

Heart Rate Variability extraction from video signals

by

Karim Alghoul

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the M.A.Sc degree in
Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa

© Karim Alghoul, Ottawa, Canada, 2015

Abstract

Heart Rate Variability (HRV) analysis has been garnering attention from researchers due to its wide range of applications. Medical researchers have always been interested in Heart Rate (HR) and HRV analysis, but nowadays, investigators from variety of other fields are also probing the subject. For instance, variation in HR and HRV is connected to emotional arousal. Therefore, knowledge from the fields of affective computing and psychology, can be employed to devise machines that understand the emotional states of humans. Recent advancements in non-contact HR and HRV measurement techniques will likely further boost interest in emotional estimation through . Such measurement methods involve the extraction of the photoplethysmography (PPG) signal from the human's face through a camera. The latest approaches apply Independent Component Analysis (ICA) on the color channels of video recordings to extract a PPG signal. Other investigated methods rely on Eulerian Video Magnification (EVM) to detect subtle changes in skin color associated with PPG.

The effectiveness of the EVM in HR estimation has well been established. However, to the best of our knowledge, EVM has not been successfully employed to extract HRV feature from a video of a human face. In contrast, ICA based methods have been successfully used for HRV analysis. As we demonstrate in this thesis, these two approaches for HRV feature extraction are highly sensitive to noise. Hence, when we evaluated them in indoor settings, we obtained mean absolute error in the range of 0.012 and 28.4.

Therefore, in this thesis, we present two approaches to minimize the error rate when estimating physiological measurements from recorded facial videos using a standard camera. In our first approach which is based on the EVM method, we succeeded in extracting HRV measurements but we could not get rid of high frequency noise, which resulted in a high error percentage for the result of the High frequency (HF) component. Our second proposed approach solved this issue by applying ICA on the red, green and blue (RGB) colors channels and we were able to achieve lower error rates and less noisy signal as compared to previous related works. This was done by using a Buterworth filter with the subject's specific HR range as its Cut-Off.

The methods were tested with 12 subjects from the DISCOVER lab at the University of Ottawa, using artificial lights as the only source of illumination. This made it a challenge for us because artificial light produces HF signals which can interfere with the PPG signal. The final results show that our proposed ICA based method has a mean absolute error (MAE) of 0.006, 0.005, 0.34, 0.57 and 0.419 for the mean HR, mean RR, LF, HF and LF/HF respectively. This approach also shows that these physiological parameters are highly correlated with the results taken from the electrocardiography (ECG).

Acknowledgments

First, I want to thank God for giving me the health and strength to be able to finish my thesis. I would also like to express my sincere gratitude to my supervisor Prof. Abdulmotaleb El Saddik for his trust, support, help and advice through my studies. I believe that everyone need a chance to shine, and Dr. EL Saddik gave me and other students this chance to work with him and learn and improve.

A special thanks to my co-supervisor and idol Prof. Hussein Al Osman who was a guide, friend and a brother to me throughout my thesis work. Words cannot describe how much I learned from him and how much I love working with him.

I would also like to say from the bottom of my heart "Thank you" to **Bahij Alghoul** my father who sacrificed a lot so I could finish my graduate studies. **Bahij Alghoul** is not just my father, he is my hero, thank you again for everything. I also want to thank all my family members, my mother Siham, my two brothers Ahmad and Amjad, my sister Katia and my lovely wife Nada for being there for me during the hard and happy moments of my life, without their support I would not have been able to accomplish this big step in my life.

Last, but not least, I would like to thank all members of the DISCOVER Lab, for their cooperation, support and being great friends.

Contents

CHAPTER 1. INTRODUCTION.....	1
1.1 PROBLEM STATEMENT	2
1.2 MOTIVATION	3
1.3 CONTRIBUTIONS	3
1.4 THESIS ORGANIZATION.....	4
CHAPTER 2. BACKGROUND AND RELATED WORK.....	5
2.1 MENTAL STRESS	5
2.2 HEART RATE VARIABILITY.....	7
2.2.1 <i>Power Spectrum Density (PSD)</i>	8
2.3 HRV MEASUREMENT	8
2.3.1 <i>Contact Measurement (ECG: Electrocardiography) Method</i>	9
2.3.2 <i>Non-Contact Measurement Methods</i>	9
2.4 SIGNAL FILTERING	10
2.5 INDEPENDENT COMPONENT ANALYSIS (ICA) TECHNIQUE.....	12
2.5.1 <i>Recovery of BVP From Video Recording</i>	13
2.5.2 <i>Filtering The Output Source</i>	16
2.6 EULERIAN VIDEO MAGNIFICATION (EVM)	17
2.6.1 <i>Basic Approach</i>	18
2.6.2 <i>Color Amplification</i>	22
CHAPTER 3. PPG SIGNAL RECOVERY METHODS	23
3.1 FIRST METHOD : RECOVERY OF PPG SIGNAL USING ICA	24
3.2 SECOND METHOD : RECOVERY OF PPG SIGNAL USING EVM	24
3.2.1 <i>Temporal Filtering</i>	26
3.3 FILTERING THE OUTPUT SOURCE	26
3.3.1 <i>Hamming Filter</i>	26
3.3.2 <i>Butterworth Band-Pass Filter</i>	28
3.4 IMPLEMENTATION	29
3.4.1 <i>Input Video</i>	29
3.4.2 <i>Choosing the Region of Interest (ROI)</i>	30
3.4.3 <i>Joint Approximate Diagonalization Eigenmatrices (JADE)</i>	31
3.4.4 <i>EVM Implementation</i>	32

3.4.5	<i>Signal Filtering</i>	32
3.5	PHYSIOLOGICAL PARAMETERS' EXTRACTION	33
3.5.1	<i>Heart Rate Estimation</i>	33
3.5.2	<i>Cubic Spline Interpolation</i>	34
3.5.3	<i>Peak Detection</i>	36
3.5.4	<i>Artifact Removal</i>	36
3.5.5	<i>Time Domain Parameters</i>	37
3.5.6	<i>Frequency Domain Parameters</i>	38
CHAPTER 4.	EVALUATION OF PROPOSED METHODS.....	40
4.1	OBJECTIVE	40
4.2	PARTICIPANTS	40
4.3	PROCEDURE	41
4.4	EVALUATION MEASUREMENTS.....	41
4.5	RESULTS.....	41
4.5.1	<i>HRV Analysis After EVM</i>	41
4.5.2	<i>HRV Analysis After ICA</i>	49
4.5.3	<i>The Root-Mean-Square-Error (RMSE)</i>	55
4.6	ANALYSIS OF THE RESULTS.....	55
CHAPTER 5.	CONCLUSION AND FUTURE WORK	57
5.1	CONCLUSION.....	57
5.2	FUTURE WORK	58

List of Tables

TABLE 3-1: SUMMARY OF THE HRV ANALYSIS PARAMETERS	39
TABLE 4-1: THE MEAN ABSOLUTE ERROR OF THE HRV PARAMETERS FOR THE THREE ATTEMPTS OF THE EVM METHODS	45
TABLE 4-2: THE MEAN ABSOLUTE ERROR OF THE HRV PARAMETERS FOR THE TWO ICA METHODS	52
TABLE 4-3: TABLE 6 3: RMSE RESULTS FOR LF/HF AND THE MEAN RR PARAMETERS OF ALL METHODS (EVM AND ICA)	55

List of Figures

FIGURE 2-1: ECG COMPLEX	9
FIGURE 2-2 : EXAMPLE ON THE TYPES OF IIR FILTERS	12
FIGURE 2-3: "RECOVERY OF THE BVP WAVEFORM. (A) FACE WITHIN THE FIRST VIDEO FRAME	16
FIGURE 2-4: EULERIAN VIDEO MAGNIFICATION STEPS [44]	18
FIGURE 2-5: LAPLACIAN PYRAMID	19
FIGURE 2-6: RESIZE DOWN WITH GAUSSIAN PYRAMID.....	20
FIGURE 2-7: "AN EXAMPLE OF USING OUR EULERIAN VIDEO MAGNIFICATION FRAMEWORK FOR VISUALIZING THE HUMAN PULSE ..	21
FIGURE 3-1:BLOCK DIAGRAM OF THE ICA BASED METHOD	24
FIGURE 3-2: BLOCK DIAGRAM OF THE EVM BASED METHODS.	25
FIGURE 3-3: EXAMPLE OF A BUTTERWORTH FILTER	28
FIGURE 3-4: EXPERIMENTAL SETUP.....	29
FIGURE 3-5: THE SIGNALS EXTRACTED FROM THE FACIAL VIDEO.	31
FIGURE 3-6: AN EXAMPLE OF USING THE EVM FOR COLOR MAGNIFICATION ONLY ON THE ROI AND NOT ALL THE FRAME	32
FIGURE 3-7: HR ESTIMATION PROCEDURE.....	33
FIGURE 3-8: HRV ANALYSIS WITH AND WITHOUT CUBIC SPLINE INTERPOLATION.	35
FIGURE 3-9: ARTIFACT REMOVAL.....	37
FIGURE 4-1: EVM WITH CUT-OFF BETWEEN 45 -180 BPM FOR TEMPORAL FILTER AND BUTTERWORTH FILTER.	42
FIGURE 4-2: EVM WITH CUT-OFF BETWEEN 45 -180 BPM FOR TEMPORAL FILTER AND THE RANGE OF THE SUBJECT'S HR AS CUT- OFF FOR BUTTERWORTH FILTER.....	43
FIGURE 4-3: THE RANGE OF THE SUBJECT'S HR AS CUT-OFF FOR THE TEMPORAL AND BUTTERWORTH FILTERS.....	44
FIGURE 4-4: SCATTER PLOTS AND PEARSON'S CORRELATION VALUES OF THE MEAN RR BETWEEN A CAMERA AND SENSOR OF THE THREE EVM BASED METHODS.....	46
FIGURE 4-5: SCATTER PLOTS AND PEARSON'S CORRELATION VALUES OF THE LF BETWEEN A CAMERA AND SENSOR OF THE THREE EVM BASED METHODS.....	47
FIGURE 4-6: SCATTER PLOTS AND PEARSON'S CORRELATION VALUES OF THE HF BETWEEN A CAMERA AND SENSOR OF THE THREE EVM BASED METHODS.....	48
FIGURE 4-7: SCATTER PLOTS AND PEARSON'S CORRELATION VALUES OF THE LF/HF BETWEEN A CAMERA AND SENSOR OF THE THREE EVM BASED METHODS.....	49
FIGURE 4-8: HRV ANALYSIS AFTER 3 DB HAMMING WINDOW WITH A CUT-OFF BETWEEN 45-180 BPM.	50
FIGURE 4-9: HRV ANALYSIS AFTER THE PROPOSED ICA METHOD	51
FIGURE 4-10:SCATTER PLOTS AND PEARSON'S CORRELATION VALUES OF THE MEAN RR BETWEEN A CAMERA AND SENSOR OF THE THREE EVM BASED METHODS.....	53

FIGURE 4-11: SCATTER PLOTS AND PEARSON'S CORRELATION VALUES OF LF BETWEEN A CAMERA AND SENSOR OF THE THREE EVM
BASED METHODS..... 53

FIGURE 4-12: SCATTER PLOTS AND PEARSON'S CORRELATION VALUES OF HF BETWEEN A CAMERA AND SENSOR OF THE THREE EVM
BASED METHODS..... 54

FIGURE 4-13: SCATTER PLOTS AND PEARSON'S CORRELATION VALUES OF LF/HF BETWEEN A CAMERA AND SENSOR OF THE THREE
EVM BASED METHODS..... 54

Glossary of Terms

HR	Heart Rate
HRV	Heart Rate Variability
ECG	electrocardiography
PPG	photoplethysmography
EVM	Eulerian video magnification
ANS	Autonomic nervous system
BR	Breath rate
BVP	Blood volume pulse
IBI	Inter-beat-intervals
SDNN	Standard deviation between NN
NN50	Total number of successive RRs which differ by more than 50 ms
PNN50	The proportion of the NN50 divided by the number of RRs
RMSD	Root mean square of successive differences
RR interval	The time between two consecutive Heart beats
VLF	Very low frequency
LF	Low frequency
HF	High frequency
PSD	Power spectral density
DFT	Discrete fourier transform
FFT	Fast fourier transform
FIR	Finite impulse response
IIR	Infinite impulse response
RGB	Red, green, blue
RGBCO	Red, green, blue, cyan, orange
OpenCV	Open computer vision
ROI	Region of interest
JADE	Joint approximate Diagonalization of Eigenmatrices
FPS	Frame per second
MAE	Mean absolute error
RMSE	Root-mean-square-error

Chapter 1.

Introduction

HR refers to the number of heart beats within a time interval, while HRV refers to the variation in the time intervals between consecutive heart beats. Changes in HRV, as well as other physiological features, are highly reflective of emotional arousal [1]. Thus, in addition to assessing cardiac problems, HRV analysis can be used to measure mental stress in individuals [2][3].

Methods for measuring HRV can be classified into either *contact* methods or *non-contact*. Contact methods are the most widely used techniques for extracting HRV signals, and often involve the use of sensors placed directly on the body. These may involve the measurement of an electrocardiography (ECG) signal obtained through an ECG sensor (a device that measures the electrical activity of the heart) or photoplethysmography (PPG) signal obtained through a pulse oximeter (a device that monitors the oxygen absorbance in the blood) . Conversely, non-

contact methods rely on optical sensors that estimate a PPG signal without directly being placed on a subject's body. Traditionally, non-contact methods involve the use of highly specialized and expensive sensors like laser and microwave Dopplers [4][5]. Recently it has been shown that a PPG signal can be measured using a standard digital camera with ambient light as an illumination source [6]. Furthermore, HRV features now can be estimated from a simple recorded facial video [7] which makes it easier, cheaper and more practical to conduct HRV analysis.

This work improves on existing methods to extracting HRV parameters using facial video captured by a standard digital camera. Such parameters can be used for medical assessment or emotion estimation.

1.1 Problem Statement

The use of non-obtrusive (i.e. non-contact) methods of measuring HR and HRV has been the focus of several research studies in recent years [8][9][5][10]. These methods of measuring physiological features can improve healthcare practices, as well as enable advancements in affective computing [9]. Many applications, such as those involved in newborn monitoring[11] and stress monitoring [12], can be enhanced by employing remote physiological measurement methods.

Contact measurement devices are still the golden standard when it comes to estimating HR and HRV, however, non-contact or remote measuring techniques do not require electrodes' attachment to the body [9], and thus may be a more convenient measurement

source. Therefore, we aim to improve the accuracy of non-contact methods in order to enable their use in the types of applications we have mentioned.

1.2 Motivation

HRV monitoring based on contact sensors may make individuals feel nervous or uncomfortable, which would create variability in the results, especially when it comes to emotion estimation. Thus, in order to use HRV analysis efficiently in affective computing and truly benefit from its ability to accurately estimate emotions, the measurement method should ideally involve non-contact or remote devices. Non-contact devices also allow measurements to be made without alerting the individual who is being measured. For example, this technique can be applied in police interrogations to monitor signs of stress and arousal.

1.3 Contributions

In this thesis we present results of HRV parameter estimation from videos of human faces using two proposed approaches. The first approach is based on the Eulerian Video Magnification (EVM) technique. To the best of our knowledge, this is the first successful attempt at using an EVM based method for this purpose. The second approach is based on Independent Component Analysis (ICA). We show that our proposed ICA based approach yields more accurate results for HRV parameters estimation using non-contact method compared to previous works.

1.4 Thesis Organization

The Thesis is organized as follows:

- Chapter 2 talks about the Related work for the topics discussed in this thesis
- Chapter 3 provides details about the two methods we introduced to extract the Heart Rate Variability parameters from a recorded facial video.
- Chapter 4 describes the Evaluation procedure and the results of the experiments done and compared with the ground truth.
- Chapter 5 provides a conclusion and an outlook on future works.

Chapter 2.

Background and Related Work

There have been many efforts to make HRV feature extraction from facial videos possible. The purpose of this chapter is to provide some backgrounds on the thesis topic which are important for understanding the rest of the work.

2.1 Mental Stress

Although the topic of stress in current literature has received much attention, researchers have still not developed a solid definition for this concept [13]. Stress leads to chemical and hormonal reactions within the body during exposure to environmental pressures or unpleasant feelings [14]. Five categories of stress are provided, based on a classification of stressors by duration and course [15]:

- Acute or time-limited stressors include limited periods of nervousness brought upon by numerous activities, such as speaking in public or engaging in mentally challenging problems.
- Brief naturalistic stressors occur naturally in a person's environment. Common examples consist of an academic examination or a job interview.
- Stressful events sequences such as the death of a parent are related to traumatic events, which may lead to more stressors.
- Person's identity or social role is changed by chronic stressors. Stability is a main character for such stressors. That is, the person does not know when the challenge is finished or recognizes that it will never end. A common case in this category is stress caused by a disability that changes the person's life.
- Distant stressors involve a cognitive or emotional response to a past distressing event which has long-lasting, adverse effects on the immune system[16]. War trauma and child abuse are examples of distant stressors.

Many common signs and symptoms of stress were outlined by The American Institute for Stress [17]. Stress is a phenomenon that varies for every person, which is why it is difficult to characterize [17]. Research has proven that many disorders are connected to stress. Examples of these disorders are depression [18], immune system disturbances [13] and heart attacks [19].

2.2 Heart Rate Variability

HRV refers to the variation in the time intervals between consecutive heart beats. This measurement was first demonstrated as useful when a relation between the decrease in HRV and fetal distress was discovered [20].

Various time and frequency domain measurements have been utilized to extract helpful parameters from an HRV signal. The Autonomous Nervous System (ANS) controls the inter-beat-intervals (IBI) of the heart, which are also referred to as NN intervals. HRV is calculated as a series of IBIs. In the time domain, the standard deviation between NN intervals (SDNN) is the most generally measured parameter. SDNN has been utilized to survey intense anxiety levels for individuals [21]. However, frequency domain parameters have received more attention from researchers [3][22][23][24]. Calculating the power spectral density (PSD) of the HRV series allows one to obtain data regarding the parasympathetic and sympathetic influences. The very low frequency power (VLF), the low frequency power (LF) and high frequency power (HF) parameters can be obtained by calculating the area under the PSD curve for the following frequency ranges:

- Very Low Frequency (VLF): 0.003 to 0.04 Hz
- Low Frequency (LF): 0.04 to 0.15 Hz
- High Frequency (HF): 0.15 to 0.4 Hz

To minimize the impact of changes in the total power, the frequency components are calculated in normalized units (n.u.). The HF band has been closely related to parasympathetic activity [25], the sympathetic and parasympathetic parts of the nervous system have an influence on the LF component [25], and the ratio between LF and HF

(LF/HF) is influenced by the balance between sympathetic and parasympathetic nervous branches [25].

2.2.1 Power Spectrum Density (PSD)

The PSD is commonly performed using two methods, Fast Fourier transform (FFT) method or the Autoregressive parametric model (AR modeling) [26]. Every method has its advantages and disadvantages, for instance FFT is simple to implement. Some of the advantages of AR can be illustrated as follows:

- Improved signal resolution especially when dealing with short samples
- Ability for signal to be factorized into dependent spectral components.

On the other hand, the complexity of AR's model order selection is the main disadvantage.

2.3 HRV Measurement

The signal from which HRV is inferred requires precise measurements of the intervals between successive heartbeats. Hence, an HRV series is typically extracted from an ECG signal. Also, an HRV series can be extracted from a PPG signal by measuring the time between cardiac cycles.

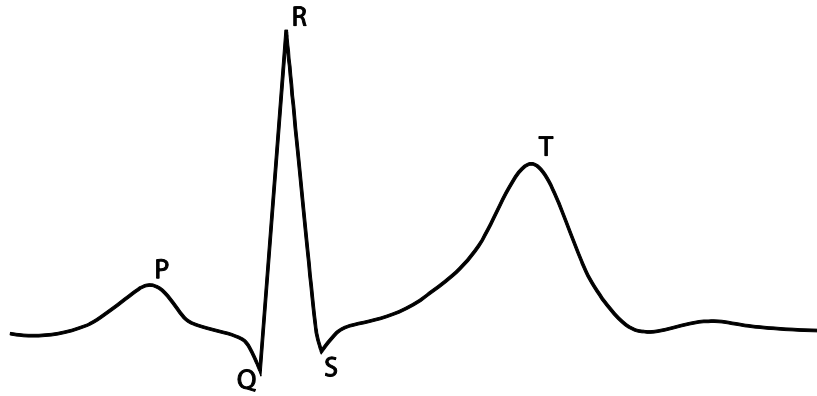


Figure 2-1: ECG complex

2.3.1 Contact Measurement (ECG: Electrocardiography) Method

Currently ECG is the golden standard when it comes to extracting the HRV signal [27]. It detects the electrical changes in the skin when the heart muscle depolarizes during each heartbeat. HRV is obtained from ECG by determining the time period between successive R peaks (also known as RR interval) (Figure 2-1). The R peak is the most prominent feature on an ECG signal and coincides with the depolarization of the ventricles. The ECG signal can be corrupted for several reasons such as power line noise or motion artifacts as a result of electrode movement [28]. Hence, corrective measures are usually applied to the extracted HRV signal in order to reduce the impact of such artifacts.

2.3.2 Non-Contact Measurement Methods

Non-contact measurement is achieved using various methods. A laser Doppler based approach [4] is one of the methods to measure heart rate (HR) and breath rate (BR). This method is not the only one however, there are several other methods for example microwave Doppler [5] and thermal imaging [29].

2.3.2.1 PPG (Photoplethysmography)

PPG is an inexpensive method for measuring the cardiovascular blood volume pulse (BVP) within the body by recording changes in the reflected light. This method may or may not involve the use of invasive instruments (i.e. external instruments which are introduced into the body) [30]. Customarily, a PPG signal is measured via specialized contact sensor and a dedicated light source like infrared light for example [9]. However, recent studies have shown that a PPG signal can be measured using a standard digital camera with ambient light as the only source of illumination [6]. Poh *et al* [8] introduced an approach to detect heart rate through facial video recordings. This approach proved the practicality of a non-contact HR measurement method using a relatively inexpensive camera. Furthermore, Poh *et al* [8] proposed a method for estimating HRV features and frequency domain parameters [7]. This technique was depending on RGB camera recording from a short distance, however it has limited frequency bands resolution (only red, green and blue). To solve this problem a novel five bands digital camera was used instead of a standard RGB camera [3][9]. Using this method, alternative frequency bands can be used to improve the accuracy of measuring physiological features. But this method requires a non-standard camera equipped with a sensor that can reveal cyan and orange bands.

2.4 Signal Filtering

A signal can be transmitted inside an assigned frequency range by an electrical network called Filter. Another term for the frequency range is "Pass Band", and where the signal is

stopped, it is referred to as "Stop Band". The frequency between the "Pass Band" and the "Stop Band" is called "cut-off" frequency.

In general terms, the filter in signal processing is a process that eliminates unwanted component from a signal. Usually filtering is used to eliminate some specific frequencies in order to reduce background noise.

Two types of filters exist, analog and digital. The digital filter is the one used in our case. The role of a digital filter in signal processing field is to execute mathematical procedures on a sampled signal in discrete-time. These mathematical operations are performed to minimize specific aspects in the signal or in other cases to maximize these aspects.

2.4.1 Types of Digital Filters

The two main types of digital filters are Finite Impulse Response (FIR) and Infinite Impulse Response (IIR).

Characteristics of FIR filters can be summed up in these three points:

- Linear phase
- High filter Order
- Stability

Characteristics of IIR filters can be summed up as such:

- Non-linear phase

- Low filter order
- A potential of instability

Due to these characteristics, the FIR filter is preferable when the phase characteristic is important because it has linear phase characteristic and FIR filters have higher order and more complexity. However, if the main interest is only the frequency response, IIR (Figure 2-2) is utilized because it has lower order meaning less complex and easier to realize.

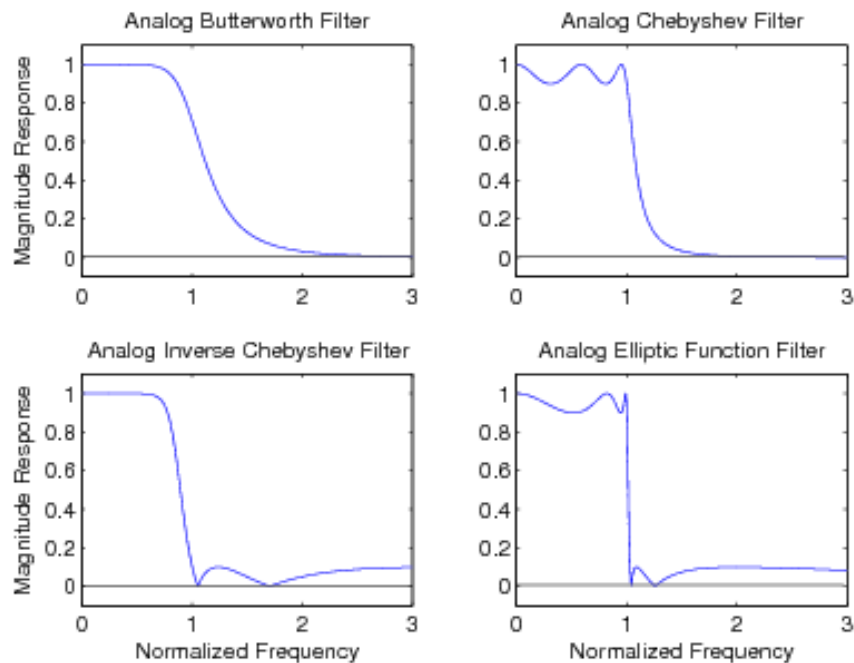


Figure 2-2 : Example on the types of IIR filters

2.5 Independent Component Analysis (ICA) Technique

The method used in [7] uses independent component analysis (ICA) for extracting the BVP waveform from the video. ICA is a technique used for isolating independent signals from a set of vectors that consist of a linear combination of these signals [31]. The difference

between ICA and other methods is that ICA searches for statistically independent and non-Gaussian signals [32].

ICA is applied to many applications in various areas [32]. For example, it can be used in brain imaging since the electrodes attached to the scalp linearly combine the signals generated from the various sources of the brain. Another application is in econometrics. Image feature extraction is another example that uses ICA to locate independent features.

When it comes to using ICA technique with PPG, the blood volume pulse (BVP) is the fundamental source that researchers [7][9] are interested in.

2.5.1 Recovery of BVP From Video Recording

Perhaps the first attempt at the development of an inexpensive method for non-contact HRV measurement was introduced by Poh *et al*[7]. According to their research, volumetric variations in the blood vessels of the face occur along the cardiac cycle. These variations alter the path length of the incident ambient light such that the following variations in the amount of reflected light demonstrate the timing of cardiovascular incidents.

When taking a facial video, RGB (red, green and blue) sensors will capture a blend of the reflected PPG signal along with other noise originating artifacts [7]. Each color sensor will record a blend of the initial source signals with a little change in their weight as a result of the differences in the hemoglobin absorptivity in the visible and near-infrared spectral range. In the method presented in [7], the detected signals from the RGB sensors are denoted as $y_1(t)$, $y_2(t)$, and $y_3(t)$ which represent the amplitudes of the saved signals at

time t . Also, it supposes that $x_1(t)$, $x_2(t)$ and $x_3(t)$ are the three fundamental source signals that have been linearly combined to generate $y_1(t)$, $y_2(t)$, and $y_3(t)$. Hence, equation (1) describes the aforementioned relationship between captured and source signals.

$$y(t) = Ax(t) \quad (1)$$

Where $y(t)$ is vector $[y_1(t), y_2(t), y_3(t)]^T$, $x(t)$ is vector $[x_1(t), x_2(t), x_3(t)]^T$ and \mathbf{A} is a square matrix (called mixture matrix) of size 3×3 that include the mixture coefficients a_{ij} . Therefore, in order to model the relationship between $y(t)$ and $x(t)$, we need to determine a separation matrix \mathbf{W} that estimates the inverse of the mixture matrix \mathbf{A} . Once \mathbf{W} is determined, the source signals (i.e. $x(t)$) can be estimated. Some of the independent random variables is more Gaussian comparing to the original variables, that is why the non-Gaussianity of all sources should be maximised by \mathbf{W} to reveal the independent sources.

In their evaluation Poh *et al* [7] used MATLAB to analyse the recorded video. Several steps of processing (Figure 2-3) were performed to recover the BVP signal from the facial video:

1. Track the face using the Open Computer Vision library [33] and a boosted cascade classifier [34].
2. Determine a region of interest (ROI) with its center being 60% of the width and full height of the facial box.
3. Extract the RGB channels from the ROI.
4. Spatially average each of the RGB channels using all pixels in the ROI to produce a red, green and blue discrete signals. These signals are denoted as $y_1(t)$,

$y_2(t)$ and $y_3(t)$ respectively and are referred to as the raw elements. Each signal contains a single data point for each video frame.

5. Detrend the raw elements using the approach of [35] with a cut-off frequency of 0.89Hz.
6. Normalize the detrended raw elements using equation (2)

$$y'_i(t) = \frac{y_i(t) - \mu_i}{\sigma_i} \quad (2)$$

Where μ_i and σ_i are the mean and standard deviation respectively for $i=1,2,3$

8. Apply ICA to break up the raw elements into three independent source signals. The joint approximate diagonalization of eigenmatrices (JADE) algorithm [36] is used in this step.
9. Select the best component after performing ICA by applying Fast Fourier Transform (FFT) on the output sources and the one with highest peak within the range of 42 and 240 bpm is chosen.

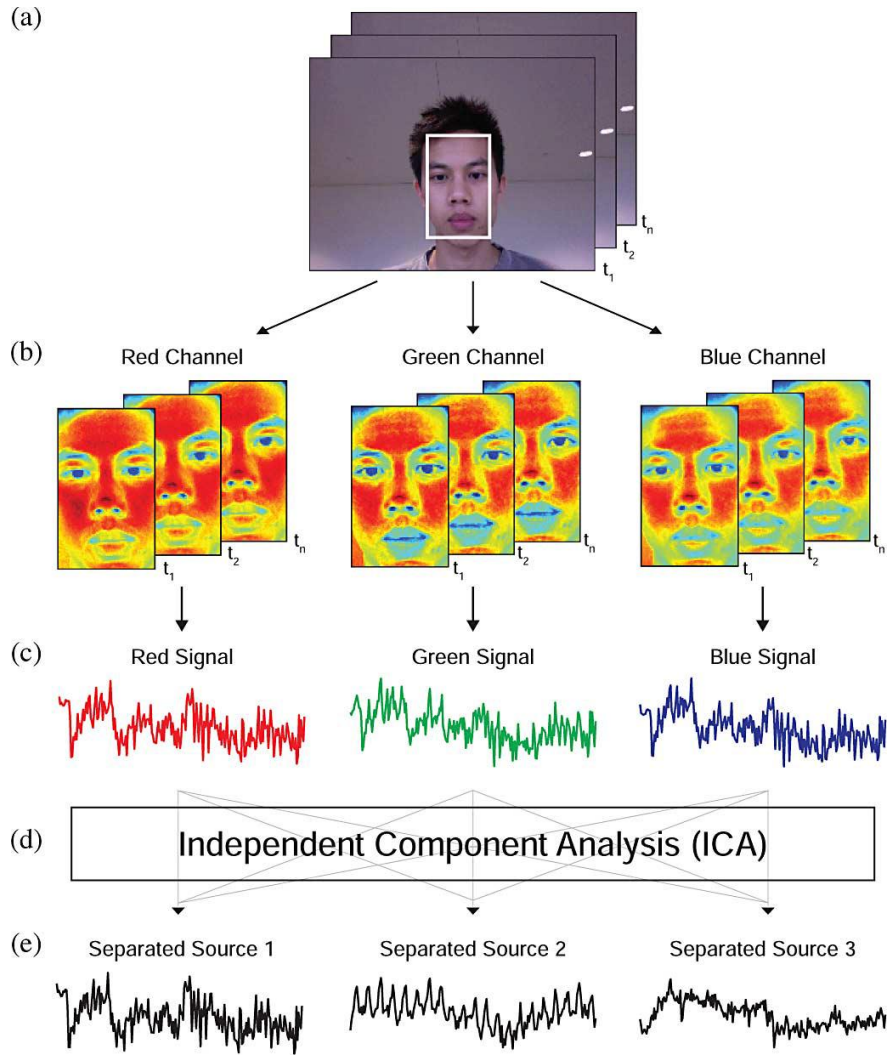


Figure 2-3: "Recovery of the BVP waveform. (a) Face within the first video frame is automatically detected to locate the ROI. (b) ROI is decomposed into red, green, and blue channels for each frame and spatially averaged to form (c) the raw signals. After the raw signals are detrended and normalized, ICA is applied to separate three independent sources. In this example, the BVP is visible in the second source signal. " [5]

2.5.2 Filtering The Output Source

In this section, we will describe the type of filters applied on the output source after the ICA.

- **Moving Average Filter:** A K-point moving average filter was used to smooth the independent source. This filter works by changing data points by the average value of the data points of its neighbors. The number of neighbors is defined in the span and in Poh *et al.*[7] case it is 5. The process of this filter is comparable to a lowpass filtering with the smoothing equation (3).

$$Y_s(i) = \frac{1}{2N + 1} (y(i + N) + y(i + N - 1) + \dots + y(i - N)) \quad (3)$$

Y_s is the notation of the output smoothed value for the i th data point, N represents the number of the neighbors. (2N+1) is 5 in Poh *et al.*[7].

- **Hamming Filter:** Following the moving average filter, a band-pass Hamming window filter is used on the signal. Poh *et al* [7] and Mcduff [9] both used this step as the last phase of processing. Poh *et al.* [7] used a 128dp hamming window filter with a cut-off between 0.7 and 4 Hz (40 - 240 bpm).

2.6 Eulerian Video Magnification (EVM)

Recovering the PPG signal using ICA is currently the most studied method. However this method is sensitive to movement and variation in light [37]. The EVM technique was shown to be effective when it comes to HR extraction [38] [39] [40] [37]. However extracting the HRV signal is still a challenge and only few attempts have been made [40] [37]. EVM differs from the Independent Component Analysis (ICA) technique, the Eulerian Video Magnification (EVM) [38] utilizes localized spatial pooling and temporal filtering to pull out the signal of the cardiac pulse [39]. This technique allows for the magnification of the subtle changes in the skin color

instigated by the blood flow in facial vessels [38]. Hence this method can be used to make this phenomenon visible to be naked eye.

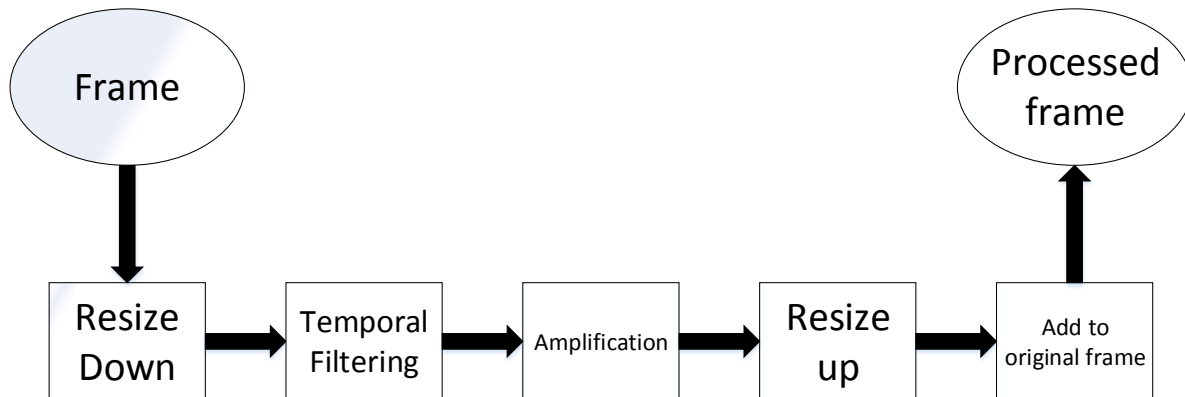


Figure 2-4: Eulerian Video Magnification steps [44]

2.6.1 Basic Approach

The work of Wu et al [38] is based on a series of previous studies on the matter [41][8][6][42]. Their approach involves taking the time series of color values at any pixel and amplifying the variations in an assigned temporal frequency band. The EVM technique (Figure 2-4) is encouraged by the Eulerian perspective where pressure and velocity develop over time and hence the term "Eulerian video magnification". When using the Eulerian perspective in videos, each pixel is processed independently and treated as a time series and signal processing is applied to it.

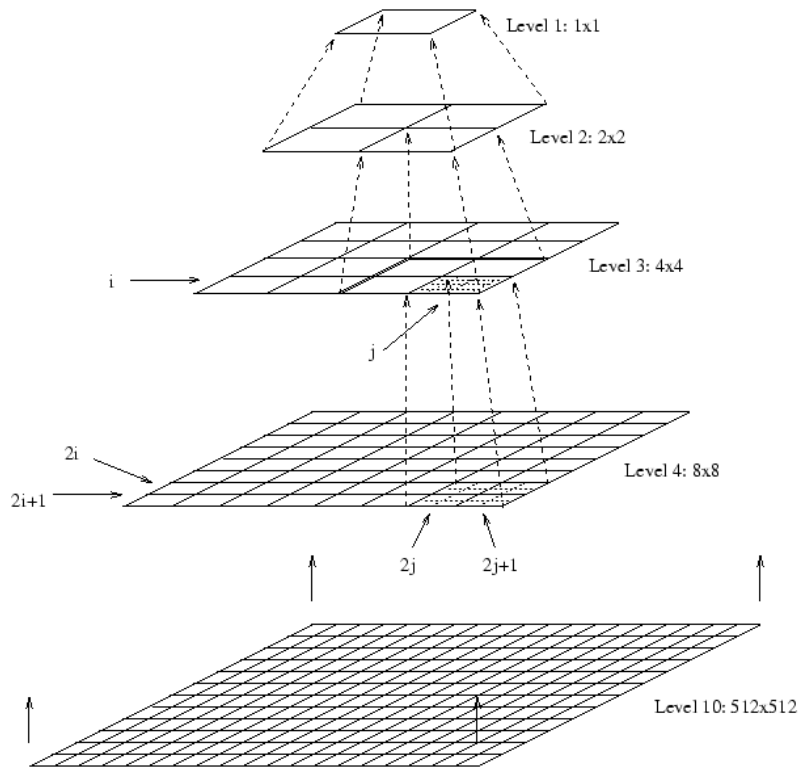


Figure 2-5: Laplacian Pyramid

2.6.1.1 Spatial Filtering (Decomposition)

As a general case for the EVM method, the full Laplacian pyramid [43] is computed when applying spatial filtering. The Laplacian pyramid (Figure 2-5) is usually used in the case of motion magnification. According to the EVM paper [38], due to different signal-to-noise ratios included in the pyramid, the extent of magnification of layers will be different.

To be able to get a better signal of a person's heart rate from their face, spatial filtering can be used to get a higher signal-to-noise ratio. In this case a Gaussian pyramid (Figure 2-6) is used instead of a Laplacian pyramid. This Gaussian pyramid is built after repeatedly computing the Gaussian average for each pixel (Resize down).

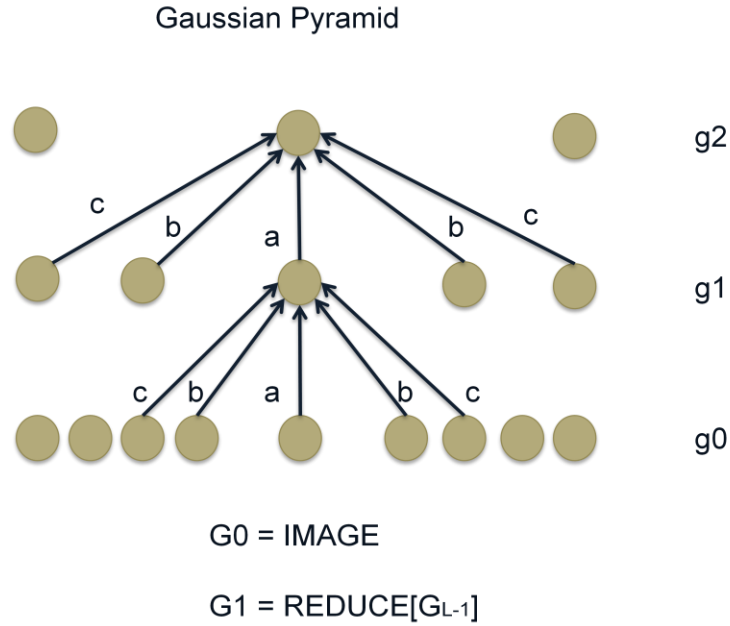


Figure 2-6: Resize down with Gaussian Pyramid

The resize down technique uses a Gaussian pyramid as a spatial filter. For this technique to be accomplished, a looping needed to be applied until the right level, such that the output of the previous loop was the input of the next loop.

$$g_l(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) g_{l-1}(2i + m, 2j + n) \quad (4)$$

The reduce method of the Gaussian Pyramid is represented in equation (4), it is defined by $g_l = REDUCE[g_{l-1}]$, l is the level of the pyramid .

2.6.1.2 Temporal Filtering

The temporal filter is applied on the spatially averaged luminance trace (time series of color values). The role of the temporal filter is to obtain the signals to be amplified. Different types of temporal filters (ideal band-pass and IIR low-pass filters) are used in this method. Thus, choosing the filter depends on the type of application. For example, to magnify motion an ideal band-pass filter with wide cut-off is better to implement. However, a narrow ideal band-pass filter will generate more accurate, less noisy results when used for color amplification of the body's blood flow. Wu *et al* [38] employed an ideal band-pass filter to benefit from its sharp Cut-Off frequencies. A low-order IIR filter is also used in their work as a viable solution for both cases: color amplification and motion magnification.

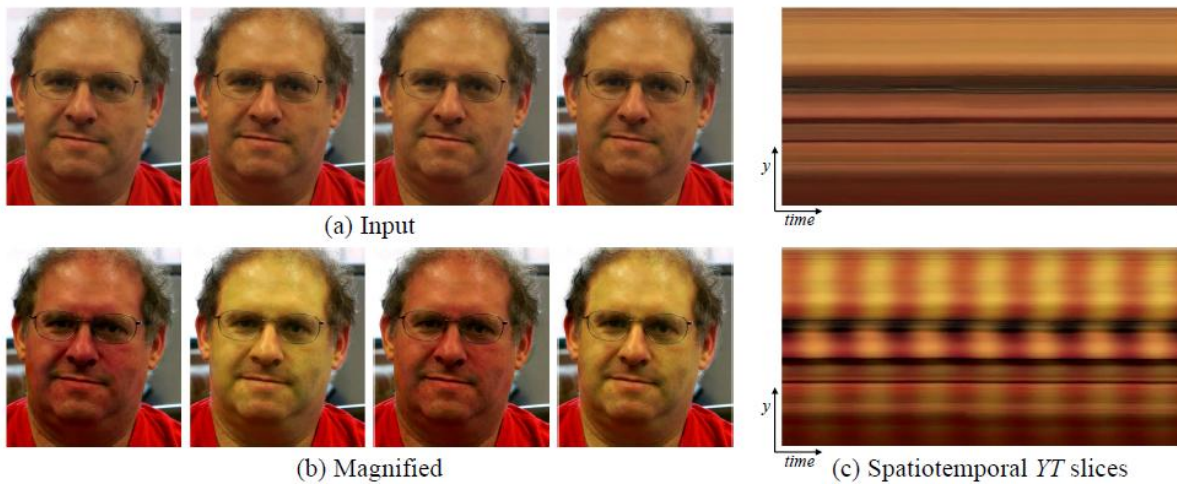


Figure 2-7: "An example of using our Eulerian Video Magnification framework for visualizing the human pulse. (a) Four frames from the original video sequence (face). (b) The same four frames with the subject's pulse signal amplified. (c) A vertical scan line from the input (top) and output (bottom) videos plotted over time shows how our method amplifies the periodic color variation. In the input sequence the signal is imperceptible, but in the magnified sequence the variation is clear. The complete sequence is available in the supplemental video." [43]

2.6.2 Color Amplification

Color amplification can be used to visualise the human pulse (Figure 2-7). For that to be done, the frequency band that contains a plausible human heart rate is amplified. After amplification, the blood flowing through the face is visible to the naked eye as a color variation. To permit the input signal to overcome the camera sensor and quantization noise, the temporal filtering used should be applied to lower spatial frequencies. Several studies have used EVM in their work [39][37][40].

Wu *et al.* [38] demonstrated that by using the correct configuration, the desired signal can be extracted. The steps to achieving this are illustrated by:

1. Choose a temporal band-pass filter
2. Choose α (the amplification factor)
3. Choose λ_c (the spatial cut-off frequency)
4. Choose the form of α 's attenuation (0 for $\lambda < \lambda_c$ or scaling α down to zero linearly)

For example, a temporal filter can be used with a cut-off frequency between 0.4 and 4 Hz (corresponding to human heart beat range 24-240). And then to emphasize the color variation, $\alpha = 100$ with $\lambda_c = 1000$ is used with 0 as the attenuation of α .

Chapter 3.

PPG signal recovery Methods

The cardiovascular pulse wave which circulates throughout the body is the fundamental source signal we are interested in. When the heart beats and pumps blood throughout the body, volumetric variations occur in the facial blood vessels . Thus, the cardiovascular blood volume pulse can be measured through the variations in the amount of reflected light.

In this thesis we propose two methods for recovering the PPG signal, one is based on the ICA technique and the second is based on the EVM technique. In this chapter we will explain these two methods in details and show our contributions.

3.1 First method : Recovery of PPG Signal Using ICA

The first method (Figure 3-1) is based on the ICA and it starts by recording a video of the subject's face and using it as an input video. We follow Poh *et al.* [7] steps (described previously in section 2.5.1) until we arrive to the filtering part where our major contribution is (explained later in section 3.3).

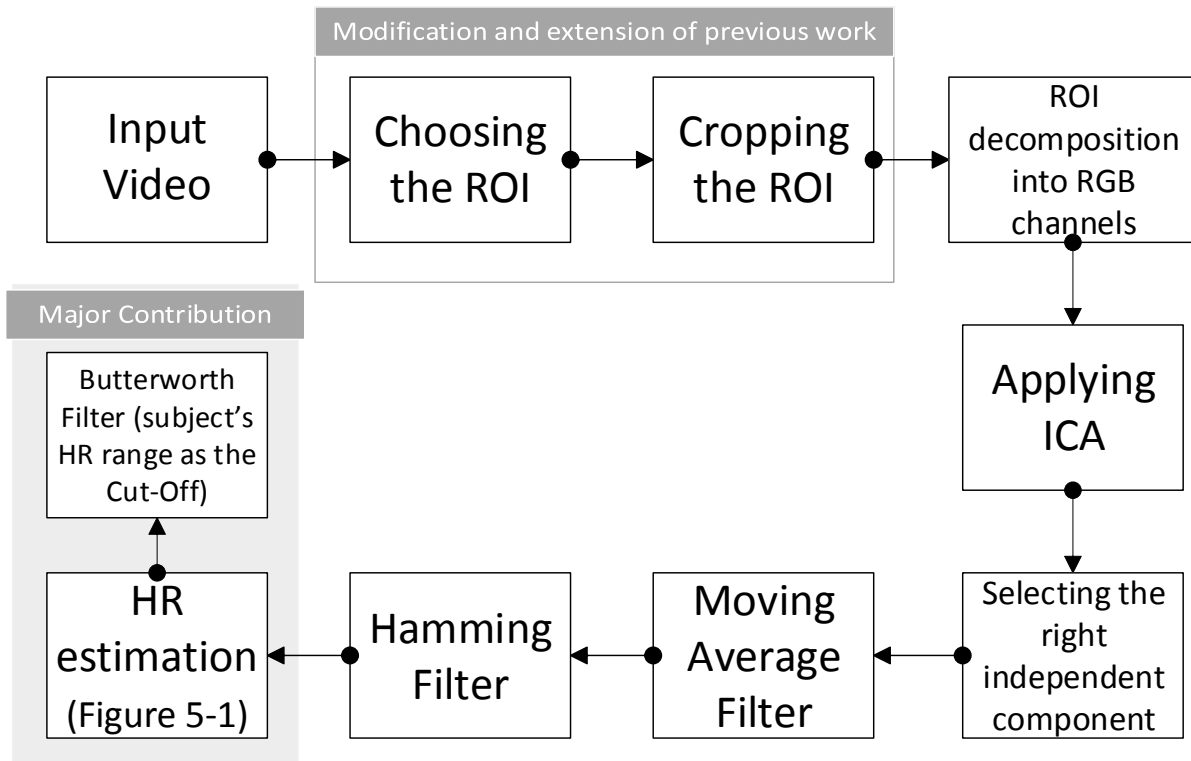


Figure 3-1:Block diagram of the ICA based method

3.2 Second Method : Recovery of PPG Signal Using EVM

The second method is based on the EVM approach. To the best of our knowledge the current work (Figure 3-2) is the first which has been able to extract good HRV features from a facial video after applying the EVM technique. If the goal is just to obtain the HR, the green channel extracted from the magnified video is enough and no further processing is necessary.

However, the HRV signal is very sensitive to noise and therefore more filtering is required. Hence, additional filters (explained later in section 3.3) are applied.

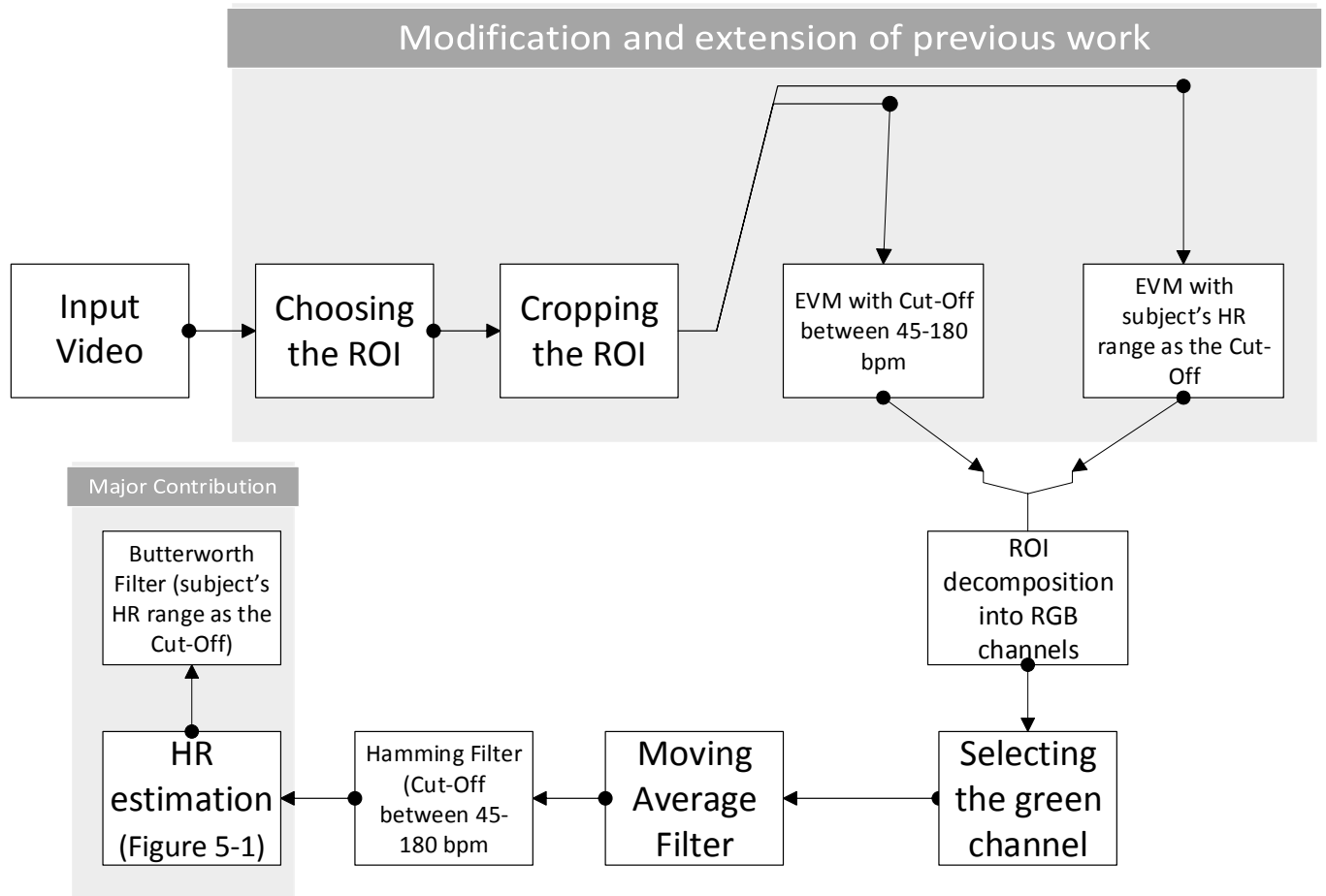


Figure 3-2: Block diagram of the EVM based methods.

The major contribution in this method is adding the filter with the subject's HR as the cut-off frequency that will be explained in section 3.3. Additional to the major contribution, couple of changes were made before arriving to the filtering part. Magnifying just the cropped

ROI instead of the whole frame is one of the things changed. The second modification is in the temporal filtering.

3.2.1 Temporal Filtering

To avoid amplifying unwanted frequencies while amplifying the color variation of pixels, an ideal band-pass filter was applied. This is the second aspect of our approach which is different from the original EVM approach. In the original approach, a Fourier transform is calculated for every pixel, and every frequency which was not between the Cut-Off (45 and 240 bpm) were set to zero and the inverse Fourier transform was applied to rebuild the frame.

To get better results we obtain the minFreq and maxFreq (explained in section 3.3) and use them as the cut-off frequencies. This makes the signal less noisy.

3.3 Filtering the Output Source

In this section, we will describe the type of filters applied on the output source after the ICA. First, we use a K-points moving average filter in our method.

3.3.1 Hamming Filter

Following the moving average filter, a band-pass Hamming window filter is used on the signal. Since our experimental setup is lit by artificial lighting sources (as described in chapter 4), the resulting signal is very noisy. Consequently, additional filtering is required to clean up the signal. If we were extracting just the HR it would not be a major issue

(since HR involves counting the number of PPG peaks regardless of the time interval between them). However, in the case of HRV estimation, in order to get the R-R interval, the signal should be very clean. Hence, we apply a narrow band pass filter with cut-off frequencies based on the specific HR range of the measurement period. This technique is especially effective for short measurement periods. It is worth noting that in numerous emotional and physiological assessment applications, HRV records of 5 minutes length are typically used [2]. So first we take the signal chosen after applying the ICA and we determine minHR (the minimum HR) and maxHR (the maximum HR) during a measurement period using a procedure that will be explained in section 3.5.1. Next, we take the minFreq (equation 5) and the maxFreq (equation 6) and used them as cut-off range with the Hamming window filter and obtained a cleaner version of the signal.

$$\text{minFreq} = (\text{minHR}_i - \alpha)/60 \quad (5)$$

where i is the period of measurement and α is a constant number $\in [0, +\infty)$

$$\text{maxFreq} = (\text{maxHR}_i + \beta)/60 \quad (6)$$

where i is the period of measurement and β is a constant number $\in [0, +\infty)$

This approach solved most of the noise problem, however the result of the frequency domain component in the HRV analysis, especially the High Frequency (HF) component, was affected after the filtering. McDuff *et al* [9] used a 3dp Hamming window filter instead of 123dp to avoid damaging the frequency component, but he used a DSLR camera to get the RGBCO channels not just the RGB colors.

To solve this problem we used a 3db hamming window filter with a Cut-Off between 0.75 and 3 HZ, which are low and high frequencies at 45 and 180 bpm. We then applied a band-pass IIR filter with the specific HR range of the subject as the Cut-Off.

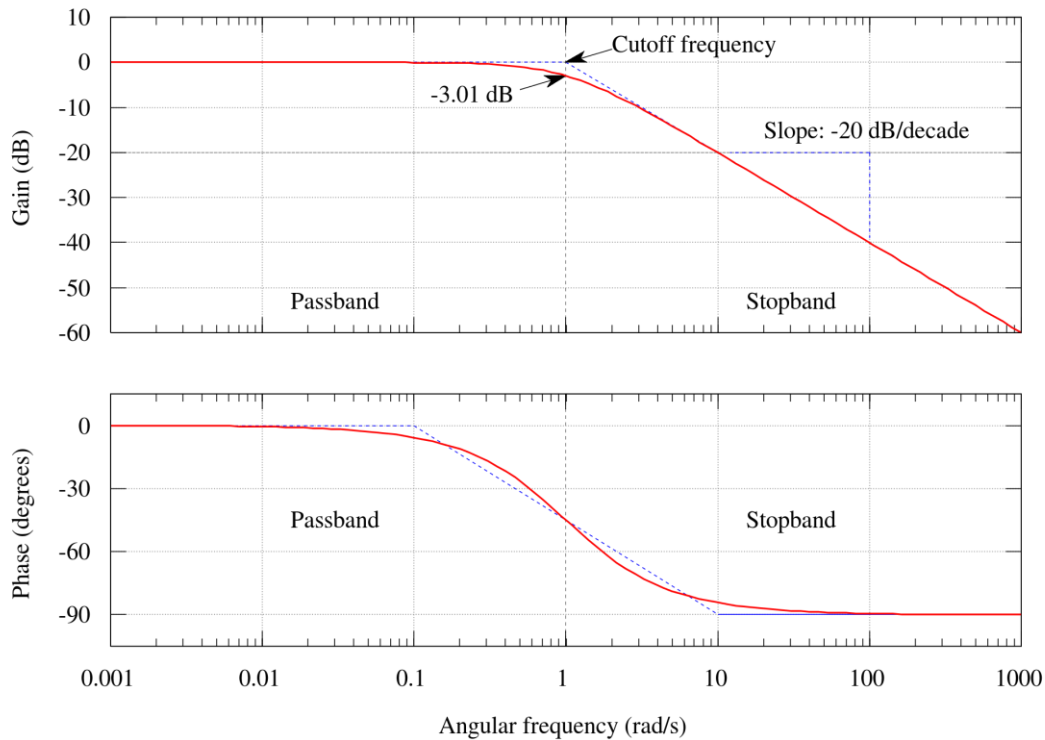


Figure 3-3: Example of a Butterworth filter

3.3.2 Butterworth Band-Pass Filter

A Butterworth filter (figure 3-3) is designed to get a frequency response as flat as possible in the pass-band. After the 3dp Hamming window filter we designed a 2nd order Butterworth filter to attenuate frequencies outside the interest band. As previously mentioned, the Cut-Off used was between the HR range of the person (minFreq as the low frequency and the maxFreq as the high frequency). Although other filters can be tested, however we preferred this filter because its flat pass-band and stop-band can help avoid

ripples like other IIR filters. This will prevent certain frequencies from being favoured over others within the valid range.

3.4 Implementation

This part will describe the steps of the implementation process of the two methods that we introduced previously.

3.4.1 Input Video

The input video is the facial recording video taken from the participants (Figure 3-4). The videos were recorded using a digital video camera inside our Lab, with artificial light as the only source of light. All the input videos were taken in true color (RGB24 with 8 bits per channel) at approximately 30 frames per second (fps) with 720 x 480 as the pixel resolution. These videos were saved in AVI format on the computer for further analysis. Further details concerning the input video will be provided in chapter 6.

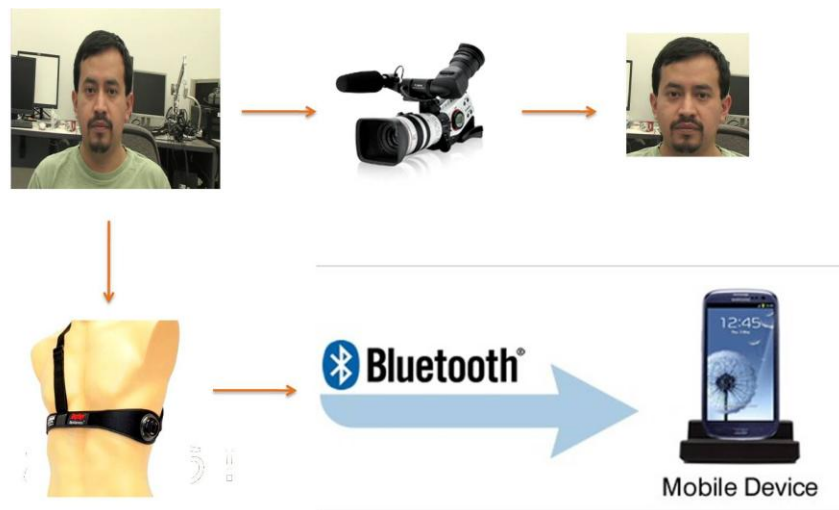


Figure 3-4: Experimental setup. Contact measurements of HR and HRV were taken from Zephyr BioHarness belt. The input video was taken from digital video camera at 29.97 fps and 720 x 480 resolution.

3.4.2 Choosing the Region of Interest (ROI)

The first step in our approach for recovering the HRV signal from a facial recorded video is to decide on the ROI that is most appropriate to use. In order to do this, we use Open Computer Vision library [33], which uses a boosted cascade classifier [34]. This enable the calculation of the facial coordinates. To be able to use OpenCV with MATLAB, a MATLAB MEX functions for OpenCV library [44] was necessary. OpenCV comes with a pre-trained classifier for the frontal face. This classifier was used in the code. A set of simple classifiers are used by the cascade, and these classifiers are used on all of the regions of interest consecutively. Boosting is the name of the weighted vote algorithm that is applied to build the classifier at each stage. If all stages pass, this indicates that there is a face within the region, otherwise the region will be rejected.

The Opencv algorithm [33] produces the coordinates, as well as the height and width that describe a box that surround the face region. In the face, the highest regions of blood flow are the forehead and the cheek [45]; thus the regions within the box which were the most relevant were the forehead and cheeks. To obtain the ROI, we select the box centered 40% width and full height of the box. The ROI taken from the first frame is used for all the frames to maintain its stability which minimize the motion artifacts.

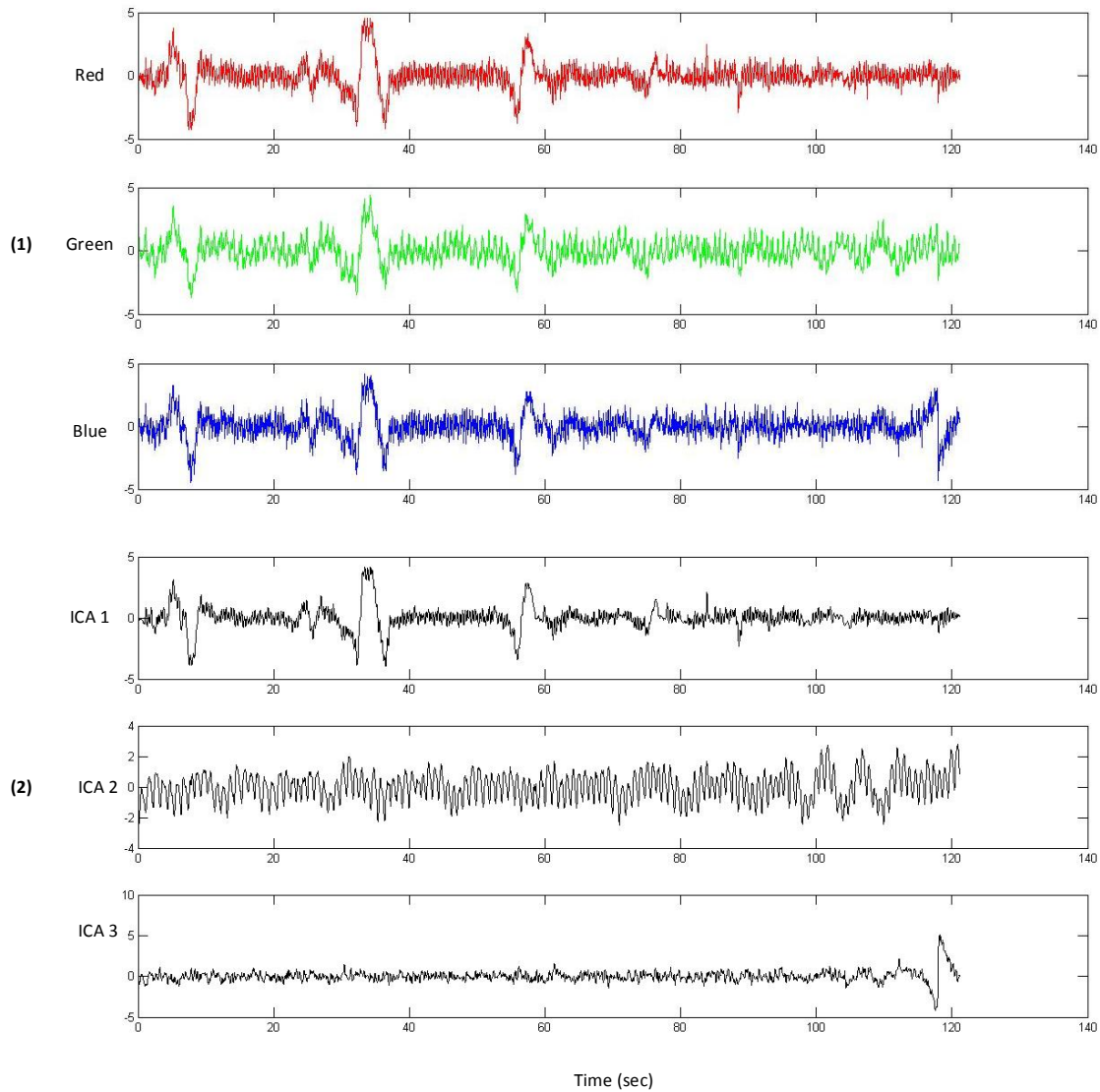


Figure 3-5: The signals extracted from the facial video. (1) The red, green and blue sources after normalization. (2) The independent components after ICA, the second component is chosen as the final output signal

3.4.3 Joint Approximate Diagonalization Eigenmatrices (JADE)

JADE module is used in this thesis to perform a blind source separation of real signals which is based on the Independent Component Analysis (ICA). JADE algorithm [46] was implemented in Matlab (Figure 3-5) .



Figure 3-6: An example of using the EVM for color magnification only on the ROI and not all the frame

3.4.4 EVM Implementation

The first difference between this approach and the original EVM approach [38] is that the magnified video was not used as the input video. Better performance was already shown to result if only the ROI rectangle is magnified, rather than the entire video frame [39] (figure 3-6). As previously mentioned different filter techniques are used in the EVM; however, the Gaussian pyramid is preferred in the case of color amplification with a narrow band-pass filter. In Matlab implementation, we use the EVM open source code and select the "amplify_spatial_Gdown_temporal_ideal.m" for this work.

3.4.5 Signal Filtering

In our two methods for recovering the PPG signal from a recorded facial video, we used a 5 points moving average ($k=5$) filter. After the moving average filter, we implemented a 3dp Hamming window filter and in the Butterworth filter, we found that $\alpha=5$ and $\beta=5$

rendered good results. All these filters were implemented using the Signal Processing Toolbox in Matlab.

3.5 Physiological Parameters' Extraction

Previously, we explained the two methods for extracting the PPG signal from a recorded video of the human face (by filtering the signal with the ICA method and by the EVM method). In this section, we detail how we extract the RR intervals from the HRV time series and estimate its associated physiological parameters (time and frequency domain parameters). This step is common for the two methods presented in section 3.2 and 3.1.

As previously mentioned, we are using the HR range of the subject as the cut-off frequency for the filters. Thus, the first section in this chapter explains the HR estimation protocol.

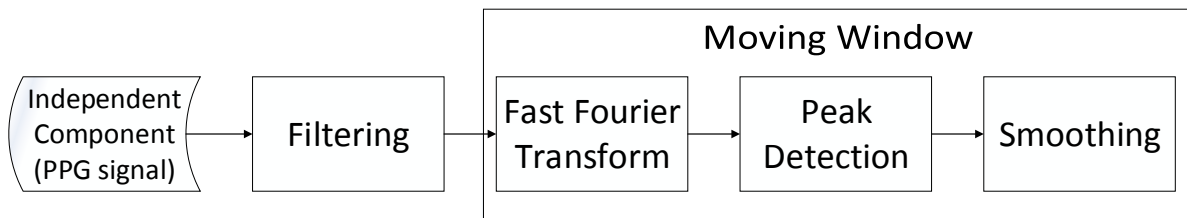


Figure 3-7: HR estimation procedure

3.5.1 Heart Rate Estimation

There are several methods to estimating HR using a facial video [8] [47] [48] [39]. In this thesis, the method we employed is based on [48] and [8] (figure 3-7). The same source video pre-processing steps explained in chapter 3 are used here. The chosen independent component outputted by the ICA based method is the input signal of this phase.

- **Filtering:** To discard the unwanted frequencies, a band-pass filter is applied to the signal. This makes the HR signal smoother and less noisy. An IIR filter, more specifically a second order Butterworth filter, is used in this case. The normal HR range for humans is between 45 and 180 bpm, therefore the cut-off frequency given to the filter is 0.75 - 3 Hz.
- **Fast Fourier Transform (FFT):** The frequency domain is the main interest; therefore, to convert the time domain signal to the frequency, the Discrete Fourier Transform (DFT) is usually used. For less processing time the FFT is applied. The FFT's complexity is much lower than DFT .
- **Moving Window:** If dynamic HR estimation is desired, a continuous estimation of HR is required. Therefore, the FFT procedure, as well as the peak detection and smoothing procedures are performed every 0.5 seconds in a moving window of the last three seconds of signal. A duration of three seconds is chosen as it gave the best results relative to other lengths that we attempted to use. These steps are performed every 0.5 seconds to increase the number of HR calculations which will be later smoothed to give a better estimation. To reduce the artificial high frequencies that appear when the signal is treated as periodic by the FFT, a Hanning window was used to bring the edges to zero.
- **Peak Detection:** After performing the FFT, the highest peak was detected in the interest band (which is the HR band between 45 and 180). The detected peak is then converted to the right frequency in the FFT vector.

3.5.2 Cubic Spline Interpolation

Due to the limited frame rate of the camera, (we use 30 fps in our data collection) the resulting PPG signal is considered to be at low resolution signal. This would drastically affect the HRV analysis. To resolve this issue, a cubic spline interpolation (Figure 3-8) is

implemented at 240 Hz (which is close to the sampling frequency of the ECG sensor used to validate our results in the experiment of chapter 4).

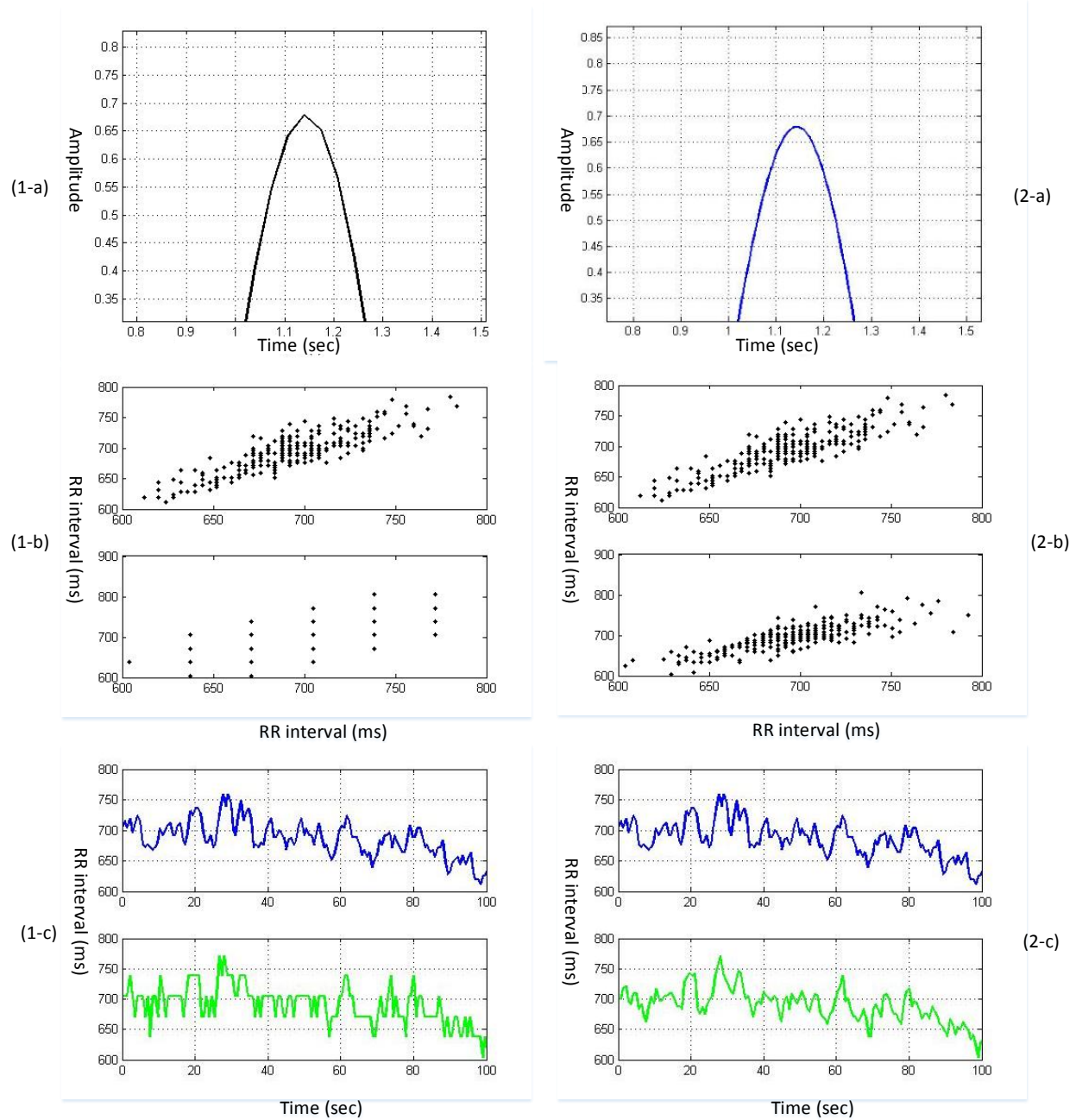


Figure 3-8: HRV analysis with and without cubic spline interpolation. (1-a) (2-a): peak of the signal without and with interpolation. (1-b) (2-b): scatter plot of the RR interval without and with interpolation. (1-c) (2-c) : RR interval without and with interpolation

3.5.3 Peak Detection

After the signal is interpolated, a peak detection algorithm is applied to it. Although there are many ways to detect the peaks, in this thesis the method used is based on McDuff *et al.* work [9].

A moving window of 0.25 seconds is used in this algorithm. The highest point is detected inside the window. If this detected point is greater than the highest point in the previous window and greater than the highest point in the next window, it will be selected as a peak. The same procedure will be repeated for the whole signal.

The length of the window was decided at 0.25 because it was proven that 0.25 seconds will give the most correlated results with the ground truth measurement [9].

3.5.4 Artifact Removal

The RR interval is constructed after the peak detection procedure by calculating the time between consecutive peaks. Head motion, variation in the illumination and other artifacts can influence the accuracy of the HRV analysis by affecting the RR interval. To prevent this from happening, the HRV series needs to be filtered. Previous studies used different filters [7] [9] [2] [49]. The interval correction used in this thesis is based on the one done in [49].

The last step before feature extraction is the interval correction (Figure 3-9). The PPG signal is very sensitive to motion artifacts[49], which is why it is mandatory to correct the internal beat intervals (IBIs) before proceeding to the HRV analysis.

A window of 10 seconds is chosen in this procedure. The analysis of the window consisted of three steps:

1. Discard all the intervals if the window only contains four intervals,
2. Discard all intervals that are greater or less than the median of all the intervals in the current window.
3. Discard all the intervals if more than half of the intervals are discarded from the window.

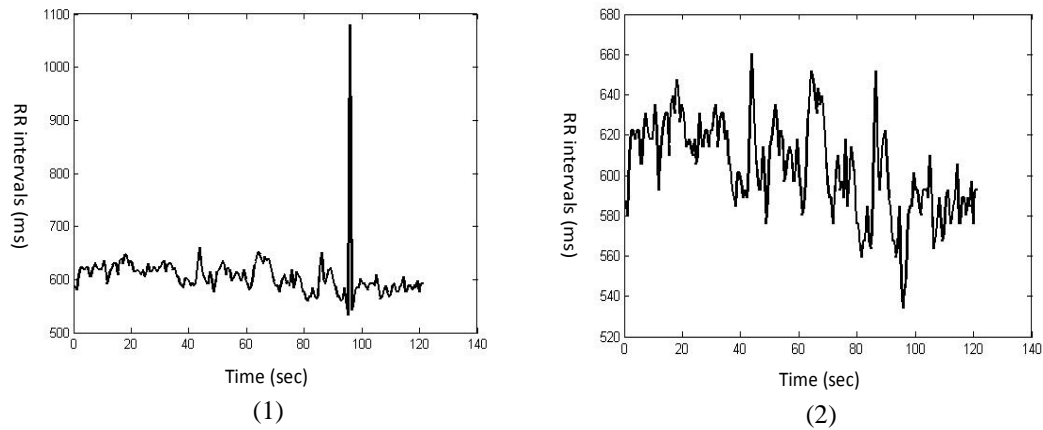


Figure 3-9: Artifact removal (1) RR interval before artifact removal (2) RR interval after artifact removal

3.5.5 Time Domain Parameters

HRV analysis produce two types of parameters. The first type is the time domain parameters. This type is simple and easy to perform because it is applied directly to the RR interval values.

The most common time domain parameter is the standard deviation of the RR interval, which is represented by SDNN.

$$SDNN = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (RR_j - \overline{RR})^2}$$

where N represents the total number of successive intervals and RR_j represent the value of the j th interval. Another parameter is the root mean square of successive differences (RMSSD) defined as follows :

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (RR_{j+1} - RR_j)^2}$$

A third measure which can be derived from the RR interval is the NN50. This parameter is the total number of successive RRs which differ by more than 50 ms. However, what is usually used is the proportion of the NN50 value divided by the number of RRs.

$$pNN50 = \frac{NN50}{N-1} \times 100$$

The last one is the mean of the RR interval which is represented by \overline{RR} .

3.5.6 Frequency Domain Parameters

The HRV analysis of the frequency domain is achieved by a powerspectrum density (PSD) estimation. There are two common methods to carry out the PSD estimation, the first one is based on FFT and the second one is based on Autoregressive parametric modeling (AR modeling) [26]. The FFT method is based on Welch's periodogram. The

Welch's periodogram method improve the estimated spectrum but would decrease the noise in the variance [50].

After obtaining the PSD of the RR interval, the frequency domain parameters could be extracted. The main parameters we are interested in are Low Frequency (LF), High Frequency (HF) and LF/HF . As previously mentioned the LF frequency range is 0.04 - 0.15 Hz; Therefore to measure the LF, the area between 0.04 and 0.15 Hz in the PSD curve is the LF parameter. For the HF, the area between 0.15 and 0.4 Hz is the area measured in the PSD curve to get the HF parameter. The values of these parameters is in normalized units (n.u.). At last the LF/HF ratio is calculated.

Parameter	Description
HR (bpm)	The heart rate
\overline{RR} (ms)	The mean of RR intervals
SDNN (ms)	The standard deviation of RR intervals
RMSSD	the root mean square of successive RR intervals difference
NN50	Number of RR intervals pairs differing more than 50 ms
pNN50	The division of NN50 over total number of RR intervals (%)
LF (n.u)	The low frequency extracted from RR intervals after applying PSD
HF (n.u)	The high frequency extracted from RR intervals after applying PSD
LF/HF	The LF frequency value divided by the HF value

Table 3-1: Summary of the HRV analysis parameters

Chapter 4.

Evaluation of proposed methods

4.1 Objective

In this thesis, a new approach to the extraction of HRV signal from a facial video is introduced. We hypothesize that by adding a Butterworth filter to the original methods and using the specific HR range of the subject to determine its cut-off frequencies, it will produce a more accurate HRV signal. In our evaluation, we will rely on HRV signals extracted from an ECG sensor as ground truth to validate our methods.

4.2 Participants

Twelve subjects participated in the experiments. The participants were different races, different ages (21 - 40), 9 males and 3 females. Some of the males had facial hair.

4.3 Procedure

The experiments took place inside a room without any windows and artificial lights were the only source of illumination. The participants were asked to sit in front of the camera and remain still while a facial video is recorded by a standard digital camera at 30 frame per second (fps) with pixel resolution of 720 x 480 and saved as AVI extension. The ground truth HEV signal was collected using the Bioharness ECG sensor.

4.4 Evaluation Measurements

To Assess the results, Pearson Correlation was used. Pearson's correlation is usually used to determine how well are the data sets related. If the value is between 0.5 and 1.0 that means there is high correlation. Also, the Mean Absolute Error (MAE) between the ground truth and calculated PPG signal is used, where a low MAE value means a low error. This formula will show the precision of the calculations. Root-mean-square-error (RMSE) was also employed .

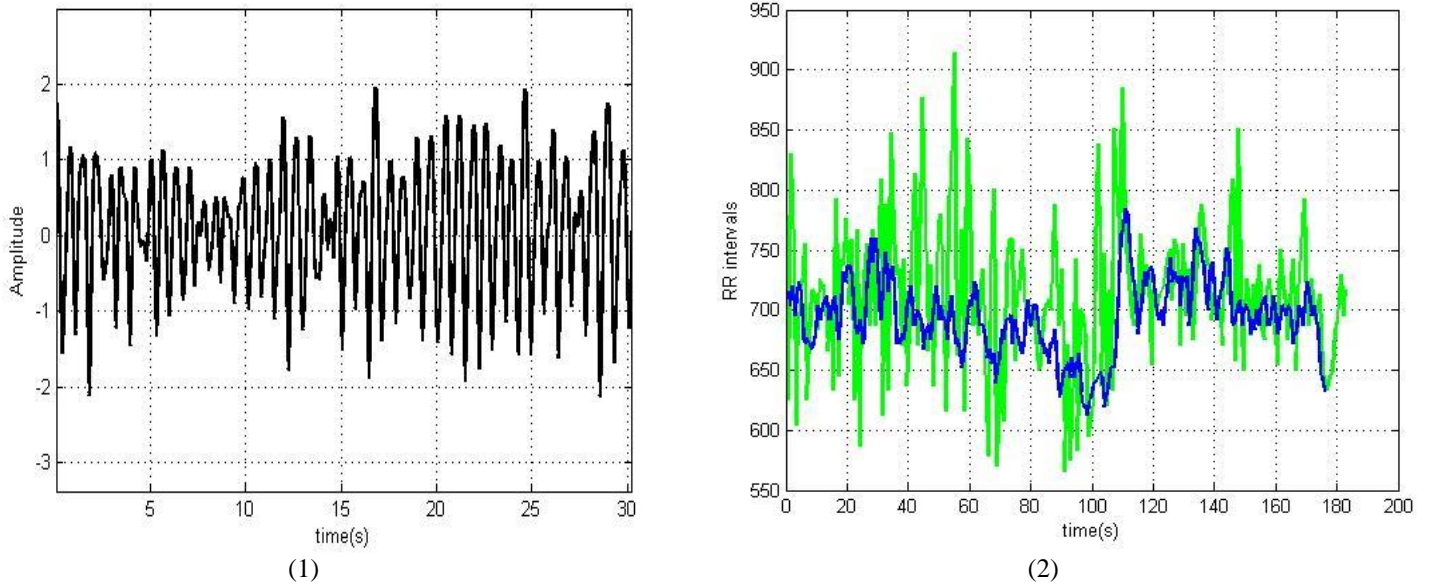
4.5 Results

After the experiments are done, the video is taken and ROI is decided as mentioned in the methodology and then the evaluation starts.

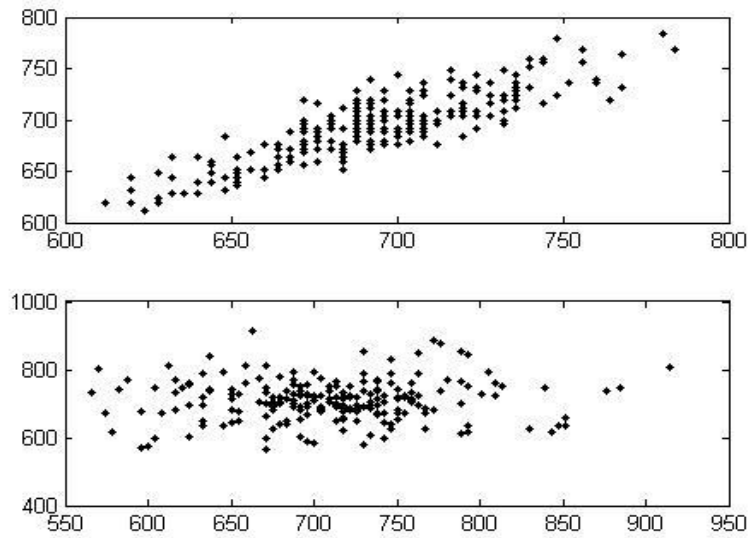
4.5.1 HRV Analysis After EVM

We started the evaluation by analysing the input video with EVM and temporal filtering Cut-Off between 45 and 180 bpm (the range of a normal human heart beat). The Butterworth filter used had the same Cut-Off range too. The Signal was very noisy

which made the RR intervals extracted noisy too and affected the results of the HRV analysis as shown in Figure 4-1. This will be referred to as EVM1.



3(a)



3(b)

Figure 4-1: EVM with Cut-Off between 45 -180 bpm for temporal filter and Butterworth filter. (1) PPG signal. (2) Blue is the RR interval from the contact sensor and Green is the RR interval from the video. (3a) Poincare plot of the RR interval from the contact sensor (3b) Poincare plot of the RR interval from the video

The second attempt was leaving the cut-off of the temporal filtering between 45 and 180 but changing the cut-off frequencies of the Butterworth filter to the values that were calculated using equations (3) and (4). This gave better results than the first attempt as shown in figure 4-2. This will be referred to as EVM2

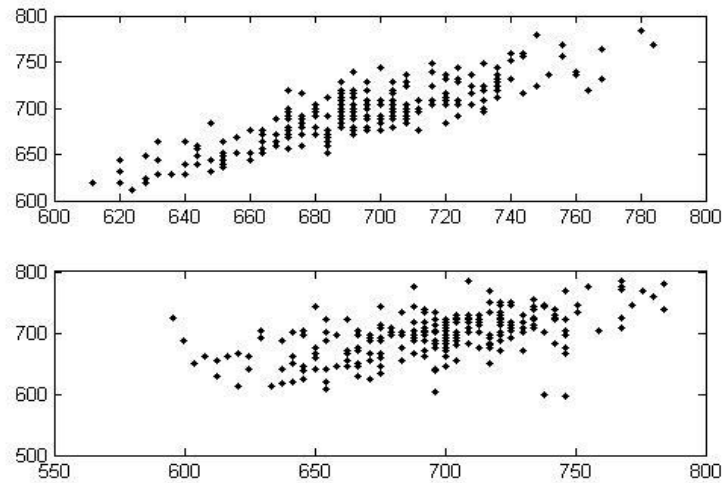
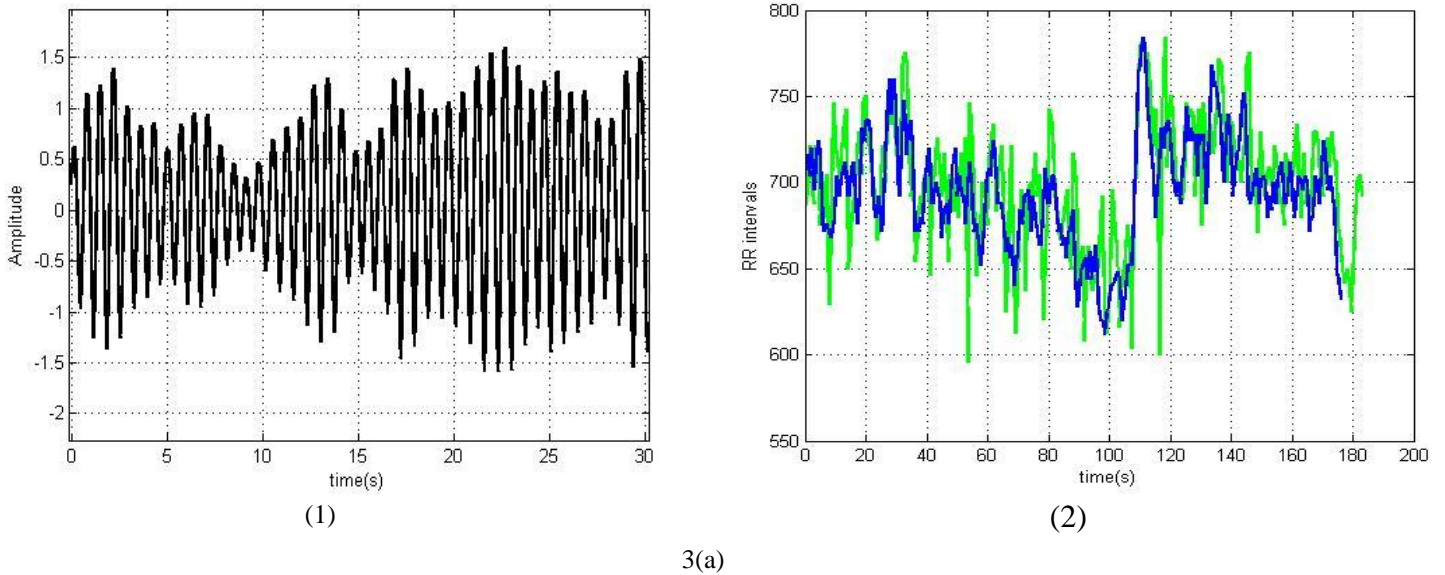
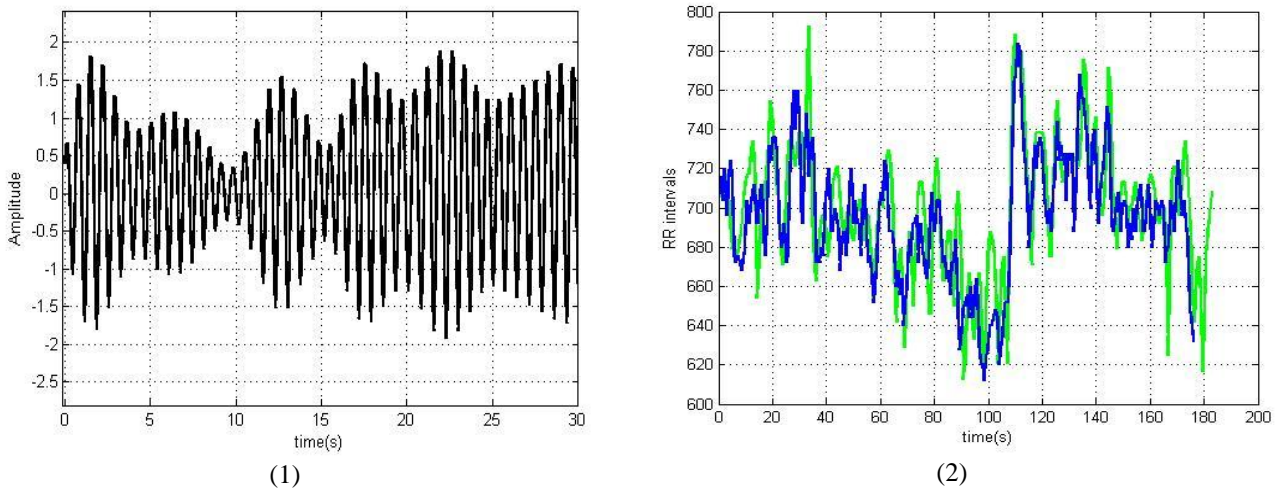
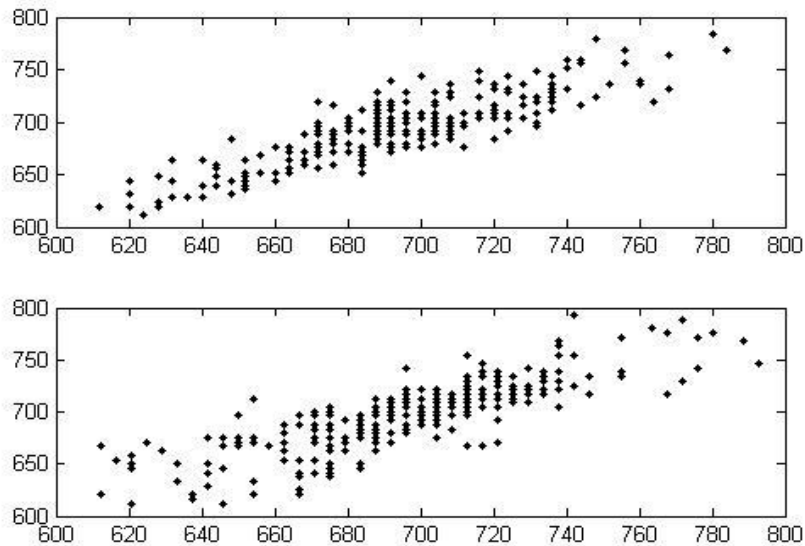


Figure 4-2: EVM with Cut-Off between 45 -180 bpm for temporal filter and the range of the subject's HR as Cut-Off for Butterworth filter. (1) PPG signal. (2) Blue is the RR interval from the contact sensor and Green is the RR interval from the video. (3a) Poincare plot of the RR interval from the contact sensor (3b) Poincare plot of the RR interval from the video

The third attempt is based on the method described in chapter 3 section 3-2. The cut-off frequency for both the temporal and Butterworth filters we calculated using equations (3) and (4). The result of this attempt is shown in Figure 4-3. This will be referred to as EVM3.



3(a)



3(b)

Figure 4-3: The range of the subject's HR as Cut-Off for the temporal and Butterworth filters. (1) PPG signal. (2) Blue is the RR interval from the contact sensor and Green is the RR interval from the video. (3a) Poincare plot of the RR interval from the contact s sensor (3b) Poincare plot of the RR interval from the video

4.5.1.1 Results of The HRV Analysis

The parameters (time and frequency) were extracted in the three attempts and compared. The Mean Absolute Error was presented in table 4-1 after performing equation 7

$$MAE = \frac{1}{N} \sum_{i=1}^n |f_i - y_i| \quad (5)$$

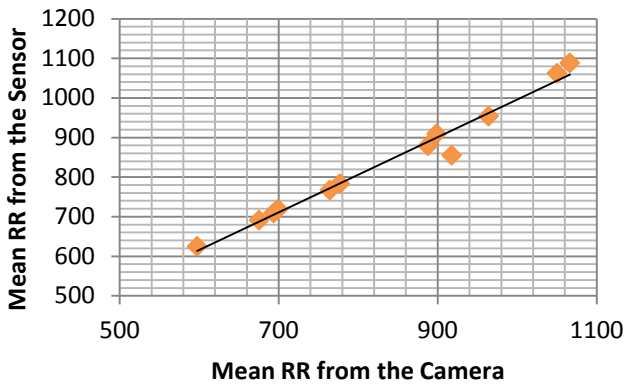
Where f_i is the prediction and y_i is the true value (the ground truth)

Parameters	MAE (EVM1)	MAE (EVM2)	MAE (EVM3)
LF (n.u)	4.34	0.31	0.28
HF (n.u)	18.58	1.88	0.76
LF/HF	0.66	0.48	1.69
SDNN (ms)	1.32	0.13	0.1
Mean RR (ms)	0.021	0.006	0.006
Mean HR (bpm)	0.02	0.012	0.006

Table 4-1: The Mean Absolute Error of the HRV parameters for the three attempts of the EVM methods

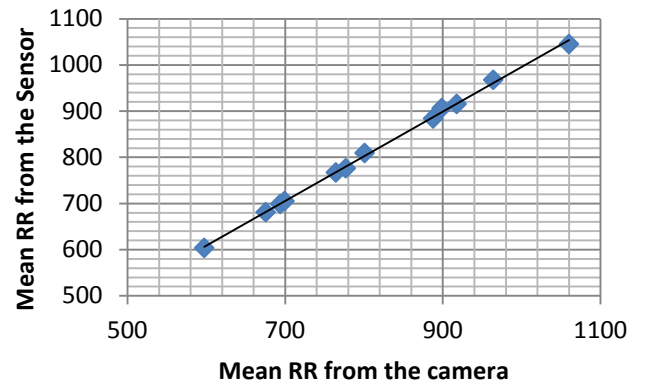
To see how well these values are related to the ground truth from the sensor, Pearson's correlation is applied. Figures 4-4,5,6,7 show scatter plots of the data and the r values (the Pearson's correlation value).

$r = 0.988$



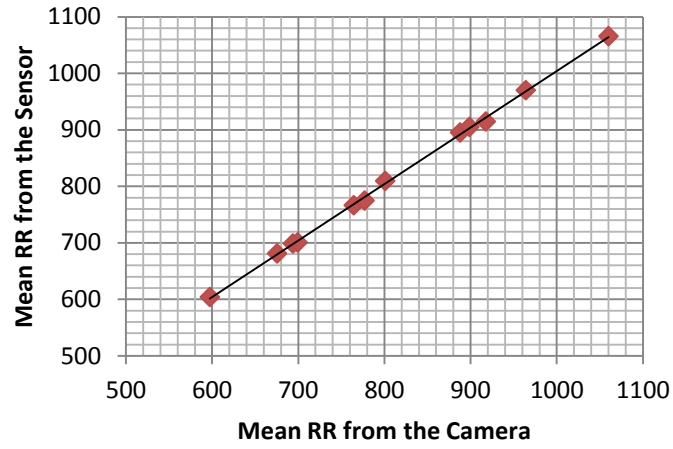
EVM 1

$r = 0.999$



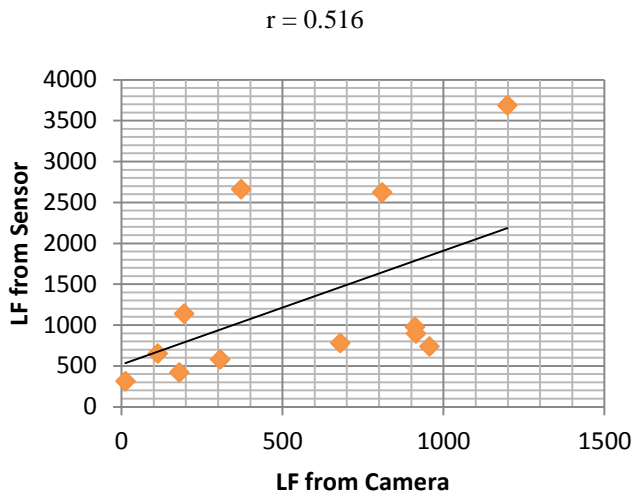
EVM 2

$r = 0.999$

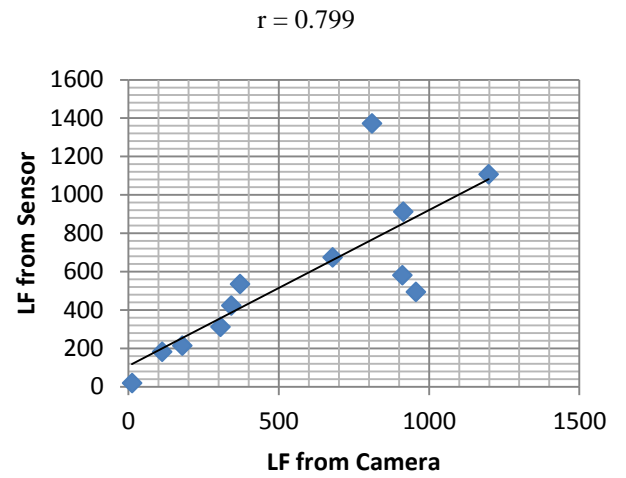


EVM 3

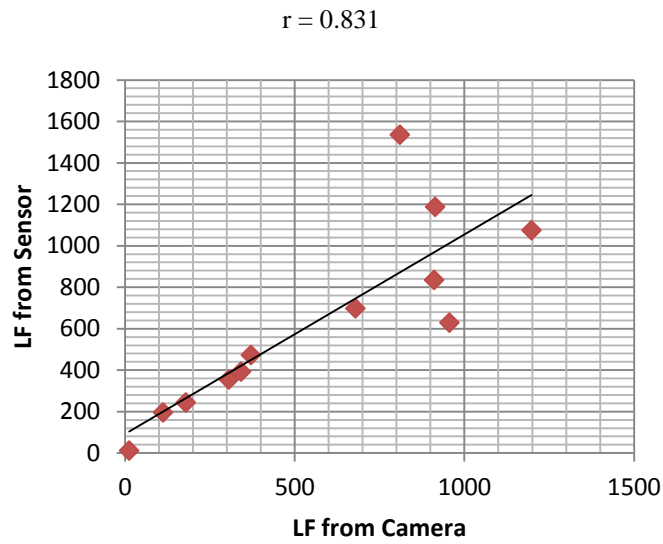
Figure 4-4: Scatter plots and Pearson's Correlation values of the Mean RR between a camera and sensor of the three EVM based methods



EVM 1

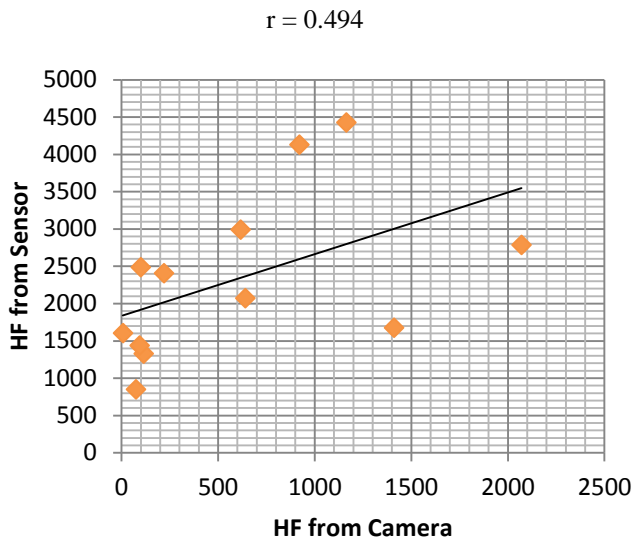


EVM 2

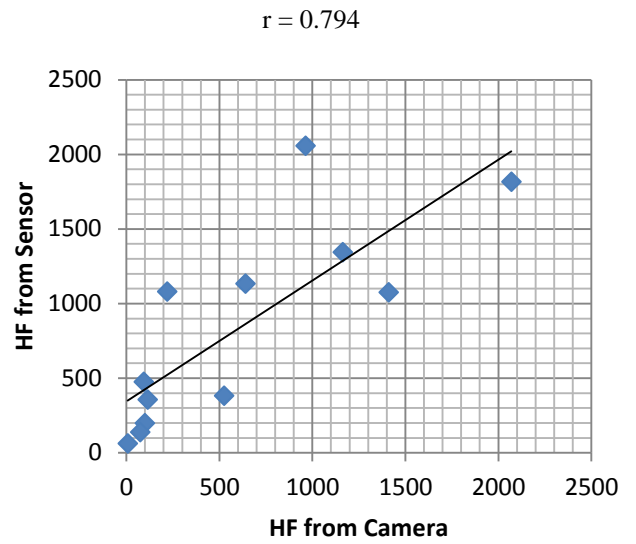


EVM 3

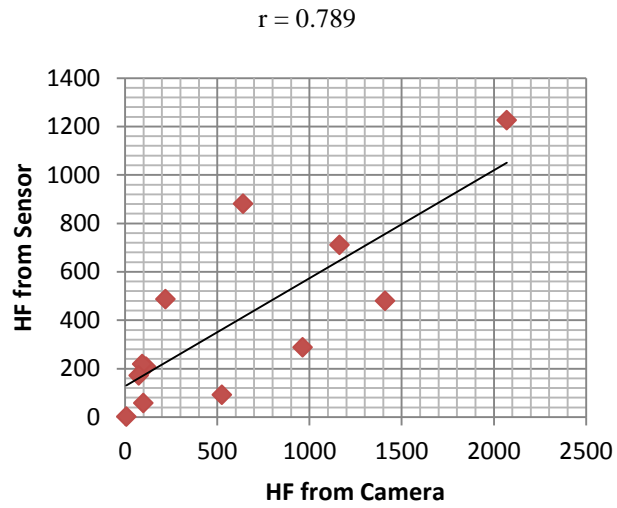
Figure 4-5: Scatter plots and Pearson's Correlation values of the LF between a camera and sensor of the three EVM based methods



EVM 1



EVM 2



EVM 3

Figure 4-6: Scatter plots and Pearson's Correlation values of the HF between a camera and sensor of the three EVM based methods

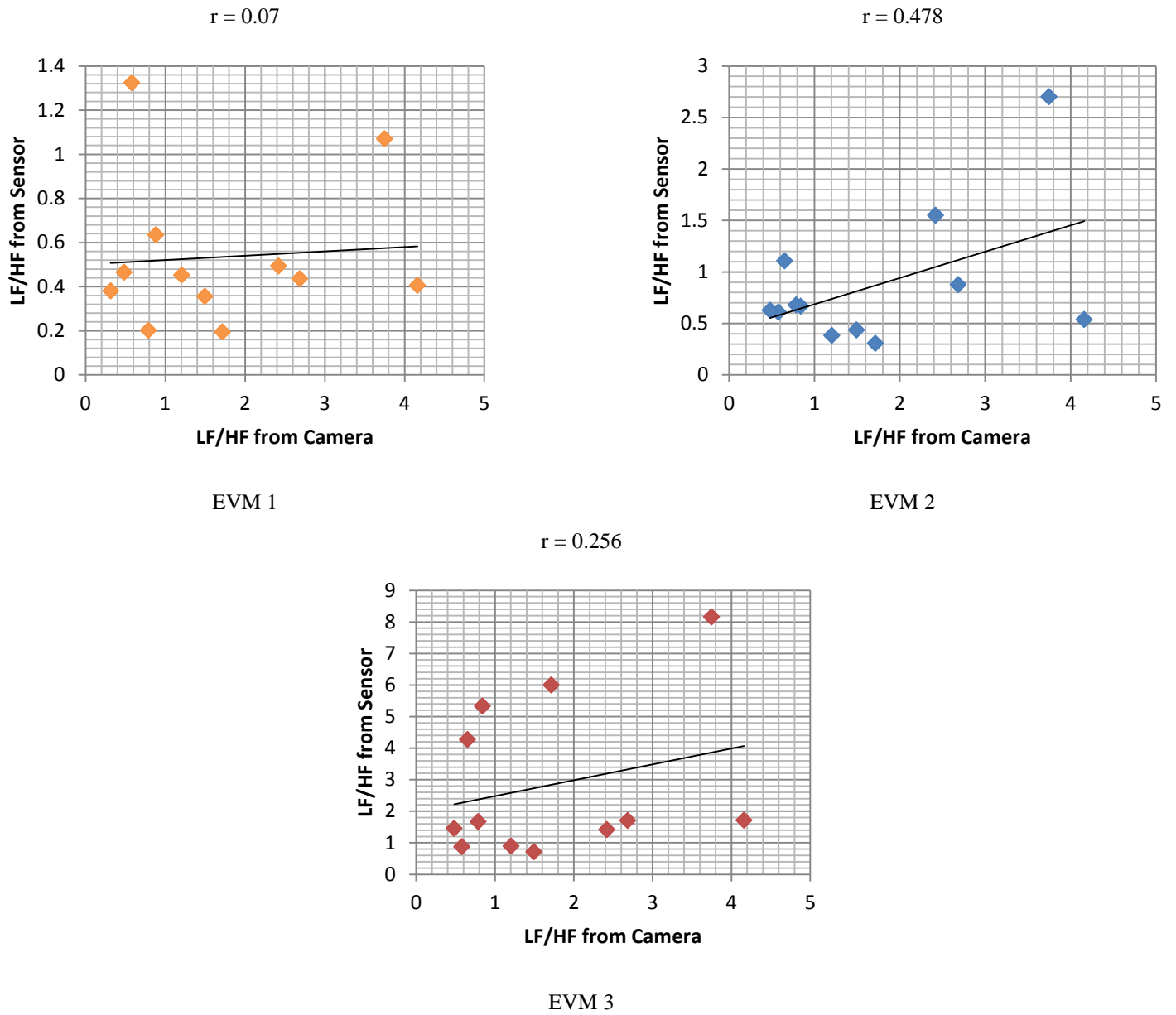
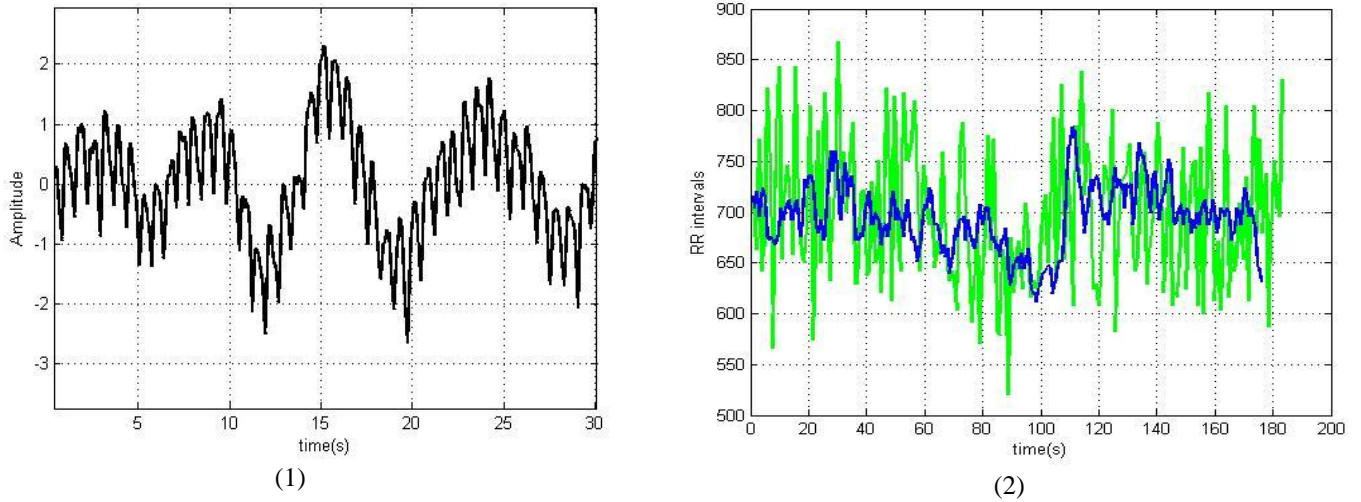


Figure 4-7: Scatter plots and Pearson's Correlation values of the LF/HF between a camera and sensor of the three EVM based methods.

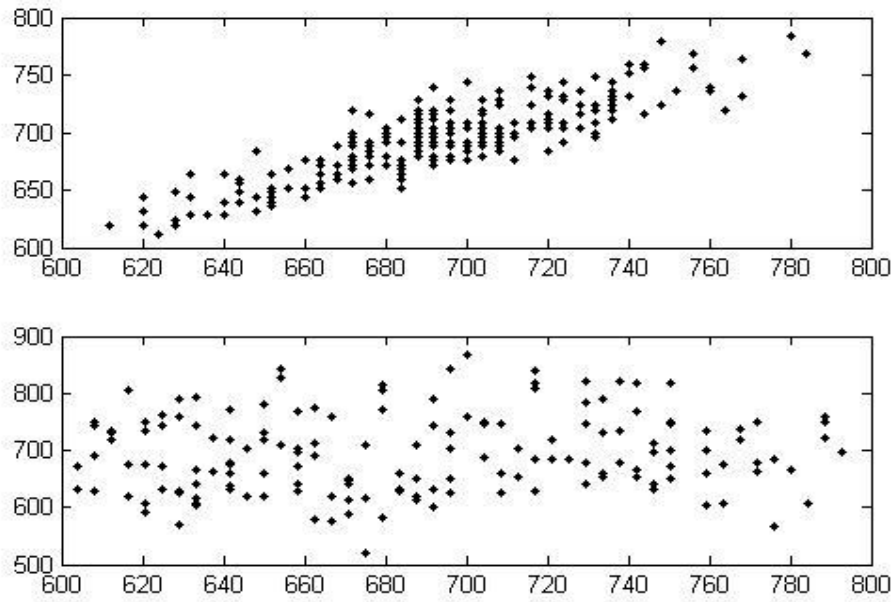
4.5.2 HRV Analysis After ICA

In this part two sets of experiments are analysed. The first set is based on McDuff *et al.*[9] methodology using just a 3 db Hamming window filter with a Cut-Off between

45 and 180 bpm (normal human HR range). The HRV analysis on this set is shown in Figure 4-8.



3(a)



3(b)

Figure 4-8: HRV analysis after 3 db Hamming window with a Cut-Off between 45-180 bpm. (1) PPG signal. (2) Blue is the RR interval from the contact sensor and Green is the RR interval from the video. (3a) Poincare plot of the RR interval from the contact sensor (3b) Poincare plot of the RR interval from the video

The second sets of experiments(Figure 4-9)is based on our approach described in section 3-1.

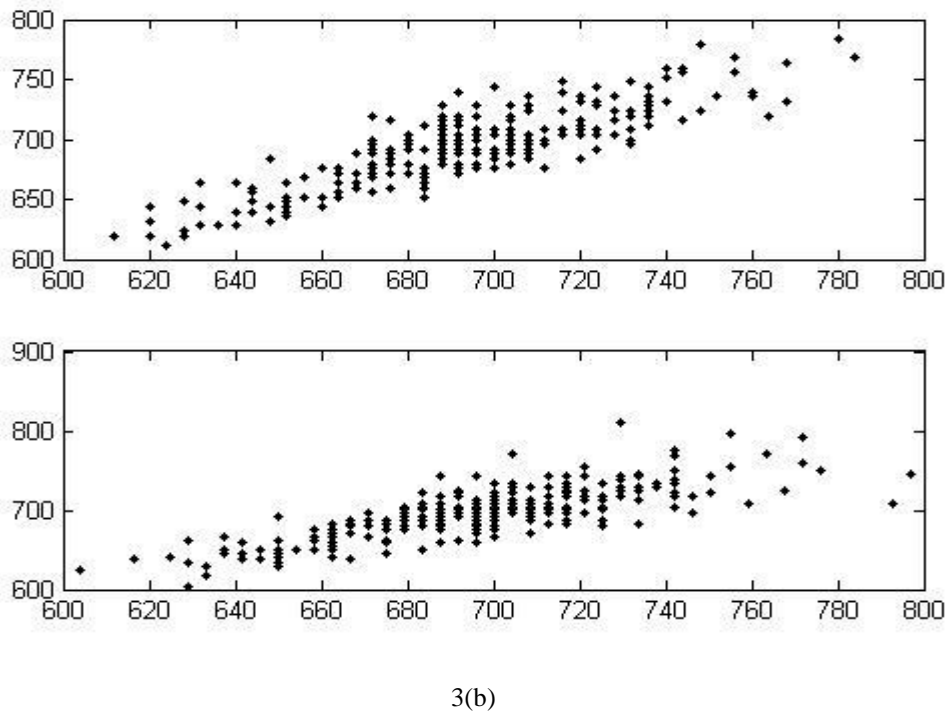
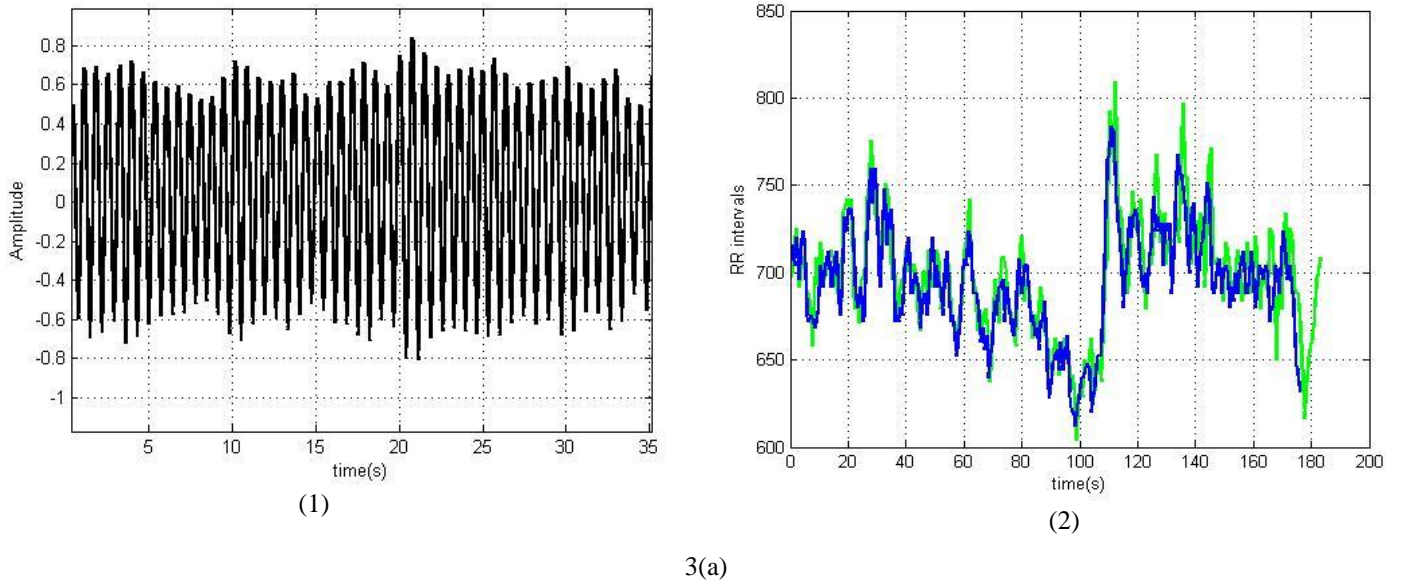


Figure 4-9: HRV analysis after the proposed ICA method (1) PPG signal. (2) Blue is the RR interval from the contact sensor and Green is the RR interval from the video. (3) Poincare plot of the RR interval from the contact sensor (3b) Poincare plot of the RR interval from the video

4.5.2.1 Results of the HRV analysis

The parameters (time and frequency) were extracted in the two sets of the experiments and compared. The Mean Absolute Error was presented in table 4-2 after performing equation 7.

Parameters	MAE (State of The Art)	MAE (Proposed ICA)
LF (n.u)	6.3	0.34
HF (n.u)	28.4	0.57
LF/HF	0.64	0.419
SDNN (ms)	1.72	0.08
Mean RR (ms)	0.012	0.005
Mean HR (bpm)	0.012	0.006

Table 4-2: The Mean Absolute Error of the HRV parameters for the two ICA methods

To see how well these values are related with the ground truth from the sensor, Pearson's correlation is applied. Figures 4-10,11,12,13 show scatter plots of the data and the r values (the Pearson's correlation value).

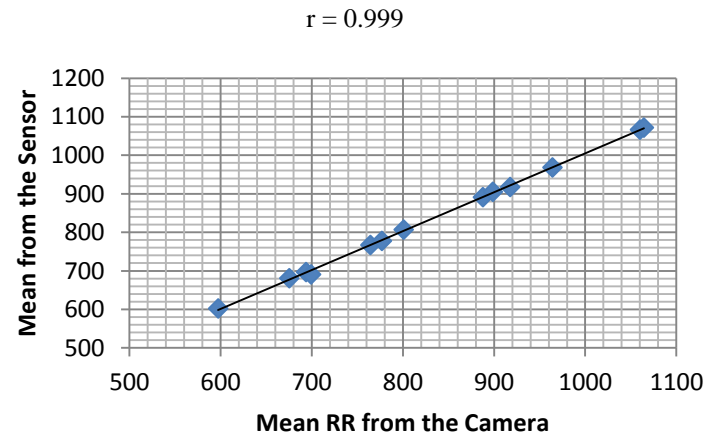
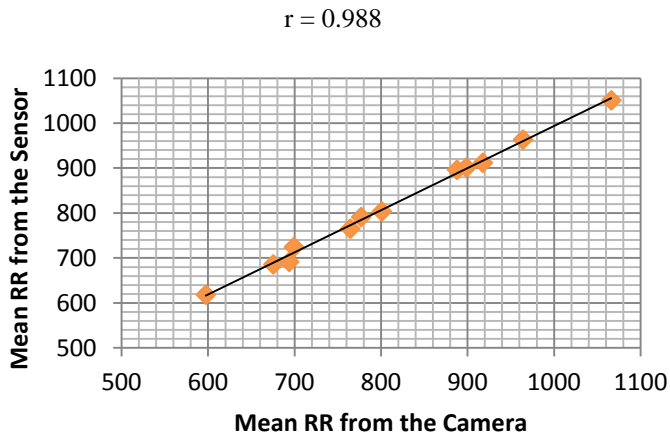


Figure 4-10: Scatter plots and Pearson's Correlation values of the Mean RR between a camera and sensor of the three EVM based methods. (1) State of the art ICA (2) Proposed ICA.

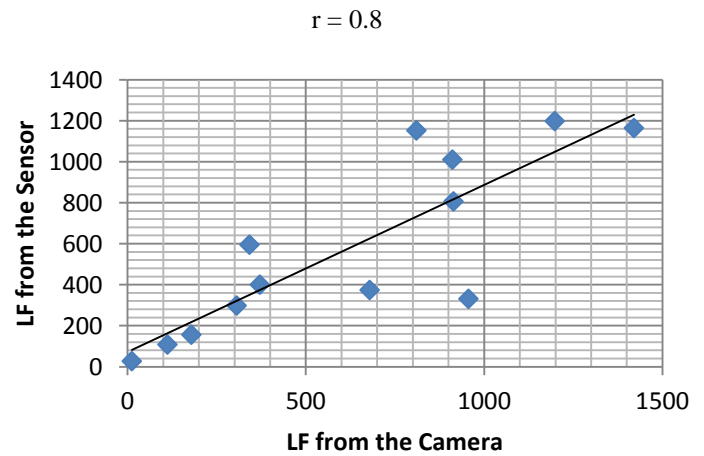
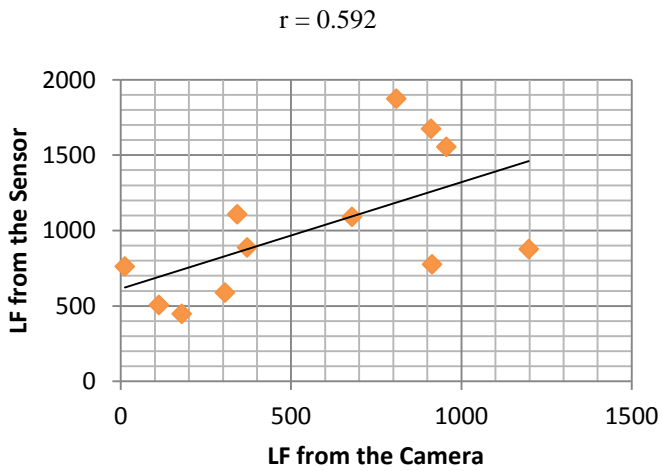
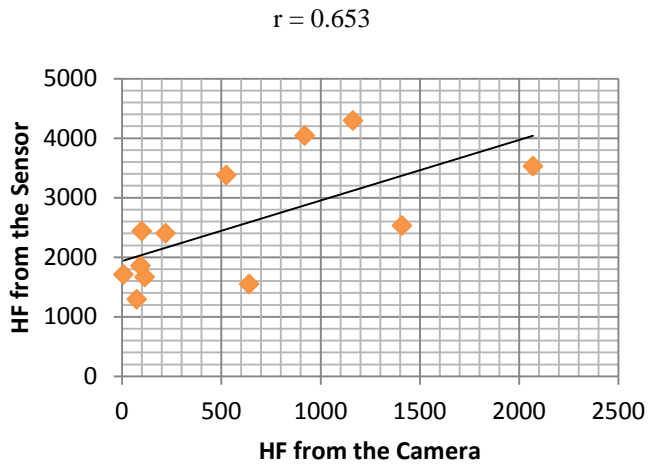
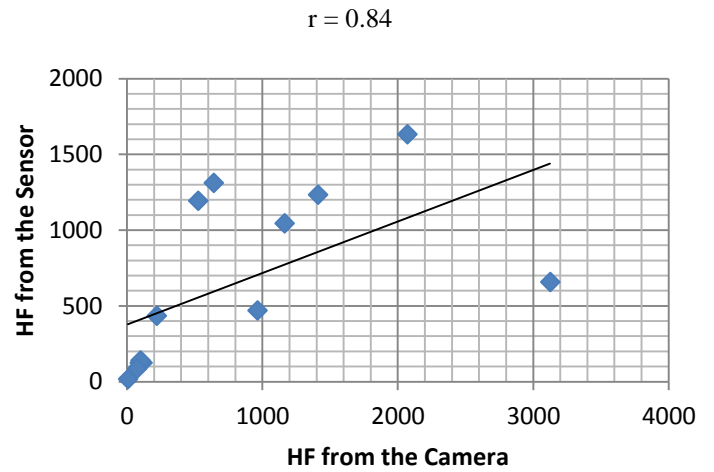


Figure 4-11: Scatter plots and Pearson's Correlation values of LF between a camera and sensor of the three EVM based methods. (1) State of the art ICA (2) Proposed ICA.

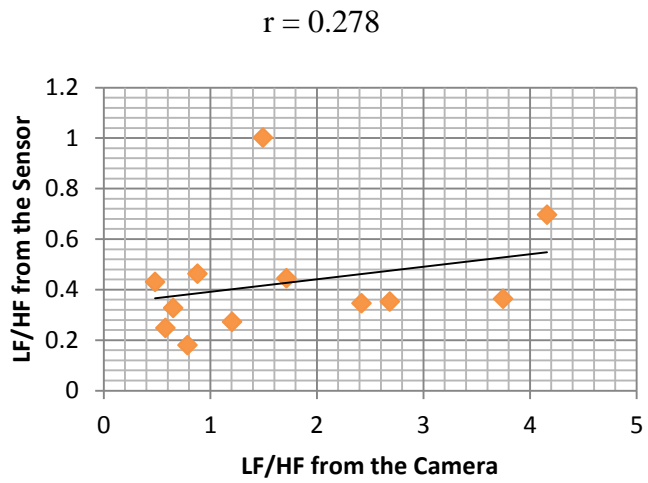


(1)

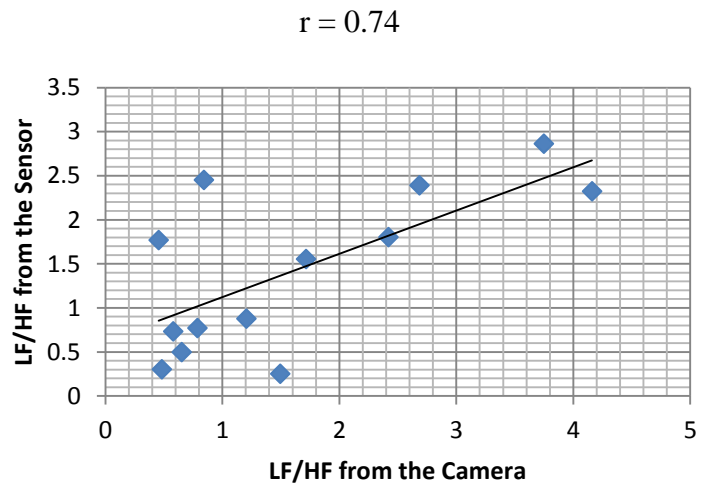


(2)

Figure 4-12: Scatter plots and Pearson's Correlation values of HF between a camera and sensor of the three EVM based methods.(1) State of the art ICA (2) Proposed ICA.



(1)



(2)

Figure 4-13: Scatter plots and Pearson's Correlation values of LF/HF between a camera and sensor of the three EVM based methods.(1) State of the art ICA (2) Proposed ICA.

4.5.3 The Root-Mean-Square-Error (RMSE)

The last measure used in the thesis to compare the results is the Root-Mean-Square-Error. Table 4-3 shows the comparison between all methods, EVM based and ICA based.

RMSE	EVM1	EVM2	EVM3	State of the art ICA	Proposed ICA
LF/HF	1.69	1.36	2.6	1.74	0.86
Mean RR	23.26	6.58	5.51	12.23	0.55

Table 4-3:Table 6 3: RMSE results for LF/HF and the mean RR parameters of all methods (EVM and ICA)

4.6 Analysis of the results

In a first approach of this work, the videos that we recorded where artificial lights were the only illumination source was tested based on previous works. As seen in the case the state of the art ICA and EVM1, the extracted signal was very noisy (Figure 4-8 and 4-1) which rendered bad results (table 4-1 and 4-2). This was due to high frequency noise that is mixed with the PPG signal. Therefore, it can be said that the previous approaches work well only when the signal is not very noisy.

The two methods this thesis proposes aim to resolve the noise problem. EVM2 and EVM3 clean up the signal more efficiently than previous methods and therefore result in good time domain results. EVM3 showed the best results when it comes to LF component with an MAE of 0.28 while the MAE of the proposed ICA was 0.34. A higher correlation

of the LF component between the EVM3 and the ground truth compared to the correlation between the proposed ICA and the ground truth is also obtained.

Also, EVM2 and EVM3 showed impressive results but the HF and LF/HF components were corrupted because of the HF frequency noise that could not be separated from the PPG signal in the EVM based methods. Therefore, the proposed ICA method is the most convenient method and produced a clean signal and very good results in time and frequency domain.

It is true that our proposed ICA showed the greatest promise in our experiment, but the EVM based methods might be more appropriate when motion is involved or when the HF component is not important.

Chapter 5.

Conclusion and Future work

5.1 Conclusion

In this thesis, we presented two new methods for extracting HRV from video signals. The first approach is based on the EVM method and the second is based on the ICA.

To the best of the author's knowledge, this is the first successful approach from HRV extracting using EVM for extracting HRV parameters from PPG signals. This has been tested on 12 subjects from different ethnicities, genders and skin colors. This approach showed great results when it comes to the time domain parameters, but we could not eliminate of the high frequency noise, which affected the HF and LF/HF parameters. The ICA based approach solved the HF noise issue. Our proposed ICA approach made enormous progress in results as compared to previous ICA based approaches. More

specifically, this approach made it possible to deal with noisy signals, especially videos taken when there is only artificial light as an illumination source. This kind of illumination produces HF signals which interfere with the PPG signals extracted from the video.

The reason why ICA was better when it came to the HF parameter is because spatial pooling is used in the EVM technique which combines the nearby pixels together, while the ICA technique separates the signal into independent components which makes the signal less noisy. This, combined with our contributions to the previous ICA approach, made us able to achieve better results and a less noisy signal.

5.2 Future Work

Since the topic of this thesis is still in its early stages, there is a lot of room for future work. One of the important areas of improvement in this approach is the motion of the subject. In the current approach the subject needed to remain still because motion can produce bad results. Allowing the subject to move will open the door for greater applications.

Another goal is to make the HRV parameters extraction software that we built in this thesis function in real time, and not just on recorded videos. This will make the software more flexible, dynamic and useful if someone wants to use it for live feedback.

The main goal for the future is to use this approach in non-medical fields. We want to take advantage of our good results which we can achieve from a non-contact method of estimating HRV, in order to estimate emotions without solely relying on facial

expressions analysis. This will allow for more accurate results for estimating emotions because the subject will not be able to their emotions with fake expressions. A great example of an application for this idea is in analysing the emotions of participants in political debates.

References

- [1] D. Purves, G. J. Augustine, D. Fitzpatrick, L. C. Katz, A.-S. LaMantia, J. O. McNamara, and S. M. Williams, "Neuroscience. 2nd edition." Sinauer Associates, 2001.
- [2] H. Al Osman, M. Eid, and A. Saddik., "U-biofeedback: a multimedia-based reference model for ubiquitous biofeedback systems," *Multimed. Tools Appl.*, pp. 1–26, 2013.
- [3] D. McDuff, S. Gontarek, and R. Picard, "Remote Measurement of Cognitive Stress via Heart Rate Variability," *Eng. Med. Biol. Soc.*, pp. 3–6, 2014.
- [4] S. S. Ulyanov and V. V. Tuchin, "OE/LASE'93: Optics, Electro-Optics, & Laser Applications in Science& Engineering," 1993, pp. 160–167.
- [5] E. F. Greneker, "Radar sensing of heartbeat and respiration at a distance with applications of the technology." IET Digital Library, pp. 150–154, 01-Jan-1997.
- [6] W. Verkrusye, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Opt. Express*, vol. 16, no. 26, p. 21434, Dec. 2008.
- [7] M. Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7–11, 2011.
- [8] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Express*, vol. 18, no. 10, p. 10762, May 2010.
- [9] D. McDuff, S. Gontarek, and R. W. Picard, "Improvements in Remote Cardio-Pulmonary Measurement Using a Five Band Digital Camera.," *IEEE Trans. Biomed. Eng.*, vol. 9294, no. c, pp. 1–8, 2014.
- [10] S. Bakhtiari, T. W. Elmer, N. M. Cox, N. Gopalsami, A. C. Raptis, S. Liao, I. Mikhelson, and A. V. Sahakian, "Compact Millimeter-Wave Sensor for Remote Monitoring of Vital Signs," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 3, pp. 830–841, Mar. 2012.
- [11] L. Aarts, V. Jeanne, and J. Cleary, "Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit—A pilot study," *Early Hum.*, 2013.
- [12] J. Delaney and D. Brodie, "Effects of short-term psychological stress on the time and frequency domains of heart-rate variability," *Percept. Mot. Skills*, 2000.
- [13] S. C. Segerstrom and G. E. Miller, "Psychological Stress and the Human Immune System: A Meta-Analytic Study of 30 Years of Inquiry."
- [14] S. Arora and J. Bhattacharjee, "Modulation of immune responses in stress by Yoga.," *Int. J. Yoga*, vol. 1, no. 2, pp. 45–55, Jul. 2008.
- [15] G. R. Elliot and C. Eisdorfer, "Stress and human health: An analysis and implications of research. A study by the institute of Medicine," *Natl. Acad. Sci.*, pp. 11–24, 1982.

References

- [16] A. Baum, L. Cohen, and M. Hall, "Psychosomatic Medicine." [Online]. Available: http://journals.lww.com/psychosomaticmedicine/Abstract/1993/05000/Control_and_intrusive_memories_as_possible.5.aspx. [Accessed: 29-May-2015].
- [17] "American Institute of Stress, effects of stress [online] Available: www.stress.org." .
- [18] H. M. van Praag, "Can stress cause depression?," *Prog. Neuropsychopharmacol. Biol. Psychiatry*, vol. 28, no. 5, pp. 891–907, Aug. 2004.
- [19] D. Trichopoulos, X. Zavitsanos, K. Katsouyanni, A. Tzonou, and P. Dalla-Vorgia, "PSYCHOLOGICAL STRESS AND FATAL HEART ATTACK: THE ATHENS (1981) EARTHQUAKE NATURAL EXPERIMENT," *Lancet*, vol. 321, no. 8322, pp. 441–444, Feb. 1983.
- [20] E. H. HON and S. T. LEE, "ELECTRONIC EVALUATION OF THE FETAL HEART RATE. VIII. PATTERNS PRECEDING FETAL DEATH, FURTHER OBSERVATIONS.," *Am. J. Obstet. Gynecol.*, vol. 87, pp. 814–26, Nov. 1963.
- [21] L. Salahuddin and D. Kim, "Detection of Acute Stress by Heart Rate Variability Using a Prototype Mobile ECG Sensor," pp. 453–459, Nov. 2006.
- [22] S. Akselrod, D. Gordon, F. Ubel, D. Shannon, A. Berger, and R. Cohen, "Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control," *Science (80-.)*, vol. 213, no. 4504, pp. 220–222, Jul. 1981.
- [23] N. Hjortskov, D. Rissén, A. K. Blangsted, N. Fallentin, U. Lundberg, and K. Søgaard, "The effect of mental stress on heart rate variability and blood pressure during computer work.," *Eur. J. Appl. Physiol.*, vol. 92, no. 1–2, pp. 84–9, Jun. 2004.
- [24] L. Bernardi, J. Wdowczyk-Szulc, C. Valenti, S. Castoldi, C. Passino, G. Spadacini, and P. Sleight, "Effects of controlled breathing, mental activity and mental stress with or without verbalization on heart rate variability," *J. Am. Coll. Cardiol.*, vol. 35, no. 6, pp. 1462–1469, May 2000.
- [25] A. Malliani, F. Lombardi, and M. Pagani, "Power spectrum analysis of heart rate variability: a tool to explore neural regulatory mechanisms.," *Br. Heart J.*, vol. 71, no. 1, pp. 1–2, Jan. 1994.
- [26] S. M. Jr, "Digital spectral analysis with applications," *Englewood Cliffs, NJ, Prentice-Hall, Inc., 1987*, ..., 1987.
- [27] S. Cerutti, A. L. Goldberger, and Y. Yamamoto, "Recent Advances in Heart Rate Variability Signal Processing and Interpretation," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 1, pp. 1–3, Jan. 2006.
- [28] J. Webster, "Medical Instrumentation," 1998.
- [29] M. Garbey, N. Sun, A. Merla, and I. Pavlidis, "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery.," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 8, pp. 1418–26, Aug. 2007.
- [30] J. Allen, "Photoplethysmography and its application in clinical physiological measurement.," *Physiol. Meas.*, vol. 28, no. 3, pp. R1–39, Mar. 2007.
- [31] P. Comon, "Independent component analysis, A new concept?," *Signal Processing*, vol. 36, no. 3, pp. 287–314, Apr. 1994.
- [32] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*. 2004.

References

- [33] A. Noulas and B. Kröse, “EM detection of common origin of multi-modal cues,” *Proc. 8th Int. Conf. ...*, 2006.
- [34] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” ... *Vis. Pattern Recognition, 2001. CVPR ...*, 2001.
- [35] M. Tarvainen, “An advanced detrending method with application to HRV analysis,” *IEEE Trans. ...*, 2002.
- [36] J. Cardoso, “High-order contrasts for independent component analysis,” *Neural Comput.*, 1999.
- [37] A. Beatson and R. Green, “Calculating a Vagal Tone Index using Webcam Photoplethysmography,” pp. 242–246.
- [38] H.-Y. Wu, M. Rubinstein, E. Shih, J. V. Guttag, F. Durand, and W. T. Freeman, “Eulerian video magnification for revealing subtle changes in the world.” Association for Computing Machinery, 01-Jul-2012.
- [39] P. Chambino, “Android-based implementation of Eulerian Video Magnification for vital signs monitoring,” 2013.
- [40] N. Miljkovi and D. Trifunovi, “Pulse Rate Assessment : Eulerian Video Magnification vs . Electrocardiography,” 2014.
- [41] C. Liu, A. Torralba, and W. Freeman, “Motion magnification,” *ACM Trans. ...*, 2005.
- [42] Philips, “Philips Vitals Signs Camera. <http://www.vitalsignscamera.com>,” 2011. .
- [43] P. Burt and E. Adelson, “The Laplacian pyramid as a compact image code,” *Commun. IEEE Trans.*, 1983.
- [44] K. Yamaguchi, “<https://github.com/kyamagu/mexopencv>,” 2011. .
- [45] D. Datcu, M. Cidota, S. Lukosch, and L. Rothkrantz, “Noncontact automatic heart rate analysis in visible spectrum by specific face regions,” in *Proceedings of the 14th International Conference on Computer Systems and Technologies - CompSysTech '13*, 2013, pp. 120–127.
- [46] J. cois Cardoso and A. Souloumiac, “BLIND BEAMFORMING FOR NON GAUSSIAN SIGNALS,” *researchgate.net*.
- [47] S. Kwon, H. Kim, and K. S. Park, “Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 2174–2177, 2012.
- [48] I. Mellado, “Measuring heart rate with a smartphone camera-<http://www.ignaciomellado.es/blog/Measuring-heart-rate-with-a-smartphone-camera>,” 2013. .
- [49] M. Altini, “Heart rate variability using the phone’s camera-<http://www.marcoaltini.com/blog/heart-rate-variability-using-the-phones-camera>,” 2014. .
- [50] P. Welch, “The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms,” *IEEE Trans. audio Electroacoust.*, 1967.