

**DYNAMIC HEART RATE ESTIMATION USING FACIAL
IMAGES FROM VIDEO SEQUENCES**

YU YONG POH

**FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2016

**DYNAMIC HEART RATE ESTIMATION USING
FACIAL IMAGES FROM VIDEO SEQUENCES**

YU YONG POH

**THESIS SUBMITTED IN FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF
PHILOSOPHY**

**FACULTY OF ENGINEERING
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2016

UNIVERSITY OF MALAYA
ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **YU YONG POH**

Registration/Matric No: **KHA090035**

Name of Degree: **DOCTOR OF PHILOSOPHY**

Title of Project Paper/Research Report/Dissertation/Thesis (“this Work”):

**DYNAMIC HEART RATE ESTIMATION USING FACIAL IMAGES
FROM VIDEO SEQUENCES**

Field of Study: **IMAGE PROCESSING**

I do solemnly and sincerely declare that:

- (1) I am the sole author/writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature

Date:

Subscribed and solemnly declared before,

Witness’s Signature

Date:

Name:

Designation:

ABSTRACT

Video images have been widely used to extract relevant information for different applications. One of the applications is the heart rate estimation using facial images from video sequences. Previous studies have focused only on heart rates that do not vary much throughout the entire video duration. However, dynamic heart rate variation is of interest since it may provide necessary information for daily application. For instance, in individual sports games such as cycling, badminton and tennis, knowing the dynamic heart rates of a player while carrying out an activity will be useful in determining the level of fatigue of that player. In this thesis, novel approaches are developed to estimate dynamic heart rate readings using facial images from video sequences. A challenge for dynamic heart rate estimation is to determine the shortest duration or length of the video sequence without compromising the accuracy of heart rate readings. To address this issue, this thesis reports two approaches: 1) Independent component analysis (ICA) combined with mutual information, 2) the decorrelation of the color components in log-space. In the first approach, ICA is used to recover the heart rate source from the color components of facial images. An important consideration in using short video sequences is that the ICA sources may have insufficient independence among themselves. Without determining the independence of the sources, there is a possibility of the heart rate signal combining with other signals to render an inaccurate reading. Hence in this study, mutual information is integrated with ICA to determine the shortest video duration needed for estimating dynamic heart rate readings accurately. In the second approach, principal component analysis (PCA) is used to recover the uncorrelated signals, including the heart rate signals. From the studies, it is found that the set of color components, namely red, green, and blue, in log-space, are correlated to each other. The principal components may have insufficient uncorrelatedness among themselves when the video duration is too short. Hence, PCA is combined with the Pearson correlation coefficient to determine the

shortest video duration that still gives acceptable accuracy. Two experiments are carried out to validate the proposed approaches. A camcorder is used to capture the facial images of seven subjects. The first experiment involves the measurement of subjects' increasing heart rates while cycling whereas the second experiment involves falling heart beats. All estimated heart rate readings are compared with readings obtained from Polar Team2 Pro. Polar Team2 Pro samples and computes the instantaneous heart rate by measuring at least one electrocardiogram (ECG) signal waveform. Overall experimental results show the proposed method can be used to measure dynamic heart rates where the root mean square error (RMSE) is less than 3 beats per minute (BPM) and the correlation coefficient is 0.99. The respective Bland-Altman plots for each approach indicate that only a small number of estimated heart rate readings are located outside the 95 % limit of agreement interval where the maximum error is less than 8 BPM.

ABSTRAK

Imej video telah digunakan secara meluas untuk mendapat maklumat yang berkaitan. Salah satu aplikasi ialah anggaran kadar denyutan jantung dengan menggunakan imej wajah dari urutan video. Kajian sebelum ini hanya menumpu pada kadar denyutan jantung yang tidak banyak berubah sepanjang tempoh video. Walau bagaimanapun, anggaran kadar denyutan jantung dinamik diberi tumpuan kerana ia dapat memberi maklumat yang sewajarnya untuk aplikasi harian. Sebagai contoh, dalam permainan sukan individu seperti berbasikal, badminton dan tenis, kadar denyutan jantung pemain yang berubah secara dinamik semasa menjalankan aktiviti-aktiviti digunakan untuk menentu tahap keletihan pemain tersebut. Dalam tesis ini, pendekatan-pendekatan baru dibangunkan untuk menganggar bacaan kadar denyutan jantung dinamik dengan menggunakan imej wajah dari urutan video. Cabaran untuk anggaran kadar denyutan jantung dinamik ialah penentuan jangka masa urutan video yang pendek tanpa menjejaskan ketepatan bacaan kadar denyutan jantung. Untuk menangani isu ini, tesis ini melaporkan dua pendekatan: 1) analisis komponen bebas (ICA) digabungkan dengan maklumat bersama, 2) nyah-korelasi komponen-komponen warna dalam ruang logaritma. Dalam pendekatan pertama, ICA digunakan untuk mendapatkan semula sumber kadar denyutan jantung daripada komponen warna imej wajah. Satu pertimbangan yang penting dalam menggunakan urutan video pendek adalah bahawa kemungkinan sumber ICA tidak mempunyai ketidagantungan yang mencukupi di antara mereka. Tanpa menentukan ketidagantungan sumber ICA, isyarat denyutan jantung mungkin masih mengandungi isyarat lain yang akan menjejaskan kejituaan bacaan yang dianggarkan. Oleh itu, dalam kajian ini, ICA digabungkan dengan maklumat bersama untuk memastikan kejituan bacaan tidak terjejas untuk video yang bertempoh pendek. Dalam pendekatan kedua, analisis komponen utama (PCA) digunakan untuk mendapatkan semula isyarat nyah-korelasi, termasuk isyarat kadar jantung. Dari kajian, didapati bahawa set komponen

warna iaitu merah, hijau, dan biru, dalam ruang logaritma, sentiasa berkorelasi antara satu sama lain. Komponen utama mungkin mempunyai nyah-korelasi yang tidak mencukupi di antara mereka sekiranya tempoh video adalah terlalu pendek. Oleh itu, PCA digabungkan dengan pekali korelasi Pearson untuk menentukan tempoh video yang paling pendek yang masih memberikan kejituan bacaan yang sepatutnya. Dua eksperimen dijalankan untuk mengesahkan pendekatan-pendekatan yang dicadangkan. Video kamera digunakan untuk menangkap imej-imej wajah sebanyak tujuh subjek. Eksperimen pertama melibatkan ukuran kadar denyutan jantung subjek yang meningkat ketika berbasikal manakala eksperimen kedua melibatkan ukuran kadar denyutan jantung yang menurun. Semua bacaan kadar denyutan jantung yang didapatkan dalam kajian dibandingkan dengan bacaan sebenar yang diukur dengan alat ukur kadar denyutan jantung Polar team2 Pro. Polar team2 Pro menyampel dan mengira kadar denyutan jantung serta-merta dengan mengukur sekurang-kurangnya satu gelombang isyarat elektrokardiogram.

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my supervisor Prof. Dr. P. Raveendran for the continuous support of my Ph.D study and related research. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor and mentor for my Ph.D study.

I thank my fellow colleagues and labmates for the stimulating discussions, and for all the fun we have had in the last four years. In particular, I am grateful to Dr. Lim Chern Loon for providing technical knowledge, discussion and review in several aspects.

Last but not least, I would like to thank my family for supporting me spiritually throughout the period of my Ph.D study and writing of this thesis.

University of Malaya

TABLE OF CONTENTS

Abstract	iii
Abstrak	v
Acknowledgements	vii
Table of Contents	viii
List of Figures	xi
List of Tables.....	xiv
List of Symbols and Abbreviations.....	xv
List of Appendices	xvi
CHAPTER 1: INTRODUCTION.....	1
1.1 Overview.....	1
1.2 Objectives	4
1.3 Scope and Organization.....	5
1.4 Contribution.....	6
CHAPTER 2:THEORIES OF COMPONENT ANALYSIS AND MATHEMATICAL MODEL FOR IMAGES OF HUMAN SKIN.....	8
2.1 Principal Component Analysis	8
2.2 Independent Component Analysis.....	11
2.3 Simplified Mathematical Model for Images of Human Skin	15
CHAPTER 3: VIDEO BASED HEART RATE ESTIMATION USING TEMPORAL INFORMATION.....	17
3.1 Overview.....	17
3.2 Proposed Method.....	17

3.2.1	Extraction of Temporal Information Using Short-time Fourier Transform	18
3.2.2	Extraction of Temporal Information Using Filter Bank.....	19
3.3	Experimental Study	21
3.3.1	Experimental Setup	21
3.3.2	Experimental Results and Analysis for STFT Approach	24
3.3.3	Experimental Results and Analysis for Filter Bank Approach	36
3.4	Chapter Conclusion	46

CHAPTER 4: DYNAMIC HEART RATE ESTIMATION FROM SHORT VIDEO SEQUENCES48

4.1	Overview.....	48
4.2	Proposed Method.....	48
4.2.1	Workflow of The Proposed Method.....	49
4.2.2	Criterion Determining the Independence of ICA sources	51
4.2.3	Significance of the Minimum Video Duration.....	54
4.3	Experimental Study	58
4.3.1	Experimental Setup	58
4.3.2	First Experiment: Observed Heart Rates Varying from Low to High	59
4.3.3	Second Experiment: Observed Heart Rates Varying from High to Low	65
4.3.4	Heart Rate Estimation for Subjects at Rest Using Proposed Method	71
4.4	Chapter Conclusion	71

CHAPTER 5: DYNAMIC HEART RATE ESTIMATION USING PRINCIPAL COMPONENT ANALYSIS73

5.1	Overview.....	73
5.2	Relationship between RGB and YCbCr Components of Facial Images	73

5.3	Proposed Method.....	76
5.3.1	Relationship between the Correlation among PCs and Video Duration ..	77
5.3.2	Block Diagram of Proposed Model.....	78
5.4	Experimental Study	80
5.4.1	Experimental Setup	80
5.4.2	Experimental Results and Analysis	80
5.4.3	Comparative Study between Proposed Method and Existing Method.....	86
5.5	Chapter Conclusion	87
 CHAPTER 6: CONCLUSION AND FUTURE WORK		88
	References	90
	List of Publications and Papers Presented	96
	Appendix	97

LIST OF FIGURES

Figure 2.1: Two independent signals (sources)	13
Figure 2.2: Two sensors (mixture of the two original sources)	14
Figure 2.3: The separation of original sources using ICA	14
Figure 3.1: Proposed method for video-based heart rate estimation using STFT.....	19
Figure 3.2: Flow chart of video-based heart rate estimation using filter bank	21
Figure 3.3: Proposed method for video-based heart rate estimation using filter bank ...	22
Figure 3.4: Subject cycling during the experiment	24
Figure 3.5: Comparison of all actual and estimated heart rate readings using STFT for the first experiment	26
Figure 3.6: Comparison of all actual and estimated heart rate readings using STFT for the second experiment.....	26
Figure 3.7: Comparison of actual and estimated heart rate readings for each subject in the first experiment	27
Figure 3.8: Comparison of of actual and estimated heart rate readings for each subject in the second experiment.....	30
Figure 3.9: Bland-Altman plot for all estimated heart rate reading using STFT for the first experiment.....	34
Figure 3.10: Bland-Altman plot for all estimated heart rate reading using STFT for the second experiment.....	34
Figure 3.11: The power spectrum distribution obtained from STFT	35
Figure 3.12: Comparison between proposed method and previous methods that used Fourier transform	36
Figure 3.13: Comparison of all actual and estimated heart rate readings using filter bank for the first experiment.....	37
Figure 3.14: Comparison of all actual and estimated heart rate readings using filter bank for the second experiment	38
Figure 3.15: Comparison of actual and estimated heart rate readings for each subject in the first experiment	38

Figure 3.16: Comparison of actual and estimated heart rate readings for each subject in the second experiment.....	42
Figure 3.17: Bland-Altman plot for all estimated heart rate reading using filter bank for the first experiment	45
Figure 3.18: Bland-Altman plot for all estimated heart rate reading using filter bank for the second experiment.....	46
Figure 4.1: Flow chart of the proposed method	50
Figure 4.2: The relationship between the averaged normalized mutual information $C(S_1;S_2;S_3)$ and the video duration and the respective computed heart rates	53
Figure 4.3: The frequency domain of the ICA sources when video duration at point A is selected.....	55
Figure 4.4: The frequency domain of the ICA sources when video duration at point B is selected.....	55
Figure 4.5: The frequency domain of the ICA sources when video duration at point C is selected.....	56
Figure 4.6: Comparison of the mean square and standard deviation of the heart rate errors for the proposed variable video intervals and fixed video intervals. T represents the video interval.....	57
Figure 4.7: Comparison of all actual and estimated heart rate readings for the first experiment.....	60
Figure 4.8: Comparison of actual and estimated heart rate readings for each subject in the first experiment	61
Figure 4.9: Bland-Altman plot for all estimated heart rate reading for the first experiment	65
Figure 4.10: Comparison of all actual and estimated heart rate readings for the second experiment.....	66
Figure 4.11: Comparison of actual and estimated heart rate readings for each subject in the second experiment.....	66
Figure 4.12: Bland-Altman plot for all estimated heart rate reading for the second experiment.....	70
Figure 4.13: Comparison of all actual and estimated heart rate readings for subjects at rest.....	71

Figure 5.1: The distribution of $\log P_R$, $\log P_G$ and $\log P_B$	75
Figure 5.2: The graph of correlation coefficient amongst PCs vs video duration for 3 PCs and 6 PCs respectively	76
Figure 5.3: The relationship of the averaged correlation coefficient R_{avg} and the video duration and the respective computed heart rate.....	78
Figure 5.4: Flow Chart of the proposed method	79
Figure 5.5: Comparison of all actual and estimated heart rate readings	82
Figure 5.6: Comparison of actual and estimated heart rate readings for each subject in the first experiment	82
Figure 5.7: Comparison of actual and estimated heart rate readings for each subject in the second experiment.....	84
Figure 5.8: Comparative study between proposed method and existing method	86

University of Malaysia

LIST OF TABLES

Table 4.1: Summary of the video durations for all heart rate readings of each subject using the proposed method	57
Table 4.2: Summary of heart rate readings results for the first experiment.....	61
Table 4.3: Summary of heart rate readings results for the first experiment.....	70
Table 5.1: Correlation coefficient among $\log P_R$, $\log P_G$ and $\log P_B$	76
Table 5.2: Summary of heart rate readings results obtained from proposed method	81
Table 5.3: Comparison of proposed method (using PCA) and method described in Chapter 4 (using ICA)	87

University of Malaysia

LIST OF SYMBOLS AND ABBREVIATIONS

For example:

BPM	:	Beat per minute
BSS	:	Blind source separation
BVP	:	Blood volume pulse
ECG	:	Electrocardiography
ICA	:	Independent component analysis
JADE	:	Joint Approximate Diagonalization of Eigenmatrices
PC	:	Principal component
PCA	:	Principal component analysis
PPG	:	Photoplethysmography
RGB	:	Red, green and blue
RMSE	:	Root mean square error
ROI	:	Region of interest
STFT	:	Short-time Fourier transform

LIST OF APPENDICES

Appendix A	97
Appendix B	100
Appendix C	102

University of Malaya

CHAPTER 1: INTRODUCTION

1.1 Overview

Human heart rate is measured as the number of heart beats per minute (BPM). It is an important parameter used to reveal the health condition of an individual. The pattern of the measured heart rate can be used to indicate levels of fitness, the presence of disease, stress or fatigue and even blockages in the artery due to diabetes or high cholesterol level.

Currently, the most common method used for human heart rate measurements is by using the Electrocardiography (ECG) machine. The electrodes are attached to the surface of the skin around the wrist and chest of the subject. The electrical activity of the human heart is captured through the attached electrodes. Heart rate measurements using ECG machine is a contact based method which might not be suitable for skin-burned patients and person with autistic disorder (sensitive to touch).

Garbey *et al.* introduced a new approach for human cardiac pulse measurement based on thermal signal analysis of the major blood vessels near the skin surface (Garbey *et al.*, 2007). The modulation of the temperature measured from these blood vessels is caused by the variations in blood flow. In the same year, Pavlidis *et al.* measured the human heart rate and breath rate through bio-heat modeling of facial imagery using a thermal camera (Pavlidis *et al.*, 2007). The cardiac pulse detection at the forehead proposed by Gatto was extracted from the video infrared thermography (Gatto, 2009). This approach is based on the principle that the variations of blood flow during the cardiac cycle will cause the fluctuation of thermal energy released by the body tissue.

Takano and Ohta developed a system to measure the human heart rate and respiratory rate based on the images from the Charge-Coupled Device camera (Takano & Ohta, 2007). The variations of the average brightness in the region of interest within the subject's skin were recorded. These data were processed through a sequence of operations which

involve interpolation, low pass filter and auto-regressive spectral analysis in order to obtain the heart rate and the respiratory rate. In the following year, Verkrusse *et al.* measured human respiration and heart rates through remote sensing of plethysmographic signals under ambient light using digital camera (Verkrusse *et al.*, 2008).

Jonathan and Leahy utilized the camera on the smartphone to capture a series of video frames of a human index finger (Jonathan & Leahy, 2010). The reflections of plethysmographic signals obtained from these video frames were used to compute the human heart rate. The engineering model created by Shi *et al.* was used for cardiac monitoring through reflection photoplethysmography (Shi *et al.*, 2010). This non-contact model is made up of a light source that consists of a Vertical Cavity Surface Emitting Laser (VCSEL) and a photo-detector that consists of a high-speed silicon PiN photodiode.

Photoplethysmography (PPG) is a non-invasive and inexpensive method to measure the variations of blood volume through the variations of light absorption or reflection (Kamshilin *et al.*, 2011). The variations of blood volume in the blood vessels are due to the contraction and relaxation of heart muscles during each cardiac cycle. The relationship between the blood volume pulses and the light in reflection PPG has been investigated by some researchers (Hertzman, 1938; Weinman *et al.*, 1977) since a few decades ago. The principle of PPG is based on the fact that body tissue is less opaque than the blood. Therefore, the increase in blood volume will reduce the intensity of the reflected light from the trans-illuminated tissue. The variations in blood volume will change the intensity of the reflectance accordingly. Therefore, the human heart rate which is the same as the frequency of cardiac cycle can be measured from the plethysmographic signals captured in the video.

Heart rate measurement from video sequences is considered as low cost since the color can be captured using any available video recording device such as video camera,

webcam or mobile phone. This remote and non-contact (without using any special device) heart rate measurement is very suitable for home-based health care applications and telemedicine.

Poh *et al.* developed a non-contact technique to estimate the heart rate of a subject whose body was stationary (Poh *et al.*, 2010; Poh *et al.*, 2011). This contact-free approach is based on automatic face tracking and the use of blind source separation on color channels within the facial region. Besides that, the proposed method is robust to motion artifacts and able to extract the heart rate of multiple people at the same time. They showed that human heart rate can be measured from video recorder, such as webcam, under ambient light.

Their model used a video with duration of 60 seconds that including the entire facial region of a subject. The Red, Green and Blue (RGB) pixel values of each video frame were used as the raw input signals. Blind source separation (BSS) method was utilized to extract the source signals (that contain the heart rate PPG signals) from the RGB input signals. The heart rate was calculated by using peak detection algorithm. The results obtained from their proposed method were compared to the ECG raw signals. Their results showed that BSS is able to extract the heart rate source signals from the facial images under stationary condition.

Poh *et al.* had shown that the human heart rate can be measured from digital color video recordings under normal ambient light. However, the whole frontal face is used as the Region of Interest (ROI) which includes the regions with less or without blood vessels such as the eyes, hair and nostrils. They used a video with duration of 60 seconds to compute the average heart rate variability for this entire duration.

Pursche *et al.* modified this technique by transforming the BSS source signals (the heart rate signals) into frequency domain (Pursche *et al.*, 2012). They divided the facial region into three parts, and concluded that the area around eye and nose (center of the face region) provides better information compared to the other two parts. The time series signals were transformed into frequency domain using Fourier transform. They concluded that this method has higher correlation compared to the peak detection algorithm.

On the other hand, Xu *et al.* designed a simplified mathematical model for images of human skin to obtain the BVP signals (Xu *et al.*, 2014). They developed a model for pigment concentration in human skin, and used it to estimate the heart rate. They computed the heart rate readings from video recordings lasting from 45s to 90s. The subjects are required to keep still during the recording. Their heart rates do not vary much.

Kumar *et al.* proposed a model, known as DistancePPG, to improve the signal-to-noise ratio of the camera-based PPG signal by combining the color change signals obtained from different regions of the face using a weighted average (Kumar *et al.*, 2015). Additionally, they introduced a method to track different regions of the face separately to extract the PPG signals under motion. The method was evaluated on people having diverse skin tones, under various lighting conditions and natural motion scenarios. Kumar *et al.* concluded that the accuracy of heart rate estimation was significantly improved using the proposed method.

1.2 Objectives

The objectives of this thesis are as follow:

- i. To design a model that is able to estimate human heart rates from video sequences, by using component analysis methods;

- ii. To estimate the instantaneous heart rates of subjects performing exercise by using the temporal information extracted from short-time Fourier transform and filter bank;
- iii. To estimate the dynamic heart rates using a short number of video frames without compromising the accuracy of heart rate readings.

1.3 Scope and Organization

This thesis contains chapters describing the research findings and experimental studies about the dynamic heart rate estimation using facial images from video sequences. The following is the summary of the content of the chapters in this thesis.

- i. **Chapter 2: Theories of Component Analysis and Mathematical Model for Images of Human Skin.** This chapter presents an overview of component analysis used for video based heart rate estimation. Applications of component analysis in various fields of image and video processing, and biomedical signal processing, are presented. Particular emphasis is given to the computation of principal component analysis (PCA) and independent component analysis (ICA). A mathematical model for image of human skin is also presented in this chapter. The use of PCA in this mathematical model is discussed in Chapter 5.
- ii. **Chapter 3: Video Based Heart Rate Estimation Using Temporal Information.** This chapter discusses the importance of temporal information in the video based heart rate estimation. The temporal information is extracted using short-time Fourier transform and filter bank to estimate the instantaneous heart rate of subjects cycling during the experiments.

- iii. **Chapter 4: Dynamic Heart Rate Estimation From Short Video Sequences.** An important consideration for dynamic heart rate estimation the video duration should be kept as short as possible without compromising the accuracy of heart rate readings. Although ICA can be used to separate the PPG signal from color components of a video clip, the amount of independence of the ICA sources may be decreased due to the short video duration. This chapter presents a method that uses ICA combined with mutual information to ensure the accuracy is not compromised in the use of short video duration.
- iv. **Chapter 5: Dynamic Heart Rate Estimation Using Principal Component Analysis.** This chapter presents another method to estimate dynamic heart rate estimation by using PCA. It is found that the color components in log-space are correlated to each other. The color components in log-space can be de-correlated using PCA to recover the PPG signal. A comparative study between the ICA (as described in Chapter 4) and PCA is included in this chapter.
- v. **Chapter 6: Conclusion and Future Work.** This chapter concludes the research works presented in the thesis. It summarizes the contents of the thesis and discuss the possibilities of the future work.

1.4 Contribution

One of the contributions of this thesis is that it proposes two methods to estimate the instantaneous heart rate using the temporal information. The proposed algorithms are able to estimate heart rates that vary rapidly.

Another contribution is the use of ICA combined with mutual information to estimate the dynamic heart rate from a short video sequence. An important consideration in

estimating the instantaneous heart rate is to use a limited number of video frames or short video duration. If the video duration is too long, the model may not report the accuracy for large heart rate variations. Therefore, the number of video frames should be taken into account when designing the heart rate estimation model.

Third contribution of this thesis is the discovery of correlation between color components of human skin in log-space. By exploiting the relationship between the color components in log-space, PCA can be then used as another model to estimate the dynamic heart rates from short video sequences.

University of Malaysia

CHAPTER 2: THEORIES OF COMPONENT ANALYSIS AND MATHEMATICAL MODEL FOR IMAGES OF HUMAN SKIN

This chapter describes the preliminary studies that had been done prior doing the experimental works and studies. It presents several fundamental theories that are closely related to the research and experimental works described in this thesis.

2.1 Principal Component Analysis

Principal component analysis (PCA) is a statistical technique that has been widely used in image and video processing (Kim, 2002; Liu *et al.*, 2010; Lu *et al.*, 2008; Pyatykh *et al.*, 2013; Vargas *et al.*, 2011; Zhang *et al.*, 2006; Zhang *et al.*, 2010). It has also, particularly found application in biomedical fields, including pulse detection (Balakrishnan *et al.*, 2013; Martis *et al.*, 2012; Martis *et al.*, 2012; Sharma *et al.*, 2012). PCA is a way of identifying the patterns in a group of high dimensional data and expressing or analyzing the data by highlighting their similarities and differences (Lindsay, 2002).

To utilize PCA, an important assumption has to be made, i.e. linearity (John, 2002). In other words, a new set of data can be formed as a linear combination of its basis vectors. Let \mathbf{A} be the original data set, \mathbf{B} be the representation of \mathbf{A} and \mathbf{T} be the linear transformation matrix that transforms \mathbf{A} into \mathbf{B} , then

$$\mathbf{B} = \mathbf{T}\mathbf{A} \quad (2.1)$$

Geometrically, \mathbf{T} is a rotation and a translation matrix which transforms \mathbf{A} into \mathbf{B} . Considering both \mathbf{A} and \mathbf{B} are a $m \times n$ matrix, then the covariance matrix of \mathbf{A} , \mathbf{C}_A can be defined as:

$$\mathbf{C}_A = \frac{1}{n-1} \mathbf{A}\mathbf{A}^T \quad (2.2)$$

where \mathbf{A}^T is the transpose matrix of \mathbf{A} . Similarly, the covariance matrix of \mathbf{B} , \mathbf{C}_B can be defined as:

$$\mathbf{C}_B = \frac{1}{n-1} \mathbf{B}\mathbf{B}^T \quad (2.3)$$

Equation (2.3) can be rewritten in terms of \mathbf{T} .

$$\mathbf{C}_B = \frac{1}{n-1} \mathbf{T}\mathbf{A}(\mathbf{T}\mathbf{A})^T \quad (2.4)$$

$$\mathbf{C}_B = \frac{1}{n-1} \mathbf{T}\mathbf{A}\mathbf{A}^T\mathbf{T}^T \quad (2.5)$$

$$\mathbf{C}_B = \frac{1}{n-1} \mathbf{T}(\mathbf{A}\mathbf{A}^T)\mathbf{T}^T \quad (2.6)$$

$$\mathbf{C}_B = \frac{1}{n-1} \mathbf{T}\mathbf{S}\mathbf{T}^T \quad (2.7)$$

where \mathbf{S} is a symmetric matrix.

A symmetric matrix can be diagonalized by an orthogonal matrix of its eigenvectors.

Therefore, symmetric matrix \mathbf{S} can be rewritten as:

$$\mathbf{S} = \mathbf{E}\mathbf{D}\mathbf{E}^T \quad (2.8)$$

where \mathbf{D} is a diagonal matrix and \mathbf{E} is a matrix of eigenvectors of \mathbf{S} . The transformation matrix \mathbf{T} is selected as a matrix where each row \mathbf{T}_i is an eigenvector of $\mathbf{A}\mathbf{A}^T$. Then,

$$\mathbf{T} = \mathbf{E}^T \quad (2.9)$$

By substituting (2.9) into (2.8), the symmetrical matrix \mathbf{S} can be expressed as

$$\mathbf{S} = \mathbf{T}^T\mathbf{D}\mathbf{T} \quad (2.10)$$

The covariance matrix of \mathbf{C}_B can be redefined in terms of diagonal matrix \mathbf{D} . By substituting (2.10) into (2.7),

$$\mathbf{C}_B = \frac{1}{n-1} \mathbf{T}(\mathbf{T}^T \mathbf{D} \mathbf{T}) \mathbf{T}^T \quad (2.11)$$

$$\mathbf{C}_B = \frac{1}{n-1} (\mathbf{T} \mathbf{T}^T) \mathbf{D} (\mathbf{T} \mathbf{T}^T) \quad (2.12)$$

Since,

$$\mathbf{T}^{-1} = \mathbf{T}^T \quad (2.13)$$

then (2.12) can be rewritten as

$$\mathbf{C}_B = \frac{1}{n-1} (\mathbf{T} \mathbf{T}^{-1}) \mathbf{D} (\mathbf{T} \mathbf{T}^{-1}) \quad (2.14)$$

$$\mathbf{C}_B = \frac{1}{n-1} \mathbf{D} \quad (2.15)$$

Therefore, \mathbf{T} diagonalizes \mathbf{C}_B . The principal components of \mathbf{A} are the eigenvectors of $\mathbf{A} \mathbf{A}^T$ and represented by the rows of \mathbf{T} . The j^{th} diagonal value of \mathbf{C}_B is the variance of \mathbf{A} along \mathbf{t}_j .

The general mathematic model of PCA was presented earlier. Currently, several enhancement and extension work of PCA have been developed for different applications in image and video processing, including Fast PCA (Mittal, 2008; Woo *et al*, 2013), weighted PCA (Chang and Yeung, 2006; Wang & Wu, 2005) and sparse PCA (Naikal, 2011).

2.2 Independent Component Analysis

Blind source separation (BSS) is used to uncover the independent signals from a set of sensor observations that are linear mixtures of statistical independent sources. Both the source signals and the ways how the signals were mixed (i.e. the sensors) were unknown. Independent Component Analysis (ICA) is a method to solve the blind source separation problem. In contrast to correlation-based transformation such as PCA, ICA does not only decorrelate the signals, but also reduces higher-order statistical dependencies (Lee *et al.*, 2000). Hence, ICA is widely used in several applications related to image and video processing (Bae *et al.*, 2003; Bartlett *et al.*, 2002; Déniz *et al.*, 2001; Hoyer & Hyvarinen, 2000; Liu & Wechsler, 2003; Wang & Chang, 2006; Yuen & Lai, 2002; Zhang & Chen, 2006). Particularly, ICA has been increasing its popularity in the field of biomedical signal processing (Beckmann & Smith, 2004; Calhoun & Adali, 2006; James & Hesse, 2004; Martis *et al.*, 2013; Salimi-Khorshidi *et al.*, 2014; Vázquez *et al.*, 2012).

Assume that there are n linear mixtures (sensors) y_1, \dots, y_n of n independent components

$$y_j = m_{j1}c_1 + m_{j2}c_2 + \dots + m_{jn}c_n, \text{ for all } j, \quad (2.16)$$

and each mixture y_j as well as the independent component c_k is a random variable, instead of a proper time signal. Let \mathbf{y} denotes the mixture y_1, \dots, y_n , \mathbf{s} denotes c_1, \dots, c_n , and \mathbf{M} denotes the m_{ij} , then (2.16) can be written as

$$\mathbf{y} = \mathbf{M}\mathbf{s}. \quad (2.17)$$

The statistical model in (2.17) is known as independent component analysis. It describes how the observed sensors y_i are generated by a process of mixing the components s_i (Hyvärinen & Oja, 2000). The mixing matrix \mathbf{M} is unknown but can be

estimated. The independent components can be obtained by computing inverse of mixing matrix \mathbf{M} , denoted by \mathbf{W} . Hence,

$$\mathbf{c} = \mathbf{W}\mathbf{y} \quad (2.18)$$

Two important conditions to use an ICA model are:

- i. The sources must have non-Gaussian distribution,
- ii. The sources are independent to each other.

To estimate the mixing matrix \mathbf{M} , the independent sources must be non-Gaussian. Consider a source \mathbf{c} consists of two independent components that have uniform pdf. The joint density of both components c_1 and c_2 is hence uniform on a square distribution. If these two independent components are mixed through a mixing matrix \mathbf{M} , then the new mixed components (sensor \mathbf{y}) will have a uniform distribution on a parallelogram. The columns of \mathbf{M} represent the direction of the edges of the parallelogram. Hence \mathbf{M} can be estimated and the ICA model can be used. However, if the source \mathbf{c} consists of two Gaussian independent components, then the joint density of both components c_1 and c_2 is completely symmetric (a circular distribution). Therefore, it does not indicate the directions of the columns of the mixing matrix \mathbf{M} . \mathbf{M} cannot be estimated. However, if at least one of the independent signals has non-Gaussian distribution, then ICA can be used.

The independence can be defined by the probability densities. Let $p(c_1, \dots, c_n)$ denotes the joint probability density of c_1, \dots, c_n , and $p_j(c_j)$ denotes the marginal *PDF* of c_j , then the elements in the \mathbf{s} are independent to each other if and only if the joint pdf is

$$p(c_1, \dots, c_n) = p_1(c_1) p_2(c_2) \dots p_n(c_n) \quad (2.19)$$

The way how ICA can recover the original sources from a set of sensors is illustrated in Figure 2.1, Figure 2.2 and Figure 2.3. Figure 2.1 shows two original signals (sources) that are independent to each other. Figure 2.2 shows two new signals (sensors) that are mixture of the two sources shown in Figure 2.1. By using ICA, the original signals can be recovered, as shown in Figure 2.3.

For the video-based heart rate measurements and monitoring, the blood volume pulse (BVP) is the independent source signal of interest. The color components of the facial images captured by the video recorder, i.e. red, green, and blue (RGB), vary in accordance to the heart rate variation, since the changes in blood volume alter the light intensity reflected from facial tissue. Each of the RGB components is actually the sensor or mixture of the reflected plethysmographic signals and other sources (as well as the artifacts).

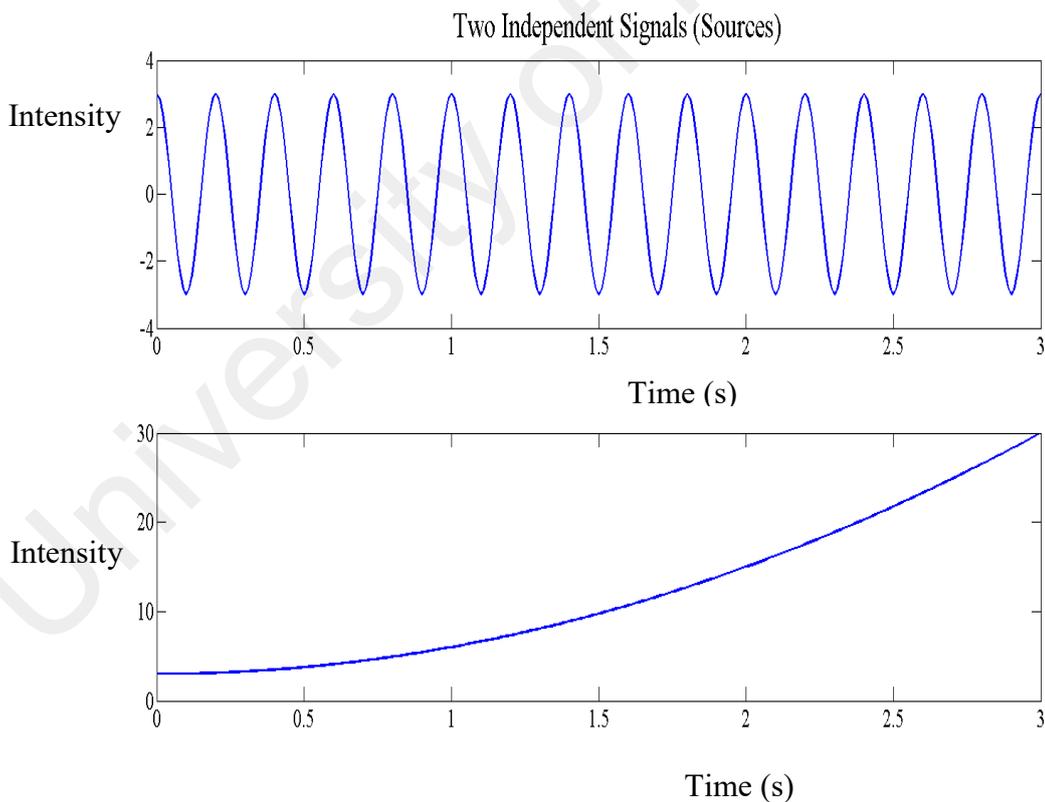


Figure 2.1: Two independent signals (sources)

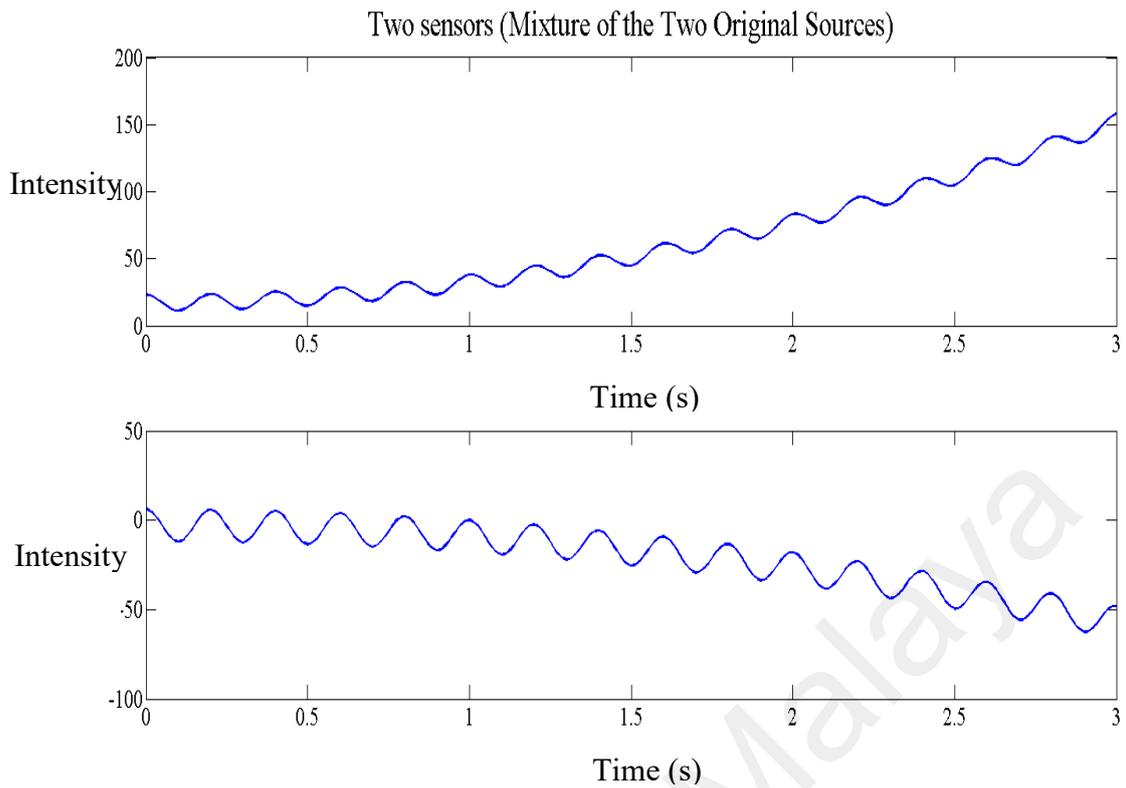


Figure 2.2: Two sensors (mixture of the two original sources)

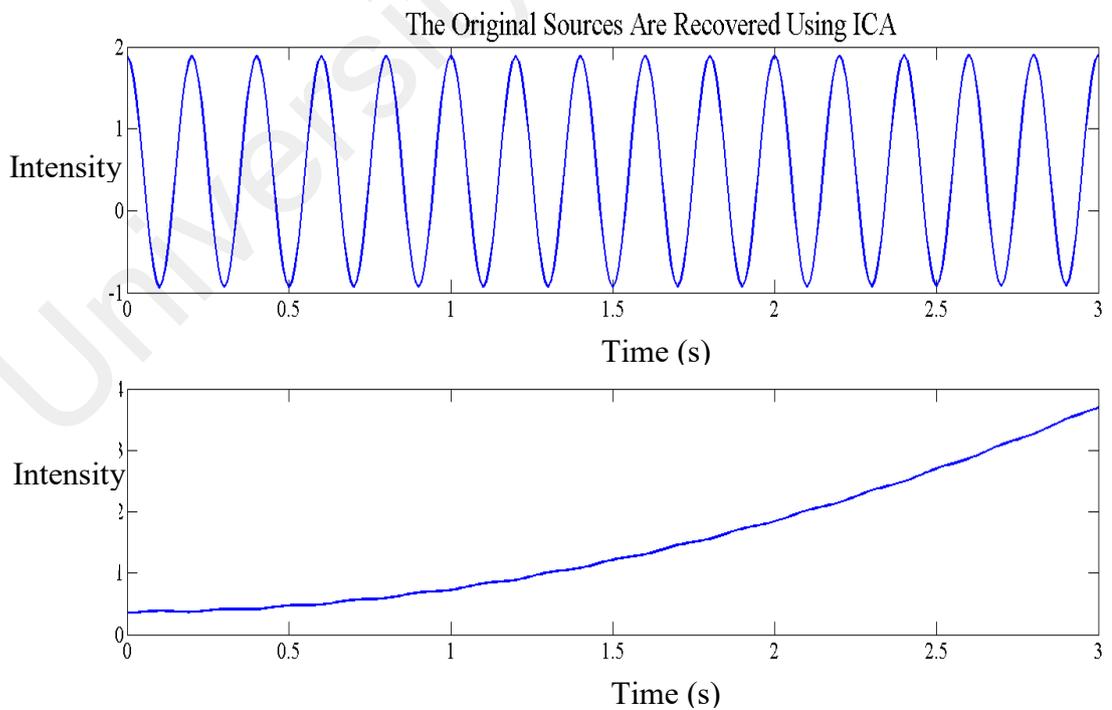


Figure 2.3: The separation of original sources using ICA

2.3 Simplified Mathematical Model for Images of Human Skin

Skin color is related to pigmentation. Melanin in the epidermis and hemoglobin in the dermis are the important factors that affect the variations in skin color (Dawson *et al.*, 1980; Tsumura *et al.*, 2003; Xu, 2008). Xu *et al.* defines the skin absorbance S at wavelength λ as

$$S(\lambda) = c_m(\lambda)p_m + c_h(\lambda)p_h + S_0(\lambda) \quad (2.20)$$

where p_m and p_h represent the pigment concentration for melanin and hemoglobin respectively, c is the product of pigment extinction coefficient and the mean path length of photons in the skin layer, and S_0 is the baseline skin absorbance (Xu *et al.*, 2014).

The absorbance can be interpreted as:

$$S = -\log\left(\frac{T}{L}\right) \quad (2.21)$$

where T and L are the power of the transmitted light and incident light respectively. The pixel intensities corresponding to skin image, I , are expressed as:

$$I = k \int T(\lambda)H(\lambda)d\lambda \quad (2.22)$$

where $H(\lambda)$ is the spectral response function for the camera sensor and k is the camera gain. The spectral response function $H(\lambda)$ can be treated as a delta function (Finlayson *et al.*, 2004; Tsumura *et al.*, 2003). Since (2.20) is equal to (2.21), then (2.22) can be rewritten as:

$$\log I_R = \log kT(R) - c_m(R)p_m - c_h(R)p_h - S_0(R) \quad (2.23)$$

$$\log I_G = \log kT(G) - c_m(G)p_m - c_h(G)p_h - S_0(G) \quad (2.24)$$

$$\log I_B = \log kT(B) - c_m(B)p_m - c_h(B)p_h - S_0(B) \quad (2.25)$$

where R , G and B represent the red, green and blue components of the image.

By defining Q as the pixel channel quotient in log space, Q can be expressed as

$$Q = \log\left(\frac{I_R}{I_G}\right) \quad (2.26)$$

At the n -th frame, the differences between the current and previous frames, ΔQ^n is given as

$$\Delta Q^n = Q^n - Q^{n-1} \quad (2.27)$$

By considering T and p_m as the dc components, then (2.23) and (2.24) can be written as

$$\log\left(\frac{I_R^n \bullet I_G^{n-1}}{I_R^{n-1} \bullet I_G^n}\right) = \Delta \log\left(\frac{T(R)}{T(G)}\right) - \Delta c_h \Delta p_h \quad (2.28)$$

Equation (2.28) can be used to represent the model for skin color in an image. To estimate the heart rate from a video sequence, a time series signal $x(n)$ has to be obtained and considered as the input data for subsequent processes, where $x(n)$ can be expressed as

$$x(n) = [\Delta Q^2, \Delta Q^3, \dots, \Delta Q^n] \quad (2.29)$$

CHAPTER 3: VIDEO BASED HEART RATE ESTIMATION USING TEMPORAL INFORMATION

3.1 Overview

Previous works focus more on the heart rate with less variation and extract the information from spatial domain only. Consider a subject is exercising and hence his/her heart rate is changing rapidly in accordance to the intensity and duration of his/her workout, then in this case the temporal information is needed. In this chapter, two different approaches are presented to indicate the heart rate of the subjects. The choice of using STFT and filter bank is their ability to provide more accurately localized temporal and frequency information, especially for the rapidly changing heart rate pattern during the exercise routine. Two experiments are carried out to validate the proposed approaches. A camcorder is used to capture the facial images of seven subjects, whose heart rate vary dynamically, between 87 and 151 BPM. The first experiment involves the measurement of subjects' increasing heart rates while cycling whereas the second experiment involves falling heart beats. Experimental results show the proposed method can provide an acceptable result where the root mean square error is less than 4.0 BPM.

3.2 Proposed Method

This section presents the way how STFT and filter bank can be used to estimate human heart rates that change dynamically. Before the STFT and filter bank are applied, the region of interest (ROI) needs to be determined. In the experiments, the area between the eyes and the upper lip of the mouth of a subject in a video frame was chosen as the ROI. The ROI of each frame for the three RGB components was extracted. For each experiment, a sixty-second video was recorded for each subject.

3.2.1 Extraction of Temporal Information Using Short-time Fourier Transform

The process began by obtaining the mean of all pixel values for each RGB color component where

μ_R : the mean of all pixel values for red component

μ_G : the mean of all pixel values for green component

μ_B : the mean of all pixel values for blue component

This was repeated for every frame obtained from each video clip. The set of RGB components were then detrended using algorithm developed by Tarvainen *et al.* (Tarvainen *et al.*, 2002). Independent component analysis was then used to separate these detrended data into their different sources. The source with the highest spike of power spectrum was selected to be used as the time-series input data for the short-time Fourier transform (STFT).

In this study, the window size of the STFT for each video is 20 seconds or 1000 frames. 950 samples or frames overlap each other for two adjacent windows. Therefore, for a sixty-second video, 40 instantaneous heart rate readings can be calculated. For any STFT window, the highest peak of the frequency components was selected as the instantaneous heart rate for that instant. Figure 3.1 summarizes the techniques used in the proposed method.

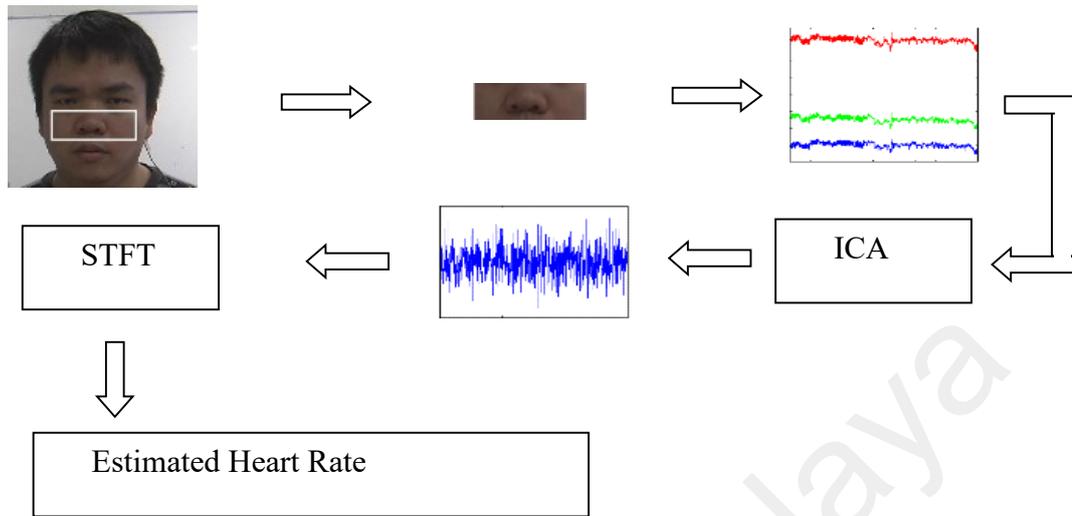


Figure 3.1: Proposed method for video-based heart rate estimation using STFT

3.2.2 Extraction of Temporal Information Using Short-time Fourier Transform

In this study, the Red, Green, and Blue (RGB) components of each ROI-cropped image were extracted. The mean of all pixel values for each component where

μ_R : the mean of all pixel values for red component

μ_G : the mean of all pixel values for green component

μ_B : the mean of all pixel values for blue component

This was repeated for every frame obtained from each video clip. The thirty-second length of times series data were used as the raw input signals for the subsequent data processing and analysis. BSS was utilized to extract the separated independent sources from the input. The Joint Approximate Diagonalization of Eigenmatrices (JADE), a member of independent component analysis (ICA) algorithm (Cardosa & Souloumiac, 1993; Cardosa, 1999) was utilized to obtain the heart rate source signals. One of the requirements of using ICA is the source itself must have a non-Gaussian distribution. It

means the heart rate signals must have a non-Gaussian distribution. Through empirical studies, it was found that actual heart rate distributions corresponding to each video clip used in the experiment are non-Gaussian. Hence 60-s of video is chosen as input to the ICA.

The ICA source signal with the highest spike of power spectrum was chosen as the best source representing the heart rate signals. This source signal was then analyzed using the filter bank. The processes were repeated for all five parts of the video. Figure 3.2 shows the flow chart of the proposed method.

Filter bank is an array of band-pass filters that separates the input signal into multiple components. Each component carries a specific frequency sub-band of the input signal. In this study, filter bank was applied to bandpass the source signal from 0.8 Hz (or 48 BPM) to 4 Hz (or 240 BPM). The difference of two adjacent frequency sub-bands was set to 0.02 Hz (or 1.2 BPM) equally. Throughout the entire 60-second length of source signals, the heart rate readings were sampled at every 1 second (or 50 points). At n -th second, the filter bank was applied to the corresponding point and its neighbouring points. The number of neighbouring points was fixed at 500 points or 10 seconds. Among the multiple components, the component with highest energy was selected and the frequency sub-band it carried was the instantaneous heart rate at n -th second. Figure 3.3 indicates how the filter bank was utilized in the proposed method.

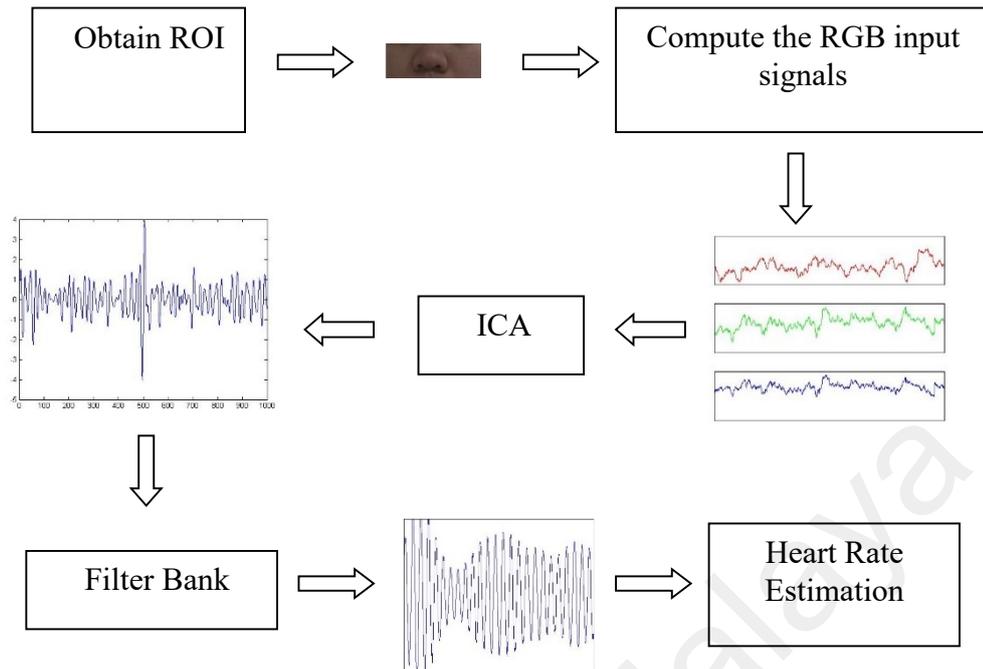


Figure 3.2: Flow chart of video-based heart rate estimation using filter bank

3.3 Experimental Study

This section outlines the experimental setup and the experimental results and analysis for both approaches. The advantage of using temporal information in estimating human heart rate that changes dynamically is described. A comparison on both approaches is also presented. At last, the drawback of using both approaches is discussed.

3.3.1 Experimental Setup

The experiments were set up under office fluorescent lights with indirect sunlight as the source of illumination. A video camera (24-Bit RGB, 8 bits per channel) with a resolution of 1440 x 1080 pixels and 50 frames per second was used in recording a subject cycling for several minutes. Figure 3.4 shows a subject cycling during the experiment. The subject was seated at a distance of about 60 cm from the camera. Two experiments were carried out. In the first experiment, seven subjects began cycling and the heart rate started to increase.

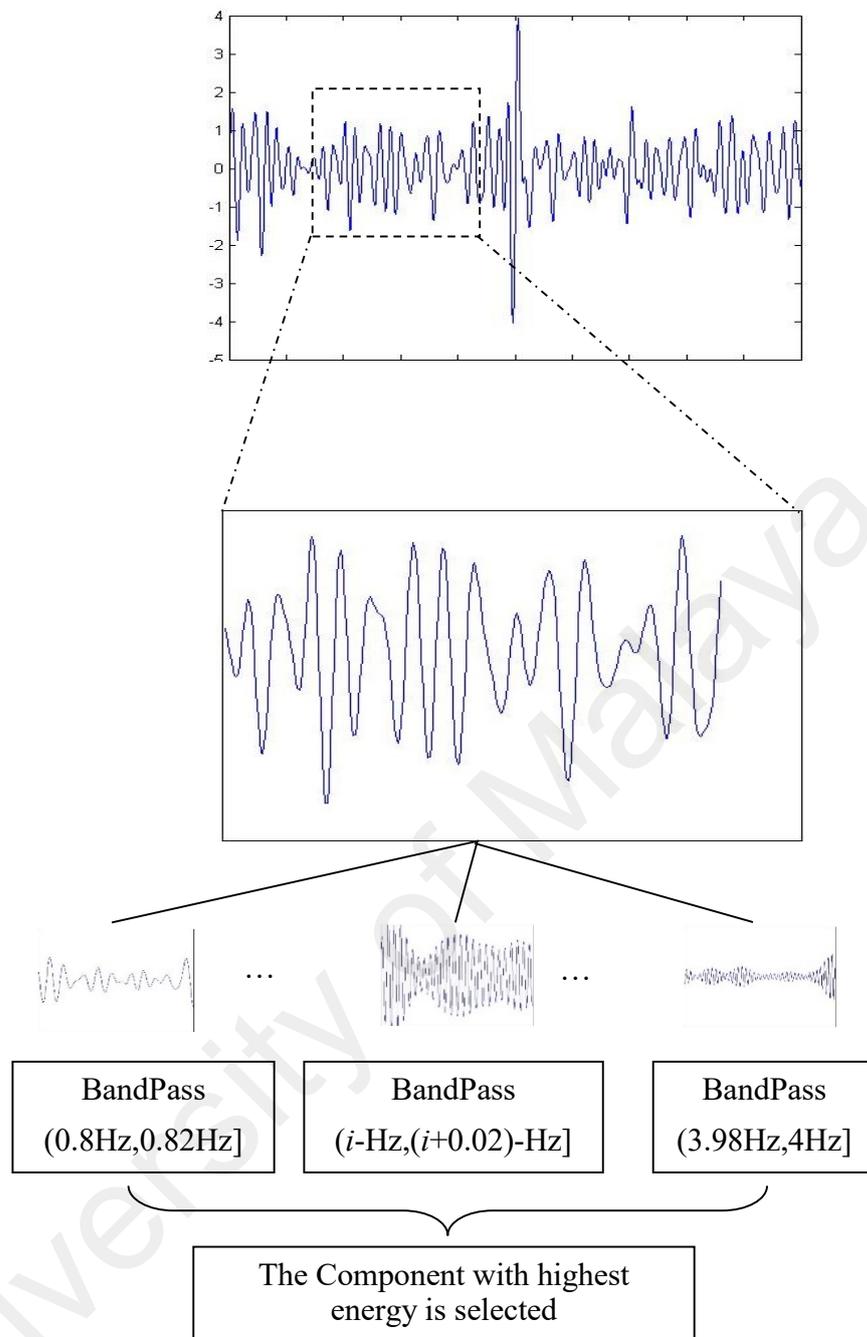


Figure 3.3: Proposed method for video-based heart rate estimation using filter bank

Videos were captured continuously for 60 seconds while the subjects were cycling. In the second experiment, seven subjects continued to increase the speed and stopped at a certain speed. The heart rates of subjects are expected to drop from higher level to a more converging and lower level. Therefore, the gradient of heart rate variation is larger and localized temporal information is needed. Similarly, videos were captured for 60 seconds

while the subjects taking rest after cycling. The data obtained from the experiments were then analyzed using the proposed approaches.

The results of the proposed methods were compared to the heart rate readings obtained from Polar Heart Rate Monitor since it is the one of the most accurate instantaneous heart rate measurement devices at this moment (Schönfelder *et al.*, 2011; Wallén *et al.*, 2012). For reference, all instantaneous heart rates of the subjects were measured using Polar Heart Rate Monitor – Polar Team² Pro. Polar Team² transmitter set records and transmits the subjects' ECG signals to its base station. The heart rate is sampled and computed by measuring at least one ECG signal waveform, as described in their patents (Heikkila, 1998; Pietila, 1997). A comparative study of the actual readings obtained from Polar Team² Pro and the computed readings from the proposed method was done.

University of Malaya



Figure 3.4: Subject cycling during the experiment

3.3.2 Experimental Results and Analysis for STFT Approach

Figure 3.5 shows the comparison between all estimated and actual heart rate readings for the first experiment using STFT approach while Figure 3.6 shows the corresponding results for the second experiment using the same approach. Overall, the results obtained from the proposed method did not vary much from the actual readings. The respective root mean square errors are 3.36 BPM (first experiment) and 2.41 BPM (second experiment). The respective correlation coefficients between actual and estimated heart rate readings are 0.99 for both experiments.

As can be observed from Figure 3.5, the readings are generally distributed below the reference line ($R = 1$). Figure 3.6 shows another trend where the data are generally distributed above the reference line ($R=1$). As a 20-second window is needed for the computation of the first reading, the estimated readings are always lagging behind the actual readings. For further discussion, the comparison of actual and estimated heart rate for each subject is presented. Figure 3.7 shows the comparison of actual and estimated rate for each subject in the first experiment while Figure 3.8 shows the comparison for the second experiment. For both figures, red color represents the actual heart rate readings while green color represents the estimated heart rate readings using proposed method. Both figures show that the estimated readings are generally lagging behind the actual readings and cause the inaccuracy. To address this issue, a short-duration video is needed for heart rate estimation. The challenge of using a short-duration video is the ICA sources may not have sufficient independence and it could render an inaccurate result. Further details are presented in Chapter 4.

In addition to this, the performance of the proposed algorithm is also evaluated using the Bland Altman plots. Bland Altman plot can be used to analyze the agreement between two different set of measurements, i.e the estimated heart rate readings from proposed method and the actual heart rate readings obtained from Polar Heart Rate Monitor. The 95 % limit of agreement for each comparison indicates that how far apart measurements by two methods are more likely to be. Figure 3.9 shows the Bland Altman plot for the heart rate readings obtained from the first experiment while Figure 3.10 shows the Bland Altman plot for the heart rate readings obtained from the second experiment. It shows that the heart rate readings obtained from the first experiment has higher range of errors as compared to the heart rate readings obtained from the second experiment. It is due to the occurrence of large amount of motion artifacts in the first experiment. As in the first

experiment, video was captured while subject was cycling. A larger amount of motion artifacts might occur.

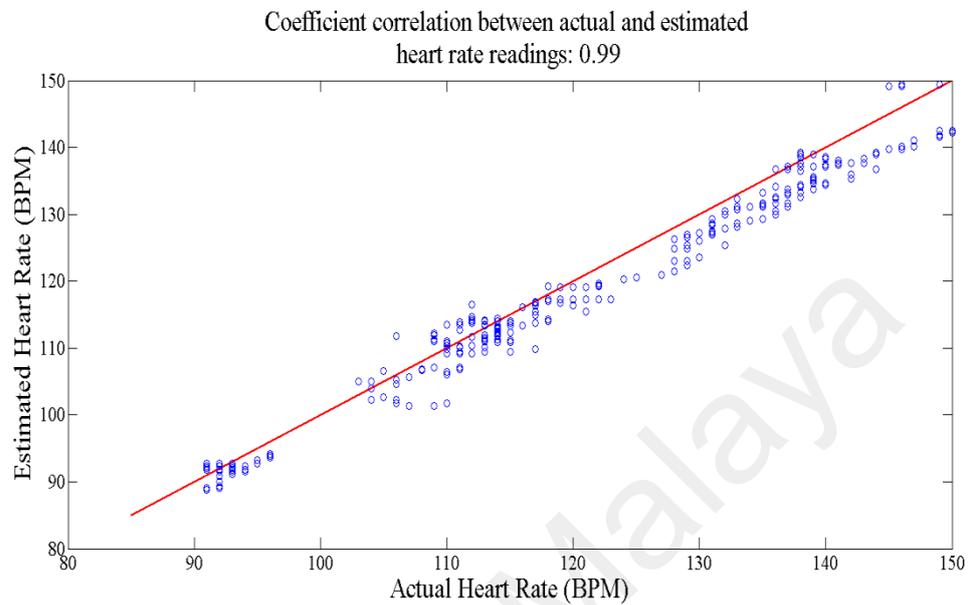


Figure 3.5: Comparison of all actual and estimated heart rate readings using STFT for the first experiment

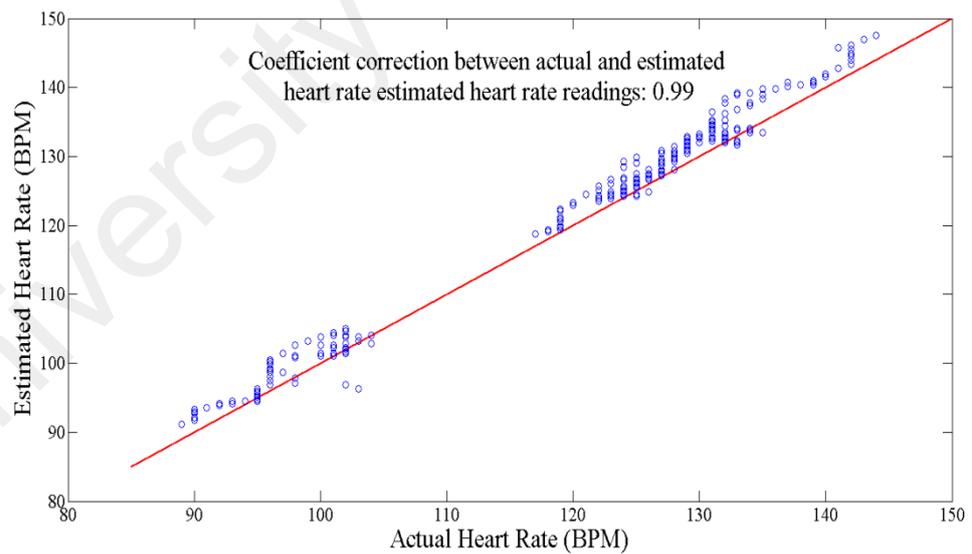
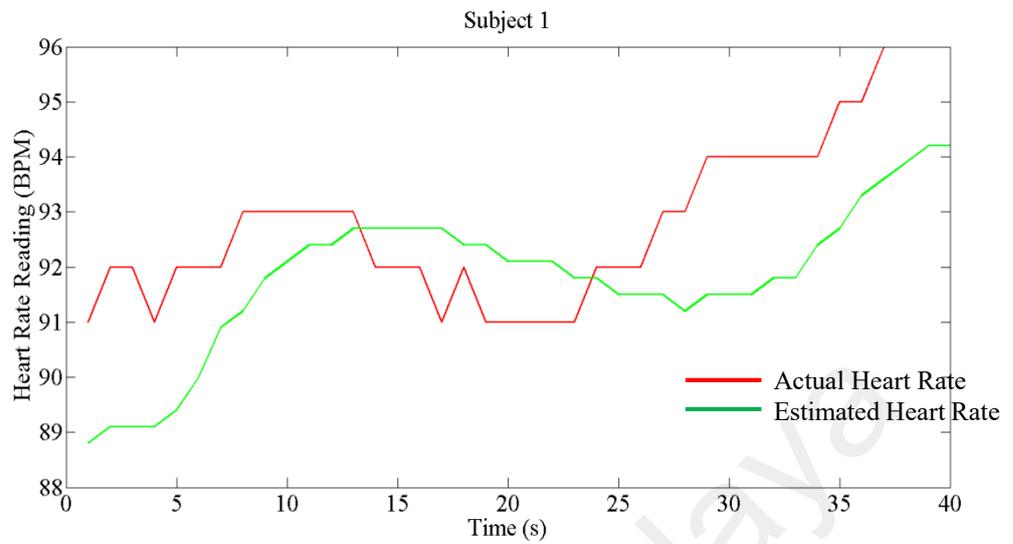
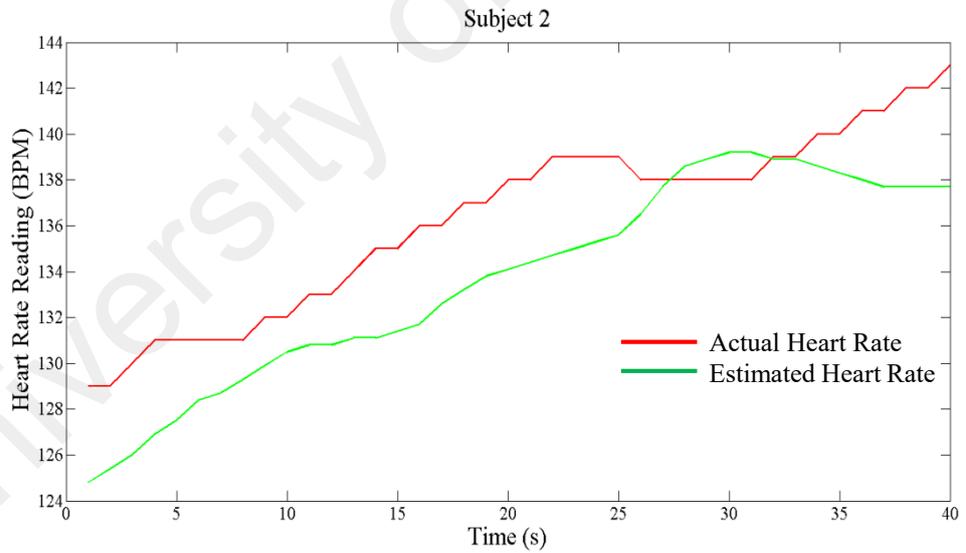


Figure 3.6: Comparison of all actual and estimated heart rate readings using STFT for the second experiment

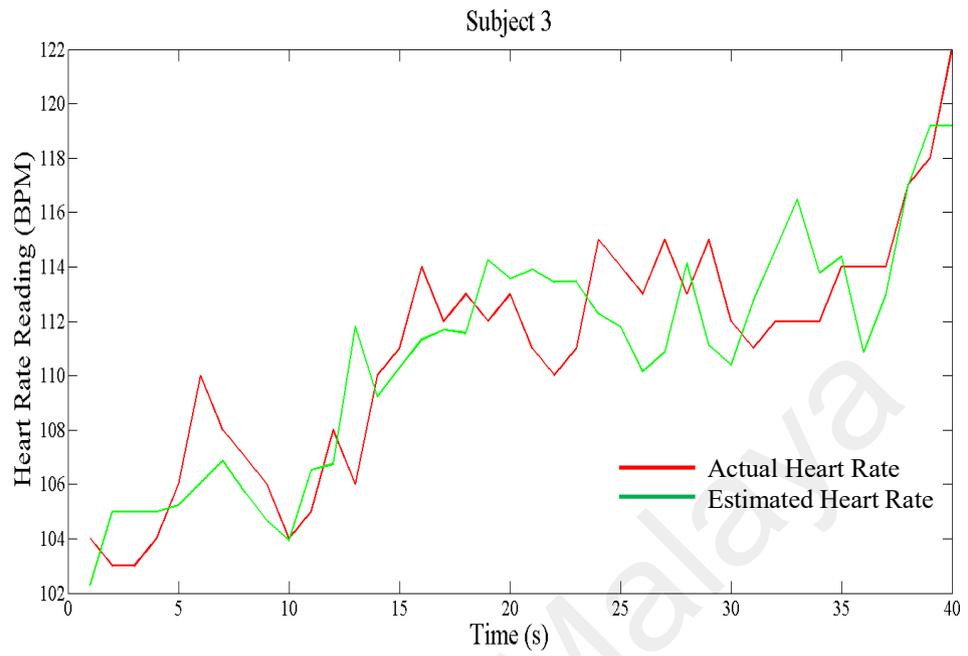


(a) First subject

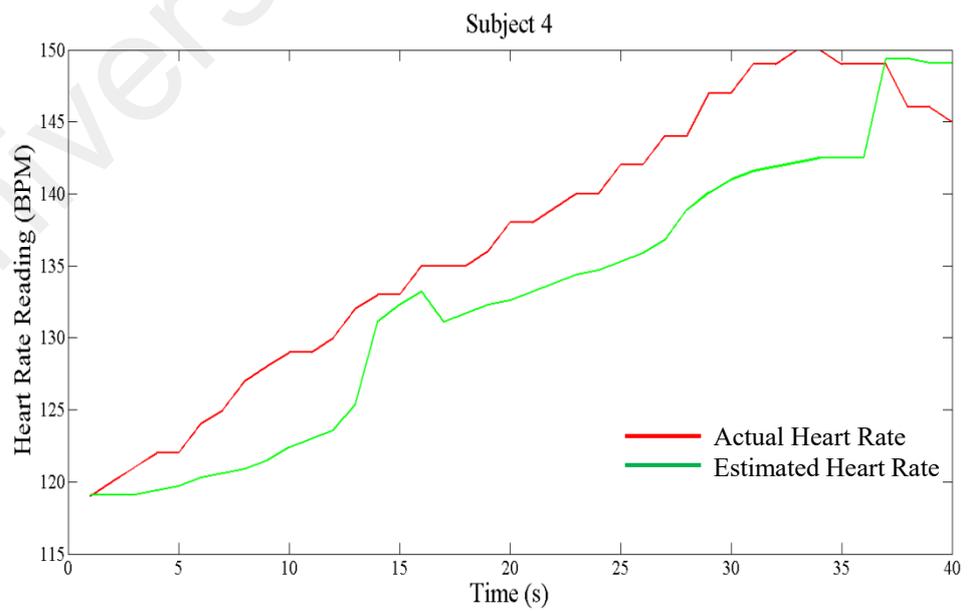


(b) Second subject

Figure 3.7: Comparison of actual and estimated heart rate readings for each subject in the first experiment

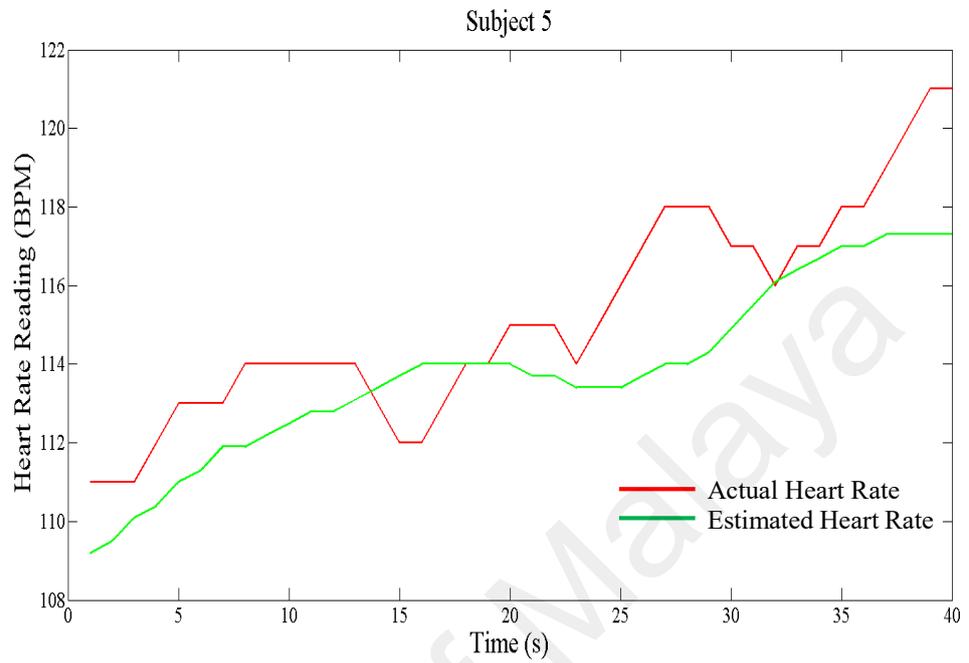


(c) Third subject

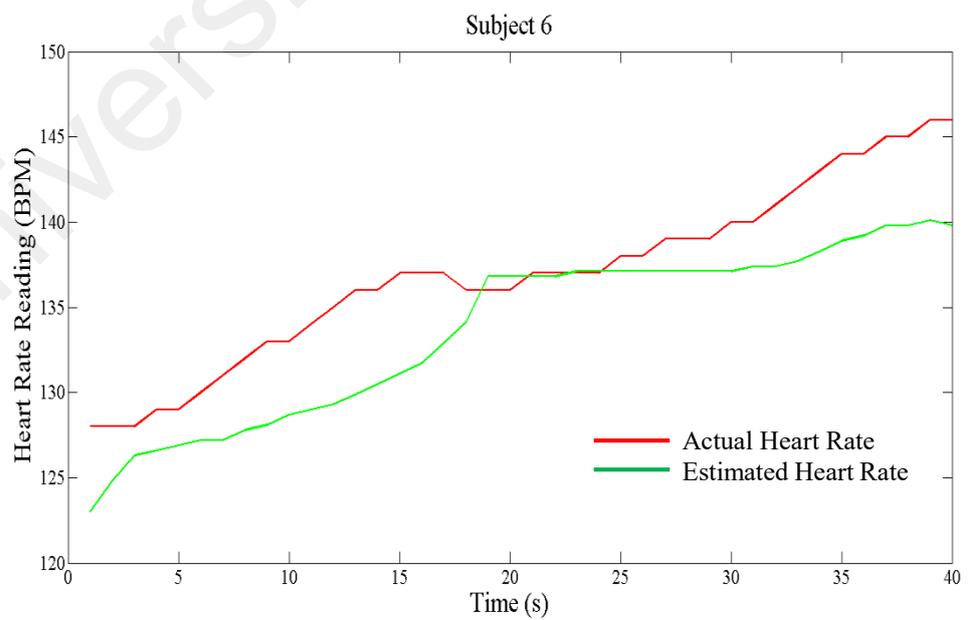


(d) Forth subject

Figure 3.7.continued

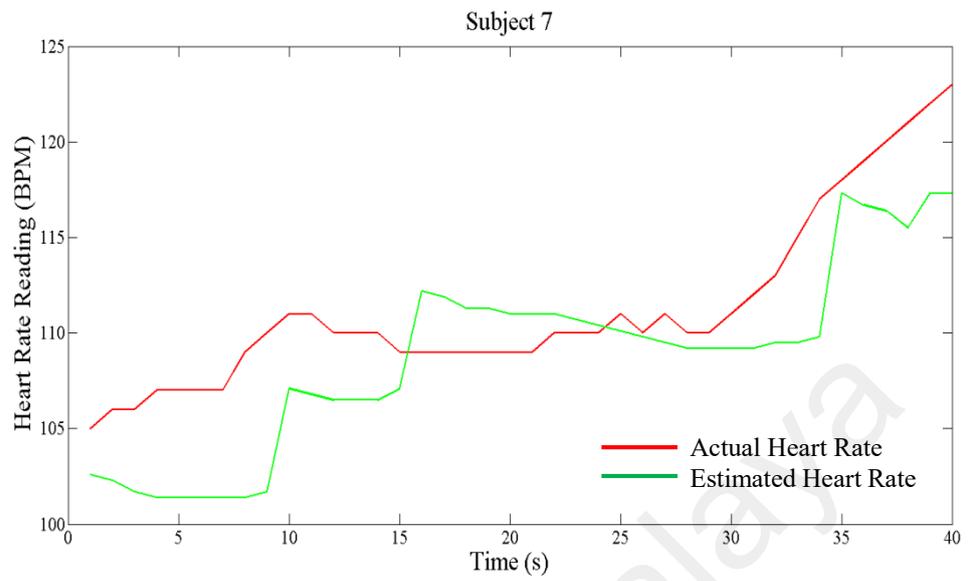


(e) Fifth subject



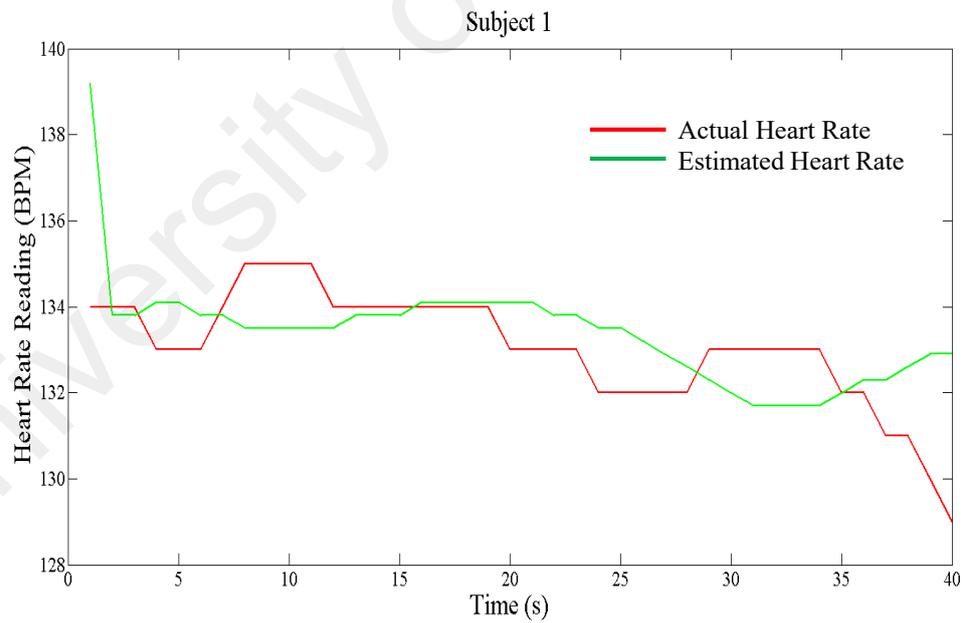
(f) Sixth subject

Figure 3.7.continued



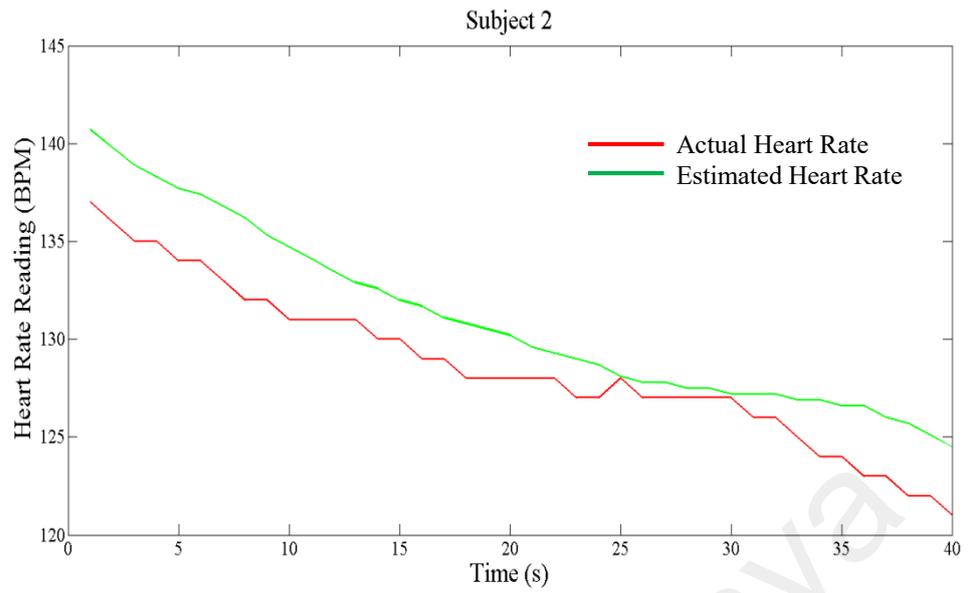
(g) Seventh subject

Figure 3.7.continued

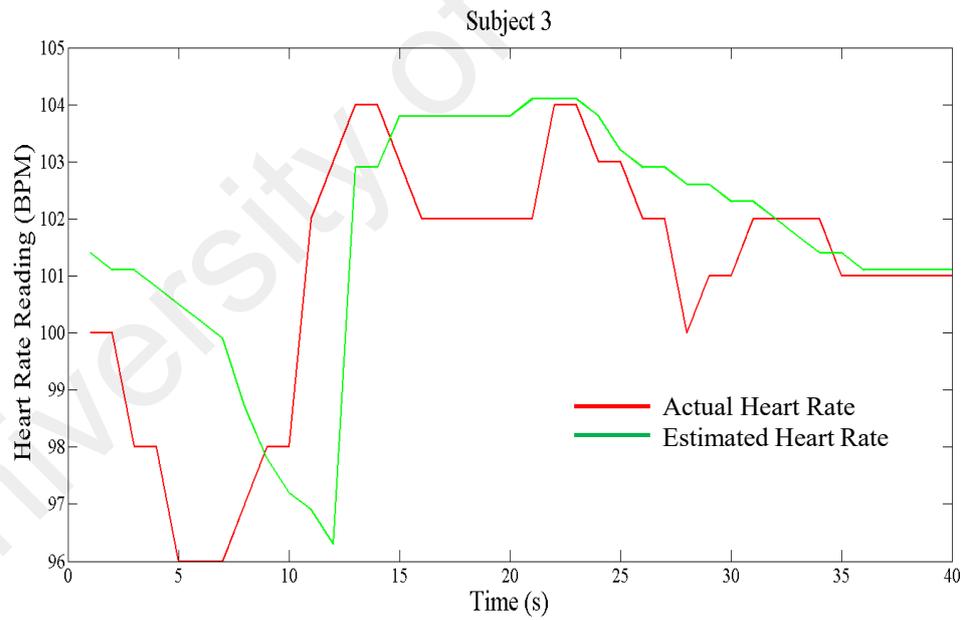


(a) First subject

Figure 3.8: Comparison of actual and estimated heart rate readings for each subject in the second experiment

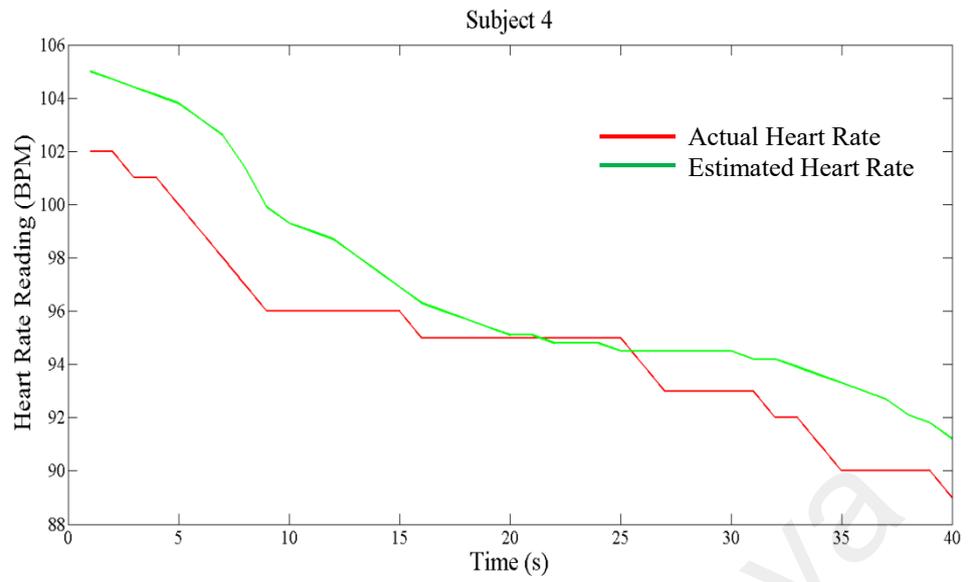


(b) Second subject

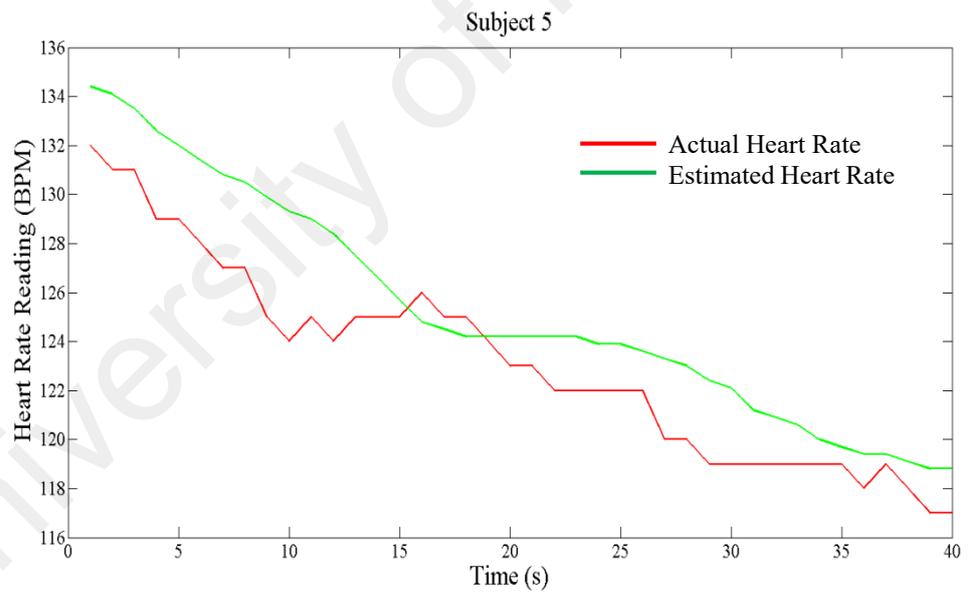


(c) Third subject

Figure 3.8.continued

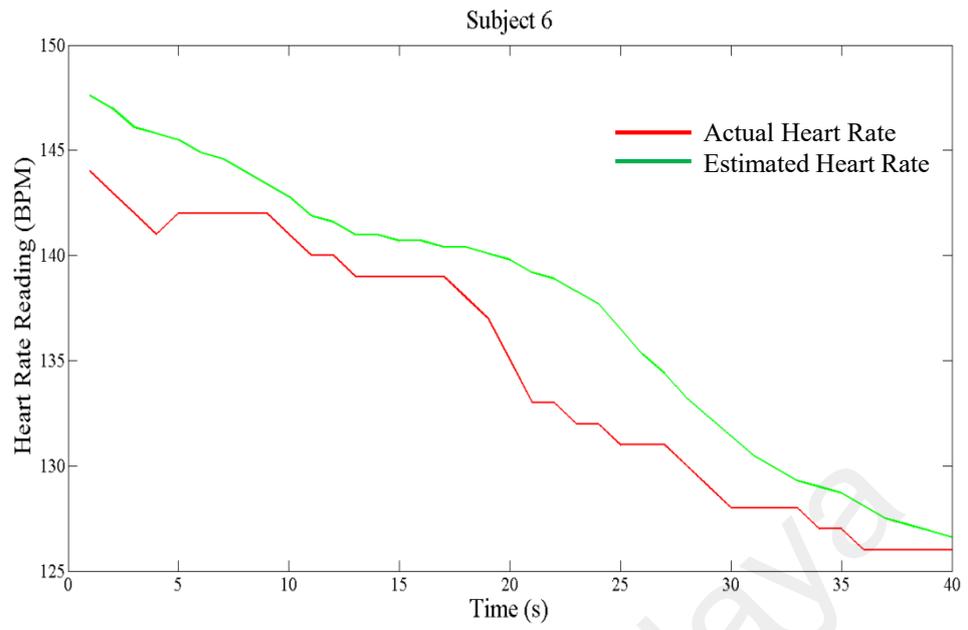


(d) Forth subject

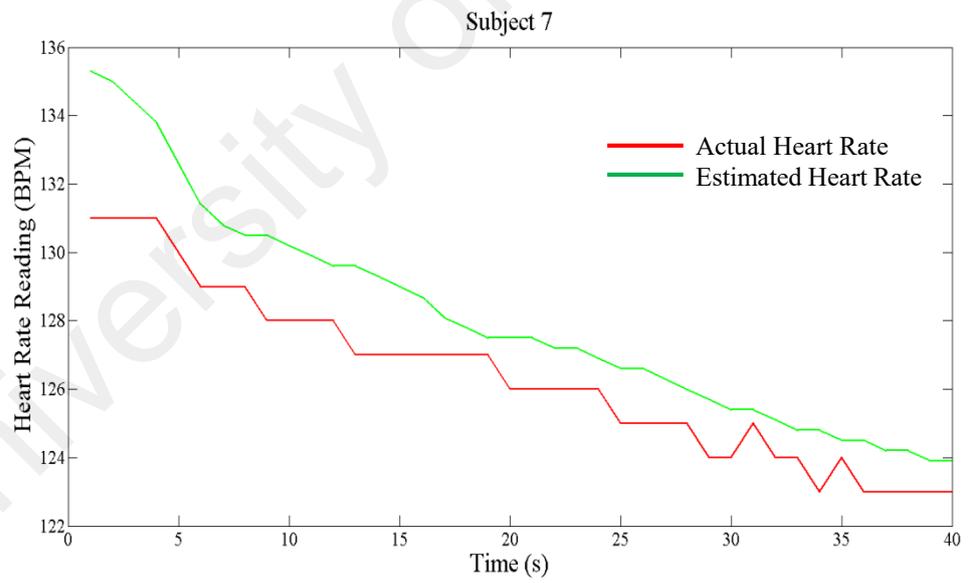


(e) Fifth subject

Figure 3.8.continued



(f) Sixth subject



(g) Seventh subject

Figure 3.8.continued

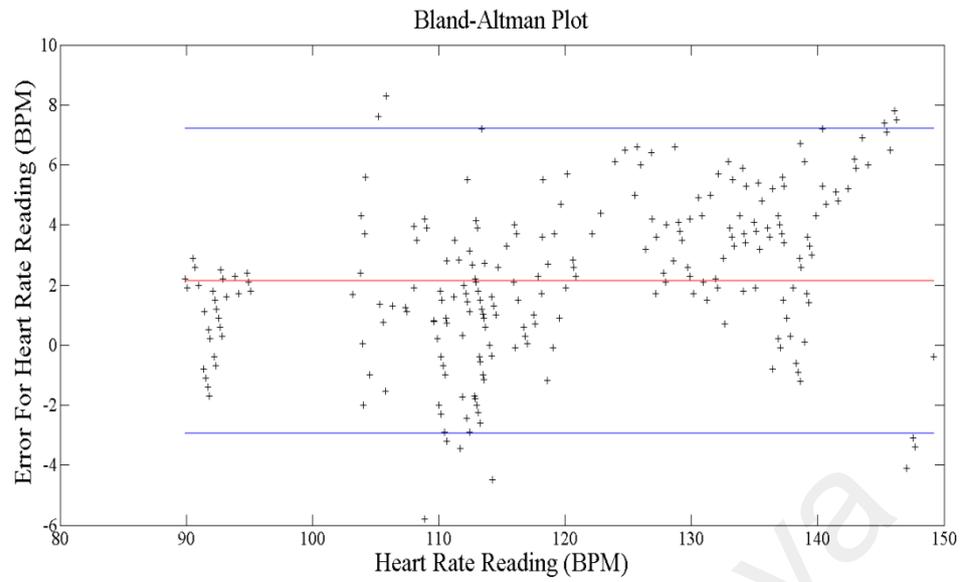


Figure 3.9: Bland-Altman plot for all estimated heart rate reading using STFT for the first experiment

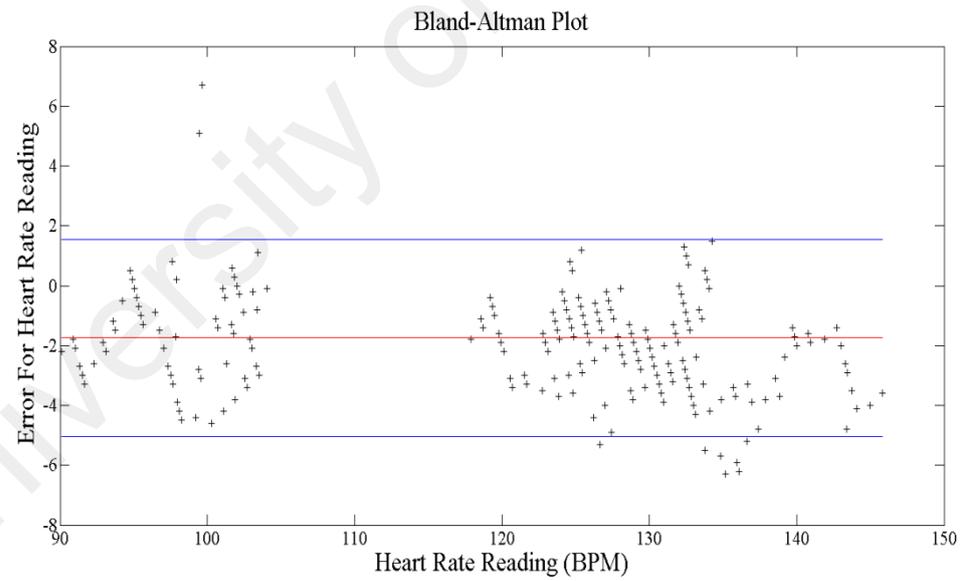


Figure 3.10: Bland-Altman plot for all estimated heart rate reading using STFT for the second experiment

In this study, the window size of the STFT was set to 1000 points, which corresponds to 20 seconds (as the frame rate is 50). The power spectrum distribution obtained from STFT is showed in Figure 3.11. On the other hand, the proposed method is compared with previous method (Pursche *et al.*, 2012) that used Fourier transform to obtain the heart rate. It is found that STFT provides better temporal information and hence better results. Figure 3.12 shows the heart rate readings obtained from STFT and Fourier transform for a particular subject. Blue line indicates the results obtained using Fourier transform. It shows a fixed value for a given time and does not show the time localization. Hence it is not suitable to measure dynamic heart rate readings.

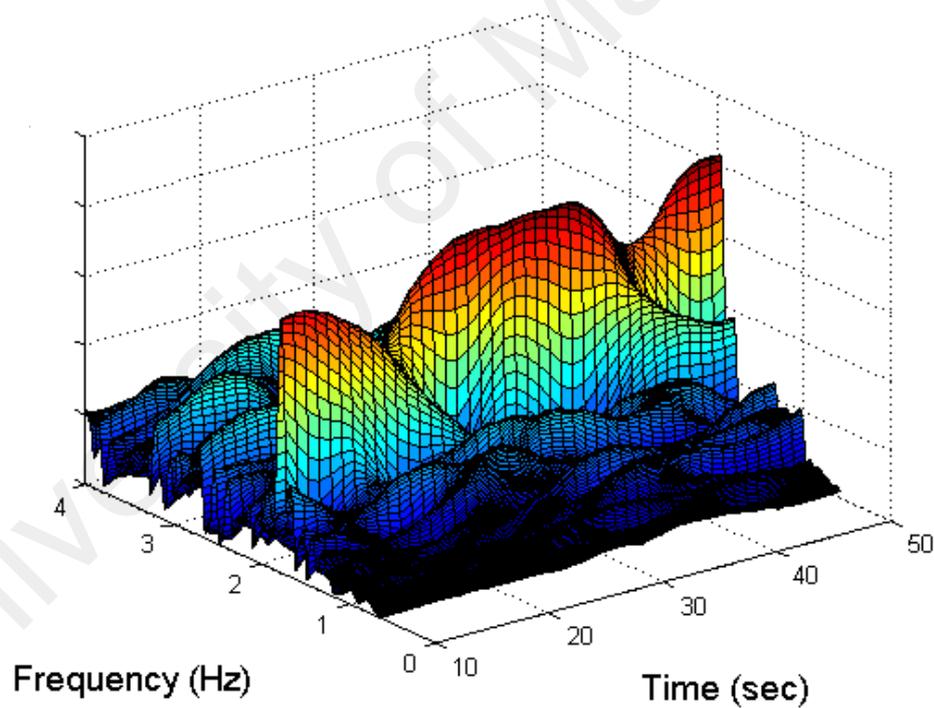


Figure 3.11: The power spectrum distribution obtained from STFT

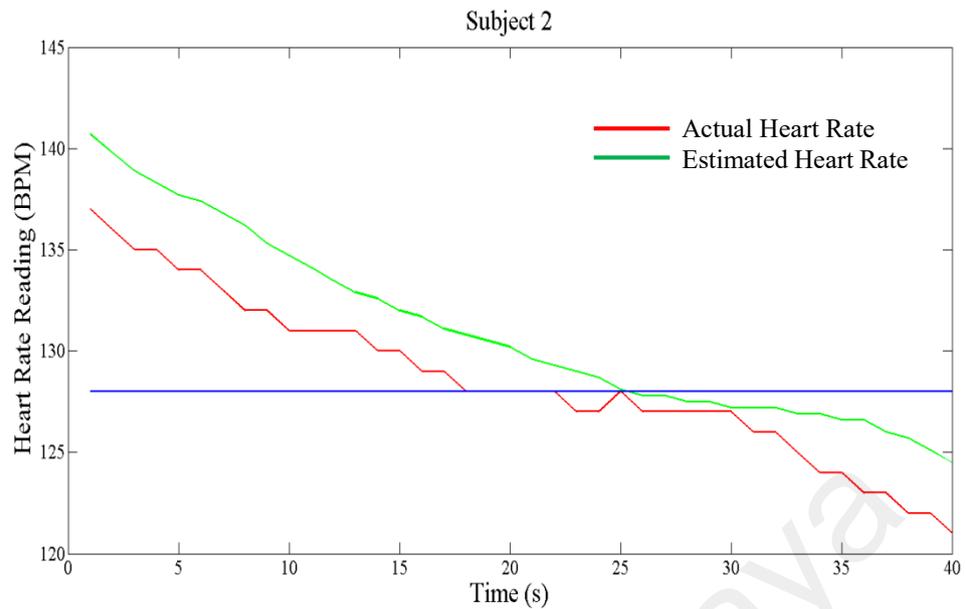


Figure 3.12: Comparison between proposed method and previous methods that used Fourier transform

3.3.3 Experimental Results and Analysis for Filter Bank Approach

Figure 3.13 shows the comparison of actual and estimated heart rate readings for all subjects in the first experiment while Figure 3.14 shows the comparison of actual and estimated heart rate readings for all subjects in the second experiment. For both figures, red color represents the actual heart rate readings while green color represents the estimated heart rate readings using the proposed method. The respective root mean square errors of the results are 2.63 BPM (first experiment) and 2.30 BPM (second experiment) while the respective correlation coefficients are 0.99 for both experiments.

The results shown in Figure 3.13 and Figure 3.14 indicate that filter bank is able to give a more accurate result as compared to the STFT. These two figures show that the readings are scattered around the reference line ($R=1$). For further discussion, the comparison of actual and estimated heart rate for each subject is presented. Figure 3.15 shows the comparison of actual and estimated rate for each subject in the first experiment while Figure 3.16 shows the comparison for the second experiment. Figure 3.15 and Figure 3.16 show that the lagging effect as shown in *Section 3.3.2* earlier is not significant at here. It

could most probably due to the longer window size of the STFT as compared to the window size of filter bank. As STFT performs the Fourier transform onto each window, the size of the window may not be too small. A very small window size gives a low-resolution frequency domain and may miss out some information.

In addition to this, the performance of the proposed algorithm (using filter bank) is also evaluated using the Bland Altman plots. Figure 3.17 shows the Bland Altman plot for the heart rate readings obtained from the first experiment while Figure 3.18 shows the Bland Altman plot for the heart rate readings obtained from the second experiment. Similarly, the heart rate readings obtained from the first experiment has higher range of errors as compared to the heart rate readings obtained from the second experiment. It is due to the occurrence of large amount of motion artifacts in the first experiment.

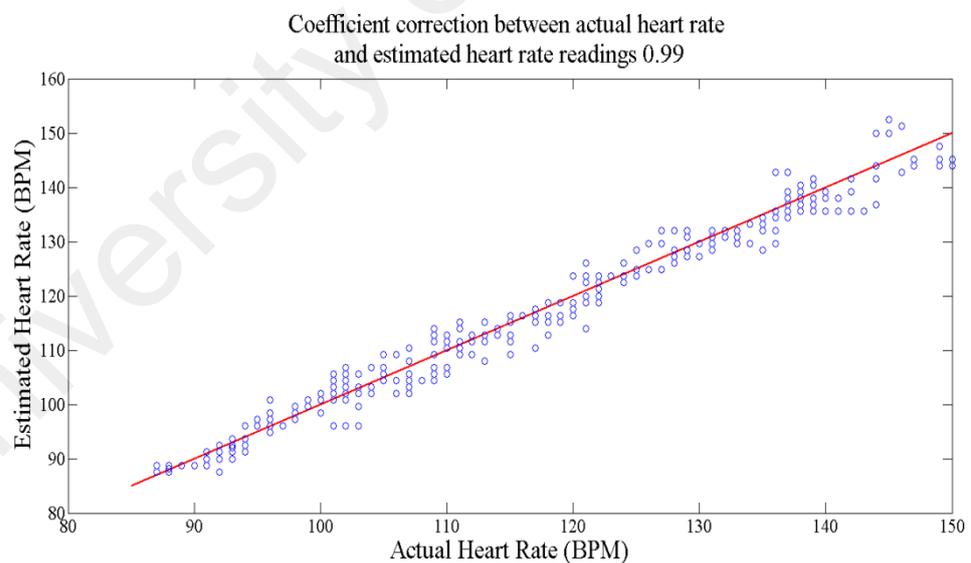


Figure 3.13: Comparison of all actual and estimated heart rate readings using filter bank for the first experiment

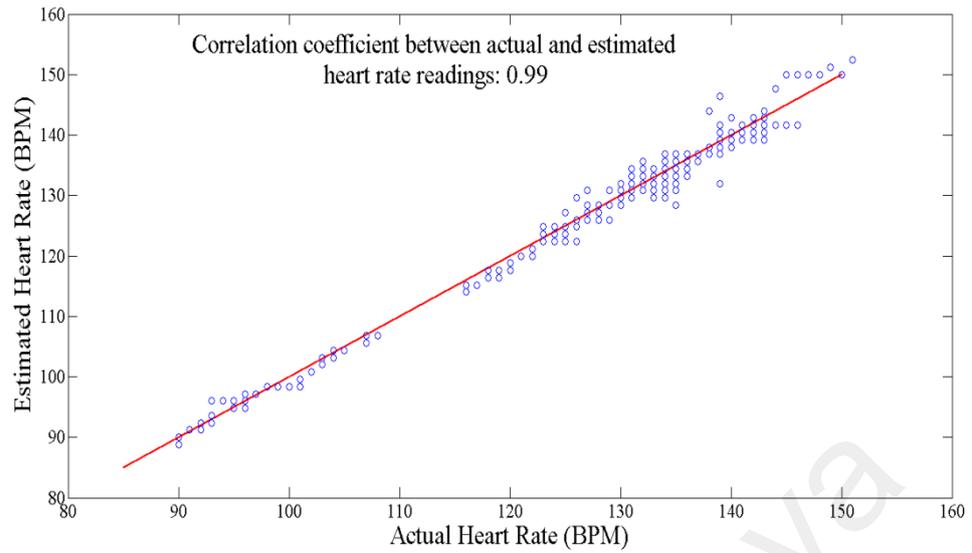
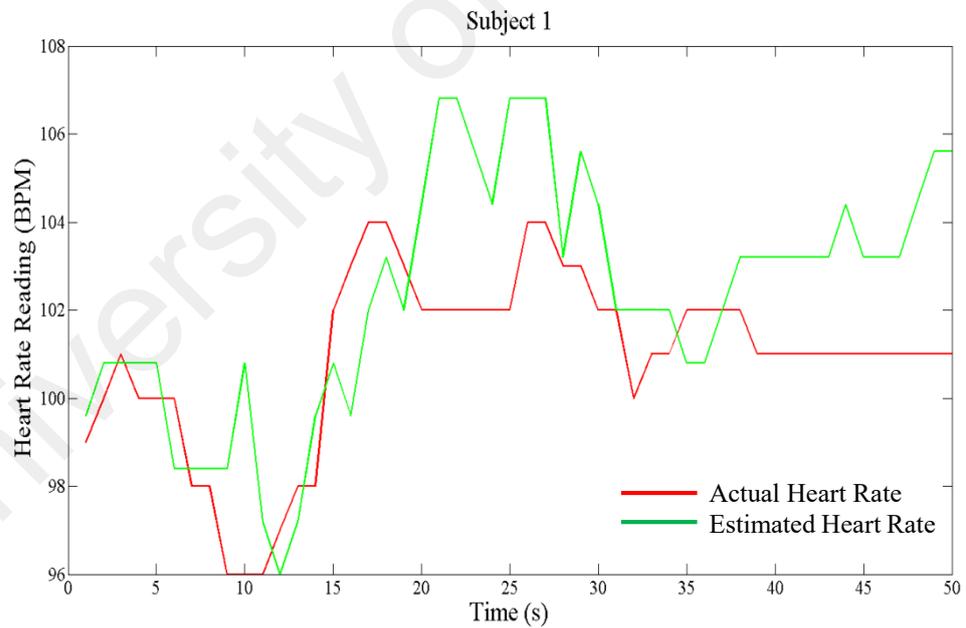
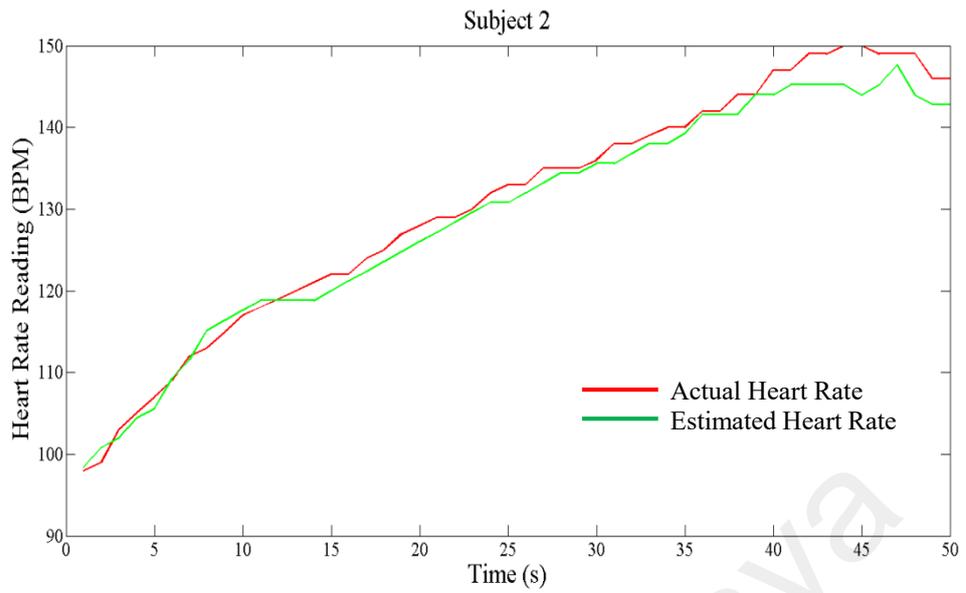


Figure 3.14: Comparison of all actual and estimated heart rate readings using filter bank for the second experiment

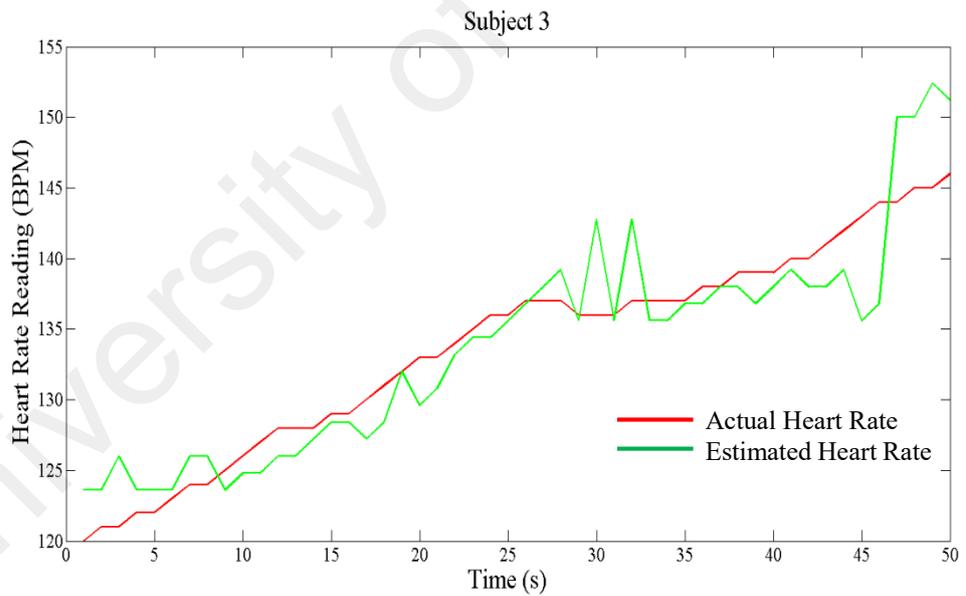


(a) First subject

Figure 3.15: Comparison of actual and estimated heart rate readings for each subject in the first experiment

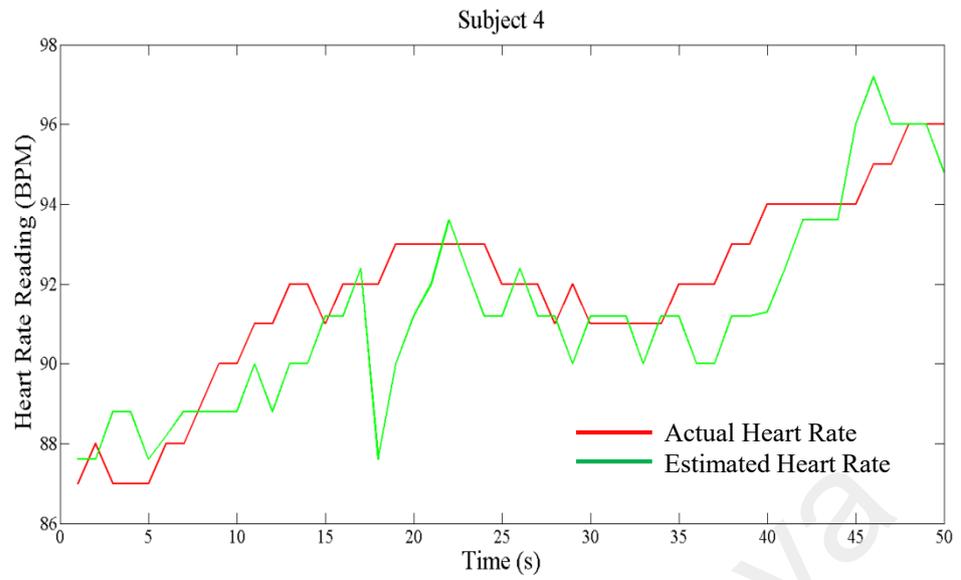


(b) Second subject

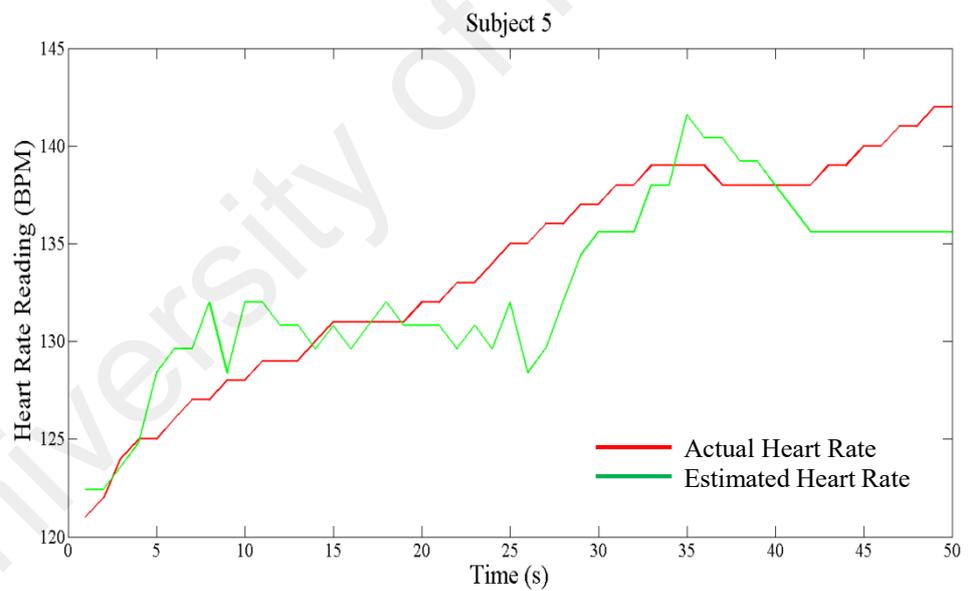


(c) Third subject

Figure 3.15.continued

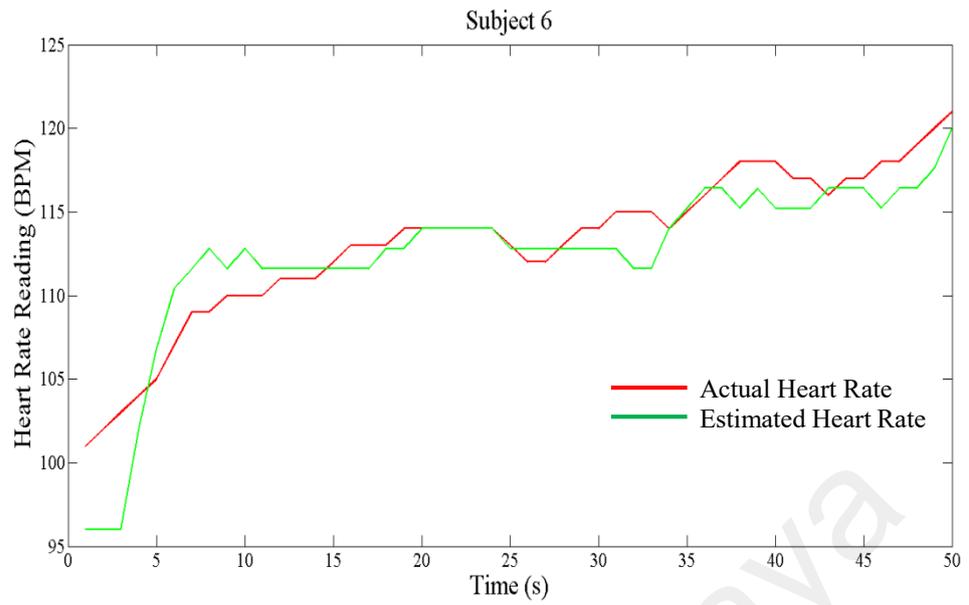


(d) Forth subject

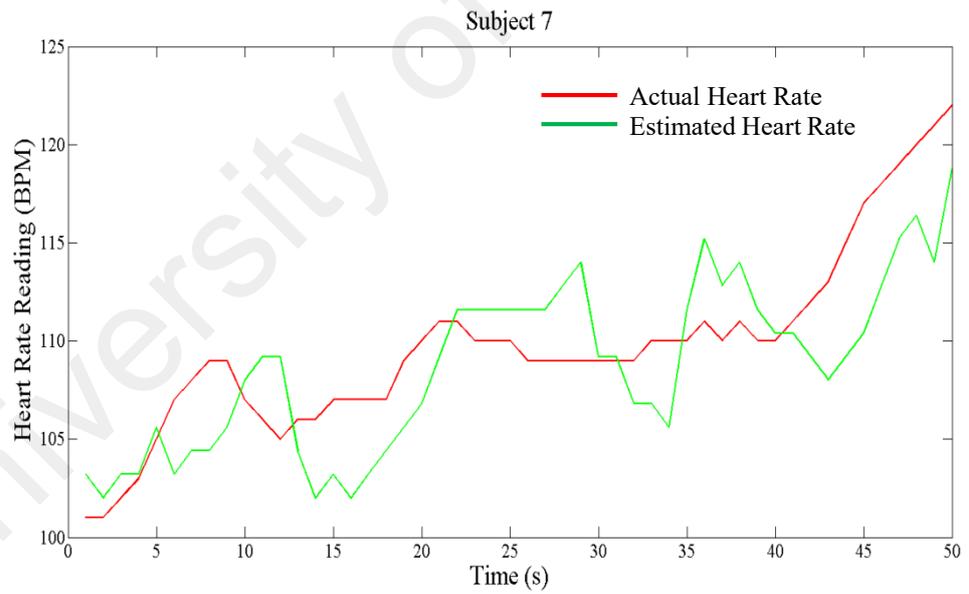


(e) Fifth subject

Figure 3.15.continued

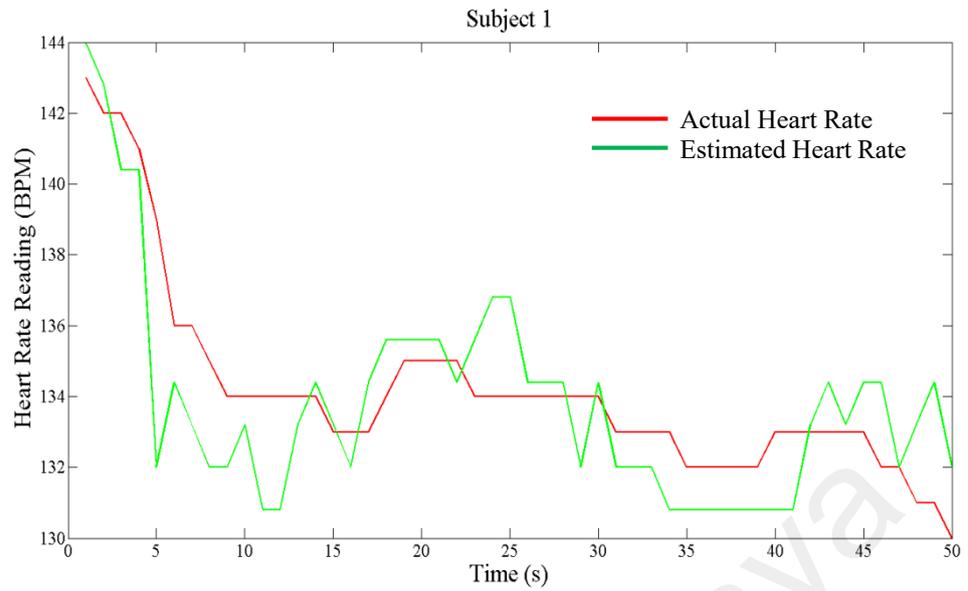


(f) Sixth subject

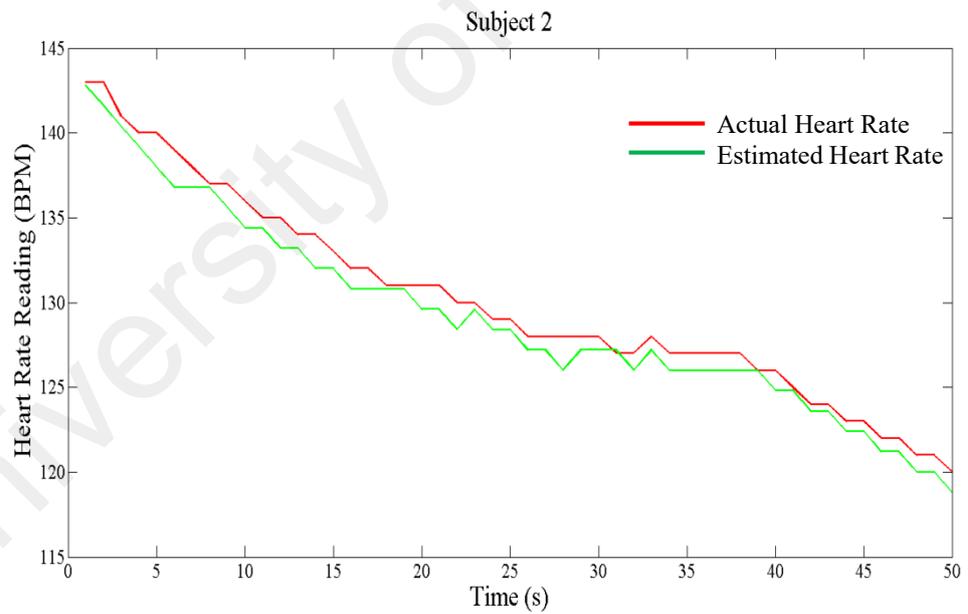


(g) Seventh subject

Figure 3.15.continued

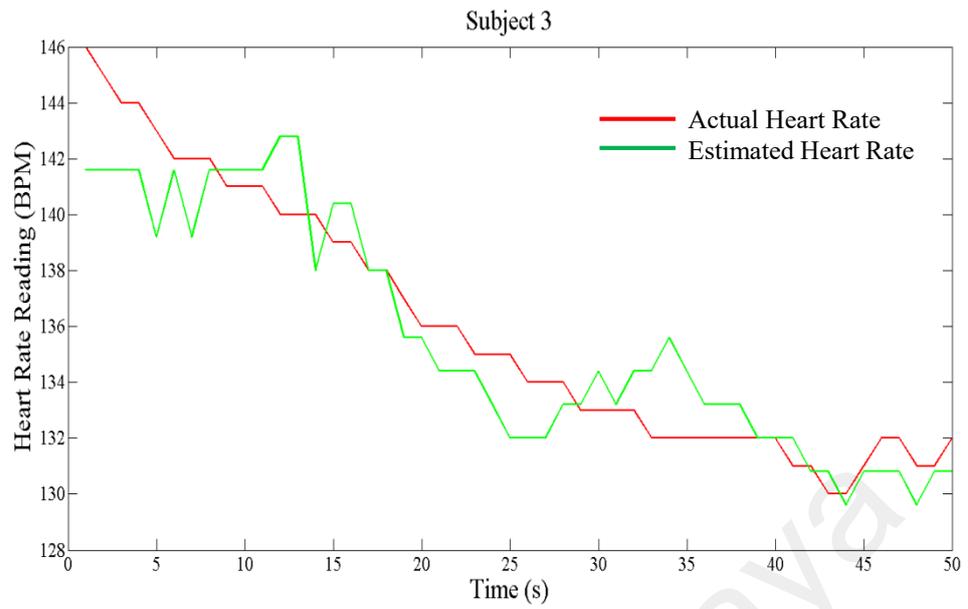


(a) First subject

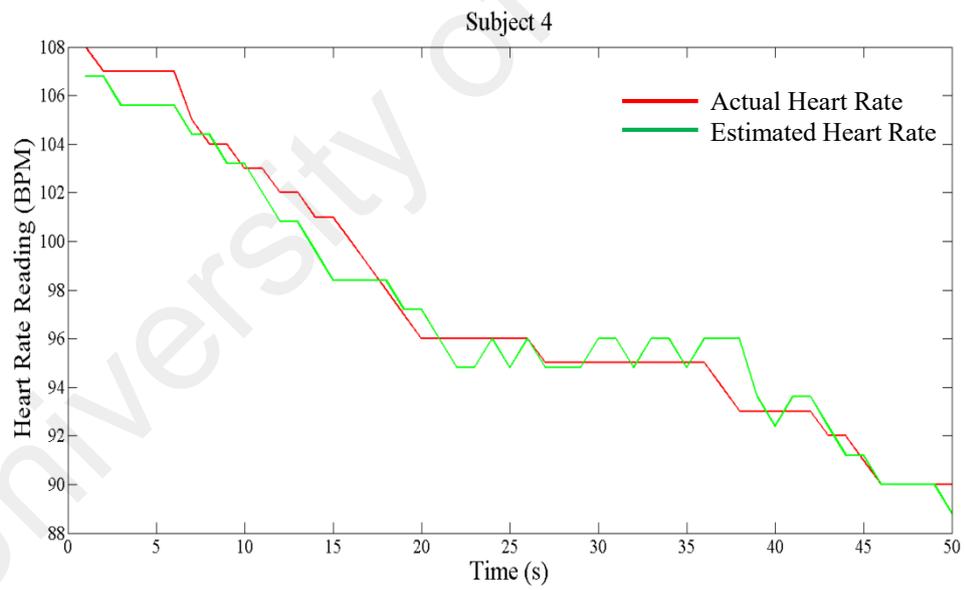


(b) Second subject

Figure 3.16: Comparison of actual and estimated heart rate readings for each subject in the second experiment

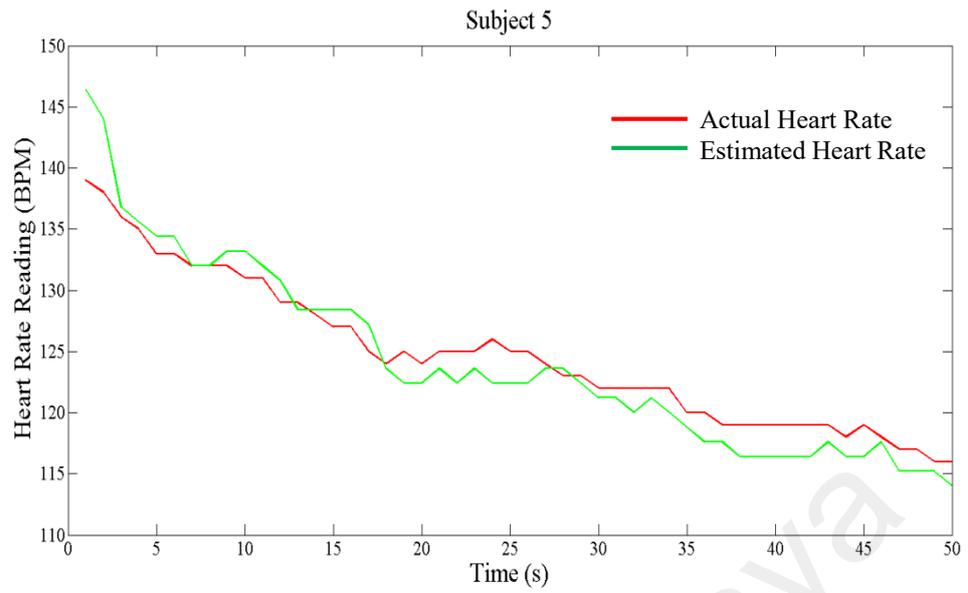


(c) Third subject

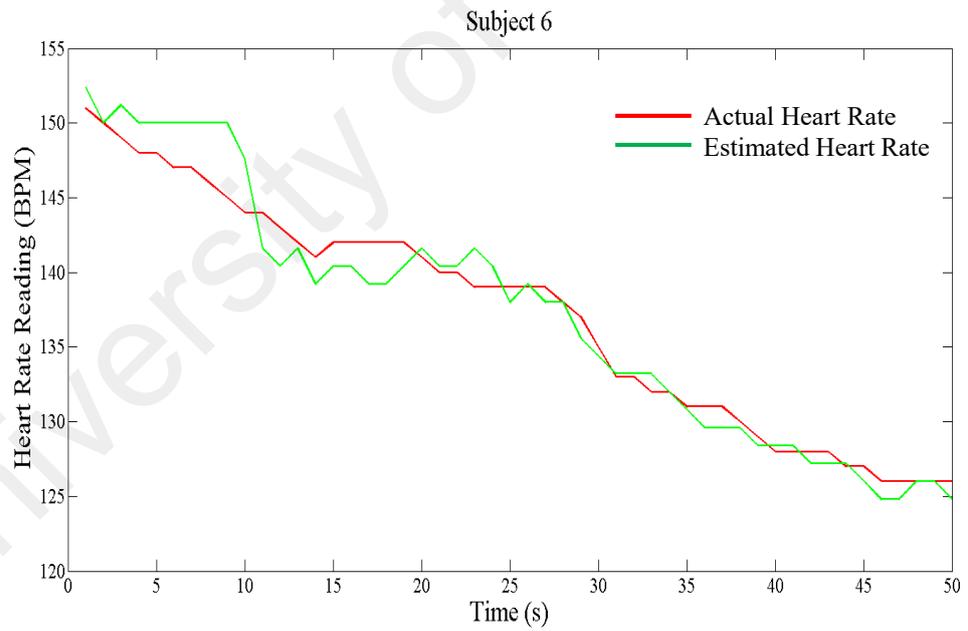


(d) Forth subject

Figure 3.16.continued

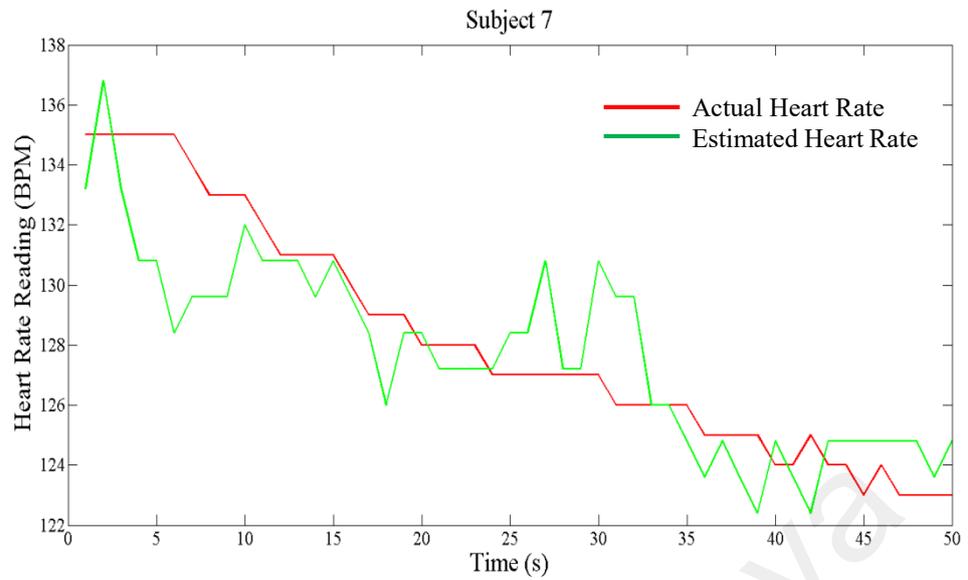


(e) Fifth subject



(f) Sixth subject

Figure 3.16.continued



(g) Seventh subject

Figure 3.16.continued

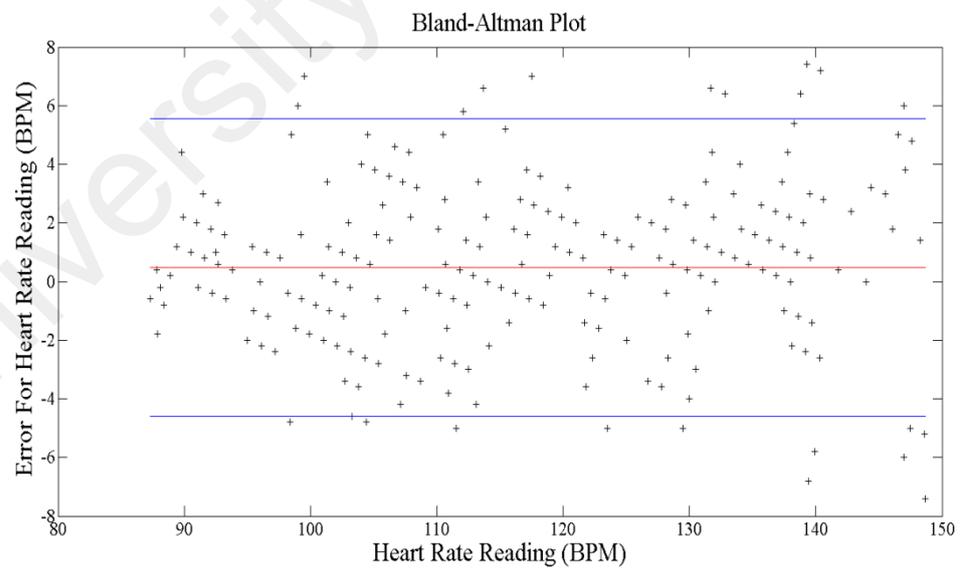


Figure 3.17: Bland-Altman plot for all estimated heart rate reading using filter bank for the first experiment

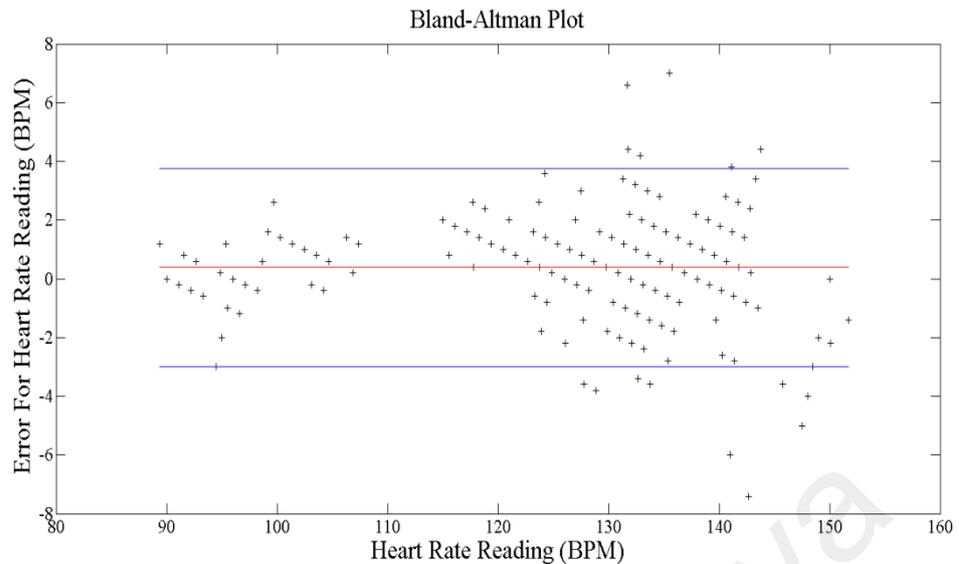


Figure 3.18: Bland-Altman plot for all estimated heart rate reading using filter bank for the second experiment

The drawback of using the approaches described above is the video duration used for heart rate estimation is quite long. For a long video duration, ICA is able to separate the sources very clearly, i.e the ICA sources are independent to each other. Hence, the source with heart rate signals can be easily obtained. However, for a very dynamic heart rate variation, short video duration is needed. An important consideration is the accuracy should not be compromised in the use of short video duration. ICA could render an inaccurate result if the video duration is too short. This issue is addressed and described in Chapter 4.

3.4 Chapter Conclusion

In this chapter, two different approaches that provide temporal information for video-based heart rate measurements are described. Previous works focus on minimal change in the heart rate variations, while this study deals with dynamic heart rate variation. For a rapidly changing heart rate patterns, the temporal information is essential to provide an accurate measurements. Both STFT and filter bank provide this information. Experimental results show that the proposed methods can give an acceptable result, for dynamic heart rate variation. However, it is found that filter bank is able to give a more

accurate answer as compared to the STFT. The common issue for both approaches is the lengthy video duration.

University of Malaya

CHAPTER 4: DYNAMIC HEART RATE ESTIMATION FROM SHORT VIDEO SEQUENCES

4.1 Overview

This chapter shows how dynamic heart rate measurements that are typically obtained from sensors mounted near to the heart can also be obtained from video sequences. As described in Chapter 3, a short video duration is needed for dynamic heart rate estimation. However, short video duration may decrease the amount of independence of ICA sources and may render to inaccurate readings. In this study, independent component analysis (ICA) is combined with mutual information to ensure accuracy is not compromised in the use of short video duration. Two experiments are carried out where a video camera captures the facial images of the seven subjects. The first experiment involves the measurement of subjects' increasing heart rates (79 to 150 beats per minute (BPM)) while cycling whereas the second involves falling heart beats (153 to 88 BPM). While both experiments are going on measures of heartbeat using the Polar heart rate monitor is also taken to compare with the findings of the proposed method. Overall experimental results show the proposed method can be used to measure dynamic heart rates where the root mean square error (RMSE) and the correlation coefficient are 1.88 BPM and 0.99 respectively.

4.2 Proposed Method

In this section, the proposed model to estimate the dynamic heart rate measurements is discussed together with how the independence of the ICA sources is established. The earliest sign of the independence of the ICA sources determined from the mutual information establishes the minimum video duration which gives the most accurate heart rate estimation is elaborated. The significance of the video duration and its relationship to the accuracy of the heart rate estimation is also discussed.

4.2.1 Workflow of the Proposed Method

In order to compute the instantaneous heart rate of a subject at n -th second instant, 50 frames of raw images (equivalent to 2 seconds) before the instant were taken as the raw input for the next processing. The block diagram of the proposed model for dynamic heart rate measurement is illustrated in Figure 4.1.

After the ROI of each frame was identified, the mean of pixel values for red (R), green (G) and blue (B) components were computed separately, where

μ_R : the mean of all pixel values for R component

μ_G : the mean of all pixel values for G component

μ_B : the mean of all pixel values for B component

The respective μ_R , μ_G , and μ_B of all these 50 continuous frames were calculated. Therefore, at that instant, a set of three raw sensors $R(n)$, $G(n)$, $B(n)$ were formed. Each raw sensor consists of 50 elements. The set of raw sensors were then detrended using algorithm developed by Tarvainen *et al.* (Tarvainen *et al.*, 2002). ICA model developed by Cardoso and Souloumiac (Cardoso, 1999; Cardoso and Souloumiac, 1993) was then used to separate one set of 3 independent sources from the set of sensors. The set of ICA sources were bandpass filtered (128-point Hamming window, 0.6-4 Hz), and the mutual information was applied to obtain the independence of the ICA sources.

The entire process was repeated by increasing the number of previous frames, one-by-one until it fulfilled the criterion. The criterion was based on the convergence of the curve fitting coefficients. It is described in details in *Section 4.2.2*.

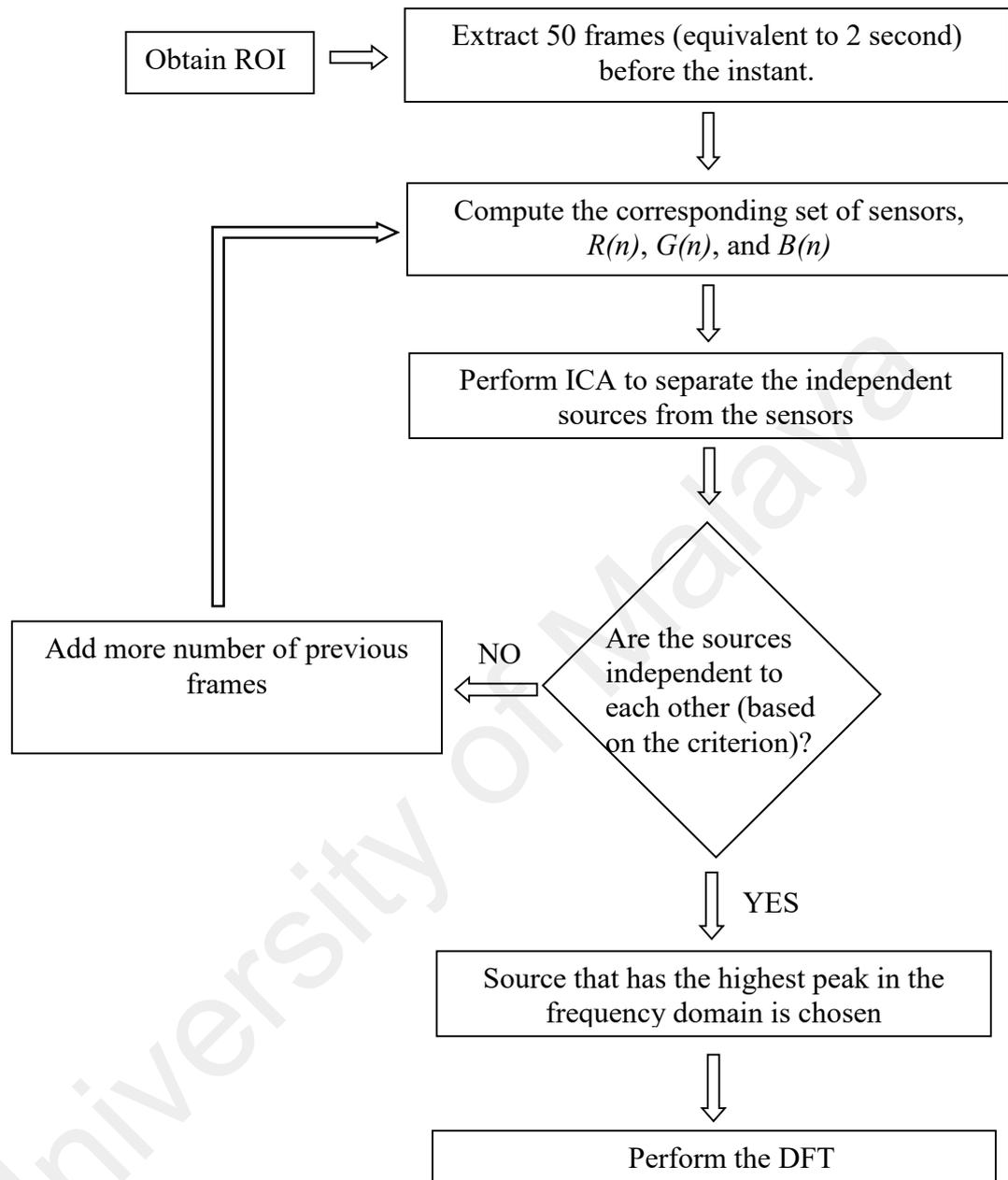


Figure 4.1: Flow chart of the proposed method

Once it fulfilled the criterion, the process was stopped. The number of frames (or video duration) at this point was chosen as the video duration needed for computing the instantaneous heart rate reading n -th second time instant. The respective sources were considered independent to each other. The source that had the highest peak in the frequency domain was chosen as the BVP or heart rate sources. The corresponding frequency was considered as the instantaneous heart rate at n -th second time instant.

4.2.2 Criterion Determining the Independence of the ICA Sources

One of the criteria or requirements to use the ICA is the sources must be independent to each other. ICA itself will separate the sources by maximizing the statistical independence among the given sensor signals. In this study, mutual information (Cover and Thomas, 2012) is used to measure the mutual dependence (and independences) of the ICA sources.

In this study, three ICA sources, i.e. S_1 , S_2 , and S_3 , are separated from each set of Red(R), Blue(B) and Green(G) sensors. The normalized mutual information of any two source S_p and S_q , $I_n(S_p; S_q)$ is expressed as

$$I_n(S_p; S_q) = \frac{I(S_p; S_q)}{\sqrt{H(S_p)H(S_q)}} \quad (4.1)$$

where $I(S_p; S_q)$ is the mutual information of the sources S_p and S_q while $H(S_p)$ and $H(S_q)$ are the respective entropies of the sources S_p and S_q .

The mutual information $I(S_p; S_q)$ can be expressed in terms of their entropy:

$$I(S_p; S_q) = H(S_p) + H(S_q) - H(S_p, S_q). \quad (4.2)$$

The entropy $H(S)$ of a source S can be computed from

$$H(S) = - \sum_{s \in S} p(s) \log p(s) \quad (4.3)$$

while the joint entropy $H(S_p, S_q)$ of the sources S_p and S_q can be computed from

$$H(S_p, S_q) = - \sum_{s_p \in S_p} \sum_{s_q \in S_q} p(s_p, s_q) \log p(s_p, s_q) \quad (4.4)$$

where $p(s_p, s_q)$ is the joint probability density function of the sources S_p and S_q .

The normalized mutual information is zero if both sources are totally independent to each other and unity if both sources are totally dependent to each other. Since there are three ICA sources, hence their normalized mutual information is averaged and it is expressed as

$$C(S_1;S_2;S_3) = \frac{I_n(S_1;S_2) + I_n(S_1;S_3) + I_n(S_2;S_3)}{3} \quad (4.5)$$

The relationship of $C(S_1;S_2;S_3)$ and video duration for estimating the instantaneous heart rate reading at a particular instant is shown in Figure 4.2. It indicates that the value of $C(S_1;S_2;S_3)$ decreases as the video duration (or number of video frames) increases. A best fit curve is used to represent the function $C(S_1;S_2;S_3)$. In this study, the best fit curve to represent the function of mutual information with the video duration is given as

$$C_t(S_1;S_2;S_3) = at^b, \quad (4.6)$$

where t is the video duration used, a and b are the coefficients of the function.

The confident bounds for the fitted coefficients are set at 95 %. The coefficients change drastically at the beginning, but remain almost constant when the video duration exceeds a specific duration and the mutual information value changes very little after this duration. The stopping criterion is set as when the difference of coefficient values for 2 continuous video frames is smaller than 2×10^{-4} . Once the stopping criterion is met, the corresponding video duration is identified as the video duration to compute the instantaneous heart rate at that particular instant.

As an example, Figure 4.2 shows that different time intervals give different values of mutual information and hence different heart rate readings. In this case, the actual heart rate reading of the subject obtained from Polar Heart Rate Monitor is 93 BPM.

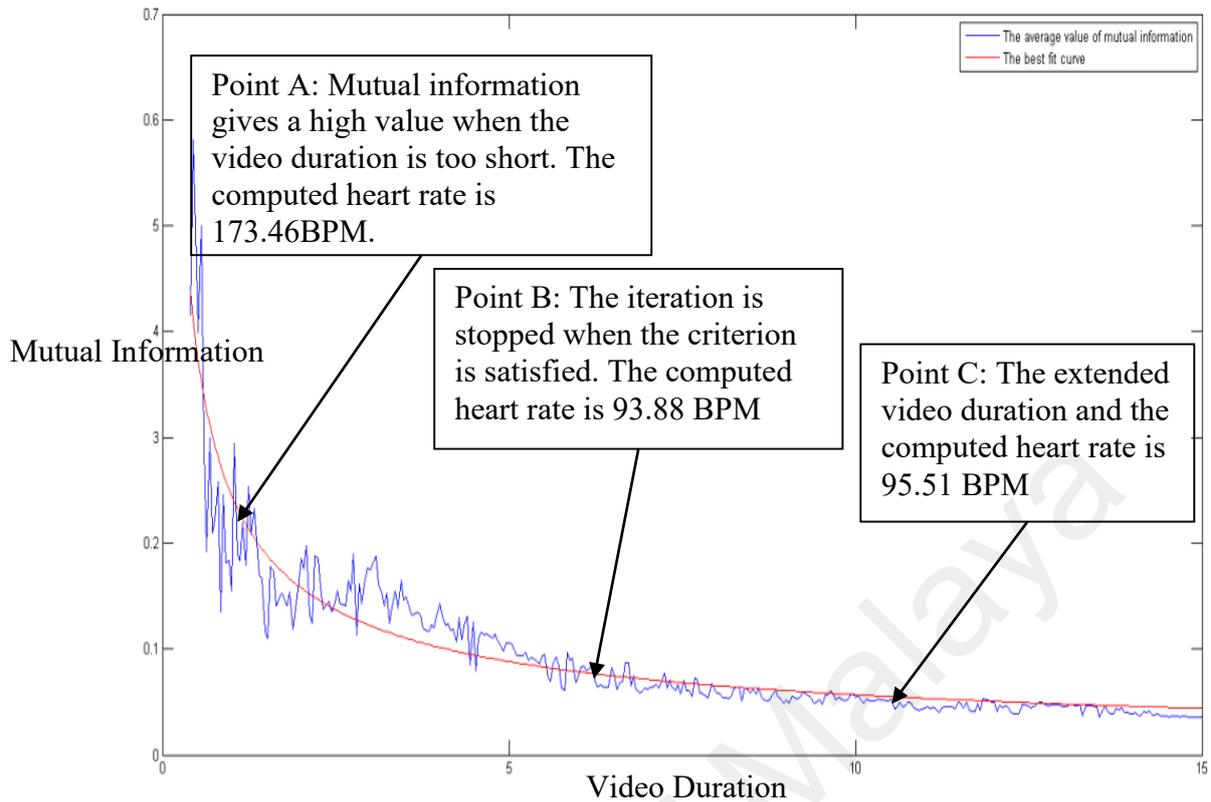


Figure 4.2: The relationship between the averaged normalized mutual information $C(S_1;S_2;S_3)$ and the video duration and the respective computed heart rates

Now consider point A in Figure 4.2 which indicates the corresponding value of $C(S_1;S_2;S_3)$ when the video duration is 2.2 seconds. Notice at this point, the mutual information gives a high value indicating the ICA sources are not independent. The heart rate computed based on these ICA sources is 173.46 BPM. As the video duration increases, correspondingly $C(S_1;S_2;S_3)$ too decreases. This trend continues till the stopping criterion is met. As can be seen in Figure 4.2, the criterion is satisfied at point B (video duration for this case is 6.88s). The heart rate reading computed at this point is 93.88 BPM, which is closer to the actual reading. Even if the video duration is extended longer than the necessary duration, the $C(S_1;S_2;S_3)$ doesn't vary much. However, the heart rate accuracy drops further if longer video duration is taken after the stopping criterion is met. For example, consider point C where the video duration is 10.8 seconds and the ICA sources are still independent to each other but it gives a less accurate heart

rate reading when compared to point B. The computed heart rate is 95.51 BPM. The details of the heart rate accuracies at different points (hence different durations) are discussed in *Section 4.2.3*.

4.2.3 Significance of the Minimum Video Duration

In this sub-section, the significance of the video duration and its relationship to the accuracy of the heart rate estimation is presented. The significance of the independent sources and how they relate to the accuracy of the heart rates is shown in Figure 4.3, Figure 4.4 and Figure 4.5. Figure 4.3 shows the separated ICA sources for point A (at 2.2 s) in the frequency domain. As can be seen in Figure 4.3, the highest peak is observed at S3 and the corresponding frequency is 173.46 BPM. The actual heart rate reading obtained from the Polar heart rate monitor for the particular instant discussed in Figure 4.2 is 93 BPM. As pointed out earlier, the ICA sources are not independent and hence the computed heart rate reading is inaccurate.

Figure 4.4 shows the separated ICA sources for point B (at 6.88 s) in frequency domain. At this point, the stopping criterion is met. As shown in Figure 4.4, the highest peak is significantly seen at S3 and the corresponding frequency is 1.561 Hz, which is same as 93.66 BPM. The computed reading is closer to the actual heart rate reading. For the video duration that is beyond point B, the accuracy of the computed heart rate readings decreases. This is shown in Figure 4.5. In this figure, point C (10.8 s), the highest peak in the frequency domain is observed at S3 and the corresponding frequency is 95.58 BPM. For dynamic heart rate measurements, a shorter video duration is preferred as it would allow a more frequent update of heart rate measurements.

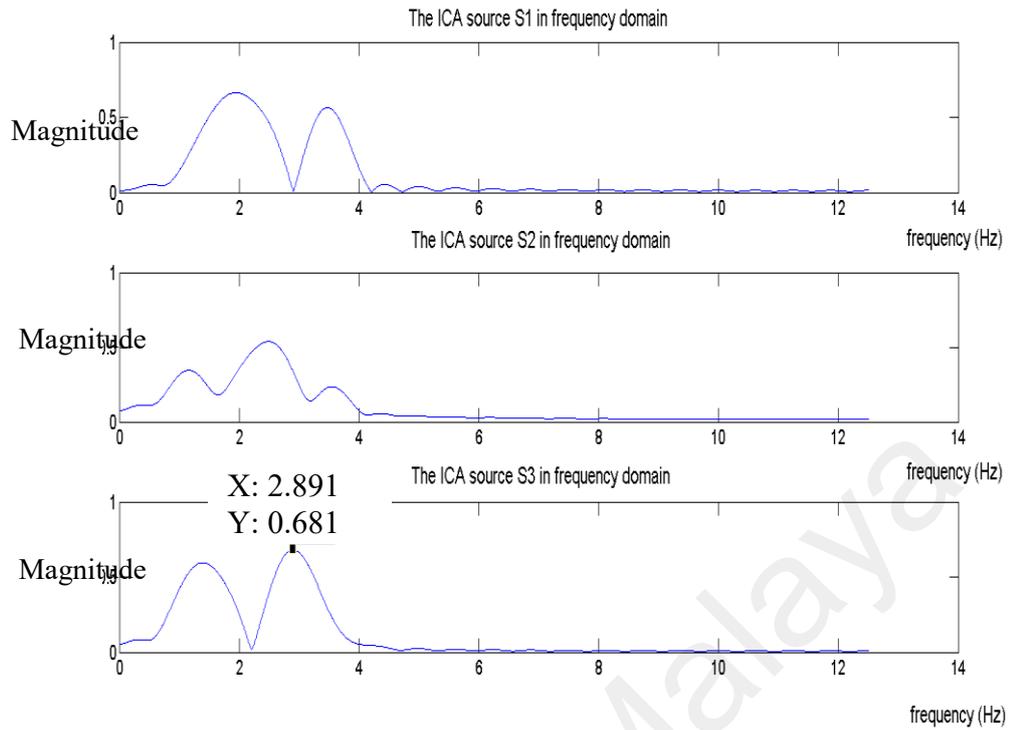


Figure 4.3: The frequency domain of the ICA sources when video duration at point A is selected

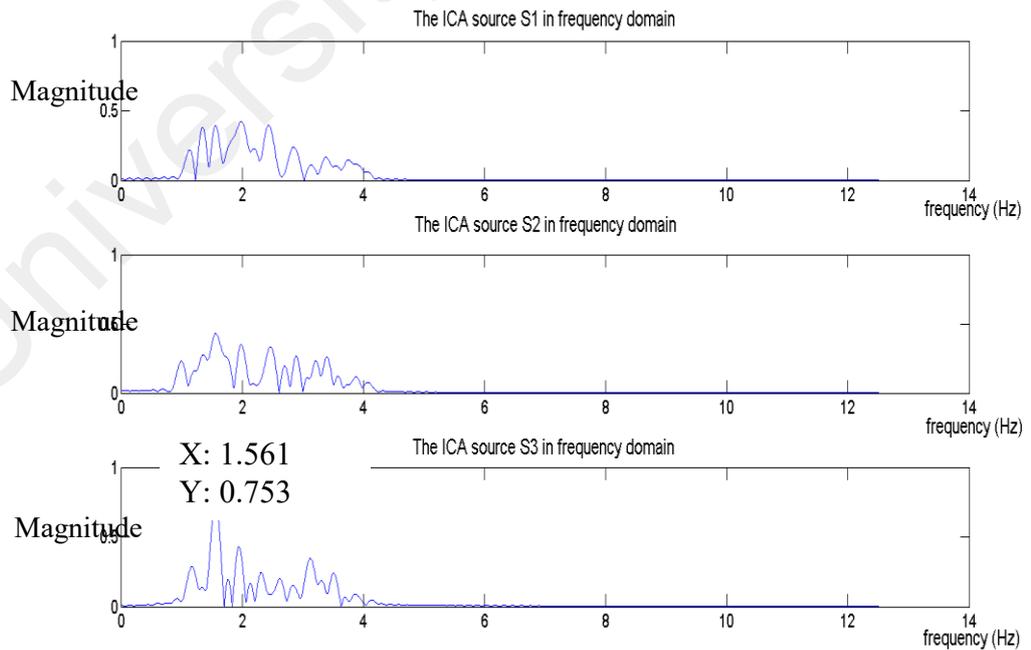


Figure 4.4: The frequency domain of the ICA sources when video duration at point B is selected

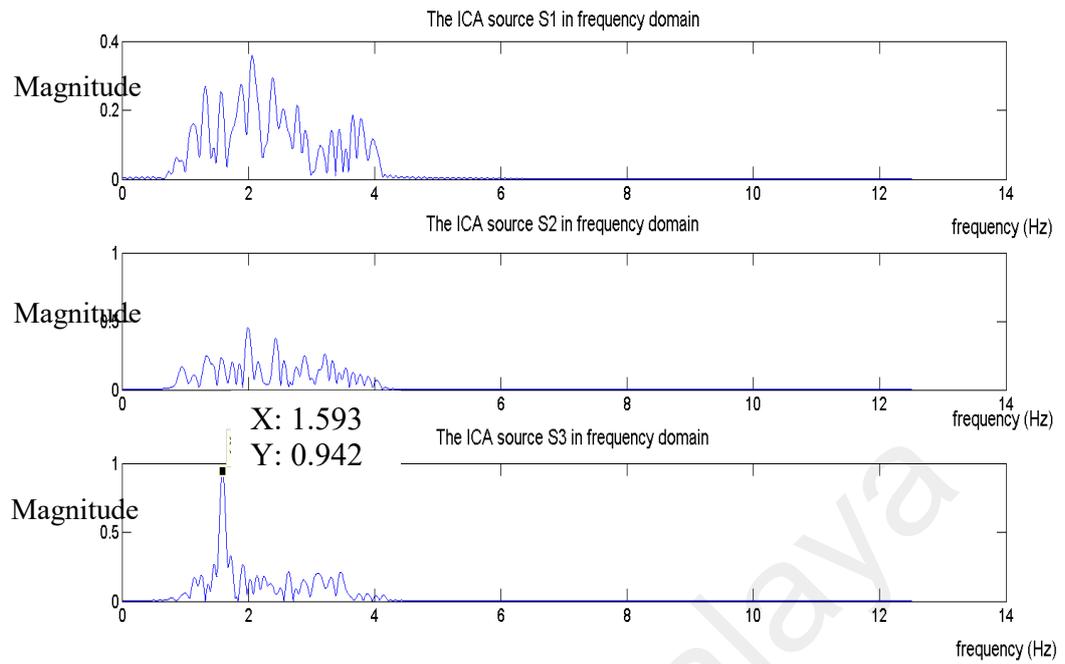


Figure 4.5: The frequency domain of the ICA sources when video duration at point C is selected

Figure 4.6 shows how the errors of the computed heart rate readings vary at different video intervals. It shows the error bars of the heart rate errors for subjects tested in the experiments. As can be seen from Figure 4.6, the error rates are considerably high if the video duration is less than 4 seconds. Even though the mean errors from 4 seconds to 5.5 seconds are relatively as close to the mean error of the proposed method, however, the standard deviation of the errors for the proposed method is much smaller when compared with those from 4.5 seconds to 5.5 seconds. This explains the need to determine the earliest sign of the independence of the ICA sources which corresponds to the minimum video duration that gives accurate heart rate estimation for all readings. Table 4.1 summarizes the video durations for all readings to estimate the heart rates of each subject using the proposed method.

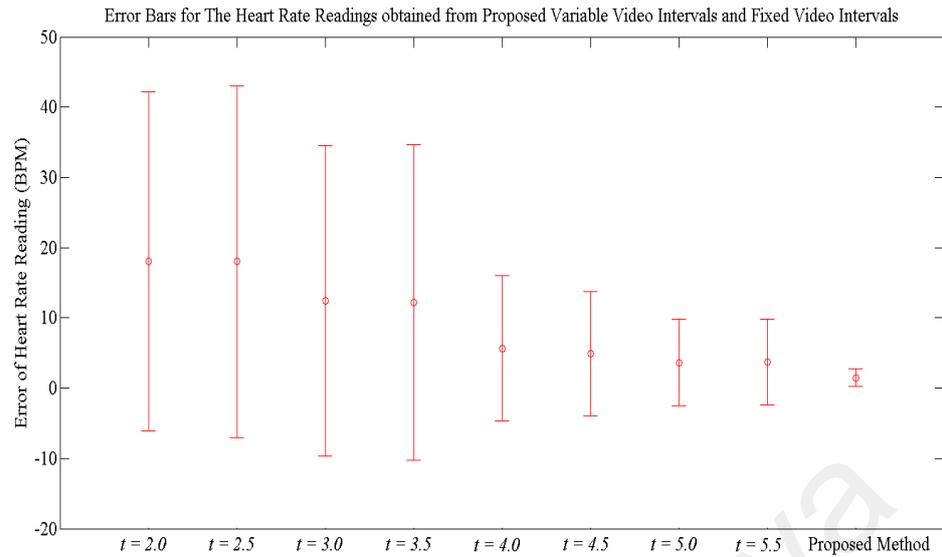


Figure 4.6: Comparison of the mean square and standard deviation of the heart rate errors for the proposed variable video intervals and fixed video intervals. T represents the video interval

Table 4.1: Summary of the video durations for all heart rate readings of each subject using the proposed method

Experiment	Subject	Video Duration (sec)		Mean (sec)	Standard Deviation (sec)
		Minimum	Maximum		
1	1	3.88	7.52	5.35	0.83
	2	3.76	7.56	5.56	0.91
	3	3.76	6.80	5.47	0.68
	4	3.80	7.32	5.63	0.90
	5	3.64	7.36	5.37	0.85
	6	3.76	7.48	5.26	0.89
	7	3.72	7.32	5.49	0.88
2	1	3.76	7.56	5.51	0.91
	2	3.72	6.80	5.17	0.68
	3	3.60	7.12	5.45	0.80
	4	3.64	7.16	5.29	0.78
	5	3.84	7.56	5.47	0.80
	6	3.96	7.20	5.42	0.76
	7	3.52	7.04	5.42	0.78

4.3 Experimental Study

In this section, the experimental setup and two experiments that relate to dynamic heart rate variation are discussed. Heart rate variation can be either increasing or decreasing. This study consists of both cases. In the first experiment, the heart rates of the subjects were increasing, ranging from 79 to 150 BPM, while in the second experiment, the heart rates of the subjects are decreasing, ranging from 153 to 88 BPM. In addition to the dynamic heart rate experiments, a sub-section is included to show that the proposed method can also be used for subjects while at rest.

4.3.1 Experimental Setup

All experiments were set up under office fluorescent lights with indirect sunlight as the source of illumination. The lighting background was homogeneous and had no significant changes or variation. All the data were processed and analyzed offline using MATLAB R2013a.

A Handycam Camcorder (Sony HDR-PJ260VE) with resolution of 1440×1080 pixels was used with 25 frames were sampled every second for video recording purpose. All videos were recorded in 24-bit RGB (with 8 bits per channel). The video camera was fixed at a position with a distance of about 0.60 m from the subject's face. In this study, the Region of Interest (ROI) is fixed at the area below eyes and above the upper lip of mouth in a video frame. As what concluded by Pursche *et al.* (Pursche *et al.*, 2012), this region gives better accuracy compared to other facial regions. The face region was detected by using existing model (Ojala *et al.*, 2002; Viola & Jone, 2001).

All subjects were asked to wear the Polar chest strap before doing the experiments. In the first experiment, seven subjects were asked to cycle at different speeds for three minutes where significant changes of the subjects' heart rates were observed after the first two minutes. The heart rate readings were taken in the last one minute for every subject.

An increasing heart rate variation was observed for every subject. Video was recorded while the subjects were cycling where their faces had minimum movement.

In the second experiment, seven subjects were asked to cycle at fast speed to raise their heart rates to a certain high level. Once this was achieved, then the subjects were asked to rest for one minute and their heart rates were observed and computed from video recorded during this period. The subjects did not move during the video recording. 60 consecutive heart rate readings (sampled at each second) were computed for every subject.

The instantaneous heart rates of all subjects for both experiments were computed based on the detailed algorithm proposed in *Section 4.2*. For reference, all instantaneous heart rates of the subjects were measured using Polar Heart Rate Monitor – Polar Team² Pro. A comparative study was done between the actual readings obtained from Polar Team² Pro and those computed readings from the proposed method.

4.3.2 First Experiment: Observed Heart Rates Varying from Low to High

In the first experiment, seven subjects' heart rates were measured while cycling and they varied from 79 BPM to 150 BPM. The video duration needed for each instantaneous heart rate reading computed using the proposed method varied from 3.64 to 7.52 seconds with a mean value of 5.45 seconds. A total of 420 instantaneous heart rate readings were obtained from the experiment. A comparison of the estimated and actual readings of the subjects in this experiment is shown in Figure 4.7. The root mean square error is 1.97 BPM while the Pearson correlation coefficient is 0.99.

The details of the heart rate variation of the seven subjects and their corresponding RMSE are shown in Table 4.2. From Table 4.2, the highest and lowest RMS errors are 2.33 and 1.64 BPM respectively. The comparison of actual and computed heart rate for

each subject is shown in Figure 4.8. For this figure, red color represents the actual heart rate readings while green color represents the estimated heart rate readings. Additionally, the performance of the proposed algorithm is evaluated using the Bland Altman plot, as shown in Figure 4.9. Bland Altman plot is used to quantify the agreement between two methods of measurements (Bland & Altman, 1999). The 95% limits of agreement, estimated by mean difference ± 1.96 standard deviation of the difference, provide an interval within which 95% of differences between the measurements by the two methods.

It can be seen from Figure 4.9 that the Bland Altman plot quantifies the agreement between the actual heart rate readings obtained from Polar Heart Rate Monitor and the computed heart rate readings using the proposed method. As it can be seen that most of the computed readings are located inside the blue boundary lines that satisfies the 95 % limits of agreement. However, there are some readings located out of the boundary lines and this is probably due to some motion artifacts.

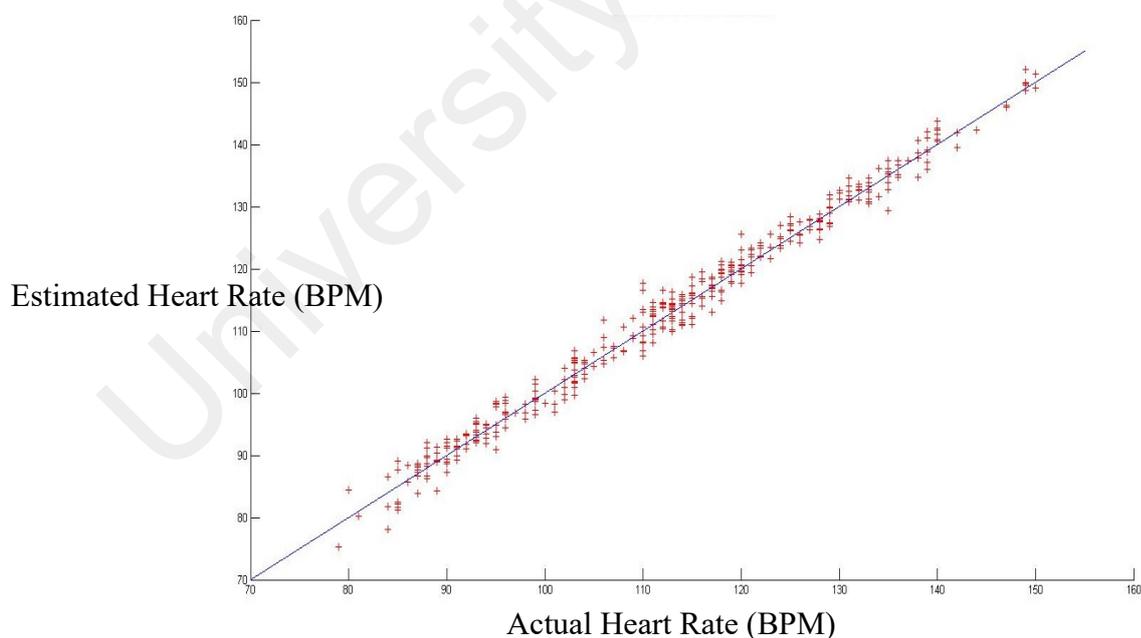
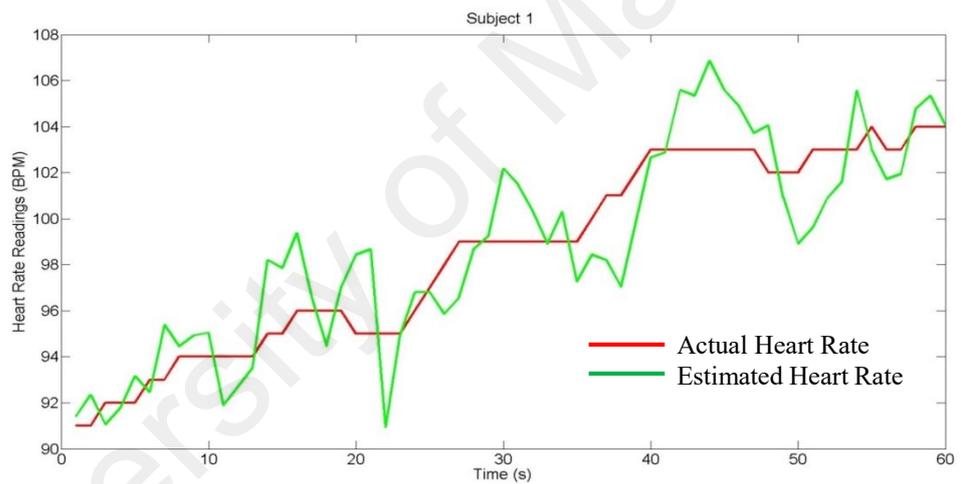


Figure 4.7: Comparison of all actual and estimated heart rate readings for the first experiment

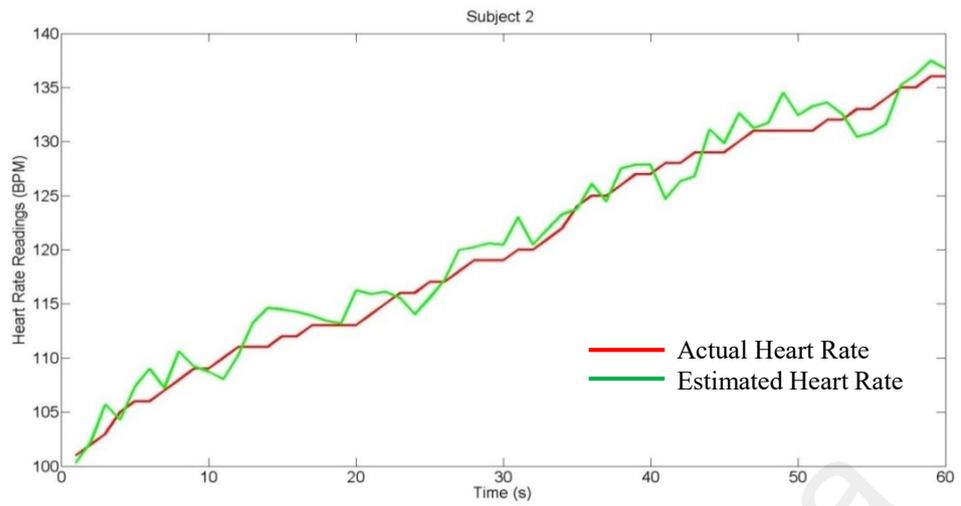
Table 4.2: Summary of heart rate readings results for the first experiment

Subject	Heart Rate Readings(BPM)		RMSE (BPM)
	Lowest	Highest	
1	91	104	2.01
2	101	136	1.76
3	104	120	2.03
4	93	150	1.64
5	110	129	1.67
6	110	140	2.33
7	79	96	2.28

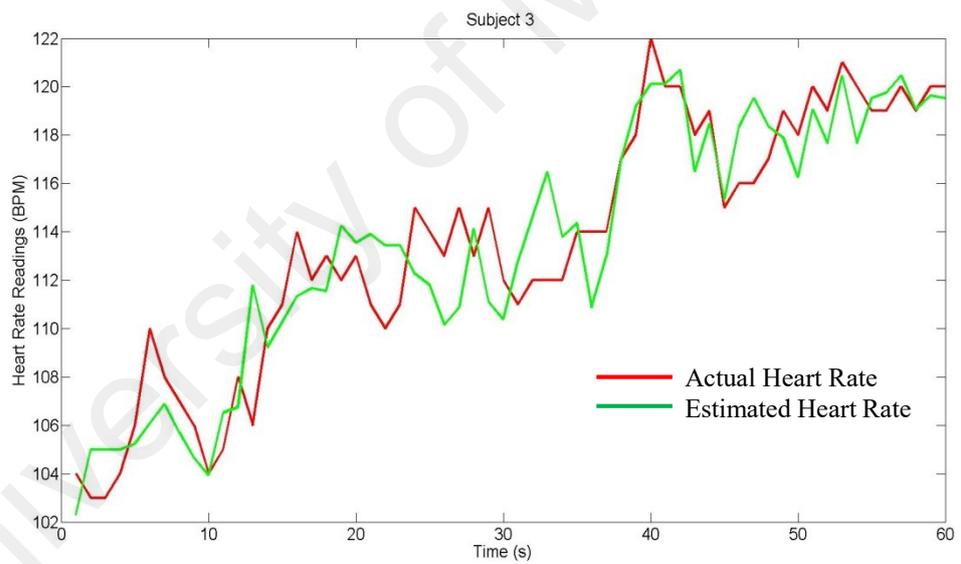


(a) First subject

Figure 4.8: Comparison of actual and estimated heart rate readings for each subject in the first experiment

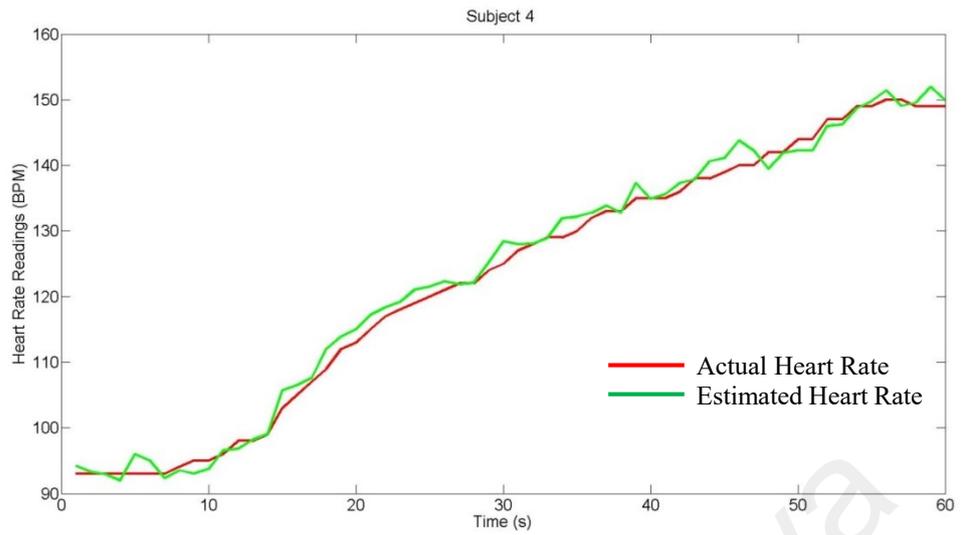


(b) Second subject

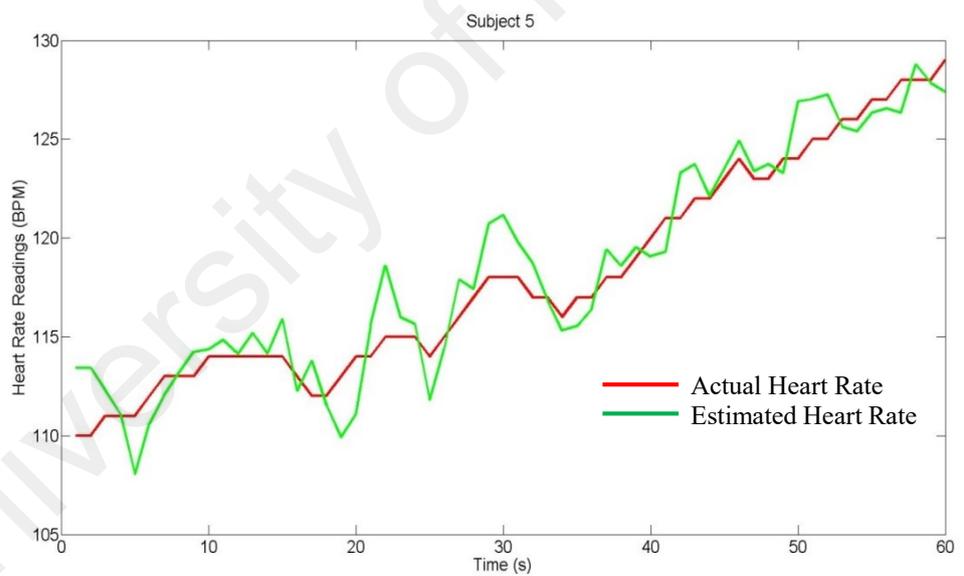


(c) Third subject

Figure 4.8.continued

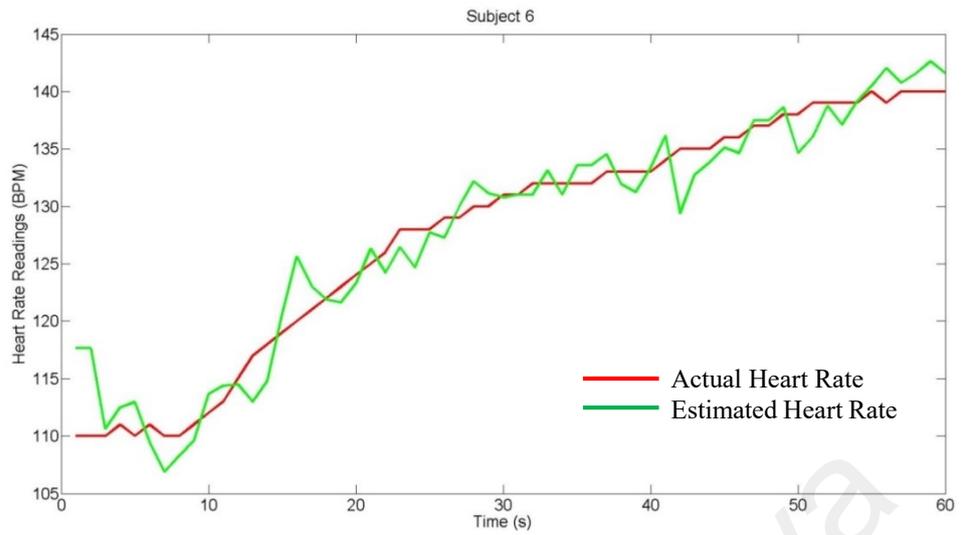


(d) Forth subject

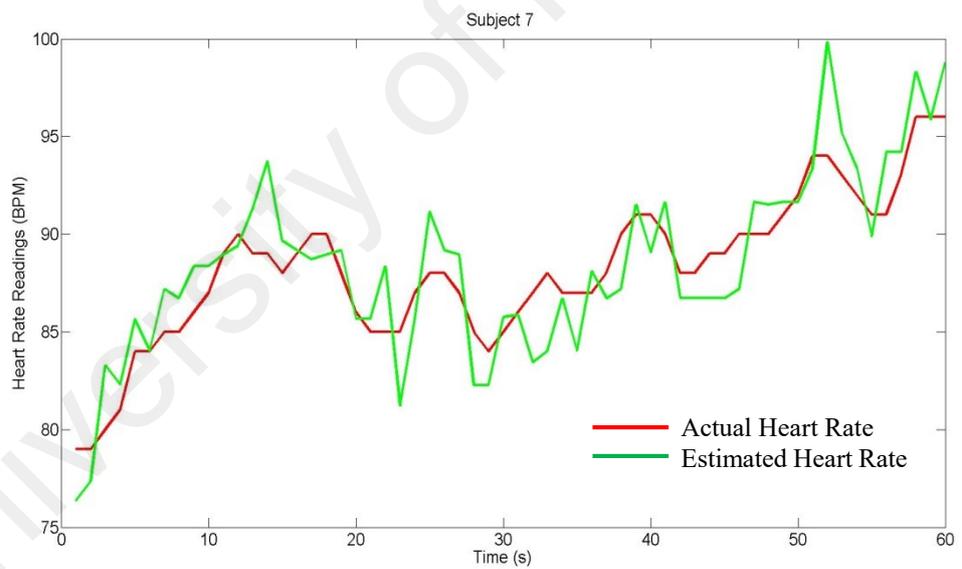


(e) Fifth subject

Figure 4.8.continued



(f) Sixth subject



(g) Seventh subject

Figure 4.8.continued

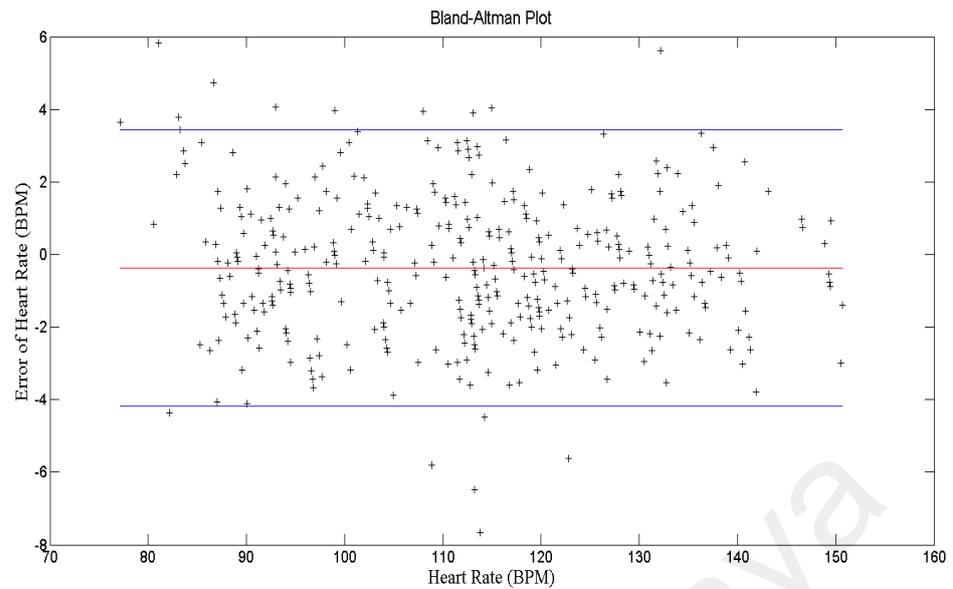


Figure 4.9: Bland-Altman plot for all estimated heart rate reading for the first experiment

4.3.3 Second Experiment: Observed Heart Rates Varying from High to Low

In the second experiment, the seven subjects' heart rates varied from 153 BPM to 88 BPM. A total of 420 instantaneous heart rate readings were obtained from this experiment. The video duration needed for each instantaneous heart rate reading computed using the proposed method varied from 3.52 to 7.56 seconds with a mean value of 5.39 seconds. Figure 4.10 shows the comparison of the estimated and actual heart rate readings of the seven subjects. The root mean square error is 1.77 BPM while the Pearson correlation coefficient is 0.99. Table 4.3 shows the details of the heart rate variations of the seven subjects and their corresponding RMSE. Just as in the first experiment, the highest and the lowest high rate RMS errors are less than 3 BPM. Figure 4.11 shows the comparison of actual and computed heart rate readings for each subject. For this figure, red color represents the actual heart rate readings while green color represents the estimated heart rate readings.

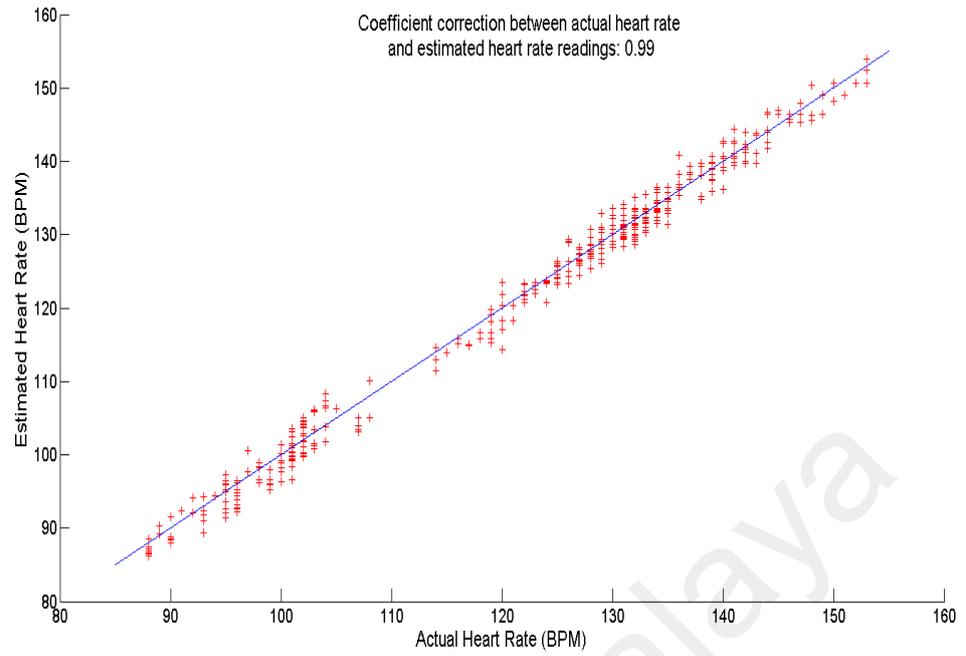
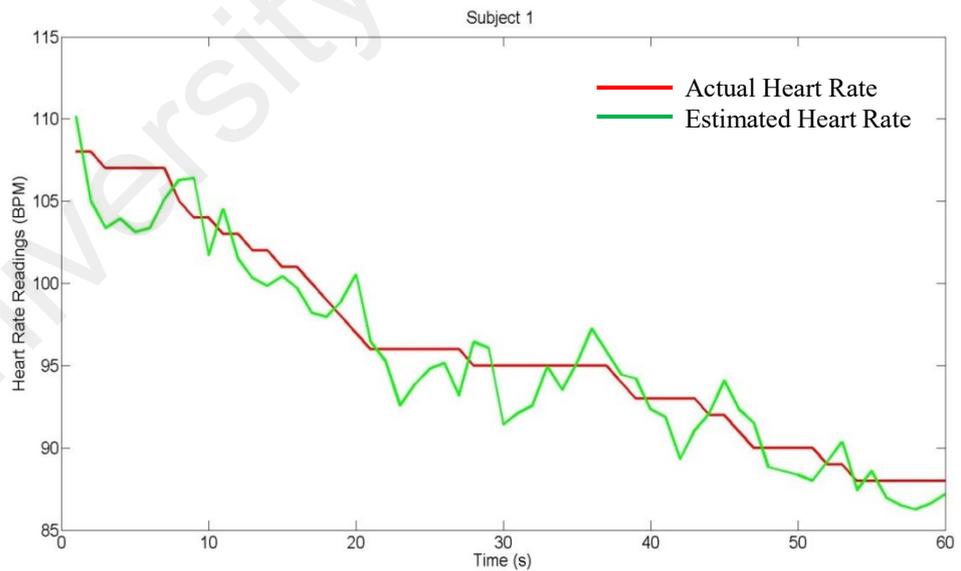
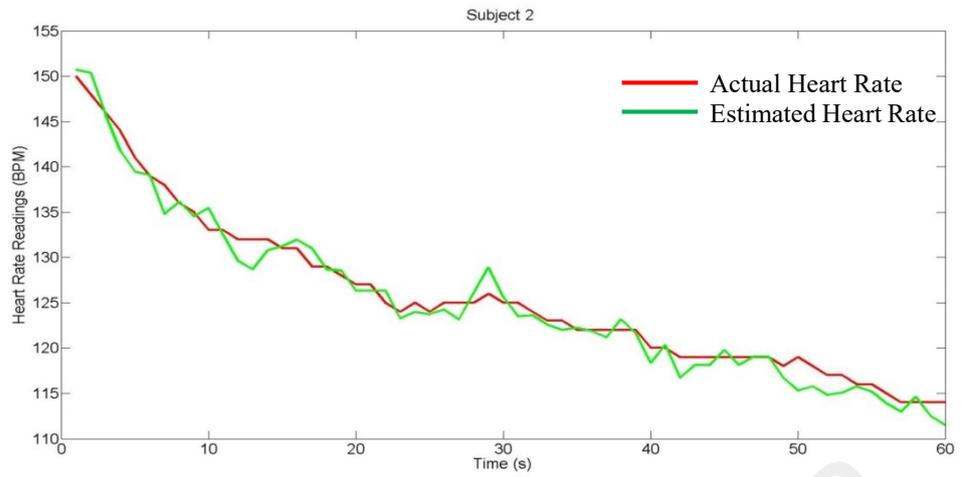


Figure 4.10: Comparison of all actual and estimated heart rate readings for the second experiment

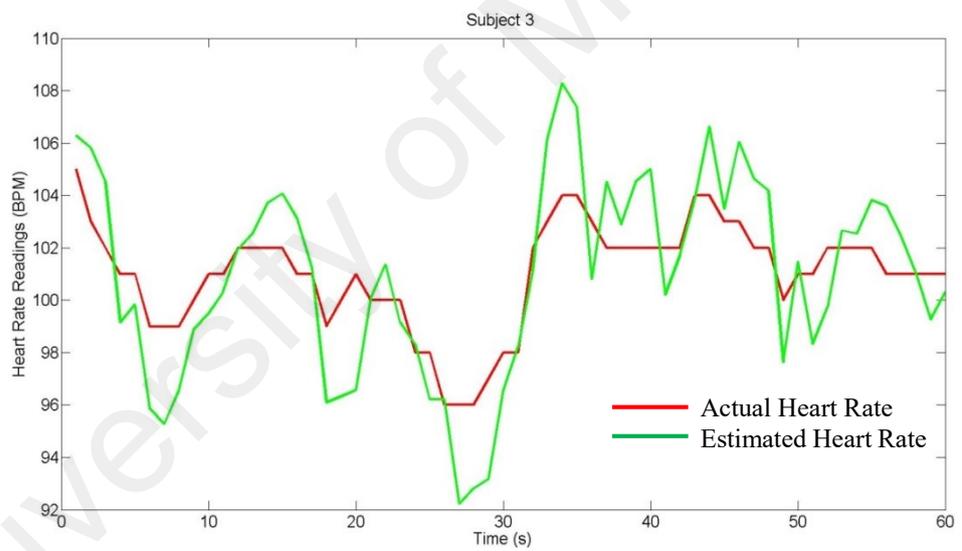


(a) First subject

Figure 4.11: Comparison of actual and estimated heart rate readings for each subject in the second experiment

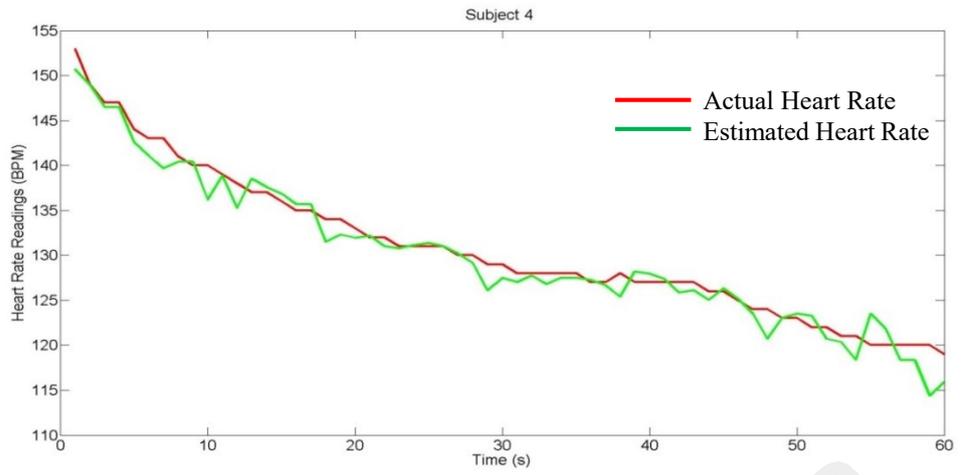


(b) Second subject

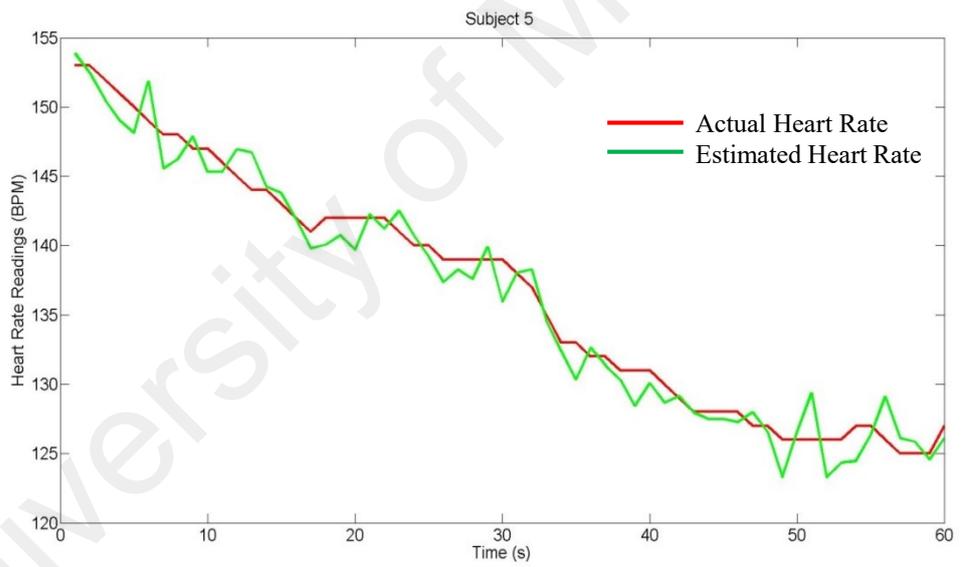


(c) Third subject

Figure 4.11.continued

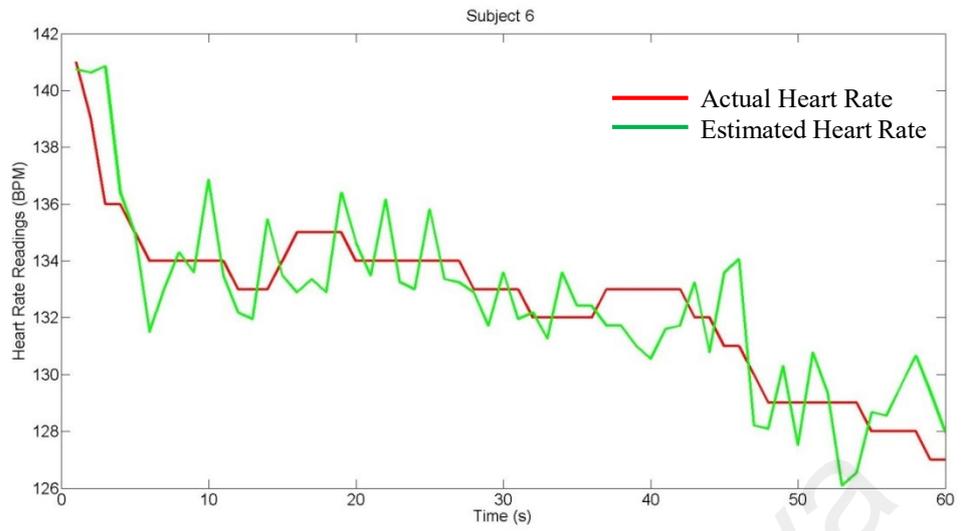


(d) Forth subject

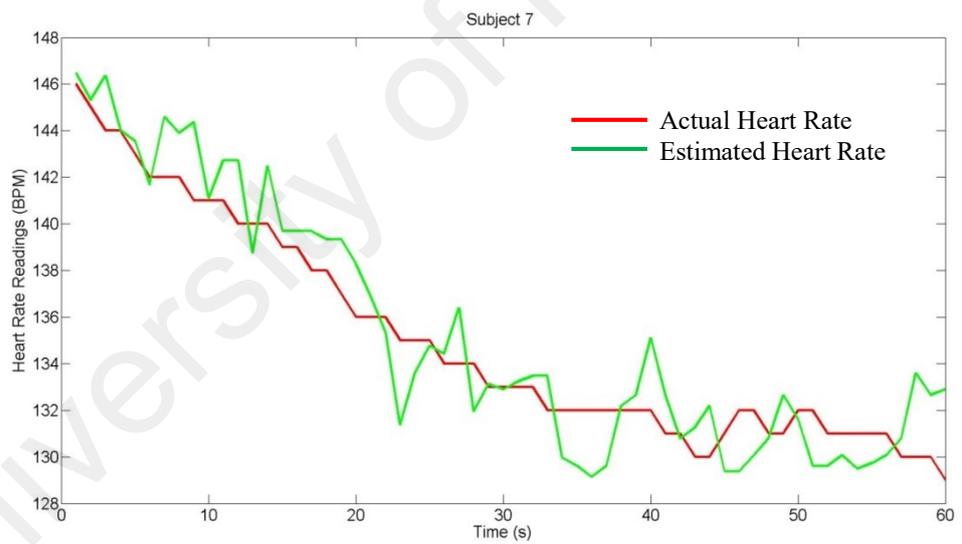


(e) Fifth subject

Figure 4.11.continued



(f) Sixth subject



(g) Seventh subject

Figure 4.11.continued

The Bland Altman plot as shown in Figure 4.12 indicates that a small number of computed heart rate readings are located outside the 95 % limit of agreement interval. The accuracy of computed heart rate readings in this experiment is better than the accuracy of the readings in the first experiment. It is because the subjects in the second experiment did not move as they were not cycling.

Table 4.3: Summary of heart rate readings results for the second experiment

Subject	Heart Rate Readings(BPM)		RMSE (BPM)
	Highest	Lowest	
1	108	88	1.97
2	150	114	1.49
3	105	98	2.14
4	153	119	1.73
5	153	127	1.58
6	141	127	1.57
7	146	129	1.82

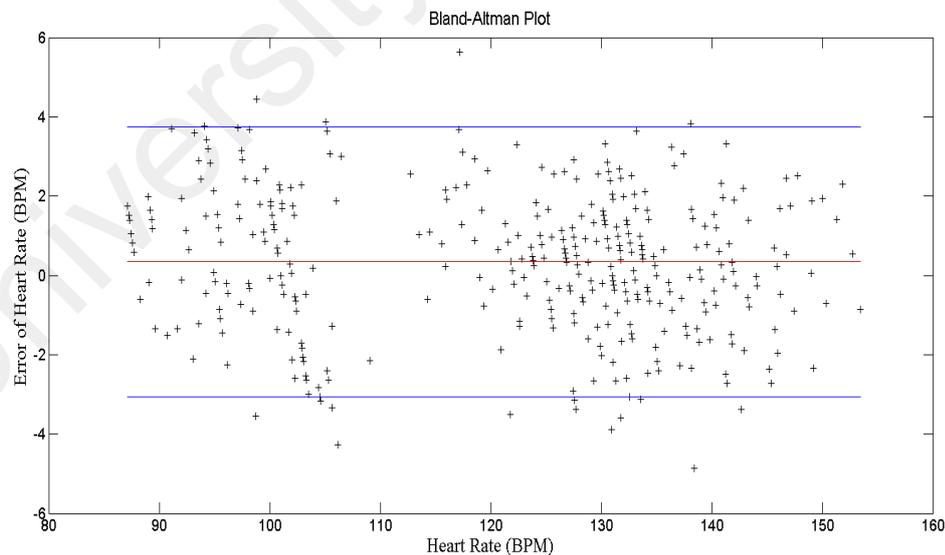


Figure 4.12: Bland-Altman plot for all estimated heart rate reading for the second experiment

4.3.4 Heart Rate Estimation for Subjects at Rest Using Proposed Method

In addition to the dynamic heart rate measurements, the proposed method can also be used for subjects whose heart rates are almost constant (non-dynamic). To do that, subjects were at rest for one minute and their measured and computed heart rate readings are shown in Figure 4.13. The root mean square error is 1.54 BPM while the Pearson correlation coefficient is 0.98.

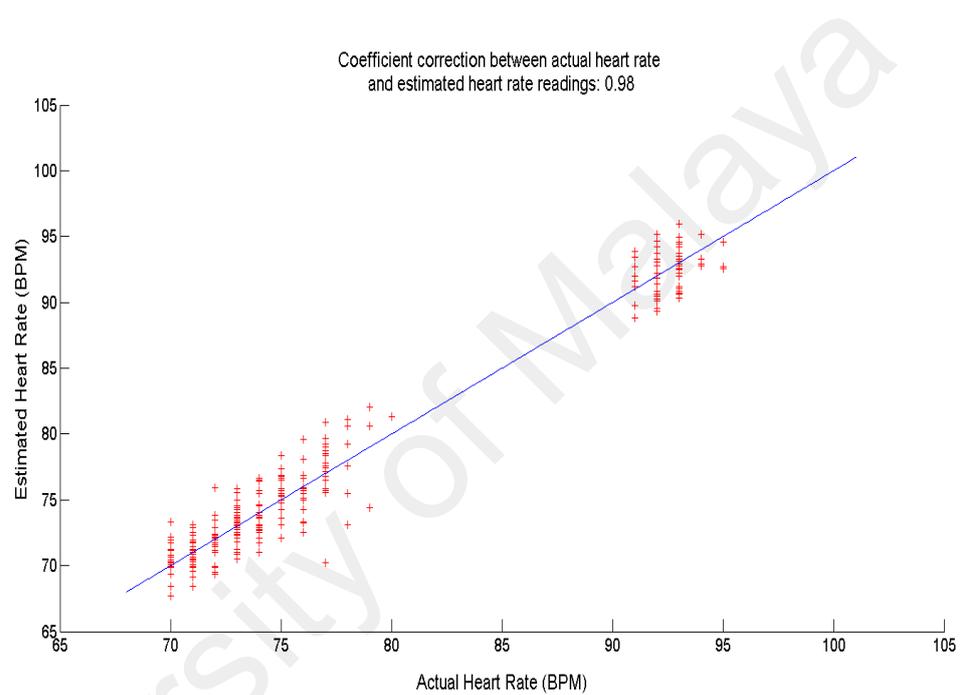


Figure 4.13: Comparison of all actual and estimated heart rate readings for subjects at rest

4.4 Chapter Conclusion

A new method of computing dynamic heart rate involving the use of short video clips has been proposed in this paper. Video clips are far easier to obtain and cheaper than the existing invasive or in-contact methods of obtaining the heart rates. In this study, it is observed that close to accurate readings can be obtained if the three ICA sources are independent of each other. The independence of the ICA sources needs to be established to ensure the reliability of the findings. For this, mutual information developed earlier was used. Two experiments were done to corroborate the validity of the proposed method

and the accuracy of its findings. These experiments show that the findings of this method agree with the findings of the established, and therefore accepted, method. The Bland-Altman plot shows that most of the findings of this study fall within the boundaries set for 95 % limit of agreement interval. The RMSE in both experiments are 1.97 BPM and 1.77 BPM, respectively.

University of Malaya

CHAPTER 5: DYNAMIC HEART RATE ESTIMATION USING PRINCIPAL COMPONENT ANALYSIS

5.1 Overview

In this study, a video camera is used to capture the facial images of seven subjects whose heart rates vary dynamically, between 81 and 153 BPM. Principal component analysis (PCA) is used to recover the blood volume pulses (BVP) which can be used for the heart rate estimation. An important consideration for accuracy of the dynamic heart rate estimation is to determine the shortest video duration that realizes it. This video duration is chosen when the six principal components (PC) are least correlated amongst them. When this is achieved, the first PC is used to obtain the heart rate. The results obtained from the proposed method are compared to the readings obtained from the Polar heart rate monitor. Experimental results show the proposed method is able to estimate the dynamic heart rate readings using less computational requirements when compared to the existing method. The mean absolute error and the standard deviation of the absolute errors between experimental readings and actual readings are 2.18 BPM and 1.71 BPM respectively.

5.2 Relationship between RGB and YCbCr Components of Facial Images

Human skin is composed of different layers (Krishnaswamy & Baranoski, 2004) and its color is highly related to melanin and hemoglobin concentrations. Xu *et al.* (Xu *et al.*, 2014) derived the relationship between the RGB pixel intensities obtained from a facial image and the hemoglobin and melanin concentrations, c_h and c_m , in the skin layer as:

$$\log P_R = -\{v_m(R)c_m + v_h(R)c_h + A_0(R)\} + \log kE(R), \quad (5.1)$$

$$\log P_G = -\{v_m(G)c_m + v_h(G)c_h + A_0(G)\} + \log kE(G), \quad (5.2)$$

$$\log P_B = -\{v_m(B)c_m + v_h(B)c_h + A_0(B)\} + \log kE(B). \quad (5.3)$$

where R,G and B represent the red, green and blue components of the image respectively while ν_h and ν_m are the products of pigment extinction coefficients of hemoglobin and melanin respectively, and their mean path length of photons in the skin layer. A_0 denotes the baseline skin absorbance while k is the constant for the camera gain and E is the power of incident light for each color component.

Considering the video is captured under constant background light and for a short duration, then the AC components of the RGB pixel intensities in log-space consist mostly of hemoglobin concentration. Since hemoglobin concentration is related to blood concentration, the frequency of this hemoglobin concentration is considered as the BVP, i.e. the heart rate pulse.

As (5.1-5.3) depend only on pigment concentration, baseline skin absorbance and the incident light, we may conclude that the RGB in log-space are correlated to each other. Figure 5.1 shows the distribution of RGB pixel intensities in log-space over a period of time used in our experiment. Table 5.1 shows the correlation among $\log P_R$, $\log P_G$, and $\log P_B$. The values in Table 5.1 show the RGB pixel intensities in log-space are highly correlated to each other. Therefore, PCA can be used to decorrelate these color components and recover the corresponding uncorrelated PCs.

However, for short video duration, the correlation amongst the PCs is still high which upon using them in the subsequent operation may lead to inaccurate heart rate readings. To address this issue, we add three more color components, i.e. luminance Y, chrominance C_b and C_r , in log-space as the input features. YCbCr components are correlated to the RGB components. These components can be derived from the RGB (Poyton, 1996) as follows:

$$Y = 16 + 65.481 \cdot R + 128.553 \cdot G + 24.966 \cdot B \quad (5.4)$$

$$C_b = 128 - 37.797 \cdot R - 74.203 \cdot G + 112 \cdot B \quad (5.5)$$

$$C_r = 128 + 112 \cdot R - 93.786 \cdot G - 18.214 \cdot B \quad (5.6)$$

With these six color components as input features, the corresponding PCs have much lower correlation as compared to the PCs whose input features are RGB components only. Figure 5.2 illustrates the correlation coefficients for six PCs and three PCs respectively. It shows that the correlation coefficient for the six PCs (green line) decreases as the video duration increases while the correlation coefficient for the three PCs (red line) is not consistent and relatively higher when compared to the six PCs. Therefore, in this study, we use PCA to recover the PPG signals from these six color components. After applying the PCA, the first PC that has the largest possible variance is considered as the PPG signal that consists of the hemoglobin concentration. The heart rate can be computed from this PPG signal.

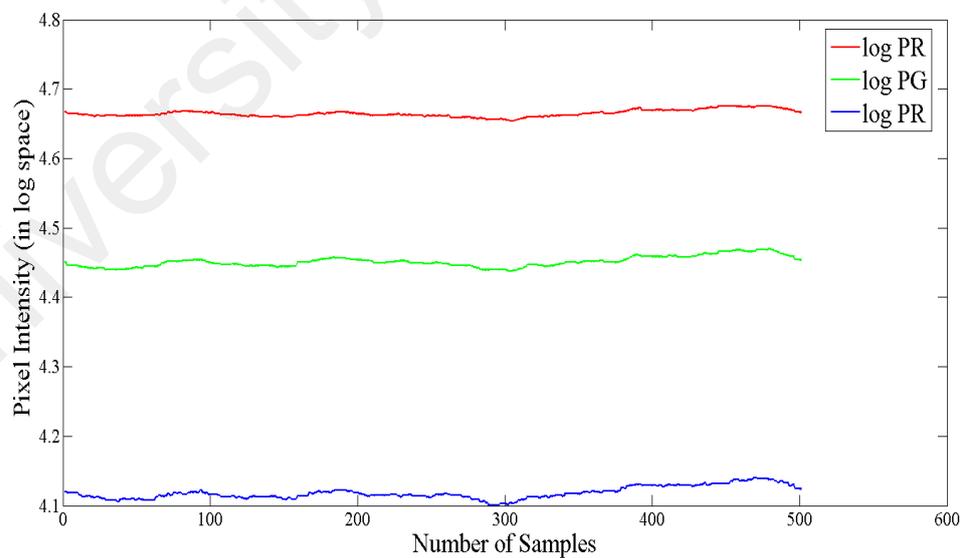


Figure 5.1: The distribution of $\log P_R$, $\log P_G$ and $\log P_B$

Table 5.1: Correlation coefficient among $\log P_R$, $\log P_G$ and $\log P_B$

	$\log P_R$	$\log P_G$	$\log P_B$
$\log P_R$	1.00	0.92	0.84
$\log P_G$	0.92	1.00	0.93
$\log P_B$	0.84	0.93	1.00

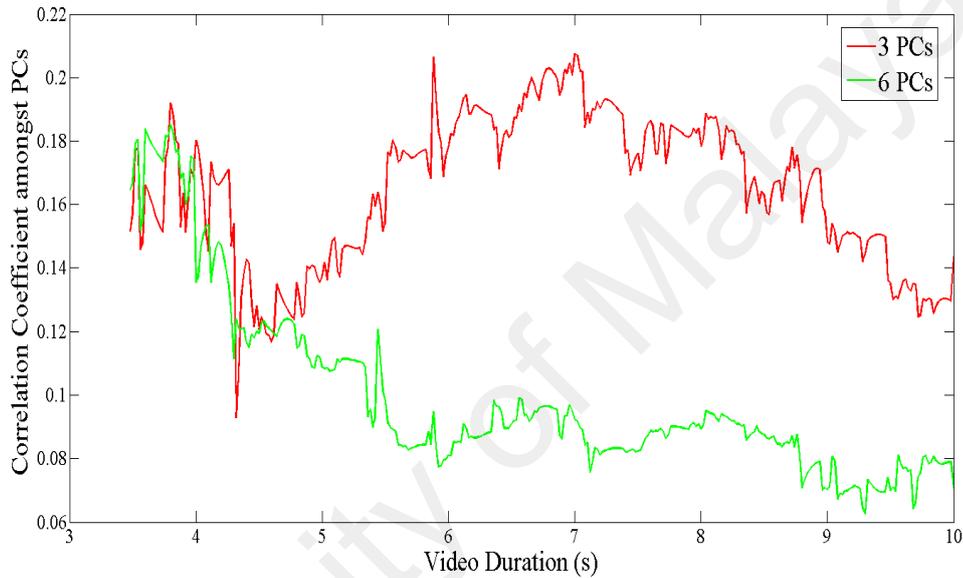


Figure 5.2: The graph of correlation coefficient amongst PCs vs video duration for 3 PCs and 6 PCs respectively

5.3 Proposed Method

In this section, the proposed model to estimate the dynamic heart rate measurements using PCA is presented. The relationship between the correlation among PCs, video duration and heart rate accuracy is also discussed. As the video duration affects the accuracy of the heart rate reading, a stopping criterion is set to determine the video duration needed for dynamic heart rate estimation. The details are described in this section.

5.3.1 Relationship between the correlation among PCS and video duration

Ideally, PCA will compute its PCs by maximizing the correlation among the input features. However, for short video duration, the PCs may still have high correlation. Hence, it is important to find out the minimum video duration that gives the least correlation. We use Pearson correlation coefficient to determine the correlation between any two PCs. For any given two PCs x and y , the Pearson correlation coefficient of these two PCs, R is given as:

$$R(x, y) = \frac{C(x, y)}{\sqrt{C(x, x)C(y, y)}} \quad (5.7)$$

where $C(x, y)$ is the covariance of PCs x and y , $C(x, x)$ and $C(y, y)$ are the variances of PCs x and y respectively. Since six PCs are recovered from the PCA, the averaged correlation coefficient amongst PCs, R_{avg} is computed using

$$R_{avg} = \frac{1}{\binom{6}{2}} \sum_{m=2}^6 \sum_{n=1}^m R(m, n) \quad (5.8)$$

Figure 5.3 illustrates the relationship between the averaged correlation coefficient R_{avg} and video duration for a particular heart rate reading used in the experiment. A power function curve is fitted to represent the function R_{avg} . It is found that the value of the R_{avg} decreases significantly at the beginning, but remains almost constant when the video duration exceeds a specific duration. The value of the R_{avg} varies very little after this duration. Hence, the stopping criterion to determine the video duration is set as when the difference of R_{avg} for 3 continuous video frames is smaller than 2×10^{-4} .

To illustrate how correlated PCs affect the accuracy of computed heart rate readings, two points X and Y are selected in Figure 5.3. The actual heart rate reading for this

particular instant is 143 BPM. Point X represents a very short video duration where the R_{avg} doesn't meet the stopping criterion. Point Y represents the suitable video duration where the R_{avg} has met the stopping criterion. Point Y gives more accurate heart rate estimation, i.e 142.38 BPM as compared to point X that gives 63.75 BPM. When the stopping criterion is met, the corresponding video duration is used to compute the instantaneous heart rate for that particular instant.

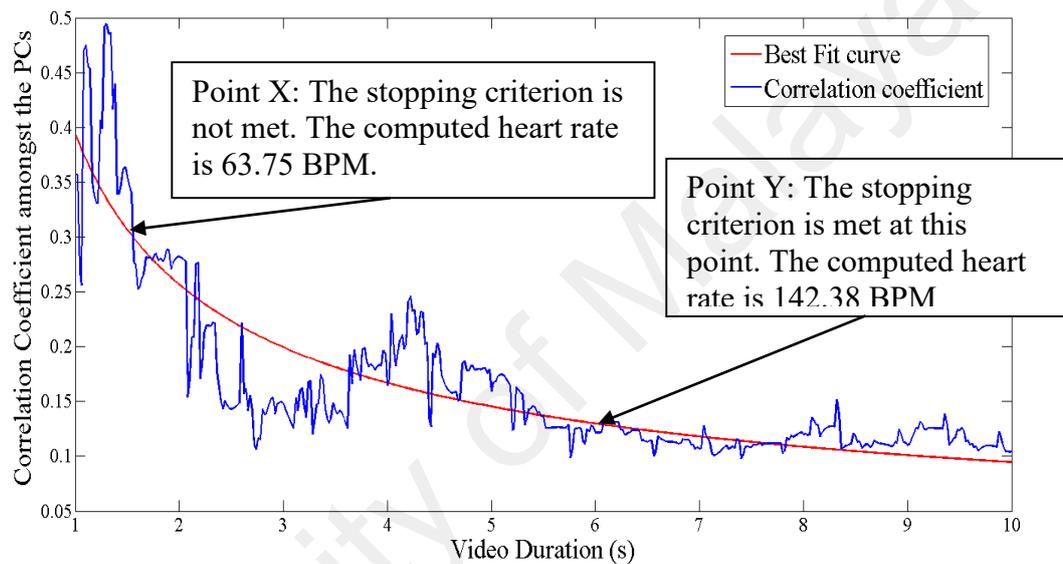


Figure 5.3: The relationship of the averaged correlation coefficient R_{avg} and the video duration and the respective computed heart rate

5.3.2 Block diagram the proposed model

The block diagram of the proposed model is illustrated in Figure 5.4. The face region is identified by using the model described in (Viola and Jones, 2001) and the region of interest (ROI) is fixed at the area below eyes and above the upper lip of mouth. For each frame, the spatially average of the RGB and YCbCr components, i.e.: $\mu_R, \mu_G, \mu_B, \mu_Y, \mu_{Cb}$, and μ_{Cr} are computed respectively. All six color components are projected into log-space. Therefore, at any time instant, a set of six input features $\log P_R, \log P_G, \log P_B, \log P_Y, \log P_{Cb}$ and $\log P_{Cr}$ are formed. The set of input features are then detrended using the model developed by Tarvainen *et al.* (Tarvainen *et al.*, 2002). PCA is then used to recover

six PCs from these six input features. The set of PCs is bandpass filtered (128-point Hamming window, 0.8-4 Hz).

The entire process is repeated by increasing the number of previous video frames, until the stopping criterion described in *Section 5.3.1* is met. At this point, the corresponding number of frames is chosen as the video duration needed to compute the instantaneous heart rate reading. The first PC is then chosen as the PPG signal. The corresponding frequency of this PPG signal is considered as the instantaneous heart rate reading for that particular instant.

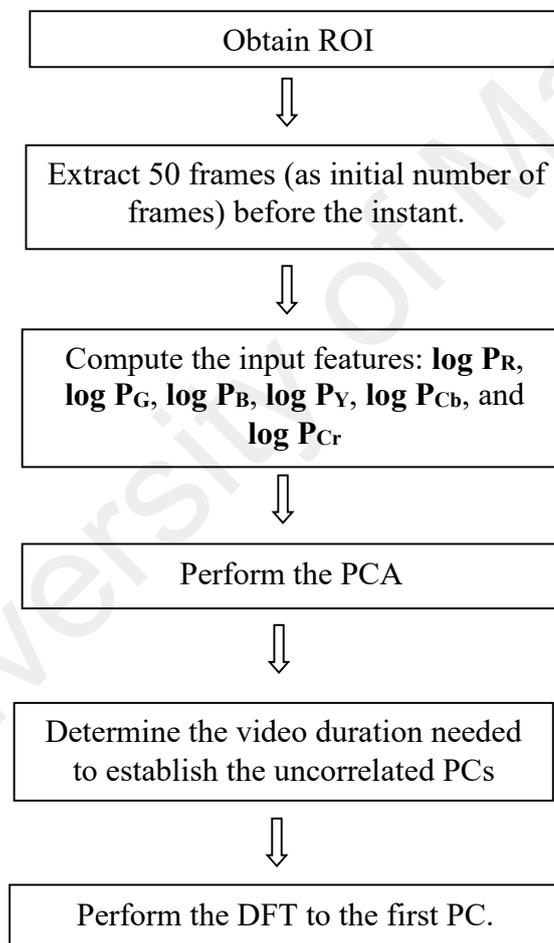


Figure 5.4: Flow Chart of the proposed method

5.4 Experimental Study

In this section, the experimental setup and the experimental results are discussed and analysed. A comparative study between the proposed method and the method described in Chapter 4 is also presented.

5.4.1 Experimental Setup

All experiments were set up under constant office fluorescent light. A Sony camcorder (HDR-PJ260VE) was used for the video recording purposes. All videos were recorded and sampled at 50 frames per second. The camcorder was fixed at a position with a distance of about 0.60 m from the subject's face. In the experiments, seven subjects were selected and requested to carry out a cycling activity. In the first experiment, four subjects were asked to cycle at different speeds for about two minutes. Then they were asked to stop for one minute. The camcorder was used to capture their facial images during that time. In the second experiment, the remaining three subjects were asked to cycle continuously and their facial images were captured by the camcorder for one minute. An increasing heart rate trend was observed. Throughout the video recordings, all subjects were asked to remain stationary. Sixty heart rate readings (sampled at each second) were computed for every subject.

As reference, the instantaneous heart rates of each subject that obtained from the proposed method were compared to the actual heart rate readings measured from Polar Heart Rate Monitor – Polar Team2 Pro. Polar Team2 Pro samples and computes the instantaneous heart rate by measuring at least one ECG signal waveform, as described in the patents (Heikkila, 1998; Pietila & Tammi, 1997).

5.4.2 Experimental Results and Analysis

A total of 420 instantaneous heart rate readings were obtained from this experiment. In the experiment, the subjects' heart rates were varying between 81 BPM and 153 BPM.

Table 5.2 summarizes the details of the computed heart rate readings of all subjects. The highest and the lowest mean absolute errors are 2.85 and 1.37 BPM. Figure 5.5 shows the scattered plot of all computed and actual heart rate readings. It shows that the computed heart rate readings are closely correlated to the actual heart rate readings. The correlation coefficient between the computed and actual heart rate readings is 0.99. The mean absolute error for all readings is 2.06 BPM while the standard deviation of absolute errors is 1.65 BPM. Figure 5.6 shows the comparison of actual and computed heart rate readings for each subject in the first experiment. Figure 5.7 shows the comparison of actual and computed heart rate readings for each subject in the second experiment. Red color represents the actual heart rate readings while green color represents the estimated heart rate readings. The Bland Altman plot is shown in Figure 5.8. It shows that only a small number of computed heart rate readings are located outside the 95 % limit of agreement interval.

Table 5.2: Summary of heart rate readings results obtained from proposed method

Subject	Heart Rate Readings (BPM)		Mean absolute error (BPM)	Standard deviation of absolute errors (BPM)
	Highest	Lowest		
1	141	127	1.37	1.02
2	134	122	1.91	2.09
3	105	96	1.88	1.21
4	153	119	1.42	0.99
5	108	81	2.32	1.54
6	153	127	2.65	2.00
7	133	105	2.85	1.77

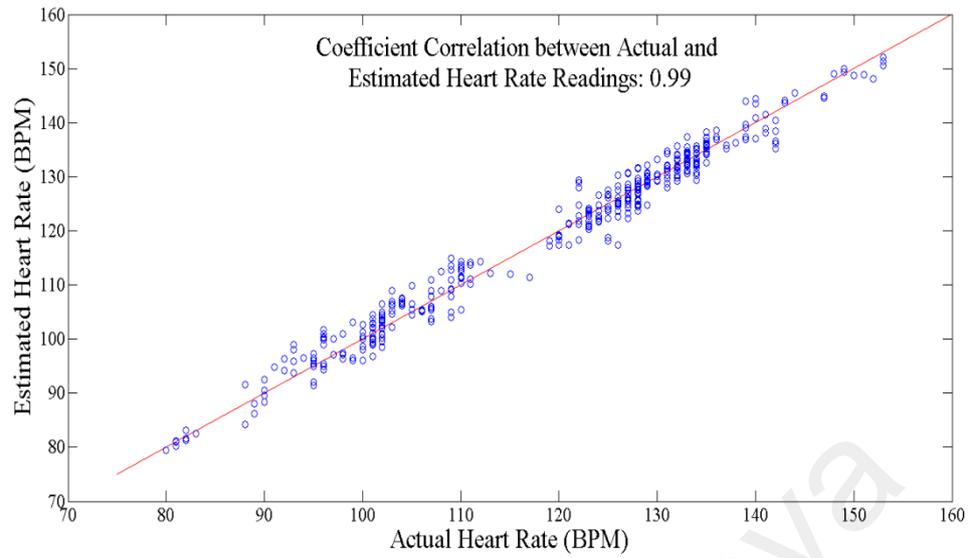
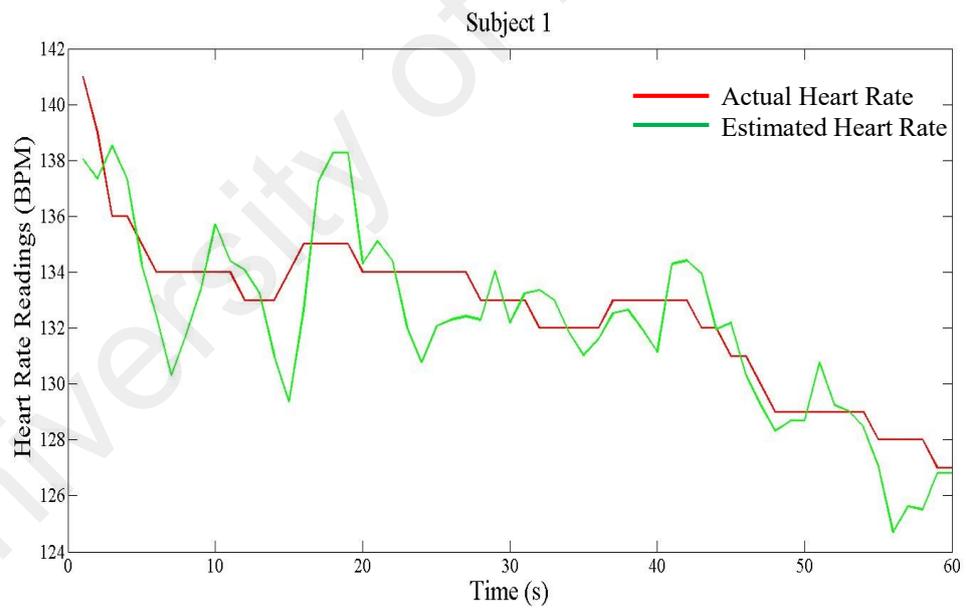
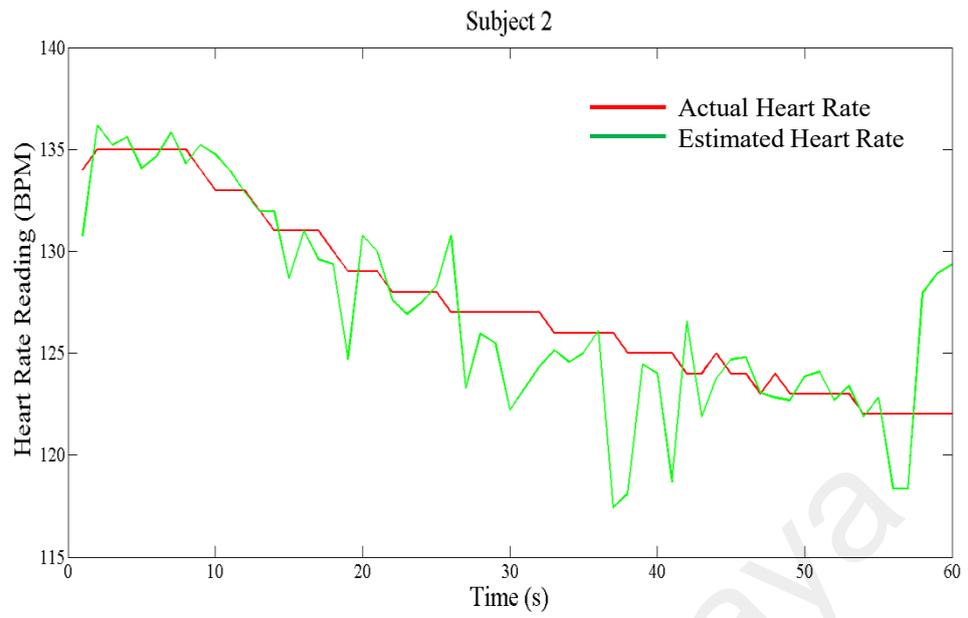


Figure 5.5: Comparison of all actual and estimated heart rate readings

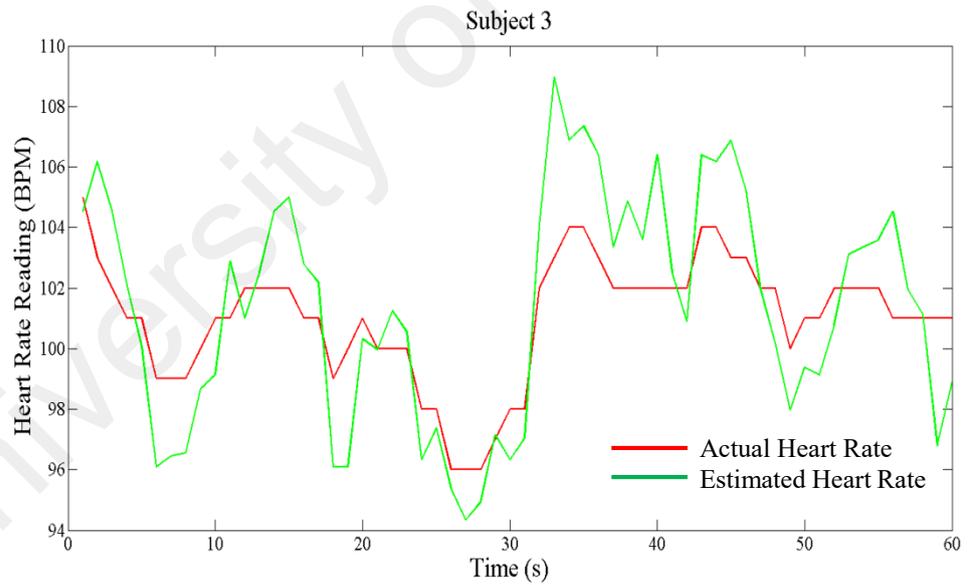


(a) First subject

Figure 5.6: Comparison of actual and estimated heart rate readings for each subject in the first experiment

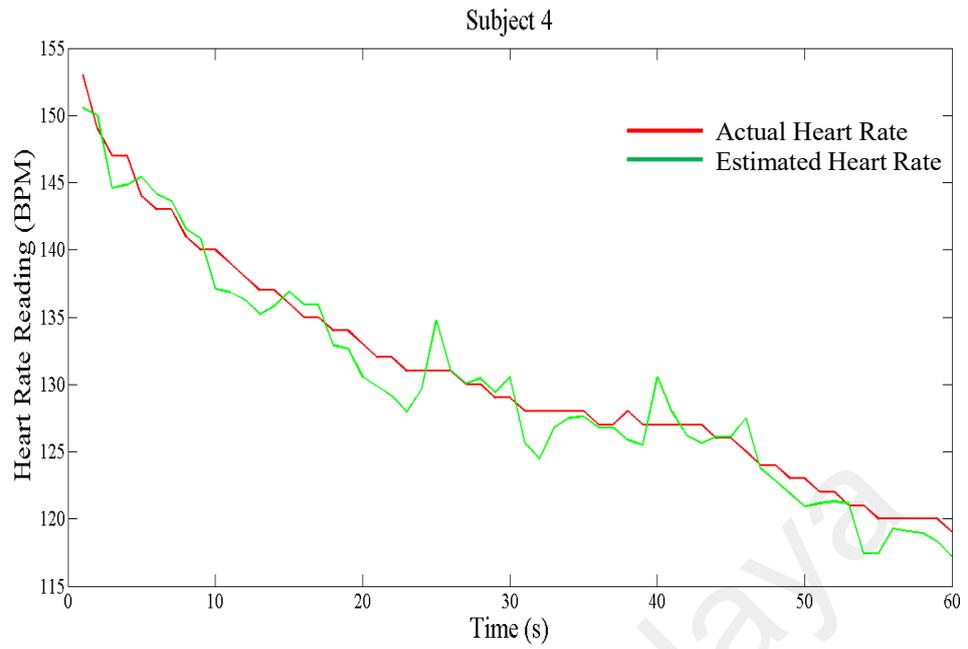


(b) Second subject

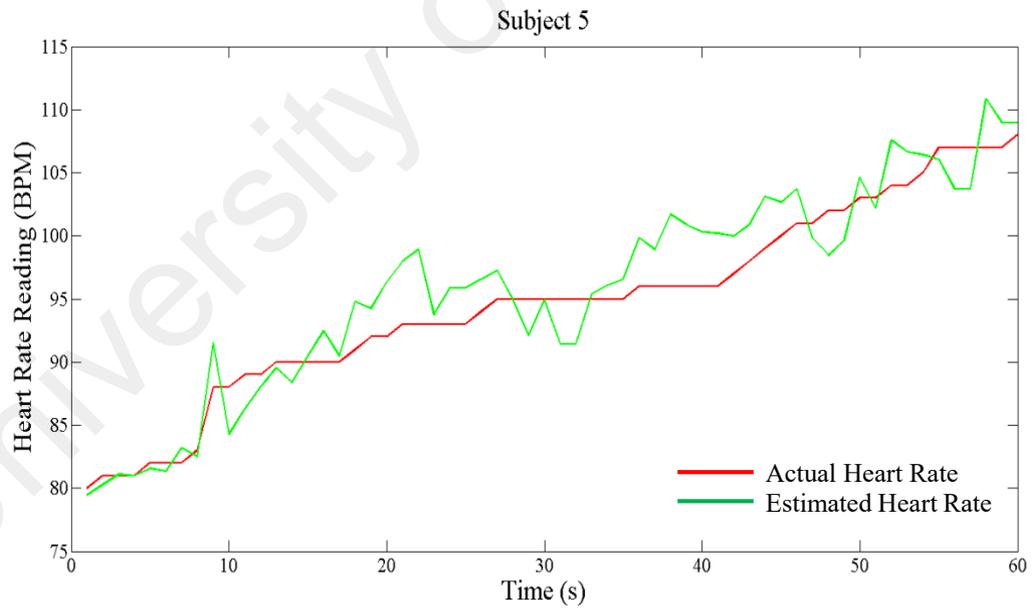


(c) Third subject

Figure 5.6.continued

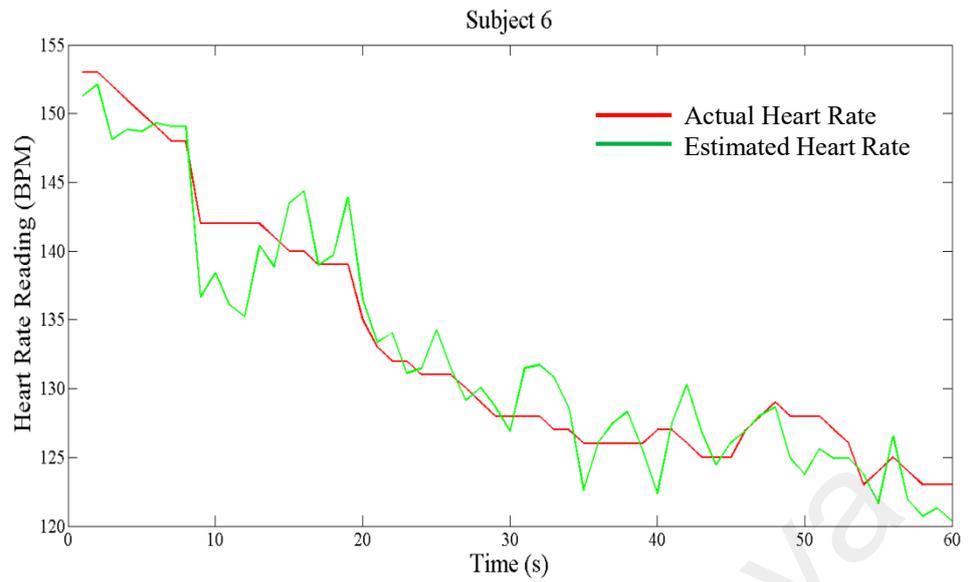


(d) Forth subject

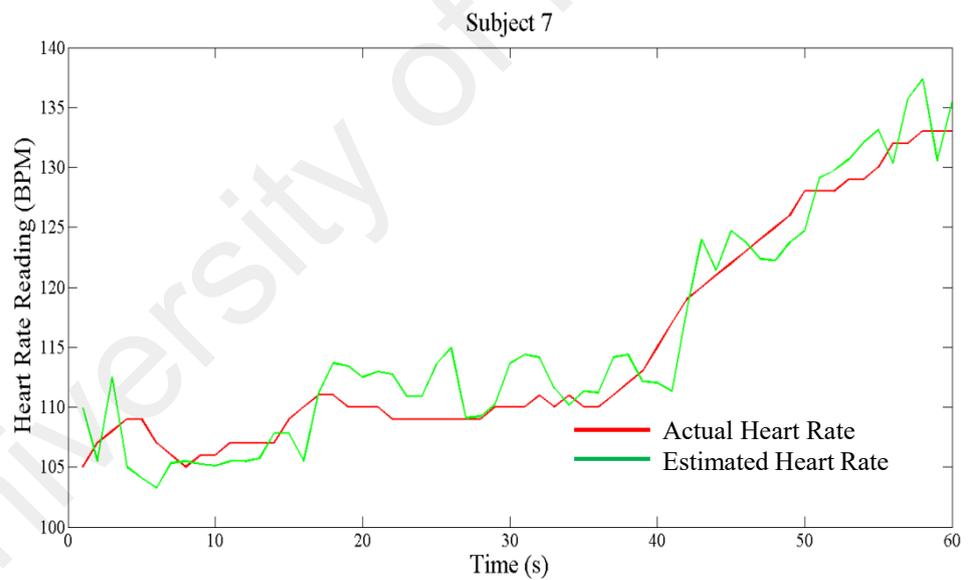


(a) Fifth subject

Figure 5.7: Comparison of actual and estimated heart rate readings for each subject in the second experiment



(b) Sixth subject



(c) Seventh subject

Figure 5.7.continued

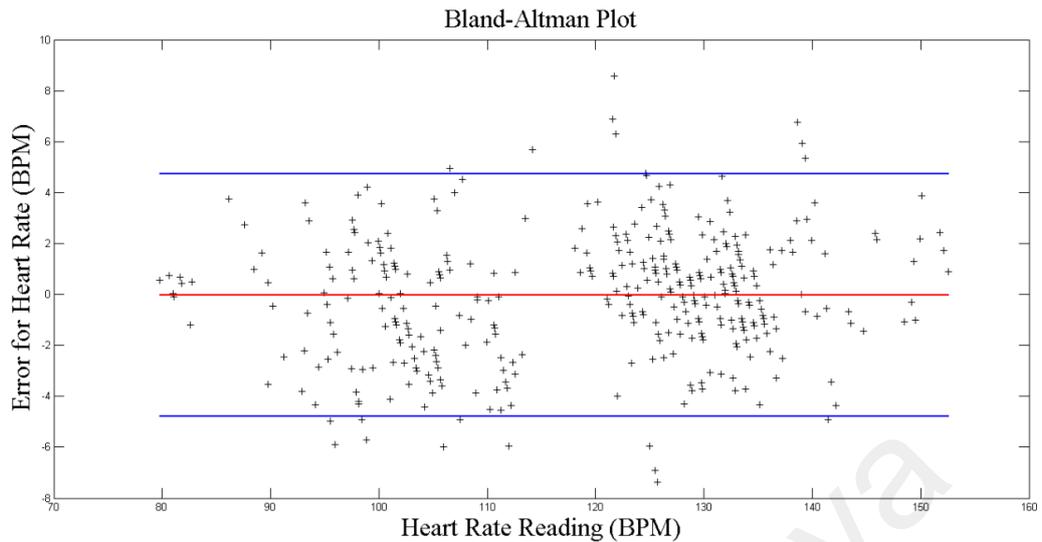


Figure 5.8: Bland-Altman plot for all estimated heart rate readings

5.4.3 Comparative study between proposed method and existing method

A comparative study has been done to compare the accuracy (mean error and standard deviation of error), video duration for the heart rate computation, and the computational cost of using the method described in this paper and the method described in Chapter 4. To calculate the computational cost, both ICA (method described in Chapter 4) and PCA computations are repeated for 1000 times and the average time taken is recorded. Table 5.3 summarizes the results of the comparative study. As can be seen in Table 5.3, both accuracy and video duration are not much different for these two methods. However, in terms of the computational cost, the proposed method is much more efficient than the method described in Chapter 4. Additionally, the proposed method directly uses the first PCs to compute the heart rate while the method described in Chapter 4 investigated all ICA sources first and then chose the source with high peak in frequency domain as the source giving heart rate information. As low computational cost and small memory resources are important factors for the eventual implementation in mobile phones, the proposed method is more efficient than the previous method.

Table 5.3: Comparison of proposed method (using PCA) and method described in Chapter 4 (using ICA)

	PCA	ICA
Mean absolute error	2.06	1.64
Standard deviation of absolute errors	1.65	1.48
Video duration (mean), in second	5.33	5.49
Computational cost (mean), in millisecond	2.25	6.62

5.5 Chapter Conclusion

In this study, it is found that heart rate readings can be obtained by applying PCA to the facial images. When the PCs are uncorrelated to each other, then an accurate reading can be obtained. An important consideration for dynamic heart rate estimation is the need for video duration. Instead of using RGB components only, three additional components, YCbCr are used. In doing so, a shorter video duration is obtained. To ensure the reliability of the heart rate estimation, the PCs must have least correlation. To validate the criterion, Pearson correlation coefficient is used. Experimental results show that this method is able to estimate dynamic heart rates from short video sequences using less computational requirements when compared to the method described in Chapter 4.

CHAPTER 6: CONCLUSION AND FUTURE WORK

This thesis presents several approaches to estimate the dynamic heart rates using facial images from video sequences. For heart rates that do not vary much throughout the entire video, the heart rate information or PPG signal can be obtained through the spatial information. However, for dynamic heart rate variation, localized temporal information is needed to determine the instantaneous heart rate reading at any particular instant. To extract the temporal information, STFT and filter bank are used. Experimental studies show that both STFT and filter bank are able to give temporal information and a better accuracy. The root mean square errors of the heart rate readings obtained from the experiments are less than 4 seconds.

However, for the above approaches, the video duration used is too long. For dynamic heart rate estimation, short video duration is needed. The challenge of using short video duration is the accuracy of heart rate readings may be compromised. To address this issue, the thesis presents a method that uses independent component analysis (ICA) combined with mutual information to estimate the dynamic heart rate readings from short video duration without compromising the accuracy. For ICA blind source separation, the ICA sources that are separated from short video duration may have insufficient independence among themselves. This thesis explains the need to determine the earliest sign of the independence of the ICA sources which corresponds to the minimum video duration that gives accurate heart rate estimation for all heart rate readings. Experimental results show the proposed methods work for both scenarios, i.e. measurement of subjects' increasing heart rates and falling heart beats. The root mean square errors are less than 3 BPM.

Another approach that uses PCA to estimate dynamic heart rate readings is also presented in this thesis. It is found that the color components (in log-space) obtained from facial images are correlated to each other. PCA is used to decorrelate the color

components in log-space and recover the PPG signals. Similarly, the issue of using short video duration is the PCs recovered using PCA may not have sufficient uncorrelatedness. To address this issue, this thesis presented a method that uses PCA with Pearson correlation coefficient to obtain the minimum video duration that still gives sufficient uncorrelatedness and accurate heart rate readings. Experimental results show the proposed methods work for both scenarios, i.e. measurement of subjects' increasing heart rates and falling heart beats, where the errors for the readings are less than 3 BPM. The advantage of using this method as compared to the method described earlier (by using ICA with mutual information) is the high efficiency of computational costs.

There are still many possibilities to extend and improve the works presented in this thesis. For all data collection and experimental works presented in this thesis, there are a few conditions and requirements. Firstly, the background light is homogenous and constant. Future work may explore and investigate the impact of heterogeneous background lighting effects onto the computation of dynamic heart rate readings. The experiments may be conducted under different environments such as under hot sun in the afternoon or under a weak light condition at night.

Second requirement for the experiments done in this study is subjects were asked to remain stationary while the video was captured. Future work may explore the impact of the motion artifact due to the large movement of subjects and the camera vibration during the video recording. A model may be designed to compensate the losses due to motion artifact and to recover the PPG signals from human facial images.

Lastly, future work may be extended to the dynamic heart rate measurements from human skin other than face, for instance, wrist, arm, or palm. It is more practical if human heart rate can be obtained using any human skin images from video sequences.

REFERENCES

- Bae, K., Noh, S., & Kim, J. (2003, June). Iris feature extraction using independent component analysis. In *Audio-and Video-Based Biometric Person Authentication* (pp. 838-844). Springer Berlin Heidelberg.
- Balakrishnan, G., Durand, F., & Guttag, J. (2013). Detecting pulse from head motions in video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3430-3437).
- Bartlett, M. S., Movellan, J. R., & Sejnowski, T. J. (2002). Face recognition by independent component analysis. *Neural Networks, IEEE Transactions on*, 13(6), 1450-1464.
- Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *Medical Imaging, IEEE Transactions on*, 23(2), 137-152.
- Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical methods in medical research*, 8(2), 135-160.
- Calhoun, V. D., & Adali, T. (2006). Unmixing fMRI with independent component analysis. *Engineering in Medicine and Biology Magazine, IEEE*, 25(2), 79-90.
- Cardoso, J. F. (1999). High-order contrasts for independent component analysis. *Neural computation*, 11(1), 157-192.
- Cardoso, J. F., & Souloumiac, A. (1993, December). Blind beamforming for non-Gaussian signals. In *IEE Proceedings F (Radar and Signal Processing)* (Vol. 140, No. 6, pp. 362-370). IET Digital Library.
- Chang, H., & Yeung, D. Y. (2006). Robust locally linear embedding. *Pattern recognition*, 39(6), 1053-1065.
- Cover, T. M., & Thomas, J. A. (2012). *Elements of information theory*. John Wiley & Sons.
- Dawson, J. B., Barker, D. J., Ellis, D. J., Cotterill, J. A., Grassam, E., Fisher, G. W., & Feather, J. W. (1980). A theoretical and experimental study of light absorption and scattering by in vivo skin. *Physics in medicine and biology*, 25(4), 695.

- Déniz, O., Castrillon, M., & Hernández, M. (2001, June). Face Recognition Using Independent Component Analysis and Support Vector Machines*. In *Audio-and Video-Based Biometric Person Authentication* (pp. 59-64). Springer Berlin Heidelberg.
- Finlayson, G. D., Drew, M. S., & Lu, C. (2004). Intrinsic images by entropy minimization. In *Computer Vision-ECCV 2004* (pp. 582-595). Springer Berlin Heidelberg.
- Garbey, M., Sun, N., Merla, A., & Pavlidis, I. (2007). Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. *Biomedical Engineering, IEEE Transactions on*, 54(8), 1418-1426.
- Gatto, R. G. (2009). *Estimation of instantaneous heart rate using video infrared thermography and ARMA models*. University of Illinois at Chicago.
- Heikkila, I. (1998). *U.S. Patent No. 5,840,039*. Washington, DC: U.S. Patent and Trademark Office.
- Hertzman, A. B. (1938). The blood supply of various skin areas as estimated by the photoelectric plethysmograph. *Am J Physiol*, 124(2), 328.
- Hoyer, P. O., & Hyvärinen, A. (2000). Independent component analysis applied to feature extraction from colour and stereo images. *Network: Computation in Neural Systems*, 11(3), 191-210.
- Hyvärinen, A., & Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural networks*, 13(4), 411-430.
- James, C. J., & Hesse, C. W. (2004). Independent component analysis for biomedical signals. *Physiological measurement*, 26(1), R15.
- John, S. (2002). A tutorial on principal components analysis. Princeton University, 1, 1-16.
- Jonathan, E., & Leahy, M. (2010). Investigating a smartphone imaging unit for photoplethysmography. *Physiological measurement*, 31(11), N79.
- Kamshilin, A. A., Miridonov, S., Teplov, V., Saarenheimo, R., & Nippolainen, E. (2011). Photoplethysmographic imaging of high spatial resolution. *Biomedical optics express*, 2(4), 996-1006.
- Kim, K. I., Jung, K., & Kim, H. J. (2002). Face recognition using kernel principal component analysis. *Signal Processing Letters, IEEE*, 9(2), 40-42.

- Krishnaswamy, A., & Baranoski, G. V. (2004). A study on skin optics. Natural Phenomena Simulation Group, School of Computer Science, University of Waterloo, Canada, Technical Report, 1, 1-17.
- Kumar, M., Veeraraghavan, A., & Sabharwal, A. (2015). DistancePPG: Robust non-contact vital signs monitoring using a camera. *Biomedical optics express*, 6(5), 1565-1588.
- Lee, T. W., Lewicki, M. S., & Sejnowski, T. J. (2000). ICA mixture models for unsupervised classification of non-gaussian classes and automatic context switching in blind signal separation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(10), 1078-1089.
- Lindsay, I. S. (2002). A tutorial on principal components analysis. University of Otago, 1, 1-27.
- Liu, C., & Wechsler, H. (2003). Independent component analysis of Gabor features for face recognition. *Neural Networks, IEEE Transactions on*, 14(4), 919-928.
- Liu, X., Wang, D., Liu, F., & Bai, J. (2010). Principal component analysis of dynamic fluorescence diffuse optical tomography images. *Optics express*, 18(6), 6300-6314.
- Lu, H., Plataniotis, K. N., & Venetsanopoulos, A. N. (2008). MPCA: Multilinear principal component analysis of tensor objects. *Neural Networks, IEEE Transactions on*, 19(1), 18-39.
- Martis, R. J., Acharya, U. R., Mandana, K. M., Ray, A. K., & Chakraborty, C. (2012). Application of principal component analysis to ECG signals for automated diagnosis of cardiac health. *Expert Systems with Applications*, 39(14), 11792-11800.
- Martis, R. J., Acharya, U. R., & Min, L. C. (2013). ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform. *Biomedical Signal Processing and Control*, 8(5), 437-448.
- Mittal, N., & Walia, E. (2008, May). Face recognition using improved fast PCA algorithm. In *Image and Signal Processing, 2008. CISP'08. Congress on* (Vol. 1, pp. 554-558). IEEE.
- Naikal, N., Yang, A. Y., & Sastry, S. S. (2011, November). Informative feature selection for object recognition via sparse PCA. In *Computer Vision (ICCV), 2011 IEEE International Conference on* (pp. 818-825). IEEE.

- Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7), 971-987.
- Pavlidis, I., Dowdall, J., Sun, N., Puri, C., Fei, J., & Garbey, M. (2007). Interacting with human physiology. *Computer Vision and Image Understanding*, 108(1), 150-170.
- Pietila, A., & Tammi, T. (1997). *U.S. Patent No. 5,622,180*. Washington, DC: U.S. Patent and Trademark Office.
- Poh, M. Z., McDuff, D. J., & Picard, R. W. (2010). Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 18(10), 10762-10774.
- Poh, M. Z., McDuff, D. J., & Picard, R. W. (2011). Advancements in noncontact, multiparameter physiological measurements using a webcam. *Biomedical Engineering, IEEE Transactions on*, 58(1), 7-11.
- Poynton, C. A. (1996). *A technical introduction to digital video*. John Wiley & Sons, Inc.
- Pursche, T., Krajewski, J., & Moeller, R. (2012, January). Video-based heart rate measurement from human faces. In *Consumer Electronics (ICCE), 2012 IEEE International Conference on* (pp. 544-545). IEEE.
- Pyatykh, S., Hesser, J., & Zheng, L. (2013). Image noise level estimation by principal component analysis. *Image Processing, IEEE Transactions on*, 22(2), 687-699.
- Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., & Smith, S. M. (2014). Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage*, 90, 449-468.
- Schönfelder, M., Hinterseher, G., Peter, P., & Spitzenpfeil, P. (2011). Scientific comparison of different online heart rate monitoring systems. *International journal of telemedicine and applications*, 2011, 6.
- Sharma, L. N., Dandapat, S., & Mahanta, A. (2012). Multichannel ECG data compression based on multiscale principal component analysis. *Information Technology in Biomedicine, IEEE Transactions on*, 16(4), 730-736.
- Shi, P., Peris, V. A., Echiadis, A., Zheng, J., Zhu, Y., Cheang, P. Y. S., & Hu, S. (2010). Non-contact reflection photoplethysmography towards effective human physiological monitoring. *J. Med. Biol. Eng*, 30(3), 161-167.

- Takano, C., & Ohta, Y. (2007). Heart rate measurement based on a time-lapse image. *Medical engineering & physics*, 29(8), 853-857.
- Tarvainen, M. P., Ranta-aho, P. O., & Karjalainen, P. A. (2002). An advanced detrending method with application to HRV analysis. *IEEE Transactions on Biomedical Engineering*, 49(2), 172-175.
- Tsumura, N., Ojima, N., Sato, K., Shiraishi, M., Shimizu, H., Nabeshima, H., ... & Miyake, Y. (2003). Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. *ACM Transactions on Graphics (TOG)*, 22(3), 770-779.
- Vargas, J., Quiroga, J. A., & Belenguer, T. (2011). Phase-shifting interferometry based on principal component analysis. *Optics letters*, 36(8), 1326-1328.
- Vázquez, R. R., Velez-Perez, H., Ranta, R., Dorr, V. L., Maquin, D., & Maillard, L. (2012). Blind source separation, wavelet denoising and discriminant analysis for EEG artefacts and noise cancelling. *Biomedical Signal Processing and Control*, 7(4), 389-400.
- Verkruyse, W., Svaasand, L. O., & Nelson, J. S. (2008). Remote plethysmographic imaging using ambient light. *Optics express*, 16(26), 21434-21445.
- Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on* (Vol. 1, pp. I-511). IEEE.
- Wallén, M. B., Hasson, D., Theorell, T., Canlon, B., & Osika, W. (2012). Possibilities and limitations of the polar RS800 in measuring heart rate variability at rest. *European journal of applied physiology*, 112(3), 1153-1165.
- Wang, H. Y., & Wu, X. J. (2005, August). Weighted PCA space and its application in face recognition. In *Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on* (Vol. 7, pp. 4522-4527). IEEE.
- Wang, J., & Chang, C. I. (2006). Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *Geoscience and Remote Sensing, IEEE Transactions on*, 44(6), 1586-1600.

- Weinman, J., Hayat, A., & Raviv, G. (1977). Reflection photoplethysmography of arterial-blood-volume pulses. *Medical and Biological Engineering and Computing*, 15(1), 22-31.
- Woo, Y., Yi, C., & Yi, Y. (2013, May). Fast PCA-based face recognition on GPUs. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on* (pp. 2659-2663). IEEE.
- Xu, S., Sun, L., & Rohde, G. K. (2014). Robust efficient estimation of heart rate pulse from video. *Biomedical optics express*, 5(4), 1124-1135.
- Xu, S., Ye, X., Wu, Y., Giron, F., Leveque, J. L., & Querleux, B. (2008). Automatic skin decomposition based on single image. *Computer Vision and Image Understanding*, 110(1), 1-6.
- Yuen, P. C., & Lai, J. H. (2002). Face representation using independent component analysis. *Pattern recognition*, 35(6), 1247-1257.
- Zhang, D., Zhou, Z. H., & Chen, S. (2006). Diagonal principal component analysis for face recognition. *Pattern Recognition*, 39(1), 140-142.
- Zhang, L., Dong, W., Zhang, D., & Shi, G. (2010). Two-stage image denoising by principal component analysis with local pixel grouping. *Pattern Recognition*, 43(4), 1531-1549.
- Zhang, X. P., & Chen, Z. (2006). An automated video object extraction system based on spatiotemporal independent component analysis and multiscale segmentation. *EURASIP Journal on Advances in Signal Processing*, 2006(1), 1-22.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

- Yu, Y. P.**, Kwan, B. H., Lim, C. L., Wong, S. L., & Raveendran, P. (2013, November). Video-based heart rate measurement using short-time Fourier transform. In *Intelligent Signal Processing and Communications Systems (ISPACS), 2013 International Symposium on* (pp. 704-707). IEEE.
- Yu, Y. P.**, Raveendran, P., & Lim, C. L. (2014, May). Heart rate estimation from facial images using filter bank. In *Communications, Control and Signal Processing (ISCCSP), 2014 6th International Symposium on* (pp. 69-72). IEEE.
- Yu, Y. P.**, Paramesran, R., & Lim, C. L. (2014, December). Video based heart rate estimation under different light illumination intensities. In *Intelligent Signal Processing and Communication Systems (ISPACS), 2014 International Symposium on* (pp. 216-221). IEEE.
- Yu, Y. P.**, Raveendran, P., & Lim, C. L. (2015). Dynamic heart rate measurements from video sequences. *Biomedical optics express*, 6(7), 2466-2480.
- Yu, Y. P.**, Raveendran, P., Lim, C. L., & Kwan, B. H. (2015). Dynamic heart rate estimation using principal component analysis. *Biomedical optics express*, 6(11), 4610-4618.

APPENDIX A

The following codes are the main parts of the algorithms used in the models described in Chapter 3.

```
%% Filter Bank

load('TimeData.mat')

% refer to (Tarvainen et al., 2002)

T = length(TimeData);
lambda = 50;
I = speye(T);
D2 = spdiags(ones(T-2,1)*[1 -2 1],[0:2],T-2,T);
AvgRGBDeTrend = (I-inv(I+lambda^2*D2'*D2))*TimeData;

Fs = 50;
NyqFreq = Fs/2;

BndPss = fir1(127,[0.8/NyqFreq 4.0/NyqFreq]);

ICASource = AvgRGBDeTrend' ;

% refer to (Cardosa, 1993) and (Cardosa, 1999)

[ICAMatrix]=jadeR(ICASource,3);
ICASourceProc = ICAMatrix*ICASource ;
ICASourceFilt = ICASourceProc;

ICASourceFilt(1,:) =
imfilter(ICASourceProc(1,:),BndPss,'replicate','same');
ICASourceFilt(2,:) =
imfilter(ICASourceProc(2,:),BndPss,'replicate','same');
ICASourceFilt(3,:) =
imfilter(ICASourceProc(3,:),BndPss,'replicate','same');

[ff1,mag1] = FFTMatlab(ICASourceFilt(1,:),Fs,50);
[ff2,mag2] = FFTMatlab(ICASourceFilt(2,:),Fs,50);
[ff3,mag3] = FFTMatlab(ICASourceFilt(3,:),Fs,50);

[magmx1,magmx1_loc]= max(mag1);
[magmx2,magmx2_loc]= max(mag2);
[magmx3,magmx3_loc]= max(mag3);
[~,magmx_mx_loc] = max([magmx1 magmx2 magmx3]);

eval(['ICASourceChosen = ICASourceFilt(' num2str(magmx_mx_loc)
',:);']);

Addition = 0 ;

for k = 500:50:3000

    Addition = Addition + 1 ;
```

```

ICASourceParticular = ICASourceChosen(:,k-499:k) ;

tambah = 0;

for j=0.8:0.02:3.98

    tambah = tambah + 1;
    BndPss = fir1(127,[j/NyqFreq (j+0.02)/NyqFreq]);

    ICASourceBandPass(tambah,:) = imfilter
(ICASourceParticular(1,:),BndPss,'replicate','same');

end

size_of_ICABP = size(ICASourceBandPass,1) ;

increase = 0 ;

clear Frequency; clear Energy ;

for i = 1:size_of_ICABP

    increase = increase + 1 ;
    Frequency (i,1) = 0.02*(i-1) + 0.8 ;
    Energy (i,1) = mean(ICASourceBandPass (i,:).^2);

end

[~,maxloc] = max(Energy) ;

HeartRate (Addition,1) = Frequency(maxloc)*60 ;

Time(Addition,1) = k/50 +NumOfSec - 1 ;

end

TotalResult = [Time HeartRate] ;

%% Short-time Fourier Transform

load('TimeData.mat')

% refer to (Tarvainen et al., 2002)

T = length(TimeData);
lambda = 50;
I = speye(T);
D2 = spdiags(ones(T-2,1)*[1 -2 1],[0:2],T-2,T);
AvgRGBDeTrend = (I-inv(I+lambda^2*D2'*D2))*TimeData;

Fs = 50;
NyqFreq = Fs/2;

```

```

BndPss = fir1(127,[0.8/NyqFreq 4.0/NyqFreq]);

ICASource = AvgRGBDeTrend' ; %(3,) ;

% refer to (Cardosa, 1993) and (Cardosa, 1999)

[ICAMatrix]=jadeR(ICASource,3);

ICASourceProc = ICAMatrix*ICASource ; % for ICA %(3,:)
ICASourceFilt = ICASourceProc;

ICASourceFilt(1,:) =
imfilter(ICASourceProc(1,:),BndPss,'replicate','same');
ICASourceFilt(2,:) =
imfilter(ICASourceProc(2,:),BndPss,'replicate','same');
ICASourceFilt(3,:) =
imfilter(ICASourceProc(3,:),BndPss,'replicate','same');

[ff1,mag1] = FFTMatlab(ICASourceFilt(1,:),Fs,50);
[ff2,mag2] = FFTMatlab(ICASourceFilt(2,:),Fs,50);
[ff3,mag3] = FFTMatlab(ICASourceFilt(3,:),Fs,50);

[magmx1,magmx1_loc]= max(mag1);
[magmx2,magmx2_loc]= max(mag2);
[magmx3,magmx3_loc]= max(mag3);
[~,magmx_mx_loc] = max([magmx1 magmx2 magmx3]);

eval(['ICASourceChosen = ICASourceFilt(' num2str(magmx_mx_loc)
',:);']);

[s,f,t,p] = spectrogram(ICASourceChosen,1000,950,10000,Fs);

s = abs(s) ;

size_of_s = size(s,2) ;

increase = 0 ;

for i = 1:size_of_s

    [~,highestpeakloc]= max (s(:,i));

    increase = increase + 1 ;

    eval(['Time (increase,1) =' num2str(i) '+NumOfSec-1+20;']);

    HeartRate (increase,1) = f(highestpeakloc)*60 ;
end

TotalResult = [Time HeartRate] ;

```

APPENDIX B

The following codes are the main parts of the algorithm used in the model described in Chapter 4.

```
load('AvgRGBTotal.mat')

for FirstInitPoint = ini:1:iniEnd

    MinPoint = 100;
    MaxPoint = 400;

    pp = 0; RTotalMin = 1;

    for mm=MinPoint:1:MaxPoint

        pp=pp+1;

        SampFreq = 50;
        NyqFreq = SampFreq/2;

        % refer to (Tarvainen et al., 2002)

        T = length(AvgRGB);
        lambda = 50;
        I = speye(T);
        D2 = spdiags(ones(T-2,1)*[1 -2 1],[0:2],T-2,T);
        AvgRGBDeTrend = (I-inv(I+lambda^2*D2'*D2))*AvgRGB;

        AvgRGBMean = mean (AvgRGBDeTrend);

        AvgRGBSTD = std (AvgRGBDeTrend);

        RedNormalized = (AvgRGBDeTrend(:,1) -
        AvgRGBMean(1,1))/AvgRGBSTD(1,1);

        GreenNormalized = (AvgRGBDeTrend(:,2) -
        AvgRGBMean(1,2))/AvgRGBSTD(1,2);

        BlueNormalized = (AvgRGBDeTrend(:,3) -
        AvgRGBMean(1,3))/AvgRGBSTD(1,3);

        RGBNormalized = [RedNormalized GreenNormalized BlueNormalized
        ]; %(:,3)
        %

        % refer to (Cardosa, 1993) and (Cardosa, 1999)

        [ICAMatrix]=jadeR(RGBNormalized',3);
        ICASource = ICAMatrix*RGBNormalized' ; % for ICA %(3,:)
```

```

b = fir1(127,[0.8/NyqFreq 4/NyqFreq]);

BndPss1 = imfilter(ICASource(1,:),b,'replicate','same');
BndPss2 = imfilter(ICASource(2,:),b,'replicate','same');
BndPss3 = imfilter(ICASource(3,:),b,'replicate','same');

ICASource = [BndPss1; BndPss2; BndPss3] ;      %(3,:)
ICASourceProcessed =ICASource';

[L1] = MutualInformation(ICASource(1,:) ', ICASource(2,:) ');
[L2] = MutualInformation(ICASource(1,:) ', ICASource(3,:) ');
[L3] = MutualInformation(ICASource(2,:) ', ICASource(3,:) ');

L(1) = L1(1,1);
L(2) = L2(1,1);
L(3) = L3(1,1);

LMeanTotal(pp,1) = mean(L);

[f1,y1] = FFTMatlab(ICASourceProcessedMin(:,1),50,50);
[f2,y2] = FFTMatlab(ICASourceProcessedMin(:,2),50,50);
[f3,y3] = FFTMatlab(ICASourceProcessedMin(:,3),50,50);

[y1Max, y1MaxLoc] = max(y1);
[y2Max, y2MaxLoc] = max(y2);
[y3Max, y3MaxLoc] = max(y3);

yTotal = [y1Max;y2Max;y3Max];
[~,yTotalMaxLoc]= max(yTotal);
yTotalCheck(pp,1:3) = [y1Max;y2Max;y3Max];
yTotalMaxLocCheck(pp,:)= yTotalMaxLoc ;
eval(['yTotalCheck(pp,4) =' num2str(yTotalMaxLoc)
';']);

eval(['HeartRateCheck(pp,1)= f1 (y1MaxLoc)*60;']);
eval(['HeartRateCheck(pp,2)= f2 (y2MaxLoc)*60;']);
eval(['HeartRateCheck(pp,3)= f3 (y3MaxLoc)*60;']);
eval(['HeartRateCheck(pp,4)= f1 (y' num2str(yTotalMaxLoc)
'MaxLoc)*60;']);

end      % for mm=MinPoint:1:MaxPoint
end % for FirstInitPoint = ini:1:iniEnd

```

APPENDIX C

The following codes are the main parts of the algorithm used in the model described in Chapter 5.

```
load('ROI_RGBMeanTotal.mat')

for NumberOfFrame=2:1:length(AvgRGBTotal)

    aa = aa + 1;
    TimeData (aa,1:3) = log (AvgRGBTotal (NumberOfFrame,:));
    TimeData (aa,4:6) = log (AvgYCbCrTotal (NumberOfFrame,:));

end

counter = 0;

for j=j1:119

    counter = counter + 1;

    tambah = 0 ;

    SamplingFrequency = 50 ;

    NyqFreq = SamplingFrequency/2 ;

    for k = 50:1:500

        Duration = k/50;

        tambah = tambah + 1;

        TestData = TimeData (j*50-k+1:j*50,:) ;

        T = length(TestData);
        lambda = 50;
        I = speye(T);
        D2 = spdiags(ones(T-2,1)*[1 -2 1],[0:2],T-2,T);
        AvgRGBDeTrend = (I-inv(I+lambda^2*D2'*D2))*TestData;

        AvgRGBMean = mean (AvgRGBDeTrend);

        AvgRGBSTD = std (AvgRGBDeTrend);

        RedNormalized = (AvgRGBDeTrend (:,1) -
        AvgRGBMean(1,1))/AvgRGBSTD(1,1);
```

```

        GreenNormalized = (AvgRGBDeTrend (:,2) -
        AvgRGBMean(1,2))/AvgRGBSTD(1,2);

        BlueNormalized = (AvgRGBDeTrend (:,3) -
        AvgRGBMean(1,3))/AvgRGBSTD(1,3);

        YNormalized = (AvgRGBDeTrend (:,4) -
        AvgRGBMean(1,4))/AvgRGBSTD(1,4);

        CbNormalized = (AvgRGBDeTrend (:,5) -
        AvgRGBMean(1,5))/AvgRGBSTD(1,5);

        CrNormalized = (AvgRGBDeTrend (:,6) -
        AvgRGBMean(1,6))/AvgRGBSTD(1,6);

        RGBNormalized = [RedNormalized GreenNormalized BlueNormalized
        YNormalized CbNormalized CrNormalized]; %(:,6)

        [~,score,latent,~,explained]= pca(RGBNormalized) ;

        score = score' ; % (6,:)

        b = fir1(127,[0.8/NyqFreq 3.5/NyqFreq]);

        BndPss1 = imfilter(score(1,:),b,'replicate','same');
        BndPss2 = imfilter(score(2,:),b,'replicate','same');
        BndPss3 = imfilter(score(3,:),b,'replicate','same');
        BndPss4 = imfilter(score(4,:),b,'replicate','same');
        BndPss5 = imfilter(score(5,:),b,'replicate','same');
        BndPss6 = imfilter(score(6,:),b,'replicate','same');

        score = [BndPss1;BndPss2;BndPss3;BndPss4;BndPss5;BndPss6] ;
        % (1,:)

        ScoreProcessed =score'; %(:,3)

        [f1,mx1] =
        FFTMatlab(ScoreProcessed(:,1),SamplingFrequency,50);

        [mxmax1, mxmaxloc1] = max(mx1);

        HeartRateFirstPC = f1(mxmaxloc1)*60 ;

        CC = zeros(6,6);
        MIValue = zeros(6,6);

        for m=1:1:length(CC)

            for n= 1:1:length(CC)

                CCSingle =
                corrcoef(ScoreProcessed(:,m),ScoreProcessed(:,n));
                CC(m,n) = abs(CCSingle(2,1));
            end
        end

```

```

end

SumOfCC = 0; jom = 0;

for m=length(CC):-1:2

    for n=m-1:-1:1

        jom = jom + 1;

        SumOfCC = SumOfCC + CC(m,n);
    end

end

for m=1:1:length(CC)

    for n= 1:1:length(CC)

        MISingle =
MutualInformation(ScoreProcessed(:,m),ScoreProcessed(:,n));
        MIValue(m,n) = MISingle;
    end

end

SumOfMI = 0; jom2 = 0;

for m=length(CC):-1:2

    for n=m-1:-1:1

        jom2 = jom2 + 1;

        SumOfMI = SumOfMI + MIValue(m,n);
    end

end

Results(tambah, 1) = Duration;

Results(tambah,2) = SumOfCC/jom;

Results(tambah,3) = SumOfMI/jom2;

Results(tambah, 4) = explained(1,1) ;

Results(tambah, 5) = HeartRateFirstPC;

end

end

```
