HEART RATE MEASUREMENT USING COMPUTER VISION

A PROJECT REPORT

Submitted by

BHUVANESHWARI P (2019202008)

submitted to the faculty of

INFORMATION SCIENCE AND TECHNOLOGY

in partial fulfilment for the award of the degree of

MASTER OF COMPUTER APPLICATIONS

DEPARTMENT OF INFORMATION SCIENCE AND TECHNOLOGY COLLEGE OF ENGINEERING, GUINDY

ANNA UNIVERSITY

CHENNAI 600 025

MAY 2022

ANNA UNIVERSITY

CHENNAI - 600 025

BONA FIDE CERTIFICATE

Certified that this project report titled **"HEART RATE MEASUREMENT USING COMPUTER VISION"** is the bona fide work of **BHUVANESHWARI (2019202008)** who carried out project work under my supervision. Certified further that to the best of my knowledge and belief, the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or an award was conferred on an earlier occasionon this or any other candidate.

PLACE: CHENNAI

DATE: 02.05.2022

Dr. L. SAIRAMESH

PROJECT GUIDE,

DEPARTMENT OF IST,

COLLEGE OF ENGINEERING,

ANNA UNIVERSITY CHENNAI

600 025

COUNTERSIGNED

Dr. S.SRIDHAR

HEAD OF THE DEPARTMENT

DEPARTMENT OF INFORMATION SCIENCE AND TECHNOLOGY

COLLEGE OF ENGINEERING, GUINDY

ANNA UNIVERSITY

CHENNAI 600025

TABLE OF CONTENTS

ABSTRACT

The heart is the most important muscular organ of a human body and its strength can be assessed by the rate at which it beats. Heart rate measurement plays an important role in human health assessment and there have been a number of methods suggested in order to monitor it remotely with more ease and comfort. Contact-free heart rate measurement is one way to make it user-friendly and can also be used for covert surveillance. The performance of previously developed touch-free methods was found to be efficient and accurate under controlled conditions. Typically, the performance is dependent on controlled lighting and limited subject movement. More realistic situations require more robust contact-free ways to measure the heart rate. This work aims to obtain a good understanding of the underlying problems of non-contact pulse estimation. This proposal is to create a Realtime application for measuring the heart rate from human faces remotely using an ordinary webcam. The Haar Cascades algorithm from Computer Vision is used for face detection and also track the faces from the recording videos which provides more reliable extraction of a region of interest (ROI) than the simple face detection. The extraction of RGB code from video frames helps in extraction of a cleaner pulse signal, which is to be further denoised and detrended. After applying Fast Fourier Transform to the signal, Power Spectral Density signal obtained which will be finally helps in Heart Rate calculation.

CHAPTER 1

INTRODUCTION

1.1 MOTIVATION

The heart is the most important muscular organ of the human body that controls the flow of blood through all parts of the body. The heart rate assesses the strength and performance of a heart and the continuous monitoring of heart rate is vital for adequate human health treatment. Heart rate can be measured by different sensory equipment such as an ECG machine, pulse oximeter, and Doppler probe etc. The performance of all these types of machinery is quite satisfactory but the way in which they are used is quite uncomfortable for the patient and as well as doctors too.

Contact-free measurement of heart rate can ease the discomfort of subjects and make it user-friendly. Many different attempts have been made in the literature to measure the heart rate contact-free using a thermal camera and/or a webcam. All methods have reliable performance in controlled conditions, but the accuracy remains a question mark if they are exposed to more challenging and realistic capturing scenarios. The subject’s motion artifacts and illumination variations can be expected in more challenging conditions, which require a robust approach to measure the pulse rate accurately and extract the pulse signal cleanly. Therefore, there is still lots of room to improve the existing methods with respect to robustness, automation, accuracy, algorithm efficiency, and the characteristics of existing methods for specific applications still can be more thoroughly investigated. The two main problems that affect the performance of contact-free heart rate measurement are subject’s movement and illumination variations. The extracted pulse signal, if obtained cleanly from the subject’s facial imagery, could have several useful applications such as human biometric recognition, lie or stress detection, sports analysis, etc.

1.2 PROBLEM STATEMENT

Traditional way of measurement of heart beat method could take lot of time consuming and in convincing of physical touching. In pandemic situations like spread of virus, we might not wish to touch a patient to find the heart rate and doing electrocardiography. So, if this happened remotely, we will be convenient to measure the heart rate i.e., blood volume pulse. This system requires only the facial video from the person who is going to be observed, from that we can able to estimate the final heart rate.

1.3 AIM AND OBJECTIVES

This project aims to investigate contact-free heart rate measurement methods, more specifically, the main problems of subject's motion and illumination noise in estimating heart rate from human face videos. We try to obtain a good understanding of these problems, the existing methods, their performance under realistic and more challenging conditions, and the possible applications of the contact-free pulse extraction from facial videos. We look into the most popular and/or the most promising methods, identify their limitations and propose possible improvements.

The main objectives of our research can be listed as follows:

- Conducting a thorough literature review of existing non-contact heart rate measurement methods, and the biometric recognition application of the extracted pulse signals. This implementation has the most promising methods for the extraction of heart rate from facial videos using an ordinary webcam in a more challenging capturing scenario.

- Improving the existing methods for contact-free heart rate measurement by addressing their limitations and introducing additional parameters to overcome these weaknesses. The main problems of almost all the previous methods are subject's movement and illumination variations. Haar Cascade is used as a face tracking technique, detrending, and robust mean color pixels algorithm into non-contact heart rate measurement system to minimize the effect of these problems and to obtain an improved accuracy of the estimated heart rate.

- Developing a Realtime application to achieve robustness in the non-contact pulse rate measurement system under more challenging and realistic capturing scenarios. This proposed method can overcome many of the problems of light reflection and subject's movements while measuring the heart rate from human faces remotely using an ordinary webcam. This application can not only improve the efficiency of heart rate measurement systems but can also result in extraction of a cleaner pulse signal, which can have many other useful applications such as human biometric recognition.

CHAPTER 2

LITERATURE REVIEW

This chapter provides the background of the heart rate measurement systems for human subjects. A wide range of heart rate estimation technologies and the terminologies used in them are introduced. Existing challenges in the area of non-contact heart rate measurement from human face videos are highlighted. Finally, the Realtime application for the contact-free pulse signal from facial videos is also discussed.

2.1 Vital Signs

The physiological measurements that assess the basic functions of a human body and are vital for the human health monitoring are called vital signs [1]. They are good indicators of the general physical health of a person, which help diagnose the possible diseases and trace the progress towards recovery. There are four main vital signs which are generally monitored by a healthcare professional which includes: body temperature, respiration rate, blood pressure, and pulse rate or heart rate.

2.1.1 Heart

The heart is a muscular organ that lies in the centre of our chest. It pumps oxygenated blood to all parts of the body and receives less oxygenated blood from the body to pass on to the lungs to absorb more oxygen and supply it back to the body. The structure of the heart comprises four main chambers. The two chambers on the left side are called the left Atrium and the left Ventricle, which supply oxygenated blood to the body through the Aorta. The other two on the right side are the right Atrium and the right Ventricle, which receive less oxygenated blood through the Ven Cava and pass it to the lungs to absorb oxygen.

These chambers are connected by valves, which ensure the one-way flow of blood through them. The veins carry the blood towards the heart and arteries take the blood away from it. The flow of blood through the heart does not happen automatically but it is driven by contractions of the heart muscle which are triggered by electrical impulses. The journey of the blood from entering the heart and leaving it is called one cardiac cycle or pulse and the heart is said to beat once [2].

2.1.2 Heart Rate

The Heart Rate (HR) refers to the number of times a heart beats in one minute. The normal resting heart rate of a healthy person ranges between 60 to 100 beats per minute (bpm). As the heart is a muscle and like all other muscles, it can also strengthen by doing exercise. Generally, a lower heart rate at rest implies more efficient heart function and better cardiovascular fitness. For example, a well-trained athlete might have a normal resting heart rate closer to 40 beats per minute. The maximum heart rate of a person is highly correlated to the age of a person [3].

2.2 HEART RATE MEASUREMENTS METHODS

The continuous monitoring of heart rate is very important to assess human health. Several different methods are currently in practice to measure the heart rate. The most common methods in practice are described in this section.

2.2.1 Phonocardiogram (PCG)

PCG (Phonocardiogram) is a heart rate measurement technique that detects and register the heart sounds which can be visually represented as a graphical chart. The heart sounds are produced due to several cardiac activities which include; opening and closing of four values of a heart, the flow of blood through these values, the blood flow through the ventricular chambers, and rubbing of heart surfaces [4]. The generated heart sound is referred to as "lub-dub".

2.2.2 Pulse Oximetry

Pulse oximetry[5] is a non-invasive method to measure the oxygen saturation level of blood in our body and estimates the heart rate. It is widely used due to its easy and relatively comfortable setup. A pulse oximeter is a small microprocessor unit comprising a peripheral probe and a display monitor. The probe which consists of a photodetector, and two light emitting diodes can be clipped to a finger, earlobe or foot of a subject, and the estimated heart rate can be displayed on the monitor screen.

2.2.3 Electrocardiogram (ECG)

ECG is the current gold standard for the accurate measurement of heart rate. As mentioned earlier, the flow of blood through the heart does not happen automatically but is driven by contractions of the heart muscle which are triggered by electrical impulses. This electrical activity of the heart with respect to each heartbeat is measured by ECG. Electrodes are placed at different locations of the subject’s body such as limbs and chest to monitor such electrical activity over a period of time, and the output can be seen in the form of an electrical signal display on the monitor or printed on the grid paper.

2.3 NON-CONTACT HEART RATE MEASUREMENT

The heart rate measurement methods are accurate and efficient but there are certain limitations associated with them and they are quite uncomfortable with the subject. It is therefore required to find contact-free ways for the estimation of the heart rate and it has been an interesting area for the researchers to work on over the past decade. In this section, the most popular and/or the most promising contact-free heart rate measurement methods, and their performance accuracy are discussed, and their limitations are identified.

Photoplethysmography (PPG)[6], is the main idea of most of the previous work related to the non-contact heart rate measurement. The term photoplethysmography is composed of two words ‘photo’ and ‘plethysmography’. Plethysmography refers to a Greek word ‘plethymos’ which means increase, and describes the measurement of volumetric changes in different parts of the body. These changes are caused by the change of blood volume triggered as a result of a specific event occurring in a body. The word ’photo’ refers to the use of light for plethysmography. This phenomenon states that when a heart beats, it produces a pulse wave called a cardiovascular pulse or the blood volume pulse (BVP), which travels through the whole body’s vascular system. When this wave reaches the face, it leaves a small variation in the volume of blood, which causes the skin color to change. The naked eye cannot view such subtle changes but we can measure the intensity of the absorbed or reflected light by means of images taken from a camera sensor. The cameras used in the previous work related to pulse extraction from videos are the thermal camera, CCD camera, and a simple computer webcam. The face of the subject is focussed by the sensors of each of these cameras.

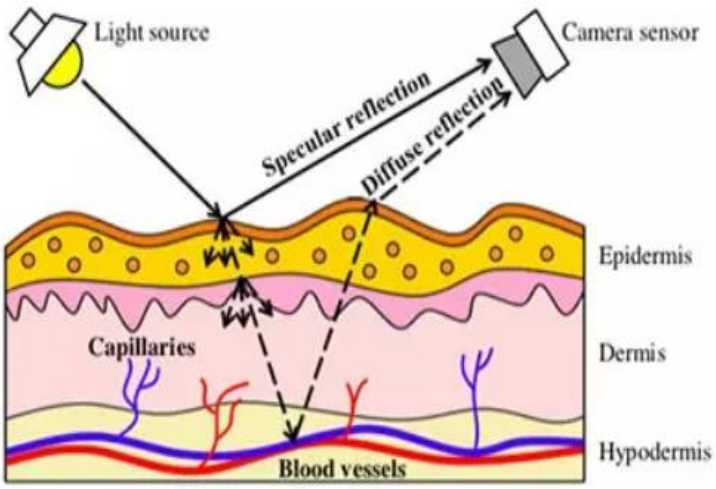


Fig. 2.1 An overview of photoplethysmography

2.3.1 Non-Contact Heart Rate Measurement Using Thermal Camera

Non-contact heart rate measurement was first introduced by Pavlidis in 2003[7], and later demonstrated by analysing thermal videos of the front face [8]. They uncover the information of blood flow, cardiac pulse, and breath rate from a bioheat modelling of facial imagery using a sensitive thermal camera. Thermal imaging is a passive modality, as it depends on the heat radiations emitted from different parts of the subject's body. Fig. 2.1 shows an overview of their pulse measurement methodology. The face was targeted as the region of interest (ROI) because it is the most sensitive and exposed part of a body. The ROI was further narrowed down to the superficial blood vessel, which they assumed to be the best-suited area for the extraction of the strongest thermal signal. The face is always in constant motion and a robust tracking technique is required to accommodate such movements. They proposed a tandem tracker (TT) to track the face from the thermal image and the measurement tracker (MT) to detect a specific vessel tissue within the tracked face. Fourier transform is applied to the extracted thermal signal to obtain dominant pulse frequency. In order to minimize the effect of noise from the pulse signal, adaptive estimation function is used, which takes into account a specific threshold in each successive heart rate measurement [8]. The foundation for contact-free heart rate measurement was laid down as they reported a good accuracy of 88% of their novel methodology but still the need for improvement in the face tracking was addressed by them. The dataset was collected in a controlled lighting condition and little subject movement was permitted, indicating that more noise could be expected when their method is applied in realistic capturing scenarios, which could affect the performance. The use of an expensive thermal camera for the data acquisition also raises concerns.

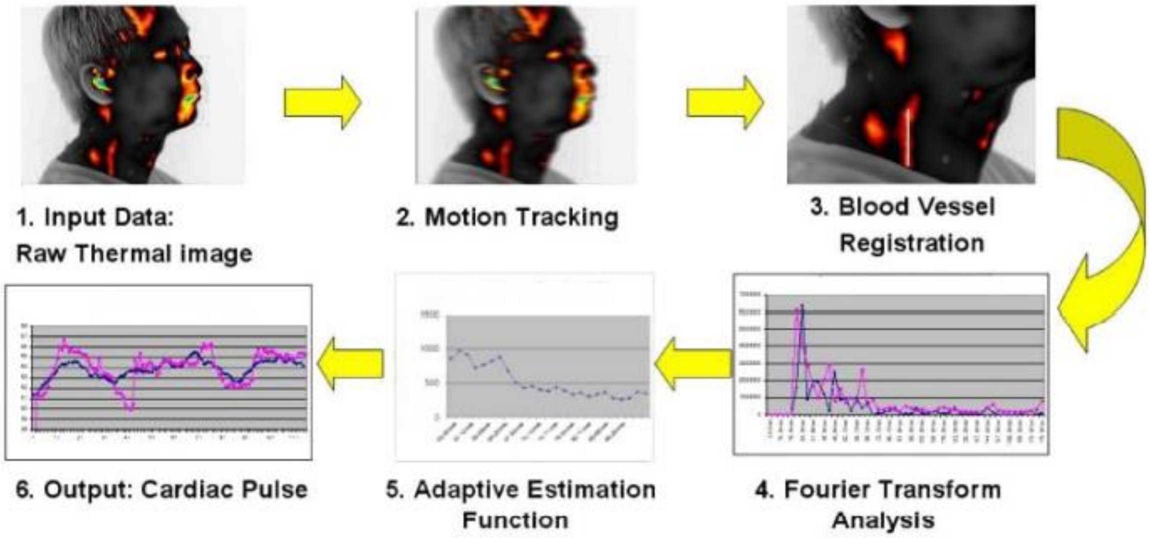


Fig. 2.2 An overview of non-contact heart rate estimation using a thermal camera

2.3.2 Non-Contact Heart Rate Measurement Using CCD Camera

Takano and Ohta [9] measured the pulse rate using 30s (s means seconds) time-lapse images captured by means of a Charge Coupled Device (CCD) camera. CCD is a small circuit etched on to a silicone surface embedded in the digital cameras to transfer electrical charge to the memory in the form of a digital value. They analysed the variations in the average brightness of a subject's facial skin images. An ROI of 30×40 pixels was extracted from the subject's cheek images and a series of processing steps were performed on the average brightness of each image, which included a first order derivative, a low pass filter of 2 Hz, and Auto-Regressive spectral analysis. The heart rate peak was chosen as the highest peak in the auto-regressive spectrum of the processed signal. The results show high accuracy in their experimental setup with the controlled illumination settings but in realistic scenarios, where rapid light variations could be expected, the performance may degrade. Such noise elimination must be considered as a challenge. The size and cost of a CCD camera also make this method less applicable to the daily life activities. There are problems of subject's motion and illumination noise, while measuring the heart rate using a CCD, and presented a chrominance-based remote pulse rate method to overcome such problems. The RGB color channels were extracted from the ROI and normalized by dividing their samples by their mean over a specific period of time, which must cover at least one heartbeat time interval. The combination of red and green color channels made a pulse signal. This worked well enough for subject's motion noise but the illumination variation was still present in the signal. The specular reflection component, which was directly reflected from the skin surface, was eliminated by using chrominance (color difference) signals. They proposed different algorithms by combining different proportions of RGB color channels and were tested on stationary and moving subjects. The results show better accuracy and signal to noise (SNR) ratio when compared to other state-of-the-art heart rate estimation methods. The use of CCD camera makes this method non-affordable to be implemented for the real-life applications.

2.3.3 Non-Contact Heart Rate Measurement Using an Ordinary Webcam

The performance of estimating the heart rate from human face videos using a thermal camera and a CCD camera was found quite accurate under controlled conditions. The performance could be affected by the subject movement and illumination variations, and the resulting noise in the extracted pulse signal can be considered as the main problem in the accurate measurement of heart rate from facial imagery. As the CCD and thermal cameras can be very expensive, in the past few years, researchers have reported on the estimation of pulse rate using an inexpensive computer webcam [10].

2.4 Research methodology

A clear consensus concerning the name of the field that is discussed in this study has yet to emerge. While researching, one could come across 15 different terms used by different authors. These typically employ lexical combinations that begin with such words as “remote”, “non-contact”, “camera-based”, “video-based”, “contactless”, “contact-free”, “imaging” and end with such terms as “photoplethysmography”, “heart rate measurement”, “heart rate estimation”, “heart rate monitoring”, as well as various abbreviations thereof. The term “remote photoplethysmography” (rPPG) is used in this proposal because it is by far the most frequently used and is an original name for this class of algorithms [11].

2.5 Background

Phenomena exploited in rPPG are closely related to the cardiac cycle. During each cycle, blood is moved from the heart to the head through the carotid arteries. It can be seen that this periodic inflow of blood affects both the optical properties of facial skin and the mechanical movement of the head, enabling researchers to measure HR remotely. The interplay of light and living tissue is complex, as many processes such as scattering, absorption, and reflection are at play. Research has shown that reflection of light is dependent on, among other factors, blood volume change and blood vessel wall movement[11]. Given suitable illumination, changes in light reflected from facial skin are thus observable, as the blood flow and variation of blood volume follow the cardiac cycle. Traditionally, dedicated light sources with red or near-infrared wavelengths have been used to obtain a (contact) photoplethysmogram. However, recent research has shown that ambient light can be sufficient to obtain a plethysmographic signal (as illustrated in Fig. 2.3).

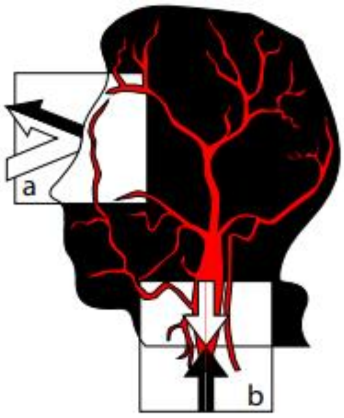


Fig. 2.3 Illustration of phenomena used in rPPG

More recently, some research has focused on remotely capturing the mechanical impact of blood flowing in through the carotid arteries at either side of the head. This approach considers the head-neck system and the trunk as a sequence of stacked inverted pendulums and surmises that the opposite reaction to blood inflow causes a displacement of the head by approximately 5 mm. Of the two approaches, that based on skin color variation, being the original, has been discussed in many more studies [11].

2.6 Early work and recent development

Hertzman and Spealman first noted in 1937 that the variation in light transmission of a finger could be detected by a photoelectric cell. The formative period of rPPG research began in 2008 with Verkruysse [11] and colleagues first showing that video recordings of a subject's face under ambient light contain a signal sufficiently rich to measure the HR. They asked volunteers to sit motionless while their faces were recorded using inexpensive consumer cameras from a distance of 1-2 m. Fig. 2.4 illustrates the typical setup of such studies.

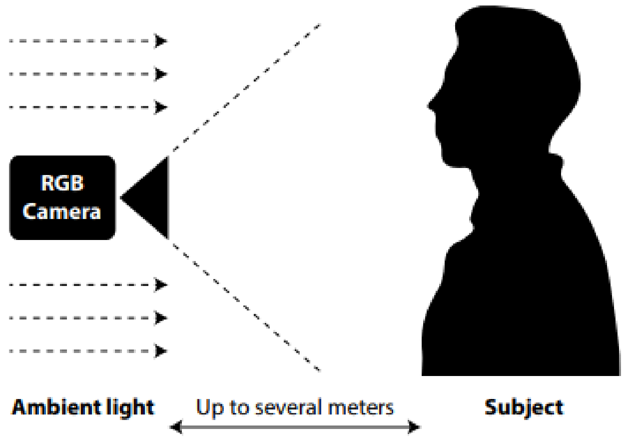


Fig. 2.4 Typical setup of an rPPG application

Verkruysse at al.[11] used color recordings of different quality. For example, a resolution of 640 x 480, which is a standard graphic mode of the video graphics array (VGA), and a frame rate of 30 frames per second (fps) were used. In these recordings, the region of interest (ROI) was manually selected. From the pixels contained in the ROI, the raw signal was computed per frame as the mean value of each of the RGB color channels. To determine power spectral density of the signal, Verkruysse et al. used the fast Fourier transform (FFT) algorithm. They showed that the signal for the green channel contains the strongest plethysmographic signal, clearly indicating the fundamental HR frequency, up to its fourth harmonic. This is consistent with the fact that haemoglobin absorbs green light better than it does red and blue.

Paving the way for future research was the first study with the explicit goal of measuring HR using video recorded with a standard laptop webcam [13]. This study by Poh et al. used a face detector to track a subject's face frame by frame, with a box containing the subject's face as the ROI and a moving window of 30 s to achieve a continuous measurement. Improving on this approach in, all three channels of RGB information were used. Blind source separation (BSS) estimated the plethysmographic signal as a linear combination of all three raw signals. The parameters for this combination were estimated using independent component analysis (ICA). However, Poh et al. always chose the second component produced by ICA as the plethysmographic signal, a shortcoming they later addressed in an improved version of their algorithm [14]. The HR was estimated as the frequency with the highest response after an FFT.

With the general feasibility of rPPG being established, an increasing number of publications have been produced in subsequent years. Initial contributions include comparing alternative methods for BSS and different selections regarding color channels, as well as adding temporal filters before BSS is performed. An approach using an ROI and neural-network-based skin detection proposed allows for more accurate measurement. Another study compared various linear and nonlinear techniques for BSS and found that Laplacian eigenmap produces the best results.

The plethysmographic signal in a subject’s face can be visualized by decomposing the video sequence into different spatial frequency bands and then magnifying a desired frequency band using bandpass filtering. When this process is applied to facial videos, slight temporal changes are detectable. This shows that HR and individual heart beats can be extracted from the amplified signal.

As other researchers [15] have found, the choice of ROI has a major influence on the quality of the plethysmographic signal, as not all areas in the face exhibit the same signal quality. The most recent studies have focused on more intelligent ROI selection and tracking to achieve motion robustness. Detection of facial landmark points is typically the basis for a more detailed ROI. Further reductions in noise were made possible by the so-called adaptive bandpass filter adopted by some authors, the cut-off frequencies for which were based on past HR estimates. Custom additional filtering steps introduced by some authors also aimed at reducing noise (e.g., used an adaptive filter to reduce noise from illumination changes using background illumination as a reference).

Further recent developments include variations in the number of used raw signals, such as the inclusion of cyan and orange frequencies. In a different approach, the facial region was divided into many small ROIs that yielded an array of signals from the green channel, each of which was later combined using a weighted average based on a goodness metric. Similarly, the researchers stochastically selected an array of points and combined them using an importance-weighted Monte Carlo approach. The use of BSS, followed by component selection, have recently been optimized using machine learning techniques [17].

Despite the fact that these recent improvements allow rPPG algorithms to be applied to more realistic situations, virtually all studies have continued to focus on proof of concept using pre-recorded videos. Although one study presented concepts for real-time applications, only one other work reported data from a real-time rPPG application. Signal-to-noise ratios and error rates have typically been reported, but comparing different approaches is difficult. This is because most authors have tended to create their own test scenarios using a variety of cameras and often have not specified the algorithms used for compression, thus making reproduction difficult. An exception to this is a study that benchmarked rPPG algorithms using videos from a publicly available database. However, no consistent practice has yet been adopted.

CHAPTER 3

DESIGN

In this section, the proposed experiment is explained in detail. First, the proposed method is described thoroughly. The system consists of three main steps—ROI detection, signal processing, and heart rate estimation—where each main step is divided into a handful of smaller steps. An overview of the system architecture can be seen in Figure 3.1. Subsequently, the experimental setup is described including the live video setup and how HR was measured. Finally, it is explained how the performance was measured.

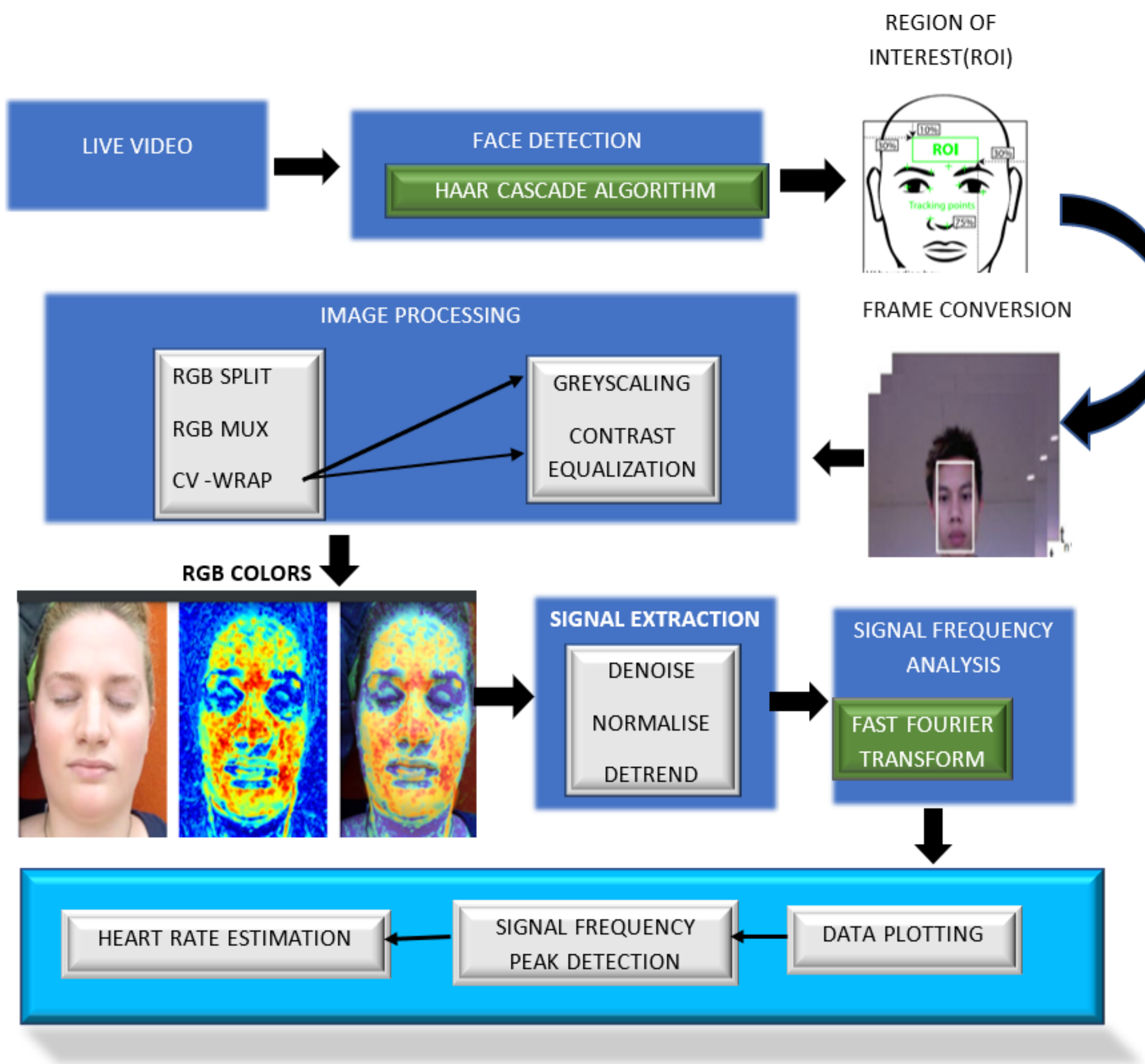


Fig. 3.1 A schematic diagram of the system architecture.

3.1 Photoplethysmography

Photoplethysmography (PPG) is a simple technique used to detect volumetric changes in different parts of the body by examining the color of the skin. Most of these changes happen because of fluctuations in blood volume in the organ, and photoplethysmography can thus be seen as a measurement of the pulse wave traveling through the body.

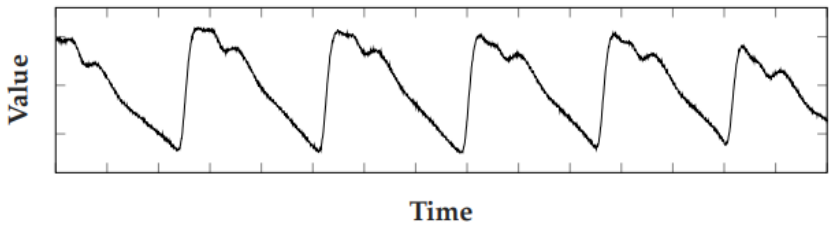


Fig. 3.2 An example of a PPG signal.

An example of a PPG signal can be seen in Fig. 3.2. By studying the time between two consecutive peaks, Δt, the heart rate (in bpm) can be calculated as 60/Δt.

3.2 HAAR CASCADE ALGORITHM

*F**ace Detection**, a widely popular subject with a huge range of applications. Modern day Smartphones and Laptops come with in-built face detection software, which can authenticate the identity of the user. There are numerous apps that can capture, detect and process a face in real time, can identify the age and the gender of the user, and also can apply some really cool filters. The list is not limited to these mobile apps, as Face Detection also has a wide range of applications in Surveillance, Security and Biometrics as well. But the origin of its Success stories dates back to* 2001, *when* Viola and Jones *proposed the first ever Object Detection Framework for Real Time Face Detection in Video Footage.*

It is an Object Detection Algorithm used to identify faces in an image or a real time video. The algorithm uses edge or line detection features proposed by Viola and Jones in their research paper “Rapid Object Detection using a Boosted

Cascade of Simple Features" published in 2001. The algorithm is given a lot of positive images consisting of faces, and a lot of negative images not consisting of any face to train on them.

3.3 FAST FOURIER TRANSFORM

The Fast Fourier Transform (FFT) is a computationally efficient method of generating a Fourier transform. The main advantage of an FFT is speed, which it gets by decreasing the number of calculations needed to analyze a waveform. FFT is an efficient algorithm that is used for converting a time-domain signal into an equivalent frequency-domain signal, based on the discrete Fourier transform (DFT). Several real-time programming examples on FFT are included.

The DFT converts a time-domain sequence into an equivalent frequency-domain sequence. The inverse DFT performs the reverse operation and converts a frequency-domain sequence into an equivalent time-domain sequence. The FFT is a very efficient algorithm technique based on the DFT but with fewer computations required. The FFT is one of the most commonly used operations in digital signal processing to provide a frequency spectrum analysis. Two different procedures are introduced to compute an FFT: the decimation-in-frequency and the decimation-in-time. Several variants of the FFT have been used, such as the Winograd transform, the discrete cosine transform (DCT), and the discrete Hartley transform.

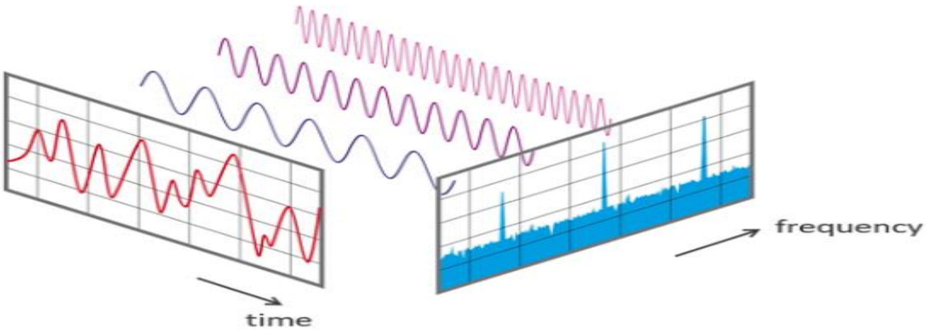


Fig 3.3 Fast Fourier Transform

CHAPTER 4

IMPLEMENTATION

4.1 Computer Vision

Computer vision is a process by which we can understand the images and videos how they are stored and how we can manipulate and retrieve data from them. Computer Vision is the base or mostly used for Artificial Intelligence. Computer-Vision is playing a major role in self-driving cars, robotics as well as in photo correction apps.

4.1.1 OpenCV

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features. It has C++, C, Python and Java interfaces and supports Windows, Linux, Mac OS, iOS and Android. When OpenCV was designed the main focus was real-time applications for computational efficiency.

OpenCV Functionality:

- Image/video I/O, processing, display (core, imgproc, highgui)
- Object/feature detection (objdetect, features2d, nonfree)
- Geometry-based monocular or stereo computer vision (calib3d, stitching, videostab)
- Computational photography (photo, video, superres)
- Machine learning & clustering (ml, Flann)

- CUDA acceleration (GPU)

4.1.2 OpenMDAO:

OpenMDAO is an open-source framework for multidisciplinary design, analysis, and optimization. This is a high-performance computing platform for systems analysis and multidisciplinary optimization written in the python programming language.

The OpenMDAO project is primarily focused on supporting gradient-based optimization with analytic derivatives to allow you to explore large design spaces with hundreds or thousands of design variables, but the framework also has a number of parallel computing features that can work with gradient-free optimization, mixed-integer nonlinear programming, and traditional design space exploration.

The OpenMDAO framework is designed to aid in linking together separate pieces of software for ***the purpose of combined analyses***. It allows users to combine analysis tools (or design codes) from multiple disciplines, at multiple levels of fidelity, and to manage the interaction between them. OpenMDAO is specifically designed to manage the dataflow (the actual data) and the workflow (what code is run when) in conjunction with optimization algorithms and other advanced solution techniques.

4.2 Face detection and tracking

A face tracking algorithm typically contains two steps. First, the face has to be located in the image, and then some feature points in it have to be followed. Here, two of the most commonly used methods to achieve this are described.

4.2.1 Detection with Viola–Jones Algorithm

One of the most known and commonly used face detection algorithms is Viola–Jones from OpenCV. The idea behind the algorithm is simple: a window is slid over the image, in which a classifier decides whether it contains an object (in this case a face) or not. The classifier, which consists of several simpler classifiers that are applied successively until a candidate is rejected or all the stages are passed, is trained on a set of face and non-face images. Haar-like features, of which examples can be seen in Figure 4.1, are used to make the decisions. To be able to find faces of not only one size, the classifier is designed so it can be resized and slid over the image repeatedly in different sizes.

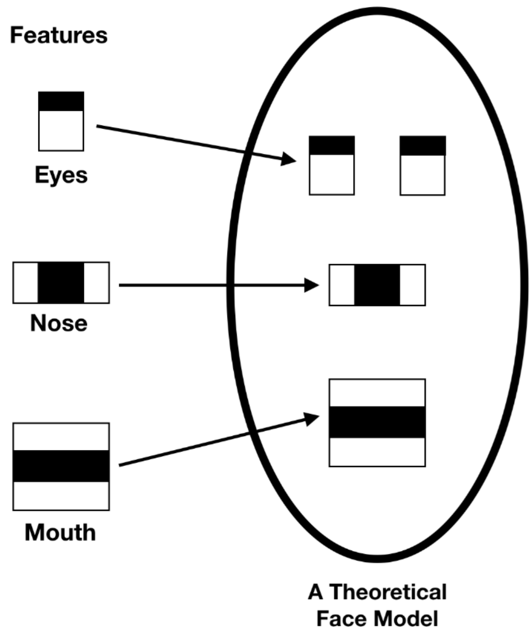


Fig. 4.1 Haar Cascades features

Once a face has been detected in the image, the known position can be used in the next frame to speed up the algorithm and to track the face. This is simply done by only searching for a face in regions close to the previous position. If a face is not found, the algorithm should preferably check the entire frame for a face.

4.2.2 Region Of Interest (ROI) Detection

Because the rPPG algorithms consider are based on the human face, ROI detection is necessary to determine the bounds of the face in a video frame. This information is typically an intermediate step from which a more accurate ROI is defined. In some of the literature, especially in earlier studies in which the subject was asked to sit motionless, the bounds of the face were selected manually from one of the first frames. The most frequently used method is the algorithm of Viola and Jones, which is a machine-learning-based approach that uses a cascade of simple features to classify faces. The popularity of this approach is partially due to its availability in the OpenCV computer vision library, which many authors have used to implement their rPPG algorithms. A bounding box of the face is returned when using the Viola-Jones algorithm. Recent approaches dealing with subject motion require more detailed information about face location. Facial landmark points provide the basis for more detailed ROI definitions as well as ROI tracking. Noise caused by subject motion may render the signal useless for rPPG. Thus, the goal of ROI tracking is to ensure that the pixels contained in the ROI belong to a skin region invariant to subject motion.

4.2.3 Frame conversion

Frame conversion is the process of converting the video into number of frames. This is done by CV2 from OpenCV Library. Processing a video means, performing operations on the video frame by frame. Frames are nothing but just the particular instance of the video in a single point of time. We may have multiple frames even in a single second. Frames can be treated as similar to an image. So, whatever operations we can perform on images can be performed on frames as well.

4.3 IMAGE PROCESSING

4.3.1 Color spaces

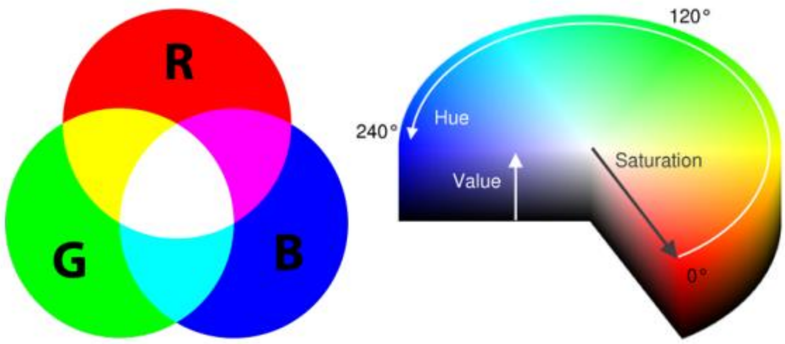
There are different models for representing colors. Two of the most used models are RGB and HSV, which are used in this thesis.

(i) RGB

RGB is a very commonly used color representation model, having three parameters—R (red), G (green), and B (blue)—each of them with a value in range [0.0, 1.0]. It is an additive color model, which means that the resulting color is the one where the parameters are mixed. For example, black is represented by the triplet (0, 0, 0), blue by (0, 0, 1) and white by (1, 1, 1). The RGB color model is visualized in Figure 4.2a and is used throughout this project.

(ii) HSV

HSV (hue, saturation, and value) is a color model that represents an RGB color in a cylindrical coordinate system, to make the representation more intuitive. The hue parameter is the angular dimension and represents the color with red at 0◦, green at 120◦ and blue at 240◦. Saturation, ranging from zero to one, states how saturated the color is. Finally, value describes the brightness in a 0–1 range. A visual explanation of the HSV color model can be seen in Figure 4.2b.



(a) The RGB color model. (b) The HSV color model.

Figure 4.2: Visualizations of the RGB and the HSV color models.

The color information of an object, which is the information of interest, is found solely in the hue parameter. This parameter does not depend on the brightness of the object and should be invariant to light changes. In RGB, this information is distributed across all three primary colors.

4.3.2 RGB Extraction

Across all pixels in the ROI, the averages for each colour channel (R, G, B) are computed and tracked (concatenated) to create three colour signals.

RGB Split: Extract the red, green, and blue channels from an (n, m,3) shaped array representing a single image frame with RGB color coding. This is straightforward NumPy slicing operation.

RGB Mux: Take three (m, n) matrices of equal size and combine them into a single RGB-coded color frame.

CV Wrapper: Generic wrapper to take the simpler functions from the cv2 or SciPy image libraries to generate connectable openMDAO components for image processing. Other functionality (like object detection, frame annotation, etc) should probably be wrapped individually. This wrapper is used for following two modules:

- **Grayscale:** This is the process of converting an image from other color spaces e.g., RGB, CMYK, HSV, etc. to shades of Gray. It varies between complete black and complete white. Turn (m, n,3) shaped RGB image frame to a (m, n) frame that discards color information to produce simple image matrix.
- **Contrast Equalization:** Automatic contrast correction. This Only works for grayscale images.

REFERENCES:

1. David L. Schriger. Cecil Medicine, 24th ed. Elsevier, 2007.
2. New Health Advisor. Diagram of Human Heart and Blood Circulation in It. http://www.newhealthadvisor.com/Heart-Diagram-Labeled.html. [accessed March 10, 2017].
3. Hirofumi Tanaka, Kevin D. Monahan, and Douglas R. Seals. Age-predicted maximal heart rate revisited. Journal of the American College of Cardiology, 37(1):153–156, 2001.
4. Tamal Chakrabarti, Sourav Saha, Sathi Roy, and Ishita Chel. Phonocardiogram signal analysis - practices, trends and challenges: A critical review. 2015 International Conference and Workshop on Computing and Communication (IEMCON), 2015.
5. Wikipedia. Pulse Oximetry. https://en.wikipedia.org/wiki/Pulse_oximetry.
6. Wim Verkruysse, Lars O. Svaasand, and J. Stuart Nelson. Remote plethysmographic imaging using ambient light. Optics Express, 16(26):21434–21445, 2008
7. Ioannis Pavlidis. Continuous physiological monitoring. Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE, pages 1084–1087, 2003.
8. Marc Garbey, Nanfei Sun, Arcangelo Merla, and Ioannis Pavlidis. Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. IEEE Transactions on Biomedical Engineering, 54(8):1418–1426, 2007.
9. Chihiro Takano and Yuji Ohta. Heart rate measurement based on a time-lapse image. Medical Engineering and Physics, 29(8):853–857, 2007.
10. Sungjun Kwon, Hyunseok Kim, and Kwang Suk Park. Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),, pages 2174–2177, 2012.

11. Verkruysse W, Svaasand LO, Nelson JS. Remote plethysmographic imaging using ambient light. Optics Express, 2008, 16(26): 21434–21445
12. Allen J. Photoplethysmography and its application in clinical physiological measurement. Physiological measurement, 2007, 28(3): R1–R39
13. Poh M-Z, McDuff DJ, Picard RW. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Optics Express, 2010, 18(10): 10762–10774
14. Poh M-Z, McDuff DJ, Picard RW. Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Transactions on Biomedical Engineering, 2011, 58(1): 1–4
15. Lempe G, Zaunseder S, Wirthgen T, Zipser S, Malberg H. Roi selection for remote photoplethysmography. In: Meinzer H-P, Deserno MT, Handels H, Tolxdorff T, eds. Informatik aktuell. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, 99–103
16. Chwyl B, Chung AG, Deglint J, Wong A, Clausi DA. Remote heart rate measurement through broadband video via stochastic bayesian estimation. Vision Letters, 2015, 1(1): 5