ABSTRACT

MOTION-RESISTANT PULSE OXIMETRY

by
Ludvik Alkhoury

The measurement of vital signs – such as peripheral capillary oxygen saturation (SpO_2) and heart rate (HR) levels – by a pulse oximeter is studied. The pulse oximeter is a non-invasive device that measures photoplethysmography (PPG) signals and extracts vital signs from them. However, the quality of the PPG signal measured by oximetry sensors is known to deteriorate in the presence of substantial human and sensor movements contributing to the measurement noise. Methods to suppress such noise from PPG signals measured by an oximeter and to calculate the associated vital signs with high accuracy even when the wearer is under substantial motion are presented in this study.

The spectral components of the PPG waveform are known to appear at a fundamental frequency that corresponds to the participant's HR and at its harmonics. To match this signal, a time-varying *comb filter* tuned to the participant's HR is employed. The filter captures the HR components and eliminates most other artifacts. A significant improvement in the accuracy of SpO_2 calculated from the comb-filtered PPG signals is observed, when tested on data collected from human participants while they are at rest and while they are exercising.

In addition, an architecture that integrates SpO_2 levels from multiple PPG channels mounted on different parts of the wearer's arm is presented. The SpO_2 levels are integrated using a Kalman filter that uses past measurements and modeling of the SpO_2 dynamics to attenuate the effect of the motion artifacts. Again, data collected from human participants while they are at rest and while they are exercising are used. The integrated SpO_2 levels are shown to be more accurate and reliable than those calculated from individual channels.

Motion-resistant algorithms typically require an additional noise reference signal to produce high quality vital signs such as HR. A framework that employs PPG sensors only – one in the green and one in the infrared spectrum – to compute high quality HR levels is developed. Our framework is tested on experimental data collected from human participants while at rest and while running at various speeds. Our “PPG-only” framework generates HR levels with high accuracy and low computational complexity as compared to leading HR calculation methods in the literature that require the availability of a noise reference signal.

The methods for SpO_2 and HR calculation presented in this study are desirable since (1) they yield high accuracy in estimating vital signs under substantial level of motion artifacts and (2) they are computationally efficient, (and therefore are capable to be implemented in wearable devices).

MOTION-RESISTANT PULSE OXIMETRY

by
Ludvik Alkhoury

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering

Helen and John C. Hartmann
Department of Electrical and Computer Engineering

May 2023

APPROVAL PAGE

MOTION-RESISTANT PULSE OXIMETRY

Ludvik Alkhoury

Moshe Kam, Dissertation Advisor Date
Professor of Electrical and Computer Engineering, NJIT

Sui-hoi Hou, Committee Member Date
Professor of Electrical and Computer Engineering, NJIT

Nirwan Ansari, Committee Member Date
Distinguished Professor of Electrical and Computer Engineering, NJIT

Ali Abdi, Committee Member Date
Professor of Electrical and Computer Engineering, NJIT

Xuan Liu, Committee Member Date
Associate Professor of Electrical and Computer Engineering, NJIT

Saikat Pal, Committee Member Date
Assistant Professor of Biomedical Engineering, NJIT

BIOGRAPHICAL SKETCH

Author: Ludvik Alkhoury

Degree: Doctor of Philosophy

Date: May 2023

Undergraduate and Graduate Education:

- Doctor of Philosophy in Electrical Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2023

- Bachelor of Science in Electrical Engineering,
  University of Balamand, El-Koura, Lebanon, 2016

Major: Electrical Engineering

Presentations and Publications:

L. Alkhoury, J. Choi, V. D. Chandran, G. De Carvalho, S. Pal, and M. Kam, "Dual Wavelength Photoplethysmography Framework for Heart Rate Calculation," *Sensors,* vol. 22, no. 24, 9955, 2022

J. Choi, L. Alkhoury, L. F. Urbano, P. Masson, M. VerMilyea, and M. Kam, "An Assessment Tool for Computer-Assisted Semen Analysis (CASA) Algorithms," *Scientific Reports,* vol. 12, no. 1, pp. 1-17, 2022

L. Alkhoury, J. Choi, C. Wang, S. Mahoney, B. S. Shender, L. Hrebien, and M. Kam, "Mitigation of Motion Arficats in Pulse Oximetry Through Redundant Sensors," *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*. Poznan, Poland: IEEE, 2022

L. Alkhoury, J. Choi, J. Bopp, M. Vermilyea, and M. Kam, "A Method of Digital Washing of Fresh Semen Samples." *Reproductive BioMedicine Online,* vol. 45, pp. e6-e7, 2022

L. Alkhoury "Heart-Rate Tuned Comb Filters for Processing Photoplethysmogram (PPG) Signals in Pulse Oximetry," *Oral Presentation, Dana Knox NJIT*, Newark, New Jersey, USA, 2021

L. Alkhoury, J. Choi, C. Wang, A. Rajasekar, S. Acharya, S. Mahoney, B. S. Shender, L. Hrebien, and M. Kam, "Heart-Rate Tuned Comb Filters for Processing Photoplethysmogram (PPG) Signals in Pulse Oximetry," *Journal of Clinical Monitoring and Computing,* vol. 35, no. 4, pp. 797-813, 2021

L. Alkhoury, J. Sodhi, and A. D. Borgaonkar, "Re-Engineering a Mini-Drone as a Project for First-Year Engineering Students," First-Year Engineering Experience (FYEE) Conference. Online: American Society of Engineering Education (ASEE), 2020

L. Alkhoury, J. Sodhi, A. D. Borgaonkar, and N. Bosca, "Work in Progress-Integration of Voice Technology into the First-Year Engineering Curriculum," *Virtual Annual Conference Content Access*. Onlune: ASEE, 2020

A. D. Borgaonkar, J. Sodhi, and L. Alkhoury, "Integrating Makerspace in First-Year Engineering Curriculum," *American Society of Engineering Education (ASEE) First-Year Engineering Experience Conference*. Penn State University, Pennsylvania, USA: ASEE, 2019

This dissertation is dedicated to my father, Toni Alkhoury.

ACKNOWLEDGMENT

As the result of endless hours of hard work and dedication, I am very proud to announce the completion of this dissertation. This experience has marked a watershed in my life. However, this would have never been possible without the help and support of some special people that I would like to take this opportunity to express my sincere gratitude to.

First, I would like to thank my research advisor, Dr. Moshe Kam for his constant support, guidance, mentorship, professionalism and work ethic, and encouragement throughout my years of study at NJIT. Dr. Kam was never a conventional advisor; he always directed me in the proper direction and led by example.

Second, I would like to thank my dissertation committee members, Dr. Sui-hoi Hou, Dr. Nirwan Ansari, Dr. Ali Abdi, Dr. Xuan Liu, and Dr. Saikat Pal for their advice and support. I would also like to thank Dr. Leonid Hrebien for his mentorship and encouragement.

I am also grateful to my lab colleagues Weiqiang Dong, Arjun Rajasekhar, Vishnu D. Chandran, and Gabriela De Carvalho for their valuable contribution to my research. A special thanks to my colleague and friend, JiWon Choi for continuously sharing his knowledge with me and for his high level of professionalism.

Moreover, I would like to thank the U.S. Office of Naval Research (ONR) for partially sponsoring the work presented in this thesis (under Grant No. N68335-16-C-0085).

Through my journey, I have been thankful for the endless love and support of my wife, Alessia Nestico; my mother, Elham Marcos; my brothers, Keven and Gustave Alkhoury; and my families: Alkhoury, Marcos, and Nestico who were always there when I needed them.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF TABLES
(Continued)

Table **Page**

LIST OF FIGURES

LIST OF FIGURES
(Continued)

LIST OF FIGURES
(Continued)

LIST OF FIGURES
(Continued)

LIST OF FIGURES
(Continued)

LIST OF SYMBOLS

Latin

A_i	Amplitude of the i^{th} sinusoidal component of the synthetic PPG signal
$BPM_{DWL}(.)$	Heart rate calculated using DWL method in beats per minute (BPM)
$BPM_{GT}(.)$	Heart rate ground truth in BPM
c_j	Concentration of the j^{th} skin layer
c_0, c_1, and c_3	Constants used in the calculation of $\Delta_n(.)$, $\Delta_m(.)$, and $\Delta_w(.)$ in Chapter 5
$\boldsymbol{\mathcal{C}}(.)$	Substance's concentration (for example, $\boldsymbol{\mathcal{C}}(Hb)$ is the concentration of Hemoglobin)
d_j	Reflective light path length of the j^{th} skin layer
D	Number of reference signals employed in DST used for computation complexity calculation
D_0	Global Detector of Chapter 5 that decides if the green PPG signal is contaminated with appreciable noise
D_1	Local Detector 1 of Chapter 5 that uses the number of dominant peaks in the frequency spectrum of the PPG signal to determine if appreciable noise is present
D_2	Local Detector 2 of Chapter 5 that uses the power of the PPG signal to determine if appreciable noise is present
D_3	Local Detector 3 of Chapter 5 that uses the Pearson Correlation between the green and IR PPG signals to determine if appreciable noise is present
$\boldsymbol{\mathcal{E}}(., \lambda)$	Extinction coefficients at wavelength λ
$f_{c,BW}$	Comb filter 3 dB bandwidth
f_{n_i}	i^{th} discrete noise frequency components in the set $\mathbb{N}_{noise}$

f_s	Sampling frequency
f_0	PPG signal's fundamental frequency that corresponds to the person's heart rate
F_1, F_2, F_3, F_A, and F_B	Points of the frequency spectrum of green and IR PPG signals used for illustration of the DWL framework in Chapter 5
$G_{PPG,\,i}$	PPG signal obtained at the end of the i^{th} stage of the C-ANC of Chapter 5
$H_c(z)$	Comb filter transfer function
$HR_{cand}(.)$	Heart rate candidate obtained from the spectrum of the green PPG signal of Chapter 5
$HR_{GT}(.)$	Heart rate ground truth
HR_{max}	Maximum heart rate value based on the Karvonen formula (Karvonen and Vuoarimaa (1988))
$HR_{pred}(.)$	Heart rate prediction obtained from the HR long-term (LT) trend using STL method of Chapter 5
HR_{rest}	Heart rate value at rest
$\hat{HR}^{(Z)}(.)$	Average of the most recent Z estimates of the heart rate
$\hat{HR}(.)$	Heart rate estimate
I_i	Intensity of the incident light
$I_{max,\lambda}$	Maximum light intensity at wavelength λ at the diastole
$I_{min,\lambda}$	Minimum light intensity at wavelength λ at the systole
I_o	Intensity of the reflected light
J	Number of different type of skin layers
$K(.)$	Kalman Gain of the SpO_2 Kalman filter model

K_c	Comb filter order
N_{infra}	Noise component of the infrared PPG signal
N_p	The number of dominant peaks threshold calculated in the frequency spectrum of the green PPG signal used by D_1
N_n	Number of elements in the set $\mathbb{N}_{noise}$
$\mathbb{N}_{noise}$	Set of motion noise frequency components obtained in Chapter 5
$\boldsymbol{\mathcal{O}}(.)$	Order of complexity of an algorithm
$P(.)$	State prediction variance of the SpO_2 Kalman filter model
$P_{desired\,signal}$	Power of the participant's heart rate component in a PPG signal
$P^G(.)$	Power of the green PPG signal
P^{ref}	Reference power threshold obtained from a noise-free green PPG signal and used by D_2
P_{noise}	Power of the motion artifact components in a PPG signal
Q_1	25^{th} percentile of data distribution
Q_3	75^{th} percentile of data distribution
r	Arbitrary optical density ratio
r_{Noise}	Ratio of the noise component of the red and infrared PPG signals
r_{SpO_2}	Optical density ratio that corresponds to the correct SpO_2 value
R^2	Coefficient of determination of the linear regression between ground truth heart rate and DWL heart rate
$\boldsymbol{\mathcal{R}}$	Optical density ratio
$S(.)$	Innovation variance of the SpO_2 Kalman filter model

$SpO_{2_{cal}}(.)$	SpO_2 level calculated using one of the tested methods (RoI or DST-based algorithm, with and without comb filtering)
$SpO_{2_{GT}}(.)$	SpO_2 ground truth
S_{infra}	Desired signal component of the infrared PPG signal
T	Number of taps used in the adaptive filter
$v(.)$	Measurement noise used for SpO_2 Kalman filter model
$w(.)$	State noise used for SpO_2 Kalman filter model
$x(.)$	SpO_2 state variable of the SpO_2 Kalman filter model
$y(.)$	Innovation of the SpO_2 Kalman filter model
$z(.)$	SpO_2 measurement of the SpO_2 Kalman filter model
$z_1(.)$	SpO_2 measurement from Channel A
$z_2(.)$	SpO_2 measurement from Channel B

Greek

α	Significance level for t-test
β_c	Comb filter parameter
γ_c	Comb filter parameter
$\delta_e(.)$	Absolute difference between heart rate candidate of the current time step and heart rate estimated from the previous time step of Chapter 5
$\delta^{R-R}(.)$	Average time difference between each two consecutive R peaks of an ECG signal present within the 8-second-long window
$\Delta(.)$	Absolute different between HR calculated using ground truth and the other tested methods in BPM
Δl	Arterial radius expansion during the systole and diastole
$\Delta_n(.)$	Narrow search range used in the motion-artifact frequency components identification process of Subsection 5.4.3
$\Delta_m(.)$	Medium search range used in the motion-artifact frequency components identification process of Subsection 5.4.3
$\Delta_w(.)$	Wide search range used in the heart rate estimation process of Subsection 5.4.5
$\Delta_w^+(.)$	Extended wide search range used in the heart rate estimation process of Subsection 5.4.5
ϵ_1	Constant used in the indicator function, $\mathbb{1}$, of Chapter 5; $\mathbb{1}$ is 1 if $\Delta(.) < \epsilon_1$ and 0 otherwise
ζ	A confidence level used in the Kalman filter model of Chapter 4 to accept/reject the SpO_2 levels calculated from PPG signals of Channels A and B
η	Constant used in the DWL framework of Chapter 5

κ	Constant used by Local Detector D_2 of Chapter 5
λ	Light wavelength
$\mu_{\lambda,j}$	Absorption coefficient of light of wavelength λ by the j^{th} skin layer
$\nu_{1,A}$	SpO_2 Mean Absolute Error (MAE) for all participants using RoI without comb filter from PPG signals of Channel A in Chapter 3
$\nu_{2,A}$	SpO_2 MAE for all participants using RoI with comb filter from PPG signals of Channel A in Chapters 3 and 4
$\nu_{2,B}$	SpO_2 MAE for all participants using RoI with comb filter from PPG signals of Channel B in Chapter 4
$\nu_{3,A}$	SpO_2 MAE for all participants using DST without comb filter from PPG signals of Channel A in Chapter 3
$\nu_{4,A}$	SpO_2 MAE for all participants using DST with comb filter from PPG signals of Channel A in Chapters 3 and 4
$\nu_{4,B}$	SpO_2 MAE for all participants using DST with comb filter from PPG signals of Channel B in Chapter 4
$\nu_{2,KF}$	MAE value of SpO_2 obtained for all participants by integrating SpO_2 calculated using RoI with comb filter PPG signals of Channels A and B in Chapter 4
$\nu_{4,KF}$	MAE value of SpO_2 obtained for all participants by integrating SpO_2 calculated using DST with comb filter PPG signals of Channels A and B in Chapter 4
$\rho_{green,\,IR}$	Pearson correlation between green and IR PPG signals
σ_v^2	Variance of the measurement noise $v(k)$ used for SpO_2 Kalman filter model
σ_w^2	Variance of the state noise $w(k)$ used for SpO_2 Kalman filter model
τ	Constant used in the DWL framework of Chapter 5

$\upsilon^{G}_{SNR,\,palm}$	SNR mean value in dB of the green PPG signals measured from the participants' palm
$\upsilon^{IR}_{SNR,\,palm}$	SNR mean value in dB of the IR PPG signals measured from the participants' palm
$\upsilon^{G}_{SNR,\,wrist}$	SNR mean value in dB of the green PPG signals measured from the participants' wrists
$\upsilon^{IR}_{SNR,\,wrist}$	SNR mean value in dB of the IR PPG signals measured from the participants' wrists

LIST OF ABBREVIATIONS

AFRL	Air Force Research Lab
ANC	Adaptive Noise Cancellation
AWGN	Additive White Gaussian Noise
BPM	Beats per Minute
C-ANC	Cascading Adaptive Noise Cancellation
$COHb$	Carboxyhemoglobin
CT	Computation Time
CWT	Continuous Wavelet Transform
DoD	Department of Defense
DST	Discrete Saturation Transform
DWL	Dual Wavelength
ECG	Electrocardiography
FDA	Food and Drugs Administration
FPGA	Field Programmable Gate Array
Hb	Deoxygenated Hemoglobin
HbO_2	Oxygenated Hemoglobin
HMAPS	Holistic Modular Aircrew Physiologic Status
HR	Heart Rate
IQR	Interquartile Range

IR	Infrared
IRB	Institutional Review Board
JOSS	Joint Sparse Spectrum Reconstruction
KF	Kalman Filter
LED	Light Emitting Diode
LOA	Limits of Agreements
LOESS	Locally Estimated Scatterplot Smoothing
MAE	Mean Absolute Error
MAEP	Mean Absolute Error Percentage
$MetHb$	Methaemoglobin
NAWCAD	Naval AirWarfare Center Aircraft Division
ONR	Office of Naval Research
PI	Performance Index
PPG	Photoplethysmography
QRD-LSL	QR-Decomposition-Based Least-Squares Lattice
RMSE	Root Mean Square Error
R	Number of SpO_2 Realizations when Testing on Synthetic Data
RoI	Red over Infrared
RR	Respiratory Rate
SaO_2	Arterial Oxygen Saturation

SD	Standard Deviation
SET	Signal Extraction Technology
SNR	Signal to Noise Ratio
SpO_2	Peripheral Capillary Oxygen Saturation
SSA	Singular Spectrum Analysis
SSR	Sparse Signal Reconstruction
STL	Seasonal Trend Decomposition with LOESS
THR	Target Heart Rate
TROIKA	Signal Decomposi**T**ion, **R**ec**O**nstruct**I**on, and Pea**K** Tr**A**cking
Var	Variance
VLSI	Very Large Scale Integration

CHAPTER 1

INTRODUCTION

Photoplethysmography (PPG) is a noninvasive [1, 2], electro-optic method for detecting the cardiovascular pulse wave generated by the elastic nature of the peripheral vascular arteries excited by the quasi-periodic contractions of the heart [3]. Vital signs such as heart rate (HR), respiratory rate (RR), and peripheral capillary oxygen saturation (SpO_2) are usually extracted from PPG waveforms.

In conventional oximetry, simple methods are employed to calculate vital signs such as SpO_2 and HR. However, when the measured individual experiences substantial motion, sensor readings become noisier, and the conventional methods for vital sign calculation sometimes fails, providing false and inaccurate readings [4].

In this dissertation, our objective is to develop a robust and computationally efficient oximetry system that:

(1) is capable of suppressing motion artifacts from the noise contaminated PPG signal,

(2) can compute high-quality vital signs such as SpO_2 and HR, and

(3) is computationally efficient, and hence, can be implemented into wearable devices.

In Chapters 2 to 5, we address the challenges we mentioned above.

Chapter 2. Background and Literature Review. This chapter provides some history and background about pulse oximetry along with the fundamentals behind this technology. The conventional methods used in oximeters to extract SpO_2 and HR are presented.

Chapter 3. ECG HR-Tuned Comb Filter for Peripheral Blood Oxygen Saturation Estimation. A motion-resistant framework for SpO_2 level calculation is presented. This method is driven by the observation that the spectral components of the PPG waveform appear at a fundamental frequency that corresponds to the person's HR and at its harmonics. An electrocardiography (ECG) heart-rate-tuned comb filter is employed to clean up the PPG waveform before extracting an SpO_2 level.

Chapter 4. Mitigation of Motion Artifacts in Pulse Oximetry through Redundant Sensors. This chapter shows, using experimental data, the benefits of integrating SpO_2 level calculated from multiple channels (two channels in this case) using a Kalman filter. Integrating multiple channels is beneficial since if the data for the two channels are collected at sufficient physical distance from one another, the motion artifacts affecting the two channels are uncorrelated. The kalman filter is then able to use past measurements and modeling of the SpO_2 dynamics to attenuate the effect of the motion artifacts.

Chapter 5. Dual Wavelength Photoplethysmography Framework for Heart Rate Calculation. The systems proposed in chapters 3 and 4 use HR levels extracted from an ECG lead, since ECG signals are less susceptible to motion artifacts than PPG signals. Therefore, the availability of an ECG signal is necessary in these systems. However, the acquisition of an ECG waveform is more challenging than a PPG waveform. Moreover, ECG readings are sometimes unavailable, especially in wrist sensors, which are widely used in the fitness industry. Chapter 5 proposes an alternative to the use of ECG sensors for HR calculation. Here, we aim to calculate high-quality HR levels from PPG signals contaminated with high intensity repetitive "macro-motion" noise. The method uses PPG signals of different wavelengths, namely, infrared and green PPG signals. We test the performance of this framework on real experimental data by comparing the HR levels we calculate using our method to (1) ground truth manually computed from an ECG lead and (2) the leading HR calculation methods in the literature.

Wearable sensors are very desirable due to their ability to monitor the health of their wearer non-invasively. This makes them widely used in medical, military, and fitness fields. The major drawback of this technology is that the quality of the vital signs extracted from these devices deteriorates in the presence of substantial motion artifacts. The systems proposed in this dissertation (chapters 3 to 5) present possible solutions that improve the accuracy of vital sign calculation from PPG sensors in the presence of substantial motion artifacts. In addition to high quality vital signs calculation, the proposed SpO_2 and HR motion-resistant methods are computationally efficient, and hence, they can be implemented into wearable devices.

CHAPTER 2

BACKGROUND OF PULSE OXIMETRY

2.1 Oxygenated and Deoxygenated Hemoglobin

Hemoglobin is a protein present in the red blood cells that plays a fundamental role in (1) transporting oxygen molecules from the lungs to tissues and organs of the body and (2) collecting carbon dioxide and returning it back to the lungs. When it is bound to an oxygen atom, hemoglobin molecules are called oxygenated hemoglobin (HbO_2). Deoxygenated hemoglobin (Hb) is a name given to hemoglobin molecules without any oxygen bound. These two hemoglobin groups, namely, HbO_2 and Hb, are referred to as functional[1] hemoglobin [5].

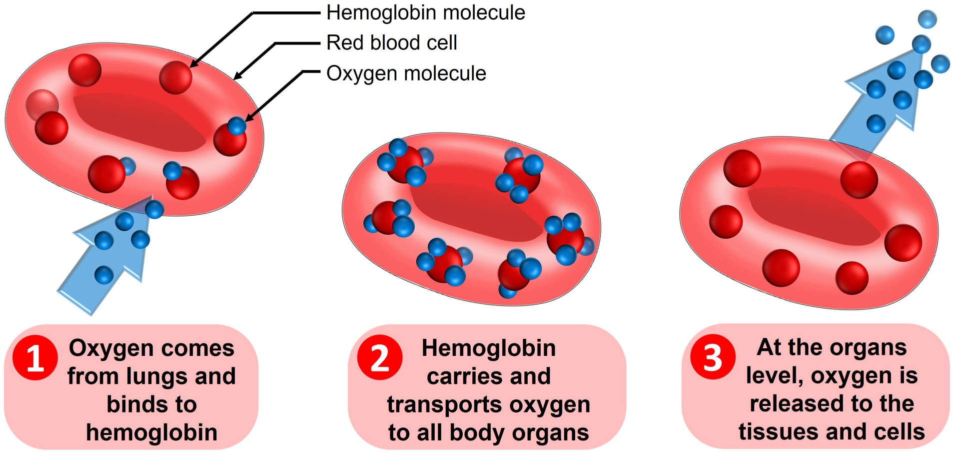


Figure 2.1 Cycle of oxygen transportation from the lungs to body organs.

We illustrate in Figure 2.1 the cycle of oxygen transportation from the lungs to body organs. Oxygen atoms arrive from the lungs and bind to hemoglobin molecules. Oxygen is then transported (via oxygenated hemoglobin) and released to the body organs, tissues, and cells (steps 2 and 3 of Figure 2.1).

---

[1] Methaemoglobin ($MetHb$) and Carboxyhemoglobin ($COHb$) are the other groups of hemoglobin which do not participate in the oxygen transport. Therefore, $MetHb$ and $COHb$ are called dysfunctional hemoglobin.

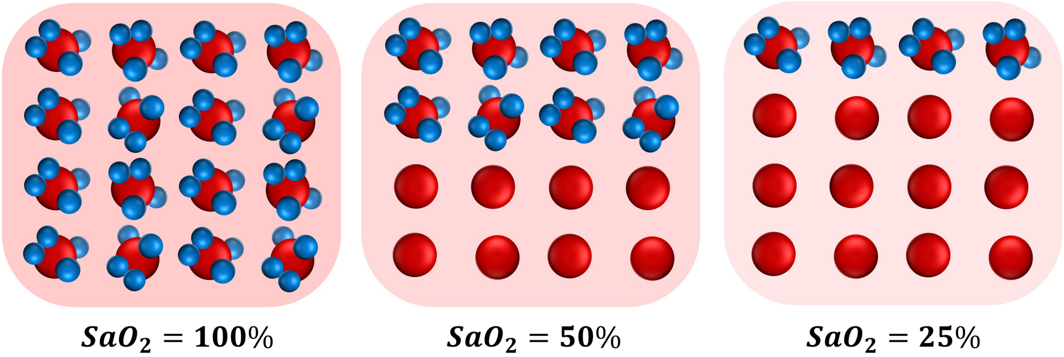


Figure 2.2 Illustration of the arterial oxygen saturation (SaO_2) calculation process.

The arterial oxygen saturation (SaO_2) is a measure of oxygen concentration in blood. This measurement is obtained by analyzing blood drawn from the patient. SaO_2 is calculated using Equation (2.1). SaO_2 calculation process is illustrated in Figure 2.2. In the left-most sample in Figure 2.2, all hemoglobin molecules are bound to oxygen. Therefore, the arterial oxygen saturation SaO_2 is 100%. When less hemoglobin molecules are bound to oxygen, the arterial oxygen saturation SaO_2 decreases. The right-most image of Figure 2.2 shows a sample with low SaO_2. In this example, SaO_2 is 25% since only four out of sixteen (4 out of 16) hemoglobin molecules are bound to oxygen atoms.

$$SaO_2(\%) = \frac{\boldsymbol{\mathcal{C}}(HbO_2)}{\boldsymbol{\mathcal{C}}(Hb) + \boldsymbol{\mathcal{C}}(HbO_2)}, \tag{2.1}$$

where $\boldsymbol{\mathcal{C}}(HbO_2)$ and $\boldsymbol{\mathcal{C}}(Hb)$ is the concentration of oxygenated and deoxygenated hemoglobin molecules.

2.2 Invention of the Pulse Oximeter

A pulse oximeter is device that detects and calculates the absorption of light by functional hemoglobin (oxygenated and deoxygenated hemoglobin) to produce an SpO_2 measurement – which is an estimate of SaO_2 [5]. The first pulse oximeter was invented by Dr. Takuo Aoyagi (see Figure 2.3), a Japanese engineering and inventor at Nihon Kohden. Dr. Aoyagi's dream was to detect oxygen saturation levels without having to draw blood. In 1974, Nihon Kohden applied for a Japanese patent[2], for its pulse oximeter – and Dr. Aoyagi's dream came true [6]. In 1975, Nihon Kohden launched "Oximeter OLV-5100" – a pulse oximeter that uses the principle developed by Dr. Aoyagi's [7]. Dr. Aoyagi died on April 18, 2020 at the age of 84.

Figure 2.3 Dr. Takuo Aoyagi, inventor of the pulse oximeter.
Source: J. Goodrich, "Takuo Aoyagi, inventor of the pulse oximeter, dies at age 84," IEEE Spectrum, 2022. [Online]. Available: https://spectrum.ieee.org/takuo-aoyagi-inventor-of-the-pulse-oximeter-dies-at-age-84. [Accessed: 20 February 2023].

2.3 Principles of Pulse Oximetry

2.3.1 Beer-Lambert Law

The principles of the pulse oximetry obey the Beer-Lambert Law shown in Equation (2.2) which is used to describe the light absorption phenomenon in biological tissues [8]. I_o and I_i are the intensities of the reflected and incident lights, respectively. $\mu_{\lambda,j}$, c_j, and d_j are the absorption coefficient of light of wavelength λ, concentration, and

---

[2] The patent was granted in 1979.

reflective light path length of layer j of the skin. Equation (2.2) assumes that the light is shining through a body that comprises J different type of layers.

$$I_o = I_i.exp\left(-\sum_{j=1}^{J} \mu_{\lambda,j} c_j d_j \right). \tag{2.2}$$

The intensity of the reflected light I_o varies as the volume of blood in the cite of measurement changes. The ratio between the maximum light intensity at diastole and the minimum at the systole at wavelength λ can be computed as following:

$$\frac{I_{max,\lambda}}{I_{min,\lambda}} = \exp\left(\Big(\boldsymbol{\mathcal{E}}(HbO_2, \lambda).\boldsymbol{\mathcal{C}}(HbO_2) + \boldsymbol{\mathcal{E}}(Hb, \lambda).\boldsymbol{\mathcal{C}}(Hb) \Big).\Delta l \right), \tag{2.3}$$

where Δl is amount by which the arterial radius expands during the systole and diastole. $\boldsymbol{\mathcal{E}}(HbO_2, \lambda)$ and $\boldsymbol{\mathcal{E}}(Hb, \lambda)$ are the extinction coefficients of HbO_2 and Hb at wavelength λ, respectively. $\boldsymbol{\mathcal{C}}(HbO_2)$ and $\boldsymbol{\mathcal{C}}(Hb)$ are the concentrations of oxygenated and deoxygenated hemoglobin, respectively [9, 10]. The Δl term of Equation (2.3) can be removed by taking the ratio $\frac{I_{max,\lambda}}{I_{min,\lambda}}$ for two different wavelengths. The choice of the two wavelengths is important since it should have some physical meaning linked to it.

Figure 2.4 shows the molar extinction (or light absorption) of the HbO_2 and Hb molecules for lights with various wavelengths ranging from ultraviolet (250 nm) to the infrared spectrum (1000 nm) [11]. The molar extinction of HbO_2 and Hb is somewhat similar for wavelengths ranging from 250 nm to 600 nm. However, we can observe a difference in the light absorption for HbO_2 and Hb in the red (around 650 nm) and infrared (around 900 nm) spectra. For instance, HbO_2 absorbs more infrared light than red light. Oppositely, Hb absorbs more red light than infrared light. Therefore, if we shine two lights – one of wavelength in the red spectrum and the other of wavelength in the infrared spectrum – we can obtain information about the functional hemoglobin molecules. This wavelength choice (red and infrared) can also remove the Δl term of Equation (2.3). We define optical density ratio '$\boldsymbol{\mathcal{R}}$' as following:

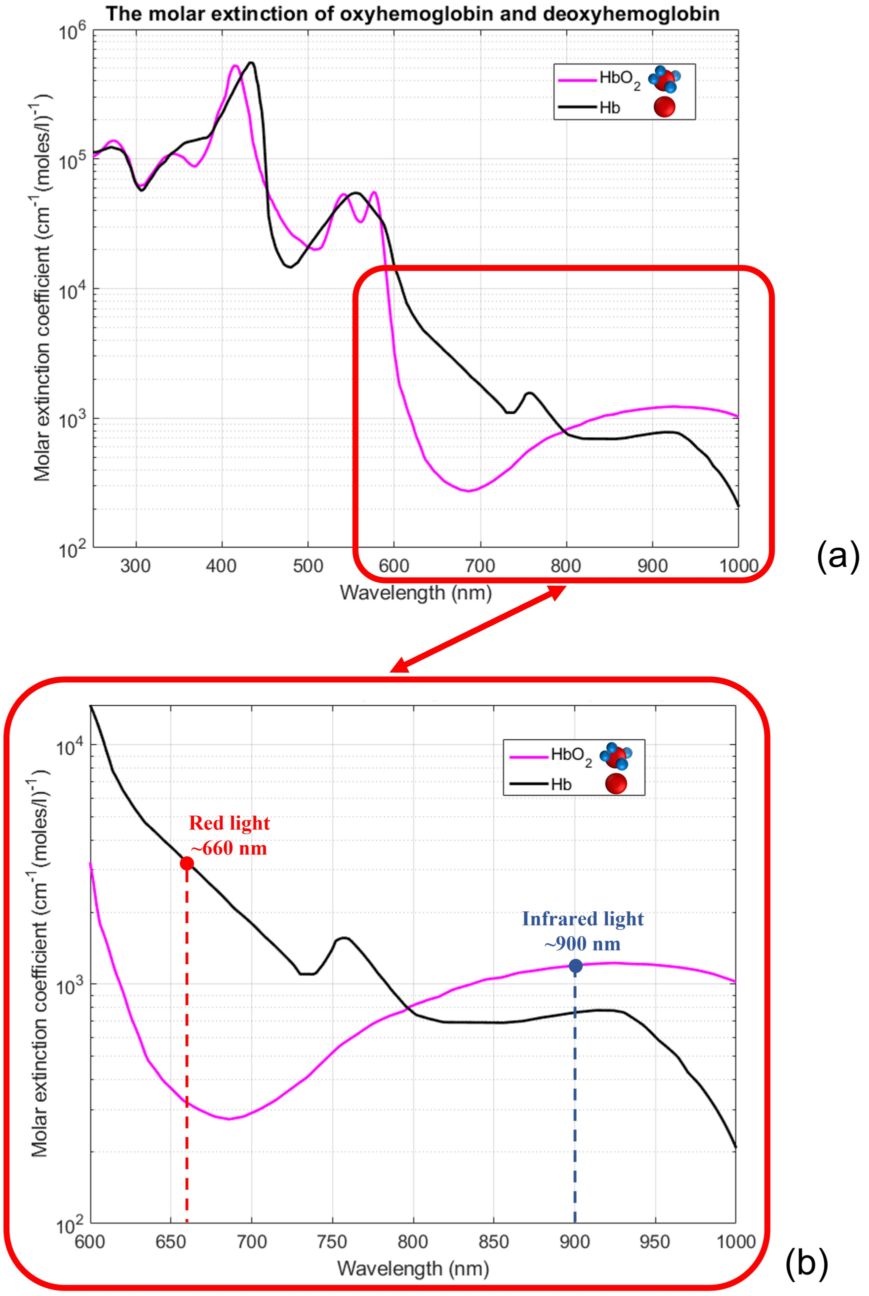


Figure 2.4 Molar extinction coefficient of HbO_2 and Hb. Subplot (a) shows the molar extinction coefficient for lights in the ultraviolet (250 nm) to the infrared spectrum (1000 nm). Subplot (b) is the zoomed version of subplot (a) for lights ranging from the red (650 nm) to the infrared spectrum (1000 nm).

$$\boldsymbol{\mathcal{R}} = \frac{log_e\Big(\dfrac{I_{max,\lambda_r}}{I_{min,\lambda_r}}\Big)}{log_e\Big(\dfrac{I_{max,\lambda_{ir}}}{I_{min,\lambda_{ir}}}\Big)} = \frac{\Big(\boldsymbol{\mathcal{E}}(HbO_2,\lambda_r).\boldsymbol{\mathcal{C}}(HbO_2) + \boldsymbol{\mathcal{E}}(Hb,\lambda_r).\boldsymbol{\mathcal{C}}(Hb)\Big).\Delta l}{\Big(\boldsymbol{\mathcal{E}}(HbO_2,\lambda_{ir}).\boldsymbol{\mathcal{C}}(HbO_2) + \boldsymbol{\mathcal{E}}(Hb,\lambda_{ir}).\boldsymbol{\mathcal{C}}(Hb)\Big).\Delta l}, \tag{2.4}$$

where λ_r and λ_{ir} are the wavelengths of red and infrared lights, respectively. SpO_2 can be obtained by plugging Equation (2.4) into Equation (2.1). SpO_2 can be written as following:

$$SpO_2(\%) = \frac{\boldsymbol{\mathcal{R}}.\boldsymbol{\mathcal{E}}(Hb,\lambda_{ir}) - \boldsymbol{\mathcal{E}}(Hb,\lambda_r)}{\boldsymbol{\mathcal{R}}.\Big(\boldsymbol{\mathcal{E}}(Hb,\lambda_{ir}) - \boldsymbol{\mathcal{E}}(HbO_2,\lambda_{ir})\Big) + \boldsymbol{\mathcal{E}}(HbO_2,\lambda_r) - \boldsymbol{\mathcal{E}}(HbO,\lambda_r)} \times 100. \tag{2.5}$$

If we use the extinction coefficients available in [11], namely, $\boldsymbol{\mathcal{E}}(Hb,\lambda_r) = 3226.6\ cm^{-1}(moles/l)^{-1}$, $\boldsymbol{\mathcal{E}}(Hb,\lambda_{ir}) = 761.84\ cm^{-1}(moles/l)^{-1}$, $\boldsymbol{\mathcal{E}}(HbO_2,\lambda_r) = 319.6\ cm^{-1}(moles/l)^{-1}$, and $\boldsymbol{\mathcal{E}}(HbO_2,\lambda_{ir}) = 1198\ cm^{-1}(moles/l)^{-1}$, then we get the relationship between SpO_2 and $\boldsymbol{\mathcal{R}}$ of Equation (2.6). Equation (2.6) associates the volume change in red and infrared lights ($\boldsymbol{\mathcal{R}}$) with an SpO_2 level. This type of relationships is called Beer-Lambert "calibration curve."

$$SpO_2(\%) = \frac{761.84\boldsymbol{\mathcal{R}} - 3226.6}{-436.16\boldsymbol{\mathcal{R}} - 2907} \times 100. \tag{2.6}$$

For simplicity, Beer-Lambert curve of Equation (2.6) is approximated by a first order polynomial. For example, the Texas Instruments AFE4490 – analog front-end for the pulse oximetry system – uses two diodes of wavelength of $\lambda_r = 660$ nm for the red light source and $\lambda_{ir} = 900$ nm for the infrared light source. They provide the calibration curve shown in Equation (2.7) in the manufacturer manual [12]. This calibration curve is an approximation of Equation (2.6) for the specific red and infrared lights they employed in their sensor. We show the Beer-Lambert calibration curve along with the AFE4490 calibration curve in Figure 2.5.

$$SpO_2(\%) = 110 - 25\boldsymbol{\mathcal{R}}. \tag{2.7}$$

Next, we discuss the calibration process that is typically used to generate a calibration curve that links an $\boldsymbol{\mathcal{R}}$ value to an SpO_2 level for an oximetry sensor.

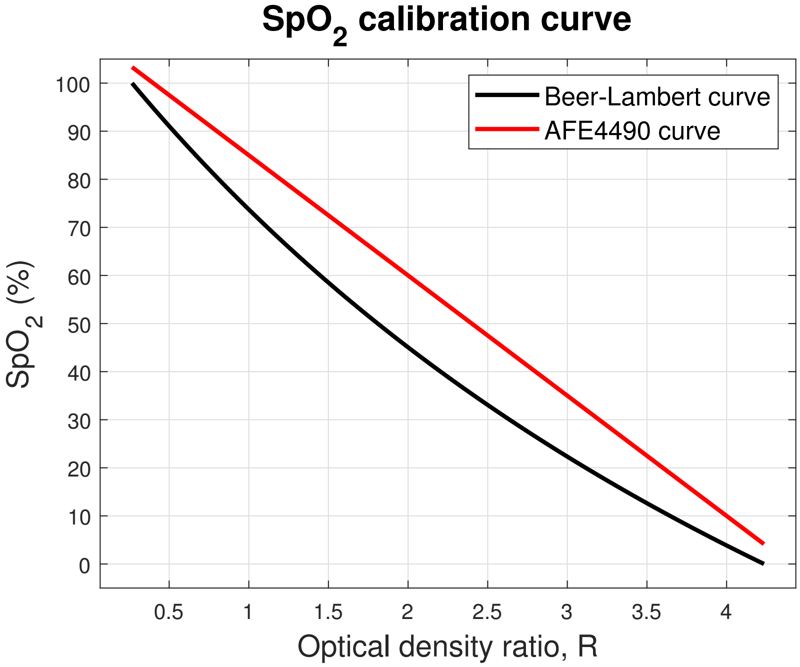


Figure 2.5 Calibration curves obtained by Beer-Lambert and AFE4490 manufacturer manual.

2.3.2 Calibrating oximetry sensors

As mentioned in Subsection 2.3.1, two lights, one of wavelength in the red spectrum and the other of wavelength in the infrared spectrum, are required in order to generate an SpO_2 value. Theoretically, the optical density ratio $\boldsymbol{\mathcal{R}}$ is related to SpO_2 using Equation (2.5). However, sensor calibration on a large population of people is required. Optimally, the calibration is achieving in an oxygen controlled environment. It is done by comparing SaO_2 obtained from blood samples drawn from a human volunteer to the optical density ratio $\boldsymbol{\mathcal{R}}$ that is simultaneously calculated from the oximetry sensor that is mounted on the volunteer. First, the volunteer is given a high

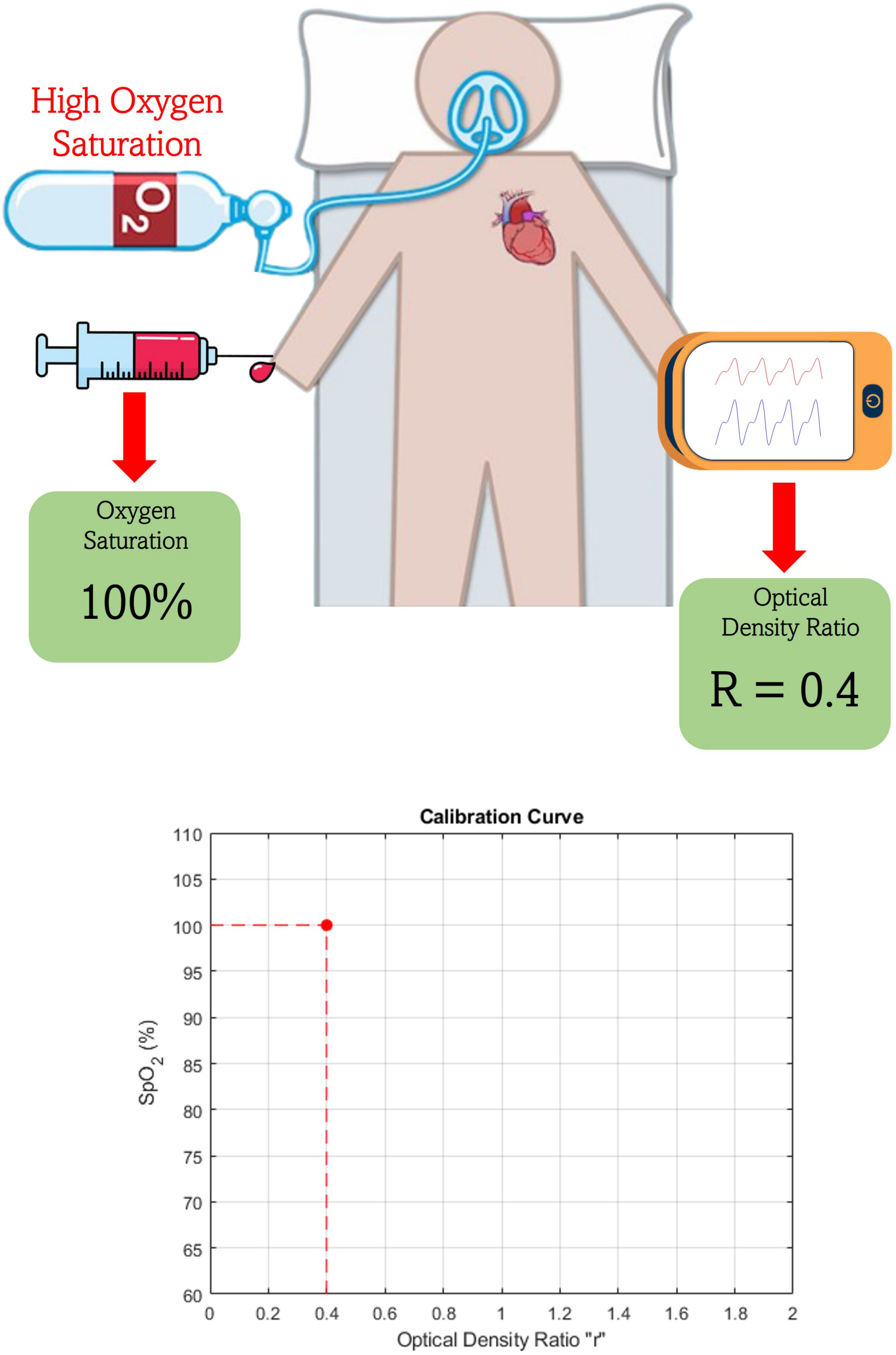


Figure 2.6 Oximetry sensor calibration. SaO_2 and optical density ration $\boldsymbol{\mathcal{R}}$ are calculated simultaneously when the volunteer is breathing a mixture with high oxygen concentration.

oxygen concentration mixture to breathe. The arterial oxygen saturation, SaO_2, is measured by analyzing a blood sample drawn from the volunteer. Concurrently, an optical density ratio $\boldsymbol{\mathcal{R}}$ value is measured from the oximetry sensor. Figure 2.6

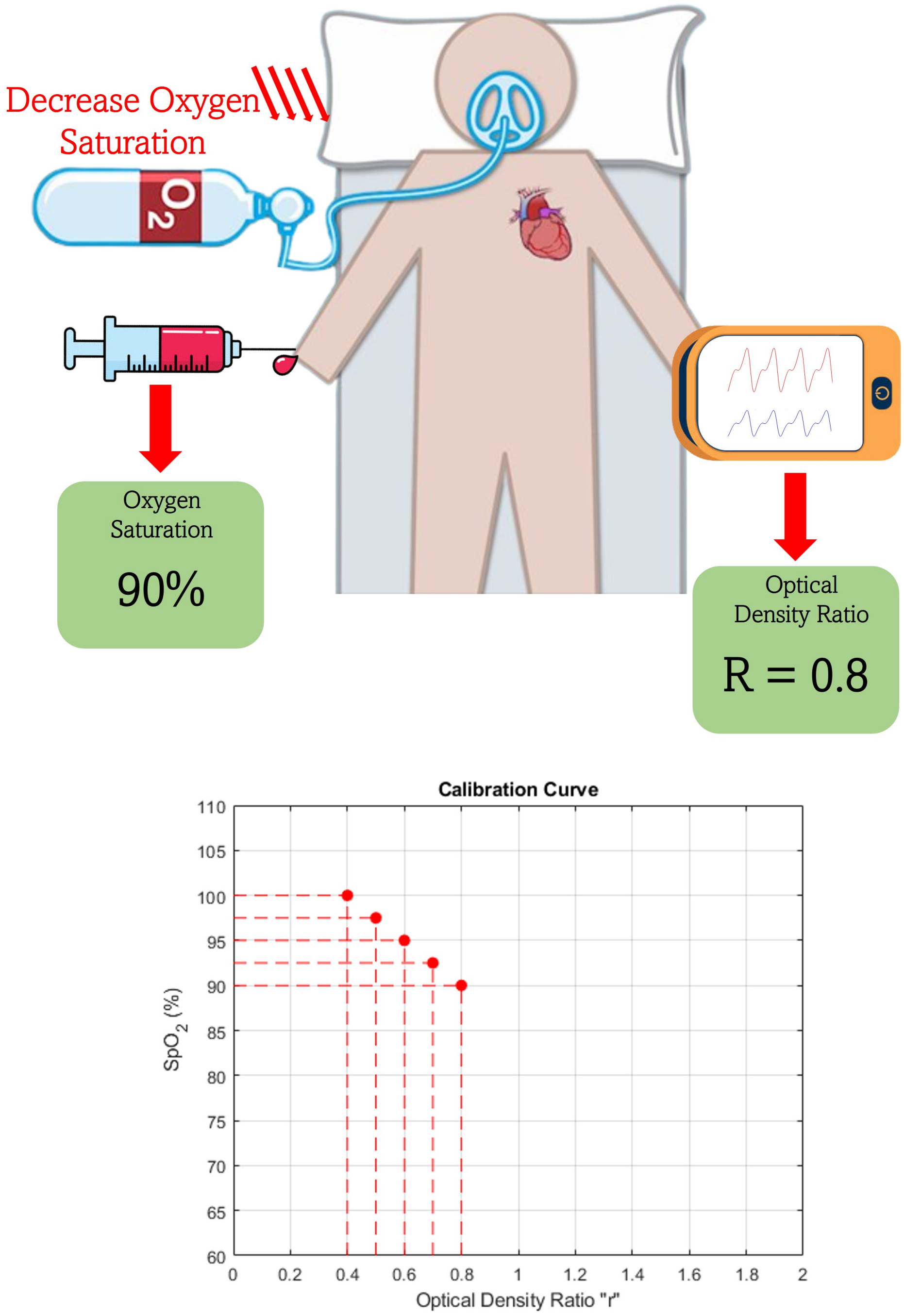


Figure 2.7 Oximetry sensor calibration. SaO_2 and optical density ration $\boldsymbol{\mathcal{R}}$ are calculated simultaneously as the concentration of oxygen in the mixture that the volunteer breathes decreases.

illustrates the initial step in the oximeter sensor calibration process. In the example for Figure 2.6, we obtained an SaO_2 value of 100% and an $\boldsymbol{\mathcal{R}}$ of 0.4. We draw a point on the graph of Figure 2.6 that corresponds to the pair ($\boldsymbol{\mathcal{R}}$, SaO_2). Next, the concentration of the oxygen in the mixture that the volunteer breathes gradually decreases (until reaching a low safe value). At each time step, a new pair ($\boldsymbol{\mathcal{R}}$, SaO_2) is generated and drawn on the same graph of Figure 2.6. Figure 2.7 continues the example of Figure 2.6. In Figure 2.7, the volunteer is breathing a mixture with a low oxygen concentration. This results in a low SaO_2 level and a high $\boldsymbol{\mathcal{R}}$ value. Finally, all the points that represent the pairs ($\boldsymbol{\mathcal{R}}$, SaO_2) are connected to form a calibration curve.

After constructing the calibration curve, the oximeter generates an SpO_2 level by first measuring the optical density ratio $\boldsymbol{\mathcal{R}}$ and then plugging its value in the calibration relationship that is obtained at the end of the sensor's calibration.

2.3.3 Transmittance vs. reflectance oximetry

There are two types of pulse oximeters – transmittance and reflectance oximeters. In both methods, a red and an infrared Light Emitting Diode (LED) are used. A photo-detector is used in order to capture the transmitted (in the case of transmittance oximetry) or back-scattered (in the case of reflectance oximetry) light. In a transmittance oximetry, the red and infrared LEDs and the photo-detector are placed on opposite sides of the measured site. In this case, the light is shone through the measured site and the photo-detector collect the light that traversed the measured side. Transmittance oximetry is commonly used in the clinical settings and the oximeter is mounted the patient's finger, earlobe, or toe. In reflectance oximetery, both the LEDs and the photo-detector are placed at the same side of the measured tissue. In this case, the light emitted by the LEDs trans-illuminates the tissue and the back-scattered light is detected by the photo-detector. Reflectance oximetery can measure SpO_2 at any place on the body. Usually, the sensor is mounted the forehead, wrist, or temple of the patient [13]. Figure 2.8 illustrates the transmittance and reflectance sensors. The signal collected by the photo-detector is called photoplethysmography (PPG) signal.

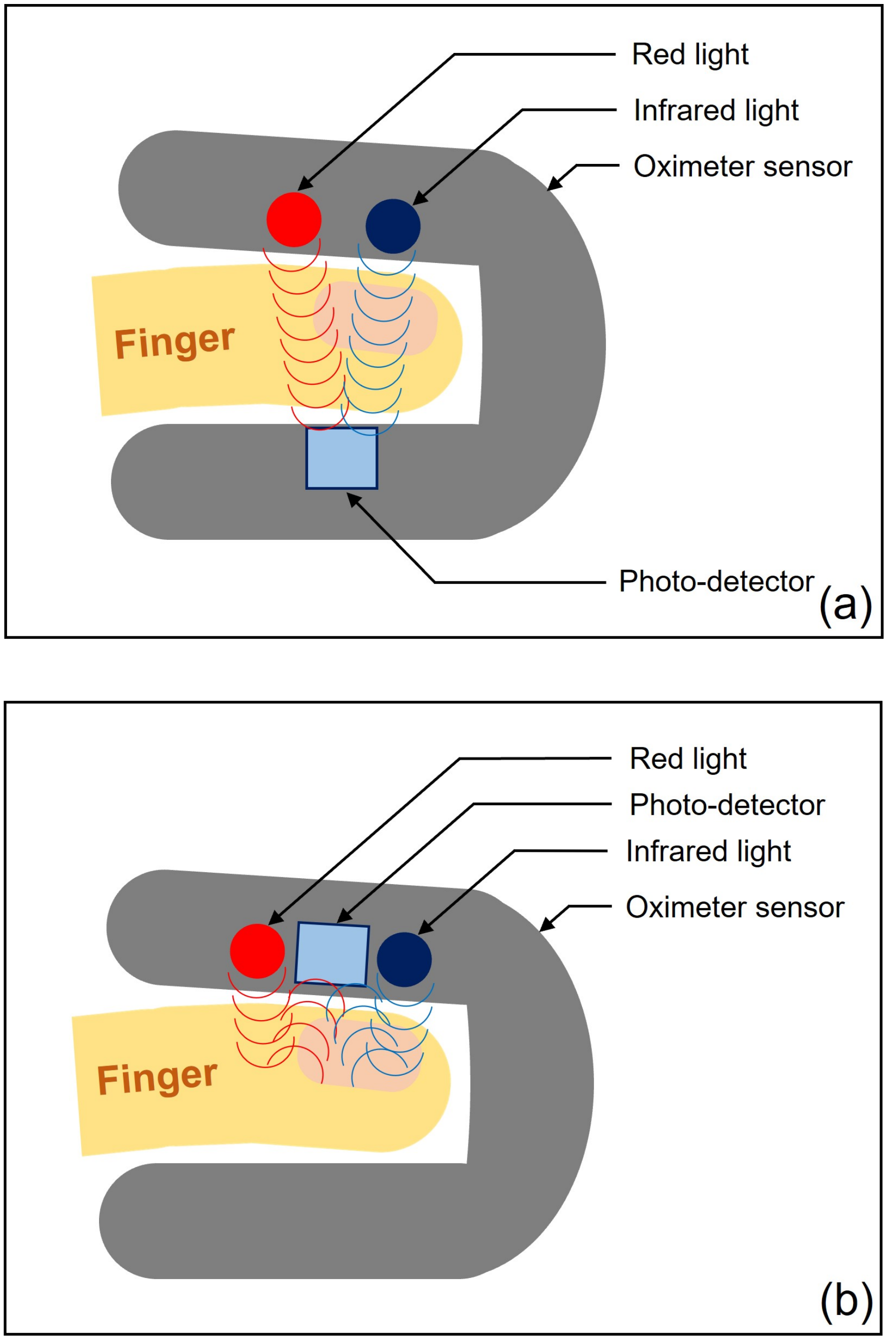


Figure 2.8 Oximetry illustration (a) transmittance and (b) reflectance.

2.3.4 Photoplethysmography (PPG) signal

PPG is a noninvasive [1, 2], electro-optic method for detecting the cardiovascular pulse wave generated by the elastic nature of the peripheral vascular arteries excited by the quasi-periodic contractions of the heart [3]. PPG signals are collected by the photo-detector of the pulse oximeter. Vital signs such as HR, respiratory rate (RR), and SpO_2 are usually extracted from PPG waveforms. In Figure 2.9, we show a clean PPG signal obtained from experimental data in the time domain (Figure 2.9 (a)) and in the frequency domain (Figure 2.9 (b)). This clean PPG signal was taken from a healthy male human in the course of intensive exercise regime[3].

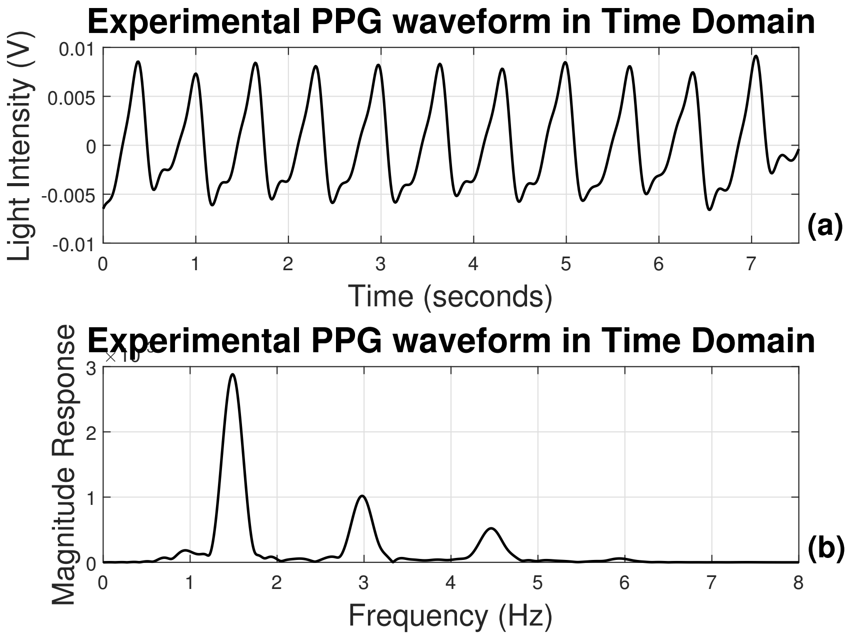


Figure 2.9 Subplot (a) is the time domain representation of a PPG signal extracted from experimental data. Subplot (b) is the frequency domain representation of a PPG signal extracted from experimental data.

We can observe in Figure 2.9 (b), that the spectral components of the PPG waveform appear at a fundamental frequency and at its harmonics (also, see Figure

---

[3] This PPG signal was collected from a exercise profile under an experimental protocol that is approved by Naval Air Warfare Center Aircraft Division IRB, protocol FWR21070114H, original approval date: 12 June 2017. Air Force Research Lab (AFRL) IRB protocols comply with DoD Directive 3216.02, Title 25, CFR 46, and are in compliance with the Declaration of Helsinki Revision 6, 2008.

6 in [14]). The fundamental frequency of the PPG signal corresponds to the person's heart rate. In order to study behavior, performance, and trade-offs in the design of SpO_2 estimators, we developed a synthetic PPG signal generator. We modeled the PPG red (Equation (2.8a)) and infrared (Equation (2.8b)) waveforms as the sum of a constant DC component and an AC component. The AC component is the sum of four sinusoids of different amplitude $A_{i\in\{1,2,3,4\}}$ (Equation (2.9)). The first sinusoid is at a frequency f_0, ranging from 0.5 to 3.5 Hz (corresponding to the person's heart rate and serving as the fundamental frequency). The three other sinusoids are its second, third, and fourth harmonics. The values of A_i were empirically derived from clean PPG waveforms taken from human volunteers at rest. The DC component was a constant value. In our simulation, we set the DC level of red and infrared PPG signals to 0.4 and 0.7, respectively. Figure 2.10 is an example of the synthetic red (red trace) and infrared (blue trace) PPG signals.

$$Red(t) = DC_{red} + AC_{red} \tag{2.8a}$$

$$Infrared(t) = DC_{infrared} + AC_{infrared} \tag{2.8b}$$

$$AC_{red}(t) = \sum_{i=1}^{4} -A_i \sin(2\pi \times i \times f_0 \times t) \tag{2.9}$$

where A_1=1.242 $\times 10^{-3}$, A_2=0.835 $\times 10^{-3}$, A_3=1.899 $\times 10^{-4}$, and A_4=0.786 $\times 10^{-4}$.

Heart Rate from PPG Signals: There are two conventional approaches to calculate HR from a PPG signal. The first method is a frequency domain approach based on the observation that the fundamental frequency of the PPG signal corresponds to the person's heart rate. Therefore, HR corresponds to the frequency that exhibit the highest magnitude in the frequency spectrum of the PPG signal. In

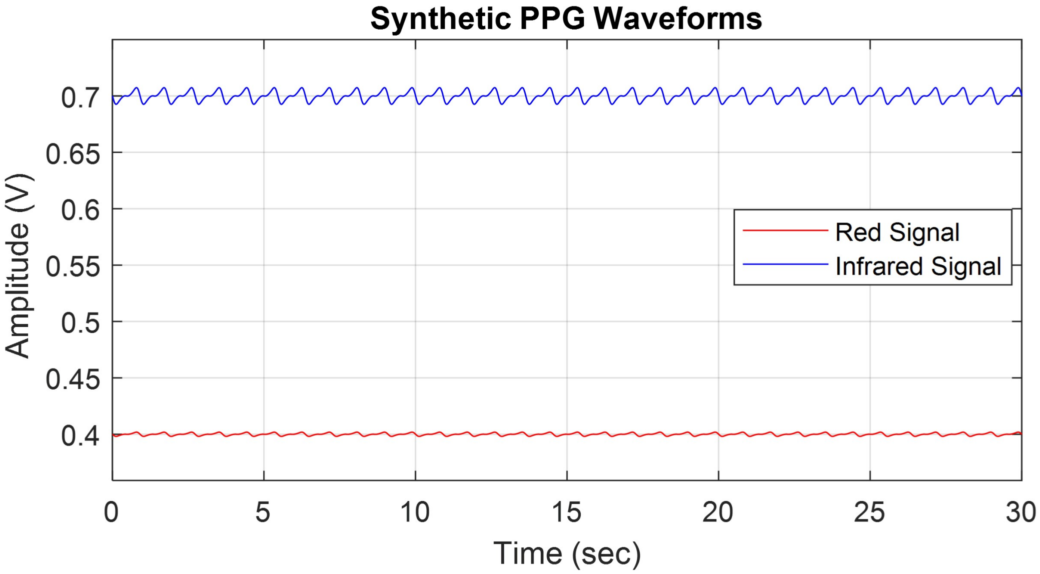


Figure 2.10 Synthetic red (red trace) and infrared (blue trace) PPG signals.

Figure 2.9, the spectral's highest magnitude appear at $f_0 = 1.47$ Hz. Therefore, the person's HR can be obtained using Equation (2.10) and is equal to 88 bpm.

$$HR\ (in\ bpm) = f_0 \times 60, \tag{2.10}$$

where f_0 is the PPG signal's fundamental frequency in Hz.

The second approach is a time domain approach. First, peaks are detected in the PPG signal during a set time interval. Thereafter, the distance between each two consecutive peaks is computed. The distance between two consecutive peaks is the period of the quasi-period PPG signal. HR is obtained by taking the average distance between two consecutive peaks. Note that in order to exhibit a high performance in HR levels calculation, the time domain approach requires the use of sophisticated peak detection methods. We show in Figure 2.11 all the peaks detected from the clean PPG signal of Figure 2.9. Red circles are added at the detected peaks. In this example, the average distance between each two consecutive peaks is 0.67 seconds – this can translate into a frequency of 1.49 Hz (or 89.4 bpm). Note that this value is very close to the heart estimated using the frequency domain.

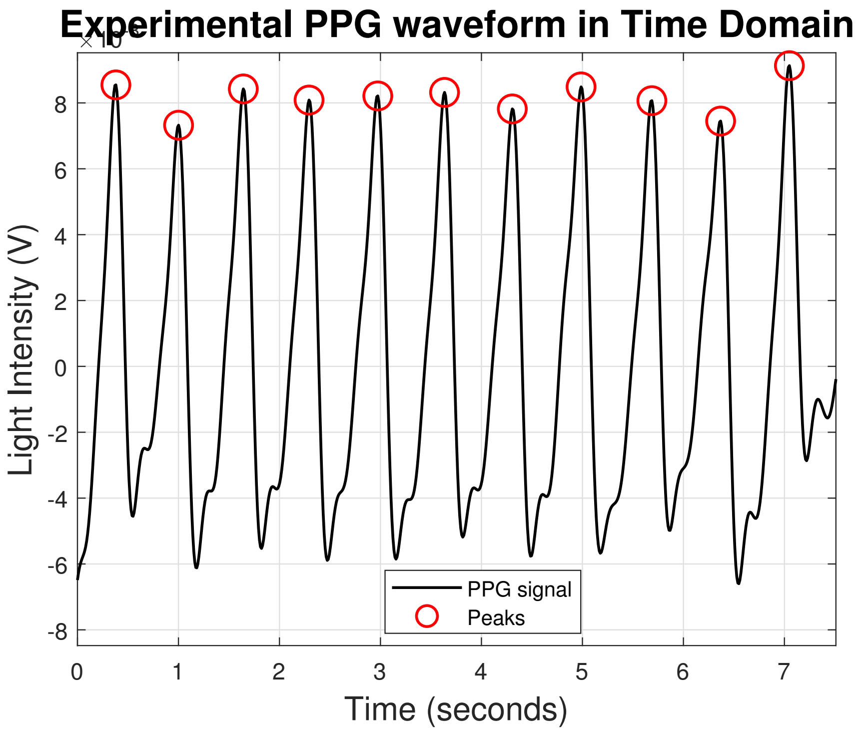


Figure 2.11 Peaks detected (red circles) from the clean PPG signal of Figure 2.9.

Peripheral Blood Oxygen Saturation from PPG Signals: Raw PPG (red and infrared) waveforms have two main components, namely: an AC component due to the light absorbed by pulsatile arterial blood, and a DC component due to the light absorbed by non-pulsatile components, such as tissues, venous, and capillary blood [15]. The AC component is obtained by applying a bandpass filter of passband frequency of 0.5 Hz to 10 Hz. The DC component is obtained by appyling a lowpass filter of bandpass filter of 0.5 Hz. Since the DC component varies from one person to another (depending on variables such as skin tone and tissue thickness), a normalization process is commonly used. The normalization is done by dividing the signal's AC component by its constant DC component. Figure 2.12 illustrates the normalization process on a red and an infrared PPG synthetic signals. As discussed previously in Subsection 2.3.1, we need two lights – one in the red spectrum and another in the infrared spectrum – in order to obtain an SpO_2 level. The optical density ratio $\mathcal{R}$, can be obtain using Equation (2.4). In practice, Equation (2.4) is estimated and replaced by a normalization technique [10]. The optical density ratio $\mathcal{R}$ can be estimated and re-written as:

$$\mathcal{R} = \frac{AC_{red}/DC_{red}}{AC_{ir}/DC_{ir}} = \frac{Normalized_{red}}{Normalized_{infrared}}. \tag{2.11}$$

The ratio $\mathcal{R}$ of Equation (2.11) is plugged into the sensor's calibration equation – for example Equation (2.7) is AFE4490 oximetry system is used – in order to obtain an SpO_2 level. This method is referred to as "Red-over-Infrared" (RoI) approach.

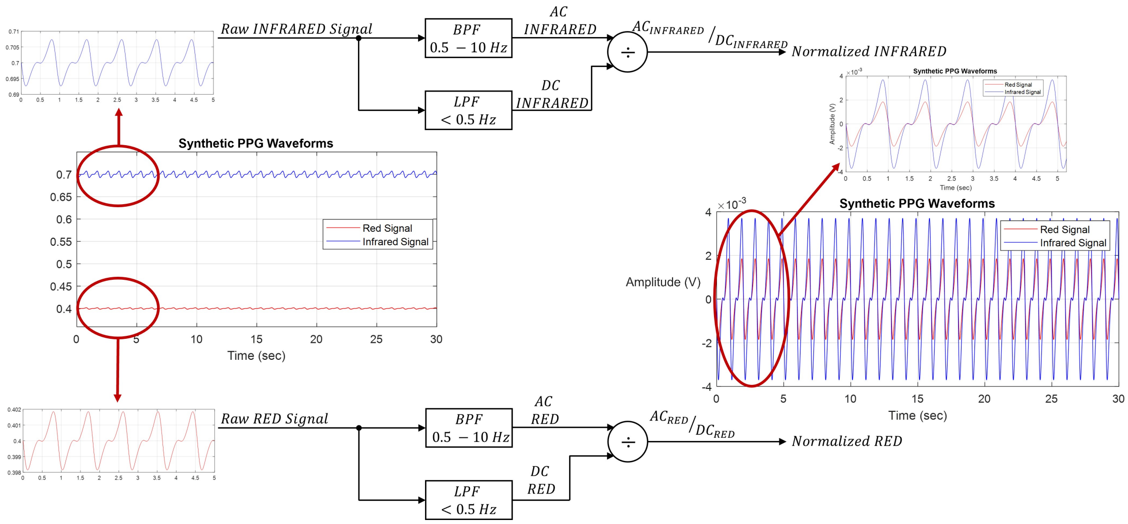


Figure 2.12 Illustration of the normalization process of synthetic red (red trace) and infrared (blue trace) PPG signals.

The time and frequency domain approaches for HR calculation and the RoI approach for SpO_2 calculation are effective when the person is steady during measurement. However, when the measurement happen while the person is experiencing substantial motion, the estimates become noisier, and the conventional approaches used for vital signs calculation sometimes fails, providing false and inaccurate readings [4]. In the following chapters, we will introduce motion-resistant methods for SpO_2 and HR calculation.

CHAPTER 3

ECG HR-TUNED COMB FILTER FOR PERIPHERAL BLOOD OXYGEN SATURATION ESTIMATION

Calculation of SpO_2 levels in humans is often made with a pulse oximeter, using PPG waveforms. However, measurements of PPG waveforms are susceptible to motion noise due to human and sensor movements. In presence of substantial motion artifacts, the conventional methods of vital signs calculations fail. Motion resistant algorithms are needed in order to provide accurate vital sign estimate in presence of motion. In this chapter, we introduce a pre-filtering method, that uses a heart-rate tuned comb peak filter to clean the PPG signals. We compare two SpO_2-level calculation techniques, and measure the effect of pre-filtering (by a heart-rate tuned comb peak filter) on their performance. These techniques are:

1. Red over Infrared (RoI) method, calculating the ratios of AC and DC components of the red and infrared PPG signals, followed by the use of a calibration curve to determine the SpO_2 level [1].
2. A motion-resistant algorithm which uses the Discrete Saturation Transform (DST) [2]. The DST algorithm isolates individual "saturation components" in the optical pathway, which allows separation of components corresponding to the SpO_2 level from components corresponding to noise and interference, including motion artifacts.

The comparison we provide here (employing the two techniques with and without pre-filtering) addresses two aspects: (1) accuracy of the SpO_2 calculations; and (2) computational complexity. We used both synthetic data and experimental data collected from human participants. The human participants were tested at rest and while exercising; while exercising, their measurements were subject to the impacts of motion. Our main conclusion is that if an uninterrupted high-quality heart rate measurement is available, then the RoI approach preceded by a heart-rate tuned comb filter provides the preferred trade-off between SpO_2-level accuracy and computational complexity. A modest improvement in SpO_2 estimate accuracy at very low SNR environments may be achieved by switching to the pre-filtered DST-based algorithm (up to 6% improvement in SpO_2 level accuracy at -10 dB over unfiltered DST algorithm and the filtered RoI approach). However, this improvement comes at a significant computational cost.

3.1 Motion Resistant Algorithms for SpO_2 Level Calculation

In presence of substantial motion, measurements become noisier, and the RoI approach – the conventional approach for SpO_2 calculation – sometimes fails, providing false and inaccurate readings [4]. A Discrete Saturation Transform (DST) based algorithm [2] that uses an adaptive noise cancellation filter [16, 17] was proposed to suppress some motion artifact effects on SpO_2 level calculations, thereby improving pulse oximetry. A 2002 study [18] reviewed the performance of twenty (20) commercial oximeters, and compared SpO_2 readings from stationary "control hand" of each of the seventy (70) healthy human participants to readings from the participant's other hand, which was in motion. In this study, a Masimo SET (Signal Extraction Technology) pulse oximeter, which uses the DST algorithm, exhibited the best performance over all other tested oximeters. Other comparisons of oximeter performances were reported in [19] (from 2016) and [20] (from 2018). The study in [20] also included a DST-based oximeter (Masimo Radical-7). It concluded that in the face of motion artifacts, the DST-based oximeter performed at a similar level to other FDA-cleared pulse oximeters[1].

In this study we are motivated by the observation that the spectral components of the PPG waveform appear at a fundamental frequency that corresponds to the person's heart rate and at its harmonics (see Figure 2.9 (b)). A comb filter tuned to these (possibly time-varying) frequencies thus may have the potential to "clean

---

[1] The pulse oximeters reported upon in [20] were the following: Masimo Radical-7, Nihon Kohden OxyPal Neo, Nellcor N-600, and Philips Intellivue MP5.

up" the PPG waveform prior to applying the SpO_2 calculation algorithm. The low computational complexity of a comb filter (when realized in software) may offer a viable alternative to the use of the more computationally complex realization of DST algorithm based systems.

3.2 Methods of ECG-HR Tuned Comb Filter

A block diagram of a processing module for PPG signals towards SpO_2 level calculation is shown in Figure 3.1. The module is subdivided into three main stages: (1) pre-processing, (2) filtering, and (3) SpO_2 calculation. The inputs are raw PPG (red and infrared signals) and ECG waveforms, and the outputs are SpO_2 levels. The virtual switch enables comparison of the performance of the SpO_2 calculation module with and without the comb filter.

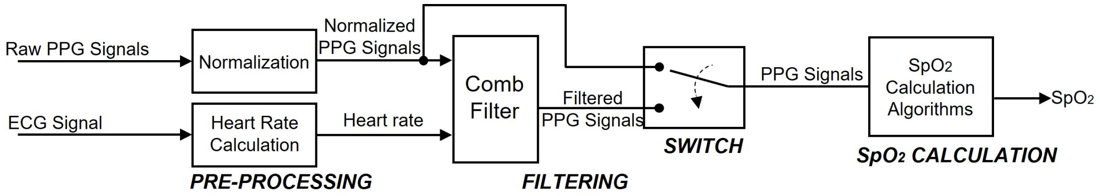


Figure 3.1 SpO_2 calculation procedure.

(1) In the pre-processing stage, raw PPG signals are normalized (Subsection 3.2.1). Concurrently, HR is calculated from an ECG waveform which is assumed to be available (Subsection 3.2.2).

(2) In the filtering stage, the normalized PPG waveforms are processed with a heart-rate tuned peak comb filter (Figure 3.2 (b)) that uses the calculated HR as a reference signal (Subsection 3.2.3). The filter presents its lowest attenuation at the HR frequency and its principal harmonics, and higher attenuation otherwise.

(3) The virtual switch (in Figure 3.1) allows us to compare the SpO_2 estimate that uses the normalized PPG signals to the estimate that uses these signals after comb filtering.

(4) In the SpO_2 calculation stage, we use one of two different algorithms RoI approach [4] or a DST-based algorithm [2].

3.2.1 Normalization

Raw PPG (red and infrared) waveforms have two main components, namely: an AC component due to the light absorbed by pulsatile arterial blood, and a DC component due to the light absorbed by non-pulsatile components, such as tissues, venous, and capillary blood [15]. Since the DC component varies from one person to another (depending on variables such as skin tone and tissue thickness), a normalization process is commonly used. The normalization is done by dividing the signal's AC component by its constant DC component.

3.2.2 Heart rate calculation

We assume that we have access to the electrocardiography (ECG) waveform of the participant whose SpO_2 level we measure. The ECG waveform is known to be less susceptible to motion noise than the PPG waveform [21, 22]. HR (in beats per minute (BPM)) was calculated in our study from an ECG signal through the Pan and Tompkins algorithm [23]. Since the fundamental frequency of the PPG signal is the HR, we use the HR to tune the comb filter. The comb filter discriminates against the portion of the PPG input signals which are not at the HR frequency or one of its principal harmonics.

3.2.3 Comb filter

The spectral components of the PPG waveform appear at a fundamental frequency (corresponding to the person's HR) and its harmonics (see Figure 2.9 (b)). The use of a comb peak filter tuned to these frequencies may therefore serve to clean up the PPG waveform. The filter exhibits low attenuation at the fundamental frequency and its harmonics, and high attenuation in the intermediate regions between these frequencies (see Figure 3.2 (b)). In this manner, the filter reduces noise that resides in the intermediate regions. In order to reject as much noise as possible, we want the 'peaks' of the filter to be narrow. On the other hand, overly narrow peaks are likely to miss the PPG harmonics if the tuning is not exact (if the filter is not tuned exactly to the HR). Therefore, a compromise is needed between tuning accuracy and noise-rejection capability. We employed an IIR comb filter with the transfer function

$$H_c(z) = \beta_c \frac{1 + z^{-K_c}}{1 - \gamma_c z^{-K_c}}, \tag{3.1}$$

where γ_c and β_c are two positive scalars and K_c is the comb filter's order. The design equations are shown in Table 3.1. We have selected the 3 dB bandwidth, $f_{c,BW}$, to be 0.2 Hz (capturing 97.5% of the total power of the signal of interest). For this $f_{c,BW}$ selection, if we tune the comb filter to the person's HR frequency (PPG signal's fundamental frequency), f_0=1 Hz, the null-to-null bandwidth of the filter's lobes is 1 Hz and the 10 dB bandwidth is 0.49 Hz. The sampling rate was $f_s = 256$ Hz.

Table 3.1 Comb Filter Design Equations

$K_c = \frac{f_s}{f_0}$ (dimensionless)	f_0 is the fundamental frequency (HR) in Hz. f_s is the sampling frequency.
$\beta_c = \frac{1-\gamma_c}{2}$ (dimensionless)	Gain at fundamental frequency and its harmonics set to 1.
$f_{c,BW} = cos^{-1}\left(\frac{2\gamma_c}{\gamma_c^2+1}\right) \times \frac{f_0}{180}$ (Hz)	3 dB bandwidth set.

Figure 3.2 shows a frequency domain plot of a noise contaminated PPG signal measured on a healthy male participant during aerobic exercise (see Subsection 3.3.2 for more details). The signal is passed through a heart-rate tuned comb filter whose transfer function (magnitude response) is shown in Figure 3.2 (b). The fundamental frequency (HR frequency) of the participant is $f_0 = 2.29$ Hz. The parameters of the comb filter are $K_c = 112$, $\gamma_c = 0.7570$, and $\beta_c = 0.1215$. Figure 3.2 (c) shows the clean PPG signal emerging from the comb filter.

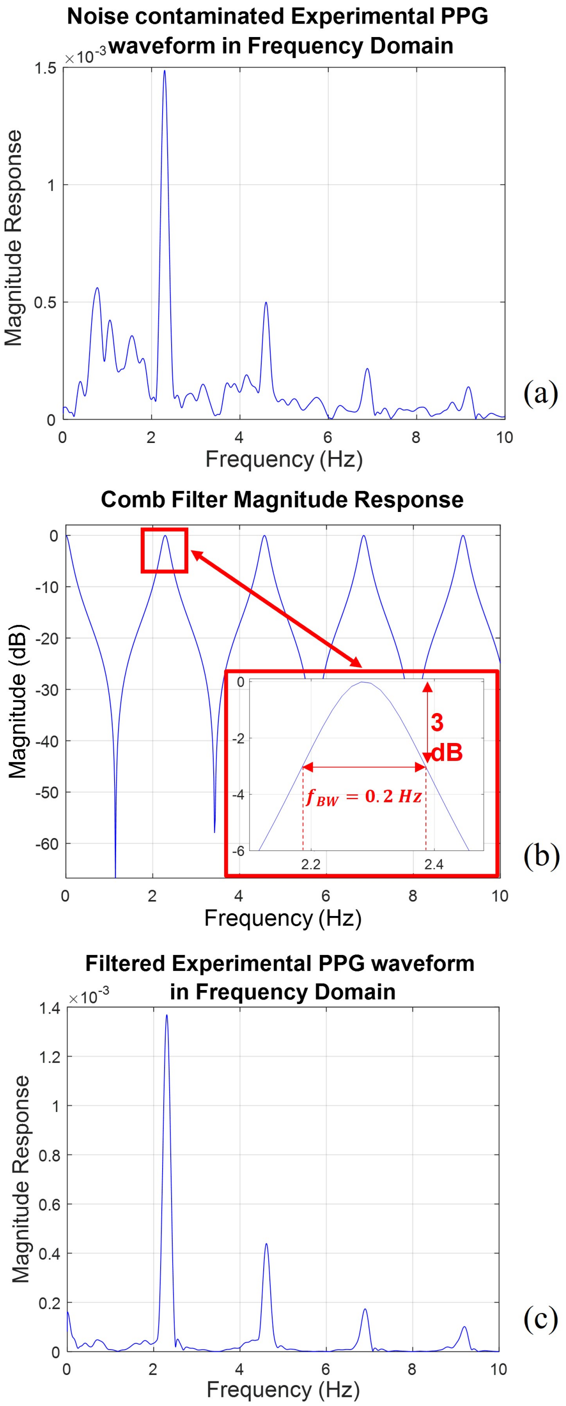


Figure 3.2 (a) Frequency domain representation of a noise contaminated experimental PPG waveform of fundamental frequency f_0=2.29 Hz. (b) Magnitude response of a tuned comb filter. (c) Frequency domain representation of the comb-filtered PPG waveform.

3.2.4 SpO_2 level calculation

The "Red over Infrared" (RoI) approach

In the RoI approach, two light sources of different wavelengths, λ_r and λ_{ir} (red and infrared light, respectively), are used. The optical density ratio $\boldsymbol{\mathcal{R}}$ is defined as the ratio of the normalized red to the normalized infrared waveforms. In our study, we have used the Texas Instruments AFE4490 as the analog front-end for the pulse oximetry system – using diodes of wavelength of $\lambda_r = 660$ nm for the red light source and $\lambda_{ir} = 900$ nm for the infrared light source. In order to calculate SpO_2 levels, we employed first the calibration curve of Equation (2.7) which was provided by the manufacturer as the standard model [12]. This calibration curve is re-shown in this chapter as Equation (3.2).

$$SpO_2(\%) = 110 - 25\boldsymbol{\mathcal{R}}. \tag{3.2}$$

To study the sensitivity of our statistical results and main conclusions (Subsection 3.4.2 and Table 3.7) to the specification of the calibration curve, we have also employed two alternate calibrations curves in this study (*viz.*, we calculated the statistics separately for each one of three different calibration curves, see Figure 3.3).

The first alternate curve is provided by the Beer-Lambert method [24], shown in Figure 3.3 as a black trace. Notably, the AFE4490 calibration curve (red trace) shows a relationship between SpO_2 and $\boldsymbol{\mathcal{R}}$ which is "to the right and above" Beer-Lambert curve (black curve of Figure 3.3)[2]. Hence, the AFE4490 calibration model overestimates the SpO_2 level when compared ot the Beer-Lambert estimate at the same value of $\boldsymbol{\mathcal{R}}$.

The second alternate curve is "to the left and below" (magenta curve of Figure

---

[2] The Beer-Lambert equation of Equation (2.6) is re-shown as Equation (3.3) in this chapter.

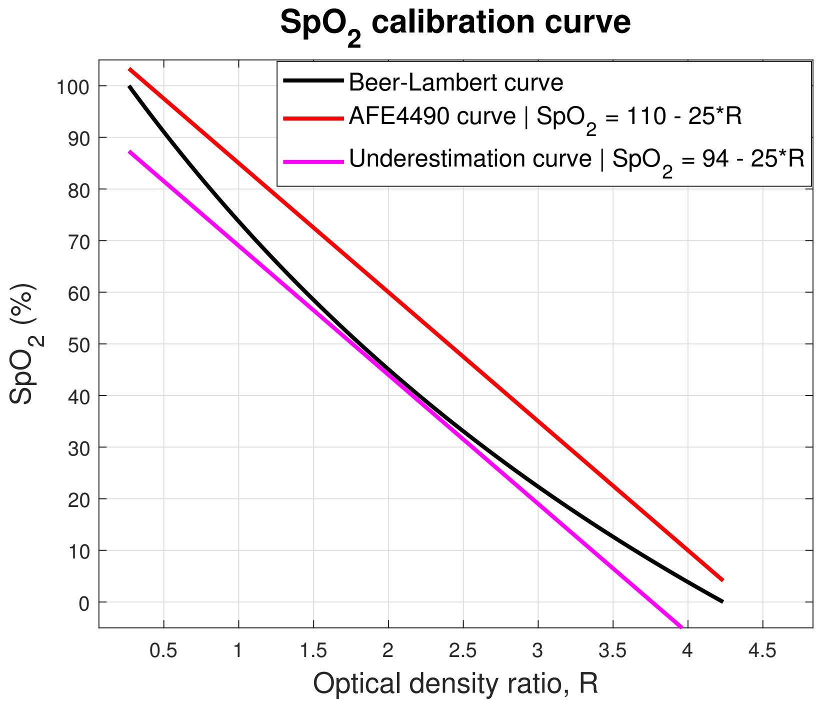


Figure 3.3 Calibration curves used for sensitivity study.

3.3) the Beer-Lambert curve, and hence underestimates the SpO_2 level when compared to the Beer-Lambert estimates. We denote this curve of Equation (3.4) "underestimation calibration curve" (it has the same slope (-25) as the standard model). The equations of the alternate calibration curves are as follows.

Beer-Lambert estimation calibration curve:

$$SpO_2(\%) = \frac{761.84\boldsymbol{\mathcal{R}} - 3226.6}{-436.16\boldsymbol{\mathcal{R}} - 2907} \times 100. \tag{3.3}$$

Underestimation calibration curve:

$$SpO_2 = 94 - 25\boldsymbol{\mathcal{R}}. \tag{3.4}$$

The Discrete Saturation Transform (DST) algorithm

The DST algorithm [2] was derived to measure SpO_2 levels in the face of motion noise. In developing the algorithm, it was assumed that the clean PPG signal of interest is contaminated by additive noise, uncorrelated with the signal. The red and infrared PPG signals are the inputs and the SpO_2 level is the output. A family of reference signals is generated for each optical density ratio corresponding to SpO_2 values ranging from 50% to 100% at a resolution of 0.5%. The reference signal is defined as

$$reference\ signal(t) = infra(t) \times r - red(t)\ . \tag{3.5}$$

Here, 'r' is an arbitrary optical density ratio value that corresponds to SpO_2 levels ranging from 50% to 100% (we use the calibration curve Equation (3.2), which gives the corresponding values of 'r' of 0.4 to 2.4). "red(t)" and "infra(t)" are the time-dependent red and infrared PPG signals collected on a range of t∈[0,T_s] (T_s is typically 10 seconds); they serve as the two inputs of the DST algorithm. Figure 3.4 is the block diagram of the DST algorithm. The inputs of the DST algorithm are a red and an infrared PPG signal. The output is an SpO_2 estimate. The red 'x' in Figure 3.4 is the SpO_2 level estimated by the DST algorithm. The blue '*' is the SpO_2 ground truth.

According to Goldman *et al.* [2], noise contaminated red and infrared PPG signals can be written as

$$red(t) = S_{red}(t) + N_{red}(t), \tag{3.6}$$

and

$$infra(t) = S_{infra}(t) + N_{infra}(t)\ . \tag{3.7}$$

Using Equations (3.6) and (3.7), we can redefine Equation (3.5) as

$$reference\ signal(t) = (r - r_{SpO_2}) \times S_{infra}(t) + (r - r_{Noise}) \times N_{infra}(t), \tag{3.8}$$

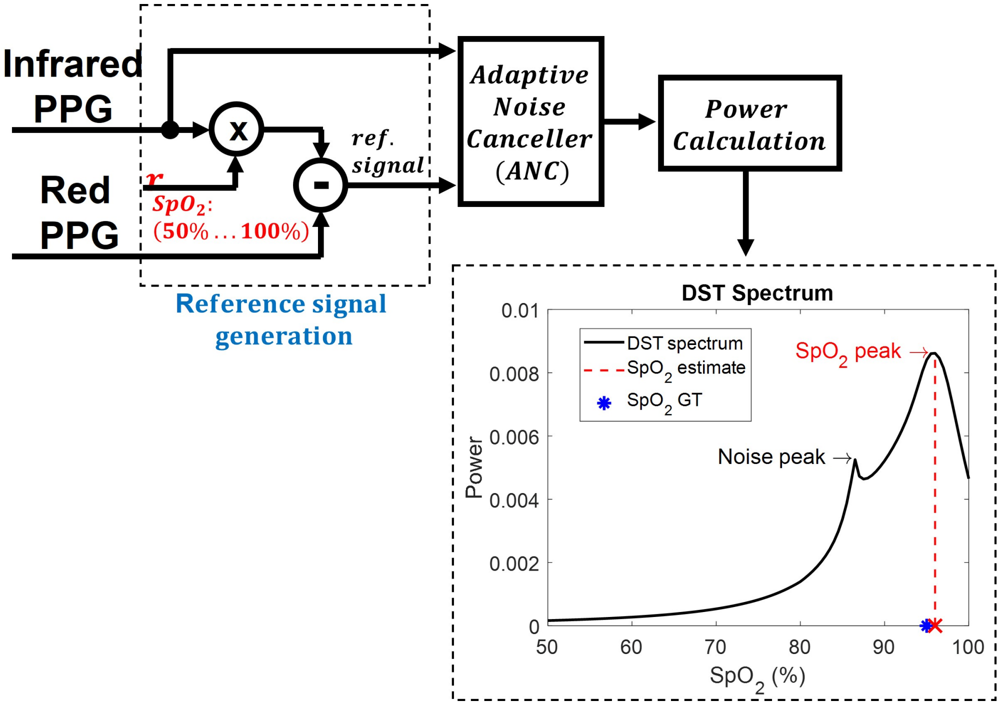


Figure 3.4 DST algorithm block diagram and DST spectrum for a noise-contaminated synthetic PPG signal of SNR of 0 dB. The SpO_2 level calculated by the algorithm corresponds to the right-most peak in the DST spectrum (output power vs. SpO_2 level) and is represented by a red 'x.' The blue '*' is the SpO_2 ground truth.

where $S_{infra}(t)$ and $N_{infra}(t)$ are the desired signal component and noise component of the infrared PPG signal, respectively. r_{SpO_2}, the optical density ratio that corresponds to the correct SpO_2 value, is defined in Equation (3.9). Additionally, r_{Noise}, the ratio of the noise component of the red and infrared PPG signals, is defined in Equation (3.10) [2].

$$r_{SpO_2} = \frac{S_{red}(t)}{S_{infra}(t)}. \tag{3.9}$$

$$r_{Noise} = \frac{N_{red}(t)}{N_{infra}(t)}. \tag{3.10}$$

The DST algorithm employs Adaptive Noise Cancellation (ANC) filters [16, 17] to remove noise, and provide a "clean" SpO_2. For each arbitrary 'r' (corresponding to an SpO_2 level between 50% and 100%), the reference signal and the infrared signal are fed into an ANC filter which identifies and removes frequency components which are in common between the two signals [2]. The power of the signal collected at the output of the ANC is calculated for each reference signal. The "DST spectrum" shows the SpO_2 values (or 'r' value) used to generate the reference signals on the abscissa, and the power of the ANC's output for each reference signal on the ordinate. As we vary the values of 'r,' we observe three distinct cases, namely;

Case 1: $r \neq r_{Noise} \neq r_{SpO_2}$,

Case 2: $r = r_{Noise}$, and

Case 3: $r = r_{SpO_2}$.

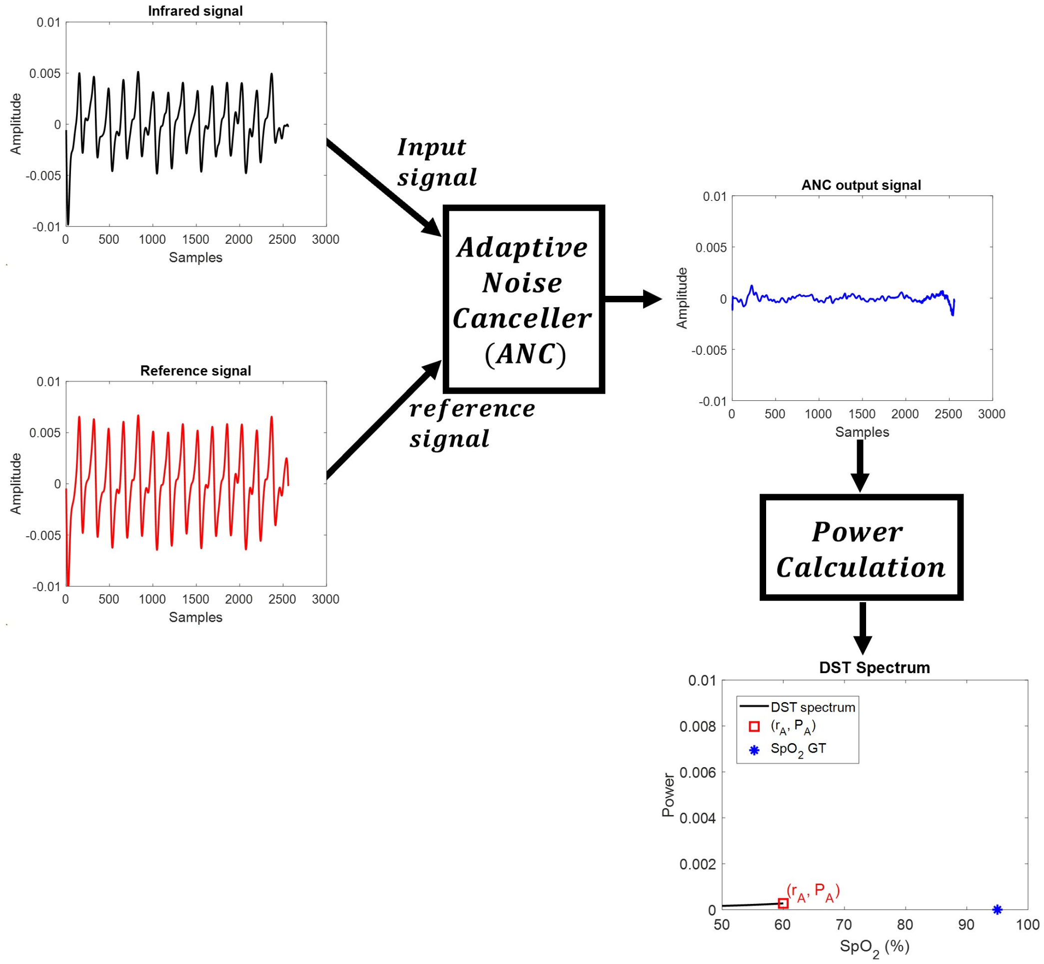


Figure 3.5 Illustration of case 1, namely, $r = r_A \neq r_{Noise} \neq r_{SpO_2}$. In this case, the noise reference signal contains the desired signal component as well as a noise component. Therefore, the output of the ANC is a signal of a weak power P_A. The point (r_A, P_A) (red square) is used to construct the DST spectrum. The blue '*' is the SpO_2 ground truth.

To better understand how the ANC filters are used in the DST algorithm, we developed an example where two PPG signals, a red and an infrared, are contaminated with an bandpass filtered (0.5 to 5 Hz) AWGN noise. The Signal-to-Noise (SNR) (see Equation (3.11)) ratio is 0 dB. The ratio between the desired component of red and infrared signals, r_{SpO_2}, is preset to 0.5, which yields into an SpO_2 ground truth of 97.5%. The ratio between the noise component of red and infrared signals, r_{Noise} is set to 1.

$$SNR = \frac{Var(S)}{Var(N)}, \tag{3.11}$$

where S is the desired signal and N is the noise component.

Case 1 is illustrated in Figure 3.5. In this illustration, the arbitrary optical density ratio taken a value $r = r_A = 2$. When plugged in the calibration curve Equation (3.2), $r = r_A = 2$ ($r_A \neq r_{Noise} \neq r_{SpO_2}$) will correspond to an SpO_2 value of 60%. At this instance, the reference signal calculated using Equation (3.8) is

$$reference\ signal(t) = (r_A - r_{SpO_2}) \times S_{infra}(t) + (r_A - r_{Noise}) \times N_{infra}(t). \tag{3.12}$$

We can see from Equation (3.12) that the reference signal contains both, desired signal components, $(r_A - r_{SpO_2}) \times S_{infra}(t)$, and noise components, $(r_A - r_{Noise}) \times N_{infra}(t)$. The input signal, $infra(t)$, and the reference signal of Equation (3.12) are used as inputs to the ANC. Since both signals contain desired signal components and noise components, the output of the ANC will be a very weak signal. We calculate the power of the output signal and denote it P_A. The pair (r_A, P_A), represented as a red square in Figure 3.5, is one of points used to construct the DST spectrum.

We keep decreasing the value of 'r' until we reach the value $r = r_B = 0.94$ which is very close to r_{noise} ($r = 0.94$ corresponds to 86.5% when plugged in Equation (3.2)).

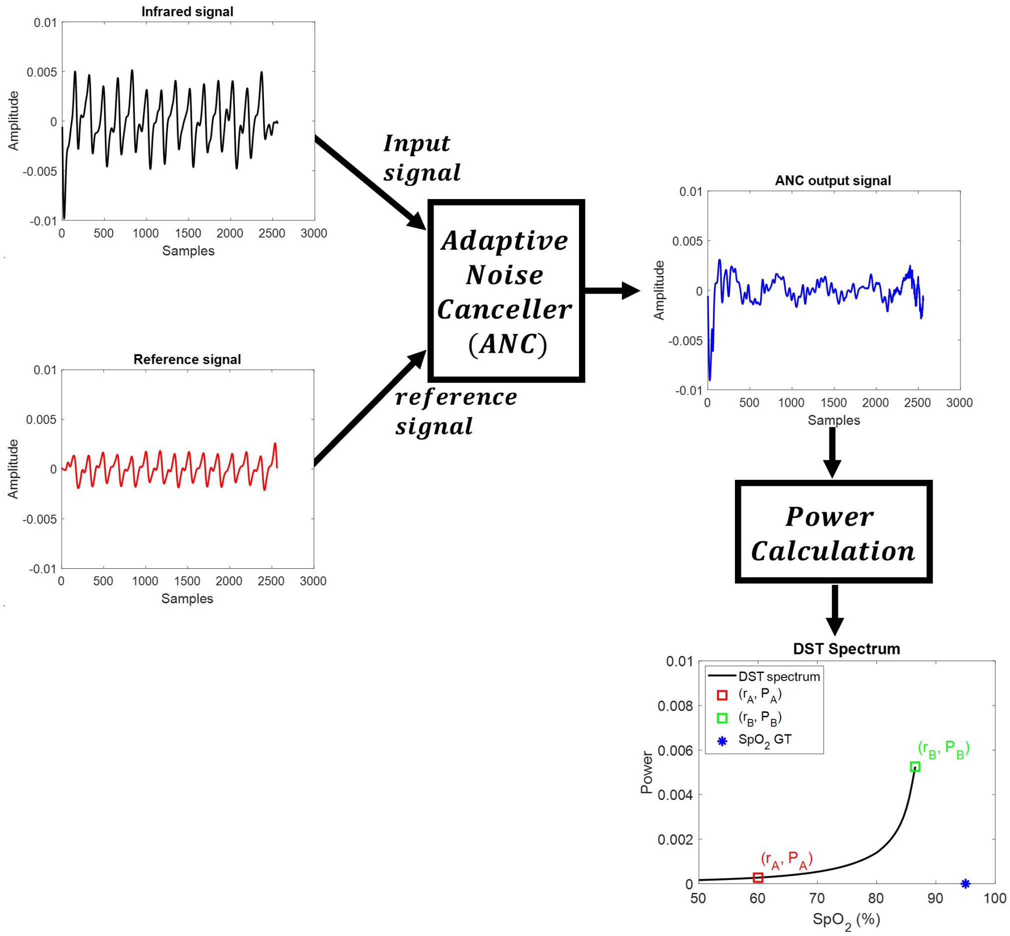


Figure 3.6 Illustration of case 2, namely, $r = r_B$ which is very close to r_{Noise}. In this case, the noise reference signal contains only the desired signal component. Therefore, the output of the ANC is the noise component of power P_B. The point (r_B, P_B) (green square) is the noise peak in the DST spectrum. The blue '*' is the SpO_2 ground truth.

Figure 3.6 is used for illustration of this case (case 2). At this instance, the reference signal calculated using Equation (3.8) is

$$reference\ signal(t) = (r_B - r_{SpO_2}) \times S_{infra}(t) + (r_B - r_{Noise}) \times N_{infra}(t), \quad (3.13)$$

and since r_B is very close to r_{Noise}, Equation (3.13) can be re-written as

$$reference\ signal(t) = (r_{noise} - r_{SpO_2}) \times S_{infra}(t). \quad (3.14)$$

Note from Equation (3.14) that the reference signal contains only the desired signal components, $(r_{noise} - r_{SpO_2}) \times S_{infra}(t)$. The ANC will remove the signal

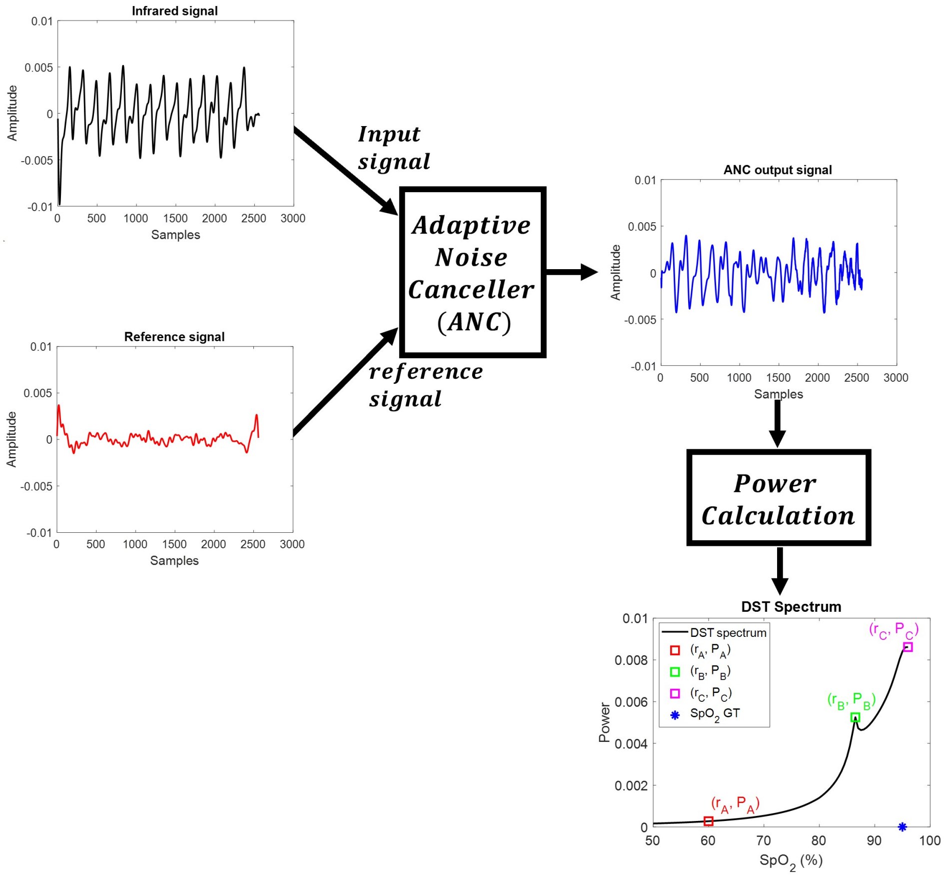


Figure 3.7 Illustration of case 3, namely, $r = r_C$ which is very close to r_{SpO_2}. In this case, the noise reference signal contains only the noise component. Therefore, the output of the ANC is the desired signal component of power P_C. The point (r_C, P_C) (magenta square) is the SpO_2 peak in the DST spectrum. The blue '*' is the SpO_2 ground truth.

component from the input signal, $infra(t)$. Therefore, the output of the ANC will be the noise component. We calculate the power of the output signal (noise component in this case) and denote it P_B. The pair (r_B, P_B), represented as a green square in Figure 3.6, correspond to the noise peak in the DST spectrum.

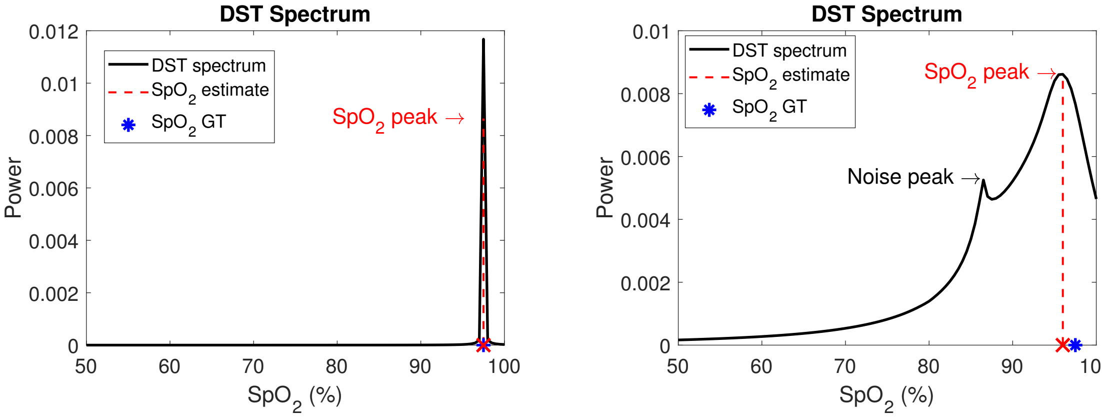
We continue decreasing the value of 'r' until we reach the value $r = r_C = 0.56$ which is very close to r_{SpO_2} ($r = 0.56$ corresponds to 96% when plugged in Equation (3.2)). Figure 3.7 is used for illustration of this case (case 3). At this instance, the reference signal calculated using Equation (3.8) is

$$reference\ signal(t) = (r_C - r_{SpO_2}) \times S_{infra}(t) + (r_C - r_{Noise}) \times N_{infra}(t), \quad (3.15)$$

and since r_C is very close to r_{SpO_2}, Equation (3.15) can be re-written as

$$reference\ signal(t) = (r_{SpO_2} - r_{Noise}) \times N_{infra}(t). \tag{3.16}$$

Note from Equation (3.16) that the reference signal contains only the noise components, $(r_{SpO_2} - r_{Noise}) \times N_{infra}(t)$. The ANC will remove the noise component from the input signal, $infra(t)$. Therefore, the output of the ANC will be the desired signal component. We calculate the power of the output signal (desired signal in this case) and denote it P_C. The pair $(r_C,\ P_C)$, represented as a blue square in Figure 3.7, correspond to the SpO_2 peak in the DST spectrum. Note that the SpO_2 estimated using DST algorithm was 96% while the ground truth was 97.5%. This error is caused by noise that contaminated the red and infrared PPG signals.



(a) (b)

Figure 3.8 (a) DST spectrum for a clean synthetic PPG signals – (b) DST spectrum for a noise-contaminated synthetic PPG signals (SNR = 0 dB). The blue '*' is the SpO_2 ground truth and the red 'x' is the SpO_2 level that the DST algorithm calculates. The SpO_2 ground truth for both subplots was 97.5%.

Figure 3.8 shows the DST spectrum for clean (Figure 3.8 (a)) and noise-contaminated (Figure 3.8 (b)) synthetic PPG signals (for the way synthetic PPG signals were generated, see Subsection 3.3.1). In the case of a clean PPG signal, the DST spectrum shows only one peak. Its location corresponds to the SpO_2 level estimate (Figure 3.8 (a)). The SpO_2 level calculated by the DST algorithm for this synthetic PPG signals (the red ‘x’ on Figure 3.8 (a)) matches the SpO_2 ground truth (blue ‘*’ on Figure 3.8 (a)). For the noisy signal, two distinct peaks will typically appear, as shown in Figure 3.8 (b). One peak corresponds to SpO_2 level and the other peak is due to noise. The right-most peak is considered to correspond the SpO_2 peak. The SpO_2 level calculated by the DST algorithm in the example used for Figure 3.8 (b) (the red ‘x’) slightly deviates from the ground truth (blue ‘*’). This difference is caused by noise.

3.3 Generation of PPG and ECG Signals

3.3.1 Synthetic data generation

Photoplethysmography (PPG) signals: As mentioned in Subsection 2.3.4, in order to study behavior, performance, and tradeoffs in the design of SpO_2 estimators, we developed a synthetic PPG signal generator. We modeled the PPG red and infrared waveforms as the sum of a constant DC component and an AC component. The AC and DC components are shown in Equations (2.8a) and (2.8b), respectively.

Electrocardiography (ECG) signals: The ECG signal is generally composed of P, QRS, T, and U waves [25]. Islam et al. [26] modeled a synthetic clean ECG waveforms by assuming that the QRS, Q, and S portions of the ECG signal can be represented by triangular waveforms, and the P, T, and U portions can be represented by the positive half period of a sinusoidal waveform. Their model was used in this study to generate the synthetic ECG waveforms [27]. Figure 3.9 shows a synthetically generated ECG signal at a fundamental frequency (HR frequency) of 1 Hz.

3.3.2 Human participant data – NAWCAD experiment

We have used human participant data collected from an experiment conducted in 2018-2019 by the United States Naval Air Warfare Center Aircraft Division (NAWCAD). Data were collected from fourteen (14) test participants who were

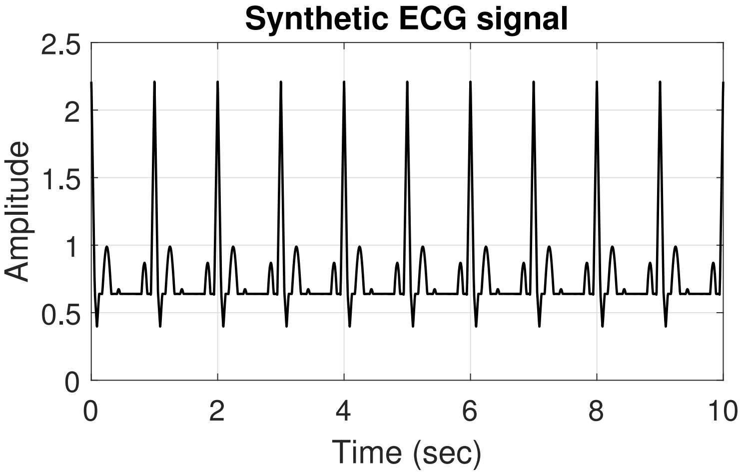


Figure 3.9 Synthetic ECG signal.

briefed and provided informed consent. All participants were volunteers, military or civil-service personnel in the employ of the US federal government. All participants were non-smokers and were reviewed by a medical monitor to ensure that they were physically fit to participate[3].

Each participant was exposed to the following profile of activities (Figure 3.10).

1. Warm-up – a warm-up and stretching session, designed to raise and maintain Target Heart Rate (THR)[4] intensity of 50% - 60% (approximately 5 minutes).
2. Aerobic – run on treadmill/bike session to raise and maintain THR of 60% - 80% (approximately 12 minutes), followed by active recovery (approximately 6 minutes), designed to decrease the heart rate by a reduced-intensity jog/walk or cycling activity.
3. Anaerobic – two sessions, each consisting of three 90-second periods of strenuous effort (treadmill/bike) designed to raise and maintain THR of 80% - 90%, separated from each other by one (1) minute of active recovery, designed to decrease heart rate by a reduced-intensity jog/walk or cycling; the two sessions were separated from each other by a period of active recovery (approximately 6 minutes). Next, the participant undertook a period of active recovery (around 4 minutes).

---

[3] Study approved by Naval Air Warfare Center Aircraft Division IRB, protocol FWR21070114H, original approval date: 12 June 2017. Air Force Research Lab (AFRL) IRB protocols comply with DoD Directive 3216.02, Title 25, CFR 46, and are in compliance with the Declaration of Helsinki Revision 6, 2008.

[4] The target Heart Rate (THR) is determine using the Karvonen formula [28] THR = $((HR_{max} - HR_{rest}) \times (\%\text{intensity})) + HR_{rest}$, where HR_{max} = 208 - 0.7 × age [29]

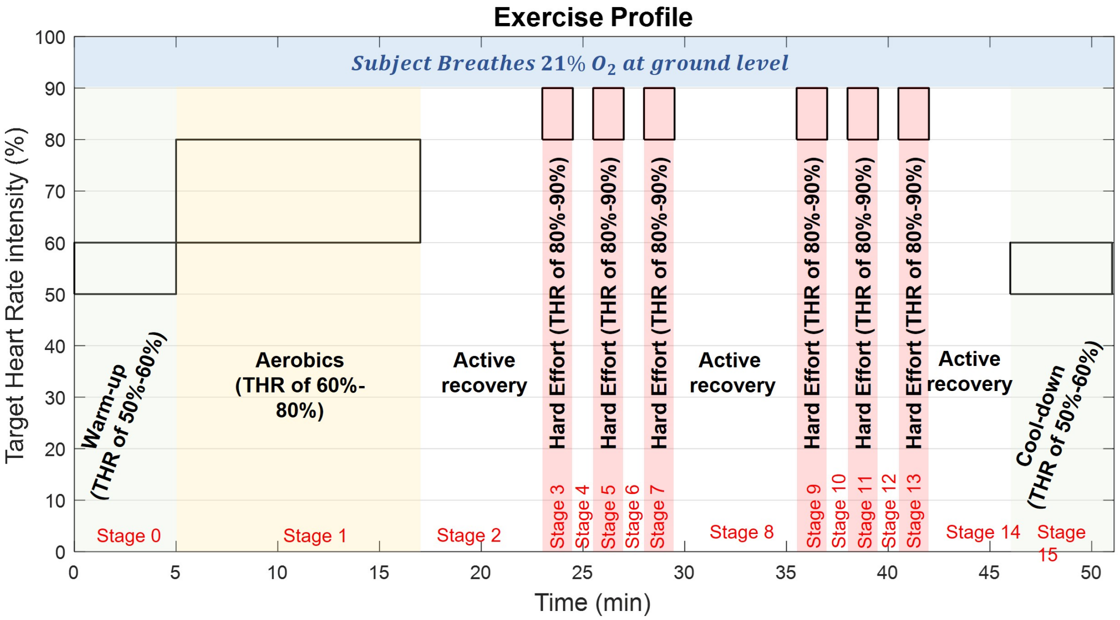


Figure 3.10 Illustration of the phases of the exercise experimental profile along with Target Heart Rate (THR) intensity. Black boxes indicate the range of THR in each exercise phase.

4. Cool-down – a cool-down and stretching session to maintain a THR of 50% - 60% (about 5 minutes).

Figure 3.10 shows all stages of this regime, marked stage 0 to stage 15. During the experimental runs, signals were recorded from a prototype dual pulse oximeter system with an accelerometer worn on the arm developed by Athena GTX (Holistic Modular Aircrew Physiologic Status (HMAPS) Monitoring System), and ECG electrodes affixed to the chest. Additionally, a Nonin 8000R reflectance pulse oximetry sensor was mounted on the participant's temple and connected by wires to the Nonin Wrist-Ox 3150 for data processing. The 8000R sensor was mounted on the participant's temple in a solid and stable manner that minimized sensor movements or dislocations on the temple surface even during strenuous exercise. SpO_2 levels calculated by the temple Nonin 8000R sensor were considered the "ground truth" due to the high accuracy of this sensor[5]. Performance of the arm-mounted oximeters was

---

[5] In general, reflectance pulse oximetry, such as the method used by Nonin 8000R is known to be much less vulnerable to artifacts (including motion artifacts). The

assessed with respect to the readings of the temple-mounted $8000R$ sensor (Subsection 3.4.2).

3.4 Impact of a Heart-Rate Tuned Comb Filter on Peripheral Blood Oxygen Saturation Measurement Performance

3.4.1 Peripheral blood oxygen saturation level calculations on synthetic data[6]

We set out to assess the effect of the heart-rate tuned comb filter on accuracy of calculated SpO_2 levels. To this end, we employed the RoI approach and the DST-based algorithm on a 10-second long synthetic PPG waveforms, with and without pre-filtering of the PPG signal by a comb filter. SpO_2 level and HR were set to 97.5% and 60 bpm, respectively. We modeled the motion artifact noise, N, as a bandpass (0.5-5 Hz) filtered AWGN (as was done in [29]), added to the clean PPG signal, S.

In Figure 3.11 we show the Root Mean Square Error (RMSE) of SpO_2 levels calculated from synthetic PPG signals for an SNR ranging from -10 dB to +10 dB. We used the RoI approach and the DST-based algorithm, both before and after processing the synthetic PPG signals with the heart-rate tuned comb filter. The RMSE is defined as

$$RMSE(s) = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(SpO_{2cal}(r) - SpO_{2GT}\right)^2}, \tag{3.17}$$

where '$SpO_{2cal}(r)$' is one value (out of the 1000 realizations) of the SpO_2 level calculated using one of the tested methods (RoI or DST-based algorithm, with and

---

manufacturer reports that SpO_2 accuracy of the Model $8000R$ sensor was determined through an induced hypoxia study on healthy participants over the range of 70% to 100% [30]. The resulting SpO_2 accuracy was $\pm 2A_{rms}$ in the range 80-100% and $\pm 3A_{rms}$ in the range 70-80%. $\pm 1A_{rms}$ encompasses 68% of the population at zero bias.

[6] In this section we have used calibration curve Equation (3.2). We have repeated the calculation for calibration curves of Equations (3.3) and (3.4) and the trends and conclusion remain the same.

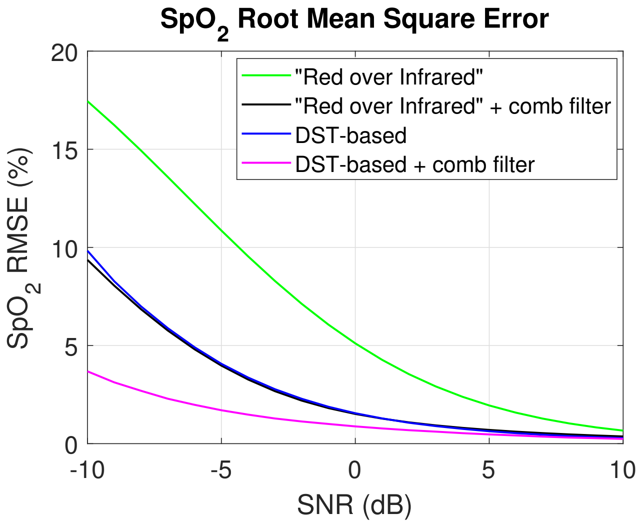


Figure 3.11 SpO_2 Root Mean Square Error using RoI approach and the DST-based algorithm, with and without a comb filter.

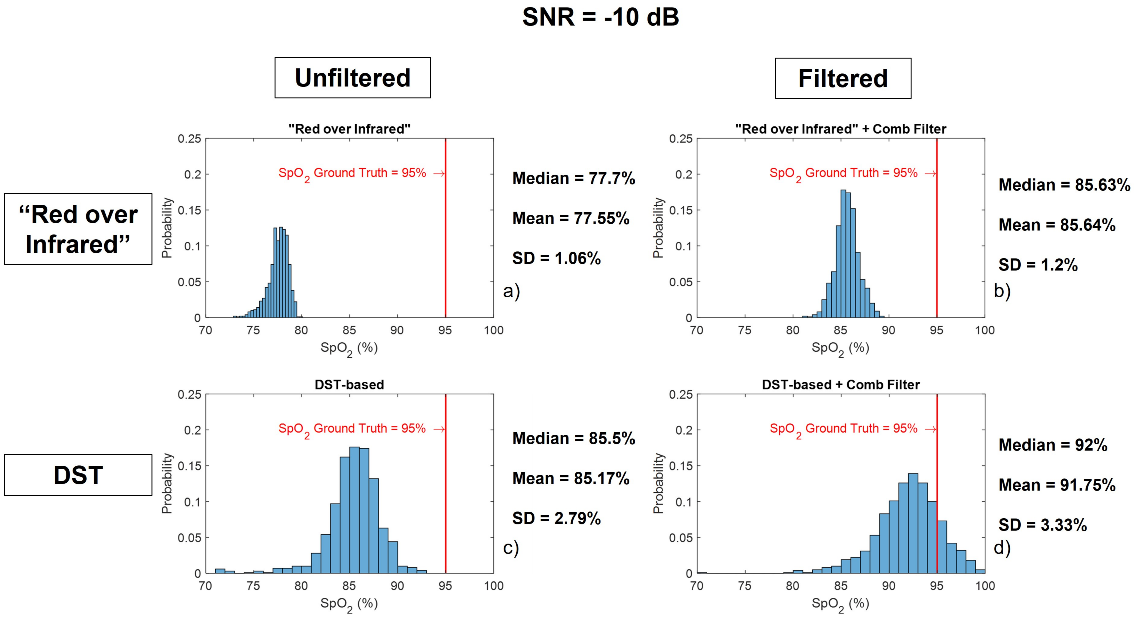


Figure 3.12 Histogram of 1000 SpO_2 levels calculated from red and infrared PPG signal with SNR=-10 dB using a) RoI approach – b) RoI approach preceded by a heart-rate tuned comb filter – c) DST-based algorithm – d) DST-based algorithm preceded by a heart-rate tuned comb filter.

without comb filtering); ‘SpO_{2GT}’ is the ground truth SpO_2 value; ‘R’ is the number of SpO_2 levels calculated for each SNR, and ‘s’ the SNR value at which the RMSE is calculated. In our study, SpO_{2GT}=95%, R=1000 realizations, and s = -10, -9, . . . , -1, 0, 1, . . . , 9, 10 (dB).

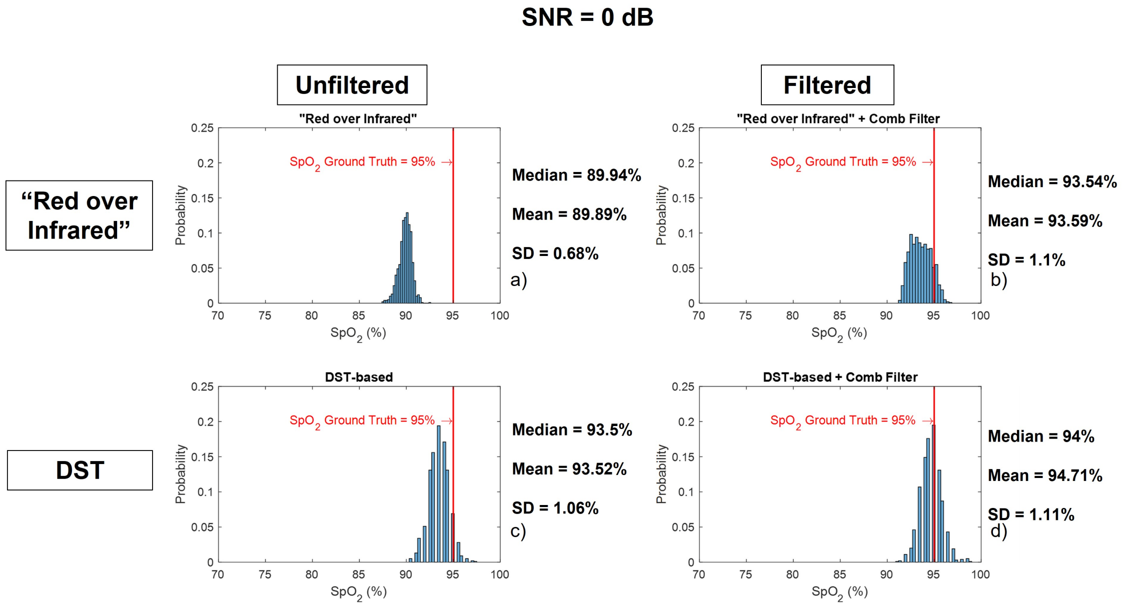


Figure 3.13 Histogram of 1000 SpO_2 levels calculated from red and infrared PPG signal with SNR=0 dB using a) RoI approach – b) RoI approach preceded by a heart-rate tuned comb filter – c) DST-based algorithm – d) DST-based algorithm preceded by a heart-rate tuned comb filter.

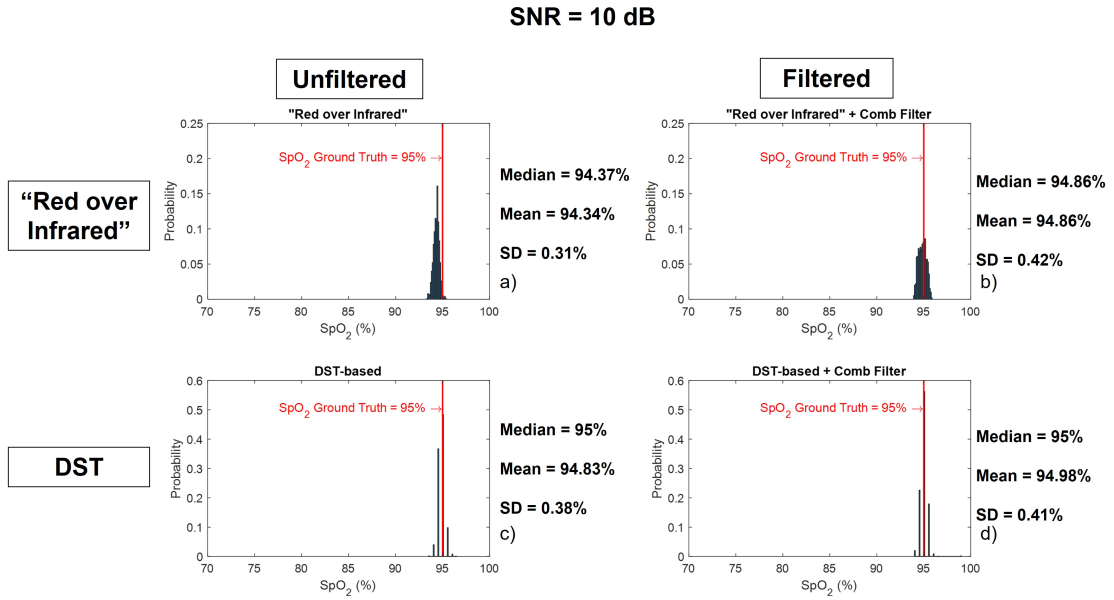


Figure 3.14 Histogram of 1000 SpO_2 levels calculated from red and infrared PPG signal with SNR=10 dB using a) RoI approach – b) RoI approach preceded by a heart-rate tuned comb filter – c) DST-based algorithm – d) DST-based algorithm preceded by a heart-rate tuned comb filter.

The main conclusion from Figure 3.11 is that the unfiltered RoI approach is inferior to the other approaches (namely filtered RoI and DST-based algorithm (filtered or unfiltered)). The filtered RoI approach and the unfiltered DST-based algorithm are comparable. The best performance was obtained by the filtered DST-based algorithm.

Figures 3.12, 3.13, and 3.14 are the histograms of SpO_2 levels (1000 SpO_2 calculations each) calculated using RoI approach and the DST-based algorithm, both with and without processing the synthetic PPG signals with the heart-rate tuned comb filter. We show results for a SNR of -10 dB (Figure 3.12), 0 dB (Figure 3.13), and 10 dB (Figure 3.14). The histograms get closer to the SpO_2 ground truth (SpO_2 of 95%) when the synthetic PPG signals are preprocessed by the heart-rate tuned comb filter for both SpO_2 calculation methods (RoI and the DST). We conclude that (1) the use of a heart-rate tuned comb filter has improved the performance of both methods (RoI and the DST), and (2) the best performance is obtained with the DST-based algorithm preceded by a heart-rate tuned comb filter. Next best is the RoI approach with comb filtering and unfiltered DST-based algorithm (SpO_2 levels estimated with these two techniques are comparable); the worst performance was the unfiltered RoI approach.

3.4.2 Peripheral blood oxygen saturation level calculations on experimental data

Data analysis using the AFE4490 manufacturer calibration curve Equation (3.2)

We tested the RoI approach and the DST-based algorithm on data collected from human participants, with and without comb filtering. Data were collected from fourteen (14) human participants for the regime described in Subsection 3.3.2 and Figure 3.10. Every two seconds, the preceding 10-second long data segment was processed. We show the results of one of the fourteen (1 of 14) participants on

Figure 3.15 (we covered the time period from 5 to 35 minutes of the exercise profile, corresponding to stages 1 to 8 in Figure 3.10). A detailed description of the traces in Figure 3.15 is provided in Tables 3.2, 3.3, 3.4, and 3.5[7]. In all cases, the standard against which the various methods were assessed was the readings of the Nonin 8000R sensor (light blue trace in Figure 12 (b), (c), (d)).

Table 3.2 Detailed Description on all Traces Shown in Subplot (a) of Figure 3.15

	Trace label	**Trace color**	**Trace name**	**Description**
Figure 3.15 (a)	a_1	Blue	X	x-axis accelerometer
	a_2	Red	Y	y-axis accelerometer
	a_3	Orange	Z	z-axis accelerometer

Table 3.3 Detailed Description on all Traces Shown in Subplot (b) of Figure 3.15

	Trace label	**Trace color**	**Trace name**	**Description**
Figure 3.15 (b)	b_1	Light blue	Nonin SpO_2	SpO_2 levels calculated by Nonin 8000R sensor
	b_2	Green (top curve)	SpO_2 “RoI”	SpO_2 levels calculated using “RoI” approach before comb filter.
	b_3	Black (top curve)	SpO_2 “RoI” + Comb Filter	SpO_2 levels calculated using “RoI” approach after comb filter.
	b_4	Green (bottom curve)	SpO_2 Difference between Nonin and “RoI”	SpO_2 Difference between levels calculated by Nonin 8000R sensor and “RoI” approach before comb filter.
	b_5	Black (bottom curve)	SpO_2 Difference between Nonin and “RoI” + Comb Filter	SpO_2 Difference between levels calculated by Nonin 8000R sensor and “RoI” approach after comb filter.

---

[7] Data for all 14 participants are available at https://github.com/moshekam/PPG-Exercise-Experimental-Data, retrieved on 20 February 2023.

Table 3.4 Detailed Description on all Traces Shown in Subplot (c) of Figure 3.15

	Trace label	**Trace color**	**Trace name**	**Description**
Figure 3.15 (c)	c_1	Light blue	Nonin SpO_2	SpO_2 levels calculated by Nonin $8000R$ sensor
	c_2	Dark blue (top curve)	SpO_2 DST-based	SpO_2 levels calculated using DST-based algorithm before comb filter.
	c_3	Magenta (top curve)	SpO_2 DST-based + Comb Filter	SpO_2 levels calculated using DST-based algorithm after comb filter.
	c_4	Dark blue (bottom curve)	SpO_2 Difference between Nonin and DST-based	SpO_2 Difference between levels calculated by Nonin $8000R$ sensor and DST-based algorithm before comb filter.
	c_5	Magenta (bottom curve)	SpO_2 Difference between Nonin and DST-based + Comb Filter	SpO_2 Difference between levels calculated by Nonin $8000R$ sensor and DST-based algorithm after comb filter.

Table 3.5 Detailed Description on all Traces Shown in Subplot (d) of Figure 3.15

	Trace label	**Trace color**	**Trace name**	**Description**
Figure 3.15 (d)	d_1	Light blue	Nonin SpO_2	SpO_2 levels calculated by Nonin $8000R$ sensor
	d_2	Magenta (top curve)	SpO_2 DST-based + Comb Filter	SpO_2 levels calculated using DST-based algorithm after comb filter.
	d_3	Black (top curve)	SpO_2 "RoI" + Comb Filter	SpO_2 levels calculated using "RoI" approach after comb filter.
	d_4	Magenta (bottom curve)	SpO_2 Difference between Nonin and DST-based + Comb Filter	SpO_2 Difference between levels calculated by Nonin $8000R$ sensor and DST-based algorithm after comb filter.
	d_5	Black (bottom curve)	SpO_2 Difference between Nonin and "RoI" + Comb Filter	SpO_2 Difference between levels calculated by Nonin $8000R$ sensor and "RoI" approach after comb filter.

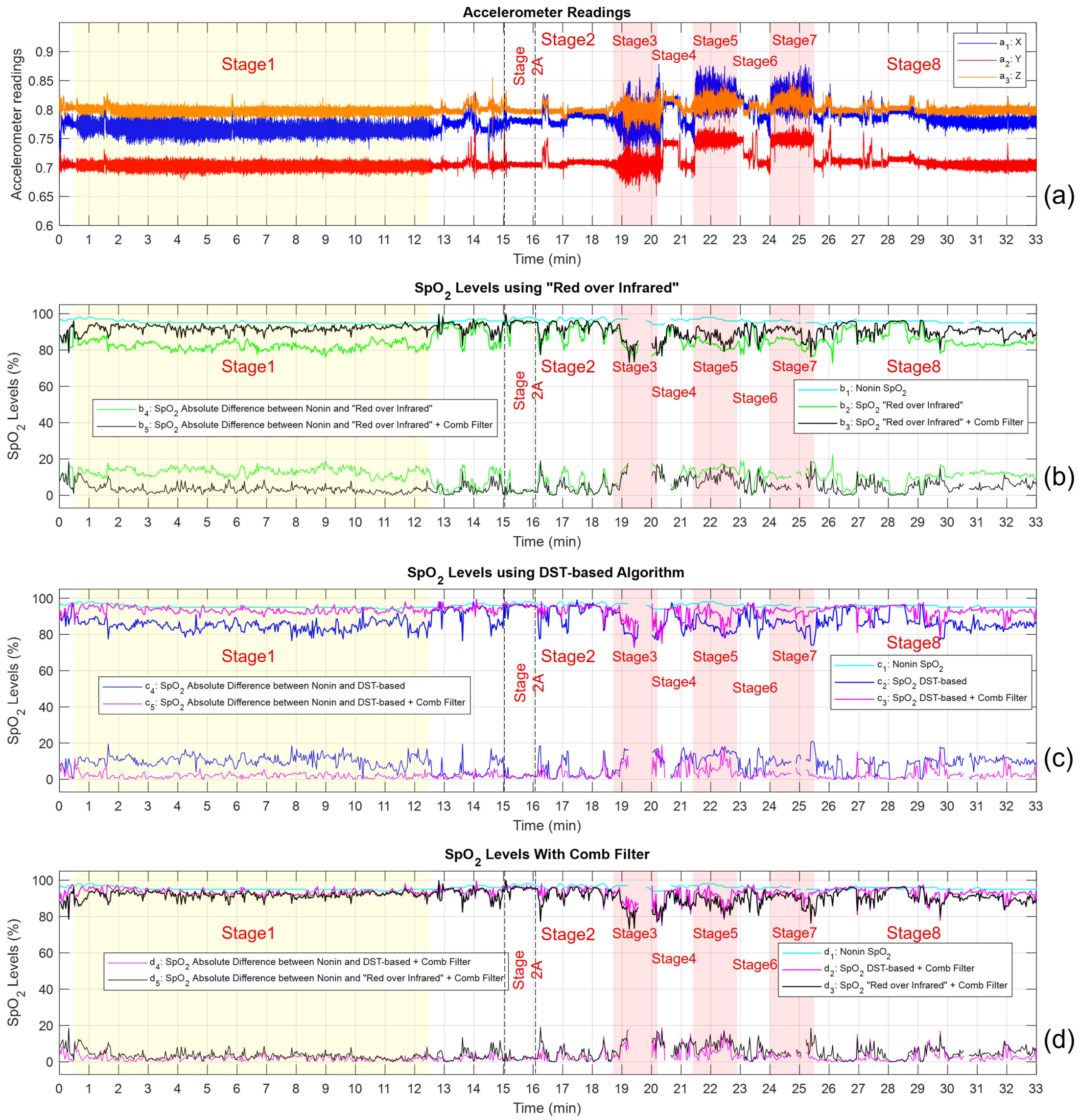


Figure 3.15 SpO_2 level calculations for an exercising participant on stages 1 to 8 (see Subsection 3.3.2 and Figure 3.10). We used the RoI approach and DST-based algorithm with and without preprocessing of the PPG signals with the heart-rate tuned comb filter. SpO_2 levels calculated using the abovementioned algorithms is compared to the SpO_2 levels calculated by Nonin $8000R$ sensor. The legends are fully explained in Tables 3.2, 3.3, 3.4, and 3.5.

The gaps in SpO_2 curves on Figure 3.15 (for example: In subplots b, c, and d from 19.5 to 20 minutes for the "Red over Infrared" approach and the DST-based algorithm, and from around time 20.5 minutes for Nonin $8000R$ sensor) indicate a failure in calculating the SpO_2 level caused by physical loss of the PPG signal. We skipped these gaps in our analysis.

Examination of Figure 3.15 suggests that the addition of a comb filter improved performance for both oximeters based on RoI and oximeters based on DST algorithm (Figure 3.15 second and third trace). Also, performance of oximeters employing RoI + comb filter was very close to performance of oximeters employing DST + comb filter (Figure 3.15 fourth trace).

Table 3.6 Overall Mean and Standard Deviation (SD) of the SpO_2 Absolute Error Calculated for RoI Approach and DST-based Algorithm With and Without Comb Filtering for all 14 Exercise Participants

	RoI		RoI + Comb		DST		DST + Comb	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Stage 1	14.03	1.54	9.71	1.72	11.34	1.52	7.56	1.64
Stage 2	7.05	1.37	4.74	1.15	4.76	1.1	3.34	0.93
Stage 3	10.49	0.87	7.59	1.04	8.82	1.02	5.64	1.04
Stage 4	9.82	1.56	6.18	1.29	5.85	1.23	4.07	1.13
Stage 5	11.6	1.11	7.49	1.09	8.32	1.03	5.01	0.99
Stage 6	9.01	1.42	4.86	1.11	5.52	1.19	3.09	0.9
Stage 7	11.2	0.96	7.72	1.49	7.91	1.02	4.46	1.03
Stage 8	8.14	1.49	5.03	1.02	6.1	1.58	3.45	0.99

Table 3.6 provides the overall mean and standard deviation of the SpO_2 error for all fourteen (14) participants we studied in all eight (8) exercise profile stages (stages 1 to 8 in Figure 3.10). The SpO_2 error is defined as absolute difference between SpO_2 levels measured by Nonin $8000R$ sensor sensor mounted on the participant's temple (the standard) and SpO_2 levels calculated by the two algorithms. We performed six (6) two-samples t significance tests with level of significance α=0.01. The tested

hypotheses and their corresponding p-values are shown in Table 3.7 and Table 3.8, respectively.

Table 3.7 Tested Hypotheses and Results

	Is the SpO_2 mean absolute error of ...		**SpO_2 mean absolute error of ...?**		
Test I:	RoI *without* comb filter ($\nu_{1,A}$)	greater than	RoI *with* comb filter ($\nu_{2,A}$)	H_0: $\nu_{1,A}$ - $\nu_{2,A} = 0$ H_a: $\nu_{1,A}$ - $\nu_{2,A} > 0$	**Yes**. (test I, Tables 3.8, 3.9, and 3.10)
Test II:	RoI *without* comb filter ($\nu_{1,A}$)	greater than	DST *without* comb filter ($\nu_{3,A}$)	H_0: $\nu_{1,A}$ - $\nu_{3,A} = 0$ H_a: $\nu_{1,A}$ - $\nu_{3,A} > 0$	**Yes**. (test II, Tables 3.8, 3.9, and 3.10)
Test III–a:	RoI *with* comb filter ($\nu_{2,A}$)	different than	DST *without* comb filter ($\nu_{3,A}$)	H_0: $\nu_{2,A}$ - $\nu_{3,A} = 0$ H_a: $\nu_{2,A}$ - $\nu_{3,A} \neq 0$	**No**[a]. (test III-a, Tables 3.8, 3.9, and 3.10)
Test III-b:	DST *without* comb filter ($\nu_{3,A}$)	greater than	RoI *with* comb filter ($\nu_{2,A}$)	H_0: $\nu_{3,A}$ - $\nu_{2,A} = 0$ H_a: $\nu_{3,A}$ - $\nu_{2,A} > 0$	**No**[b]. (test III-b, Tables 3.8, 3.9, and 3.10)
Test IV:	RoI *with* comb filter ($\nu_{2,A}$)	greater than	DST *with* comb filter ($\nu_{4,A}$)	H_0: $\nu_{2,A}$ - $\nu_{4,A} = 0$ H_a: $\nu_{2,A}$ - $\nu_{4,A} > 0$	**Yes**. (test IV, Tables 3.8, 3.9, and 3.10)
Test V:	DST *without* comb filter ($\nu_{3,A}$)	greater than	DST *with* comb filter ($\nu_{4,A}$)	H_0: $\nu_{3,A}$ - $\nu_{4,A} = 0$ H_a: $\nu_{3,A}$ - $\nu_{4,A} > 0$	**Yes**. (test V, Tables 3.8, 3.9, and 3.10)

[a] For calibration curve of Equation (3.2), the answers are:
For stages 1, 2, 4, 5, 6, 7, and 8, the answer is No. For stage 3, the answer is Yes. Overall answer: No

For calibration curve of Equations (3.3) and (3.4), the answer is No for all stages.

[b] For calibration curve of Equation (3.2), the answers are:
For stages 2, 4, 5, 6, 7, and 8, the answer is No. For stages 1 and 3 the answer is Yes. Overall answer: No

For calibration curve of Equations (3.3) and (3.4), the answers are:
For stages 2 through 8, the answer is No. For stage 1, the answer is Yes. Overall answer: No

We concluded from Table 3.7 and Table 3.8 the following:

- The mean absolute error using RoI was larger than the mean absolute error using RoI + Comb in all eight (8) stages;
- The mean absolute error using RoI was larger than the mean absolute error using DST in all eight (8) stages;
- The mean absolute error using RoI + Comb was larger than the mean absolute error using DST + Comb in all eight (8) stages;
- The mean absolute error using DST was larger than the mean absolute error using DST + Comb in all eight (8) stages.

Additionally, we concluded from Table 3.7 and Table 3.8 (tests III-a and III-b) that the mean absolute errors using RoI + Comb and using DST are comparable during most stages.

Table 3.8 P-values of all Six Tests

	p-value (significance level was α = 0.01)					
	Test I	Test II	Test III-a	Test III-b	Test IV	Test V
Stage 1	< 0.001	<0.001	0.020	<0.010	0.002	< 0.001
Stage 2	< 0.001	< 0.001	0.010	0.482	0.002	0.001
Stage 3	< 0.001	< 0.001	0.008	0.004	<0.001	< 0.001
Stage 4	< 0.001	< 0.001	0.251	0.250	< 0.001	<0.001
Stage 5	< 0.001	< 0.001	0.059	0.029	<0.001	< 0.001
Stage 6	< 0.001	< 0.001	0.152	0.077	<0.001	<0.001
Stage 7	< 0.001	< 0.001	0.351	0.350	< 0.001	< 0.001
Stage 8	0.0011	0.002	0.026	0.265	< 0.001	<0.001

Data analysis using alternate calibration curves Equations (3.3) and (3.4)

We replicate Table 3.8 (p-values) for the alternate calibration curves, namely; the Beer-Lambert calibration curve of Equation (3.3); and the underestimation calibration curve of Equation (3.4). The results are shown in Tables 3.9 and 3.10, respectively. Tables 3.9 and 3.10 support the conclusions in Table 3.7 for the alternate calibration curves as well (see Table 3.7 footnotes a and b). The relative performance of the SpO_2 calculation techniques and impact of the comb filter are qualitatively the same, regardless of the specific calibration curve (whether we used Equation (3.2), (3.3), or (3.4)).

Table 3.9 P-values of all Six Tests for "Beer-Lambert Calibration Curve" (Equation (3.3))

	p-value (significance level was $\alpha = 0.01$)					
	Test I	Test II	Test III-a	Test III-b	Test IV	Test V
Stage 1	< 0.001	< 0.001	0.011	0.006	<0.001	< 0.001
Stage 2	< 0.001	< 0.001	0.961	0.481	< 0.001	< 0.001
Stage 3	< 0.001	< 0.001	0.079	0.040	< 0.001	< 0.001
Stage 4	< 0.001	< 0.001	0.515	0.258	< 0.001	< 0.001
Stage 5	< 0.001	< 0.001	0.110	0.055	< 0.001	< 0.001
Stage 6	< 0.001	< 0.001	0.268	0.134	< 0.001	< 0.001
Stage 7	< 0.001	< 0.001	0.164	0.082	< 0.001	< 0.001
Stage 8	< 0.001	< 0.001	0.289	0.145	< 0.001	< 0.001

Table 3.10 P-values of all Six Tests for "Underestimation Calibration Curve" (Equation (3.4))

	p-value (significance level was $\alpha = 0.01$)					
	Test I	Test II	Test III-a	Test III-b	Test IV	Test V
Stage 1	< 0.001	< 0.001	0.010	0.005	0.001	< 0.001
Stage 2	< 0.001	< 0.001	0.983	0.491	< 0.001	< 0.001
Stage 3	< 0.001	< 0.001	0.073	0.036	< 0.001	< 0.001
Stage 4	< 0.001	< 0.001	0.494	0.247	< 0.001	< 0.001
Stage 5	< 0.001	< 0.001	0.156	0.078	< 0.001	< 0.001
Stage 6	< 0.001	< 0.001	0.284	0.142	< 0.001	< 0.001
Stage 7	< 0.001	< 0.001	0.189	0.095	< 0.001	< 0.001
Stage 8	< 0.001	< 0.001	0.244	0.122	< 0.001	< 0.001

3.4.3 Computational complexity

Red over Infrared approach: For red and infrared signals of length B (B sampled values) the number of operations required for a typical RoI calculation is $2B$ multiplications, $2(B-1)$ additions, three divisions and two square root operations. For each B (in our case in the thousands), the required computation time is proportional to B.

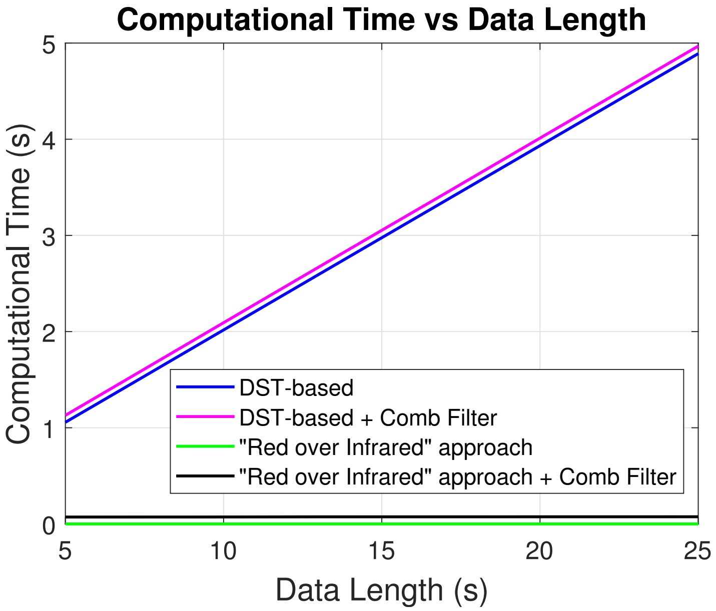


Figure 3.16 Comparison of computational time of the RoI approach and the DST-based algorithm with and without comb filtering.

DST algorithm: At every time instance, the DST algorithm generates a family of D reference signals (in our study we used $D = 101$ reference signal for each optical density ratio of SpO_2 level, ranging from 50% to 100% with an increment of 0.5%). For each one of the D SpO_2 levels, the DST algorithm uses an ANC filter and additional computations to create one point on the DST graph of power vs. SpO_2 level (*e.g.*, Figure 3.8 (b) in this document and Figure 3 in [2]). For the ANC we have used the QR-decomposition-based least-squares lattice (QRD-LSL) adaptive filter algorithm, on account of its relative computational efficiency. The QRD-LSL requires $\boldsymbol{\mathcal{O}}(T)$ operations per time instance (we use B instances for block of data), where T is the number of taps in the adaptive filter [31]. The computational times (in seconds) of RoI approach and the DST-based algorithm (with and without comb filtering) are shown on the ordinate of Figure 3.16. The abscissa of Figure 3.16 is the red and infrared signals' length in second. In Table 3.11 we show the computational time and the SpO_2 RMSE (in %) calculated on a 10-second long synthetic PPG signals. We use the RoI approach and the DST-based algorithm, both before and after comb filtering in each case, for different SNR values (-10dB, 0dB, and 10dB)[8]. Results presented in Figure 3.16 and Table 3.11 indicate that the computational cost of the DST-based algorithm is considerably larger than that of the RoI approach. The effects of adding the comb filter are much smaller.

Table 3.11 Computational Time and SpO_2 RMSE Calculated on a 10-second Long PPG Signals Using RoI Approach and DST-based Algorithm Before and After Comb Filtering for a SNR of -10dB, 0dB, and 10dB

	SpO_2 RMSE (%)			**Comput. Time (s)**
	SNR = -10 dB	SNR = 0 dB	SNR = 10 dB	
RoI	17.4811	5.1564	0.7301	2.463×10^{-5}
RoI + Comb	9.4388	1.7844	0.4425	7.248×10^{-2}
DST-based	10.2173	1.8191	0.4135	2.017
DST-based + Comb	4.6482	1.1431	0.4056	2.092

3.5 Discussion and Conclusion

Preprocessing PPG signals with a heart-rate tuned comb filter improved the performance of the two tested SpO_2 calculation algorithms (namely, RoI approach and DST-based algorithm). We tested both algorithms on synthetic and experimental data.

- The most accurate technique was the filtered DST-based algorithm. At very low signal to noise (SNR) environments the filtered DST-based algorithm performed somewhat better on synthetic data compared to the other methods (up to 6% improvement in accuracy at minus 10 dB SNR over the unfiltered DST algorithm and the filtered RoI approach). However, this technique was costly in computations.
- Next best were the filtered RoI approach and the unfiltered DST-based algorithm, which provided similar accuracies. However, the DST-based algorithm was much costlier in computations compared to the filtered RoI approach.

---

[8] The results were generated by MATLAB R2018a on a personal computer, with an Intel Core^{TM} i5-8500 CPU running at 3.00 GHz, 8GB RAM and Windows 10 operating system.

- The least accurate performance was of the unfiltered RoI approach.

The overall conclusion is that if an uninterrupted high-quality heart rate measurement is available, then the pre-filtered RoI approach (using a heart-rate tuned comb filter) provides a preferred trade-off between SpO_2-level accuracy and computational complexity.

CHAPTER 4

MITIGATION OF MOTION ARTIFACTS IN PULSE OXIMETRY THROUGH REDUNDANT SENSORS

In this chapter, using field data collected from human participants during rest and while exhibiting medium to high-intensity physical exercises, we show how sensor redundancy can mitigate motion effects by employing two channels of red and infrared (IR) PPG signals. We demonstrate the advantage of redundancy using two SpO_2 calculation methods. These methods are widely used “Red over Infrared” (RoI) method and the Discrete Saturation Transform (DST) algorithm both, assisted by pre-filtering by a heart-rate tuned comb filter. The data were collected from individuals while at rest and while exercising. SpO_2 levels are calculated from two channels, each using a red and an IR PPG signal. We integrated the SpO_2 computed from the two channels with a Kalman filter (KF) and computed the SpO_2 mean absolute error (MAE) from each channel individually and after integrating the two channels. The SpO_2 MAEs of each one of the two channels separately were greater (and statistically significant) than the SpO_2 MAE obtained from the two channels integrated with a KF. We observed that when two channels of PPG (red and IR) signals were available, then combining SpO_2 levels obtained from these two channels exhibited a smoother and more reliable estimate of the blood oxygen saturation level as compared the SpO_2 levels computed from each channel separately. The reason is that if the data for the two channels are collected at sufficient physical distance from one another, the motion artifacts affecting the two channels are uncorrelated. The KF is then able to use past measurements and modeling of the SpO_2 dynamics to attenuate the effect of the motion artifacts.

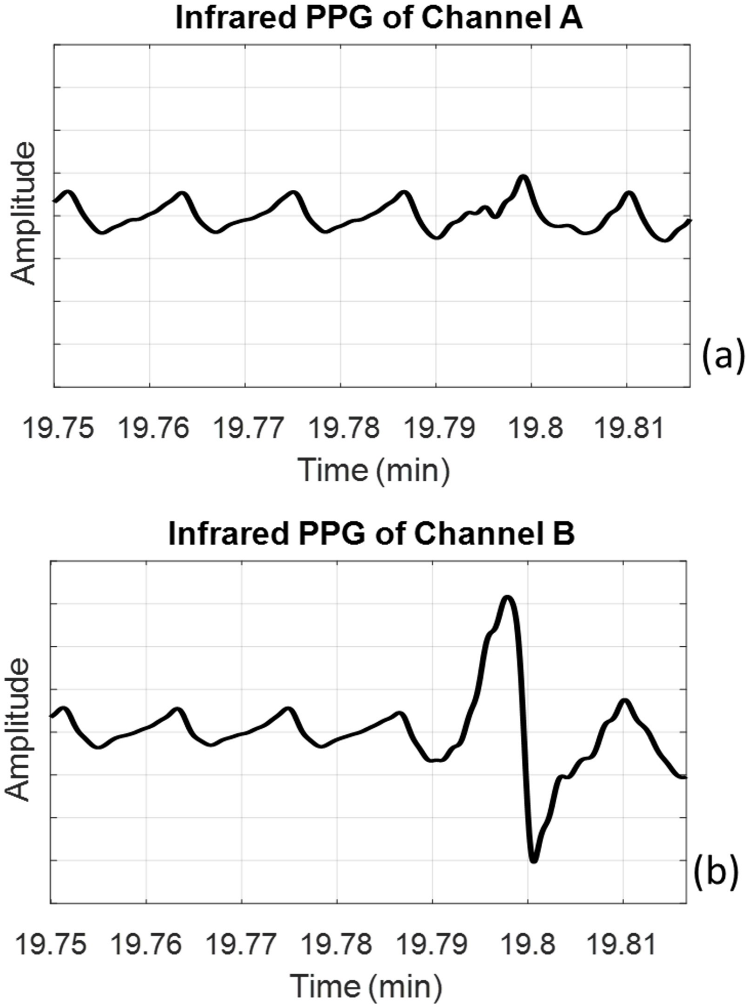


Figure 4.1 The impacts of motion artifacts on PPG signals when measured from two different sites on the person's upper arm region.

4.1 Motivation Behind Integrating Two Channels of SpO_2 Levels

The impact of person motion manifests itself differently in PPG sensors that are mounted on different parts of the body. The impacts are only weakly correlated, for example, when measured on the two different locations on the person's arm. For instance, we show on Figure 4.1 the IR PPG signal collected from two different channels (channels A and B) mounted on two different locations of a person's upper arm where the person was in motion. We see in Figure 4.1 that the IR signal of channel B was corrupted by noise around time 19.79 minutes. However, at the same time instance, no disruption appears in the IR signal of channel A. In general, integrating

SpO_2 levels computed by multiple sensors, each mounted on different part of the body, may result in a more reliable and smoother estimates as compared to individual sensor readings. In a different but relevant setting, Acharya *et al.* [32] showed that integrating SpO_2 levels measured from three commercial oximeters leads to a more reliable detection of the onset of Hypoxia. In Acharya *et al.* [32], participants were placed in an altitude chamber and exposed to a varying altitude profile ranging from 0 to 18,000 ft. SpO_2 levels were measured using three commercial oximeters, namely, Respironics Novametrix 515B, Nonin forehead pulse oximeter 9847, and Masimo Rad-87. The authors integrated the SpO_2 levels measured from the three sensors using a KF. In Acharya *et al.* [32], however, the effect of motion artifact was not studied, and the system's noise was dominated by other mechanisms having to do with the error profile of commercial oximeters measuring low SpO_2 levels. In this chapter, we study the effect of sensor redundancy on the impact of motion artifacts that contaminates the PPG signals when the person is engaged in medium to high intensity exercises (aerobic and anaerobic).

4.2 Methods of SpO_2 Integration

The methods we propose to compute SpO_2 levels from PPG waveforms are RoI method and DST-based algorithm, both preceded by a heart rate tuned comb filter (per Chapter 3). We assume the availability of an ECG waveform. The inputs are two channels of PPG red and IR signals and an ECG waveform. The output is SpO_2 levels. Figure 4.2 illustrates the SpO_2 calculation methods and is divided into four stages: (1) pre-processing, (2) filtering, (3) SpO_2 calculation, and (4) data fusion.

1. In the pre-processing stage, raw PPG signals of each channel are normalized (see Subsection 3.2.1 titled "normalization"). Simultaneously, HR is computed from an ECG waveform (see "heart rate calculation" Subsection 3.2.2).
2. In the filtering stage, a heart-rate tuned peak comb filter is applied to the normalized PPG signals. The comb filter exhibits its lowest attenuation at the heart rate frequency (which is computed from the ECG signal) and its principal

harmonics, and higher attenuation otherwise (see Subsection 3.2.3 titled "comb filter").

3. SpO_2 levels are computed from RoI method and DST-based algorithm, both preceded by a heart rate tuned comb filter (see Subsection 3.2.4 titled "SpO_2-level calculation").

4. At stage four, SpO_2 levels computed from the two channels are integrated using a KF. The integrated SpO_2 levels are obtained at the output of this stage (see Subsection 4.2.1 titled "Data Integration using a KF").

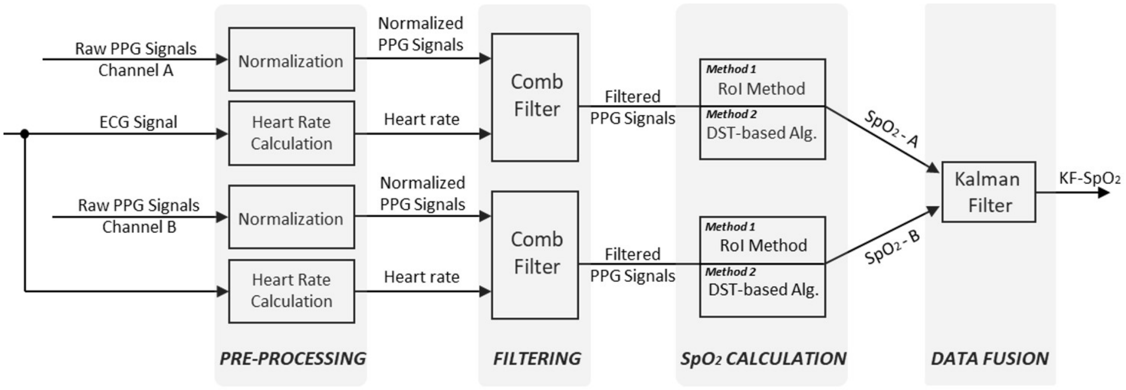


Figure 4.2 Block diagram of the proposed SpO_2-level calculation method for two PPG channels.

4.2.1 Data integration using a Kalman filter

We adopt the same model used by Galli *et al.* [21] to represent the dynamic behavior of human SpO_2. The model is based on discrete-time random walk process. Equations (4.1) and (4.2) represent the state and measurement equations, respectively, of the KF used to integrate the SpO_2 levels computed from two PPG channels.

$$x(k) = x(k-1) + w(k). \tag{4.1}$$

$$z(k) = x(k) + v(k). \tag{4.2}$$

In Equations (4.1) and (4.2), $x(k)$ is the current SpO_2 value, $z(k)$ is the measured SpO_2, $w(k)$ is the state zero-mean white Gaussian noise with variance σ_w^2, and $v(k)$ is the measurement zero-mean white Gaussian noise with variance σ_v^2. We tune the KF parameters, namely, σ_w^2 and σ_v^2, to obtain the best performance on our experimental data.

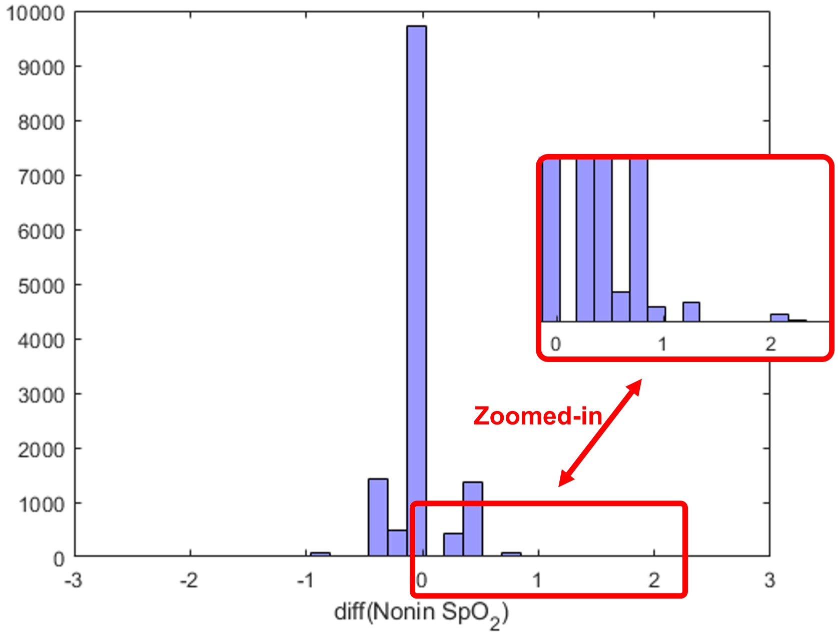


Figure 4.3 Histogram of the difference time series of the Nonin SpO_2 ground truth calculated for all fourteen (14) participants.

The state noise variance, σ_w^2, is obtained by examining the Nonin SpO_2 ground truth. The difference time series of the Nonin SpO_2 ground truth is calculated for all fourteen (14) participants. The distribution of this series is shown in the histogram of Figure 4.3. The majority of the values of the difference time series range between -1 and 1 (values greater than 1 or less than -1 occur with lower frequency – see the red box of Figure 4.3 which is a magnification of the right hand-side of Figure 4.3). We assume that the data of Figure 4.3 range between -1.45 and 1.45. As proposed by Galli *et al.* [21], we take a conservative approach and assume that the data of Figure 4.3 are distributed uniformly. The variance of the uniform distribution, denoted $\sigma_{U.D.}^2$, is obtained as following:

$$\sigma^2_{U.D.} = \frac{1}{12}(b-a)^2, \tag{4.3}$$

where b is the maximum value in the distribution and a is the minimum value. By applying Equation (4.3) to the data of Figure 4.3, we obtained $\sigma^2_w = \frac{1}{12}(1.45 - (-1.45))^2 = 0.7$.

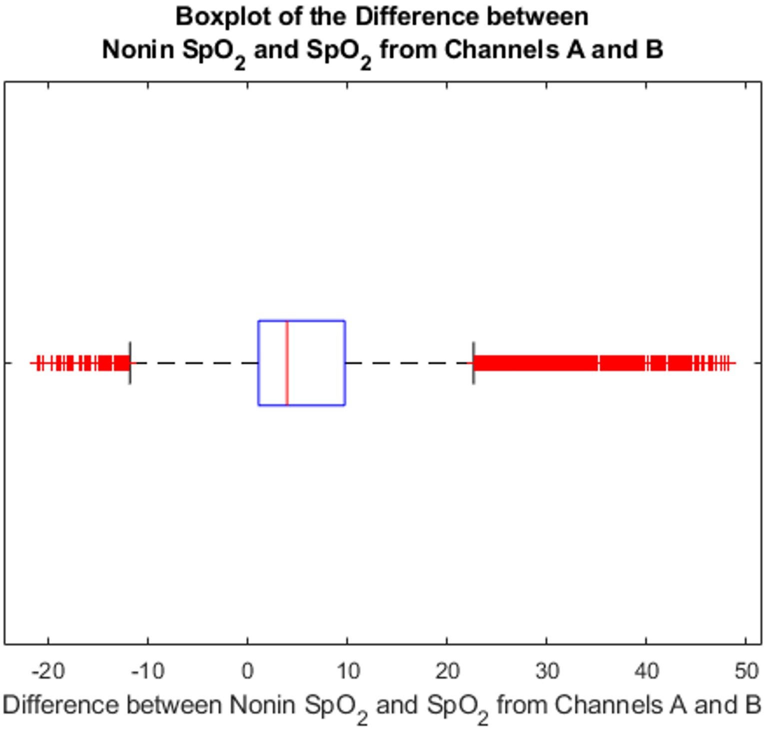


Figure 4.4 Boxplot of the time series that corresponds to the difference between the ground truth SpO_2 and the SpO_2 calculated from channels A and B for all fourteen (14) participants.

The measurement noise variance, σ^2_v, is estimated in a similar fashion by considering the time series that corresponds to the difference between the ground truth SpO_2 measured by the Nonin sensor and the SpO_2 levels obtained from channels A and B for all fourteen (14) participants. Figure 4.4 shows the boxplot of the time series that corresponds to the difference between the ground truth SpO_2 and the SpO_2 calculated from channels A and B. We calculate the 25th and 75th percentiles and found them to be 1.1 and 9.7, respectively. We consider the data ranging from

-3 and 12. As proposed by Galli *et al.* [21], we take a conservative approach and assume that the data of Figure 4.4 are distributed uniformly. The variance of the measurement noise is estimated using Equation (4.3) and found to be $\sigma_v^2 = 18.7$.

In summary, we found that $\sigma_w^2 = 0.7$ and $\sigma_v^2 = 18.7$ are the optimal values for our experiment data. The KF equations used for SpO_2 estimation at each iteration are summarized in Table 4.1.

Table 4.1 KF Equations Used for SpO_2 Estimation

Equation	**Function**
$x^-(k) = x^+(k-1)$	State prediction
$P^-(k) = P^+(k-1) + \sigma_w^2$	State prediction variance
$y(k) = z(k) - x^-(k)$	Innovation
$S(k) = P^-(k) + \sigma_v^2$	Innovation variance
$G(k) = \dfrac{P^-(k)}{P^-(k) + \sigma_v^2}$	Kalman gain
$x^+(k) = x^-(k) + G(k)y(k)$	Updated state estimate
$P^+(k) = G(k)\sigma_v^2$	Updated state estimate variance

At each KF iteration, one SpO_2 raw value is obtained from channel A (SpO_2-A) and another SpO_2 raw value is obtained from channel B (SpO_2-B). SpO_2-A and SpO_2-B at iteration k are denoted $z_1(k)$ and $z_2(k)$, respectively. As shown in Galli *et al.* [21], $z_1(k)$ and $z_2(k)$ go through a selection process where only the best measurement is passed onto the KF for smoothing. First, the variance of $z_1(k)$ and $z_2(k)$ values that had been computed over the last 3 minutes. If either one of the two variances is greater than two times the other variance, then the SpO_2 level computed from the channel with the highest variance is discarded and not considered from the following steps. Next, the innovation of each one of the two measurements, $z_1(k)$ and $z_2(k)$, is calculated. The measurement is accepted if and only if it satisfies the following condition:

$$|z_n(k) - x^-(k)| \leq \zeta\sqrt{S(k)},\ where\, n = 1, 2. \tag{4.4}$$

Here, "ζ" corresponds to a confidence level. In this work we used $\zeta = 2$ which corresponds to a confidence level of 95% under the assumption that the innovation follows a Gaussian distribution. After checking the condition in Equation (4.4) for both measurements $z_1(k)$ and $z_2(k)$, one of three cases will take place:

Case 1 – if both measurements satisfy the condition in Equation (4.4): The measurement with the smaller innovation is passed to the KF for smoothing.

Case 2 – if only one of the measurements satisfy the condition in Equation (4.4): The measurement that satisfies the condition in Equation (4.4) is passed to the KF for smoothing. The other measurement is discarded.

Case 3 – if none of the measurements satisfies the condition in Equation (4.4): Both measurements are discarded, and the updated state estimate is passed to the next iteration (e.g., $x^+(k) = x^-(k)$ and $P^+(k) = P^-(k)$).

If case 3 reoccurs five times consecutively, then the updated state estimate ($x^+(k)$) is obtained by averaging the predicted state estimate ($x^-(k)$) and all available new measurements. This step is performed to bring KF estimates progressively back to correct values [21].

4.3 Results

4.3.1 SpO_2 level calculation on experimental data

Data were collected from fourteen (14) participants for the regime described in Subsection 3.3.2 titled "Human Participant Data – NAWCAD Experiment." Every 2 seconds, the preceding 10-second-long data segment is used to compute an SpO_2 value.

SpO_2 levels are computed using the RoI approach and the DST-based algorithm, both preceded by a heart-rate tuned comb filter, from two channels of PPG signals – channel A and channel B – where each channel contains a set of one red and one IR sensors. In this chapter, SpO_2 values are calculated using the calibration Equation (3.2).

We show in Figure 4.5 the SpO_2 levels computed for two of the fourteen (2 out of 14) participants that covered the period from 5 to 35 minutes of the exercise profile (that correspond to stages 1 to 8 in Figure 3.10) described in Subsection 3.3.2 titled "Human Participant Data – NAWCAD Experiment." SpO_2 levels are calculated using the RoI approach preceded by a heart-rate tuned comb filter. The red and blue curves in Figure 4.5 are SpO_2 levels computed from channels A and B, respectively. The black curve is the result of integrating SpO_2 levels of channels A and B using the Kalman filter. The cyan curve is the readings of the Nonin 8000R sensor (our ground truth). We show the results of participant 1 and 2, in Figure 4.5 (a) and (b), respectively. The gaps in SpO_2 level curves on Figure 4.5 indicate that the PPG signal was physically lost and hence we were not able to computer an SpO_2 level at these times. By examining the traces of Figure 4.5 we conclude that the SpO_2 levels computed by integrating the SpO_2 levels of channels A and B are closer to the ground truth and exhibit less fluctuations as compared to the SpO_2 levels of channels A and B separately. We compute the overall mean and standard deviation of the SpO_2 error for all fourteen (14) participants in all eight (8) exercise profile stages (stages 1 to 8 in Figure 3.10). We defined SpO_2 error as the absolute difference between SpO_2 levels measured by Nonin 8000R sensor mounted on the participant's temple (the standard) and SpO_2 levels calculated by channel A, B, and KF. Tables 4.2 and 4.3 summarize the overall mean and standard deviation of the SpO_2 error for RoI approach and DST-based algorithm both preceded by a ECG-HR-tuned comb filter, respectively.

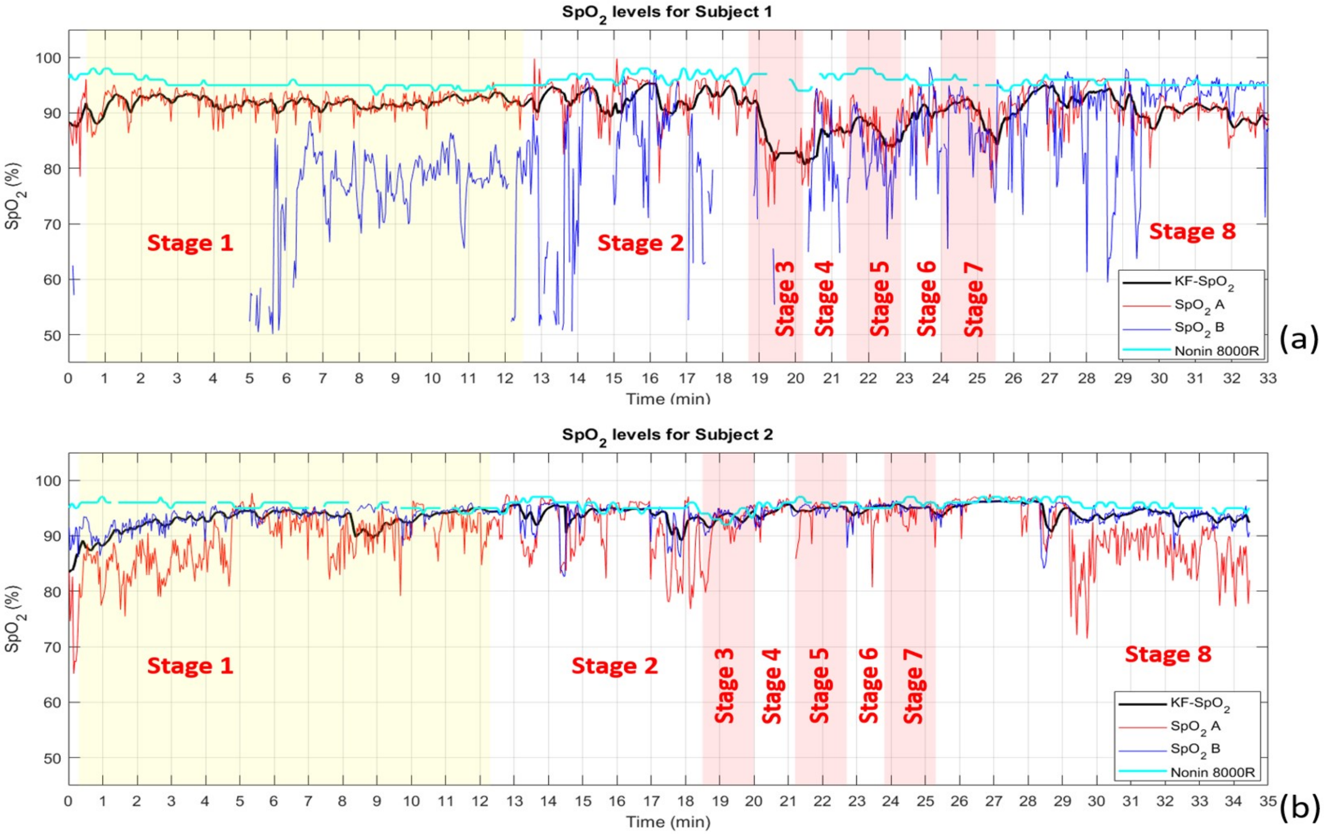


Figure 4.5 SpO_2 levels computed for two exercise participants. The method used for SpO_2 calculation is RoI preceded by a heart rate tuned comb filter. SpO_2 levels are computed from two channels of PPG signals (channel A in red and Channel b in blue). SpO_2 levels are then integrating using a KF (black trace). The cyan trace if the ground truth measured by Nonin 8000R sensor. Subplot (a) and (b) are the results of participants 1 and 2, respectively.

Table 4.2 Overall Mean and Standard Deviation (SD) of the SpO_2 MAE Calculated Using RoI + Comb for Channel A, B, and KF for all 14 Participants

	RoI + Comb SpO_2 MAE channel A		RoI + Comb SpO_2 MAE channel B		RoI + Comb SpO_2 MAE KF	
	Mean	SD	Mean	SD	Mean	SD
Stage 1	9.71	1.72	8.55	1.37	8.13	1.47
Stage 2	4.74	1.15	4.88	1.55	3.17	0.62
Stage 3	7.59	1.04	7.92	0.93	5.54	0.73
Stage 4	6.18	1.29	6.28	1.23	5.05	0.58
Stage 5	7.49	1.09	6.79	0.89	5.9	0.35
Stage 6	4.86	1.11	4.37	1.31	3.62	0.45
Stage 7	7.72	1.49	6.11	0.97	5.11	0.47
Stage 8	5.03	1.02	4.16	1.23	3.2	0.49

Table 4.3 Overall Mean and Standard Deviation (SD) of the SpO_2 MAE Calculated Using DST + Comb for Channel A, B, and KF for all 14 Participants

	DST + Comb SpO_2 MAE channel A		DST + Comb SpO_2 MAE channel B		DST + Comb SpO_2 MAE KF	
	Mean	SD	Mean	SD	Mean	SD
Stage 1	7.56	1.64	6.74	1.39	6.64	1.26
Stage 2	3.34	0.93	3.35	1.29	1.87	0.42
Stage 3	5.64	1.04	5.35	1.1	4.2	0.82
Stage 4	4.07	1.13	4.13	1.15	2.6	0.45
Stage 5	5.01	0.99	4.98	0.76	3.88	0.39
Stage 6	3.09	0.9	2.4	0.78	2.15	0.36
Stage 7	4.46	1.03	4.18	0.66	3.24	0.42
Stage 8	3.45	0.99	2.89	0.91	1.99	0.4

We perform four (4) two-samples t significance tests, with level of significance α=0.05. Test I is on the SpO_2 error for RoI between channel A and the KF, Test II is on the SpO_2 error for RoI between channel B and the KF, Test III is on the SpO_2 error for DST between channel A and the KF, and Test IV is on the SpO_2 error for DST between channel B and the KF. The conclusions of hypothesis testing (using α=0.05) and the p-values are shown in Tables 4.4 and 4.5, respectively.

Table 4.4 Tested Hypotheses and Results

	Is SpO_2 MAE ...		the SpO_2 MAE ... ?		
Test I	computed from channel A using RoI + Comb ($\nu_{2,A}$)	greater than	computed from the KF ($\nu_{2,KF}$)	$H_0 : \nu_{2,A} - \nu_{2,KF} = 0$ $H_a : \nu_{2,A} - \nu_{2,KF} > 0$	Yes.
Test II	computed from channel B using RoI + Comb ($\nu_{2,B}$)	greater than	computed from the KF ($\nu_{2,KF}$)	$H_0 : \nu_{2,B} - \nu_{2,KF} = 0$ $H_a : \nu_{2,B} - \nu_{2,KF} > 0$	Yes. Except for stage 1 (yellow region of Figure 3.10) where we fail to reject the null hypothesis.
Test III	computed from channel A using DST + Comb ($\nu_{4,A}$)	greater than	computed from the KF ($\nu_{4,KF}$)	$H_0 : \nu_{4,A} - \nu_{4,KF} = 0$ $H_a : \nu_{4,A} - \nu_{4,KF} > 0$	Yes. Except for stage 1 (yellow region of Figure 3.10) where we fail to reject the null hypothesis.
Test IV	computed from channel B using DST + Comb ($\nu_{4,B}$)	greater than	computed from the KF ($\nu_{4,KF}$)	$H_0 : \nu_{4,B} - \nu_{4,KF} = 0$ $H_a : \nu_{4,B} - \nu_{4,KF} > 0$	Yes. Except for stages 1 and 6 (Figure 3.10) where we fail to reject the null hypothesis.

Table 4.5 P-values of All Tests of Table 4.4

	p-value (significance level was α = 0.05)			
	Test I	Test II	Test III	Test IV
Stage 1	0.011	0.225	0.06	0.422
Stage 2	< 0.001	0.001	< 0.001	< 0.001
Stage 3	< 0.001	< 0.001	< 0.001	0.004
Stage 4	0.005	0.002	< 0.001	< 0.001
Stage 5	< 0.001	0.002	< 0.001	< 0.001
Stage 6	< 0.001	0.032	0.002	0.148
Stage 7	< 0.001	0.002	< 0.001	< 0.001
Stage 8	< 0.001	0.009	< 0.001	0.002

From Tables 4.4 and 4.5, we have strong evidence to reject the null hypothesis for test I which means that the SpO_2 mean absolute error of channel A (calculated using RoI proceded by a HR-tuned comb filter) is greater than the SpO_2 mean absolute error of the KF (obtained by integrating both channels A and B). We also have strong evidence to reject the null hypothesis for test II which means that the SpO_2 mean absolute error of channel B (calculated using RoI proceded by a HR-tuned comb filter) is greater than the SpO_2 mean absolute error of the KF, except for stage 1 where we fail to reject the null hypothesis. Additionally, for Test III, the SpO_2 MAE computed from channel A (using the DST-based algorithm preceded by a heart rate tuned comb filter) is greater than the SpO_2 MAE computed from the KF. The only exception is stage 1 (yellow region of Figure 3.10). As for Test IV, SpO_2 MAE computed from channel B (using the DST-based algorithm preceded by a heart rate tuned comb filter) is greater than the SpO_2 MAE computed from the KF (obtained by integrating both channels A and B), except for stages 1 and 6 (check Figure 3.10). Therefore, overall, integrating SpO_2 levels computed from the two channels of PPG signals (red and IR) produces a more reliable estimate as compared the SpO_2 levels computed from each channel separately.

4.4 Discussion and Conclusion

We study the potential advantage of integrating SpO_2 levels calculation from PPG signal (red and IR) measured from two separate sensors (as opposed to using the separate sensors without integration). The methods we use for SpO_2-level calculation are the RoI approach and the DST-based algorithm, both preceded by a heart rate tuned comb filter. SpO_2 levels from both channels were integrated using a KF. We show the performance of integrating two SpO_2 sources on experimental real exercise data, where we compute the SpO_2 MAE for fourteen participants studied over eight exercise stages. We conclude that when SpO_2 levels are calculated using RoI approach preceded by a heart rate tuned comb filter, the SpO_2 MAE of channel A is greater (and statistically significant) than the SpO_2 MAE of the KF. Similarly, the SpO_2 MAE of channel B is greater (and statistically significant) than the SpO_2 MAE of the KF, except for one of the stages (stage 1) where both errors are the same. We drew similar conclusion when we calculated SpO_2 levels using DST-based algorithm preceded by a heart rate tuned comb filter. The overall conclusion is that if two channels of PPG (red and IR) signals are available, then combining SpO_2 levels obtained from these channels is likely to exhibits a smoother and more reliable estimate of the blood oxygen saturation level as compared to the SpO_2 levels computed from each channel separately.

CHAPTER 5

DUAL WAVELENGTH PHOTOPLETHYSMOGRAPHY FRAMEWORK FOR HEART RATE CALCULATION

5.1 Introduction

Much effort has been exerted to suppress motion artifacts in order to extract high-quality vital signs, especially heart rate (HR), from noise contaminated PPG signals [2, 33, 34, 35, 21]. In this chapter, we contribute to this effort by proposing a PPG-only motion-resistant framework for HR calculation.

There are two main sources of motion artifacts that could contaminate a PPG signal collected from a human in motion [35]. The first source of noise is the sensor displacement relative to its original point of contact with the skin. This displacement could change the path of light, and hence modify the signal collected by the photodetector [36]. The second source of noise is skin and tissue deformations caused by the sensor's movement.

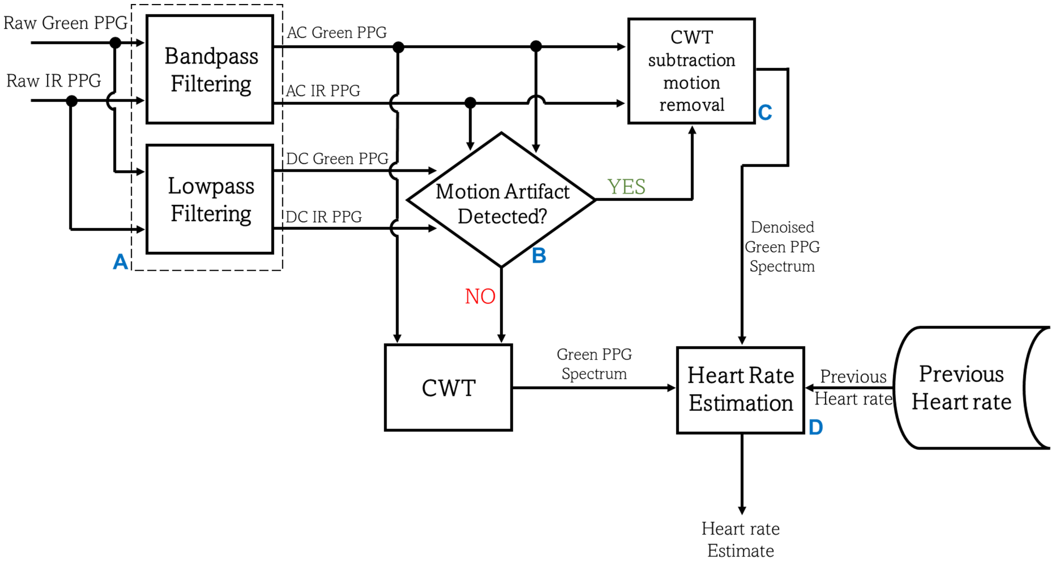


Figure 5.1 Block diagram of a possible implementation of the HR algorithm by Zhang *et al.* (2019).

Zhang [35] proposed an HR calculation method that uses a dual-wavelength sensor that comprises a green and an infrared (IR) PPG signal. The IR PPG signal was used to develop a noise source. A block diagram of an implementation of Zhang's algorithm is shown in Figure 5.1. The inputs are green and IR PPG signals. Both green and IR signals are collected from a wrist-worn PPG sensor unit. The output is a heart rate estimate. Zhang's algorithm provides a clean version of the raw green signal as another output. In [35], the raw green and IR PPG signals are first filtered with a second order IIR bandpass filter (0.4-4 Hz) to suppress noise components that are outside the normal range of human heart rate (block A of Figure 5.1). The presence of motion in the system is obtained using two parameters namely; (1) peak to noise ratio of the green PPG signal in the frequency domain; and (2) the ratio (AC/DC) between the AC and DC components of the green PPG signal and the AC/DC ratio of the IR PPG signal (block B of Figure 5.1 and Section 4.1 in [35]). If noise is detected in the PPG signals, the noise frequencies are removed from the green PPG signal by subtracting the scaled frequency spectrum of the IR signal from the frequency spectrum of the green signal. Frequency spectra are generated using Continuous Wavelet Transform (CWT) (block C in Figure 5.1 and Section 4.2 in [35]). An HR estimate is calculated from the green frequency spectrum (bandpass filtered green spectrum when there is no noise and denoised green spectrum if noise if detected in the signals) using a mechanism that integrates the heart rate values calculated from the current and previous time intervals (block D in Figure 5.1 and Section 4.3 in [35]).

The HR calculation algorithm presented in [35] was tested on "micromotion artifacts" such as *finger tapping* and *fist opening and closing.* In the current study, we examined the applicability of a related approach for more substantial movements and dynamic scenarios. Motivated by the sensor architecture proposed in [35], we expanded the HR calculation technique to high-intensity full body repetitive "macro-motion" exercise data. The resulting Dual Wavelength (DWL) method collects green and IR PPG data from a dual-wavelength wrist unit and processes them to estimate the participant's heart rate. Performance of DWL was documented in an extensive motion experiment involving fourteen (14) human participants[1].

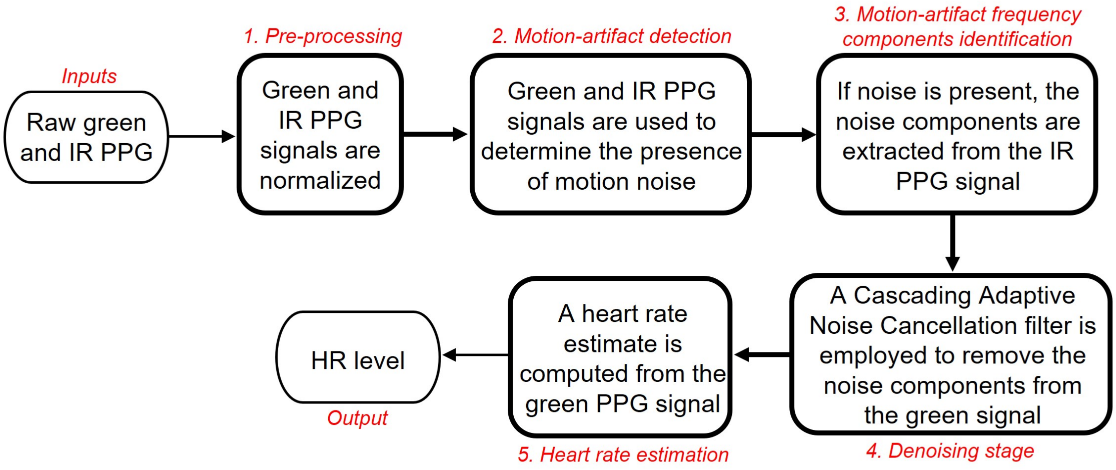


Figure 5.2 Summary of the DWL method.

Figure 5.2 shows the essentials of the DWL method. It consists of five (5) stages; *1. Pre-processing*, *2. Motion-artifact detection*, *3. Motion-artifact frequency components identification*, *4. Denoising*, and *5. Heart rate estimation*. The inputs to the DWL system are green and IR PPG channels measured from wrist-unit constructed for this study (see Section 5.2). The output is an HR level. First, the green and IR PPG signals are normalized by dividing the signal's AC component by its DC component. We then check if significant motion noise is present in the PPG signals (Subsection 5.4.2). If the signals appear noise-free, the normalized green PPG signal is directly used to calculate an HR value. If the signals appear noise contaminated, we then extract the noise components from the IR PPG signal. These noise components are removed from the noisy green PPG signal. We employ a Cascading Adaptive Noise

---

[1] There were three separate experiments. In the first (SNR experiment), we used all fourteen (14) participants. In the second experiment (wrist-based heart rate calculation), we used eleven (11) participants due to sensor failure on three of the participants. In the third experiment (palm-based heart rate calculation), we used twelve (12) participants due to sensor failure on two of the participants.

Cancellation (C-ANC) architecture that uses a QR-decomposition-based least-squares lattice (QRD-LSL) algorithm [37] to denoise the green PPG signal before it is used for HR calculation. A separate decision mechanism validates the HR estimate, and corrects it when noise levels are too high to produce a meaningful estimate.

The rest of this chapter is organized as follows. Section 5.2 describes the experimental settings along with the sensors suite. In Section 5.3, we use experimental data to present the rationale for choosing the IR PPG signal as noise reference signal. In Section 5.4, we introduce a method for (1) denoising the green PPG using the noise components extracted from an IR PPG signal, and (2) computing HR levels. Section 5.5 reviews alternative HR calculation methods that use auxiliary sensors, namely, accelerometers, as a noise source. These methods are TROIKA [33] and JOSS [34]. In Section 5.6, we compare the performance of the DWL method to that of our implementation of TROIKA and JOSS. The comparison we provide is with respect to (1) the HR ground truth computed from an electrocardiography (ECG) signal, and (2) the HR levels obtained using TROIKA and JOSS. In Section 5.7, we validate our framework by testing its performance on experimental data collected from the palms (instead of wrists) of the same participants during a second run (validation run). We conclude that the DWL method provides several desirable features, including; (1) the DWL framework uses only PPG signals; auxiliary signals (such as accelerometers used by TROIKA and JOSS) are not needed; and (2) the DWL framework appears to exhibit high accuracy and lower computational burden in the presence of motion artifacts as compared to TROIKA and JOSS.

5.2 Experimental Protocol and Sensors Suite

We conducted a high-intensity full body exercise experiment where we collected PPG, electrocardiography (ECG), and tri-axial accelerometer data. Accelerometers measured accelerations in three orthogonal directions X, Y, and Z, simultaneously

[38]. Readings were obtained from fourteen (14) human participants while they were standing or running on a split-belt instrumented treadmill (Bertec Corp., Columbus, OH) [39]. First, a multi-wavelength wrist oximeter unit was strapped around the participant's wrist. The wrist unit encloses two blue LEDs (of wavelength $\lambda_B = 460$ nm), two green LEDs (of wavelength $\lambda_G = 520$ nm), and two IR LEDs (of wavelength $\lambda_{IR} = 940$ nm), as well as a photo-detector. Additionally, a tri-axial accelerometer sensor was placed on the participant's arm (right above the PPG wrist-unit) and secured in place using athletic tapes. Lastly, an ECG sensor was mounted onto the participant's chest using adhesive electrodes. Athletic tape was wrapped around each participant's chest to ensure sensor's stability and good skin contact. Table 5.1 shows all the instruments and sensors used in the experiment. Both ECG and accelerometer data were recorded using the Delsys EMGworks Software. Multi-wavelength PPG wrist-unit data were recorded using an Arduino UNO. All signals were sampled at 100 Hz. Raw data were processed using MATLAB 2022b (Mathworks, Natick, MA) [40]. All raw data are available through the GitHub repository in [41].

We observed that the quality of the blue PPG signals collected from the wrist was very poor for some participants, even at rest (when the participant was not performing any physical activities). The poor quality of blue PPG signals was also reported in [42], and two possible reasons were presented. The first reason is the use of a photo-detector with low sensitivity to lights in the blue spectrum (which is probably the case for the photo-detector used in our study). The second reason is that signals collected from measuring sites closer to the periphery (such as fingertips, toes, and ears) exhibit more stable readings as compared to sites that are farther from the periphery (such as the wrist and forehead). Due to the unstable readings from the blue PPG sensor, the blue PPG signals we collected were not used in our study. However, they are reported in the GitHub repository [41].

The ECG signal was used to calculate the HR "ground truth" values. We manually labeled the R-peaks for all ECG signals. The HR ground truth at time step l, $HR_{GT}(l)$, is obtained using the relationship

$$HR_{GT}(l) = \frac{1}{\delta^{R-R}(l)}, \tag{5.1}$$

where $\delta^{R-R}(l)$ is the average time difference between each two consecutive R peaks present within the 8-second-long window, at time step l.

Table 5.1 Instruments Used for Data Collection in the Exercise Experiment

Instrument/Sensor	**Manufacturer**	**Reference**
Split-belt Instrumented Treadmill	Bertec Corp. (Columbus, OH, USA)	Catalog in [39]
IR LED (TSAL6100)	Vishay Intertechnology Inc. (Malvern PA. USA)	Datasheet in [43]
Green LED (A-U5MUGC12)	Light House LEDs LLC (Medical Lake, WA, USA)	Datasheet in [44]
Blue LED (A-U5MUBC12)	Light House LEDs LLC (Medical Lake, WA, USA)	Datasheet in [45]
Photo-detector (OPT101)	Texas-Instrument Inc. (Dallas, TX, USA)	Datasheet in [46]
Delsys Trigno Avanti (tri-axial accelerometer)	Delsys Inc. (Natick, MA, USA)	Catalog in [47]
Trigno EKG Biofeedback sensor (ECG)	Delsys Inc. (Natick, MA, USA)	Catalog in [48]

The experimental protocol we followed during data collection was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of the New Jersey Institute of Technology (protocol code 2108010504; approved on September 14^{th}, 2021). All participants were physically fit, healthy, and athletic volunteers. Each participant was exposed to the following profile of treadmill activities:

- **Stage 1**: The participant stood steady on the treadmill for 1 minute. During this stage, clean physiological signals were collected.
- **Stage 2**: The participant ran at a speed of 6 km/h (about 3.7 mph) for 1 minute.
- **Stage 3**: If the participant was comfortable, the speed was increased gradually to 12 km/h (about 7.5 mph), for 1 minute. At any time, if the participant was not comfortable, the speed was reduced to the participant's comfort zone.
- **Stage 4** (same as stage 2): The participant ran at a speed of 6 km/h (about 3.7 mph) for 1 minute.
- **Stage 5** (same as stage 3): If the participant was comfortable, the speed was increased gradually to 12 km/h (about 7.5 mph), for 1 minute. At any time, if the participant was not comfortable, the speed was reduced to the participant's comfort zone.
- **Stage 6**: The participant stood steady on the treadmill for a duration of 1 minute.

5.3 Infrared PPG Signal as Noise Reference Signal

According to [35], IR PPG signals are more affected by motion artifacts than green PPG signals. To verify this behavior in our experiment, we calculated the Signal-to-Noise (SNR) ratios for both the green and IR PPG signals. The SNR is defined as

$$SNR(in\ dB) = 10\ log_{10}\left(\frac{P_{desired\,signal}}{P_{noise}}\right), \tag{5.2}$$

where $P_{desired\,signal}$ and P_{noise} are the power of the participant's heart rate component and motion artifact components, respectively. In order to calculate an SNR value, the desired and noise signal components should be identified and separated. At this stage, we used the participant's HR ground truth (obtained from an ECG signal, collected simultaneously with the PPG signals) using Equation (5.1), in order to determine the desired signal component.

The desired signal and noise components were obtained, respectively, from the green and IR PPG signals. First, the green and IR signals were normalized by dividing their AC component by their DC component. The desired signal component (the component that contains heart rate information) of the normalized PPG signal was obtained by applying two bandpass filters centered at the participant's heart rate frequency and its first harmonic [35]. During this step the participant's heart rate was obtained from the ECG signal. The noise component was obtained by subtracting the desired signal component from the normalized signal.

We calculated the SNR values of the green and IR PPG signals for all fourteen (14) subjects in the following manner. Every 2 seconds, the preceding 8-second-long PPG segment was used to obtain an SNR value. In total, each subject had between 175 and 177 SNR values for each PPG signal (green and IR signals). The first and last minute of the collected PPG data were omitted since these data segments were noise-free. SNR values for all subjects were grouped together – their distribution is presented in Figure 5.3 as boxplots. A boxplot depicts the distribution of a variable [49] (in this case we show the distribution of SNR values computed for IR and green PPG signals for all fourteen (14) experimental subjects). The lower and upper limit of the black box in Figure 5.3 represent the 25^{th} percentile (denoted Q_1) and 75^{th} percentiles (denoted Q_3) of the data, respectively. The difference between Q_3 and Q_1 defines the interquartile range (IQR). The red bar is the 50^{th} percentile (denoted Q_2) or the median value of the distribution. The black bar at the end of the lower and upper whiskers determine the minimum and maximum limits, respectively. Typically, the minimum and maximum values are $Q_1 - 1.5 \times IQR$ and $Q_3 + 1.5 \times IQR$, respectively. If an SNR value falls outside these limits, it is considered an outlier and shown as a red plus '+' sign on the boxplot. The green dot (green PPG boxplot) and black dot (IR PPG boxplot) are the mean values of the corresponding distributions.

In our experiment, the SNR mean value of the IR PPG, $\upsilon^{IR}_{SNR,\,wrist} = -8.5\ dB$ (black dot in Figure 5.3), was less than the SNR mean value of the green PPG signal, $\upsilon^{G}_{SNR,\,wrist} = -4.8\ dB$ (green dot in Figure 5.3). These results are statistically significant for a level of significance $\alpha = 0.01$. This difference supports the choice of IR PPG as a noise reference signal using experimental data.

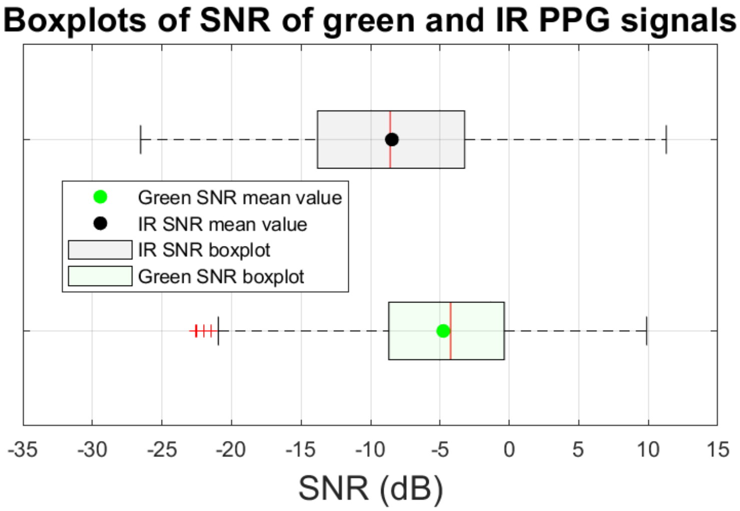


Figure 5.3 SNR values of IR and green PPG signals, respectively, calculated from all fourteen (14) subjects. The dots represent the mean value of SNR. The red bars represent the median value of SNR. The red '+' signs represent outliers.

5.4 DWL Framework

The proposed DWL framework consists of the following stages (Figure 5.2), *A. Preprocessing*, *B. Motion-artifact Detection*, *C. Motion-artifact Frequency Components Identification*, *D. Denoising*, and *E. Heart Rate Estimation*. The inputs to the system are raw green and IR PPG signals measured using the dual-wavelength PPG wrist-unit sensor (described in Section 5.2). The output is an HR estimate, $\hat{HR}(l)$ at time step l (the initial time step is $l = 1$). We refer to the average of the latest Z estimates of the heart rate as $\hat{HR}^{(Z)}(l)$, namely

$$\hat{HR}^{(Z)}(l) = \frac{1}{Q}\sum_{q=0}^{Q-1} \hat{HR}(l-q) \mid Q = min\{Z, l\}. \tag{5.3}$$

Figure 5.4 is a block diagram of the DWL method. The system produced a new estimate of HR at every time step ($\hat{HR}(l)$ at time step l). The time between two subsequent windows in our study was 2 seconds. In addition, the system produces three *search ranges*. They are; the "narrow search range," $\Delta_n(l+1)$; the "medium search range," $\Delta_m(l+1)$; and the "wide search range," $\Delta_w(l+1)$. Ranges $\Delta_m(l+1)$ and $\Delta_n(l+1)$, which are used in the motion-artifact frequency components identification process of Subsection 5.4.3, are centered at $\hat{HR}(l)$. The range $\Delta_w(l+1)$, which is used in the heart rate estimation process of Subsection 5.4.5, is centered at $\hat{HR}^{(6)}(l)$, the average of the 6 previous heart rate estimates. The ranges satisfy $\Delta_n(l+1) < \Delta_m(l+1) < \Delta_w(l+1)$. Moreover, $\Delta_n(l+1) = \frac{\Delta_m(l+1)}{2}$. The three ranges are calculated as following:

Wide search ranges, $\Delta_w(l+1)$: $\Delta_w(l+1)$ is adopted from [50] and employed in the heart rate estimation process (Subsection 5.4.5) as the range in which we search for the participant's heart rate. At time step $l+1$, the search for $\hat{HR}(l+1)$ is confined to

$$[\hat{HR}^{(6)}(l) - \frac{\Delta_w(l+1)}{2},\ \hat{HR}^{(6)}(l) + \frac{\Delta_w(l+1)}{2}]\ Hz\ , \tag{5.4}$$

$\Delta_w(l)$ is defined as

$$\Delta_w(l) = \begin{cases} c_1 + 2 \times max\Big(\hat{HR}(j) - \hat{HR}(j-1)\ \Big|\ l - c_3 \le j \le l-1\Big) & ,\text{if}\, l > c_3 \\ c_0 & ,\text{if}\, l \le c_3\ , \end{cases} \tag{5.5}$$

where $max(.)$ is the maximum value. $\Delta_w(l)$ is updated based on the maximum value of the differences between two consecutive heart rate values calculated over c_3 previous time steps. For the first c_3 iterations, $\Delta_m(l)$ is equal to c_0. In this work, c_0 was set to 0.33 Hz (or 20 BPM), c_1 was set to 0.37 Hz (or 22 BPM), and $c_3 = 15$

(which is equivalent to 30 seconds).

Medium and narrow search range, $\Delta_m(l+1)$ **and** $\Delta_n(l+1)$: $\Delta_m(l+1)$ and $\Delta_n(l+1)$ are employed in the motion-artifact frequency components identification process (Subsection 5.4.3). At time step $l+1$, the *medium search range* is the range confined to

$$[\hat{HR}(l) - \frac{\Delta_m(l+1)}{2},\ \hat{HR}(l) + \frac{\Delta_m(l+1)}{2}]\ Hz\ , \tag{5.6}$$

where $\hat{HR}(l)$ is the HR estimated at the time step l. $\Delta_m(l)$ is defined as

$$\Delta_m(l) = \begin{cases} c_1 + 2 \times SD\Big(\hat{HR}(j) - \hat{HR}(j-1)\ \Big|\ l - c_3 \leq j \leq l-1\Big) & ,\text{if}\, l > c_3 \\ c_0 & ,\text{if}\, l \leq c_3 \end{cases}, \tag{5.7}$$

where $SD(.)$ is the standard deviation. $\Delta_m(l)$ is updated based on the standard deviation of the differences between two consecutive heart rate values calculated over c_3 previous time steps. For the first c_3 iterations, $\Delta_w(l)$ is equal to constant c_0. In this work, the parameters c_0, c_1, and c_3 are same defined for the *wide search range*, namely; $c_0 = 0.33$ Hz (or 20 BPM), $c_1 = 0.37$ Hz (or 22 BPM), and $c_3 = 15$ (which is equivalent to 30 seconds).

The *narrow search range* is defined as $\Delta_n(l+1) = \dfrac{\Delta_m(l+1)}{2}$.

In addition to the three search ranges, we calculate a short-term 3-point-average heart rate, $\hat{HR}^{(3)}(l)$, that we provide to the users and employ in Section 5.6 for assessing the performance of DWL.

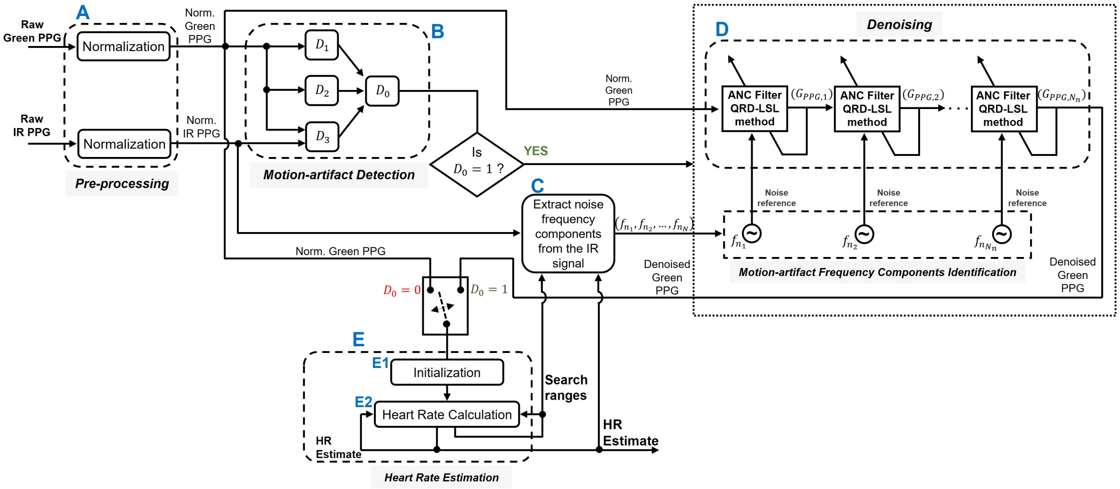
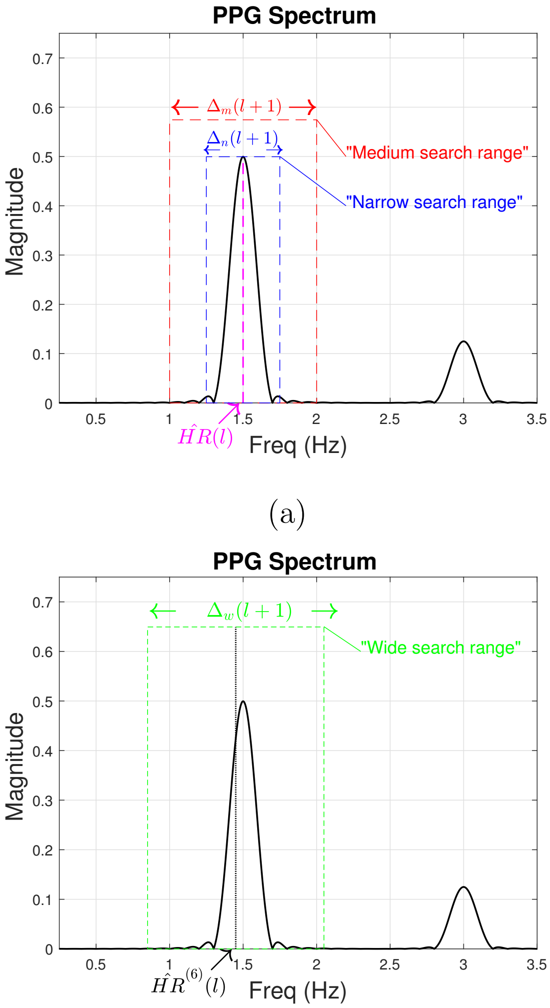


Figure 5.4 Block diagram of DWL method. The inputs for calculating heart rate are raw green and IR PPG signals. The output is an estimate of the HR.

Figure 5.5 is an illustration of a typical IR PPG spectrum. The magenta dashed line in Figure 5.5 (a) is the heart rate estimated at time step l, $\hat{HR}(l)$. The black dotted line in Figure 5.5 (b) is the average of the 6 previous heart rate estimates at time step l, $\hat{HR}^{(6)}(l)$. In this example, $\hat{HR}(l)$ is 1.5 Hz and $\hat{HR}^{(6)}(l)$ is 1.45 Hz. Additionally, we present in Figure 5.5 the "wide search range," $\Delta_w(l+1)$, as a green dashed rectangle, the "medium search range," $\Delta_m(l+1)$, as a red dashed rectangle, and the "narrow search range," $\Delta_n(l+1)$, as a blue dashed rectangle.

5.4.1 Pre-processing

First, both green and IR PPG signals are normalized (block A of Figure 5.4). Normalization is done by dividing the signal's AC component by its DC component [51]. The AC component is obtained by passing the raw PPG signal through a Chebyshev Type II bandpass filter of order 5 and bandpass frequency range of 0.5 to 10 Hz. The DC component is obtained by passing the raw signal through a Chebyshev Type II lowpass filter of order 5 and passband frequency of 0.5 Hz.



(a)

(b)

Figure 5.5 Illustration of the frequency spectrum of typical IR PPG signal. In this example, $\hat{HR}(l) = 1.5$ Hz and $\hat{HR}^{(6)}(l) = 1.45$ Hz. Three search ranges are declared at time step l. They are the narrow search range, $\Delta_n(l+1)$ (blue dashed rectangle); the *medium search range*, $\Delta_m(l+1)$ (red dashed rectangle); and the *wide search range*, $\Delta_w(l+1)$ (green dashed rectangle). The narrow and medium search ranges are centered at $\hat{HR}(l)$ (magenta dashed line) and used for noise frequency component search. The wide search range is centered at $\hat{HR}^{(6)}(l)$ (black dotted line) and is used to search, at time step $l+1$, for $\hat{HR}(l+1)$.

5.4.2 Motion-artifact detection

Motion artifact detection is used to find out whether the PPG signals are contaminated by motion noise (if they are not, we can bypass noise suppression operations that are not needed). The PPG signals go through the following three

(3) local detectors to determine if appreciable levels of noise motion are present (block B of Figure 5.4):

Local Detector 1 (D_1) – number of peaks: The number of dominant peaks (whose magnitude exceeds 30% of the maximum peak for this example) in the frequency spectrum of the green PPG signal, denoted N_p, is calculated. If N_p exceeds two (2), D_1 indicates that the signal is contaminated with motion noise. If N_p is 1 or 2, then we conclude that no appreciable motion noise is present, since the frequency of the heart rate and sometimes its first harmonic component are typically observed in the spectrum of a clean PPG signal.

Local Detector 2 (D_2) – power of green signal: The power of the green PPG signal calculated at the beginning of the experiment (when the participant is at rest) is considered the reference power, denoted P^{ref}. At each time step l, the power of the green PPG, $P^G(l)$, is calculated and compared to the reference power P^{ref}. If $P^G(l)$ is more than $(1 + \kappa)P^{ref}$, D_2 indicates that the green PPG signal is contaminated with motion noise[2]. In this study we used $\kappa = 0.2$.

Local Detector 3 (D_3) – Pearson correlation between green and IR PPG signals: The correlation between the green and IR PPG signals is also used to assess noise contamination in the green signal. If the correlation between the green and IR PPG signals, $\rho_{green,\,IR}$, is below a certain threshold (we used 0.8), then D_3 will decide that the green PPG signal is contaminated with motion noise.

---

[2] The amplitude of the PPG signal might change over time [52]. Therefore, the reference power P^{ref} is updated whenever no motion is detected in the system for five (5) consecutive time steps (global detector D_0 return '1'). In this case, the updated value of P^{ref} is set to the power of the green PPG signal calculated at the current time step, l.

Global Detector – noise detector: The decisions of the three local detectors are fed into a global detector that will decide whether the signal is noise contaminated or not. The global detector is shown as:

$$D_0 = D_1 \vee D_2 \vee D_3 = \begin{cases} 1 \ (noise\ is\ present) & \text{, if } D_1 \vee D_2 \vee D_3 = 1 \\ 0 \ (no\ noise) & \text{, if } D_1 \vee D_2 \vee D_3 = 0 \end{cases} \tag{5.8}$$

where "$\vee$" represents the OR logic operator.

5.4.3 Motion-artifact frequency components identification

If motion artifacts are detected in the normalized green PPG signal, we use the normalized IR signal to build the motion noise component set $\mathbb{N}_{noise}$ (block C of Figure 5.4). $\mathbb{N}_{noise}$ can be written as $\mathbb{N}_{noise} = \{f_{n_i} | 1 \leq i \leq N_n\}$ where f_{n_i} is the i^{th} discrete noise frequency components and N_n is the number of elements in the set $\mathbb{N}_{noise}$. The set $\mathbb{N}_{noise}$, which contains all the noise frequency components that we aim to remove from the normalized green PPG signal, is obtained using the following five (5) steps in sequence. The first three steps capture noise with relatively high intensity, usually harmonically-related frequency pairs that contaminate the PPG signals. The last two steps compare the IR and green signal spectra to discover additional noise components of reduced-intensity presence in the IR spectrum.

Step 1 – Identification of dominant frequency components. First, we capture the dominant frequency components in the spectrum of the normalized IR PPG signal. Those are the frequencies (between 0.5 to 4 Hz) whose magnitude exceeds 50% of the highest peak in the IR PPG spectrum. Figure 5.6, which is an image that was created for illustration purposes, depicts how we capture dominant peaks from a typical IR signal. In this scenario, the highest peak (which actually corresponds to the participant's heart rate) is F_1. Two other dominant peaks are shown as red circles (F_2 and F_3). Typically, the peaks captured in step 1 include the frequency of the participant's heart rate, as well as the frequencies of dominant noise components. We add all of them (F_1, F_2, and F_3 in our example) to $\mathbb{N}_{noise}$ with the understanding that one of them may correspond to the participant's HR and may therefore need to be removed from $\mathbb{N}_{noise}$ later.

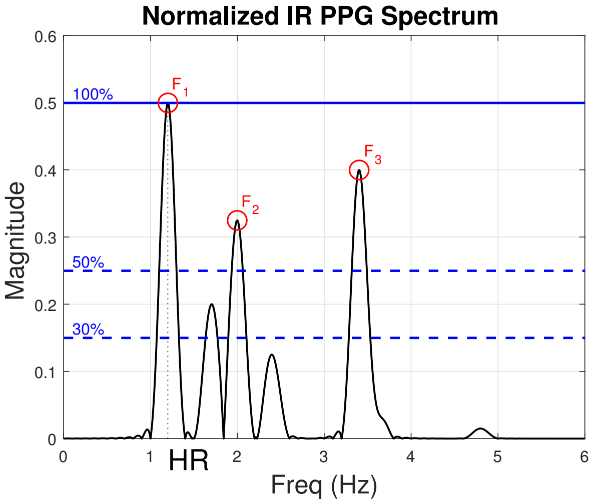


Figure 5.6 Illustration of the frequency spectrum of a typical IR PPG signal. The red circles correspond to the dominant peaks, denoted F_1, F_2, and F_3 (extracted in step 1 of Subsection 5.4.3). The highest peak, F_1, corresponds to the participant's HR.

Step 2 – Identification of harmonic frequency components. Noise components created by repetitive motion (*e.g.*, when the participant is walking or running) typically occur in harmonically-related pairs [53]. It is possible, however, that the PPG signal contains pairs of harmonically-related noise components whose magnitude is smaller than the 50% threshold used in step 1 to identify dominant frequencies. Step 2 is used to capture pairs of fundamental frequencies and their first harmonics present in the spectrum of the normalized IR PPG signal. Here, we look at all peaks whose magnitudes are above 30% of the highest peak in the IR PPG spectrum. For each such peak, we search for a harmonic at double its frequency. If a pair of harmonically-related frequencies is thus discovered, its component(s) that were not flagged in step 1 are added to the noise frequency set $\mathbb{N}_{noise}$. Again, $\mathbb{N}_{noise}$ may still contain at this stage a component that corresponds to the participant's true HR. Figure 5.7 uses the same spectrum shown in Figure 5.6 to illustrate how a pair of harmonically-related components (F_A, $F_B = F_3$) was discovered. Of this pair, F_B was known to us already from step 1 (it is the same as F_3 in Figure 5.6), and F_A, discovered by step 2, is added to $\mathbb{N}_{noise}$. So now, $\mathbb{N}_{noise} = \{F_1, F_2, F_3, F_A\}$.

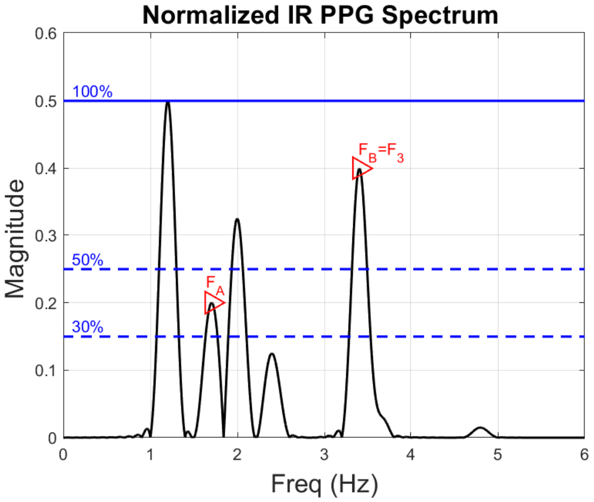


Figure 5.7 Illustration of the frequency spectrum of a typical IR PPG signal. The red triangles correspond to the pair of frequencies, F_A and F_B, that has a harmonic relationship. Frequency F_B is the same as frequency F_3 from Figure 5.6.

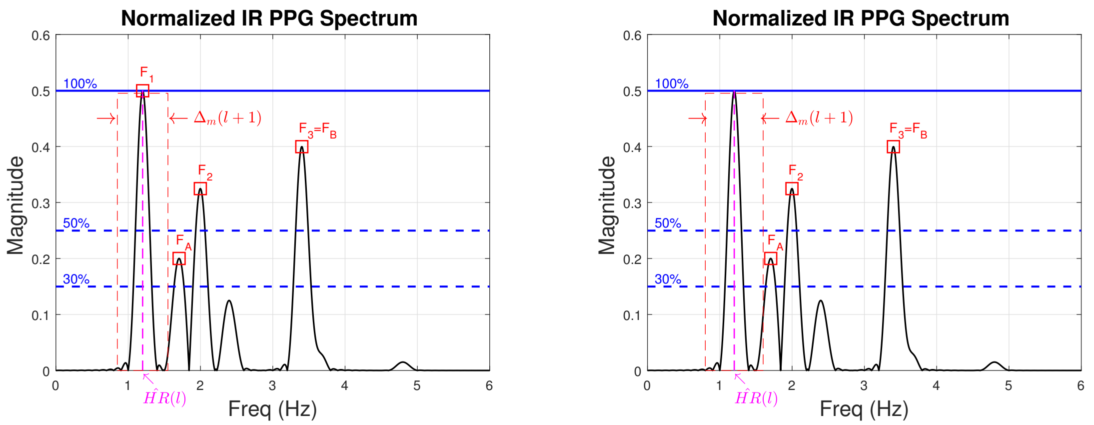
Step 3 – Removal of the heart rate from noise set. As mentioned in our setting in Section 5.4, our system creates a new estimate of the heart rate, $\hat{HR}(l)$ at every time step l. A new time step starts every 2 seconds when l is incremented by 1. Moreover, in step $l + 1$ we calculate $\Delta_w(l + 1)$ (the "wide search range") which is where we search for $\hat{HR}(l + 1)$.

Next, frequency components in $\mathbb{N}_{noise}$ which we captured during steps 1 and 2, and are close to the heart rate estimated at time step l ($\hat{HR}(l)$) are removed from $\mathbb{N}_{noise}$, as we suspect they do not represent noise but rather represent the participant's heart rate. To be precise, at time step $l + 1$, we remove from $\mathbb{N}_{noise}$ all the noise components in the "medium search range" $\Delta_m(l + 1)$.

Figure 5.8 continues the examples of Figure 5.6 and Figure 5.7 to illustrate step 3. In Figure 5.8 (a) and Figure 5.8 (b), we show the estimate of the participant's heart rate at time step l, denoted $\hat{HR}(l)$. We also show $\Delta_m(l+1)$, the "medium search range", $[\hat{HR}(l) - \Delta_m(l+1)/2, \hat{HR}(l) + \Delta_m(l+1)/2]$, from which we remove dominant frequencies deposited earlier into $\mathbb{N}_{noise}$. The red squares in Figure 5.8 (a) represent the frequency components that we obtained from steps 1 and 2 all of which are currently in $\mathbb{N}_{noise} = \{F_1, F_2, F_3, F_A\}$. We now discard the frequency around 1.2 Hz (labeled F_1) since it falls in $\Delta_m(l+1)$, the "medium search range" (region represented by a red dashed rectangle in 5.8). Figure 5.8 (b) show (in red squares) the noise frequency components that are left in the noise set $\mathbb{N}_{noise} = \{F_2, F_3, F_A\}$. $\mathbb{N}_{noise}$ no longer contains the participant's heart rate.

The next two steps seek additional noise components, often attributed to repetitive movements by the participant, through comparison of the IR and green spectra.

Step 4: Step 4 focuses on instances where the noise set $\mathbb{N}_{noise}$, after step 3, has only one noise component, f_{n_1}. In this case, we look at the green spectrum. If we find a component at half f_{n_1} ($f_{n_1}/2$) or twice f_{n_1} ($2 \times f_{n_1}$) in the green spectrum, we add this component to $\mathbb{N}_{noise}$. The only exception is if the component we seek to add falls into the *narrow search range*, $\Delta_n(l+1)$, around $\hat{HR}(l)$, $[\hat{HR}(l) - \Delta_n(l+1)/2, \hat{HR}(l) + \Delta_n(l+1)/2]$; in this case, we refrain from adding it to set $\mathbb{N}_{noise}$.

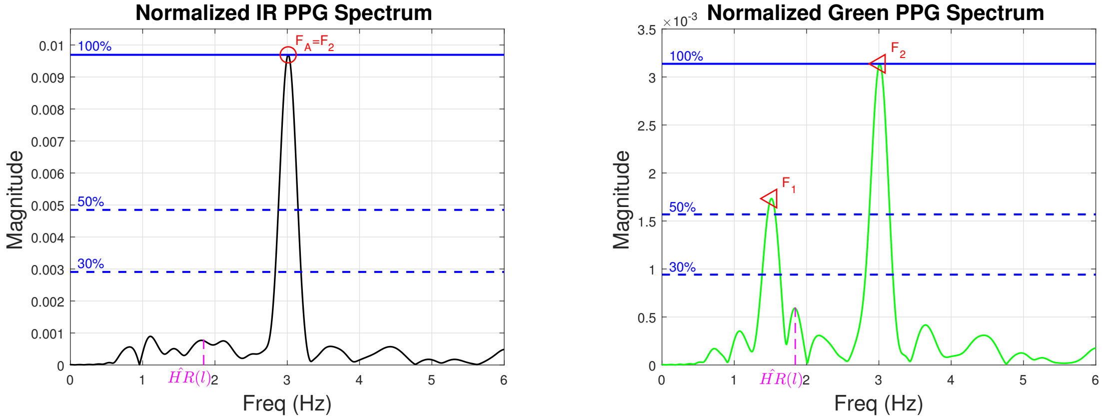


(a) (b)

Figure 5.8 Frequency spectrum of a typical IR PPG signal. $\hat{HR}(l)$ is the heart-rate estimate at time step l. $\Delta_m(l+1)$ is the "medium search range" represented by a red dashed rectangle. The frequency components we obtained from step 1 and 2, namely, F_1, F_A, F_2, and $F_3 = F_B$, are represented by red squares. In Figure 5.8 (a) frequency F_1 falls within $\Delta_m(l+1)$. In Figure 5.8 (b) we discard the frequency F_1 since it falls within $\Delta_m(l+1)$ and leave the rest in $\mathbb{N}_{noise}$ (F_A, F_2, and $F_3 = F_B$).

Step 5: This step addresses spectra that are dominated by vigorous limb swinging by the participant, which may cause displacement of the sensor. In this scenario, the green PPG signal is typically dominated by two high intensity harmonically-related noise frequencies which may dwarf the component at the heart rate frequency. If these frequency components are not already placed in $\mathbb{N}_{noise}$ after steps 1-3, they are added to $\mathbb{N}_{noise}$ at this step. This step is automatically triggered when all the following conditions are met, namely; (a) the IR spectrum contains only one significant frequency component that dominates the spectrum; (b) the green spectrum contains only one pair of significant harmonically-related frequencies; and (c) the dominant frequency component present in the IR spectrum matches with one of the harmonically-related frequencies discovered in the green spectrum.

Figure 5.9 is a real-life example that illustrates this scenario (signals were collected from participant 10 in our experiment, around time 136 seconds). We show the spectrum of participant 10's IR signal in Figure 5.9 (a) and green signal in Figure 5.9 (b). We show in magenta the heart rate estimate at time step l, $\hat{HR}(l)$. The



(a) (b)

Figure 5.9 IR and green spectra from participant 10 around 136 seconds. We show in magenta the heart rate estimate at time step l, $\hat{HR}(l)$. (a) Frequency spectrum of participant 10's IR PPG signal. The red circle labeled $F_A = F_2$ represent the dominant noise frequency. (b) Frequency spectrum of participant 10's green PPG signal around time 136 seconds. The two red triangles labeled F_1 and F_2 represent high intensity harmonically-related frequencies. Note that the frequency F_A from Figure 5.9 (a) is the same of the frequency F_2 from subplot 5.9 (b). Both F_1 and $F_A = F_2$ are put into $\mathbb{N}_{noise}$.

green signal captures the high intensity harmonically-related frequency pairs F_1 and F_2 of Figure 5.9 (b). The IR spectrum (Figure 5.9 (a)) is dominated by the frequency F_A that is equal to frequency F_2 from the green spectrum, but does not capture a noise component at F_1. Here, frequencies F_1 and $F_A = F_2$ are put into $\mathbb{N}_{noise}$.

At the end of this stage, the set $\mathbb{N}_{noise}$ will contain N_n elements that correspond to the noise frequencies that we wish to remove from the normalized green PPG signal.

5.4.4 Denoising stage

Adaptive Noise Cancellation (ANC) filters are often employed to eliminate in-band motion artifacts [54, 55]. In-band noise in our case occurs when the spectra of motion artifacts overlap significantly with that of the PPG signal [56]. An ANC filter for our environment would use as inputs (1) a noise contaminated signal, and (2) a noise reference signal. The ANC filter seeks to eliminate the noise components (measured by the reference signal) from the input noise contaminated signal and provide a noise-free version of the input signal.

Motivated by the architecture in [50], we employ a Cascading Adaptive Noise Cancellation (C-ANC) architecture to remove all the elements of the set $\mathbb{N}_{noise} = \{f_{n_i} | 1 \leq i \leq N_n\}$ (developed in Subsection 5.4.3) from the green PPG signal, *one element at the time*. The block diagram of the proposed C-ANC is shown in Figure 5.10. We show the frequency spectrum of the input signal in Figure 5.10 (spectrum A). This is the green signal collected from participant 3 around time 66 seconds. The spectrum contains three noise frequency components that we wish to eliminate from the signal. The signal collected at the output of the C-ANC (spectrum D in Figure 5.10) does not contain any of the noise components; only the heart rate frequency component remained in the spectrum.

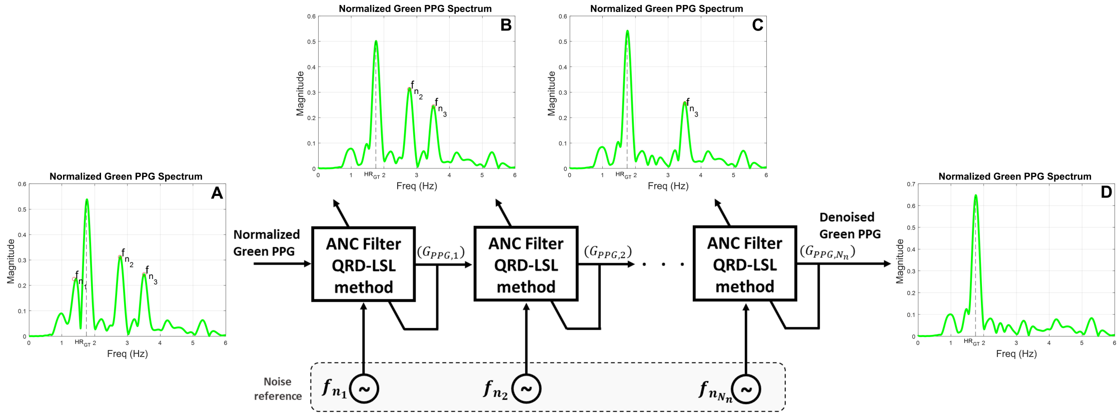


Figure 5.10 Cascading Adaptive Noise Canceler (C-ANC) block diagram.

A total of N_n C-ANC were used to remove the noise components of $\mathbb{N}_{noise}$ from the green PPG signal. At the i^{th} stage ($1 \leq i \leq N_n$), the noise reference signal is a pure sinusoid of frequency f_{n_i}. For instance, the first ANC filter block shown in Figure 5.10 removes the first noise frequency component f_{n_1} from the normalized green PPG signal (see spectrum B of Figure 5.10). The output of the first block is denoted $G_{PPG,\,1}$. $G_{PPG,\,1}$ is fed to the next block where the second noise frequency component f_{n_2} is removed (see spectrum C of Figure 5.10). The process is repeated until all noise components are removed from the normalized green PPG signal. The final output, $G_{PPG,\,N_n}$, is a noise-free version of the green PPG signal. In the proposed method, the QR-decomposition-based least-squares lattice (QRD-LSL) adaptive filter algorithm was used to remove noise components from the green PPG signal [57]. The method incorporates the desirable features of recursive least-square estimation (fast convergence rate), QR-decomposition (numerical stability), and lattice structure (computational efficiency) [37]. The implementation of the QRD-LSL filter in our study used the built-in MATLAB function "*AdaptiveLatticeFilter*" [58] with 10 filter taps and forgetting factor of 0.99.

5.4.5 Heart rate estimation

In this stage (see block E of Figure 5.4), the green PPG signal is used to compute an HR value. If no noise was detected in the green PPG ($D_0 = 0$), then the normalized green PPG is used for heart rate calculation. When noise was detected in the green PPG signal, an HR value is obtained from the denoised green signal (obtained at the output of block D in Figure 5.4, also shown in Figure 5.10). The "Heart Rate Estimation" stage comprises two steps, namely, "Initialization" and "Heart Rate Calculation."

Initialization (block E1 of Figure 5.4). This is a process of capturing a baseline heart rate at rest. In our experiment, it was a one-minute phase during which participants were asked to remain steady in order to capture noise-free green and IR PPG signals. To calculate the initial HR estimate, $\hat{HR}(1)$ at time step $l = 1$, we used the frequency spectrum of the normalized green PPG signal. $\hat{HR}(1)$ corresponds to the highest peak within the initial search range 0.5 to 3 Hz (which corresponds to 30 to 180 BPM).

Heart rate calculation (block E2 of Figure 5.4).

At time step $l+1$, the heart rate calculation method we propose employs the following variables in order to generate an HR estimate, $\hat{HR}(l+1)$:

1. The heart rate estimated from the previous time step l, $\hat{HR}(l)$.
2. A heart rate candidate $HR_{cand}(l+1)$ which is obtained from the spectrum of the green PPG signal.
3. A heart rate prediction, $HR_{pred}(l+1)$ which is obtained from the long-term (LT) trend of the past six (6) HR estimates. The LT trend is obtained using the Seasonal-Trend decomposition using LOESS[3] (STL) method [59]. In this study, we used the MATLAB implementation, *trenddecomp*.

First, we seek to find a heart rate candidate, $HR_{cand}(l+1)$, within the wide search range $\Delta_w(l+1)$, which corresponds to the highest peak in the green spectrum ($HR_{cand}(l+1) \in [\hat{HR}^{(6)}(l) \pm \Delta_w(l+1)/2]$). If $HR_{cand}(l+1)$ is available, we calculate $\delta_e(l+1)$, which is the absolute difference between $HR_{cand}(l+1)$ and $\hat{HR}(l)$ (in Hz) at time step $l+1$. We distinguish between four (4) cases.

Case 1: If a peak was found in $[\hat{HR}^{(6)}(l) \pm \Delta_w(l+1)/2]$ and $D_0 = 0$ ("no noise")

or

If a peak was found in $[\hat{HR}^{(6)}(l) \pm \Delta_w(l+1)/2]$ and $D_0 = 1$ ("noise is present") and $\delta_e(l+1) < 0.1$ Hz

In this case, $HR_{cand}(l+1)$, corresponds to the highest peak in the green spectrum, within the wide search range $\Delta_w(l+1)$ ($HR_{cand}(l+1) \in [\hat{HR}^{(6)}(l) \pm \Delta_w(l+1)/2]$). The estimated heart rate, $\hat{HR}(l+1)$ is calculated as

$$\hat{HR}(l+1) = HR_{cand}(l+1). \tag{5.9}$$

---

[3] Locally estimated scatterplot smoothing (LOESS).

Case 2: If a peak was found in $[\hat{HR}^{(6)}(l) \pm \Delta_w(l+1)/2]$ **and** $D_0 = 1$ **and** $\delta_e(l+1) > 0.1$ **Hz.**

In this case, we follow the procedure recommended in [33] to consider at most three dominant peaks in the green spectrum, whose magnitude exceed 50% of the maximum peak. Here, $HR_{cand}(l+1)$ is obtained by averaging all the peaks that we considered. The estimated heart rate, $\hat{HR}(l+1)$ is calculated as

$$\hat{HR}(l+1) = \eta \times HR_{cand}(l+1) + (1-\eta) \times HR_{pred}(l+1), \tag{5.10}$$

where η is a constant we set to 0.9.

Case 3: If no peak was found in $[\hat{HR}^{(6)}(l) \pm \Delta_w(l+1)/2]$**.**

In this case, we extend the wide search range, $\Delta_w(l+1)$. The extended wide search range is $\Delta_w^+(l+1) = (1+\tau) \times \Delta_w(l+1)$ ($\tau = 0.25$ in this study). We seek to find at most three dominant peaks within the extended wide search range, $\Delta_w^+(l+1)$ (the range $[\hat{HR}^{(6)}(l) \pm \Delta_w^+(l+1)/2]$). If we find at least one peak, we consider at most three dominant peaks, whose magnitude exceed 50% of the maximum peak. $HR_{cand}(l+1)$ is obtained by averaging all the peaks that we considered. The estimated heart rate, $\hat{HR}(l+1)$ is calculated as

$$\hat{HR}(l+1) = \eta \times HR_{cand}(l+1) + (1-\eta) \times HR_{pred}(l+1), \tag{5.11}$$

where η is a constant we set to 0.9.

Case 4: If no peak was found in $[\hat{HR}^{(6)}(l) \pm \Delta_w(l+1)/2]$ **or in** $[\hat{HR}^{(6)}(l) \pm \Delta_w^+(l+1)/2]$**.**

In this case, $\hat{HR}(l+1)$ is calculated as

$$\hat{HR}(l+1) = HR_{pred}(l+1). \tag{5.12}$$

The heart rate calculation process we used requires the availability of the previous six heart rate estimates in order to generate a heart rate prediction, $HR_{pred}(l+1)$ at time step $l+1$. Therefore, from time steps $l = 2$ to $l = 6$, the heart rate estimates $\hat{HR}(2)$ through $\hat{HR}(6)$ corresponds to the highest peak in the green spectrum, within the wide search range $\Delta_w(l+1)$ ($\hat{HR}(l+1) \in [\hat{HR}^{(6)}(l) \pm \Delta_w(l+1)/2]$)[4]. In this case, $\hat{HR}^{(6)}(l)$ is the average of all the previously calculated heart rate estimates (see Equation (5.3)).

5.5 Alternative HR Calculation Methods

In most studies involving PPG signals collected from humans in motion, suitable reference signals, representing motion artifacts, were obtained through additional hardware [56]. For example, when the PPG sensor is mounted on the wrist of a running participant, accelerometer sensors mounted on the participant's wrist are often used as noise reference signals [60, 61, 62].

TROIKA is an *HR* calculation framework proposed by Zhang *et al.* [33]. TROIKA is based on Singular Spectrum Analysis (SSA) [63] followed by Sparse Signal Reconstruction (SSR) [64] to eliminate the noise dominant components present in PPG signals. The inputs to TROIKA are a green PPG signal and X, Y, and Z accelerometer data. The output is a heart rate estimate. In our implementation of TROIKA, the noise components were obtained from a tri-axial accelerometer. In [33], TROIKA was tested on data collected from a wrist-worn sensor (that encloses a green PPG channel and X, Y, and Z accelerometer data) from twelve (12) participants, during fast running at peak speed of 15 km/h. The heart rate average absolute error of TROIKA in this test was 2.34 beat per minutes (BPM).

A related method is based on Zhang's Joint Sparse Spectrum Reconstruction (JOSS). It was shown in [34] to exhibit a heart rate average absolute error as small as

---

[4] If no such peak is detected, we increment $\Delta_w(l+1)$ by 0.02 Hz (or 1.2 BPM) and we search again for a peak. This process repeats until a peak is found.

1.28 BPM when tested on the same twelve (12) participants used in Zhang's TROIKA study [33]. In JOSS, the input signals are a green PPG signal and X, Y, and Z accelerometer data. The accelerometer data are considered the noise signals. The output is a heart rate estimate. Compared to TROIKA where PPG and accelerometer signals were sampled at 125 Hz, JOSS's low sampling rate, namely 25 Hz, is an attractive feature that gives JOSS the potential to be implemented in Very Large-Scale Integration (VLSI) or Field Programmable Gate Array (FPGA) in wearable devices [34].

Figure 5.11 is a block diagram of a possible implementation of TROIKA (Figure 5.11-a) and JOSS (in Figure 5.11-b).

The heart rate calculation mechanism of the DWL method was inspired by that of TROIKA and JOSS. We compare the quality of HR calculated by the DWL method which does not require accelerometers, to our implementation of the accelerometer-dependent TROIKA and JOSS. The TROIKA and JOSS experimental results were obtained from the same participants that we employed in the analysis of the DWL method.

5.6 Results

5.6.1 Performance metrics

To assess, evaluate, and compare the heart rate estimation performance of DWL method to TROIKA and JOSS, we used four metrics, namely; Mean Absolute Error (MAE) (Equation (5.13)); Mean Absolute Error Percentage ($MAEP$) (Equation (5.14)); a specific performance index (PI) (Equation (5.15)) which is the frequency, in percent, of obtaining an HR estimate that is within ± 5 BPM of the HR ground truth; and computation time (CT). We defined CT to be the total time duration (in seconds) that an algorithm takes to generate heart rate levels from the entire 360-second-long off-line data that has already been collected during the experimental

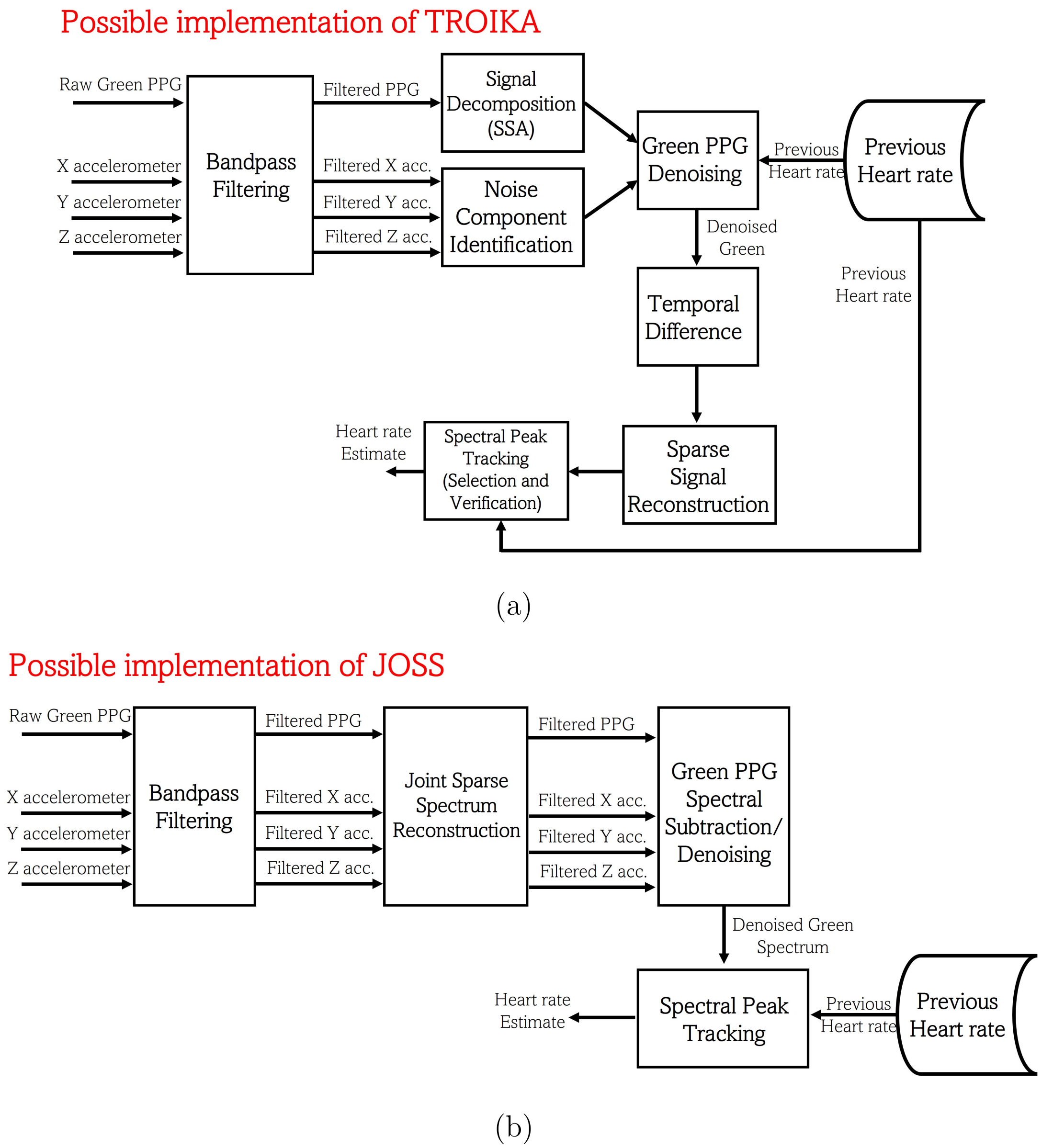


Figure 5.11 Block diagram of a possible implementation of (a) TROIKA algorithm and (b) JOSS algorithm.

run. We compare the HR values calculated by the three tested methods to ground truth values obtained from an ECG signal that is simultaneously recorded, hence synced, with the green and IR PPG waveforms and the X, Y, and Z accelerometer data. All R peaks in the ECG signal were manually labeled. The ground truth HR was obtained using Equation (5.1). The relevant definitions are:

$$MAE\ (in\ BPM) = \frac{1}{L}\sum_{l=1}^{L}\Delta(l) \qquad \text{(ideally 0 BPM)} \tag{5.13}$$

$$MAEP\ (in\ \%) = \frac{1}{L}\sum_{l=1}^{L}\frac{\Delta(l)}{BPM_{GT}(l)} \times 100 \qquad \text{(ideally 0 \%)} \tag{5.14}$$

$$PI\ (in\ \%) = \sum_{l=1}^{L}\frac{\mathbb{1}_{(\Delta(l)<\epsilon_1)}}{L} \times 100 \qquad \text{(ideally 100\%)} \tag{5.15}$$

In Equations (5.13), (5.14), and (5.15), $\Delta(l)$ is defined as

$$\Delta(l)\ (in\ BPM) = \left|BPM_{HR\,method}(l) - BPM_{GT}(l)\right|, \tag{5.16}$$

where $\left|.\right|$ is the absolute value. $BPM_{HR\,method}(l)$ is the heart rate in beat per minutes (BPM) calculated using each one of the tested methods (DWL, TROIKA, and JOSS) at time step l. $BPM_{GT}(l)$ is the HR ground truth value in BPM obtained as $BPM_{GT}(l) = HR_{GT}(l) \times 60$, where $HR_{GT}(l)$ is calculated using Equation (5.1). In Equation (5.15), $\mathbb{1}$ is the indicator function that returns 1 if $\Delta(l) < \epsilon_1$ and 0 otherwise. ϵ_1 was set to 5 BPM. The heart rate estimated using DWL at time step l (in BPM) is calculated as $BPM_{DWL}(l) = \hat{HR}^{(3)}(l) \times 60$, where $\hat{HR}^{(3)}(l)$ is the 3-point-averaged heart rate estimate (see Equation (5.3)).

5.6.2 DWL performance on wrist data

Data were collected from fourteen (14) participants while standing, walking, and running on the treadmill, following the experimental protocol described in Section 5.2. In this section, we analyze data collected from participants 1 to 11. Data from participants 12, 13, and 14 are not included in our analysis since for these participants, the system suffered from physical malfunction (intermittent readings due to loss of sensor contact). However, we still provide the data for these participants in the repository in [41]. Every 2 seconds, the preceding 8-second-long green and IR PPG data were used to generate a short-term 3-point-average HR estimate, $\hat{HR}^{(3)}(l)$, using the DWL method. Heart rate levels obtained using DWL are compared to those of TROIKA[5] [33] and JOSS[6] [34].

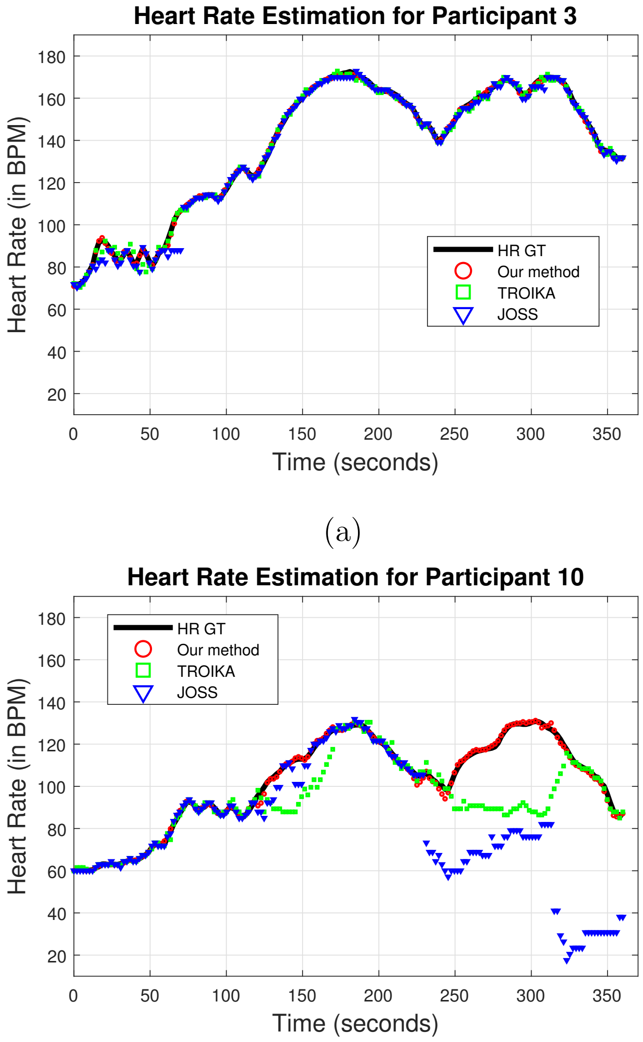
As examples, we show in Figure 5.12 the HR calculated for the whole experimental run for two participants, participant 3 (Figure 5.12 (a)) and participant 10 (Figure 5.12 (b)). We use red circles, green squares, and blue triangles to represent the HR values calculated using DWL, TROIKA, and JOSS, respectively. The ground truth HR is the solid black line. In Figure 5.12 (a) all three methods generate accurate HR estimates (the magnitude of the noise level present in the signals of participant 3 was small). For participant 10 (see Figure 5.12 (b)), however, TROIKA lost track of the correct heart rate from 120 to 175 seconds and from 250 to 325 seconds. This phenomenon (losing track of the correct heart rate) is referred to as Lock Loss. Similary, JOSS suffered from a Lock Loss from 225 seconds until the end of the experimental run. During these intervals, the DWL method was still able to estimate the participants' HR accurately (see red circles of Figure 5.12 (b)).

We calculate MAE, $MAEP$, PI, and CT for all eleven (11) experimental participants and present them in Tables 5.2, 5.3, 5.4, and 5.5, respectively. In Table 5.2, we show the MAE for DWL, TROIKA, and JOSS. We calculate and report the MAE mean and standard deviation for each method in the second to last row of

---

[5] For the TROIKA's implementation, we used a sampling rate of 100 Hz. We recreated the TROIKA code using MATLAB. Our code was tested on the same dataset of the TROIKA paper and compared to the results presented in [33]. The results using our code are very close to the results presented in the TROIKA paper [33].

[6] For the JOSS's implementation, we used a sampling rate of 25 Hz, as suggested in the JOSS paper [34]. We recreated the JOSS code using MATLAB. Our code was tested on the dataset used in JOSS paper and compared to the results presented in [34]. The results using our code are very close to the results presented in the JOSS paper [34].



(a)

(b)

Figure 5.12 HR calculated for the whole experimental run using DWL method (red circles), TROIKA (green squares), and JOSS (blue triangles). (a) HR values for participant 3. DWL, TROIKA, and JOSS were able to calculate accurate heart rate levels. (b) HR values for participant 10. TROIKA lost track of the correct heart rate from 120 to 175 seconds and from 250 to 325 seconds. Similary, JOSS lost track of the correct heart rate from 225 seconds until the end of the experimental run. DWL methods was able to estimate the participant's HR accurately during the whole experimental run.

Table 5.2. In the last row of Table 5.2, we calculate the MAE mean and standard deviation of all participants that do not suffer from Lock Loss. Lock Loss happens if MAE exceeds 5 BPM. participants who suffer Lock Loss are underlined. Table 5.3 summarizes the $MAEP$ for DWL, TROIKA, and JOSS. We calculate and report the

$MAEP$ mean and standard deviation for each method in the last row of Table 5.3. Moreover, we calculate PI for DWL, TROIKA, and JOSS, and report it in Table 5.4. In the last row of Table 5.4, we calculate the PI mean and standard deviation of all participants.

Table 5.2 MAE in BPM for all Eleven (11) Experimental Participants, using DWL, TROIKA, and JOSS (ideal MAE is 0)

	***HR* Calculation Methods**		
Participants	**TROIKA**[1]	**JOSS**[1]	**DWL**[2]
Participant 1	1.09	1.39	0.74
Participant 2	4.3	<u>87.61</u>	1.48
Participant 3	1.21	1.41	0.63
Participant 4	1.96	1.72	2.36
Participant 5	<u>7.85</u>	<u>5.49</u>	1.86
Participant 6	2.57	2.73	1.18
Participant 7	1.83	2.03	1.64
Participant 8	1.08	0.84	0.61
Participant 9	1.73	1.86	0.76
Participant 10	<u>9.34</u>	<u>21.8</u>	0.85
Participant 11	2.72	4.87	1.31
Average*	3.24|2.82 BPM	11.98|25.79 BPM	1.22|0.57 BPM
Average without Lock Loss**	2.05|1.03 BPM	2.11|1.24 BPM	1.22|0.57 BPM

[1] JOSS and TROIKA use tri-axial accelerometer data as noise reference.

[2] DWL method uses an IR PPG signal as noise reference.

* The second to last row shows the MAE average of all eleven (11) participants shown as "mean|standard deviation."

** The last row shows the MAE average of participants that do not suffer from Lock Loss (MAE value is less than 5 BPM). All MAE values that exceed 5 BPM are underlined and not included into the calculation of the average performance.

As shown in Tables 5.2 the average MAE for all eleven participants using the DWL method is MAE of 1.22|0.57 BPM ("mean|standard deviation") (see Table 5.2), which is smaller than average MAE of TROIKA (3.24|2.82 BPM) and JOSS (11.98|25.79 BPM), respectively. When we exclude participants who suffer from Lock

Table 5.3 $MAEP$ in % for all Eleven (11) Experimental Participants, Using DWL, TROIKA, and JOSS (Ideal $MAEP$ is 0%)

	HR **Calculation Methods**		
Participants	**TROIKA**	**JOSS**	**DWL**
Participant 1	0.92	1.19	0.66
Participant 2	3.88	59.22	1.47
Participant 3	1.03	1.23	0.5
Participant 4	1.43	1.31	1.61
Participant 5	4.96	3.41	1.19
Participant 6	2.06	2.24	0.85
Participant 7	1.38	1.49	1.16
Participant 8	0.77	0.62	0.42
Participant 9	1.94	2.08	0.81
Participant 10	7.91	19.53	0.83
Participant 11	2.07	3.19	1.00
Average[*]	2.58|2.19 %	8.68|17.6 %	0.95|0.38 %

[*] The last row shows the $MAEP$ average of all eleven (11) participants shown as "mean|standard deviation."

Loss (shown in the last row of Table 5.2), DWL (with the same average $MAE = 1.22|0.57$ BPM) still yields a smaller average MAE than that of TROIKA (with average $MAE = 2.05|1.03$ BPM) and JOSS (with average $MAE = 2.11|1.24$ BPM). Note that the MAE calculated using DWL method did not exceed 5 BPM for any of the participants. However, this was not the case for TROIKA and JOSS method. participant 10 is an example where the MAE of TROIKA (9.34 BPM) and JOSS (21.8 BPM) exceeds 5 BPM, whereas the MAE of the DWL method is 0.85 BPM.

In addition to MAE, we calculate average $MAEP$ of all three methods for all eleven participants. Average $MAEP$ of DWL method of 0.95|0.38 % is smaller than average $MAEP$ of TROIKA (2.58|2.19 %) and JOSS (8.68|17.6 %) (see Table 5.3).

Table 5.4 summarized the PI values for all eleven (11) participants. The PI of DWL is larger than the PI of TROIKA and JOSS. For instance, on average, the PI

Table 5.4 PI in % for all Eleven (11) Experimental Participants, Using DWL, TROIKA, and JOSS (Ideal PI is 100%)

	HR **Calculation Methods**		
Participants	**TROIKA**	**JOSS**	**DWL**
Participant 1	96.02	93.75	98.36
Participant 2	72.57	6.86	94.29
Participant 3	96.02	93.75	100
Participant 4	88.07	90.91	86.93
Participant 5	61.58	80.79	88.7
Participant 6	80.68	84.66	100
Participant 7	90.96	88.7	91.53
Participant 8	97.18	98.31	100
Participant 9	90.91	89.2	100
Participant 10	63.84	57.63	99.44
Participant 11	84.75	80.23	95.48
Average[*]	83.87|12.75 %	78.62|26.16 %	95.88|4.9 %

[*] The last row shows the PI average of all eleven (11) participants shown as “mean|standard deviation.”

of DWL method is 95.88|4.9 % that is larger than that of TROIKA with 83.87|12.75 % and JOSS with 78.62|26.16 %.

The CT is an indication of the algorithm’s computational complexity. In order to be implement in wearable devices, the algorithm should be able to run in real-time and be energy efficient. A desirable algorithm should have a small CT. Table 5.5 shows the CT of DWL, TROIKA, and JOSS for participants 1 to 11. The average CT of DWL is smaller than that of TROIKA and JOSS. For instance, the average CT of DWL is 3.0|0.3 seconds is smaller than the average CT of TROIKA with 247.7|43.8 seconds and JOSS with 8.5|0.24 seconds.

Additionally, we show the *Bland-Altman plot* (Figure 5.13 (a)) of the HR values computed using DWL method for participants one (1) through eleven (11). The Bland-Altman plot describes the agreement between two quantitative measurements

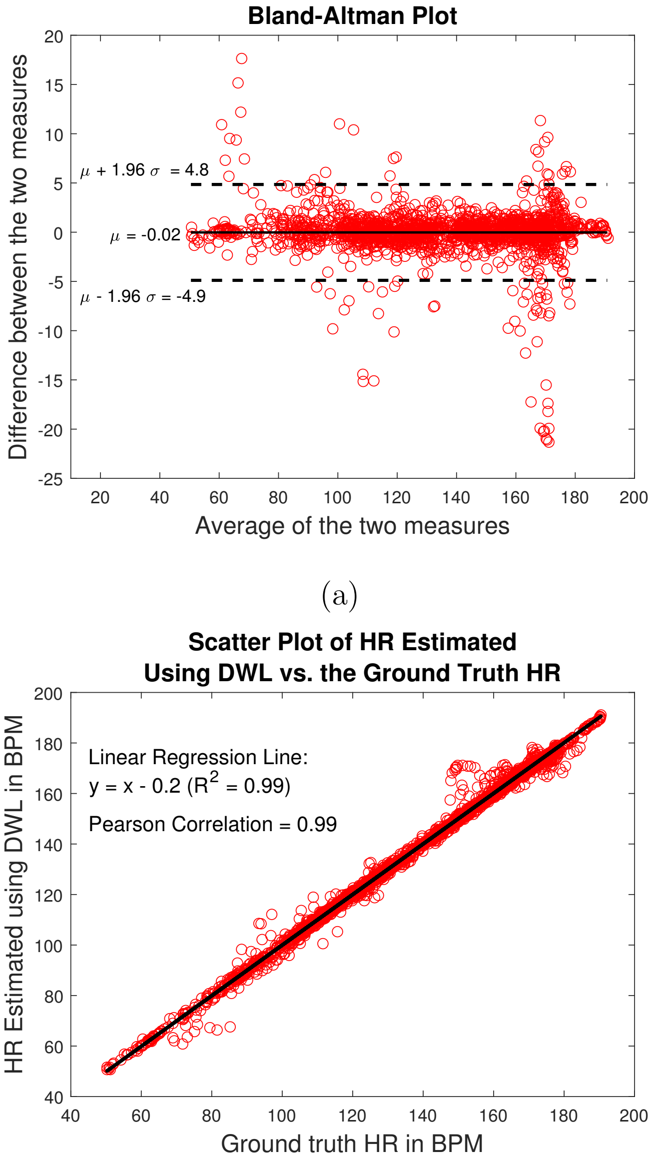
Table 5.5 CT in Seconds for all Eleven (11) Experimental Participants, Using DWL, TROIKA, and JOSS

	HR **Calculation Methods**		
Participants	**TROIKA**	**JOSS**	**DWL**
Participant 1	243.6	8.5	2.8
Participant 2	238.0	8.3	3.0
Participant 3	294.6	8.6	3.1
Participant 4	259.7	8.4	2.7
Participant 5	246.5	8.5	3.1
Participant 6	239.7	9.2	3.7
Participant 7	237.0	8.5	2.7
Participant 8	326.4	8.5	2.8
Participant 9	278.3	8.3	2.8
Participant 10	194.3	8.5	3.3
Participant 11	166.9	8.5	3.1
Average*	**247.7 | 43.8 s**	**8.5 | 0.24 s**	**3.0 | 0.3 s**

Note: the results shown in this table were generated by MATLAB R2022b on a personal computer, with an Intel®Core^{TM} i9-10900K CPU running at 3.70 GHz, 32GB RAM, and Windows 11 operating system.

* The last row shows the CT average of all eleven (11) participants shown as "mean|standard deviation."

(A and B) by constructing the Limits of Agreements (LOA). These statistical limits are calculated by using the mean and the standard deviation of the differences between the two measurements. The resulting graph is a scatter plot, in which the y-axis shows the difference between the two paired measurements (A-B) and the x-axis represents the average of these measures ((A+B)/2) [65]. The LOA we use is $[\mu - 1.96 \times \sigma \ , \ \mu + 1.96 \times \sigma]$ where μ is the average difference between each HR estimate and the associated ground-truth HR against their average, and σ is the standard deviation (95% of all differences are within this range) [33]. The LOA in Figure 5.13 (a) is [-4.9, 4.8] BPM. Moreover, we construct the scatter plot of the HR estimated using DWL method versus the associated ground truth HR for participants one (1) through eleven (11). The scatter plot is shown in Figure 5.13 (b). We construct a linear regression



(a)

(b)

Figure 5.13 (a) Bland-Altman plot of HR estimated using DWL method and the ground truth HR for participants one (1) to eleven (11). The LOA = [-4.9, 4.8] BPM. (b) Scatter plot of HR estimated using DWL method (on the y-axis) vs. the ground truth HR (x-axis) for participants one (1) to eleven (11). The linear regression line that fits the data is shown in black. The line is y = x - 0.2 ($R^2 = 0.99$). The Pearson correlation is found to be 0.99.

for the data points of Figure 5.13 (b). The fitted line is y = x - 0.2 ($R^2 = 0.99$), where x is the ground truth HR and y is the HR estimated using DWL method. The Pearson correlation between the HR estimated using DWL method and ground truth HR is also calculated and found to be 0.99. The high R^2 value and Pearson correlation indicate that DWL method is able to compute accurate HR levels.

5.7 Validation of the DWL Method on Palm Data

In order to validate the performance of the DWL framework, we ran a second experiment (validation run). During the second experiment, we asked the same volunteers who participated in our previous experiment to run on the treadmill again, following the experimental protocol described in Section 5.2. We reused the same ECG, accelerometer, and PPG sensors. The only difference was that we mount the dual wavelength sensor onto the participant's palm (instead of wrist). Both wrist and palm experiments took place on the same day. There was a break of approximately 15 minutes between the first and the second run during which the dual-wavelength PPG sensor was relocated from the wrist to the palm of the participant. Data for all participants are provided in the repository in [41].

We observed that the quality of the PPG signals collected during the second run (palm run) is better than that of the PPG signals collected during the first run (wrist run). For instance, the SNR mean value of the palm IR PPG is $\upsilon^{IR}_{SNR,\, palm} = -7\ dB$ and the SNR mean value of the palm green PPG signal, $\upsilon^{G}_{SNR,\, palm} = -2.2\ dB$ (as opposed to $\upsilon^{IR}_{SNR,\, wrist} = -8.5\ dB$ and $\upsilon^{G}_{SNR,\, wrist} = -4.8\ dB$ for wrist-collected PPG signals). Data for all participants are provided in the repository in [41].

Participants 5 and 13 deviated from the data collection protocol by interfering with the sensor during collection. Their measurements were excluded from the analysis we provide (but are available in the repository in [41]). In the following analysis, the DWL framework parameters derived in Sections 5.4 are kept unchanged.

MAE, $MAEP$, and PI were calculated from the twelve (12) participants of the "palm run" for DWL, TROIKA, and JOSS, and presented in Tables 5.6, 5.7, and 5.8, respectively. As shown in Tables 5.6, 5.7, and 5.8, DWL succeeds in generating accurate HR estimates when tested on the PPG data collected from the second run (palm run) for all the twelve (12) subjects.

Table 5.6 MAE in BPM for all Twelve (12) Participants During the Palm Run, Using DWL, TROIKA, and JOSS (Ideal MAE is 0)

	HR **Calculation Methods**		
Participants	**TROIKA**	**JOSS**	**DWL**
Participant 1	0.84	0.89	0.7
Participant 2	1.06	1.01	0.71
Participant 3	0.78	0.64	0.68
Participant 4	1.2	1.32	0.79
Participant 6	2.64	3.18	1.76
Participant 7	2.34	2.16	1.89
Participant 8	1.22	1.11	0.9
Participant 9	1.42	1.56	0.68
Participant 10	1.34	1.21	0.82
Participant 11	2.96	2.65	2.05
Participant 12	2.08	<u>52.32</u>	1.57
Participant 14	3.6	<u>86.52</u>	3.1
Average*	1.79|0.92 BPM	12.88|27.41 BPM	1.3|0.77 BPM
Average without Lock Loss**	1.79|0.92 BPM	1.57|0.83 BPM	1.3|0.77 BPM

* The second to last row shows the MAE average of all eleven (11) participants shown as "mean|standard deviation."

** The last row shows the MAE average of participants that do not suffer from Lock Loss (MAE value is less than 5 BPM). All MAE values that exceed 5 BPM are underlined and not included into the calculation of the average performance.

As shown in Tables 5.6 the average MAE for all twelve 'palm run' participants using the DWL method 1.3|0.77 BPM ("mean|standard deviation") (see Table 5.6), which is smaller than average MAE of TROIKA (1.79|0.92 BPM) and JOSS (12.88|27.41 BPM), respectively. When we exclude participants who suffer from Lock Loss (shown in the last row of Table 5.6), DWL (with the same average MAE = 1.3|0.77 BPM) still yields a smaller average MAE than that of TROIKA (with average MAE = 1.79|0.92 BPM) and JOSS (with average MAE = 1.57|0.83 BPM). Note that the MAE calculated using DWL and TROIKA method did not exceed 5 BPM for any of the participants. However, this was not the case for JOSS method. For articipants 12 and 14, the MAE of JOSS exceeds 5 BPM.

Table 5.7 $MAEP$ in % for all Twelve (12) Participants During the Palm Run, Using DWL, TROIKA, and JOSS (Ideal $MAEP$ is 0%)

	HR **Calculation Methods**		
Participants	**TROIKA**	**JOSS**	**DWL**
Participant 1	0.74	0.81	0.58
Participant 2	0.81	0.73	0.52
Participant 3	0.57	0.47	0.49
Participant 4	0.96	1.06	0.61
Participant 6	2.17	2.63	1.38
Participant 7	1.8	1.66	1.39
Participant 8	0.85	0.77	0.63
Participant 9	1.47	1.54	0.7
Participant 10	1.38	1.25	0.77
Participant 11	2.1	1.96	1.37
Participant 12	1.48	32.72	1.13
Participant 14	2.84	56.51	2.56
Average*	1.43|0.69 %	8.51|17.62 %	1.01|0.6 %

* The last row shows the $MAEP$ average of all Eleven (11) participants shown as "mean|standard deviation."

In addition to MAE, we calculate average $MAEP$ of all three methods for all eleven participants. Average $MAEP$ of DWL method of 1.01|0.6 % is smaller than average $MAEP$ of TROIKA (1.43|0.69 %) and JOSS (8.51|17.62 %) (see Table 5.7).

Table 5.8 summarized the PI values for all twelve (12) participants of the 'palm run.' The PI of DWL is larger than the PI of TROIKA and JOSS. For instance, on average, the PI of DWL method is 95.33|6.46 % that is larger than that of TROIKA with 90.23|8.94 % and JOSS with 80.93|29.18 %.

Moreover, we show the summary of the performance metrics (MAE, $MAEP$, and PI) obtained for the first run (the "wrist run") in Table 5.9 and for the second run (the "palm run") in Table 5.10. The results are presented as "mean|standard deviation."

Table 5.8 PI in % for all Twelve (12) Participants During the Palm Run, Using DWL, TROIKA, and JOSS (Ideal PI is 100%)

	HR **Calculation Methods**		
Participants	**TROIKA**	**JOSS**	**DWL**
Participant 1	98.29	97.71	100
Participant 2	94.89	98.3	100
Participant 3	100	99.43	100
Participant 4	96.59	94.89	98.3
Participant 6	82.49	80.79	89.83
Participant 7	88.7	89.83	94.35
Participant 8	96.59	97.16	97.16
Participant 9	92.66	92.66	100
Participant 10	95.45	96.02	99.43
Participant 11	80.11	83.52	88.64
Participant 12	86.93	30.68	97.16
Participant 14	70.06	10.17	79.1
Average*	90.23|8.94 %	80.93|29.18 %	95.33|6.46 %

* The last row shows the PI average of all eleven (11) participants shown as "mean|standard deviation."

Tables 5.9 and 5.10 show that the DWL method performs as well when the measurements were taken from the wrist as when they were taken from the palm.

5.8 Conclusion

We presented a framework for heart rate calculation under motion using a dual-wavelength (green and IR) PPG sensor. We used PPG data collected from fourteen (14) individuals engaged in high-intensity full body exercise. Analysis of green and IR PPG signals indicates that the IR PPG signal is a good noise reference signal. We used this observation to develop a motion-resistant HR calculation method derived from

Table 5.9 Summary of Performance Metrics for Run 1 (Wrist Run) – For Run 1, the Average Performance of Eleven (11) Participants is Shown; Results Are Represented as "Mean|Standard Deviation"

	Run 1 (wrist run)		
	TROIKA	**JOSS**	**DWL**
Average MAE (BPM) of all participants	3.24|2.82	11.98|25.79	1.22|0.57
Average MAE (BPM) of participants without Lock Loss	2.05|1.03	2.11|1.24	1.22|0.57
Average MAEP (%) of all participants	2.58|2.19	8.68|17.6	0.95|0.38
Average PI (%) of all participants	83.87|12.75	78.62|26.16	95.88|4.9

Table 5.10 Summary of Performance Metrics for Run 2 (Validation Palm Run) – For run 2, the Average Performance of Twelve (12) Participants is Shown; Results Are Represented as "Mean|Standard Deviation"

	Run 2 (palm run)		
	TROIKA	**JOSS**	**DWL**
Average MAE (BPM) of all participants	1.79|0.92	12.88|27.41	1.3|0.77
Average MAE (BPM) of participants without Lock Loss	1.79|0.92	1.57|0.83	1.3|0.77
Average MAEP (%) of all participants	1.43|0.69	8.51|17.62	1.01|0.6
Average PI (%) of all participants	90.23|8.94	80.93|29.18	95.33|6.46

[35] that measures noise components from the IR PPG signal. Afterwards, a green PPG signal is denoised and used for HR calculation. The proposed method, Dual Wavelength (DWL), was tested on experimental data collected from participants' wrists while the participants were standing, walking, and running on a treadmill. The performance of the method, using several measures of accuracy and computational effort, was then compared to popular methods in the literature that use data from a tri-axial accelerometer for denoising, namely, TROIKA and JOSS. Using the experimental wrist-data we collected, we showed that the DWL method exhibits good performance in the face of motion artifacts. For instance, DWL yielded a Mean Absolute Error (MAE) of 1.22|0.57 BPM, Mean Absolute Error Percentage

($MAEP$) of 0.95|0.38 %, and performance index (PI) (which is the frequency in percent of the event that we obtain a HR estimate that is within ± 5 BPM of the HR ground truth) of 95.88|4.9 %. Moreover, DWL yielded a short computation period of 3.0|0.3 seconds to process a 360-second-long run. We validated the performance of the DWL method by testing it on data collected from the participants' palms, obtaining similar behavior. The DWL method is desirable since (1) it performed well under high-intensity full body repetitive "macro-motion," exhibiting high accuracy in the presence of motion artifacts (as compared to the leading accelerometer-dependent HR calculation techniques TROIKA and JOSS); (2) it used only PPG signals; auxiliary signals such as accelerometer signals were not needed; and (3) it was computationally efficient, hence implementable in wearable devices.

CHAPTER 6

CLOSING REMARKS – ADDRESSING MOTION ARTIFACTS IN PULSE OXIMETRY

In this dissertation, we studied the problem of motion artifacts in the pulse oximetry systems. We developed motion-resistant methods for high-quality SpO_2 and HR calculation, under significant level of motion noise. We first employed a comb filter, tuned to the participant's heart rate (as a pre-processing step), to capture the heart rate components in the PPG signal and suppress motion artifacts. We then showed the benefits of integrating vital signs calculated from multiple PPG channels (when available). Lastly, we developed a motion-resistant HR calculation framework that uses PPG signals only.

Study and testing of the motion-resistant methods we developed in this thesis can be expanded. Next, we list a few challenges that call for future research.

- All tested datasets were collected from healthy and physically fit participants. Therefore, irregular heart rhythms and their effect on SpO_2 level calculation using our methods were not studied.
- The SpO_2 ground truth level we employed in chapters 3 and 4 was obtained using the highly-accurate Nonin $8000R$ sensor [30], mounted on the participant's temple. It may have been preferable to draw blood samples from the participants and measure arterial oxygen saturation directly. However, this invasive procedure was not permitted under the IRB-approved experimental protocol.
- The ECG-derived HR of chapter 3 can be replaced by the PPG-derived HR of chapter 5. In this case, the ECG-HR-tuned comb filter will be substituted by a PPG-HR-tuned comb filter. The proposed system will generate high quality HR and SpO_2 levels using green, red, and infrared PPG signals only.

REFERENCES

[1] J. G. Webster, *Design of Pulse Oximeters*, 1st ed. Boca Raton, Florida: CRC Press, 1997.

[2] J. M. Goldman, M. T. Petterson, R. J. Kopotic, and S. J. Barker, "Masimo signal extraction pulse oximetry," *Journal of Clinical Monitoring and Computing*, vol. 16, no. 7, pp. 475–483, 2000.

[3] H. W. Lee, J. W. Lee, W. G. Jung, and G. K. Lee, "The periodic moving average filter for removing motion artifacts from PPG signals," *International Journal of Control, Automation and Systems*, vol. 5, no. 6, pp. 701–706, 2007.

[4] G. Clarke, "Signal Quality Analysis in Pulse Oximetry: Modelling and Detection of Motion Artifact," Ph.D. dissertation, Carleton University, Ottawa, Ontario, Canada, 2015.

[5] M. J. Grap, "Pulse Oximeter," *Critical Care Nurse*, vol. 22, no. 3, pp. 69–74, 2002.

[6] K. Bhattacharya, "Takuo Aoyagi — a Tribute to the Brain Behind Pulse Oximetry," *Indian Journal of Surgery*, vol. 82, no. 6, pp. 1332–1333, 2020.

[7] N. Kohden, "Takuo Aoyagi and the pulse oximeter." [Online]. Available: https://www.nihonkohden.com/topics/aoyagi.html. Accessed on 20 February 2023.

[8] P. Renevey, R. Vetter, J. Krauss, P. Celka, and Y. Depeursinge, "Wrist-located pulse detection using IR signals, activity and nonlinear artifact cancellation," in *Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 3, Istanbul, Turkey, 2001, pp. 3030–3033.

[9] S. Kästle, F. Noller, S. Falk, A. Bukta, E. Mayer, and D. Miller, "A new family of sensors for pulse oximetry," *Hewlett-Packard Journal*, vol. 48, no. 1, pp. 39–61, 1997.

[10] J. Ghairat and H. Mouhsen, "Motion artifact reduction in PPG signals," Ph.D. dissertation, Lund University, Lund, Sweden, 2015.

[11] S. Prahl, "Tabulated Molar Extinction Coefficient for Hemoglobin in Water," 1998. [Online]. Available: https://omlc.org/spectra/hemoglobin/summary.html. Accessed on 20 February 2023.

[12] S.-S. Oak and P. Aroul, "How to Design Peripheral Oxygen Saturation (SpO2) and Optical Heart Rate Monitoring (OHRM) Systems Using the AFE4403," Tech. Rep. March, 2015. [Online]. Available:

http://www.ti.com/lit/an/slaa655/slaa655.pdf. Accessed on 20 February 2023.

[13] P. Kyriacou, K. Budidha, and T. Y. Abay, “Optical techniques for blood and tissue oxygenation,” *Encyclopedia of Biomedical Engineering*, vol. 1-3, no. April, pp. 461–472, 2019.

[14] J. M. Cho, Y. K. Sung, K. W. Shin, D. J. Jung, Y. S. Kim, and N. H. Kim, “A preliminary study on photoplethysmogram (PPG) signal analysis for reduction of motion artifact in frequency domain,” in *IEEE-EMBS Conference on Biomedical Engineering and Sciences.* Langkawi, Malaysia: IEEE, 2012, pp. 28–33.

[15] B. M. Jayadevappa, G. H. Kiran Kumar, L. H. Anjaneya, and S. H. Mallikarjun, “Design and Development of Electro-Optical System for Acquisition of Ppg Signals for the Assessment of Cardiovascular System,” *International Journal of Research in Engineering and Technology*, vol. 03, no. 06, pp. 520–525, 2014.

[16] B. Widrow, C. S. Williams, J. R. Glover, J. M. McCool, R. H. Hearn, J. R. Zeidler, J. Kaunitz, E. Dong, and R. C. Goodlin, “Adaptive Noise Cancelling: Principles and Applications,” *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692–1716, 1975.

[17] J. C. Stapleton and S. C. Bass, “Adaptive noise cancellation for a class of non-linear IIR filters,” *IEEE Transactions on circuits and systems*, vol. 32, no. 2, pp. 143–150, 1985.

[18] S. J. Barker, “”Motion-resistant” pulse oximetry: A comparison of new and old models,” *Anesthesia and Analgesia*, vol. 95, no. 4, pp. 967–972, 2002.

[19] M. S. Lipnick, J. R. Feiner, P. Au, M. Bernstein, and P. E. Bickler, “The Accuracy of 6 Inexpensive Pulse Oximeters Not Cleared by the Food and Drug Administration,” *Anesthesia and Analgesia*, vol. 123, no. 2, pp. 338–345, aug 2016. [Online]. Available: http://journals.lww.com/00000539-201608000-00009

[20] A. Louie, J. R. Feiner, P. E. Bickler, L. Rhodes, M. Bernstein, and J. Lucero, “Four Types of Pulse Oximeters Accurately Hypoxia during Low Perfusion and Motion,” *Anesthesiology*, vol. 128, no. 3, pp. 520–530, 2018.

[21] A. Galli, C. Narduzzi, and G. Giorgi, “Measuring Heart Rate during Physical Exercise by Subspace Decomposition and Kalman Smoothing,” *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 5, pp. 1102–1110, 2018.

[22] Q. Zhang, X. Zeng, W. Hu, and D. Zhou, “A Machine Learning-Empowered System for Long-Term Motion-Tolerant Wearable Monitoring of Blood Pressure and Heart Rate With Ear-ECG/PPG,” *IEEE Access*, vol. 5, pp. 10 547–10 561, 2017. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7933339

[23] J. Pan and W. J. Tompkins, “A Real-Time QRS Detection Algorithm,” *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, 1985.

[24] S. Kästle, F. Noller, S. Falk, A. Bukta, E. Mayer, and D. Miller, “A new family of sensors for pulse oximetry,” *Hewlett-Packard Journal*, vol. 48, no. 1, pp. 39–61, 1997.

[25] S. Saechia, J. Koseeyaporn, and P. Wardkein, “Human identification system based ECG signal,” in *IEEE Region 10 Annual International Conference (TENCON).* Melbourne, VIC, Australia: IEEE, 2005, pp. 2–5.

[26] M. K. Islam, A. N. M. M. Haque, G. Tangim, T. Ahammad, and M. R. H. Khondokar, “Study and Analysis of ECG Signal Using MATLAB and LABVIEW as Effective Tools,” *International Journal of Computer and Electrical Engineering*, vol. 4, no. 3, pp. 404–408, 2012.

[27] K. Raviprakash, “ECG simulation using MATLAB,” 2006. [Online]. Available: www.mathworks.com/matlabcentral/fileexchange/10858. Accessed on 20 February 2023.

[28] J. Karvonen and T. Vuoarimaa, “Heart Rate and Exercise Intensity During Sports Activities,” *Sports Medicine*, vol. 5, no. 5, pp. 303–311, 1988. [Online]. Available: https://link.springer.com/article/10.2165/00007256-198805050-00002

[29] H. Tanaka, K. D. Monahan, and D. R. Seals, “Age-predicted maximal heart rate revisited,” *Journal of the American College of Cardiology*, vol. 37, no. 1, pp. 153–156, 2001. [Online]. Available: http://dx.doi.org/10.1016/S0735-1097(00)01054-8

[30] “Model 8000R Reflectance Pulse Oximeter Sensor,” 2016. [Online]. Available: https://www.nonin.com/support/8000r/. Accessed on 20 February 2023.

[31] Y. S. Yan and Y. T. Zhang, “An efficient motion-resistant method for wearable pulse oximeter,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 3, pp. 399–405, 2008.

[32] S. Acharya, A. Rajasekar, B. S. Shender, L. Hrebien, and M. Kam, “Pulse oximeter signal modeling and fusion for hypoxia monitoring,” in *17th International Conference on Information Fusion (FUSION).* Salamanca, Spain: IEEE, 2014.

[33] Z. Zhang, Z. Pi, and B. Liu, “TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise,” *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 522–531, 2015.

[34] Z. Zhang, “Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction,” *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 8, pp. 1902–1910, 2015.

[35] Y. Zhang, S. Song, R. Vullings, D. Biswas, N. Simões-Capela, N. Van Helleputte, C. Van Hoof, and W. Groenendaal, “Motion artifact reduction for wrist-worn photoplethysmograph sensors based on different wavelengths,” *Sensors (Switzerland)*, vol. 19, no. 3, 2019.

[36] D. Castaneda, A. Esparza, M. Ghamari, C. Soltanpur, and H. Nazeran, “A review on wearable photoplethysmography sensors and their potential future applications in health care,” *Physiology and behavior*, vol. 176, no. 1, pp. 139–148, 2016.

[37] S. O. Haykin, “Adaptive Filter Theory.” Pearson, 2002, ch. 12, pp. 571–599.

[38] J. M. Paros, “Triaxial Accelerometer Assembly and In-Situ Calibration Method for Improved Geodetic and Seismic Measurements,” 2017. [Online]. Available: https://patents.google.com/patent/US9645267B2/en?oq=9645267

[39] Bertec, “Bertec Instrumented Treadmill.” [Online]. Available: https://www.bertec.com/products/instrumented-treadmills. Accessed on 20 February 2023.

[40] The Math Works Inc., “MALAB 2022b.” [Online]. Available: www.mathworks.com/. Accessed on 20 February 2023.

[41] L. Alkhoury, “DWL Method,” *GitHub Repository*, 2022. [Online]. Available: https://github.com/ludvikalkhoury/DWL-Method.git. Accessed on 20 February 2023.

[42] S. Han, D. Roh, J. Park, and H. Shin, “Design of multi-wavelength optical sensor module for depth-dependent photoplethysmography,” *Sensors (Switzerland)*, vol. 19, no. 24, 2019.

[43] Vishay, “High Power Infrared Emitting Diode, 940 nm, GaAlAs, MQW.” [Online]. Available: https://www.farnell.com/datasheets/2049675.pdf. Accessed on 20 February 2023.

[44] LightHouseLEDs, “5mm Round Top Pure Green LED - Ultra Bright.” [Online]. Available: https://lighthouseleds.com/5mm-led-green-ultra-bright-15-000-mcd.html. Accessed on 20 February 2023.

[45] ——, “5mm Round Top Pure Blue LED - Ultra Bright.” [Online]. Available: https://lighthouseleds.com/5mm-led-blue-ultra-bright-8-000-mcd.html. Accessed on 20 February 2023.

[46] Digi-Key, “OPT101PG4.” [Online]. Available: https://www.digikey.com/en/products/detail/texas-instruments/OPT101PG4/1899803. Accessed on 20 February 2023.

[47] "Delsys Trigno Avanti Sensor." [Online]. Available: https://delsys.com/trigno-avanti/. Accessed on 20 February 2023.

[48] "Trigno EKG Biofeedback Sensor." [Online]. Available: https://delsys.com/trigno-ekg-biofeedback/. Accessed on 20 February 2023.

[49] M. Krzywinski and N. Altman, "Visualizing samples with box plots," *Nature Methods*, vol. 11, no. 2, pp. 119–120, 2014.

[50] M. Wang, Z. Li, Q. Zhang, and G. Wang, "Removal of motion artifacts in photoplethysmograph sensors during intensive exercise for accurate heart rate calculation based on frequency estimation and notch filtering," *Sensors (Switzerland)*, vol. 19, no. 15, 2019.

[51] R. S. Kasbekar and Y. Mendelson, "Evaluation of key design parameters for mitigating motion artefact in the mobile reflectance PPG signal to improve estimation of arterial oxygenation," *Physiological Measurement*, vol. 39, no. 7, 2018.

[52] U. Rubins, V. Upmalis, O. Rubenis, D. Jakovels, and J. Spigulis, "Real-time photoplethysmography imaging system," *IFMBE Proceedings*, vol. 34 IFMBE, pp. 183–186, 2011.

[53] R. W. Wijshoff, M. Mischi, and R. M. Aarts, "Reduction of periodic motion artifacts in photoplethysmography," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 196–207, 2017.

[54] A. R. Relente and L. G. Sison, "Characterization and adaptive filtering of motion artifacts in pulse oximetry using accelerometers," *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, vol. 2, pp. 1769–1770, 2002.

[55] K. W. Chan and Y. T. Zhang, "Adaptive Reduction of Motion Artifacts from Photoplethysmographic Recordings using a Variable Step-Size LMS Filter," in *IEEE Sensors*, 2002, pp. 0–3.

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

[57] A. H. Sayed; and T. Kailath, "A State-Space Approach to Adaptive RLS Filtering," *IEEE Signal Processing Magazine*, vol. 11, no. 3, 1994. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/295229

[58] "dsp.AdaptiveLatticeFilter," 2022. [Online]. Available: https://www.mathworks.com/help/dsp/ref/dsp.adaptivelatticefilter-system-object.html. Accessed on 20 February 2023.

[59] R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. Terpenning, "STL: A Seasonal-Trend Decomposition Procedure Based on Loess," *Journal of Official Statistics*, vol. 6, no. 1, pp. 3–73, 1990.

[60] L. B. Wood and H. H. Asada, "Noise cancellation model validation for reduced motion artifact wearable PPG sensors using MEMS accelerometers," *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pp. 3525–3528, 2006.

[61] S. S. Chowdhury, R. Hyder, M. S. B. Hafiz, and M. A. Haque, "Real-Time Robust Heart Rate Estimation from Wrist-Type PPG Signals Using Multiple Reference Adaptive Noise Cancellation," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 2, pp. 450–459, 2018.

[62] P. Gibbs and H. H. Asada, "Reducing motion artifact in wearable bio-sensors using MEMS accelerometers for active noise cancellation," *Proceedings of the American Control Conference*, vol. 3, pp. 1581–1586, 2005.

[63] H. Hassani, "Munich Personal RePEc Archive Singular Spectrum Analysis: Methodology and Comparison Singular Spectrum Analysis: Methodology and Comparison," *Journal of Data Science*, vol. 4991, no. 4991, pp. 239–257, 2007.

[64] L. Stanković, E. Sejdić, S. Stanković, M. Daković, and I. Orović, "A Tutorial on Sparse Signal Reconstruction and Its Applications in Signal Processing," *Circuits, Systems, and Signal Processing*, vol. 38, no. 3, pp. 1206–1263, 2019.

[65] D. Giavarina, "Understanding Bland Altman Analysis," *Biochemia Medica*, vol. 25, no. 2, pp. 10–16, 2015. [Online]. Available: https://hrcak.srce.hr/file/206210