *European Respiratory Journal*, 42(6):1696–1705, 2013.

[47] C. D. Hanning and J. M. Alexander-Williams. Pulse oximetry: a practical review. *The BMJ*, 311(7001):367–370, Aug 1995.

[48] U. Hassan and M. S. Anwar. Reducing noise by repetition: introduction to signal averaging. *European Journal of Physics*, 31(3):453–465, 2010.

[49] M. J. Hayes and P. R. Smith. Artifact reduction in photoplethysmography. *Applied Optics*, 37(31):7437–7446, Nov 1998.

[50] S. O. Heard, A. Lisbon, I. Toth, and R. Ramasubramanian. An evaluation of a new continuous blood pressure monitoring system in critically ill patients. *Journal of Clinical Anesthesia*, 12(7):509–518, Nov 2000.

[51] B. K. P. Horn and B. B. Schunck. Determining Optical Flow. *Artificial Intelligence*, 17:185–203, 1981.

[52] L. S. Howard, R. A. Barson, B. P. Howse, T. R. McGill, M. E. McIntyre, D. F. O'Connor, and P. A. Robbins. Chamber for controlling end-tidal gas tensions over sustained periods in humans. *Journal of Applied Physiology*, 78(3):1088–1091, Mar 1995.

[53] S. Hu, V. A. Peris, A. Echiadis, J. Zheng, and P. Shi. Development of effective and photoplethysmographic measurement techniques: from contact to non-contact and from point to imaging. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009:6550–6553, 2009.

[54] M. Hülsbusch. *Ein bildgestütztes, funktionelles Verfahren zur optoelektronischen Erfassung der Hautperfusion (A functional imaging technique for optoelectronic assessment of skin perfusion).* PhD thesis, RWTH Aachen University, 2008.

[55] M. Hülsbusch and V. Blažek. Contactless mapping of rhythmical phenomena in tissue perfusion using PPGI. *Proceedings of SPIE*, 4683:110–117, Apr 2002.

[56] K. Humphreys, T. Ward, and C. Markham. A CMOS Camera-Based Pulse Oximetry Imaging System. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, volume 4, pages 3494–3497. IEEE, Jan 2005.

[57] K. Humphreys, T. Ward, and C. Markham. Noncontact simultaneous dual wavelength photoplethysmography: A further step toward noncontact pulse oximetry. *Review of Scientific Instruments*, 78(4):044304, 2007.

[58] D. Iakovlev, V. Dwyer, S. Hu, and V. Silberschmidt. Noncontact blood perfusion mapping in clinical applications. *Proceedings of SPIE*, 9887:988712–988712–11, 2016.

[59] B. P. M. Imholz, G. A. Van Montfrans, J. J. Settels, G. M. A. Van Der Hoeven, J. M. Karemaker, and W. Wieling. Continuous non-invasive blood pressure monitoring: Reliability of Finapres device during the Valsalva manoeuvre. *Cardiovascular Research*, 22(6):390–397, 1988.

[60] G. M. Janelle and N. Gravenstein. An accuracy evaluation of the T-Line Tensymeter (continuous noninvasive blood pressure management device) versus conventional invasive radial artery monitoring in surgical patients. *Anesthesia & Analgesia*, 102(2):484–490, Feb 2006.

[61] I. C. Jeong and J. Finkelstein. Introducing Contactless Blood Pressure Assessment Using a High Speed Video Camera. *Journal of Medical Systems*, 40(4):77, 2016.

[62] S. Jones, W. Jaffe, and R. Alvi. Burns associated with electrocardiographic monitoring during magnetic resonance imaging. *Burns*, 22(5):420–421, 1996.

[63] A. A. Kamshilin, S. Miridonov, V. Teplov, R. Saarenheimo, and E. Nippolainen. Photoplethysmographic imaging of high spatial resolution. *Biomedical Optics Express*, 2(4):996–1006, 2011.

[64] A. A. Kamshilin, E. Nippolainen, I. S. Sidorov, P. V. Vasilev, N. P. Erofeev, N. P. Podolian, and R. V. Romashko. A new look at the essence of the imaging photoplethysmography. *Scientific Reports*, 5:10494, 2015.

[65] A. A. Kamshilin, I. S. Sidorov, L. Babayan, M. A. Volynsky, R. Giniatullin, and O. V. Mamontov. Accurate measurement of the pulse wave delay with imaging photoplethysmography. *Biomedical Optics Express*, 7(12):5138, Dec 2016.

[66] A. A. Kamshilin, V. Teplov, E. Nippolainen, S. Miridonov, and R. Giniatullin. Variability of microcirculation detected by blood pulsation imaging. *PLOS ONE*, 8(2):e57117, Jan 2013.

[67] W. Karlen, H. Gan, M. Chiu, D. Dunsmuir, G. Zhou, G. A. Dumont, and J. M. Ansermino. Improving the accuracy and efficiency of respiratory rate measurements in children using mobile devices. *PLOS ONE*, 9(6):e99266, 2014.

[68] B. Kasravi, J. P. Boehmer, and U. A. Leuenberger. A noninvasive method for estimating cardiac output using lung to finger circulation time of oxygen. *American Journal of Cardiology*, 82(7):915–917, 1998.

[69] E. S. Katz, J. Lutz, C. Black, and C. L. Marcus. Pulse transit time as a measure of arousal and respiratory effort in children with sleep-disordered breathing. *Pediatric Research*, 53(4):580–588, 2003.

[70] B. Kaur, E. Tarbox, M. Cissel, S. Moses, M. Luthra, M. Vaidya, N. Tran, and V. N. Ikonomidou. Remotely detected differential pulse transit time as a stress indicator. In *Independent Component Analyses, Compressive Sampling, Large Data Analyses (LDA), Neural Networks, Biosystems, and Nanoengineering XIII*, volume 9496, page 949604, 2015.

[71] KDOQI. Part 4. Definition and Classification of Stages of Chronic Kidney Disease. *American Journal of Kidney Diseases*, 39(2, Supplement 1):S46–S75, 2002.

[72] A. Khasnis and Y. Lokhandwala. Clinical signs in medicine: pulsus paradoxus. *Journal of Postgraduate Medicine*, 48(1):46–49, 2002.

[73] G. Koklu, J. Ghaye, R. Beuchat, G. De Micheli, Y. Leblebici, and S. Carrara. Quantitative comparison of commercial CCD and custom-designed CMOS camera for biological applications. In *ISCAS 2012 – 2012 IEEE International Symposium on Circuits and Systems*, pages 2063–2066, 2012.

[74] L. Kong, Y. Zhao, L. Dong, Y. Jian, X. Jin, B. Li, Y. Feng, M. Liu, X. Liu, and H. Wu. Non-contact detection of oxygen saturation based on visible light imaging device using ambient light. *Optics Express*, 21(15):17464, July 2013.

[75] D. J. Korteweg. Ueber die Fortpflanzungsgeschwindigkeit des Schalles in elastischen Rhren. *Annalen der Physik*, 241(12):525–542, 1878.

[76] J. Kranjec, S. Beguš, G. Geršak, and J. Drnovšek. Non-contact heart rate and heart rate variability measurements: A review. *Biomedical Signal Processing and Control*, 13(0):102–112, 2014.

[77] M. Kumar, A. Veeraraghavan, and A. Sabharval. DistancePPG: Robust non-contact vital signs monitoring using a camera. *Biomedical Optics Express*, 6(5):1565–1588, Apr 2015.

[78] J. Landgraf, S. H. Wishner, and R. A. Kloner. Comparison of automated oscillometric versus auscultatory blood pressure measurement. *The American Journal of Cardiology*, 106(3):386–388, Aug 2010.

[79] J. D. Lane, L. Greenstadt, D. Shapiro, and E. Rubinstein. Pulse transit time and blood pressure: an intensive analysis. *Psychophysiology*, 20(1):45–49, Jan 1983.

[80] P. Lantelme, C. Mestre, M. Lievre, A. Gressard, and H. Milon. Heart rate: an important confounder of pulse wave velocity assessment. *Hypertension*, 39(6):1083–1087, Jun 2002.

[81] C. Laurent, B. Jonsson, M. Vegfors, and L. G. Lindberg. Non-invasive measurement of systolic blood pressure on the arm utilising photoplethysmography: development of the methodology. *Medical & Biological Engineering & Computing*, 43(1):131–135, Jan 2005.

[82] L. W. Lehman, M. Saeed, D. Talmor, R. Mark, and A. Malhotra. Methods of blood pressure measurement in the ICU. *Critical Care Medicine*, 41(1):34–40, Jan 2013.

[83] T. Lestari, S. Ryll, and A. Kramer. Microbial contamination of manually reprocessed, ready to use ECG lead wire in intensive care units. *GMS Hygiene and Infection Control*, 8(1):Doc07, 2013.

[84] Q. Li and G. G. Belz. Systolic time intervals in clinical pharmacology. *European Journal of Clinical Pharmacology*, 44(5):415–421, 1993.

[85] A. Lima and J. Bakker. Noninvasive monitoring of peripheral perfusion. *Intensive Care Medicine*, 31(10):1316–1326, Oct 2005.

[86] H. Liu, K. Ivanov, Y. Wang, and L. Wang. A novel method based on two cameras for accurate estimation of arterial oxygen saturation. *BioMedical Engineering OnLine*, 14(1):52, Dec 2015.

[87] B. D. Lucas and T. Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence – Volume 2*, IJCAI'81, pages 674–679, San Francisco, CA, USA, 1981. Morgan Kaufmann Publishers Inc.

[88] F. C. Luft. Renal disease as a risk factor for cardiovascular disease. *Basic Research in Cardiology*, 95 Suppl 1:I72–76, 2000.

[89] C. MacEwen. *Can data fusion techniques predict adverse physiological events during haemodialysis?* PhD thesis, University of Oxford, 2016.

[90] Y. Maeda, M. Sekine, and T. Tamura. The advantages of wearable green reflected photoplethysmography. *Journal of Medical Systems*, 35(5):829–834, Oct 2011.

[91] R. Maestri, G. D. Pinna, E. Robbi, S. Capomolla, and M. T. La Rovere. Noninvasive measurement of blood pressure variability: Accuracy of the Finometer monitor and comparison with the Finapres device. *Physiological Measurement*, 26(6):1125–1136, 2005.

[92] E. Mambelli, E. Mancini, and A. Santoro. A continuous and non-invasive arterial pressure monitoring system in dialysis patients. *Nephron Clinical Practice*, 107(4):c170–176, 2007.

[93] K. Matthys and P. Verdonck. Development and modelling of arterial applanation tonometry: a review. *Technology and Health Care*, 10(1):65–76, 2002.

[94] B. M. McCarthy, B. O'Flynn, and A. Mathewson. An Investigation of Pulse Transit Time as a Non-Invasive Blood Pressure Measurement Method. *Journal of Physics: Conference Series*, 307(1):012060, 2011.

[95] C. M. McEniery, J. R. Cockcroft, M. J. Roman, S. S. Franklin, and I. B. Wilkinson. Central blood pressure: current evidence and clinical importance. *European Heart Journal*, 35(26):1719–1725, Jul 2014.

[96] R. Mehrotra, Y. Chiu, K. Kalantar-Zadeh, J. Bargman, and E. Vonesh. Similar outcomes with hemodialysis and peritoneal dialysis in patients with end-stage renal disease. *Archives of Internal Medicine*, 171(2):110–118, 2011.

[97] G. F. Mitchell, H. Parise, E. J. Benjamin, M. G. Larson, M. J. Keyes, J. A. Vita, R. S. Vasan, and D. Levy. Changes in arterial stiffness and wave reflection with advancing age in healthy men and women: the Framingham Heart Study. *Hypertension*, 43(6):1239–1245, Jun 2004.

[98] A. I. Moens. *Over de voortplantingssnelheid van den pols*. PhD thesis, University of Leiden, 1877.

[99] A. I. Moens. *Die Pulskurve*. Brill, E. J., 1878.

[100] E. Monte-Moreno. Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques. *Artificial Intelligence in Medicine*, 53(2):127–138, Oct 2011.

[101] G. B. Moody and L. H. Lehman. Predicting acute hypotensive episodes: The 10th annual PhysioNet/Computers in Cardiology Challenge. *Computing in Cardiology*, pages 541–544, Sep 2009.

[102] D. Moran, Y. Epstein, G. Keren, A. Laor, J. Sherez, and Y. Shapiro. Calculation of mean arterial pressure during exercise as a function of heart rate. *Applied Human Science*, 14(6):293–295, Nov 1995.

[103] J. L. Moran, J. V. Peter, P. J. Solomon, B. Grealy, T. Smith, W. Ashforth, M. Wake, S. L. Peake, and A. R. Peisach. Tympanic temperature measurements: are they reliable in the critically ill? A clinical study of measures of agreement. *Critical Care Medicine*, 35(1):155–164, Jan 2007.

[104] J. Muehlsteff, X. L. Aubert, and M. Schuett. Cuffless estimation of systolic blood pressure for short effort bicycle tests: the prominent role of the pre-ejection period. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1:5088–5092, 2006.

[105] R. Mukkamala, J. O. Hahn, O. T. Inan, L. K. Mestha, C. S. Kim, H. Toreyin, and S. Kyal. Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice. *IEEE Transactions on Biomedical Engineering*, 62(8):1879–1901, Aug 2015.

[106] K. Murakami, M. Yoshioka, and J. Ozawa. Non-contact Pulse Transit Time Measurement using Imaging Camera, and its Relation to Blood Pressure. In *14th IAPR International Conference on Machine Vision Applications (MVA)*, pages 414–417, Tokyo, Japan, 2015.

[107] T. Nakamura, K. Fukuda, K. Hayakawa, I. Aoki, K. Matsumoto, T. Sekine, H. Ueda, and Y. Shimizu. Mechanism of burn injury during magnetic resonance imaging (MRI)–simple loops can induce heat injury. *Frontiers of Medical and Biological Engineering*, 11(2):117–129, 2001.

[108] K. Nakano, R. Satoh, A. Hoshi, R. Matsuda, H. Suzuki, and I. Nishidate. Non-contact imaging of venous compliance in humans using an RGB camera. In *Biomedical Imaging and Sensing Conference*, 2014.

[109] J. A. Nijboer, J. C. Dorlas, and H. F. Mahieu. Photoelectric plethysmography–some fundamental aspects of the reflection and transmission method. *Clinical Physics and Physiological Measurement*, 2(3):205–215, Aug 1981.

[110] M. Nitzan, B. Khanokh, and Y. Slovik. The difference in pulse transit time to the toe and finger measured by photoplethysmography. *Physiological Measurement*, 23(1):85–93, Feb 2002.

[111] M. Nitzan, A. Patron, Z. Glik, and A.T. Weiss. Automatic noninvasive measurement of systolic blood pressure using photoplethysmography. *BioMedical Engineering OnLine*, 8:28, 2009.

[112] E. O'Brien, B. Waeber, G. Parati, J. Staessen, and M. G. Myers. Blood pressure measuring devices: recommendations of the European Society of Hypertension. *The BMJ*, 322(7285):531–536, Mar 2001.

[113] National Institute of Diabetes, Digestive, and Kidney Diseases. Diagram of haemodialysis. http://www.niddk.nih.gov/health-information/health-topics/kidney-disease/hemodialysis. Accessed 2017-01-06.

[114] B. F. Palmer and W. L. Henrich. Recent advances in the prevention and management of intradialytic hypotension. *Journal of the American Society of Nephrology*, 19(1):8–11, 2008.

[115] B. F. Palmer and W. L. Henrich. Recent advances in the prevention and management of intradialytic hypotension. *Journal of the American Society of Nephrology*, 19(1):8–11, 2008.

[116] J. Pan and W. J. Tompkins. A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32(3):230–236, Mar 1985.

[117] G. Parati, R. Casadei, A. Groppelli, M. Di Rienzo, and G. Mancia. Comparison of finger and intra-arterial blood pressure monitoring at rest and during laboratory testing. *Hypertension*, 13(6 Pt 1):647–655, Jun 1989.

[118] G. Parati, G. Ongaro, G. Bilo, F. Glavina, P. Castiglioni, M. Di Rienzo, and G. Mancia. Non-invasive beat-to-beat blood pressure monitoring: New developments. *Blood Pressure Monitoring*, 8(1):31–36, 2003.

[119] M. K. Park, S. W. Menard, and C. Yuan. Comparison of auscultatory and oscillometric blood pressures. *Archives of Pediatrics and Adolescent Medicine*, 155(1):50–53, 2001.

[120] A. Patzak, Y. Mendoza, H. Gesche, and M. Konermann. Continuous blood pressure measurement using the pulse transit time: Comparison to intra-arterial measurement. *Blood Pressure*, 24(4):217–221, 2015.

[121] R. A. Payne, C. N. Symeonides, D. J. Webb, and S. R. Maxwell. Pulse transit time measured from the ECG: an unreliable marker of beat-to-beat blood pressure. *Journal of Applied Physiology*, 100(1):136–141, Jan 2006.

[122] M. E. V. Petersen, T. R. Williams, and R. Sutton. A comparison of non-invasive continuous finger blood pressure measurement (Finapres) with intra-arterial pressure during prolonged head-up tilt. *European Heart Journal*, 16(11):1647–1654, 1995.

[123] J. Peňáz. Photoelectric measurement of blood pressure, volume, and flow in the finger. *Digest of the 10th International Conference on Medical and Biomedical Engineering; Dresden, Germany*, page 104, 1973.

[124] 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*, 18(10):10762–10774, May 2010.

[125] M. Z. Poh, D. J. McDuff, and R. W. Picard. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Transactions on Biomedical Engineering*, 58(1):7–11, Jan 2011.

[126] C. C. Poon and Y. T. Zhang. Cuff-less and noninvasive measurements of arterial blood pressure by pulse transit time. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 6:5877–5880, 2005.

[127] S. Prahl. Tabulated Molar Extinction Coefficient for Hemoglobin in Water (Oregon Medical Laser Center). http://omlc.org/spectra/hemoglobin/summary.html, 2008. Accessed 2017-01-06.

[128] M. Prgomet, M. Cardona-Morrell, M. Nicholson, R. Lake, J. Long, J. Westbrook, J. Braithwaite, and K. Hillman. Vital signs monitoring on general wards: clinical staff perceptions of current practices and the planned introduction of continuous monitoring technology. *International Journal for Quality in Health Care*, pages 1–7, 2016.

[129] S. S. Prichard. Comorbidities and their impact on outcome in patients with end-stage renal disease. *Kidney International*, 57(S74):S100–S104, Jan 2000.

[130] M. Ravera, M. Re, L. Deferrari, S. Vettoretti, and G. Deferrari. Importance of blood pressure control in chronic kidney disease. *Journal of the American Society of Nephrology*, 17(4 Suppl 2):98–103, Apr 2006.

[131] U. Rubins, V. Upmalis, O. Rubenis, D. Jakovels, and J. Spigulis. *15th Nordic-Baltic Conference on Biomedical Engineering and Medical Physics (NBC 2011): 14-17 June 2011, Aalborg, Denmark*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[132] J. C. Ruiz-Rodríguez, A. Ruiz-Sanmartín, V. Ribas, J. Caballero, A. García-Roche, J. Riera, X. Nuvials, M. de Nadal, O. de Sola-Morales, J. Serra, and J. Rello. Innovative continuous non-invasive cuffless blood pressure monitoring based on photoplethysmography technology. *Intensive Care Medicine*, 39(9):1618–1625, 2013.

[133] R. L. Sacco, E. J. Benjamin, J. P. Broderick, M. Dyken, J. D. Easton, W. M. Feinberg, L. B. Goldstein, P. B. Gorelick, G. Howard, S. J. Kittner, T. A. Manolio, J. P. Whisnant, and P. A. Wolf. American Heart Association Prevention Conference. IV. Prevention and Rehabilitation of Stroke. Risk factors. *Stroke*, 28(7):1507–1517, Jul 1997.

[134] W. M. Saltzman. *Biomedical Engineering: Bridging Medicine and Technology*. Cambridge University Press, 2009.

[135] B. Saugel, F. Fassio, A. Hapfelmeier, A. S. Meidert, R. M. Schmid, and W. Huber. The T-Line TL-200 system for continuous non-invasive blood pressure measurement in medical intensive care unit patients. *Intensive Care Medicine*, 38(9):1471–1477, Sep 2012.

[136] B. Scheer, A. Perel, and U. J. Pfeiffer. Clinical review: complications and risk factors of peripheral arterial catheters used for haemodynamic monitoring in anaesthesia and Intensive Care Medicine. *Critical Care*, 6(3):199–204, Jun 2002.

[137] A. E. Schutte, H. W. Huisman, J. M. Van Rooyen, N. T. Malan, and R. Schutte. Validation of the Finometer device for measurement of blood pressure in black women. *Journal of Human Hypertension*, 18(2):79–84, 2004.

[138] C. G. Scully, J. Lee, J. Meyer, A. M. Gorbach, D. Granquist-Fraser, Y. Mendelson, and K. H. Chon. Physiological parameter monitoring from optical recordings with a mobile phone. *IEEE Transactions on Biomedical Engineering*, 59(2):303–306, Feb 2012.

[139] A. Secerbegovic, J. Bergsland, P. S. Halvorsen, N. Suljanovic, A. Mujcic, and I. Balasingham. Blood pressure estimation using video plethysmography. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 461–464, 2016.

[140] D. Shao, Y. Yang, C. Liu, F. Tsow, H. Yu, and N. Tao. Noncontact monitoring breathing pattern, exhalation flow rate and pulse transit time. *IEEE Transactions on Biomedical Engineering*, 61(11):2760–7, Nov 2014.

[141] R. P. Smith, J. Argod, J.-L. Pépin, and P. A. Lévy. Pulse transit time: An appraisal of potential clinical applications. *Thorax*, 54(5):452–457, 1999.

[142] J. R. Sowers, M. Epstein, and E. D. Frohlich. Diabetes, hypertension, and cardiovascular disease: an update. *Hypertension*, 37(4):1053–1059, Apr 2001.

[143] J. A. Staessen, E. T. O'Brien, L. Thijs, and R. H. Fagard. Modern approaches to blood pressure measurement. *Occupational and Environmental Medicine*, 57(8):510–520, Aug 2000.

[144] A. Steptoe, H. Smulyan, and B. Gribbin. Pulse Wave Velocity and Blood Pressure Change: Calibration and Applications. *Psychophysiology*, 13(5):488–493, 1976.

[145] A. Stockman and L. T. Sharpe. The spectral sensitivities of the middle- and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, 40(13):1711–1737, 2000.

[146] O. Such. Near infrared CCD imaging of hemondynamics. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 18:2105–2106, 1996.

[147] O. Such, S. Acker, and V. Blažek. Mapped hemondynamic data acquisition by near infrared CCD imaging. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 19:637–639, Nov 1997.

[148] Y. Sun, C. Papin, V. Azorin-Peris, R. Kalawsky, S. Greenwald, and S. Hu. Use of ambient light in remote photoplethysmographic systems: comparison between a high-performance camera and a low-cost webcam. *Journal of Biomedical Optics*, 17(3):037005, Mar 2012.

[149] Y. Sun and N. Thakor. Photoplethysmography Revisited: From Contact to Noncontact, From Point to Imaging. *IEEE Transactions on Biomedical Engineering*, 63(3):463–477, Mar 2016.

[150] S. Suzuki and K. Oguri. Cuffless and non-invasive Systolic Blood Pressure estimation for aged class by using a Photoplethysmograph. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008:1327–1330, 2008.

[151] S. Suzuki and K. Oguri. Cuffless blood pressure estimation by error-correcting output coding method based on an aggregation of AdaBoost with a photoplethysmograph sensor. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009:6765–6768, 2009.

[152] C. Takano and Y. Ohta. Heart rate measurement based on a time-lapse image. *Medical Engineering & Physics*, 29(8):853–857, Oct 2007.

[153] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. A. Clifton, and C. Pugh. Non-contact video-based vital sign monitoring using ambient light and auto-regressive models. *Physiological Measurement*, 35(5):807–831, May 2014.

[154] J.-L. Teboul, B. Saugel, M. Cecconi, D. De Backer, C. K. Hofer, X. Monnet, A. Perel, M. R. Pinsky, D. A. Reuter, A. Rhodes, P. Squara, J.-L. Vincent, and T. W. Scheeren. Less invasive hemodynamic monitoring in critically ill patients. *Intensive Care Medicine*, pages 1–10, 2016.

[155] V. Teplov, E. Nippolainen, A. A. Makarenko, R. Giniatullin, and A. A. Kamshilin. Ambiguity of mapping the relative phase of blood pulsations. *Biomedical Optics Express*, 5(9):3123–3139, Sep 2014.

[156] J. G. Thomas. A method for continuously indicating blood pressure. *Journal of Physiology*, 129(3):75–76, Sep 1955.

[157] C Tomasi and T. Kanade. Detection and Tracking of Point Features. Technical report, International Journal of Computer Vision, 1991.

[158] A. Trumpp, J. Schell, H. Malberg, and S. Zaunseder. Vasomotor assessment by camera-based photoplethysmography. *Current Directions in Biomedical Engineering*, 2(1):199–202, 2016.

[159] M. van Gastel, S. Stuijk, and G. de Haan. New principle for measuring arterial blood oxygenation, enabling motion-robust remote monitoring. *Scientific Reports*, 6:38609, Dec 2016.

[160] W. Verkruysse, M. Bartula, E. Bresch, M. Rocque, M. Meftah, and I. Kirenko. Calibration of Contactless Pulse Oximetry. *Anesthesia & Analgesia*, 124(1):136–145, Jan 2017.

[161] W. Verkruysse, L. O. Svaasand, and J. S. Nelson. Remote plethysmographic imaging using ambient light. *Optics Express*, 16(26):21434–21445, Dec 2008.

[162] M. Villarroel, A. Guazzi, J. Jorge, S. Davis, P. Watkinson, G. Green, A. Shenvi, K. McCormick, and L. Tarassenko. Continuous non-contact vital sign monitoring in neonatal intensive care unit. *Healthcare Technology Letters*, 1(3):87–91, Sep 2014.

[163] P. Viola and M. 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–I–518 vol.1, 2001.

[164] M. A. Volynsky, O. V. Mamontov, I. S. Sidorov, and A. A. Kamshilin. Pulse wave transit time measured by imaging photoplethysmography in upper extremities. *Journal of Physics: Conference Series*, 737:012053, 2016.

[165] M. Ward and J. A. Langton. Blood Pressure Measurement. *Continuing Education in Anaesthesia, Critical Care and Pain*, 7(4):122–126, 2007.

[166] M. G. Weaver, M. K. Park, and D.-H. Lee. Differences in blood pressure levels obtained by auscultatory and oscillometric methods. *The American Journal of Diseases of Children*, 144(8):911–914, 1990.

[167] T. Wibmer, K. Doering, C. Kropf-Sanchen, S. Rudiger, I. Blanta, K. M. Stoiber, W. Rottbauer, and C. Schumann. Pulse transit time and blood pressure during cardiopulmonary exercise tests. *Physiological Research*, 63(3):287–296, 2014.

[168] F. P. Wieringa, F. Mastik, R.H. Boks, A Visscher, A. J. J. C. Bogers, and A. F. W. Van der Steen. In vitro demonstration of an SpO2-camera. *Computing in Cardiology*, pages 749–751, Sep 2007.

[169] F. P. Wieringa, F. Mastik, and A. F. van der Steen. Contactless multiple wavelength photoplethysmographic imaging: a first step toward "SpO2 camera" technology. *Annals of Biomedical Engineering*, 33(8):1034–1041, Aug 2005.

[170] M. C. Wilkins. Residual bacterial contamination on reusable pulse oximetry sensors. *Respiratory Care*, 38(11):1155–1160, Nov 1993.

[171] E. S. Winokur, D. D. He, and C. G. Sodini. A wearable vital signs monitor at the ear for continuous heart rate and pulse transit time measurements. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012:2724–2727, 2012.

[172] E. C. Wong, R. B. Buxton, and L. R. Frank. Implementation of quantitative perfusion imaging techniques for functional brain mapping using pulsed arterial spin labeling. *NMR in Biomedicine*, 10(4-5):237–249, 1997.

[173] H. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman. Eulerian Video Magnification for Revealing Subtle Changes in the World. *ACM Transactions on Graphics (Proceedings SIGGRAPH 2012)*, 31(4), 2012.

[174] T. Wu. PPGI: New development in noninvasive and contactless diagnosis of dermal perfusion using near infrared light. *Journal of the Gesellschaft Chinesischer Physiker in Deutschland e.V.*, 7(1):17–24, 2003.

[175] T. Wu, V. Blažek, and H. J. Schmitt. Photoplethysmography imaging: a new noninvasive and noncontact method for mapping of the dermal perfusion changes. *Proceedings of SPIE*, 4163:62–70, Nov 2000.

[176] A. Yamashina, H. Tomiyama, T. Arai, Y. Koji, M. Yambe, H. Motobe, Z. Glunizia, Y. Yamamoto, and S. Hori. Nomogram of the relation of brachial-ankle pulse wave velocity with blood pressure. *Hypertension Research*, 26(10):801–806, Oct 2003.

[177] J. Yang, B. Guthier, and A. El Saddik. Estimating two-dimensional blood flow velocities from videos. *2015 IEEE International Conference on Image Processing (ICIP)*, pages 3768–3772, 2015.

[178] G. Zhang, D. Xu, N. B. Olivier, and R. Mukkamala. Pulse arrival time is not an adequate surrogate for pulse transit time in terms of tracking diastolic pressure. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011:6462–6464, 2011.

[179] G. R. Zubieta-Calleja, G. Zubieta-Castillo, P.-E. Paulev, and L. Zubieta-Calleja. Non-invasive measurement of circulation time using pulse oximetry during breath holding in chronic hypoxia. *Journal of Physiology and Pharmacology*, 56(SUPPL. 4):251–256, 2005.