Preprocessing Realistic Video for Contactless Heart Rate Monitoring Using Video Amplification Methodologies

By

Ahmed Alzahrani

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs

in partial fulfillment of the requirements for the degree of

Master of Applied Science

In

Ottawa-Carleton Institute for Electrical and Computer Engineering

Carleton University

Ottawa, Ontario

Abstract

This research seeks to improve the outcomes of Eulerian Video Magnification in real life scenarios. We address the core requirement in Eulerian Magnification that the person in the video be completely still. The proposed system preprocesses the video in multiple stages using subject targeting and stabilization. The resulting video is better suited to Eulerian Magnification restrictions. Our method enables the use of magnification in a variety of applications where motion is present such as monitoring the heart rate of a person using a treadmill.

Stabilization, which is the core element of our research, was achieved through two methods. First, we used face tracking to generate a stabilized video with limited motion. Second, feature detection, extraction, and matching with skin selection were used to produce a stabilized video that is ready to be processed for measuring heart rate. However, skin tone and illumination in the environment adversely affected the results. Since heart rate is monitored by counting the subtle changes in skin redness related to blood flow, managing the skin's redness helps to produce more accurate results. As a result, we propose ways to eliminate, decrease, or control those parameters to increase the accuracy of heart rate estimations.

Acknowledgment

I would like to express my sincere gratitude to my supervisor Professor Anthony Whitehead for his constant support, guidance, and encouragement. I truly appreciate his vast knowledge and expertise.

Second, I would like to thank the financial support of the government of Saudi Arabia in the form of a generous academic scholarship.

From my heart I thank my dear wife Azizah. She continually supports me and has always been my source of motivation. I would also like to express my gratitude to my beloved parents Mohammed and Jummah, for without their support I would not be here. I thank my daughters Sadeem and Taleen and my son Abdullah who have been patient with me being away from them during my studies. Finally, thanks to "GOD" the almighty for giving me the help I needed to reach this goal.

Table of Contents

List of Figures

List of Tables

List of Appendices

Chapter 1: Introduction

Contactless heart rate measuring methods are important and in demand. There are many motivations for obtaining vital life signs without touching the subjects. For instance, such measuring methods are necessary for health issues related to quarantine, sports activity monitoring, and other applications where touching the subject is not possible. Methods of measuring life signs such as heart rate, respiratory rate, and other information remotely were introduced as early as 1960. First, volume changes of the heart were captured by microwaves, followed by detection of the heart rate and respiration rate [1]. Doppler radar was used to measure heart rate within a 30 cm distance from the subject [2]. Detecting mechanical vibrations was the key to obtaining the heart rate. Doppler radar was also used by Matthews and Gregory to detect the chest motion that was caused by breathing and heart beat [3]. Moreover, thermal and optical based systems were used to obtain the same results. Systems equipped with laser Doppler vibrometer captured neck movement to measure heart rate [4, 5]. In addition, sensors were installed in a bed to measure subjects' life signs through motion [6]. Overall, the majority of contact free methods for obtaining a subject's life signs have depended on certain hardware capabilities such as specialized sensors.

Alternatively, video analysis has been used in a contactless measuring system to read vital sign information that is unreadable from the original unprocessed frames. Video analysis has been used in two main types of systems. The first type captures certain frames using specific equipment (complex hardware) followed by analysis of the

videos to obtain life sign readings. For instance, a thermal camera captures thermal images of the subject's face or neck which are then analyzed based on thermal changes on the face or the neck to obtain the respiratory rate or the heart rate [4]. Another example of complex hardware was introduced by Kenneth and Tomas in capturing pulse oximetry without touching the subject by using a camera capable of simultaneously capturing two photoplethysmographic signals at two different wavelengths [7]. These two systems used sophisticated hardware to produce videos that were analyzed for determination of the vital signs. However, the cost and complexity of such hardware could potentially reduce the mobility and accessibility of the system. This type of system is suitable for use in research labs or highly equipped hospitals or organizations. The second type of system uses regular cameras such as cell phone cameras, tablet cameras, or webcams to produce normal every day videos. Video magnification, a form of video analysis, has been used to investigate videos for small changes in motion and colour [8-10]. However, the experiment setup for different studies has shown some constraints that could affect the application of such systems. Figure 1.1 shows usage of a web camera and an iPad built-in camera to capture subjects' faces with limited motion [11-13]. The cameras were stationed in front of the subjects to prevent unwanted motion in the captured videos.

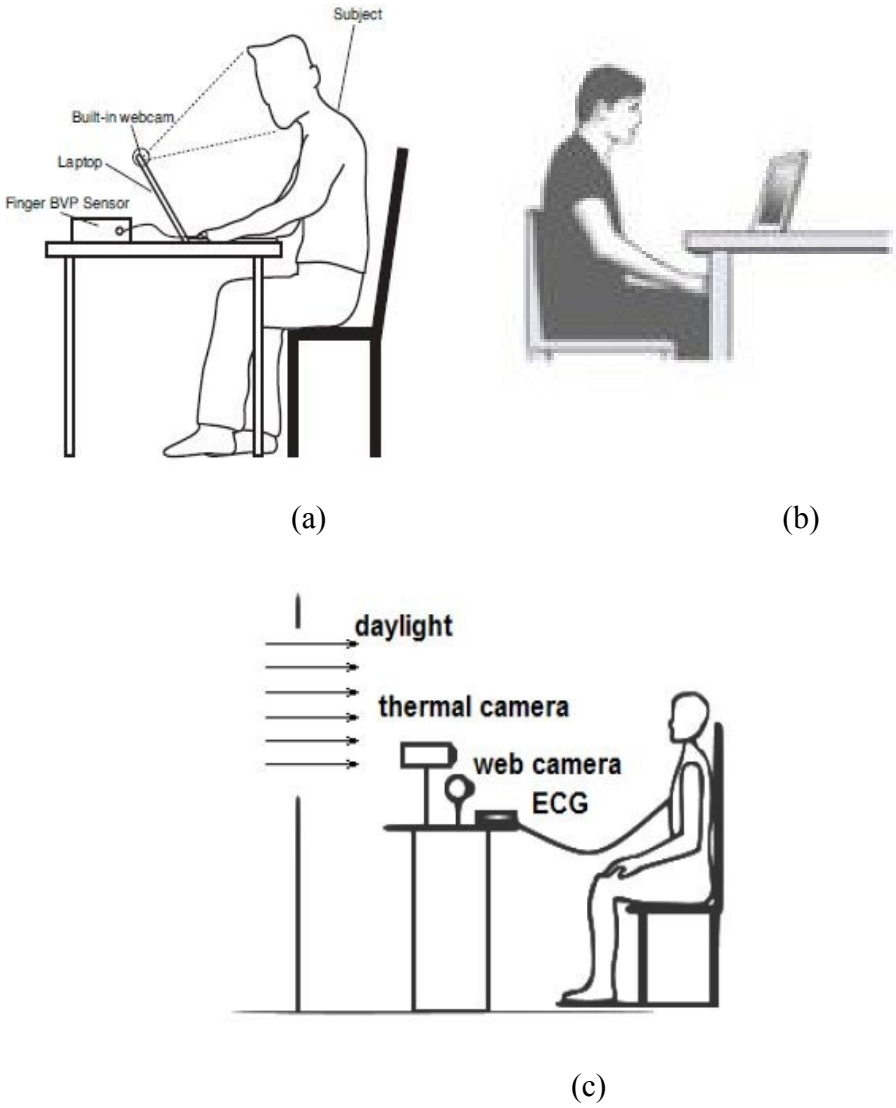


Figure 1.1. (a) Experiment setup where the camera is a built in webcam; (b) iPad camera is used in the experiment setup while the subject is sitting; and (c) Experiment setup for heart rate measuring using web camera that is installed in front of the subject. © [2007] IEEE. Copied from [12, 11, 4].

In summary, the majority of contactless methods to measure life signs are dependent on the following:

1. Changes in skin colour and brightness;
2. Thermal changes in neck skin and nose areas;
3. Movements of head or neck skin; and

4. Volume changes of the chest.

This thesis focuses on improving methods that fall into category 1: Changes in skin colour and brightness.

1.1. Motivation

It is difficult to obtain critical information about users' life signs while they are in emergency situations or participating in sport activities where connecting regular ECG electrodes is inconvenient or not possible. Resolving this issue is the main motivator for our research. Contact free applications which measure vital signs in sports help to monitor body performance and warn athletes when a certain level has been reached. Moreover, in cases such as arrhythmias where life is threatened, contact free measurement is ideal [14]. Furthermore, contact free measurement of life signs using video cameras is a promising application for ambient assisted living [15]. As well, a contact free system that interacts with computers using the users' life signs (HCI) may improve many applications that are related to users' experiences and health [16]. Users are more likely to accept having their heart rate monitored through regular cameras compared to sitting in a computer user chair that is equipped with a heart rate monitoring device [17].

The main goal of this research is to increase the ability of reading information or applying accurate measurements for subtle changes in videos captured under realistic everyday conditions. These subtle changes cannot be seen by the naked eye, nor can they be seen by increasing the video resolution. In response, some research in this area over

the last few years has developed different algorithms and constraints [8, 12] that make seeing subtle changes possible and determining the heart rate a viable task.

Eulerian video magnification has been shown to magnify motion and colour information so that subsequent analyses can reveal biomedical information about a human being imaged with a typical camera [8]. Heart rate, for example, is one key element that can be detected. However, for heart rate detection to be successful, it is expected that the subject remains very still during the video capture process. This requirement makes the system fragile and difficult to use in real life scenarios.

The primary focus of this research has been to enhance the Eulerian method [8] to provide better results in videos that have motion, such as video sequences of a person running on a treadmill. Stabilizing the video as a preprocessing technique before applying the magnification method has improved the results and we show that the Eulerian magnification is more accurate when we stabilize the input video. In addition, we compared the output results on different stabilization techniques applied in the video. The results of this effort provided an accurate idea of the challenges that the proposed methods will need to improve upon. The stabilization layer that was used is based either on face tracking or optical flow using feature detection. The resulting sequence of frames after stabilization using face tracking will only keep the face, which will provide more stability in the Eulerian magnification stage.

However, stabilizing using feature detection in a scenario of a running person on a treadmill where the camera is fixed results in a moving face with a still background does not work as required for this research, because the goal of this research is to stabilize the human in the video, rather than the background. Since feature-based stabilizing will work

with all features in the image, the resulting output will be a state of overall stabilization based on the background features and not the human being. Face and skin extraction methods were also used which focused on the face to prepare the sequence of images to be stabilized.

The other unexpected yet challenging aspect this research uncovered is the effect of skin colour and illumination on the accuracy of Eulerian Magnification techniques. This research demonstrates a relationship between skin tone and the accuracy of the result provided by Eulerian Magnification. By using subjects with different levels of skin colour, we demonstrate the impact on the results. Treating the skin colour and illumination effects in the video is a critical step toward enhancing the overall accuracy.

1.2. Problem Definition and Challenges

Using Eulerian magnification on unstable objects in videos leads to poor results of estimating the heart rate [8]. Requiring very stable objects in video frames limits the applications of contactless heart rate monitoring and makes it difficult for contactless heart-rate monitoring to be used in real-life situations. Moreover, monitoring the heart rate of people in an idle state while sitting or sleeping is less important compared to monitoring the heart rate of people who are engaged in physical activity, or during critical times such as waiting for medical attention in hospital emergency rooms. In these types of situations, Eulerian magnification does not work; therefore, the video needs to be preprocessed by stabilizing the targeted object. The experiments have been run over subjects with a variety of skin colours, which introduced a new challenge, as the heart rate measurements of lighter skin tones were more accurate. The accuracy of results was lower for darker skin

tones. Thus, a method for handling this factor is required for a more general solution. Figure 1.2 illustrates the overall proposed system complete with positioning the pre-processing step after capturing the video. Figure 1.3 shows the methods used to pre-process the video in our research.

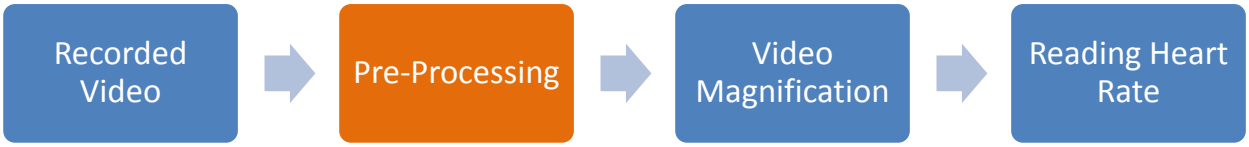


Figure 1.2. The pre-processing step, which is the core of this thesis, with respect to the overall proposed sytem

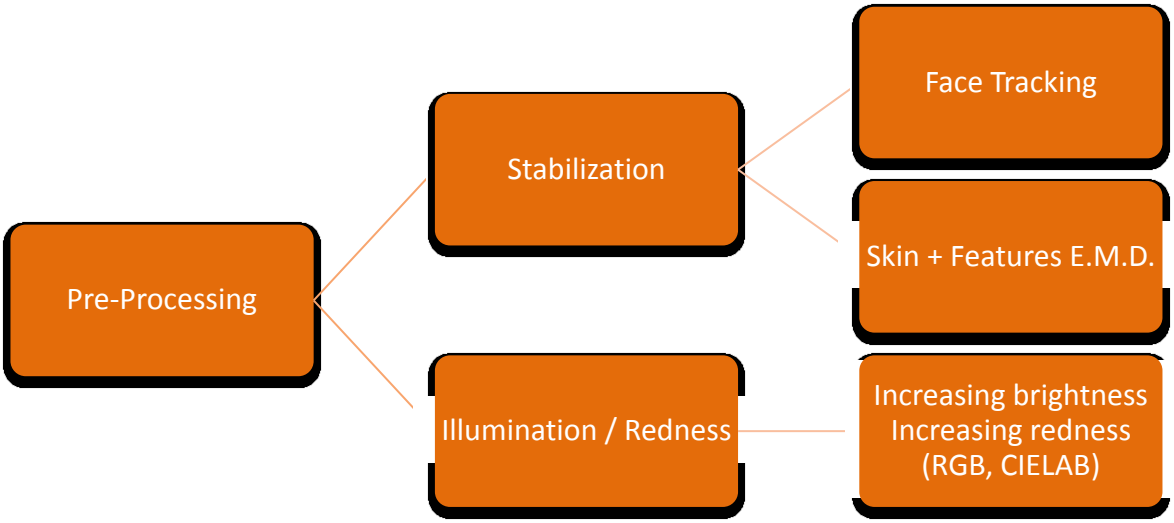


Figure 1.3. Proposed methods to enhance video magnification accuracy

1.3. Contributions

The primary contribution of this thesis is the development of a multi-stage preprocessing system to precondition the input video so that Eulerian magnification has a greater opportunity of being more accurate.

The specific research contributions of this thesis are as follows:

- Enhancing the contactless heart rate monitoring by using Eulerian video magnification and adding layers of preprocessing. The preprocessing system has two major subcomponents which are stabilization and controlling skin redness and brightness. This thesis objectively demonstrates the problem and proposes a solution.
- Demonstrating the negative impact of skin colour and exposure to light in the Eulerian method and presenting a skin and light exposure reduction preprocessing system to improve the overall results compared to the original method.

Table 1.1 presents Eulerian Video Magnification capabilities and the T letter indicate the research enhancements this thesis makes.

Table 1.1. Method (left) and method capabilities (top).

	Skin colour awareness	Computation Complexity	Low level of Movement (------)	High level of Movement (running)	Handling of illumination effect	Simple hardware
Eulerian Video Magnification	T	✓	T	T	T	✓

1.4. Thesis Outline

Chapter 2 presents a literature review of the field of video magnification and video stabilizing. Preprocessing for illumination and skin detection as well as noise reduction are also discussed in this chapter. Chapters 3 and 4 address the two main phases

of this research. Chapter 3 investigates the problem of using the existing Eulerian method to measure heart rate while subjects are in motion and presents the proposed solution. Chapter 4 discusses the effect of skin tone on the Eulerian video magnification method and presents the proposed solution. Chapters 3 and 4 contain the experiment details and discuss and analyze the results. Finally, Chapter 5 offers concluding remarks and presents the important areas, problems, and challenges for future work.

Chapter 2: Related Work

This chapter presents background information for the thesis and establishes the required knowledge related to the main research contributions. We first highlight the stages of Eulerian video magnification and then review other contactless heart rate monitoring systems such as face tracking and thermal imaging analysis. We then highlight the basic information about the pre-processing phases and the methods used to precondition input data. Face tracking, video stabilization using optical flow or feature detection, and skin colour enhancement were used in this research to prepare the video sequence for processing. Finally, we introduce the effect of illumination as a critical step in image analysis. The last section in this chapter reviews some existing methods of dealing with illumination effect.

2.1 Contactless Heart Rate Monitoring

A sphygmomanometer is a popular and important tool to gauge pulse rate. However, researchers have recently been focusing their attention on the ability to read life signs remotely without touching the subjects or patients. Contactless heart rate monitoring can be done using a special camera and light source [12], such as through thermal imaging or by using normal cameras with digital image processing layers.

The following section presents Eulerian video magnification, a magnification method that can read more information than the naked eye alone.

2.1.1 Eulerian Video Magnification

Eulerian Video Magnification is a method that allows us to see small colour changes in videos which are hidden from the naked eye. This process depends on the core assumption of optical flow which uses a brightness constancy assumption as a base for the algorithm [8]. The framework that was used in the Eulerian method is presented in two main stages: spatial decomposing and temporal filtering. The video sequence was decomposed into different spatial frequency bands. A temporal filter is applied to all bands. Subsequently, the filtered band for the targeted frequency range is amplified by a predetermined factor α. Finally the amplified downsampled bands are added to the original signal and then the pyramid is collapsed to create the magnified signal. Frequencies within the range of 0.4 – 4 Hz are used to target the beating heart rate of 24 to 240 beats every 60 seconds. The amplifying factor used in colour magnification is between 0 and 100.

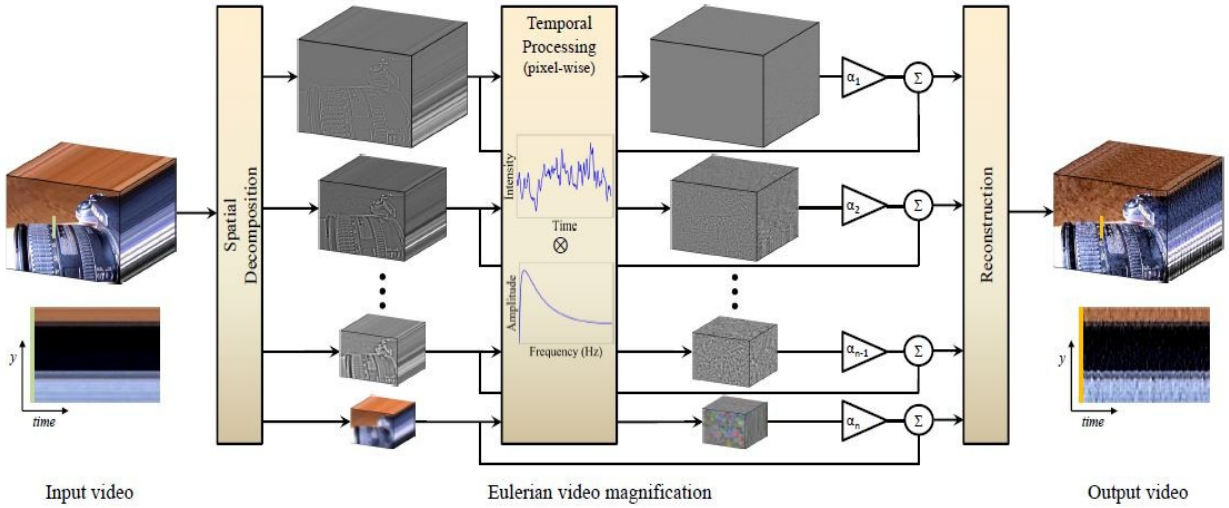


Figure 2.1. Framework of Eulerian Video Magnification: a. Spatial decomposing; b. Temporal filtering; and c. Reconstructing the output video. Reprinted with permission [8] © [2012].

This proposed method works accurately with very low motion in the image. Since our goal is to use this method with image sequences that have more motion, and because of the brightness constancy assumption which was first introduced by Lucas and Kanade [18] which evaluates the movement vector of pixels over the entire image, this method is insufficient for spotting the colour changes necessary to compute heart rate. The changes in pixel values in this scenario are caused by the movement of the subject and the illumination effect on the subject's surface. This high frequency change in colour is not the targeted change we desire to compute heart rate.

The Eulerian method has a set of constraints to effectively compute accurate results. Since the targeted change by magnification is colour, the subjects on the original video need to be still. Moreover, the method has been applied over two faces with different skin colour. Figure 2.2 shows the lighter skin of face1 and the darker skin of face2. However, the effect of skin colour on the accuracy of pulse readings has not been addressed.

(a) (b)

Figure 2.2. (a) face1 with lighter skin colour; and (b) face2 with darker skin colour. Reprinted with permission [8] © [2012].

2.1.2 Philips Vital Signs Camera

Philips has developed an app which appears to use Eulerian Video Magnification to measure heart rate and breathing rate by using cameras in iPad or iPhone [11]. The software recognizes the small changes of face colour that are caused by heartbeat, and subsequently records these changes as heart rate. Certain conditions are necessary to ensure good measurements. First, the device must be placed or held to make sure that a steady video is captured. The subject should place the face and the chest in the area that is required by the app and remain steady for the software to be able to accurately measure the heart rate. Results will be inaccurate in cases of unbalanced illumination such as when there is a source of light behind the subject. The background should be clear of moving objects or other people. In the case of placing the iPad on a table, the person should not touch the device or the table because the software is sensitive to any kind of motion source. The app requires that the whole face is placed in the marked area to ensure that enough skin areas from the face of the person are captured. All of these constraints lead us to believe that this app uses Eulerian Video Magnification as its method, yet the issue of motion is not directly addressed and they leave it to the user to make up for the deficiencies of the system. Figure 2.3 shows the proper positioning of a subject with respect to the device and the light in the room.

Figure 2.3. The top image shows how the subject should position himself in front of the camera without touching the iPad or the table. The bottom image shows the location of the light source with respect to the subject. Copied from [11].

2.1.3 Video Imaging for Targeted Subject

This section presents two methods that use video images of the targeted subjects to obtain the heart rate. Both methods are explored deeply from two sides. First, examination of the experiment setup shows whether or not the subjects were in motion. Second, we study the core of the method that obtained the heart rate readings.

2.1.3.1 Video Imaging and Blind Source Separation

Ming-Zher proposed a contactless method to measure cardiac pulse rate using a simple webcam to record the video [12]. The subject's face was tracked on the recorded video by Open Computer Vision library which detects the face location using a box with x- and y-coordinates. The area of interest is the centre 60% of the width and full height of the box. In the case of not detecting a face in a single frame, the algorithm uses information from the previous location to minimize the effects of this type of error on the overall performance. Then, the area of interest is analyzed according to the following steps. First, the captured images are separated into the three RGB channels. Then, the three measurement points (red, green, and blue) are found after spatially averaging to all pixels in each

channel from the area of interest. Finally, the raw traces are formed and normalized, as per the equation below (Equation 2.1):

$$x_i'(t) = \frac{\mathcal{X}_i(\mathrm{t}) - \mu_i}{\sigma_i} \tag{2.1}$$

Where μ_i, σ_i are the mean and standard deviation of $\mathcal{X}_i(\mathrm{t})$ for each i where i = 1,2,3.

Second, Independent Component Analysis (ICA) was used to decompose the normalized raw traces into three independent source signals. After that, the second component was selected and transformed to frequency domain using FFT. Selecting the signal with highest power on the spectrum gives the heart beat rate. The operational frequency is 0.74 – 4 Hz or 45 – 240 beats per minute. Historical estimations for pulse frequency were used to avoid the effect of noise on the ICA computation. This step helped to reject the artifacts by allowing change in the pulse frequency when there is enough time for the heart rate to increase (time > 1 second).

Figure 2.4 shows the experiment setup where the subject faces the laptop's built-in webcam. In this experiment the subject is allowed to move normally, as the person using the computer might move. Subjects are allowed to perform slow movements such as touching the laptop, tilting or nodding the head, and turning the head; however, quick movements are not considered. The illumination is controlled. Dark illumination (less than 61 lumins) was not addressed in this research.

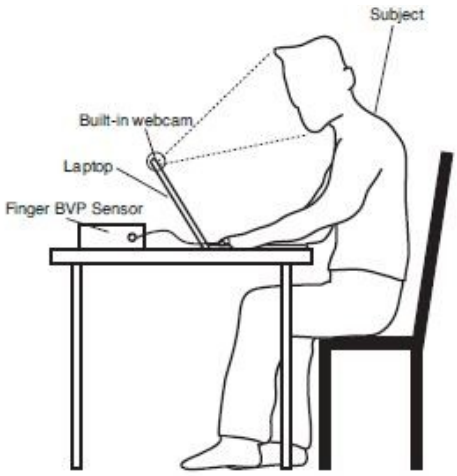


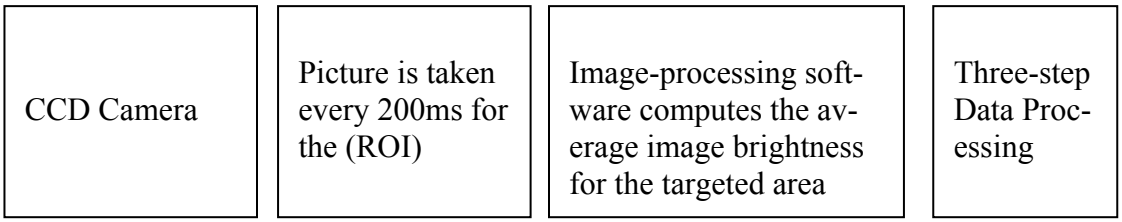
Figure 2.4. Experiment setup shows the subject's location in relation to the camera. Copied from [12].

2.1.3.2 Using Time-Lapse Imaging to Measure Heart Rate

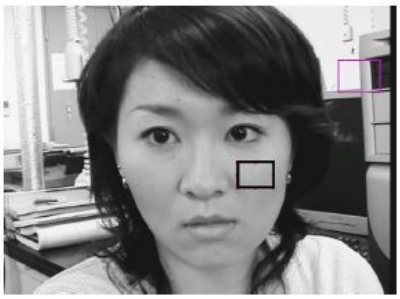
Chihiro and Yuji proposed time-lapse imaging as a contactless heart rate measurement method that is capable of simultaneously measuring respiratory and pulse rates [19]. First, this research depends on the capability of capturing the average brightness every 200ms for the subjects. Then, the results are processed mathematically to produce the frequency peaks which are determined by the heart and respiratory rates. The process uses the first-order derivative, a 2 Hz low pass filter, and Auto-Regressive spectral analysis.

The experiment was performed using a CCD camera and a typical PC for recording and analyzing the video. The subjects were located in front of the CCD camera at a distance that allowed the camera to capture an ROI area (3 cm by 4 cm) from the subjects' cheeks. Figure 2.5 (a) shows the resulting image that was captured and the black rectangle shows the region of interest. Figure 2.5 (b) presents a block-diagram that shows the time-lapse method to measure pulse rate remotely. In addition, the subject's skin colour is light and the illumination is critical to the methods success. The CCD camera used

in the experiment is able to tolerate sudden changes of illumination level within the range of 270 and 1500 lux (lumens per square meter).



(a)



(b)

Figure 2.5. (a) Block-diagram shows the time-lapsed method to measure pulse rate; and (b) Resulting image captured with the black square showing the region of interest. Copied from [19].

In conclusion, simple calculations have been used to process the collected data from the camera. Moreover, the hardware is not overly complex. On the other hand, the research does not address the effect of skin colour differences on the accuracy of measurements. However, the illumination issues are addressed by the use of a camera that is capable of handling illumination change.

2.1.4 Complex Hardware Systems

This section presents methods that use complex hardware as part of contactless heart rate measuring systems. Systems that use thermal changes in the face to capture un-

seen information and systems that use a laser to invasively measure pulse rate are presented.

2.1.4.1 Reading Vital Signs using Thermal Imaging

Garbey and Marc proposed a novel contactless method to compute heart rate [4]. This research is based on the main assumption that blood flow produces high thermal changes on the blood vessels of the face. These changes are captured using a thermal camera. The camera is connected to a typical computer, and the computer's resources are used to perform the heart rate measurements. The method is able to extract blood flow, pulse rate, and breathing rate.

In the experiment setup, persons are generally positioned in front of a computer. The user faces a Mid-Wave Infra-Red (MWIR) sensor which is the thermal camera. The sensor is sensitive to spectral range 3-5 um. Since the computer user's face is always in motion and the system needs to capture the same part of the face for a given period of time, tandem tracking is required to track the target (simultaneously tracking central facial region and a small region of interest).

The pulse is retrieved through the following six steps. The first three steps localize the selection of the pixels with respect to time. Figure 2.6 (a) shows the raw thermal images and Figure 2.6 (b) shows the tracked images. Next, the tracked frames are ready for blood vessel registration. Figure 2.6 (c) shows a white line in the subject's neck which presents the tracked blood vessel. Fourier transform analysis is then applied on the captured signals to obtain its power spectrum. By using Equation 2.2 to average all power

spectra to obtain the composite power spectrum, the heart rate frequency is found by finding the dominant frequency.

$$\bar{P} = \frac{1}{R_y} \sum_{y=0}^{R_y} P_y \tag{2.2}$$

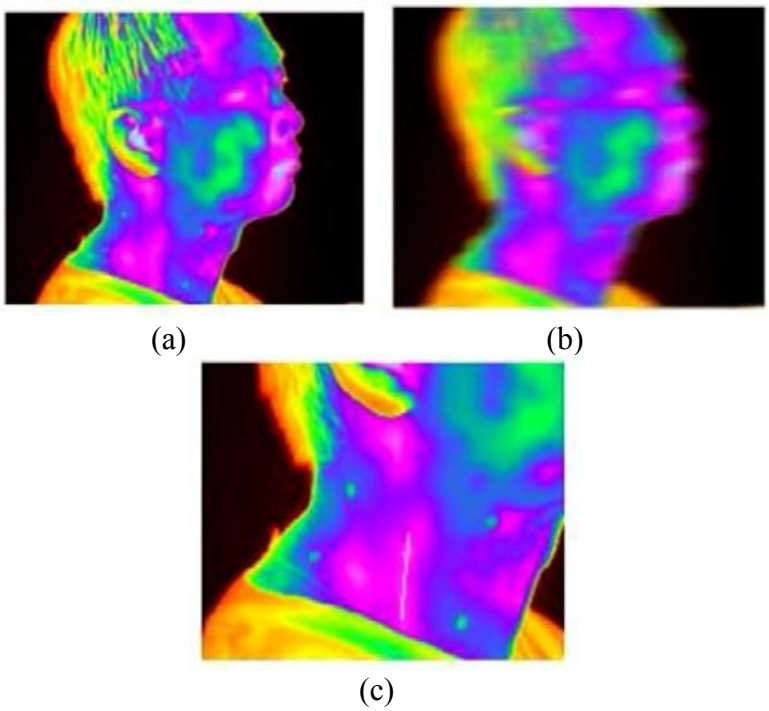


Figure 2.6. (a) Raw thermal images; (b) Tracked images; and (c) Selected blood vessel. Reprinted with permission from [4] © [2007] IEEE.

In terms of measuring a consistent pulse with respect to time, this method is limited by tracking the area of interest when the subject is sitting and the head is frequently moving. Moreover, change in temperature is weak in the time domain, and this difficulty has been overcome by using the frequency domain to enhance the signal instead of weakening it. On the other hand, the use of personal computers to read the user's vital signs is an important advantage of this method, since this type of resource is widely available.

The researcher used the same technique to measure breathing. The capturing of this targeted thermal signal in this case is from the nose and was simpler and faster in the

case of subject movement. Recorded video using a basic camera does not have the required information to obtain breathing rate, because the air is not visible. However, tandem tracking tracks the central facial region and the small region of interest at the same time is possible. This pre-processing step is investigated in Section 2.3.

2.1.4.2 Reading Vital Signs using Direct Beam Laser Doppler Vibrometer

The laser Doppler vibrometer was used by Lorenzo to measure subjects' heart rate remotely [5]. The pulse was obtained by measuring the vibration of the main vessels on a subject's neck using an optical laser. Figure 2.7 depicts the setup of the Lorenzo experiment where the laser is directed at the subject's neck from a distance of 1.5 meters.

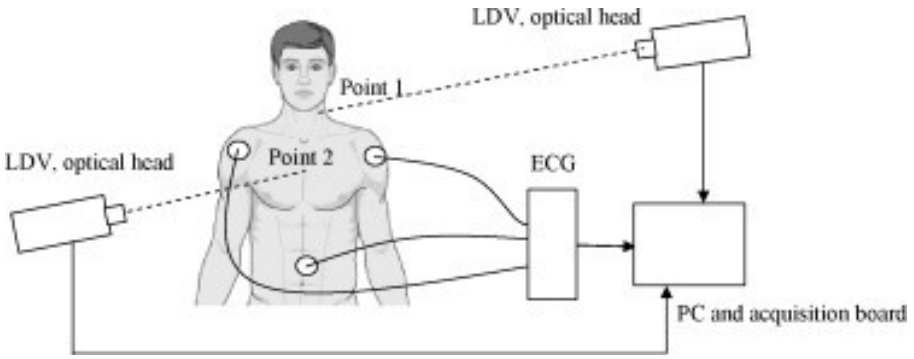


Figure 2.7. Experiment setup of laser Doppler vibrometer. Copied from [5].

The proposed method is a contactless heart rate measurement system; however, the factors relating to the subject's positioning and the equipment necessary are very difficult to achieve for an average user.

2.1.4.3 Using Simultaneous Dual Wavelength Photoplethysmography (PPG) to Measure Pulse Oximetry Remotely

Hu and Wa addressed the capturing of pulse oximetry (a process to measure blood oxygen level) without contacting the subject by using a camera that is capable of simultaneously capturing two photoplethysmography (simple optical method to measure blood volume change in the microvascular bed of tissue) signals at two different wavelengths [7]. The research presented a contactless method by using sophisticated equipment, and the information was acquired with high accuracy.

The imaging system consists of a camera equipped with c-mount zoom lens and a 36 LED light source. The camera is zoomed manually to the targeted area and captures a 20 second video. The camera is 30 cm from the subject's skin. In addition, the light source is located exactly beside the camera to illuminate the rest of the surface, as shown in Figure 2.8. The other parts of this system, including the computer connection and zoom lens, are shown in this same diagram.

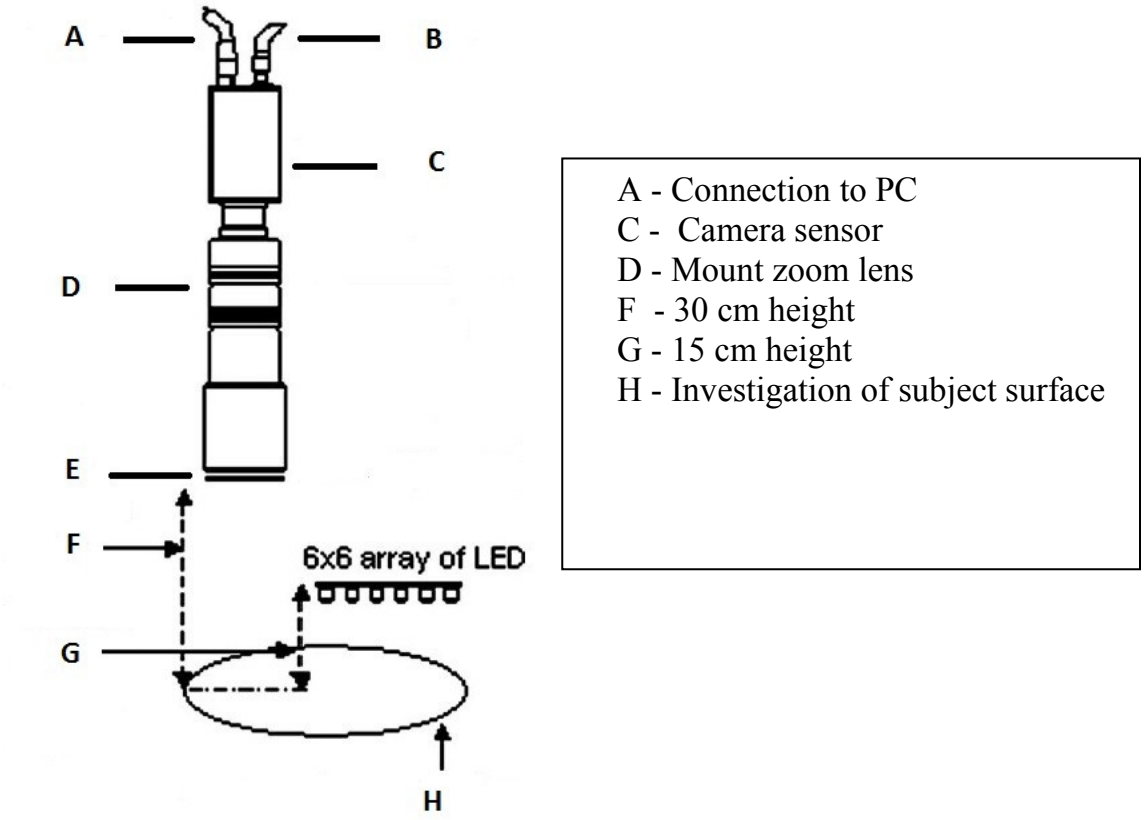


Figure 2.8. Pulse oximetry system parts and setup. Copied from [7].

The resulting videos are processed to locate the presence of the reflected light source in the skin. The research showed that the resulting measurements are exactly equal to the conventional method in terms of measuring pulse oximetry using a patient's finger-tip. However, the setup is difficult and costly. The subject should be in a still state where movement is very limited. These constraints do not make this method optimal good choice for contactless measurements. The need for a special camera and a source of light to emit the dual wavelengths makes the system costly in comparison to other systems.

Contactless heart rate measuring techniques are summarized in the following table which shows each method's capability under certain conditions. The T in the first row of

the table presents the limitations our proposed system is intended to overcome in the Eulerian Video Magnification method.

Table 2.1. Comparison of methods (left) and method capabilities (top).

	Skin colour awareness	Computation complexity	Movement Low level (------)	Movement High level (running)	Handling of illumination effect	Simple hardware
Eulerian Video Magnification	T	✓	T	T	T	✓
Philip Vital Signs Camera		✓	✓			✓
Video Imaging & Blind Src Separation			✓			✓
Time-lapse Imaging		✓		✓		
Thermal Images Analysis	✓		✓		✓	

2.2 Preprocessing Video

Preprocessing the video for Eulerian Video Magnification can be categorized by two major parts: stabilization and colour compensation. Stabilizing deals with the subject's position and spatial location in the scene with respect to time. On the other hand, colour effects are caused by the subjects' skin colour or illumination around the subject.

Eliminating or minimizing the colour effect plays a major role in increasing the accuracy of the proposed system.

2.2.1 Stabilizing

Video stabilization is used in the field of image processing in many applications such as removing unwanted motion caused by hands shaking while recording regular home videos [20]. The resulting stabilized videos using look more like they were captured using a tripod. Furthermore, photographers often face the problem of shaky or blurry images. Image stabilization or removal of blurriness can be accomplished by using a single frame or a pair of frames [21]. In addition, hardware like CMOS sensors or optically stabilized lenses can be used to stabilize images [22].

Disadvantages and advantages of video stabilization methods relate to the application in which they have been used. Using video stabilizing, which is dependent on hardware features, saves the camera from needing a motion estimating step in its software. However, at present, using hardware to stabilize images or videos must be accompanied by using a device that has been equipped with such complex hardware. In the future, it is likely that such advanced hardware will be embedded in normal cameras [22]. All of the stabilization methods that have been described for images or videos are for cases where the movements are quite small. Using these methods to remove a larger motion, such as running, is more complicated.

The process of tracking an object and then recreating the video based on the tracking results produces a video where the object is motionless. There are two main types of object tracking: recognition based and motion based [23]. Coupled object detection and

tracking were used to enhance the regular objects' tracking outputs in [24]. In addition, tandem tracking, which simultaneously tracks the central facial region and the small region of interest, has been proposed by [16] that minimizes false tracking. Optical flow methods are motion based method of face tracking.

This section highlights the existing methods to stabilize videos with subjects in motion, such as when subjects are running. The following methods are presented: stabilizing using face tracking; features detection, extraction, and matching; and optical flow.

2.2.1.1 Face Tracking

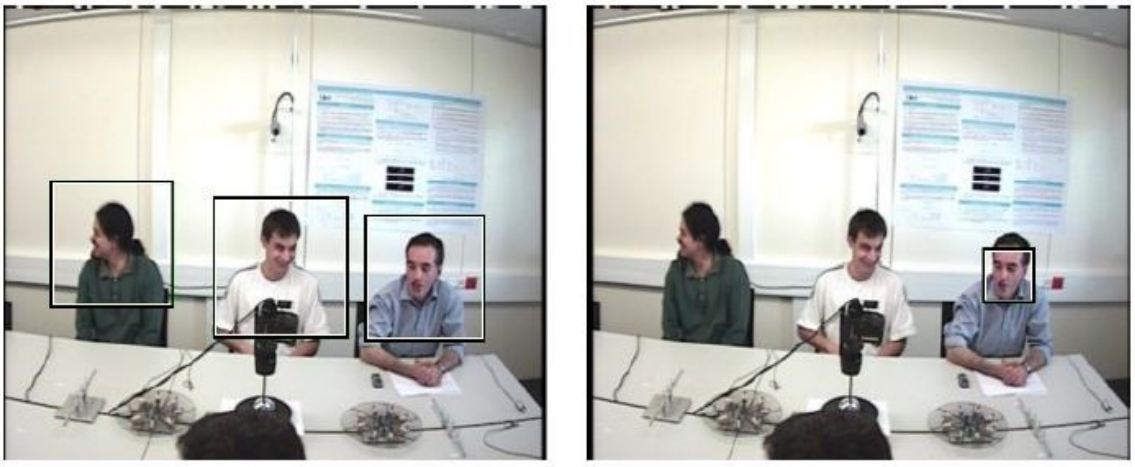
This section presents face detection and tracking as an important method of preparing the raw video sequence. Using video stabilizing by feature detection results in poorly stabilizing the targeted subjects, as the system will try to stabilize using the non-moving features. This scenario gives face tracking and detection an advantage over other stabilizing methods, as face detection helps to select the targeted area of skin within the face more robustly. Many approaches detect faces in images, including feature invariant, knowledge-based, and template matching approaches [25].

2.2.1.1.1 Face Tracking using Viola-Jones Algorithm

The Matlab function vision Cascade Object Detector System which uses the Viola-Jones algorithm was used to detect objects in each frame [26]. By default, the detector is configured to detect faces. The detector achieves the results through integral images, Adaboost algorithm, and cascading [27].

The Viola-Jones algorithm can be summarized by the following stages [28]. First, integral images are used to quickly and easily find the features in the images. Second, AdaBoost machine learning mechanism is used to detect faces through simple and efficient classifier. Third, the cascade algorithm is used to reduce the computational complexity and quickly give better results in detecting objects. In addition, when the object is a face, the AdaBoost step helps to produce a proper classification. The proposed algorithm is fast and accurate in detect objects.

To detect the face even if the face front side is not presented clearly in the frame, Kruppa and Hannes proposed the local context approach which is able to detect faces with different positioning without adding new data to the training set [29]. Images in Figure 2.9 show this scenario. The first image (a) shows the results of the local context approach and the second image is the result without using this approach. The location of the face in the local context of the subject's upper body part that has the face is determined by (W/4, H/10) with dimension of width equal to half of W and dimension of height equal to H/2. Figure 2.10 shows the dimension and the location of the local context of the face by the large square H by W.



(a) (b)

Figure 2.9. (a) Shows the result of the local context approach, and (b) Shows the result without using the proposed approach. Reprinted with permission from [29].

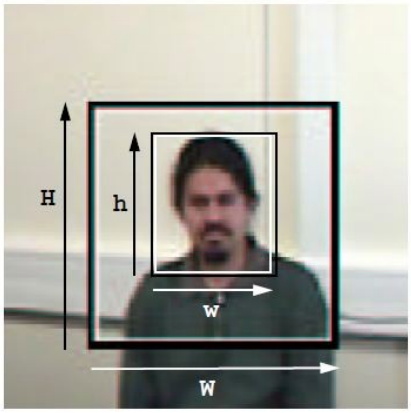


Figure 2.10. Relation between the local context of the face and the face location on that context. Reprinted with permission from [29].

2.2.1.1.2 Tandem Tracking

Tandem tracking, which simultaneously tracks the central facial region and the small region of interest, was proposed by [16]. The goal of this method is to minimize the false tracking and thereby increase the tracking quality. The main idea is that two regions are tracked. The first region is selected as template for tracking and is typically a region with high contrast that allows for easy tracking. The second region, is the region of

evaluation that is typically of low contrast, or more difficult to track. The two regions are spatially separated, and this separation does not change over time. By tracking the easy to track first region, we are able to determine the location of our second region upon which we perform the evaluations.

2.2.1.2 Stabilizing using Feature Detection, Extraction, and Matching

Video stabilizing can be achieved through many different frameworks, but they mainly use the same three layer process of motion estimation, motion smoothing, and image composition [30]. This section presents four ways to achieve video stabilizing. The first framework finds features using Harris Corner detection as a main step. Next, point feature matching, or geometric transformations, is presented. After that, the probabilistic estimation framework is presented to characterize the motion between frames. Finally, a compensation method is described that takes care of moving objects by stabilizing through the use of feature trajectories.

2.2.1.2.1 Video Stabilizing using Point Feature Matching

LabeebMo, Noori, and Mustfa presented a stabilizing technique for video frames using point feature matching [31]. The algorithm depends on choosing matching points between two frames and determining the affine transformation between them. Harris Corner detection was used to detect the corners around the important features in the image [32]. The method can be summarized by the following steps:

1. Read the consecutive frames from the original video.

2. Find matching points around the salient features between the consecutive frames by detecting detection the corners.

3. Each selected feature point in the first frame is matched to a selected feature in the second frame. Matching scores are calculated using Sum of Squared Differences (SSD) of a small window around the corner features. Random sample and consensus (RANSAC) algorithm was used to find the strongest and most accurate correspondence [33].

4. The transformation computed above is applied to each pair of frames creating a stabilized video output.

2.2.1.2.2 Video Stabilizing using Point Feature Matching-Geometric Transformations

Matlab used FAST (Features from Accelerated Segment Test) algorithm to find corners in images to be used as features [34]. Among these features the interest point descriptors were selected using the SIFT feature descriptor [35]. Next, the matched features were used to estimate the geometric transformation between the two frames. Finally, the estimated transformation is applied to the frames to produce a new stabilized frame pair. Geometric transformations were used to produce a single frame each time from a pair of frames where the first frame is the fixed frame and the second frame is the frame with motion to be stabilized. An intensity based image registration is available in Matlab to produce a stabilized image from each pair of images [26]. The algorithm uses a metric to compare the moving frame after transformation with the fixed frame to optimize the stabilization. The optimizer stops the iterative process of the algorithm when the metric is minimized and the transformation computed is determined to be the optimal transformation. However, this method is easily affected by background details and face location.

Since the existing algorithm is not able to target one area apart from others, unwanted parts of the image must be removed manually. Figure 2.11 shows the iterative view of the algorithm.

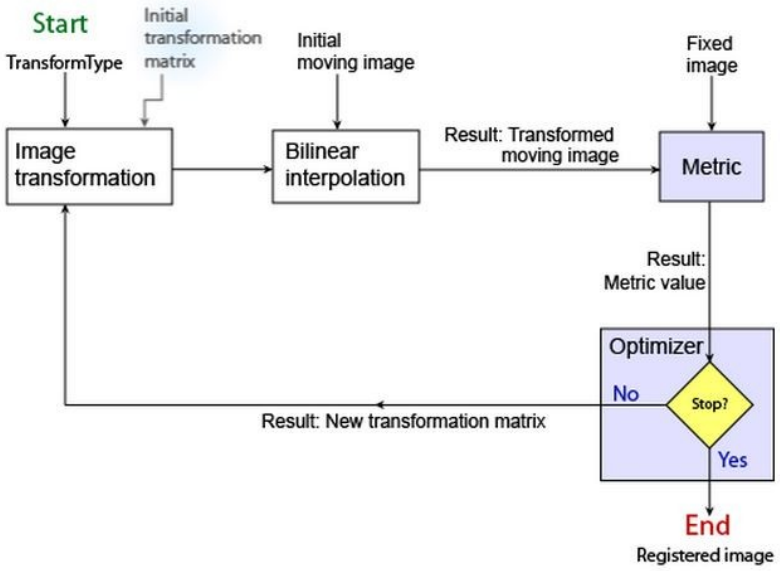


Figure 2.11. Intensity-based image registration algorithm. Copied from [26].

2.2.1.2.3 Probabilistic Video Stabilizing using Kalman Filtering and Mosaicking

Litvin and Andrey present a video stabilization framework using probabilistic estimation [36]. The block diagram in Figure 2.12 shows two main steps for the probabilistic stabilization algorithm. The first step estimates the pair-wise transformation, the intentional motion parameters, and frame warping. The second step uses mosaicking to reconstruct the undefined regions. This step is performed by transformation estimation and warping of distant frames.

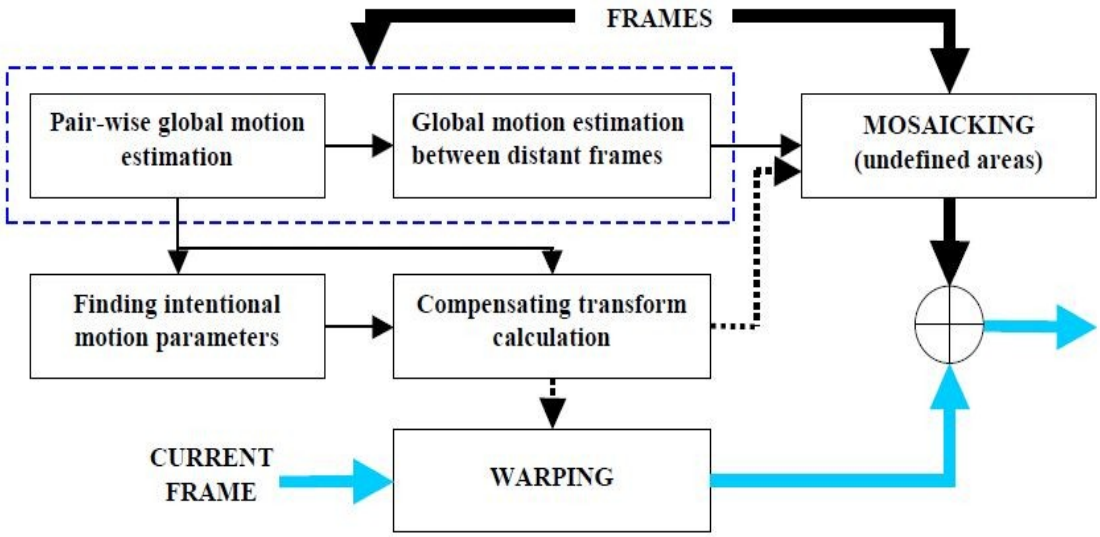


Figure 2.12. Probabilistic video stabilizing flow of frames. Copied from (34).

The algorithm extends further than just shifting or rotating the images. In addition, the algorithm is capable of estimating the intentional motion (intended camera motion) which will lead to removing unwanted motion (jitter and hand shaking). Instead of using frame cropping and magnification to recover the undefined region in the registered images, the algorithm uses mosaicking to recover these areas using neighboring frames.

2.2.1.2.4 Video stabilization using Robust Feature Trajectories

It is challenging for video stabilization methods to deal with moving objects in a video or in a pair of frames without focusing on the main object or the larger object that dominates the frame. Lee used feature trajectories to stabilize these types of videos [37]. Although the camera motion is not estimated, the algorithm captures the trajectories of the main features which help to create very smooth results.

In cases with more than one similar feature matched between frames, a best match was chosen by using an optimization step. The algorithm proceeded to find similar features and trajectories and false matching for feature trajectories was solved by using spa-

tial motion consistency constraint. Moreover, temporal motion similarity of trajectories was used for long trajectory sequences.

2.2.1.3 Video Stabilization using Optical Flow

Optical flow algorithms produce a stabilized video that contains subjects in motion. The resulting vectors from optical flow estimation algorithm express the relation between each pair of frames in the video. Horn and Schunck used global brightness changes to compute the flow vectors that provide the direction and magnitude of change for each pixel [38]. This information helps any algorithm to reconstruct a stabilized video. The optical flow vectors of the pixels are used to remove the unwanted movements from the scene. The process is repeated and the calculation is done for each pair of frames. Furthermore, colour optical flow makes it easy to directly deal with colour frames [39].

Lucas and Kanade used the local neighborhood with constant flow assumption to estimate pixel displacement vectors [18]. Both algorithms face problems in cases of huge displacement and illumination changes. Table 2.2 shows the advantages and disadvantages of the discussed optical flow algorithms.

Table 2.2. Advantages and disadvantages of optical flow methods

	Horn and Schunck	Lucas and Kanade
Advantages	smooth flow + global information + using more than two frames	easy and fast calculation
Disadvantages	Slow – iterative	errors on boundaries

2.3.2. Skin and Illumination

This section presents a background of areas that are related to background effects. Illumination and skin colour are factors that affect the accuracy of results. Understanding these factors and how they affect the result is important for eliminating or reducing these impacts.

2.3.2.1. Skin Detection

Kakumanu and Sokratis presented the use of colour spaces to detect skin by thresholding [40]. A set of boundaries is used to define skin from non-skin. Each component in the colour space has a range of values that determine whether the inspected area is skin or not. In the YIQ colour space, the I component helps to define yellow skin. On the other hand, in the HSV colour space the skin colour values are in the ranges of RH = [0, 50], and RS = [0.20, 0.68]. Seo and Kap-Ho used a RGB space colour to detect face region [41]. The following values were used to find human skin: 0.353<= r <= 0.465, 0.27 <= g <= 0.363 where r = R/(R+G+B) , g = G/(R+G+B).

However, Albid proved that all colour spaces can provide the same level of performance in skin detection [42]. CIE LUV colour space minimizes the effect of illumination in the images or frames.

2.3.2.2. Illumination

Illumination effects in the frames that need to be processed in video magnification produce a level of noise that affects the final results. In particular, over or under exposure creates limitations. The effect of illumination on the subject can be caused directly from a source of light, or might be caused by reflected light. This uncertainty makes the pre-

diction and detection of illumination issues a complex step. Illumination detection has been achieved previously by using a chromogenic camera that captures two normal images through a coloured filter. The method depends on finding the shadows' edges by capturing the change in the scene [43]. Chihiro and Yuji proposed a system which controlled the illumination [19]. The sudden change of illumination level was tolerated using the auto iris function of the CCD camera within the range of 270 and 1500 lux (lumens per square meter). Furthermore, equalization of illumination was used to enhance images by applying homomorphic filtering [44]. The researchers concluded that this filter will work with RGB colour model as well as with other colour models to increase the illumination in the darker areas.

Chapter 3: Enhancing Video Magnification using video stabilization

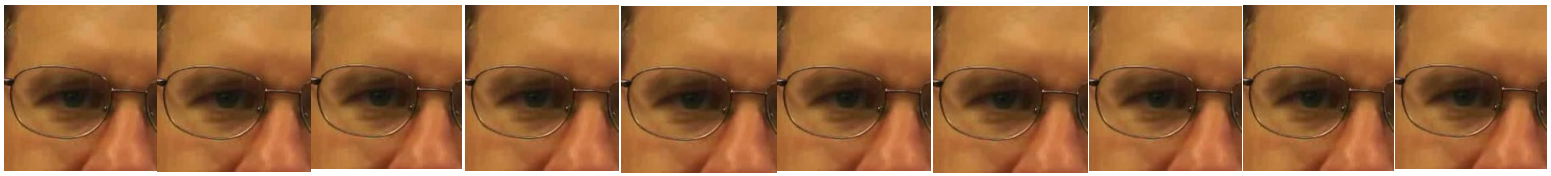
3.1. Introduction

The video is required to be pre-processed using stabilization methods to eliminate or minimize the motion of the subject as much as possible. By comparing the output results that come from using the Eulerian algorithm alone with different stabilization techniques applied on the video, we can evaluate the effectiveness of the stabilization layer. The stabilization layer in this evaluation will use either face tracking or optical flow using feature detection, extraction and matching. The resulting sequence of frames after stabilization using face tracking will only have the face which will give more stability.

The proposed scenario in which the Eulerian method will be evaluated has important setup conditions. The camera was stationary with respect to the subjects, and the background in this scenario was stable (i.e. non-moving). The camera will capture the subject's face by targeting only the upper body of the person. Figure 3.1 A is showing a stable subject's location over 10 frames which is unrealistic in everyday action while Figure 3.1 B shows a more realistic situation. Figure 3.2 is plotting the face location changes from Figure 3.1. The x-axis is the frame number and the y-axis is the displacement of the selected area in cm. The plot is showing subject's face location change in period of 0.033 second. The blue line is presenting the unrealistic motion-less face with stable position that was presented in Figure 3.1 A. The red line is showing the subject face

location change because of walking or running that was presented in Figure 3.1 B. The displacement was in the range of 2-10 pixels.

A



B

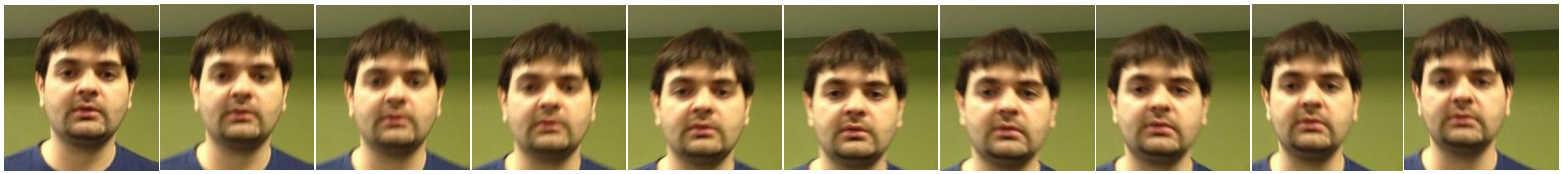


Figure 3.1. Location of background and subject's face with respect to the camera are presented on 0.33 second of frames. A images present the 10 frames which are motion-less face used by Wu [8]. B images present the first 10 frames of Subject3 (Mustfa) while he is running on the treadmill.

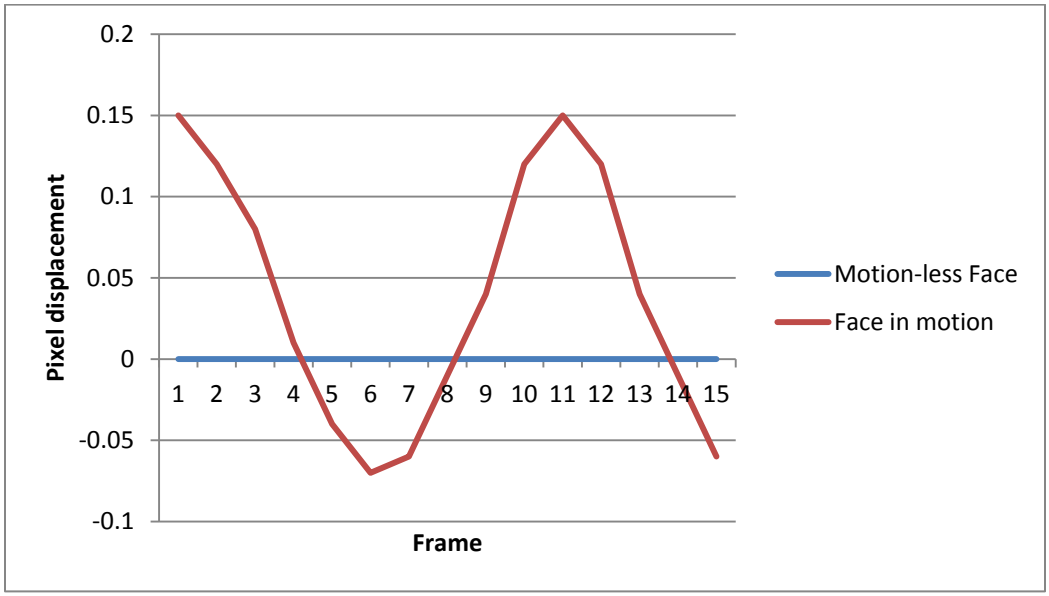


Figure 3.2. 15 frames of a subject's face location change. The blue line is presenting the unrealistic motionless face used in Eulerian research [8]. The red line is showing the subject face location change because of walking or running.

The next section will describe the methodology used to test the proposed system to enhance Eulerian Video Magnification on different scenarios that have motion. The pre-processing layer will use two methods of stabilization which are video stabilizing using feature detection, extraction and matching, and stabilizing using face tracking. Section 3.3 will present the results and the discussion of the proposed method. The conclusion and the future work will be presented in section 3.4.

3.2. Method Overview

The Eulerian methods capability is clear on videos with limited motion affect [8]. Other contactless methods to measure heart rate suffer the same problem of subject movement in video. This movement must be removed to accurately obtain the heart rate measurements from skin color changes [11, 19]. Our research experiment will demonstrate the capability of the proposed system to improve Eulerian video magnification.

The camera used to capture the subject's upper body while on the treadmill is an iPhone5 cell phone camera. The subjects were asked to wear a heart rate monitor device with two sensors on it to establish ground truth. We acknowledge that the error range on the heart rate monitor is unknown, but our goal would be to be at least as accurate as the commercial heart rate monitor system. Detailed setup of the experiment is presented in section 3.3. There was a pilot of the experiment to insure everything would work as expected. The experiment data will be recorded in two parts. The first part is the video capture which was recorded by the iPhone5 device. The second part is the heart rate readings which were recorded by a Garmin 45 watch. The readings of both devices were synchronized to start the recording in the same time. The acquired information was saved in the devices and then moved to the experiment computer, a 3.4 GHz Quad Core with 12 GB RAM. The recorded videos were transferred and converted to AVI video file format. Each video was 2 minutes long and was divided into 6 parts, each one 20 seconds in length. The proposed system was implemented in Matlab.

Figure 3.3 shows the experiment setup. We fixed the camera in front of the subjects using a tripod stand. The iPhone5 camera, which is an 8 megapixel iSight camera, was used to capture the subject's face as the predominant part in the scene. The subjects

were asked to wear the heart rate monitor- chest strap directly touching the subject's skin. The contact pads in the chest strap need to get warmed up by the body heat to capture the changes accurately as the manual instructed. The heart rate monitor was connected wirelessly to the Garmin-Forerunner 410 watch that was held near the subject with short distance range to avoid miscommunication.

The subjects' were selected randomly with no prior selection for skin to be lighter or darker. The next chapter will introduce the skin tone effect in results accuracy.

1 is the recording device.
2 is the Garmin watch.
3 is showing the location of the heart rate monitor chest strap under the subject t-shirt.

Figure 3.3. Experiment setup shows location of camera that used to record subject videos while they are running over treadmill.

Figure 3.4. Chest wire wirelessly connects to the Garmin watch.

Using the Eulerian method on the raw videos without pre-processing produced a bad magnified version due to the high frequency changes in pixels color caused by displacement of the pixels, not by the blood flow in face vessels [8]. Next, the proposed system will be introduced with the pre-processing step used in to eliminate or minimize the motion. Section 3.3 will highlight this claim using the data collected through the experiment. The estimated heart rates will be compared to real heart rate reading obtained by Garmin device to evaluate the error percentage.

3.2.1 Video Stabilizing Using Features Detection, Extraction and Matching

Feature detection and matching stabilization method was used to remove unwanted movements in video [31, 36]. The main framework of the method is based on three main stages which are motion estimation, motion smoothing, and image composition. In this research the method was used to remove large scale movements in the scene

such as a person who is running. The nature of the algorithm is to detect features on the video frame where the background and other many features are not moving. Unwanted feature selection will produce unwanted results. Figure 3.5 shows the unwanted results such as scaling which is a result of set features that need to be scaled while our proposed situation does not require it. This leads us to make the subjects the direct target for stabilization.

Figure 3.5. The effect of unwanted features detection and matching – scaling.

A skin detection step was added before applying the stabilization to perform the targeting step. However, skin detection might capture the subject's hand (or other person in the background such as in gym setting) which needs to be removed in such cases. Figure 3.6 is showing the proposed system in a block diagram. The figure shows the pre-processing step that was constructed from two steps which are skin extraction and video stabilization. The raw video will be processed as the following steps:

1. Extracting the face skin of the subject by using thresholding in each frame and recreate the video to show only skin of the subject and black everywhere else.
2. Deleting other parts of the body if they exist in the frames
3. Stabilizing the face area only using feature detection and matching methods.

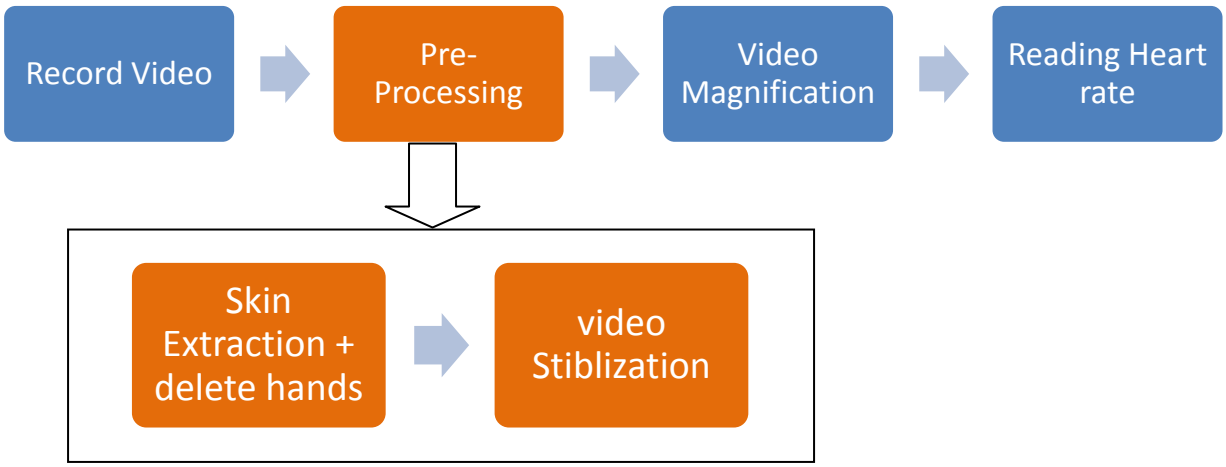


Figure 3.6. The pre-processing block consists of skin extraction and video stabilization using feature detection and matching.

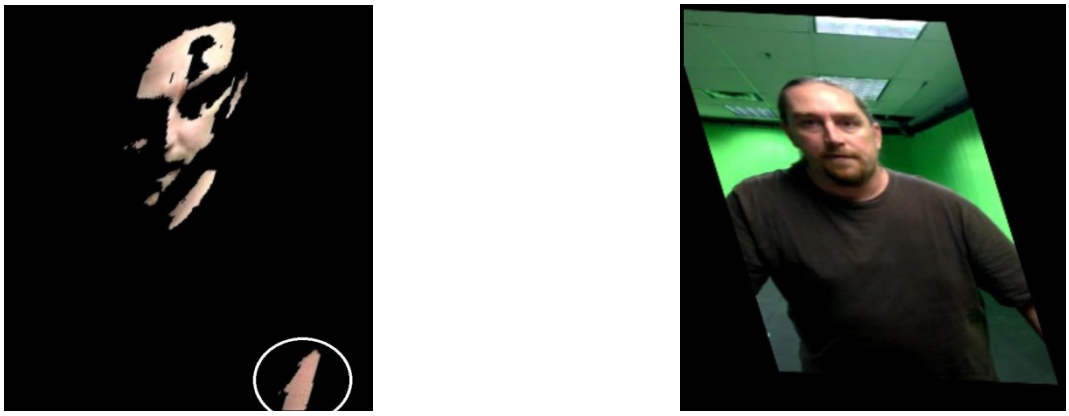
Color thresholding is one of the techniques to obtain a subjects skin. It is fast and computationally efficient. Equation 3.1 is showing the set of u values and v values that define the skin color in the LUV color system.

$$\text{Skin} = u>12 \text{ and } v>12 \tag{3.1}$$

where LUV color space in use

Figure 3.7 (a) shows the effect of other parts of the body in the scene (in this case, the hand) on the stabilization output. Since the existing algorithm will extract the skin then the background will be determined and eliminated. The resulting skin version of the video might have small parts like subject's hands which were removed manually for simplicity. However, the face can be automatically targeted in future work using real-time

face detection [28]. In addition, Figure 3.7 (b) shows the effect of the background on the stabilization output. The background, with a high number of corners will eventually force the stabilization algorithm to stabilize the frame according to these features rather than the subject. As a result, the frame will be transformed to match the fixed frame but with unwanted result on the subject face as Figure 3.7 (b). Figure 3.5 shows the effect of scaling as part of the last stage of the stabilization method which is images composition.



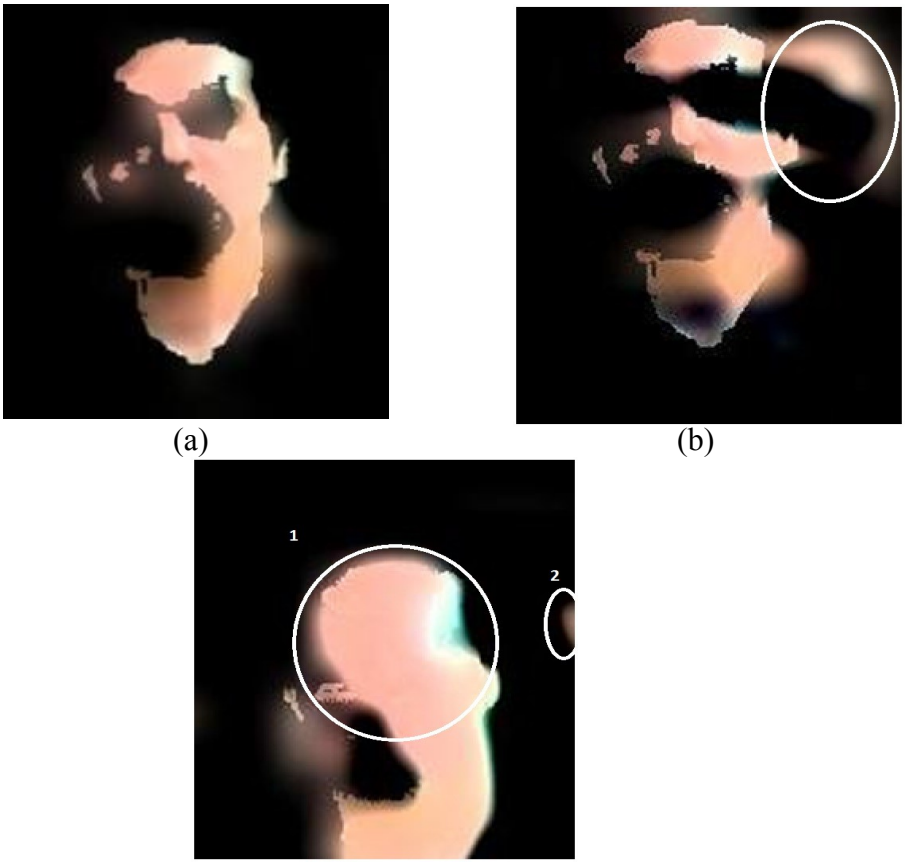
(a) (b)

Figure 3.7. (a) The effect of background on the stabilization output. (b) The effect of unwanted background features used in stabilization.

Figure 3.8 is showing six rows where each row presents 10 frames of videos used by the proposed system. First four rows (A, B, C and D) are showing the absence of stabilization on the magnification results. Row A is showing the original face of the subjects. Row B is showing the result of using Eulerian method on the raw video without any pre-processing. Row C and D are presenting the non-stabilized magnified face skin result which is showing many high frequency changes in color at the face edges. However, frames in row C look fuzzier around the face area where some areas of the background or hand are still in the scene. Row E shows the stabilized face skin without any processing. The last row presents the result of the proposed system which used feature detection in the stabilization layer.

Figure 3.8 the first row shows the original face of the subjects. Row B shows the result of using Eulerian method without any pre-processing. Row C shows the face skin magnified where some area of the background or hand are in the scene. Row D is the non-stabilized magnified face result after deleting the extra parts from the scene. Row E shows the stabilized face skin only without magnification. Last row F presents the final outcomes of the proposed system.

The three images in Figure 3.9 are showing the results of video magnification method on a video with and without the stabilization layer. The outcomes of video magnification process for non-stabilized video will be fuzzy and not clear like Figure 3.9 (b) and (c). However, Figure 3.9 (a) is showing better features and face details that will be targeted for further investigation. The ability to read the redness change because of blood flow is better when the stabilization step is used. Section 3.3 will show in detail how the proposed system achieved that objective.



(a) (b)

(c)

Figure 3.9. The results of video magnification method on the video with stabilization layer (a) and without the stabilization layer (b) and (c).

3.2.2 Video Stabilizing Using Face Tracking

The use of face tracking in contactless heart rate measurements was considered by I. Pavlidis [16] . The region of interest was targeted by tandem face tracking then the method applies the required steps to obtain the heart rate. Our research differs in that we used face tracking to locate the face in each frame of the recorded video. The face was located and tracked using Viola-Jones algorithm [26]. The resulting location of the subject's face was used to create a stabilized version of the original video. However, the assumption that the face size will not change significantly between consecutive frames will help to create a smooth stabilized video.

Figure 3.10 is showing the block diagram of the overall system with the face tracking stabilization step in the pre-processing block. This block is going to create a video from the detected faces from the original video using the following procedure:

1. Face detection algorithm will obtain the face location of the subject if a face was detected in the processed frame.
2. In case of no face detected by the algorithm in the processed frame then we use the location of the last face detected in the previous frame as a face location of the current frame. Table 3.1 shows the total number of frames that had no face detected and the previous good location was used.
3. Using the face location obtained from the previous steps a new video will be created using the faces' locations and the area around them.

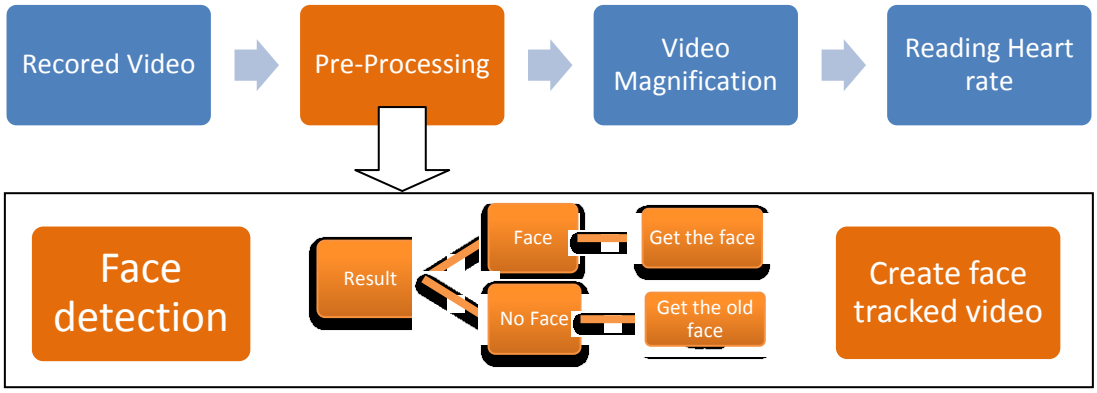


Figure 3.10. Face detection, tracking stabilization in a block diagram.

Subject 1	57 / 3552	1.6 %
Subject 2	98 / 3582	2.7%
Subject 3	103 / 3522	2.9%
Subject 4	52 / 3570	1.5%
Subject 5	86 / 3510	2.5%

Table 3.1 The number of no face detected/total number of frames for 2 minute videos. The biggest number for subject 3 was showing 3% frames only the face was not fully in the stabilized version which is very low number.

Figure 3.11 is showing four rows where each row presents 10 frames of a video. The first two rows, A and B, were presented in Figure 3.8 and used here to make the comparison easy and clear. Row A is showing the original face of the subjects. Second row is presenting the resulting frames of using Eulerian video magnification without any pre-processing. Next, row C is showing the video frames of the created face tracking video. Row D is showing the frames of the face tracked video after applying Eulerian video magnification.

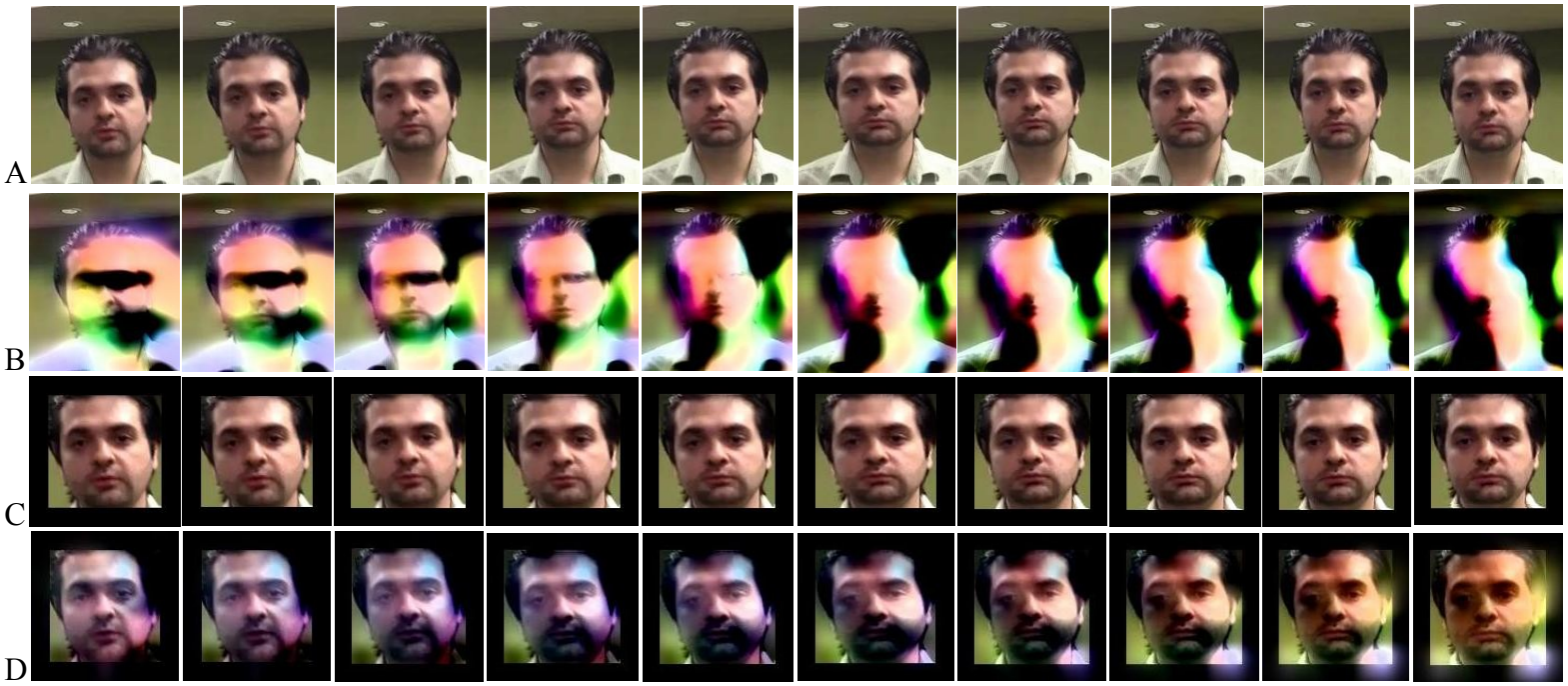


Figure 3.11. Row A is showing the original face of the subjects. Second, row B is presenting the resulted frames of using Eulerian video magnification without any pre-processing. Row C is presenting the video frames of the created face tracking video. Finally, row D frames are the outcome of the face tracked video after applying Eulerian video magnification.

3.2.3 Reading heart rate after using the proposed system

The heart rate measures were captured from the proposed system by two ways. The first method used manual counting of peaks that presented on redness changes plots such as Figure 3.12. The second method used a Fourier Transform (FFT) analysis of the redness change indicator used to automatically capture the heart rate frequency which is the highest peak in the power spectrum.

The following steps describe the full process of the first method that used Manual Counting to measure heart rate:

1. Each 20 second video was classified by the subject motion before stabilizing and magnifying. Since, the experiment was done with subjects walking or running on a treadmill, we used walking, low speed running and high speed running as three main classifications.
2. The classification will be used to set the frequency bands that will be magnified.
3. The video will be stabilized and magnified to be processed for heart rate measurements.
4. Each 20s video was processed in the proposed system to magnify the face skin color changes. The ranges of frequency targeted by magnifying were 0.83 – 1.67 Hz (for ranges of 50-100 bpm) and 1.33 – 2.17 Hz (80 -130 bpm) as needed.
5. Each frame will be accessed to select the targeted area (25 pixels) and compute the average of the Red channel values over all the pixels. This information will be used to :

A. Plot the average redness change through all frames. The overall plot will present the redness change over 20s. Counting the peaks on the produced plot will be used to figure the heart rate per minutes. Figure 3.12 presents the plots and the blue small circles are the peaks.

B. Apply FFT then plotting the magnitude spectrum which will show the highest peak as the heart rate frequency.

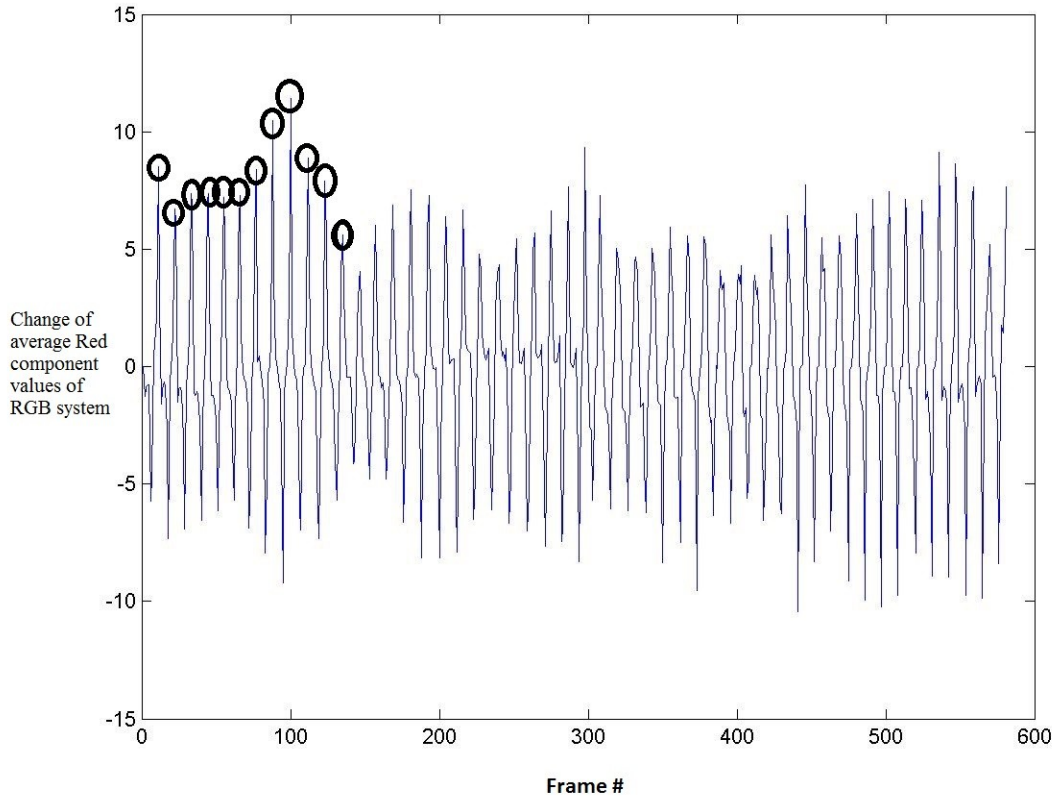


Figure 3.12. The plots of average redness change for 20 seconds video length. The small circles are the peaks.

Selecting the proper frequency limits in video magnification that match the blood flow in the face will increase the accuracy. Comparing different magnified videos for the

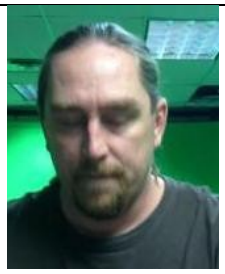
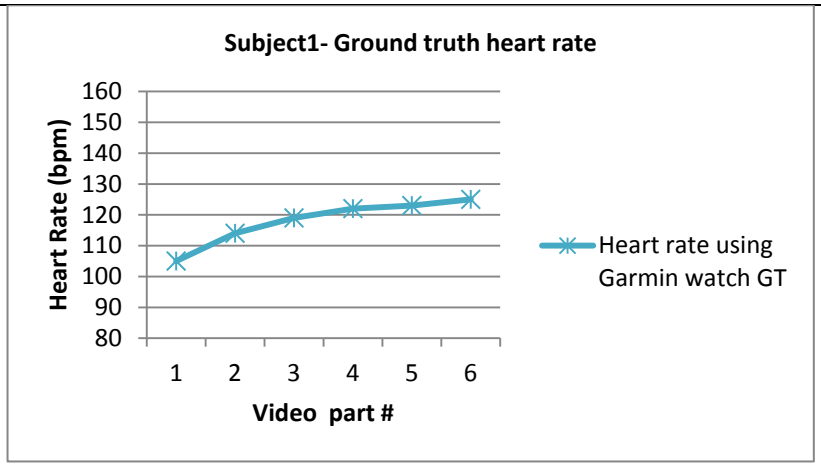
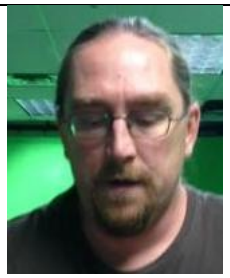
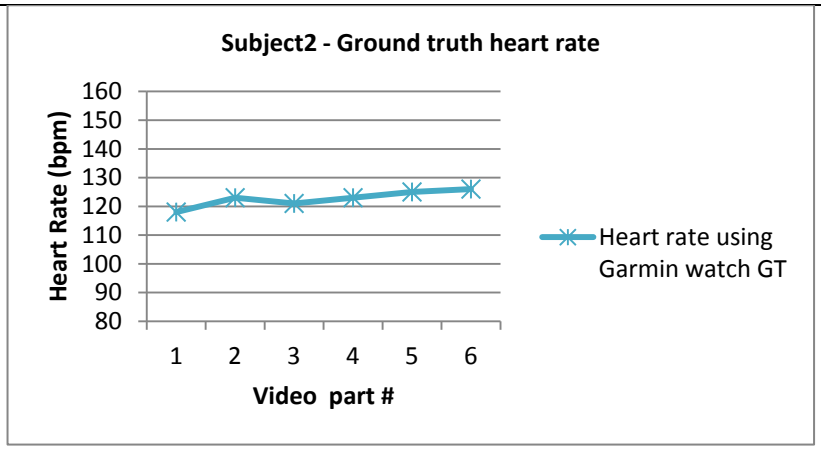
same subject but with different band-pass filters is the goal. Appendix B shows all the frequencies band used in the experiment. To obtain those measurements the magnified videos were processed as follow:

1. Using the video classification described above to determine six frequency band-pass filters that target the actual heart rate. This will produce six videos each one stabilized and magnified for the corresponding frequency band.
2. Selecting an area of interest with a size of 25 pixels (The number selected is small to avoid noise as much as possible)
3. Computing the average of the red channel values (RGB) for the selected area above for each frame of the video.
4. Computing the average of the highest 40 (The average heart rate over all the 30 videos of the five subjects is 120 bpm. This will create 120 peaks of redness which means 40 peaks on each video with 20 seconds length) values resulting from step 3 for each video to be used as redness change indicator.
5. Comparing the redness change indicators values and the highest value is supposedly found in the video that was magnified with an optimal choice of frequency band-pass filter. The results will show values that match the ground truth and other produce bad readings with respect to real heart rate.
6. The highest value determines the heart rate frequency that matches the center of the frequency band filter.

3.3. Experimental Results and Discussion

Table 3.2 presents five persons who were the experiment subjects. The table presents the average heart rate readings captured by the Garmin Forerunner 410 heart rate monitor. Since we process the captured videos in small sections with a period of 20 seconds each, we need to find the average heart rate values that correspond with each video section. Table 3.3 presents the subjects' video sections and the corresponding heart rate average (bpm). Appendix A presents the full data gathered with the time stamps.

Table 3.2. Presents five subject's first frame and the average real heart rate readings.

Name	Original first frame	Two minutes real heart rate reading (average of readings in each 20s video part)
Subject1		
Subject2		

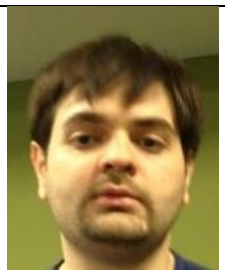
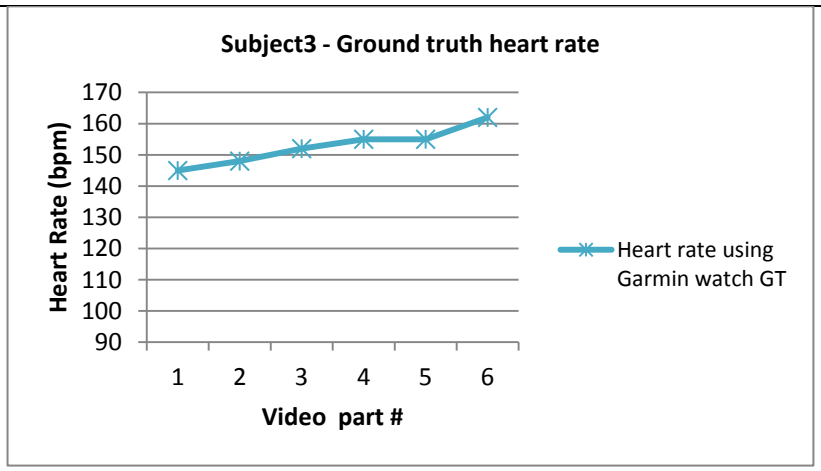
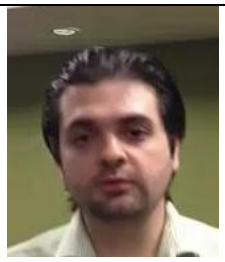
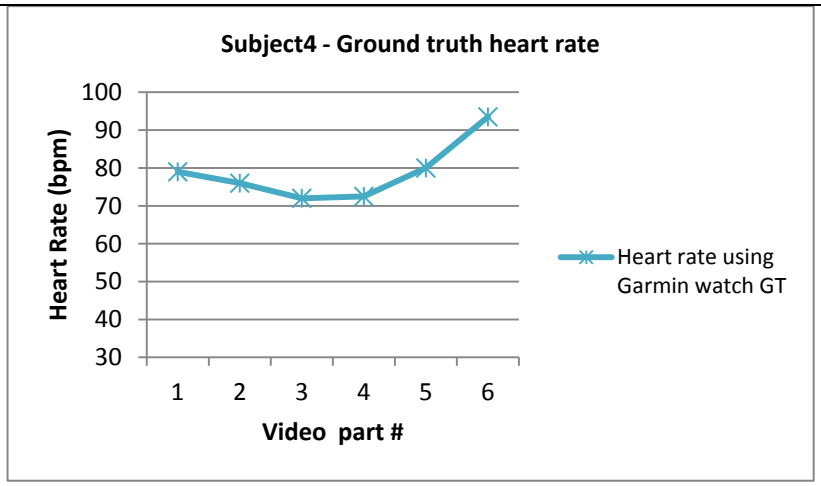
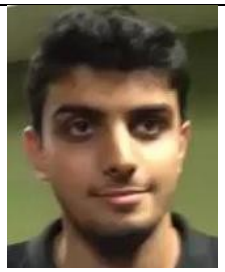
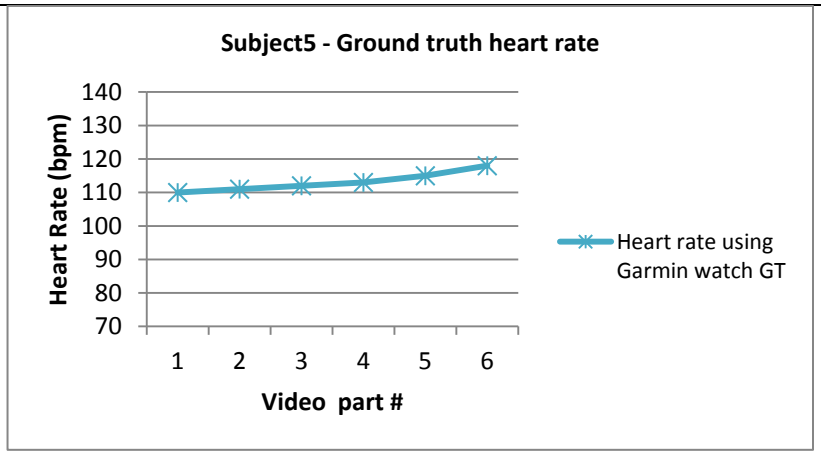
Subject3		
Subject4		
Subject5		

Table 3.3. Subjects corresponding heart rate average (bpm) of each section of video
***there is a high heart rate change in this video section.**

	Video Part 1	Video Part 2	Video Part 3	Video Part 4	Video Part 5	Video Part 6
Subject 1	105	114	119	122	123	125
Subject 2	118	123*	121	123	125	126
Subject 3	145	148	152	155	155	157
Subject 4	79	76	72	72.5	80	93.5
Subject 5	113	112	112	113	115	118

Manual Counting Method

The following Figures 3.13 – 3.17 are the results of heart rate estimates of the proposed system to enhance Eulerian Video Magnification on moving subjects. The heart rate estimates were acquired using Manual Counting method which is presented in Section 3.2.4. The sky blue line presents the actual heart rate of the subjects using Garmin watch our ground truth,. The heart rate readings using video magnification with face tracking (FT) stabilization layer are presented by dark blue line. The third and red line is presenting the heart rate readings using video magnification with skin extraction and features detection and matching stabilizing layer (SE&FDM). It is important for the reader to understand that the method that is closest to the ground truth is the method that is performing best.

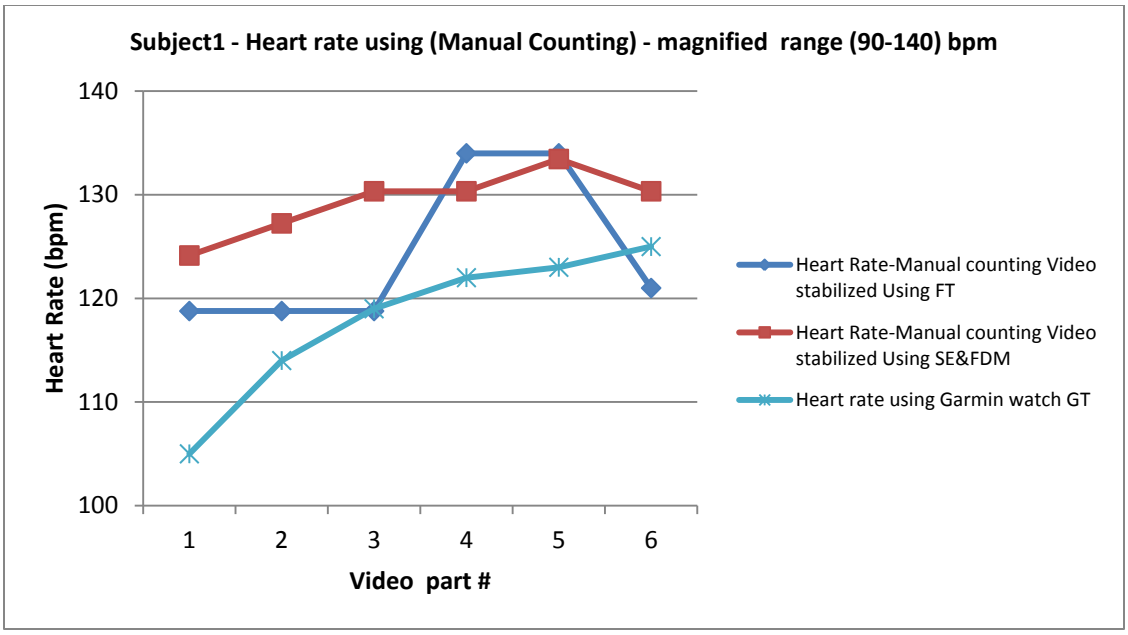


Figure 3.13. Subject 1 heart rate using proposed systems with different stabilization methods.

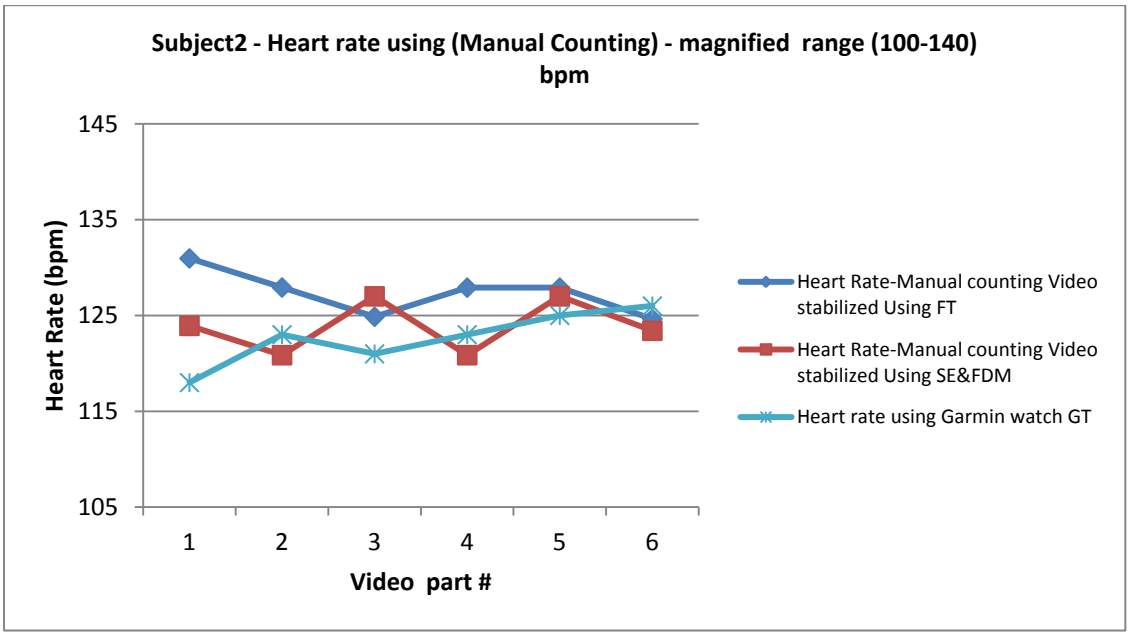


Figure 3.14. Subject 2 heart rate using proposed systems with different stabilization methods.

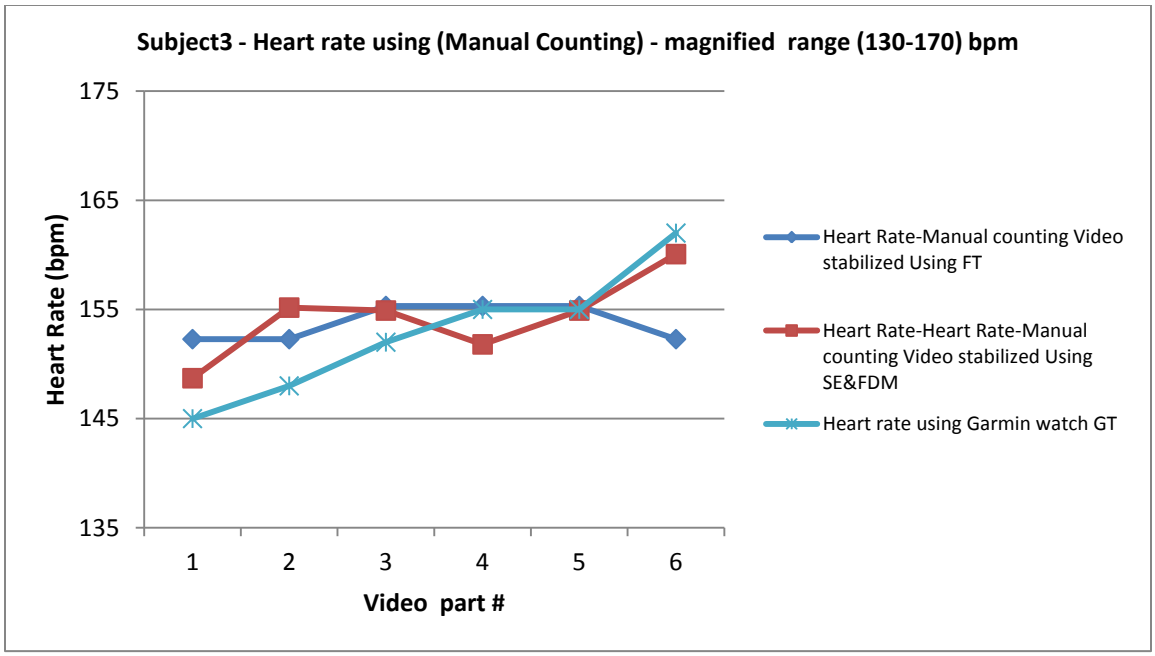


Figure 3.15. Subject 3 heart rate using proposed systems with different stabilization methods.

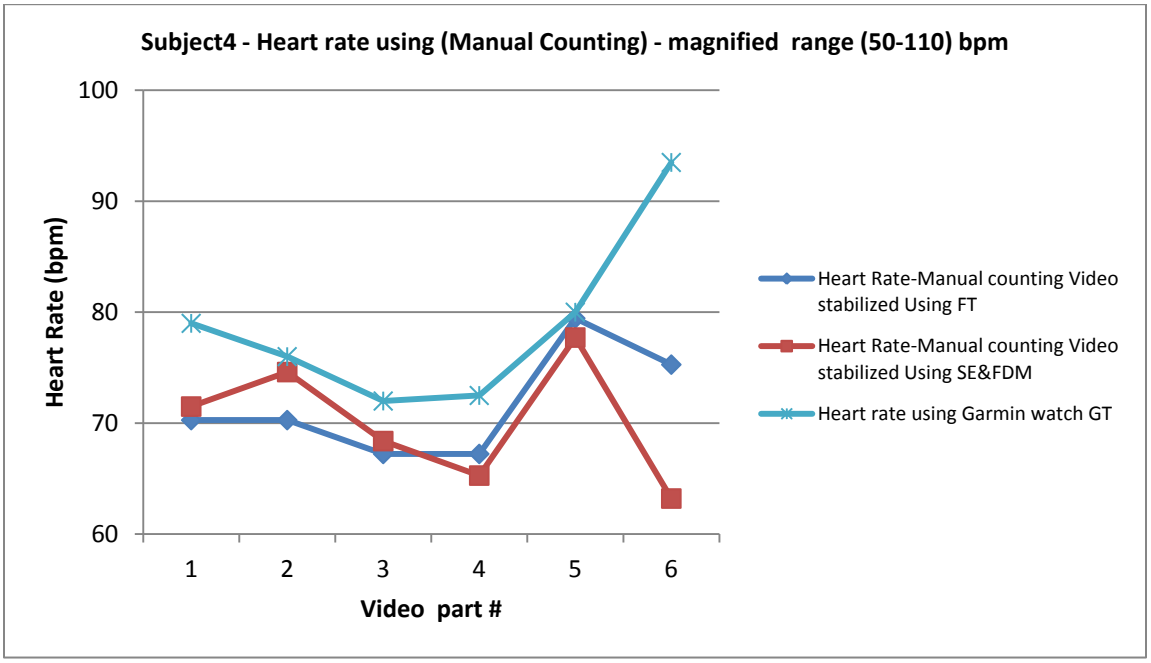


Figure 3.16. Subject 4 heart rate using proposed systems with different stabilization methods.

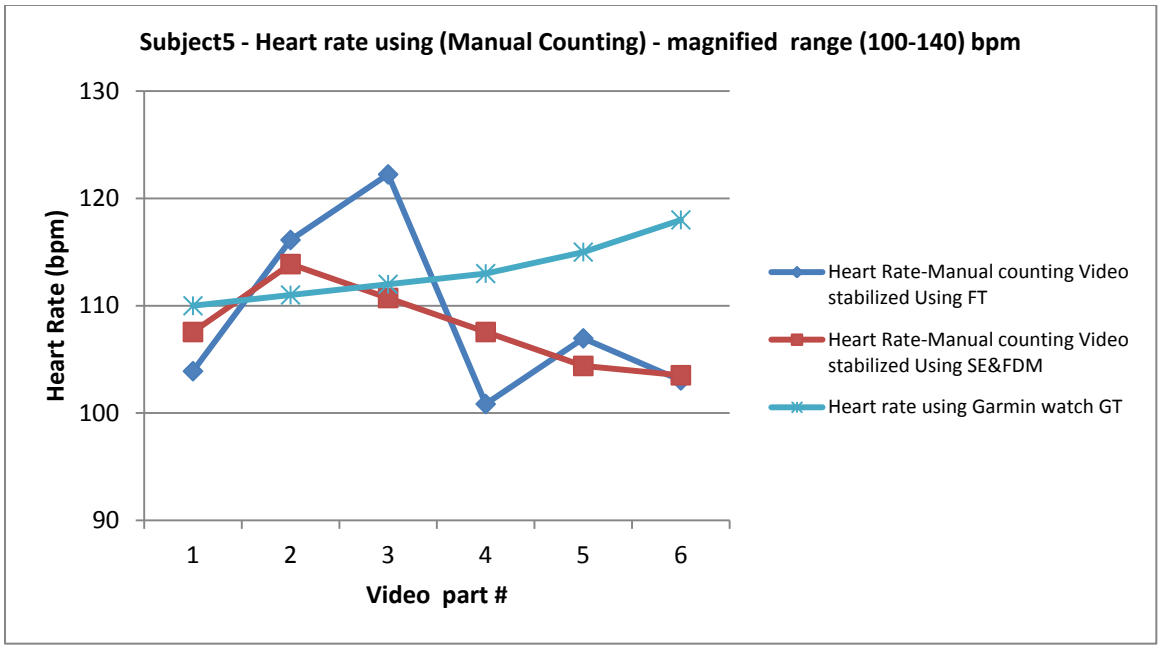


Figure 3.17. Subject 5 heart rate using proposed systems with different stabilization methods.

Figure 3.14, 3.15 and 3.16 present better results than the values in figure 3.13 and 3.17. We posit that this is because the skin is brighter. We examine skin issues in more detail in the next chapter. Figures 3.14, 3.15 and 3.16 are presenting higher accuracy heart rate estimations with error rates less than 5% for 80% of the readings as appendix C shows.

Data presented in the above plots were analysed and the following table shows the average errors on the heart rate estimates using our proposed system is lower. The average error is highest when the videos were not pre-processed for stabilization. The lowest error rates are bolded in Table 3.4 and stabilization is better than the original methods in all cases. The full data presented in Appendix C.

Table 3.4 Average error on the estimated heart rate which counted manually.

	Average Error Stabilization using Face tracking	Average Error Stabilization using SE&FDM	Average Error No Stabilization
Subject 1	**6.65**	9.57	14.59
Subject 2	4.27	**2.88**	5.06
Subject 3	2.30	**2.11**	3.11
Subject 4	**9.14**	10.20	15.87
Subject 5	8.20	**5.21**	9.86

Frequency Band Comparing Method

The heart rate readings which were captured by Garmin watch (ground truth) and were presented in Table 3.2 and 3.3 are forming the purple lines in Figure 3.18-3.22. The figures showed ambiguous change in the heart rate readings that are not related to any source of correct information which are corresponding to not stabilized videos.

As Figures 3.18 - 3.22 will show, the most error in the system occurs when there is no preprocessing of the video data performed. To summarize figures 3.18 – 3.22 are presenting four lines in each figure and they are:

1. Real heart rate reads measured by Garmin watch (purple line).
2. Heart rate reads measured by the proposed system using Face Tracking as stabilizing step (blue line).
3. Heart rate reads measured by the proposed system using skin extraction and stabilizing by features detection and extraction (red line).
4. Heart rate reads measured using video magnification in absent of stabilization pre-processing (green line).

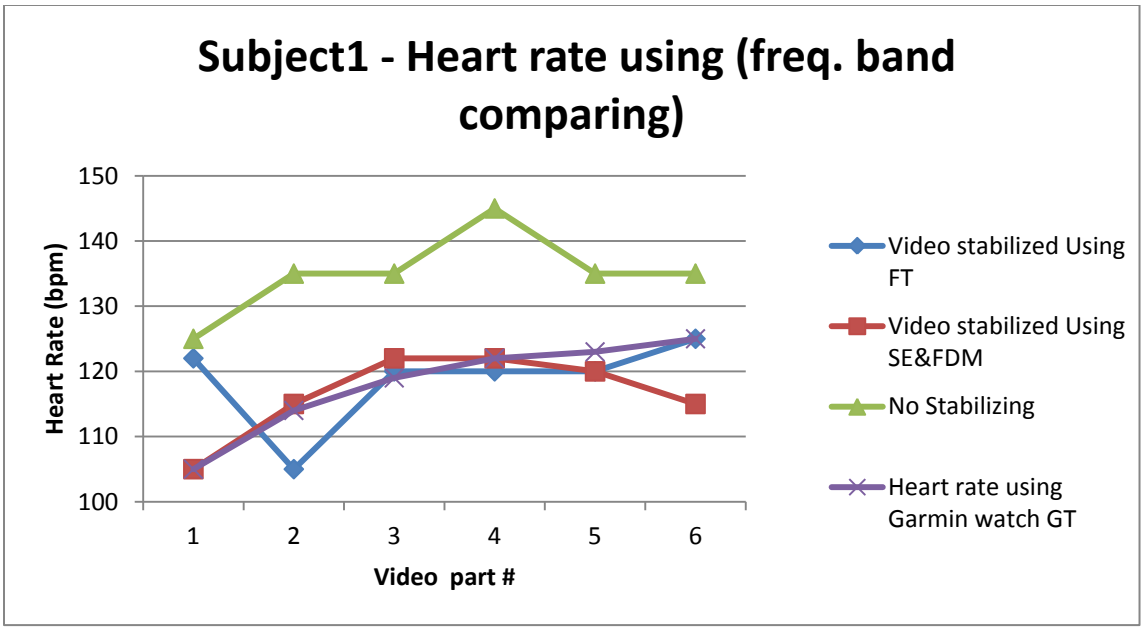


Figure 3.18. The heart rate values of subject 1 after pre-processing in the proposed system. Moreover, the purple line presents the ground truth of heart rate

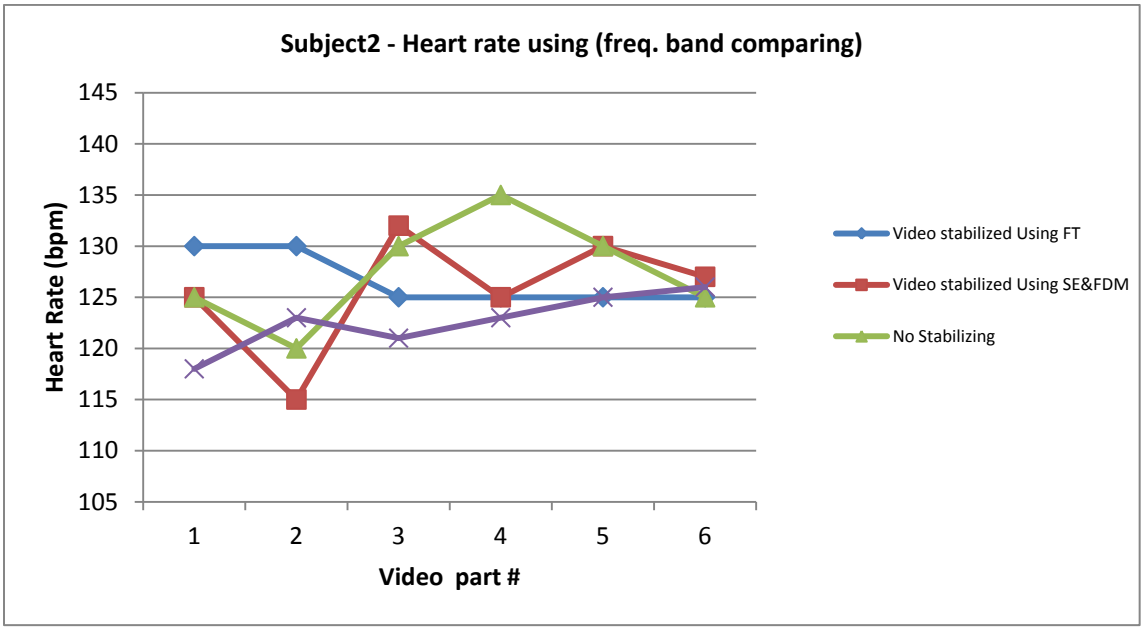


Figure 3.19. The heart rate values of subject 2 after pre-processing in the proposed system

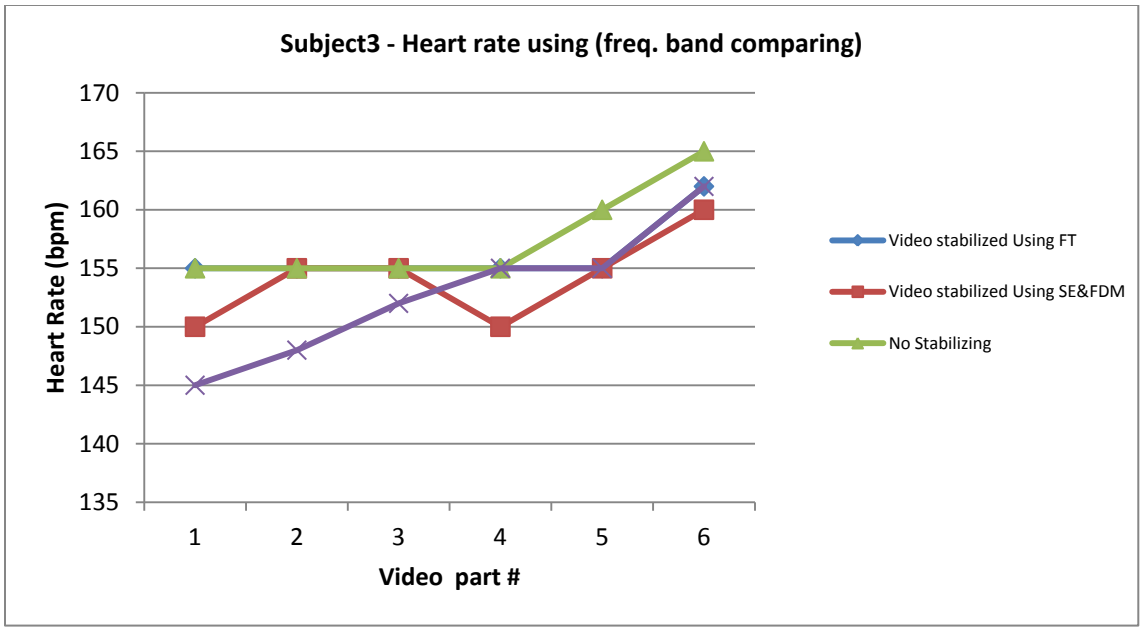


Figure 3.20. The heart rate values of subject 3 after pre-processing in the proposed system

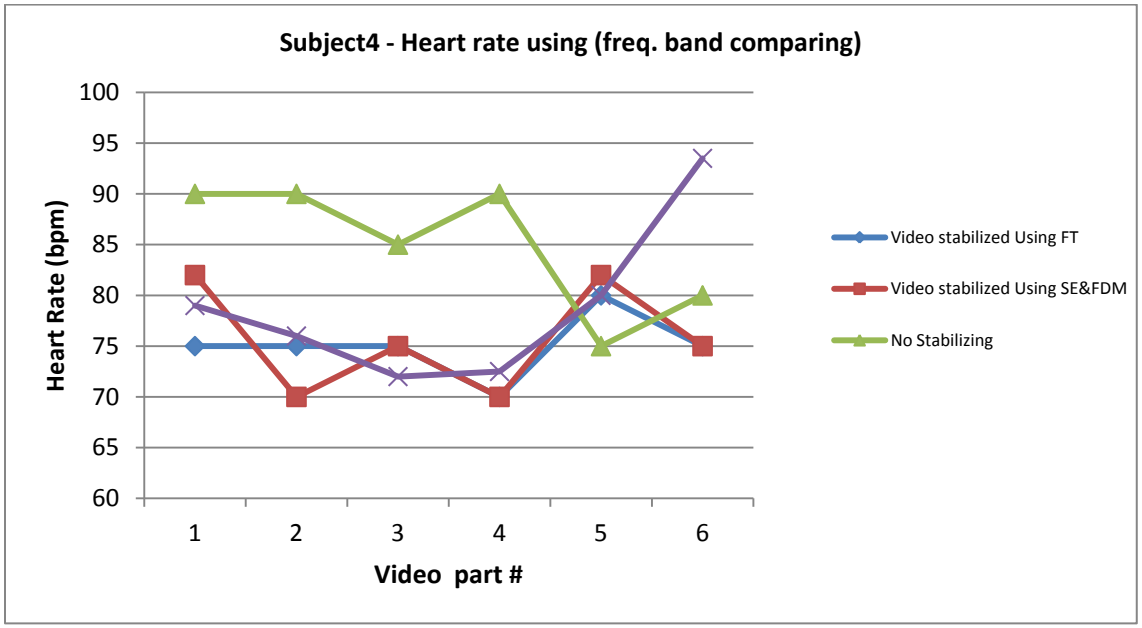


Figure 3.21. The heart rate values of subject 4 after pre-processing in the proposed system

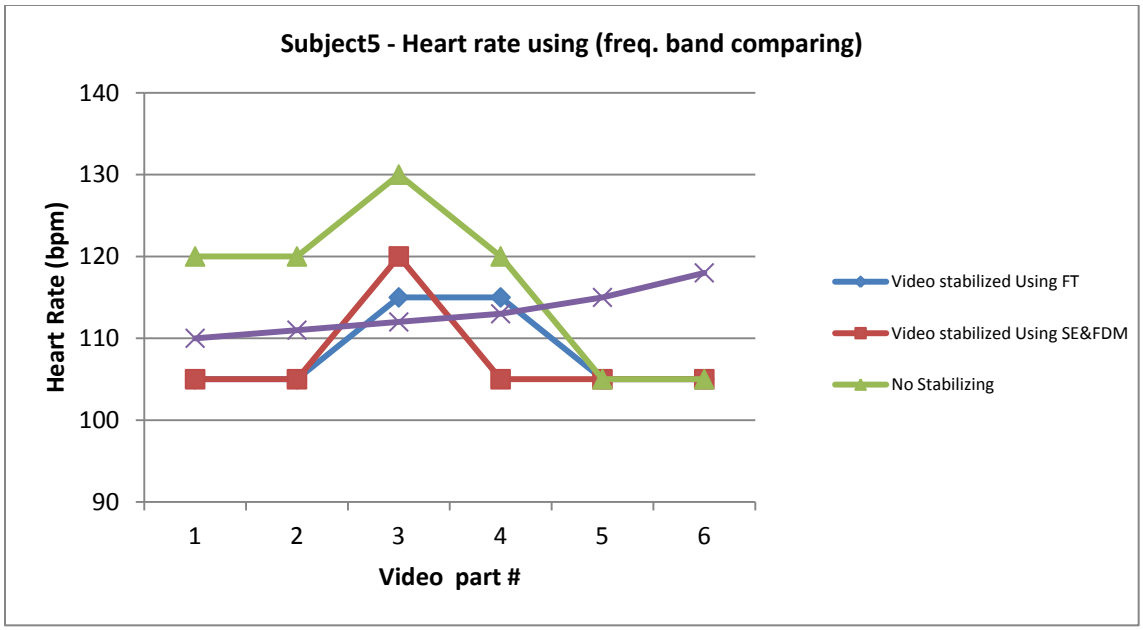


Figure 3.22. The heart rate values of subject 5 after pre-processing in the proposed system

Table 3.5 shows the average errors on the heart rate estimates using our proposed system. The readings were evaluated using the frequency band comparison method. The average error is highest when the videos were not pre-processed for stabilization. Moreover, replacing the bad readings reduces the error rate. The bad reading mean the heart rate reading spike suddenly to a higher or lower value which is not suitable with previous and after reading. The full data presented in Appendix C.

Table 3.5 Average error on the estimated heart rates uaing frequancy bands comparison.

	Average Error Stabilization using Face tracking		Average Error Stabilization using SE&FDM		Average Error No Stabilization
	Original data	With bad data removed	Original data	With bad data removed	
Subject 1	4.83	2.56	2.31	1.17	14.59
Subject 2	3.60	2.28	4.66	3.77	5.06
Subject 3	2.27	1.34	2.44	1.98	3.11
Subject 4	5.63	2.80	6.93	4.36	15.87
Subject 5	5.69	4.62	7.31	6.57	9.86

FFT Counting Method

Transforming the redness change indicator values for each of the frames to the frequency domain helps to indicate the redness frequencies in the videos automatically. Figures 3.18-3.22 clearly show the benefits of stabilizing the video before applying video magnification.

Figure 3.23 – 3.27 are the results of using FFT counting method to spot the heart rate frequency. Clearly the skin color brightness affects the accuracy of reading precise heart rates, but we address these issues in the next chapter. The following five figures 3.23 – 3.27 present heart rate results all subjects independent of the skin tone. Figure 3.23 and 3.24 are presenting the subjects with the lightest skin tone (light). Figure 3.25 and 3.26 present results of the second level of skin tone (mid-tone). The third level skin tone is presented in Figure 3.27 (dark).

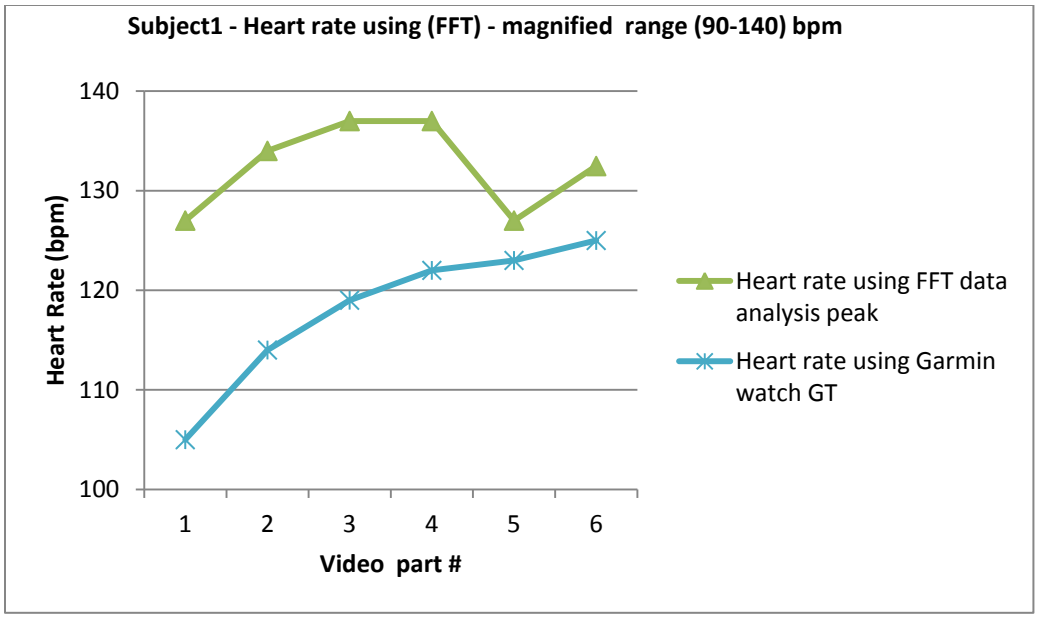


Figure 3.23. Subject 1 heart rate which shows high difference for the first four readings.

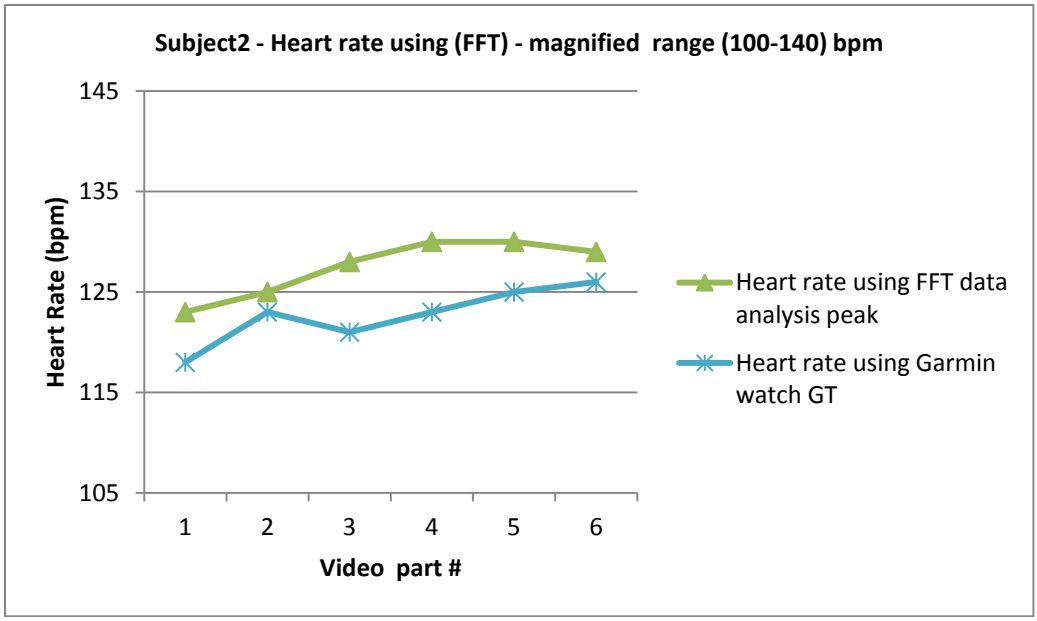


Figure 3.24. Subject 2 heart rate using FFT data analysis estimation.

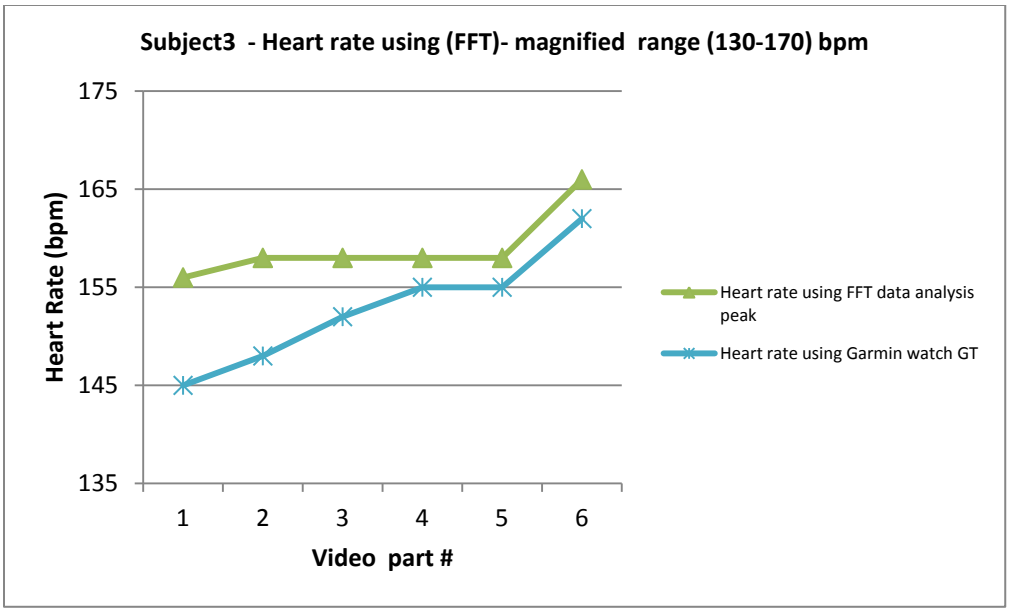


Figure 3.25. Subject 3 heart rate using FFT data analysis estimation.

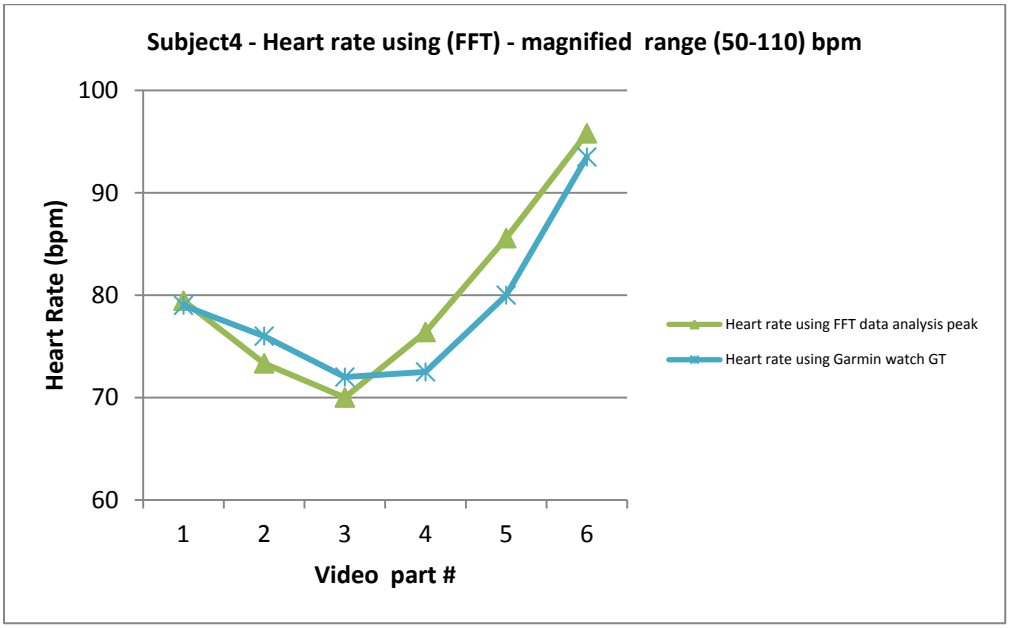


Figure 3.26. Subject 4 heart rate using FFT data analysis estimation

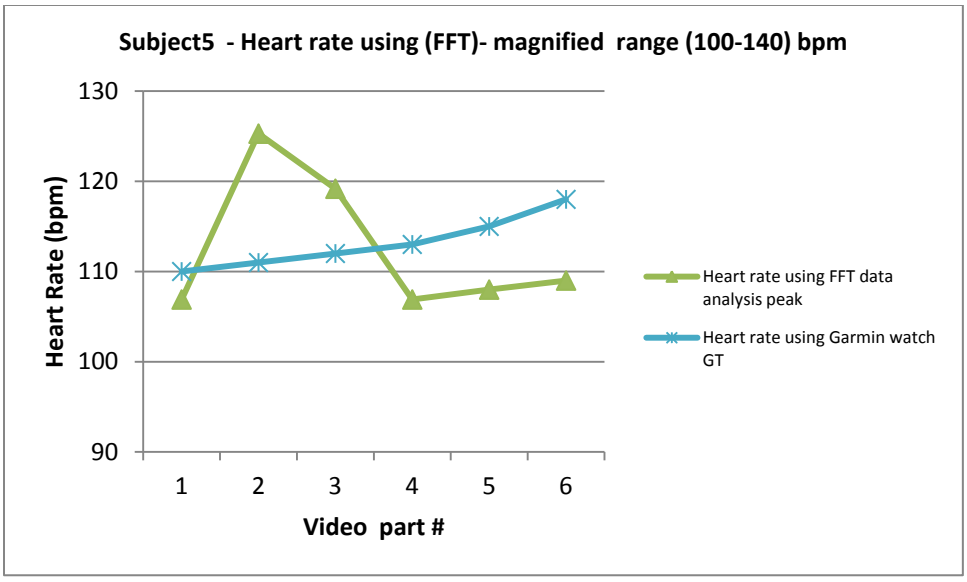


Figure 3.27. Subject 5 heart rate using FFT data analysis estimation.

Results in Figure 3.27 which are less accurate are affected by the skin tone of subject 5. Even though, the stabilization improved the results as per the experiment outcomes, but the subjects skin tone is important factor in the results accuracy. Detailed results of the tests appear in Appendix C. The table of appendix C is showing the absolute heart rate estimation error rate. The lowest error rate will be highlighted.

3.4. Conclusion

The proposed system results showed the importance of pre-processing the realistic videos which mainly the subjects are freely moving. The ability to read more information from such videos is going to expose the Eulerian Video Magnification to wider range of contact-less heart rate monitor applications. We found that lighter skin leads to increasing accuracy of the estimate of the heart rate. Table 3.4 and 3.5 showed lowest error rate for subjects 1 and 2 who have the brighter skin colour. Controlling parameters significantly affects the results. For instance, targeting subjects face skin area that is less affected by high illumination reduces the errors. The next chapter will highlight the proposed system results when illumination and skin tone controlled.

Chapter 4: Skin Color and Illumination

4.1 Introduction

Chapter 3 showed the importance of enabling the existing contactless method to measure heart rate for subjects in motion by presenting experimental results that demonstrate better results with Eulerian Video Magnification on videos with moving subjects. However, many factors like skin color and illumination were affecting the overall results. Furthermore, our examination showed a relationship between skin tone and accuracy of heart rate measurements. Previous work used indoor lighting with level of normal illumination which is evaluated in the range of 270 and 1500 1x (lumens per square meter) but without addressing the direct effect in their proposed methods [8, 11, 12]. In addition, subject motion was causing illumination changes in the targeted skin areas, in particular the appearance and disappearance of specular highlights. Therefore, illumination and skin color are key factors that affect the accuracy. Understanding the factors and how they are going to affect the results is important to eliminate or reduce their impact on the results.

Illumination influence was noticeable when we selected different areas in the subject's face because we would achieve different results. The effect of illumination on the

subject can be directly from a source of light or might be caused by reflected light and this will make the prediction and detection of illumination a complex step. Previously, Finlason and Graham proposed a method to detect the illumination using a chromogenic camera [43]. This camera will capture two images, one normal and one through a colored filter. The method depends on finding the shadow edges by capturing the changes in the scene. Using a method such as this will help to automate the selection of a good location on the face which is ideally illuminated.

The following section will show the methodology used to verify the relation between skin tone and the accuracy of results obtained by our proposed system. Moreover, the outcomes of the proposed system will get better with applying color filters. Section 4.3 will present the results and the discussion for controlling illumination and skin tone enhancement. The conclusion and the future work will be presented in Section 4.4.

4.2 Methodology

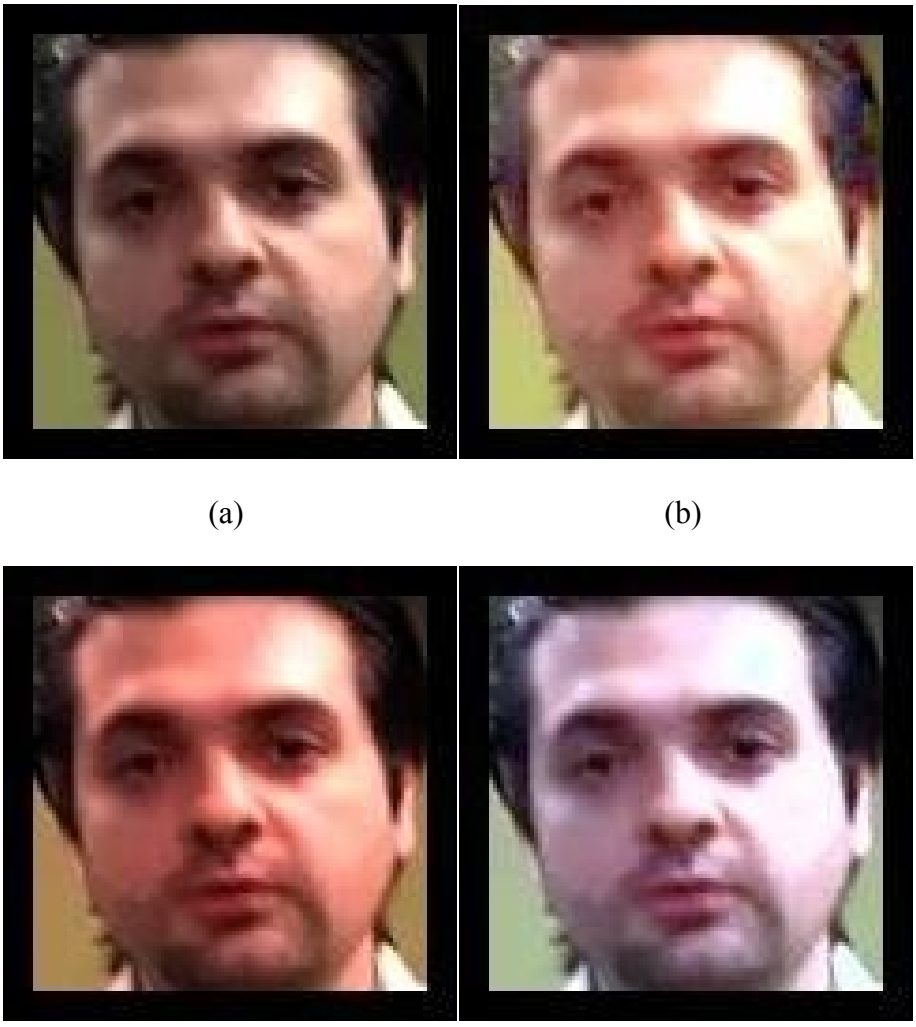
The existence of variable illumination in the frames that need to be processed by Eulerian video magnification produces a level of noise that will affect the final results. Finding areas with less variability of illumination or dealing with this kind of illumination change or eliminating the issue altogether is an important step in the pre-processing phase for the original video.

Our experiment categorizes the skin into light, mid-tone, and dark skin. We show a direct relationship between the skin tone and the accuracy of the proposed system and are encouraged to investigate further. We use the same experimental setup as described in Chapter 3.

4.2.1 Skin tone (brightness – redness) and illumination

Adding brightness to darker skin increases the differentials and makes the redness changes more detectable. Subjects with darker skin were targeted by increasing the brightness on the video by brightening all of the frames. The brightness was added after the pre-process stage, i.e. the brightness adaptation was added to the stabilized video. Figure 4.2 shows a block diagram for the proposed system after adding brightness/redness control layer. In the experiment we used the HSV (H - hue, S - Saturation and V – value of brightness) color system to brightening the frames. The following steps showing how we increase the brightness:

1. Convert the RGB color space frames to HSV and sum the targeted part area of frames V component values.
2. Adding half the average to each pixel V component. Figure 4.1 (b) shows the brightness deference if compared with figure 4.1 (a).
3. Convert the frames back to RGB color space to be processed through the proposed system.



(a) (b)

(c) (d)

Figure 4.1. Four images are of the same frame. A is the original frame , B is the frame with brightness up effect , C is the frame with RGB red channel level up and D is the increase of redness using LAB.

An alternative method examined has the redness changed using two simple techniques. The first method we add 30% more to the red channel in the RGB color system. Figure 4.1 (c) shows the resulting frame after red channel increase. The second used the CIELAB (CIE - International Commission on Illumination, L - lightness, A and B - colour component dimensions) color system to increase the redness of the subject skin. Figure 4.1 (d) presents the output of this method

where the pinkish degree of skin is elevated. This was done by adding 11 to L component and decreasing the B component by 11. This number was selected arbitrarily. After adapting the color, the resulting frame will be converted back to RGB to produce the video with redness increased and fed back into the Eulerian method. Figure 4.2 shows the block diagram of the proposed system in chapter 3 with a step added which is presenting the brightness/redness control.

Figure 4.2. Block diagram shows the proposed method to enhance heart rate estimation through controlling skin brightness and redness.

Illumination influence on result will be shown in the next section. The selection of different areas in the subject's face which is not over saturated will increases the results accuracy. Section 4.3 presents invalid results obtained when the targeted face area was over saturated by light. In our experiment we chose the part of the face that is less affected by saturation manually. Furthermore, balancing the illumination over the image will reduce the affect of unbalanced areas. Figure 4.3 shows the proposed system with an extra step that used to reduce the affect of illumination by selecting areas in the subject's face that are less expose to light.

Figure 4.3. Block diagram shows the proposed enhancement method.

4.3 Experimental Results and Discussion

Changing the brightness or redness in the frames was able to enhance the overall accuracy of the results and Figure 4.5, Figure 4.6 and Figure 4.8 are showing that. The system as the previous section demonstrated was equipped by redness or brightness filters to control these parameters. Moreover, variable illumination was affecting the result accuracy. Selecting different areas of interest to be used by our proposed system was also positively affecting the results. For instance, the result will get better if we select an area with normalized illumination. Furthermore, the subject's position with respect to the light source and the camera will cause high reflection of light in some areas of skin. For instance, Figure 4.4 shows the subject with two different selections for the targeted area of skin: left and right. By selecting the area less affected by illumination variability will be presented as Case1. Case 2 will present the affect of increasing the brightness. Furthermore, Case 3 and 4 will demonstrate the redness affect.

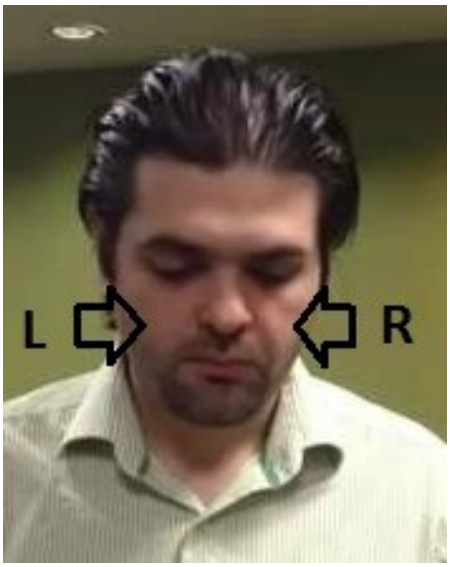


Figure 4.4. Subject's face and arrows pointing to the right and left areas. The right area is exposed more to the room light.

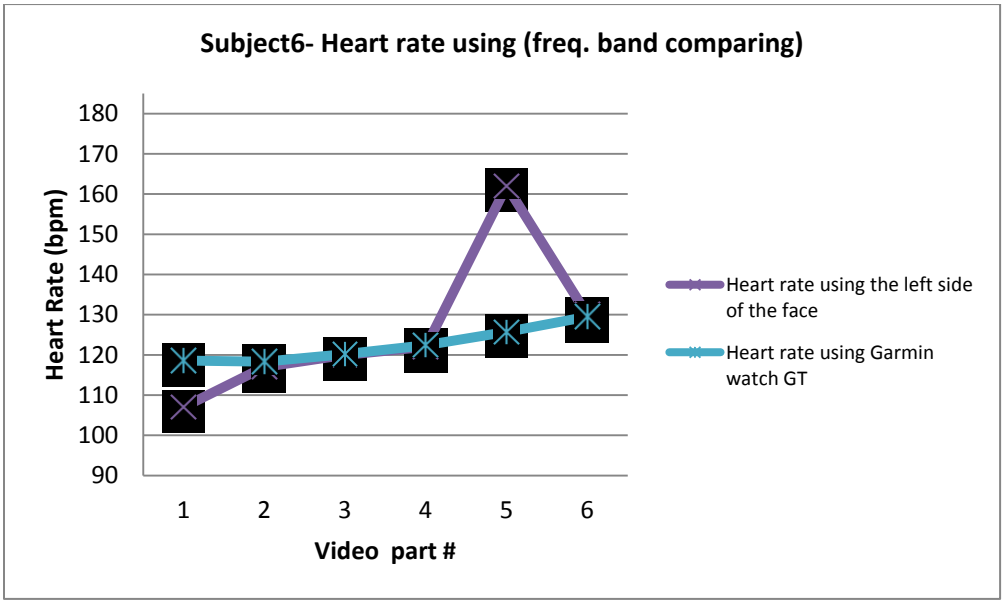


Figure 4.5. Subject 6 heart rate using the left side of the face with less effect of light on the targeted area.

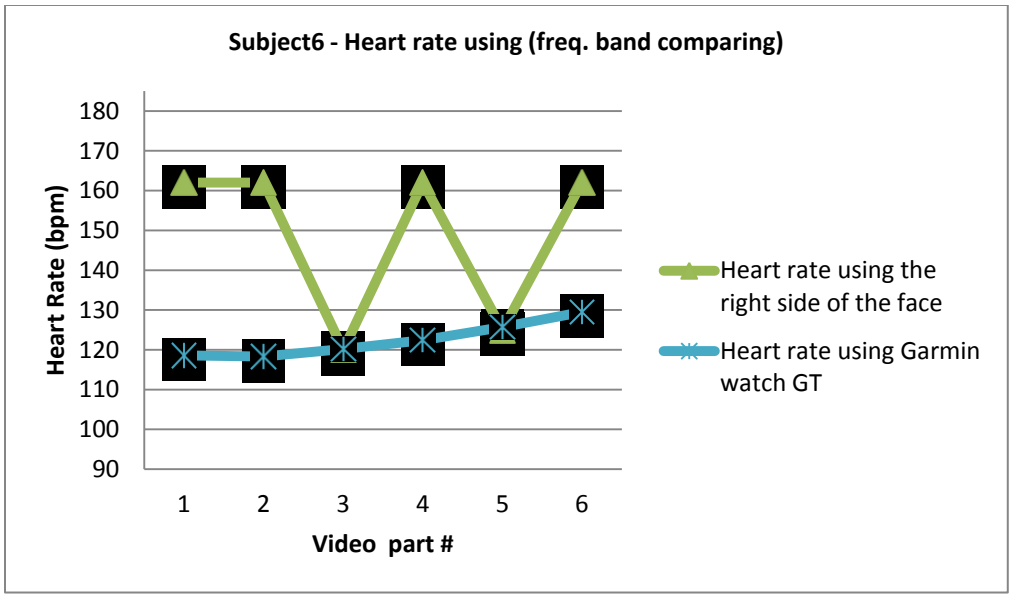


Figure 4.6. Subject 6 heart rate using the right side of the face and the effect of light on the results are clear.

Case 1 - Controlling illumination by selecting less affected area:

Figure 4.5 shows improved readings in the targeted magnified video with band-pass frequency filters that matched the real heart rate. Even though, it is the same skin, the effect of illumination is clear. Figure 4.6 shows the results of the right side of the face which was more exposed to light. The heart rate estimates are fluctuating with high error if compared to the ground truth readings. Furthermore, using the same technique in the subject with darker skin does not improve the results we experienced that with subject number 7.

The location of the light source and more importantly the subjects' movement is causing unbalanced illumination. The values of redness in the targeted area when the illumination is high, the redness values are low then the heart rate reading will be less accurate. Choosing properly balanced illuminated area produces accurate results as Figure 4.5 shows. Furthermore, while we extracting the skin we face similar problems for some

parts of skin that were over saturated by light and, hence that parts were not defined as skin. Figure 4.7 shows those areas in red circles and shows the targeted areas with green circles.

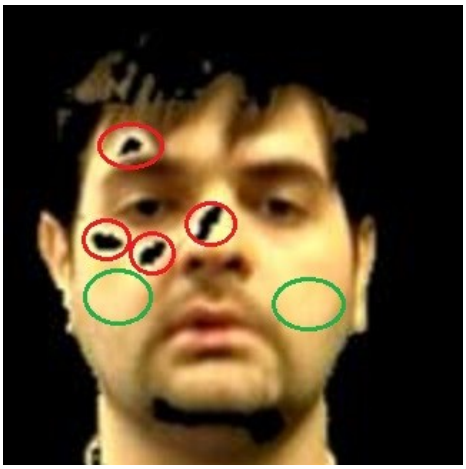


Figure 4.7. Subject's face skin that has some areas not defined as skin because of light.

Case 2 - Increasing the brightness:

Increasing the brightness produces frames that are more readable by our proposed system. Figure 4.8 shows more accurate heart rate reads in video part # 5 when the Face Tracking stabilization preprocessing. Video part # 6 heart rate become more accurate after lighting up the frame for preprocessing system that used SE&FDM (skin extraction + features detection, extraction and matching) to stabilize the video. The average error rate was 7.31 % as Table 3.5 showed with respect to ground truth reads and become 5.90 %. In contrast, subject 4 with light skin is different, the heart rate estimation mostly will improve and Figure 4.9 presents that. Overall the method of increasing the brightness does not seem to be the best solution.

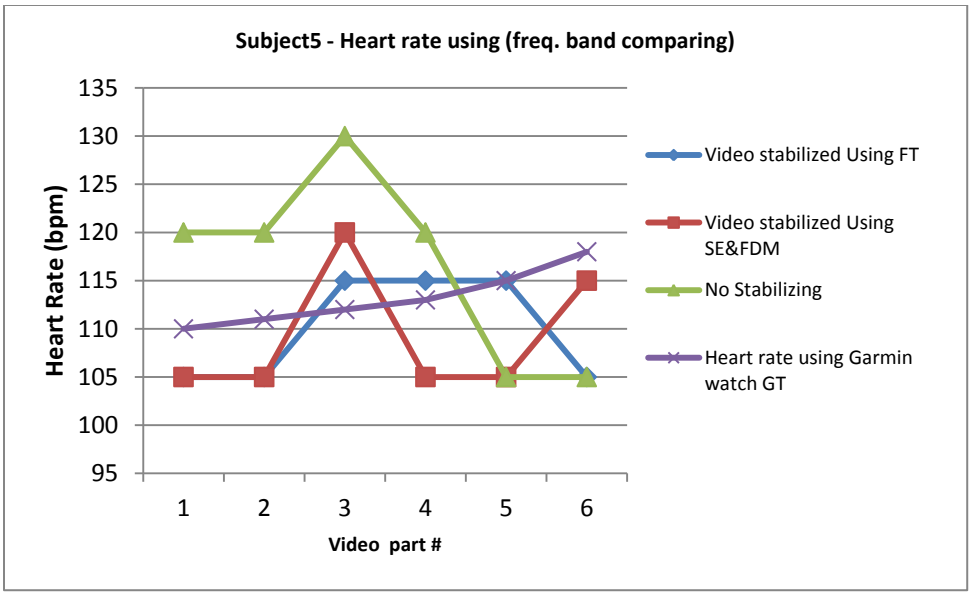


Figure 4.8. Subject 5 heart rate using frequency bands comparison after brightness up the frames.

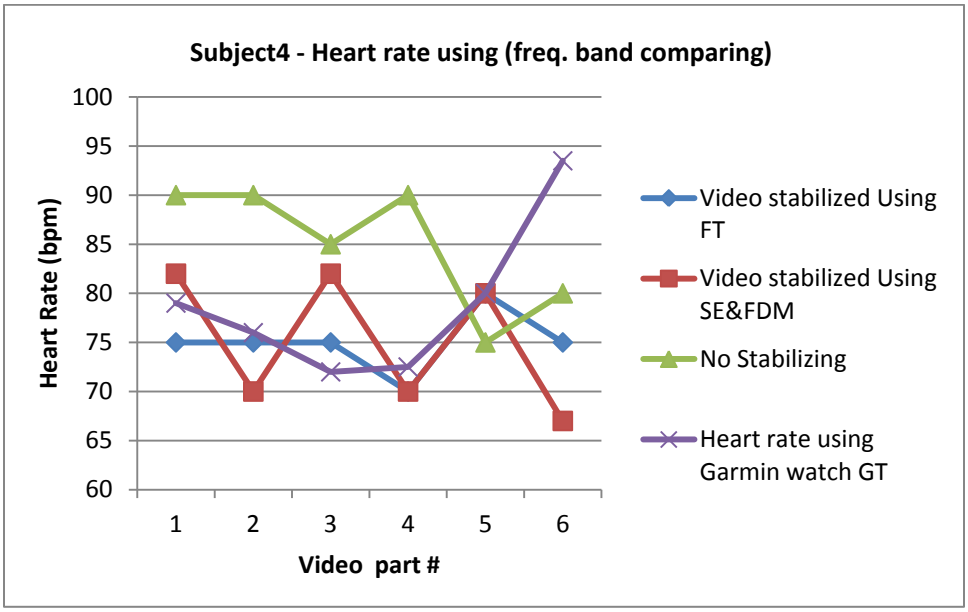


Figure 4.9. Subject 4 heart rate using frequency bands comparison after brightness up the frames.

Case 3 - Increasing the redness using red channel in the RGB color model:

Increasing the redness using red channel in the RGB color model produced more redness in frames as shown in Figure 4.1 (c). However, the results were not promising. The overall values of redness change are lower if compared with values presented in Chapter 3. Figure 4.10 shows the readings mostly unchanged and even with worse accuracy in one reading. Moreover, the darker skin of subject 5 caused the reading not to change. Figure 4.11 shows unchanged values of heart rate readings in presence of SE&FDM (skin extraction + features detection, extraction and matching) stabilizing layer. Results changed in five reads when FT (face tracking) as stabilization layer was used but only one reading become more accurate. The average error rate increased by 20% when video stabilized using SE&FDM (Subject 4). The results of (Subject 5) get changed by 40% increased in the error rate with video stabilized using face tracking.

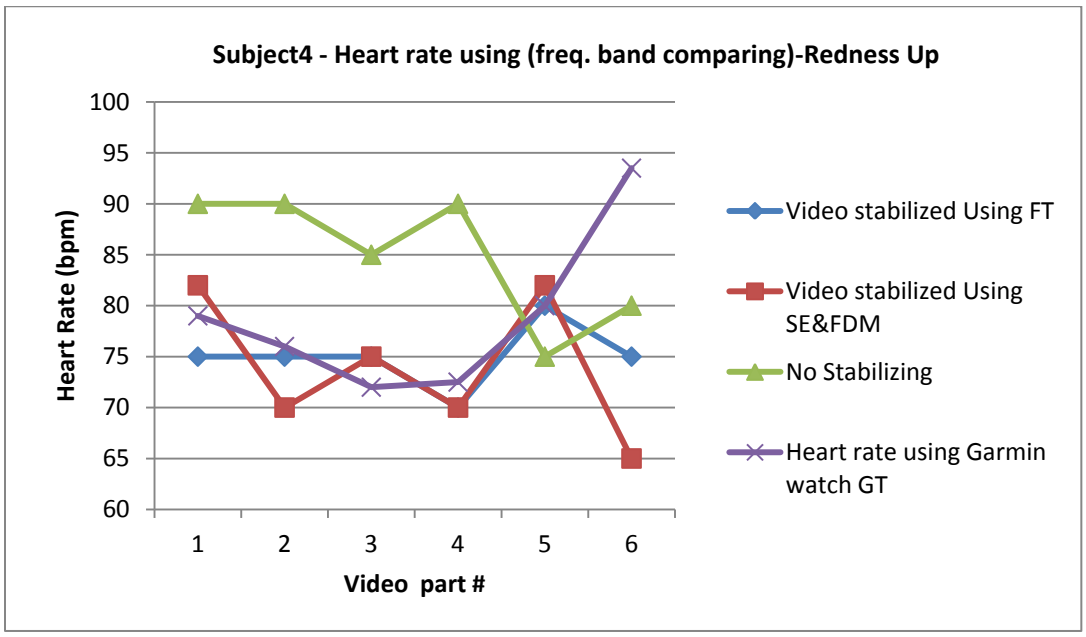


Figure 4.10. Subject4 heart rate reading after increasing the redness in the R channel.

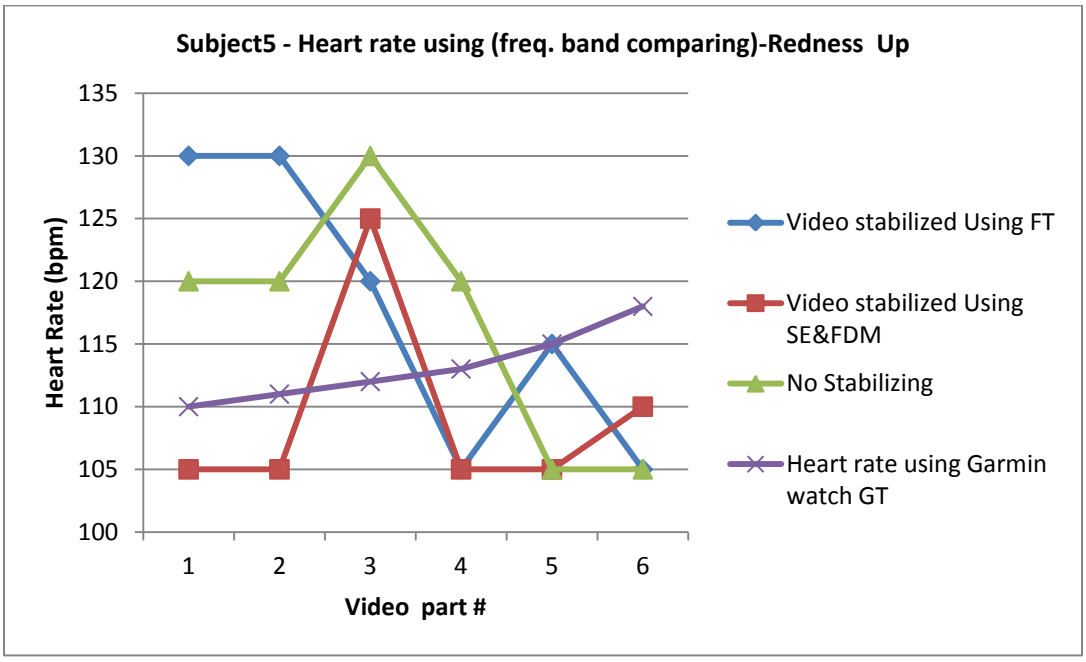


Figure 4.11. Subject 5 heart rate reading after increasing the redness in the R channel.

Overall, the method of increasing the red channel (in RGB space) does not seem to be a workable solution either.

Case 4 - Increasing the redness using CIELAB color system:

Using CIELAB color system to increase the redness in the subject's face skin leads to enhance the existing system to obtain high accuracy results. The experiment was done on Subject 3 using SE&FDM (skin extraction + features detection, extraction and matching) in stabilization. However the estimation of heart rate did not change at most of the videos indicating that it doesn't necessarily decrease the systems capabilities.

4.4 Conclusion

The experiments show the results of the proposed system in this research can be improved when the illumination and skin tone are controlled or modified. Adding light or controlling the selection of the system to choose well lit areas improved the accuracy of heart rate reading through the proposed system. On the other hand, increasing redness using CIELAB color system exhibits promise that the method may help solve some of the skin tone issues..

The future work related to our research will be in the area of increasing the performance of video magnification for subtle color changes when the subjects are in motion. Moreover, controlling the majority of parameters that affect the performance or minimize it will be a good area to investigate. In addition, adding a source of light that might be controlled by our proposed system will make the system adaptable to the environmental illumination change that is caused by the subject motion. Automated selection of face parts that are less affected by illumination is helping to automate the whole system.

Chapter 5: Conclusion and Future Work

5.1. Summary

This thesis proposed systems that used stabilization as pre-processing to enhance Eulerian Video Magnification. The results showed the importance of pre-processing realistic videos which are mainly subjects who are freely moving. The ability to read more information from such videos is going to help Eulerian Video Magnification to be successful in a wide range of contactless heart rate monitor applications. Furthermore, the results demonstrate a proportional relation between skin tone and heart rate measurement accuracy as chapter 4 results suggest. Moreover, other parameters such a variable illumination is playing a major role in the systems accuracy. Thus, controlling these parameters is significantly affecting the results.

Also examined were the skin tone, illumination and skin redness. For instance, targeting a subject's face skin area that is less affected by high illumination changes reduces the errors. The experiment aimed to compare the results of controlling such parameters and their affect on accuracy of readings. Clearly the results showed better readings when the brightness was increased. However, if the skin was bright enough the filter did not affect the results.

Our motivation to use contact free vital signs measuring applications in sports activities is driven by using simple hardware and ability to implement it easily. The results suggest that the proposed system will present significant enhancements to Eulerian Video Magnification when the subject is in motion.

5.2. Future Work

The experiment used average heart rate reading for each 20 seconds video. The proposed system will be evaluated in the future with presence of real time video feeds of heart rate and automated face localization. Moreover, the system will automatically select the area of the face that is less affected by illumination. This will allow the system to make the estimated readings quickly real-time.

Appendix A

Subject 1 heart rate readings from Garmin Connect web page.

Time Stamp	Heart Rate	Video #
2013-12-03T19:10:07.000Z	NoReading	1
2013-12-03T19:10:08.000Z	100	1
2013-12-03T19:10:10.000Z	101	1
2013-12-03T19:10:11.000Z	102	1
2013-12-03T19:10:13.000Z	103	1
2013-12-03T19:10:14.000Z	105	1
2013-12-03T19:10:16.000Z	106	1
2013-12-03T19:10:19.000Z	107	1
2013-12-03T19:10:20.000Z	108	1
2013-12-03T19:10:21.000Z	109	1
2013-12-03T19:10:24.000Z	110	1
2013-12-03T19:10:28.000Z	111	2
2013-12-03T19:10:32.000Z	113	2
2013-12-03T19:10:36.000Z	114	2
2013-12-03T19:10:40.000Z	115	2
2013-12-03T19:10:41.000Z	116	2
2013-12-03T19:10:44.000Z	115	2
2013-12-03T19:10:45.000Z	116	2
2013-12-03T19:10:49.000Z	117	3
2013-12-03T19:10:52.000Z	119	3
2013-12-03T19:10:54.000Z	118	3
2013-12-03T19:10:59.000Z	119	3
2013-12-03T19:11:02.000Z	120	3
2013-12-03T19:11:04.000Z	119	3
2013-12-03T19:11:07.000Z	120	3
2013-12-03T19:11:17.000Z	121	4
2013-12-03T19:11:22.000Z	122	4
2013-12-03T19:11:25.000Z	123	4
2013-12-03T19:11:29.000Z	122	5
2013-12-03T19:11:30.000Z	123	5
2013-12-03T19:11:32.000Z	122	5
2013-12-03T19:11:42.000Z	123	5
2013-12-03T19:11:45.000Z	124	5
2013-12-03T19:11:47.000Z	125	5
2013-12-03T19:11:49.000Z	124	6
2013-12-03T19:11:50.000Z	125	6
2013-12-03T19:11:52.000Z	124	6
2013-12-03T19:11:54.000Z	125	6
2013-12-03T19:12:01.000Z	126	6
2013-12-03T19:12:07.000Z	127	6

Subject 2 heart rate readings from Garmin Connect web page.

Time Stamp	Heart Rate	Video #
2013-12-03T19:13:01.000Z	116	1
2013-12-03T19:13:02.000Z	118	1
2013-12-03T19:13:18.000Z	117	1
2013-12-03T19:13:22.000Z	118	2
2013-12-03T19:13:25.000Z	128	2
2013-12-03T19:13:27.000Z	127	2
2013-12-03T19:13:28.000Z	124	2
2013-12-03T19:13:29.000Z	123	2
2013-12-03T19:13:34.000Z	122	2
2013-12-03T19:13:36.000Z	121	2
2013-12-03T19:13:40.000Z	120	3
2013-12-03T19:13:42.000Z	121	3
2013-12-03T19:14:04.000Z	122	4
2013-12-03T19:14:05.000Z	123	4
2013-12-03T19:14:07.000Z	122	4
2013-12-03T19:14:08.000Z	123	4
2013-12-03T19:14:12.000Z	124	4
2013-12-03T19:14:18.000Z	125	4,5
2013-12-03T19:14:43.000Z	124	5,6
2013-12-03T19:14:50.000Z	125	6
2013-12-03T19:14:51.000Z	126	6
2013-12-03T19:14:54.000Z	127	6
2013-12-03T19:14:55.000Z	128	6

Subject 3 heart rate readings from Garmin Connect web page.

Time Stamp	Heart Rate	Video #
2014-04-12T21:50:47.000Z	145	1
2014-04-12T21:50:54.000Z	146	2
2014-04-12T21:50:56.000Z	147	2
2014-04-12T21:50:59.000Z	148	2
2014-04-12T21:51:01.000Z	149	2
2014-04-12T21:51:06.000Z	148	2
2014-04-12T21:51:08.000Z	147	2
2014-04-12T21:51:10.000Z	148	2
2014-04-12T21:51:13.000Z	149	3
2014-04-12T21:51:19.000Z	150	3
2014-04-12T21:51:23.000Z	151	3
2014-04-12T21:51:24.000Z	152	3
2014-04-12T21:51:27.000Z	153	3
2014-04-12T21:51:28.000Z	154	3
2014-04-12T21:51:31.000Z	155	3
2014-04-12T21:51:35.000Z	156	4
2014-04-12T21:51:37.000Z	155	4
2014-04-12T21:51:56.000Z	156	5
2014-04-12T21:51:57.000Z	155	5
2014-04-12T21:52:05.000Z	154	5
2014-04-12T21:52:12.000Z	155	5
2014-04-12T21:52:16.000Z	156	5
2014-04-12T21:52:20.000Z	157	5
2014-04-12T21:52:23.000Z	158	5
2014-04-12T21:52:26.000Z	159	5
2014-04-12T21:52:32.000Z	160	6
2014-04-12T21:52:34.000Z	161	6
2014-04-12T21:52:37.000Z	162	6
2014-04-12T21:52:39.000Z	163	6
2014-04-12T21:52:43.000Z	164	6
2014-04-12T21:52:49.000Z	165	6
2014-04-12T21:52:51.000Z	NoReading	6

Subject 4 heart rate readings from Garmin Connect web page.

Time Stamp	Heart Rate	Video #
2013-04-02T07:31:48.000Z	NoReading	1
2013-04-02T07:32:00.000Z	79	1,2
2013-04-02T07:32:32.000Z	74	2,3
2013-04-02T07:32:34.000Z	74	3
2013-04-02T07:32:36.000Z	70	3,4
2013-04-02T07:33:14.000Z	75	4,5
2013-04-02T07:33:16.000Z	75	5
2013-04-02T07:33:21.000Z	79	5
2013-04-02T07:33:23.000Z	83	5
2013-04-02T07:33:24.000Z	83	5
2013-04-02T07:33:32.000Z	85	6
2013-04-02T07:33:34.000Z	89	6
2013-04-02T07:33:35.000Z	89	6
2013-04-02T07:33:40.000Z	95	6
2013-04-02T07:33:42.000Z	97	6
2013-04-02T07:33:45.000Z	101	6
2013-04-02T07:33:46.000Z	102	6

Subject 5 heart rate readings from Garmin Connect web page.

Time Stamp	Heart Rate	Video #
2013-04-02T16:43:17.000Z	NoReading	1
2013-04-02T16:43:19.000Z	110	1
2013-04-02T16:43:20.000Z	110	1
2013-04-02T16:43:23.000Z	111	1
2013-04-02T16:43:25.000Z	112	1
2013-04-02T16:43:27.000Z	114	1
2013-04-02T16:43:30.000Z	113	1
2013-04-02T16:43:31.000Z	114	1
2013-04-02T16:43:33.000Z	113	1
2013-04-02T16:43:34.000Z	114	1
2013-04-02T16:43:40.000Z	113	2
2013-04-02T16:43:41.000Z	112	2
2013-04-02T16:43:45.000Z	113	2
2013-04-02T16:43:46.000Z	114	2
2013-04-02T16:43:51.000Z	113	2
2013-04-02T16:43:52.000Z	112	2
2013-04-02T16:43:54.000Z	111	2
2013-04-02T16:43:55.000Z	112	2
2013-04-02T16:43:57.000Z	113	2
2013-04-02T16:44:00.000Z	112	3
2013-04-02T16:44:07.000Z	111	3
2013-04-02T16:44:14.000Z	112	3
2013-04-02T16:44:19.000Z	113	4
2013-04-02T16:44:21.000Z	112	4
2013-04-02T16:44:31.000Z	113	4
2013-04-02T16:44:33.000Z	112	4
2013-04-02T16:44:43.000Z	113	5
2013-04-02T16:44:45.000Z	114	5
2013-04-02T16:44:48.000Z	115	5
2013-04-02T16:44:50.000Z	116	5
2013-04-02T16:44:53.000Z	115	5
2013-04-02T16:45:04.000Z	116	6
2013-04-02T16:45:05.000Z	117	6
2013-04-02T16:45:10.000Z	118	6
2013-04-02T16:45:13.000Z	119	6
2013-04-02T16:45:14.000Z	118	6
2013-04-02T16:45:16.000Z	119	6
2013-04-02T16:45:17.000Z	118	6

Subject 6 heart rate readings from Garmin Connect web page.

Time Stamp	Heart Rate	Video #
2014-04-12T21:45:24.000Z	NoReading	1
2014-04-12T21:45:26.000Z	118	1
2014-04-12T21:45:29.000Z	119	1
2014-04-12T21:45:33.000Z	120	1
2014-04-12T21:45:35.000Z	121	1
2014-04-12T21:45:39.000Z	120	1
2014-04-12T21:45:41.000Z	119	1
2014-04-12T21:45:42.000Z	117	1
2014-04-12T21:45:43.000Z	115	1
2014-04-12T21:45:45.000Z	116	2
2014-04-12T21:45:46.000Z	118	2
2014-04-12T21:45:48.000Z	119	2
2014-04-12T21:45:53.000Z	118	2
2014-04-12T21:45:54.000Z	119	2
2014-04-12T21:45:55.000Z	120	2
2014-04-12T21:46:02.000Z	118	2
2014-04-12T21:46:05.000Z	119	3
2014-04-12T21:46:06.000Z	120	3
2014-04-12T21:46:08.000Z	121	3
2014-04-12T21:46:09.000Z	120	3
2014-04-12T21:46:21.000Z	121	3
2014-04-12T21:46:27.000Z	122	4
2014-04-12T21:46:28.000Z	123	4
2014-04-12T21:46:30.000Z	122	4
2014-04-12T21:46:33.000Z	123	4
2014-04-12T21:46:42.000Z	122	4
2014-04-12T21:46:46.000Z	123	5
2014-04-12T21:46:48.000Z	124	5
2014-04-12T21:46:49.000Z	125	5
2014-04-12T21:46:59.000Z	126	5
2014-04-12T21:47:00.000Z	127	5
2014-04-12T21:47:02.000Z	128	5
2014-04-12T21:47:03.000Z	127	5
2014-04-12T21:47:10.000Z	128	6
2014-04-12T21:47:21.000Z	129	6
2014-04-12T21:47:23.000Z	130	6
2014-04-12T21:47:24.000Z	131	6

Subject 7 heart rate readings from Garmin Connect web page.

Time Stamp	Heart Rate	Video #
2014-03-17T02:02:16.000Z	no reading	1
2014-03-17T02:02:19.000Z	68	1
2014-03-17T02:02:32.000Z	68	1
2014-03-17T02:02:35.000Z	65	1
2014-03-17T02:02:38.000Z	71	2
2014-03-17T02:02:40.000Z	71	2
2014-03-17T02:02:48.000Z	67	2
2014-03-17T02:02:54.000Z	66	2
2014-03-17T02:02:55.000Z	68	2
2014-03-17T02:03:03.000Z	67	3
2014-03-17T02:03:06.000Z	66	3
2014-03-17T02:03:25.000Z	64	4
2014-03-17T02:03:33.000Z	60	4
2014-03-17T02:03:41.000Z	66	5
2014-03-17T02:03:42.000Z	66	5
2014-03-17T02:03:52.000Z	62	5
2014-03-17T02:03:57.000Z	61	6
2014-03-17T02:04:04.000Z	71	6
2014-03-17T02:04:05.000Z	89	6
2014-03-17T02:04:06.000Z	95	6
2014-03-17T02:04:08.000Z	95	6
2014-03-17T02:04:11.000Z	104	6
2014-03-17T02:04:12.000Z	104	6
2014-03-17T02:04:13.000Z	108	6

Appendix B

The frequencies used to produce different magnified videos to do comparison on.

	Band pass filter 10 Hz ranges	Band pass filter 5 Hz ranges 1	Band pass filter 5 Hz ranges 2
Subject1	90 -100	100-105	98-102
	100-110	105-110	103-107
	110-120	110-115	108-112
	120-130	115-120	113-117
	130-140	120-125	118-122
	140-150	125-130	123-127
Subject2	90-100	105-110	108-112
	100-110	110-115	113-117
	110-120	115-120	118-122
	120-130	120-125	123-127
	130-140	125-130	128-132
	140-150	130-135	133-137
Subject3	120-130	135-140	138-142
	130-140	140-145	143-147
	140-150	145-150	148-152
	150-160	150-155	153-157
	160-170	155-160	158-162
	170-180	160-165	163-167
Subject4	50-60	65-70	63-67
	60-70	70-75	68-72
	70-80	75-80	73-77
	80-90	80-85	78-82
	90-100	85-90	83-87
	100-110	90-95	88-92
Subject5	90-100	100-105	103-107
	100-110	105-110	108-112
	110-120	110-115	113-117
	120-130	115-120	118-122
	130-140	120-125	123-127
	140-150	125-130	128-132

Appendix C

The following table shows all subjects heart rates (Ground truth, Face Tracking stabilization method, Error rate, Skin Extraction and feature detection , extraxtion and matching, Error rate, Not stabilized, Error rate) the video magnified results are obtained using comparing technique. Error rate column corresponds to the estimated heart rate if compared with real heart rate.

	Video #	Ground Truth (Garmin)	Face Tracking	Error %	Skin E,M & D	Error %	Not Stabilized	Error %
Subject 1	1	105	122	16.19048	105	0	125	19.04762
	2	114	105	7.89474	115	0.877193	135	18.42105
	3	119	120	0.840336	122	2.521008	135	13.44538
	4	122	120	1.63934	122	0	145	18.85246
	5	123	120	2.43902	120	2.43902	135	9.756098
	6	125	125	0	115	8	135	8
Subject 2	1	118	130	10.16949	125	5.932203	125	5.932203
	2	123	130	5.691057	115	6.50407	120	2.43902
	3	121	125	3.305785	132	9.090909	130	7.438017
	4	123	125	1.626016	125	1.626016	135	9.756098
	5	125	125	0	130	4	130	4
	6	126	125	0.79365	127	0.793651	125	0.79365
Subject 3	1	145	155	6.896552	150	3.448276	155	6.896552
	2	148	155	4.72973	155	4.72973	155	4.72973
	3	152	155	1.973684	155	1.973684	155	1.973684
	4	155	155	0	150	3.22581	155	0
	5	155	155	0	155	0	160	3.225806
	6	162	162	0	160	1.234568	165	1.851852
Subject 4	1	79	75	5.063291	82	3.797468	90	13.92405
	2	76	75	1.315789	70	7.894737	90	18.42105
	3	72	75	4.166667	75	4.166667	85	18.05556
	4	72.5	70	3.44828	70	3.44828	90	24.13793
	5	80	80	0	82	2.5	75	6.25
	6	93.5	75	19.7861	75	19.7861	80	14.4385
Subject 5	1	110	105	4.54545	105	4.54545	120	9.090909
	2	111	105	5.40541	105	5.40541	120	8.108108
	3	112	115	2.678571	120	7.142857	130	16.07143
	4	113	115	1.769912	105	7.07965	120	6.19469
	5	115	105	8.69565	105	8.69565	105	8.69565
	6	118	105	11.0169	105	11.0169	105	11.0169

Face tracking result 8 cases out of 30 where heart rate measurement error is lowest if compared to ground truth. 9 cases are lowest error rate using skin extraction + features detection, extraction and matching stabilization. 10 cases are the same error rate produced of using both methods of stabilization. 3 Cases are presenting the lowest error rate when no stabilization is used.

The following table shows all subjects heart rates (Ground truth, Face Tracking stabilization method, Error rate, Skin Extraction and feature detection , extraction and matching, Error rate FFT peak, Error rate) the video magnified results are obtained using manual counting technique. Error rate column corresponds to the estimated heart rate if compared with real heart rate.

	Video #	Ground Truth (Garmin)	FT (MC)	Error % FT (MC)	Skin E,M & D	Error % Skin	FFT peak	Error % FFT
Subject1	1	105	119	13.33333	124	18.09524	127	20.95238
	2	114	119	4.385965	127	11.40351	134	17.54386
	3	119	119	0	130	9.243697	137	15.12605
	4	122	134	9.836066	130	6.557377	137	12.29508
	5	123	134	8.943089	133	8.130081	127	3.252033
	6	125	121	3.2	130	4	132	5.6
Subject 2	1	118	131	11.01695	124	5.084746	123	4.237288
	2	123	128	4.065041	121	1.62602	125	1.62602
	3	121	125	3.305785	127	4.958678	128	5.785124
	4	123	128	4.065041	121	1.62602	130	5.691057
	5	125	128	2.4	127	1.6	130	4
	6	126	125	0.79365	123	2.38095	129	2.380952
Subject 3	1	145	152	4.827586	149	2.758621	156	7.586207
	2	148	152	2.702703	155	4.72973	158	6.756757
	3	152	155	1.973684	155	1.973684	158	3.947368
	4	155	155	0	152	1.93548	158	1.935484
	5	155	155	0	155	0	158	1.935484
	6	157	155	1.27389	160	1.910828	166	5.732484
Subject 4	1	65	70	7.692308	72	10.76923	79	21.53846
	2	65	70	7.692308	75	15.38462	73	12.30769
	3	70	67	4.28571	68	2.85714	70	0
	4	72.5	67	7.58621	65	10.3448	76	4.827586
	5	80	79	1.25	78	2.5	86	7.5
	6	93.5	75	19.7861	63	32.6203	96	2.673797
Subject 5	1	110	104	5.45455	108	1.81818	107	2.72727
	2	111	116	4.504505	114	2.702703	125	12.61261
	3	112	122	8.928571	111	0.89286	119	6.25
	4	113	101	10.6195	108	4.42478	107	5.30973
	5	115	107	6.95652	104	9.56522	108	6.08696
	6	118	103	12.7119	104	11.8644	109	7.62712

References

[1] J. C. Lin, "Microwave sensing of physiological movement and volume change: A review," *Bioelectromagnetics,* vol. 13, pp. 557-565, 1992.

[2] F. Mohammad-Zadeh, F. Taghibakhsh and B. Kaminska, "Contactless heart monitoring (CHM)," in *Electrical and Computer Engineering, 2007. CCECE 2007. Canadian Conference On,* 2007, pp. 583-585.

[3] G. Matthews, B. Sudduth and M. Burrow, "A non-contact vital signs monitor," *Critical Reviews™ in Biomedical Engineering,* vol. 28, 2000.

[4] M. Garbey, N. Sun, A. Merla and I. Pavlidis, "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery," *Biomedical Engineering, IEEE Transactions On,* vol. 54, pp. 1418-1426, 2007.

[5] L. Scalise and U. Morbiducci, "Non-contact cardiac monitoring from carotid artery using optical vibrocardiography," *Med. Eng. Phys.,* vol. 30, pp. 490-497, 2008.

[6] M. Brink, C. H. Müller and C. Schierz, "Contact-free measurement of heart rate, respiration rate, and body movements during sleep," *Behavior Research Methods,* vol. 38, pp. 511-521, 2006.

[7] K. Humphreys, T. Ward and C. Markham, "Noncontact simultaneous dual wavelength photoplethysmography: a further step toward noncontact pulse oximetry," *Rev. Sci. Instrum.,* vol. 78, pp. 044304, 2007.

[8] H. Wu, M. Rubinstein, E. Shih, J. V. Guttag, F. Durand and W. T. Freeman, "Eulerian video magnification for revealing subtle changes in the world." *ACM Trans.Graph.,* vol. 31, pp. 65, 2012.

[9] J. Achten and A. E. Jeukendrup, "Heart rate monitoring," *Sports Medicine,* vol. 33, pp. 517-538, 2003.

[10] G. Balakrishnan, F. Durand and J. Guttag, "Detecting pulse from head motions in video," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference On,* 2013, pp. 3430-3437.

[11] PHILIPS. (2014). *Philips Vital Signs Camera* [Online]. Available: http://www.vitalsignscamera.com/.

[12] M. Poh, D. J. McDuff and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Optics Express,* vol. 18, pp. 10762-10774, 2010.

[13] M. Lewandowska, J. Ruminski, T. Kocejko and J. Nowak, "Measuring pulse rate with a webcam—a non-contact method for evaluating cardiac activity," in *Computer Science and Information Systems (FedCSIS), 2011 Federated Conference On,* 2011, pp. 405-410.

[14] G. D. Clifford and D. Clifton, "Wireless technology in disease management and medicine," *Annu. Rev. Med.,* vol. 63, pp. 479-492, 2012.

[15] F. Cardinaux, D. Bhowmik, C. Abhayaratne and M. S. Hawley, "Video based technology for ambient assisted living: A review of the literature," *Journal of Ambient Intelligence and Smart Environments,* vol. 3, pp. 253-269, 2011.

[16] I. Pavlidis, J. Dowdall, N. Sun, C. Puri, J. Fei and M. Garbey, "Interacting with human physiology," *Comput. Vision Image Understanding,* vol. 108, pp. 150-170, 2007.

[17] J. Anttonen and V. Surakka, "Emotions and heart rate while sitting on a chair," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* 2005, pp. 491-499.

[18] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision." in *IJCAI,* 1981, pp. 674-679.

[19] C. Takano and Y. Ohta, "Heart rate measurement based on a time-lapse image," *Med. Eng. Phys.,* vol. 29, pp. 853-857, 2007.

[20] M. L. Gleicher and F. Liu, "Re-cinematography: Improving the camerawork of casual video," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP),* vol. 5, pp. 2, 2008.

[21] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis and W. T. Freeman, "Removing camera shake from a single photograph," in *ACM Transactions on Graphics (TOG),* 2006, pp. 787-794.

[22] M. Fuchs, T. Chen, O. Wang, R. Raskar, H. Seidel and H. Lensch, "Real-time temporal shaping of high-speed video streams," *Comput. Graph.,* vol. 34, pp. 575-584, 2010.

[23] D. Murray and A. Basu, "Motion tracking with an active camera," *Pattern Analysis and Machine Intelligence, IEEE Transactions On,* vol. 16, pp. 449-459, 1994.

[24] B. Leibe, K. Schindler, N. Cornelis and L. Van Gool, "Coupled object detection and tracking from static cameras and moving vehicles," *Pattern Analysis and Machine Intelligence, IEEE Transactions On,* vol. 30, pp. 1683-1698, 2008.

[25] M. Yang, D. Kriegman and N. Ahuja, "Detecting faces in images: A survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions On,* vol. 24, pp. 34-58, 2002.

[26] MathWorks. (2015, Jan 25). ***Intensity-Based Automatic Image Registration*** [Online]. Available: http://www.mathworks.com/help/images/intensity-based-automatic-image-registration.html.

[27] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference On,* 2001, pp. I-511-I-518 vol. 1.

[28] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision,* vol. 57, pp. 137-154, 2004.

[29] H. Kruppa, M. Castrillon-Santana and B. Schiele, "Fast and robust face finding via local context," in *Joint IEEE Internacional Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS),* 2003, pp. 157-164.

[30] Y. Matsushita, E. Ofek, W. Ge, X. Tang and H. Shum, "Full-frame video stabilization with motion inpainting," *Pattern Analysis and Machine Intelligence, IEEE Transactions On,* vol. 28, pp. 1150-1163, 2006.

[31] L. M. Abdullah, N. M. Tahir and M. Samad, "Video stabilization based on point feature matching technique," in *Control and System Graduate Research Colloquium (ICSGRC), 2012 IEEE,* 2012, pp. 303-307.

[32] C. Harris and M. Stephens, "A combined corner and edge detector." in *Alvey Vision Conference,* 1988, pp. 50.

[33] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun ACM,* vol. 24, pp. 381-395, 1981.

[34] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Computer Vision–ECCV 2006*Anonymous Springer, 2006, pp. 430-443.

[35] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision,* vol. 60, pp. 91-110, 2004.

[36] A. Litvin, J. Konrad and W. C. Karl, "Probabilistic video stabilization using kalman filtering and mosaicing," in *Electronic Imaging 2003,* 2003, pp. 663-674.

[37] K. Lee, Y. Chuang, B. Chen and M. Ouhyoung, "Video stabilization using robust feature trajectories," in *Computer Vision, 2009 IEEE 12th International Conference On,* 2009, pp. 1397-1404.

[38] B. K. Horn and B. G. Schunck, "Determining optical flow," in *1981 Technical Symposium East,* 1981, pp. 319-331.

[39] R. J. Andrews and B. C. Lovell, "Color optical flow," in *Workshop on Digital Image Computing,* 2003, pp. 135-139.

[40] P. Kakumanu, S. Makrogiannis and N. Bourbakis, "A survey of skin-color modeling and detection methods," *Pattern Recognit,* vol. 40, pp. 1106-1122, 2007.

[41] K. Seo, W. Kim, C. Oh and J. Lee, "Face detection and facial feature extraction using color snake," in *Industrial Electronics, 2002. ISIE 2002. Proceedings of the 2002 IEEE International Symposium On,* 2002, pp. 457-462.

[42] A. Albiol, L. Torres and E. J. Delp, "Optimum color spaces for skin detection." in *ICIP (1),* 2001, pp. 122-124.

[43] G. Finlayson, C. Fredembach and M. S. Drew, "Detecting illumination in images," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference On,* 2007, pp. 1-8.

[44] L. I. Voicu, H. R. Myler and A. R. Weeks, "Practical considerations on color image enhancement using homomorphic filtering," *Journal of Electronic Imaging,* vol. 6, pp. 108-113, 1997.