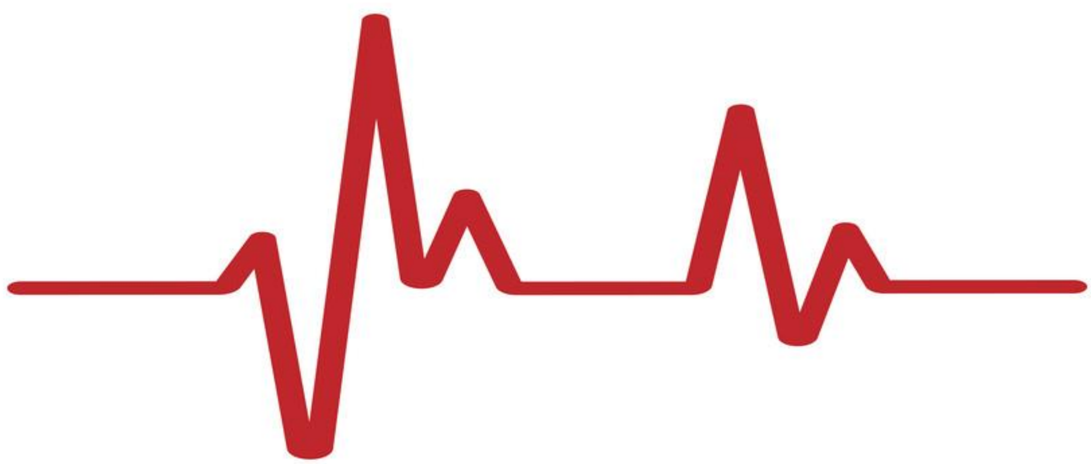
Helwan University

Faculty of Engineering

Department of Communications and Electronics

Videoscope



July 2021

Under the Supervision of Dr. Samir Gaber

Students' names and authors

Mohamed Hussein Mohamed Ibrahim
Mohamed Ayman Abdel Sabour Abdul Hamed
Mohamed Gouda Ahmed Gouda
Viola Laqi Luqa Dous
Amira Amr Mohamed Hassan Abd Al Rasoul

Content

List of figures

List of tables

1. Abstract

Video magnification reveals subtle variations that would be otherwise invisible to the naked eye. Current techniques that make be the changes in color be visible to us

Eulerian Video Magnification is a method, recently presented at SIGGRAPH 2012, capable of revealing temporal variations in videos that are impossible to see with the naked eye. Using this Method

it is possible to visualize the flow of blood as it fills the face And to assess The main field of this research work is image processing and computer vision, whose main The purpose is to translate dimensional data from the real world in the form of images into numerical or symbolic information.

Other fields include medical applications, software development for mobile devices, digital signal processing.

2.Introduction

Helwan University – Faculty of Engineering is aiming to be one of the best colleges known for their leadership regionally and internationally in engineering education and scientific research through interdisciplinary and unique academic programmer that meet the needs of the community and contribute to sustainable development.

It aims for preparation of distinguished graduates capable of keeping pace with global technology in various disciplines that meet the needs of local and regional markets and can conduct scientific research.

This is applied through the creation of appropriate conditions for faculty members and their assistants and students, and through providing educational programmers in advanced undergraduate studies, as well as establishing advisory centers and research labs which include sophisticated contributions to community service and to meet its needs.

Engineering Programmers at the Faculty of Engineering -Helwan University is one of the outstanding models for engineering education in Arab Republic of Egypt, as it seeks to provide high-quality engineering education based on interdisciplinary programmers and the application of international standards of mainstream systems followed in the most prestigious universities in the world.

Learning environment at Helwan university focused on the graduation engineers equipped with skills, knowledge, and the ability to life-long learning.

3.The project

Aiming to help in the medical deficiency happening due to corona pandemic; to minimize its effect on those who need follow ups remotely with their specialists, making a user-friendly application easy to use without the need to a medical background to know their vital rates.

The purpose of our project is that we can help anyone from a distance. Neither he needs to find a doctor, nor he goes to one, so actually it was inspired from the Corona pandemic as we all were quarantined in our homes so it was so much hard and very dangerous to go outside.

So, we shall be doing some project to help anyone who is facing the same problem if he cannot go outside of his home, as well we can help anyone who needs urgent help.

For example, if we found someone just had had an accident so with the help of a smartphone, we can just capture him a video and then we can tell what he has and what we should know, as our project can inform the person the heart rate of the person who was captured a video of his face.

It can be done by detecting the small changes that take place in our face and entire body but cannot be noticed by our eyes, and we shall see in the upcoming chapters everything in detail.

3.1 Technical details

This will focus on the heart rate estimation from a person's face captured through a simple webcam or mobile cam. User will load a video which will be 10-15 sec to the application, algorithm will work in the video frames, to work on the frames it is needed to deal with the pixels frequency so the algorithm will start with image pyramids "Gaussian & Laplacian pyramid decomposition technique" to decompose the video sequence according to spatial frequency, which is followed by temporal filtering then we will multiply this frequencies by a factor this factor is known as magnification factor and added back to the signals that are entered to the temporal filter. Then we will reconstruct the frame with the added frequencies and get the changes in RGB then heart rate will be calculated.

Our project consists of 2 steps:

1. Algorithm

2. User interface

- Website
- Mobile application

3.2 Project aims and objectives

As we mentioned before this program has been made specifically to help anyone who need any kind of medical help especially the ones who have had an accident and waiting for an ambulance to come save them, so in this time of waiting they can use our program to know what to do to help themselves while the help from a professional come.

In this work, an Android application for monitoring vital signs based on the Eulerian Video Magnification method will be developed, which should include the following features:

- Display the magnified blood flow, obtained from the Eulerian Video Magnification method
- Heart rate detection and assessment based on the Eulerian Video Magnification method

It should be noted that a straightforward implementation of the Eulerian Video Magnification method is not possible, due to various reasons. First, the Eulerian Video Magnification method provides motion magnification along with color magnification which will introduce several problems with artifacts' motion. Second, the requirement of implementing a real-time smartphone

4.The idea of the Video Magnification

Eulerian Video Magnification is a recently presented method capable of revealing temporal variations in videos that are impossible to see with the naked eye.

Using this method, it is possible to visualize the flow of blood as it fills the face. From its result, a person's heart rate is possible to be extracted.

This research work was developed at Fraunhofer Portugal and its goal is to test the feasibility of the implementation of the Eulerian Video Magnification method on smartphones by developing an Android application for monitoring vital signs based on the Eulerian Video Magnification method.

There has been some successful effort on the assessment of vital signs, such as, heart rate, and breathing rate, in a contact-free way using a web camera and even a smartphone. However, since the Eulerian Video Magnification method was recently proposed, its implementation has not been tested in smartphones yet. Thus, the Eulerian Video Magnification method performance for color amplification was optimized in order to execute on an Android device at a reasonable speed.

The Android application implemented includes features, such as, detection of a person's cardiac pulse, dealing with artifacts' motion, and real-time display of the magnified blood flow.

In spite of the fact that the visual system of a human has a limited perceptibility in either spatial or temporal domains, technology can reveal many of the imperceptible signals that fall outside human perceptual range.

Many of these signals carry important information, such as a slight difference in human skin color according to blood circulation. Although this difference is concealed to our eyes, it can be used to estimate the heartbeat. Similarly, motion, which is invisible to the naked eye and has low spatial energy, can be revealed by magnification, which allows us to exploit interesting behavior.

Figures 1 and 2 provide examples of magnification for color and motion variations, respectively. Figure 1(a) shows a man's face as it appears to the naked eye, and no changes can be seen. The processed images in Figure 1(b) clearly reveal the variation in the color of his face that is caused by his heartbeat. Similarly, Figure 2(a) shows three frames of an eye that show no change, and Figure 2(b) shows a magnification of the small pattern of the eye.

Our environment is crowded by small and significant temporal variations. Several approaches and schemes have been evolved to visualize these variations in either motion or color, thereby resulting in the so-called computerized microscope.

Computerized microscopes depend on computation rather than optical magnification to amplify subtle color and motion changes in high-speed or ordinary videos. The success of algorithms has supported the development of modern techniques that discover unnoticeable signals in videos.

The capability to magnify unnoticeable variations in an imaging video has opened the door for applications in biology, healthcare and mechanical and material engineering.

Figure 1. (a) source frames

Figure 2 .(b) Amplified frames

Figure 3(a) source frames

Figure 4. (b) amplified frames

4.1 History of Video Magnification

In 2005, a research group from the laboratory of computer science and artificial intelligence in the Massachusetts Institute of Technology (MIT) proposed a video magnification algorithm based on cluster trajectories. The algorithm was used to amplify subtle color or motion changes over time, which made invisible changes visible.

The authors proposed using a cartoon animation filter to create visible motion construction. These algorithms depend on a Lagrangian perspective, in which the pixel path is temporally tracked over video frames. However, these algorithms are computationally costly because they are based on accurate motion estimation. Moreover, they cannot be verified artefact-free at areas of dense motions. By contrast, Eulerian based methods amplify the intensities of pixel variation over time in a multiple scale manner.

In Eulerian methods, motion magnification does not clearly estimate motion but rather extends it through amplifying temporal color changes at fixed positions. The Eulerian methods are similar to optical flow algorithms in using differential approximation.

The approach of Eulerian video magnification (EVM) was first proposed by the MIT research group in 2012. The basic methodology of EVM considers the time series of pixels and amplifies any variation in a specified interest band of temporal frequency. For example, the selected frequencies in Figure 1 consist of plausible human heart rates to amplify a temporal band. The amplification exposes the dissimilarity of redness as blood flows through the face.

Since 2012, EVM has been a popular research area in several interesting applications, such as extracting the depth and velocity of hot air, human feeling detection, Android smartphone software, plasma physics, sound reconstruction from a distance by the vibration measurement of an item in a high-speed video, rescue, biology, mechanical engineering and civil engineering. EVM may also have significant potential in diagnostic and monitoring in medical applications. Most medical applications depend on the recovery of temporal features using the capabilities of color amplification of EVM, such as measuring vital signs without touching patients and revealing otitis media in infants.

The EVM technique can be classified into linear- and phase-based EVM. In linear approaches, motion in video is linearly proportional to the intensity variation over a first-order expansion of Taylor series.

A video sequence is considered an input, spatial decomposition is applied, and frames are filtered by a temporal filter. The produced temporal region is amplified to expose hidden information. Although this procedure is simple and rapidly detects small motion variations, it easily breaks down when the magnification factor is large because the Taylor approximation becomes inaccurate.

To overcome this problem, phase-based magnification (PVM) replaces the linear approximation with Fourier decomposition by complex steerable pyramids. The variations in phase of the pyramid coefficients over time are proportional to the motion in different video frames.

These variations can be temporally processed and then amplified to visualize the motion. In contrast to the linear-based method, the phase-based method has a higher complexity and longer processing time, but it can support larger amplification of the motion.

Eulerian linear and phase-based approaches operate faster and suffer less noise than the Lagrangian-based video magnification. However, EVM approaches do not work well for large and arbitrary motions. In this paper, we present a comprehensive study of EVM methods and review the latest work in this area. We also compare the existing works in terms of quality, speed and amplification factor. The rest of the paper is organized as follows

1.Detecting Pulse from Head Motions in Video (2013)

Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on A method to measure pulse from the Newtonian motion of the head as blood flows into its. Structural Modal Identification through High Speed Camera Video: Motion Magnification

2.Proceedings of the 2nd International Modal Analysis Conference (2014)

A validation that the motion magnified motions are indeed real and a way to compute the mode shapes of a cantilevered beam from video.

3.The Visual Microphone: Passive Recovery of Sound from Video (2014)

ACM Transactions on Graphics, Volume 33, Number 4 (Proc. SIGGRAPH), 2014.A technique to recover sound from videos of objects subtly vibrating in response to sound.

4.Refraction Wiggles for Measuring Fluid Depth and Velocity from Video (2014)

Proc. of European Conference on Computer Vision (ECCV), 2014. A method to recover the velocity and depth of hot air from video.

5.isual Vibrometry: Estimating Material Properties from Small Motions in Video (2015)

IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2015 A method to estimate the material properties of small motions in videos.

5. EVM

EVM can reveal and amplify the small motions and changes in videos. EVM is applied to every level within a pyramid and not over the original pictures because the goal is to amplify the pyramid levels that contain the movement frequencies. This technique then multiplies the desired frequencies by a factor known as magnification factor α, which is defined by the user. EVM amplifies the actual motion and allows us to recognize movements that are undetectable to the naked eye.

The magnified values of the desired frequency are added back to the not magnified ones of the same level to obtain the final video with exaggerated motion. Although many methods use the same general principle as that of EVM, they differ in the way they work.

One of the major differences amongst these methods is the type of pyramid utilized in the algorithm. The linear video magnification (LVM), which was presented by Wu et al, applies the Laplacian pyramid decomposition technique to the input video to decompose the video sequence according to spatial frequency, which is followed by temporal filtering. The resulting output signals of this operation are then magnified by a factor and added back to the signals that are entered to the temporal filter. However, this method supports low magnification factors.

To solve this problem, Wadhwa et al proposed an Eulerian method based on a technique called complex steerable technique, which was inspired by phase-based optical flow. This method accepts large magnification factors, possesses fewer artefacts and introduces less noise than LVM.

However, this method requires a longer time processing due to the complexity of piping representation of the steerable pyramid. This method can be over 21 times as long as the LVM technique.

The authors developed their previous work by using a new pyramid, which they called the Resize pyramid, to reduce execution time. Liu et al proposed a way to improve LVM after processing, which is called enhanced EVM (E2VM). E2VM supports magnification factors with greater and less noise than LVM.

The efficient motion magnification system (EMMS) method has been developed to improve processing speed, which depends on wavelet decomposition.

This method supports large amplification factors, improves the speed of implementation and reduces noise.

5.1 Signal post-processing

After obtaining the raw pixel values (red, green, and blue channels), a combination of the following methods may be used to extract and improve the reflected plethysmography signal. However, each The method introduces complexity and expensive computation.

5.2 Linear Based EVM

Small motion amplification can be achieved through processing. In optical flow, motion magnification can be produced via temporal processing by using the first-order Taylor series expansions.

The goal of EVM is to process the time series of color values for each pixel in the spatial domain independently by applying a standard 1D temporal signal processing to each time series to amplify a specific band of interest. The input video frame is decomposed into distinct spatial frequency bands using a full Laplacian pyramid.

The Laplacian pyramid is a data structure in which an image is down sampled at successively sparser densities until no further down sampling is possible.

The Laplacian pyramid depends on an analysis pyramid for videos that is based on a Gaussian pyramid. The Laplacian pyramid method for analyzing video processing time has become less popular than before. Figure 3 shows a working LVM mechanism.

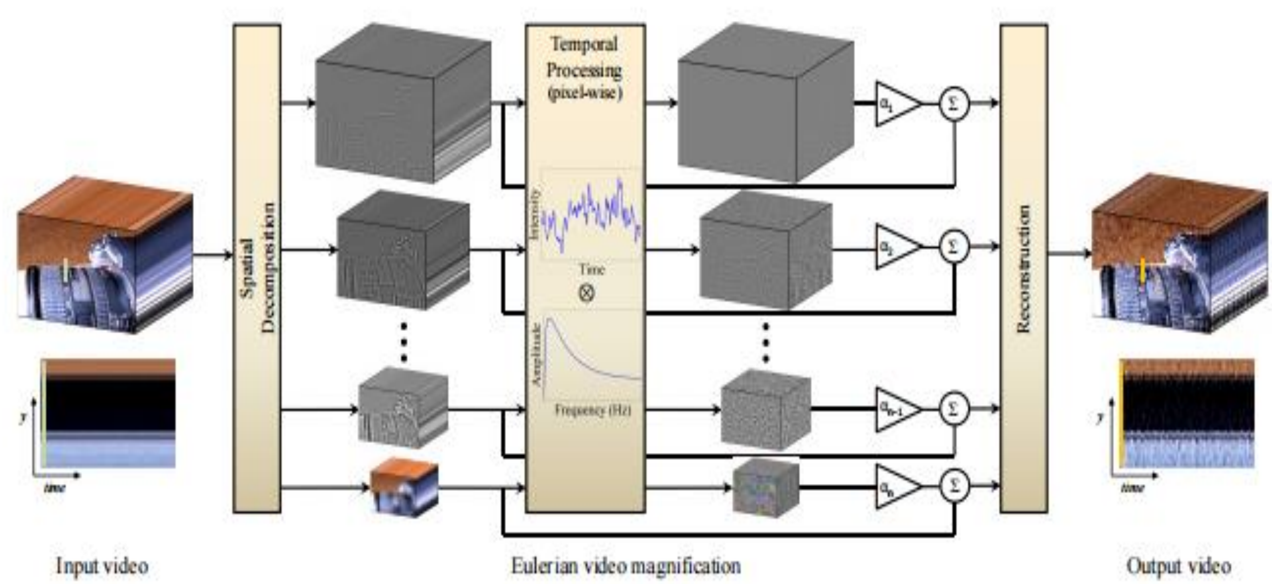


Figure 5. working LVM mechanism

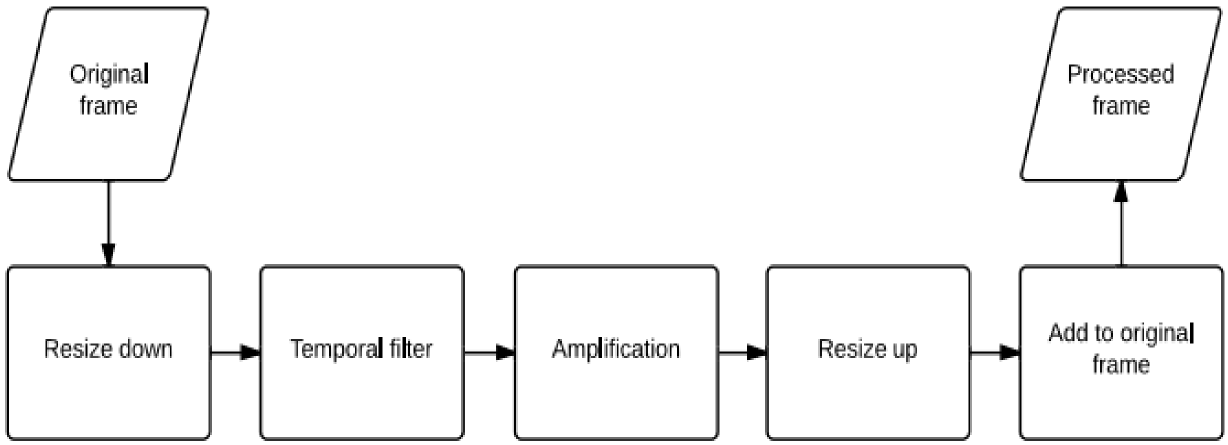


Figure 6 . Overview of Eulerian video magnification mechanism

The purpose of implementing multiple variants of the method was to study how the method worked and select which spatial and temporal filters would better fit the application goal: amplify color variation in real-time.

shows generic steps of the method which will be detailed on each of the following sections.

The final step, added to the original frame, however, remains the same in all implementations.

Which is when the magnified values are added back to the original frame in order to obtain the processed frame.

5.2.1 Resize down

This step applies a spatial filter by calculating a level of the Gaussian pyramid. This is achieved by looping to the desired level where the input to the next loop is the result from

the previous loop, starting with the original frame. A Gaussian pyramid level is calculated

by first, convolving the input frame with the kernel, *K*:

$$K = \frac{1}{256}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

and then, down sampling the frame by rejecting even rows and columns

5.2.2 Temporal filter

An ideal bandpass filter was used to remove any amplification of undesired frequency from the color variation of each pixel.

To construct this ideal filter, the Fourier transform was

calculated for each pixel over the sliding window of 30 frames. Then, frequencies below 45 and above 240 were set to zero, and the frame was rebuilt using the inverse Fourier Transform.

5.2.3 Amplification

In this step, the result of the temporal filter is multiplied by an α value, which results in the magnification of the color variation selected by the temporal filter.

5.2.4 Resize up

This step performs the inverse operation of the resize down step, where it resembles the frame by inserting even rows and columns with zeros, and then, convolves the input frame with the same kernel multiplied by 4. However, when the original frame is not multiple of two, an additional resize operation has to be done in order for the up sampled frame to match the original frame's size.

5.3 Pyramid Decomposition

A pyramid is the process of decompositions of images in different scales spatially. The two most popular pyramid decomposition approaches are Gaussian and Laplacian.

Gaussian pyramid is constructed by smoothing the original image with a Gaussian filter and then scaling it down. A Gaussian pyramid consists of a sequence of lowpass, down sampled pictures. The Gaussian pyramid is similar to a Laplacian pyramid; at a specific level of each image of the Laplacian pyramid is the difference between two corresponding neighboring levels of the Gaussian pyramid.

The smallest level is merely preserved. As a result, the different images can be used to reconstruct the original image. However, the Laplacian pyramid can be assumed to be a sequence of bandpass, downsampled images. Figure 5 and figure 6 illustrate the Gaussian and Laplacian pyramid decompositions.

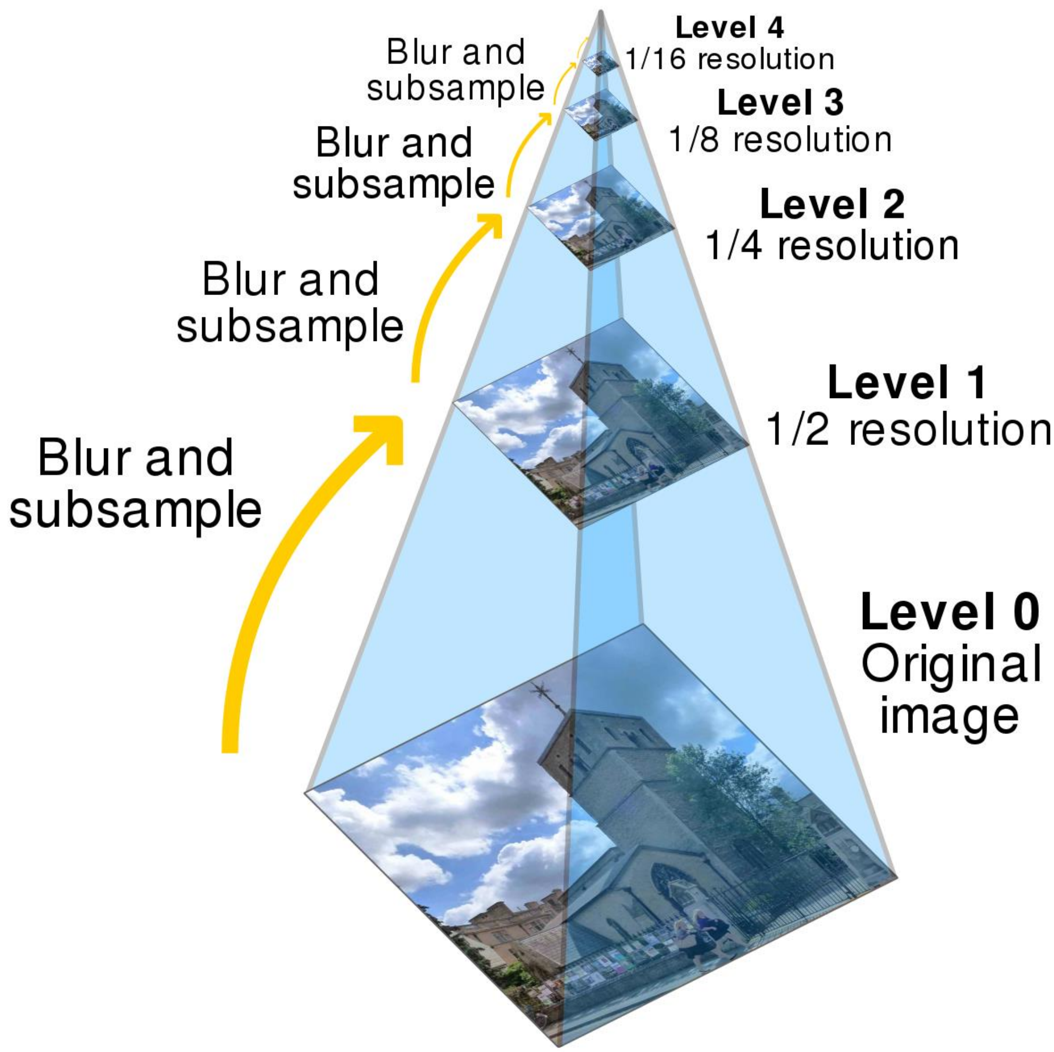


Figure 7. Gaussian and Laplacian pyramids

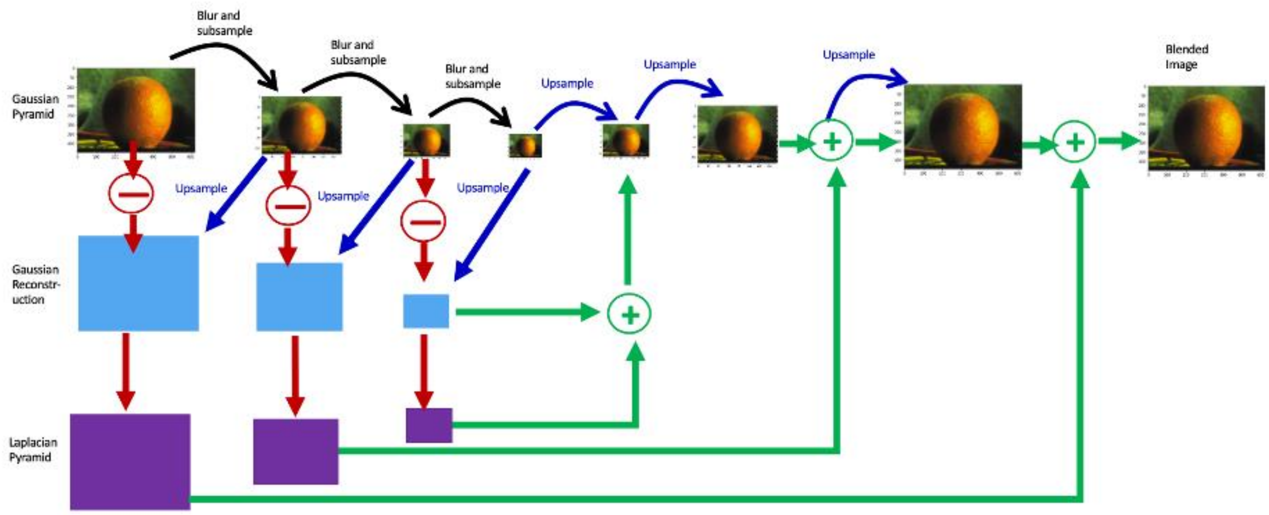


Figure 8. Gaussian and Laplacian pyramids 2

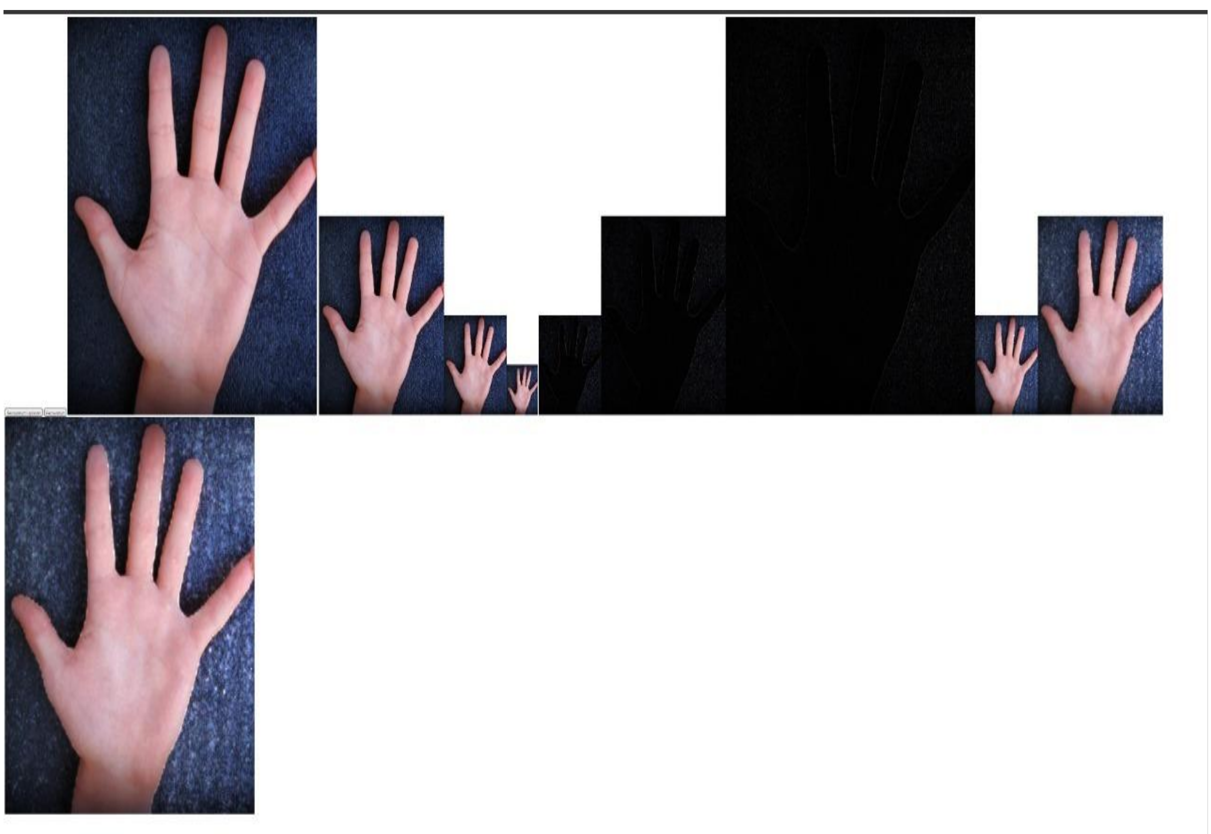


Figure 9. Result by applying gaussian and Laplacian pyramids

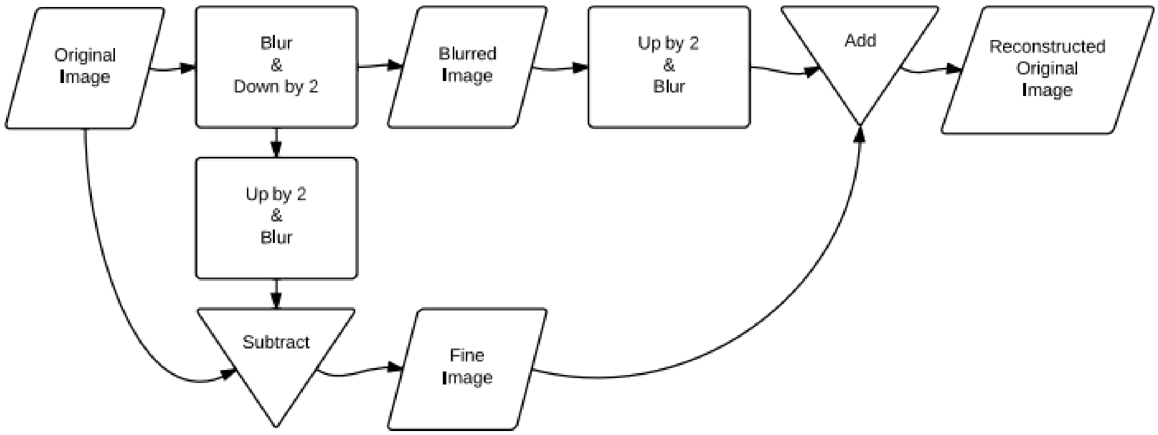


Figure 10. Overview of image deconstruction and reconstruction for building a Laplacian pyramid

by applying any type of spatial lowpass filter and scaling the image down or up by 2. In this case, a Gaussian filter was applied as described on steps resize down and resize up of the first implementation. Further levels of the pyramid can be computed by decomposing the blurred image in the same manner

5.4 Temporal filtering

Temporal filtering is used to extract the motions or signals to be amplified. Thus, the filter choice is application dependent. For motion magnification, a broad bandpass filter, such as the Butterworth filter, is preferred. A narrow bandpass filter produces a more noise-free result for color amplification of blood flow. An ideal bandpass filter is used due to its sharp cutoff frequencies. Alternatively, for a real-time implementation low-order IIR filters can be useful for both: color amplification and motion magnification.

A temporal filtering process is applied to the series of the temporal pixels in each spatial band of the spatial pyramid construction to extract the interest frequency bands. A temporal bandpass filter is used to extract motions or signals that are intended to be amplified.

According to the application utilized in the algorithm, users should be able to control the frequency band interest. However, the frequency band can be automatically selected in some cases.

Filter selection also depends on the type of the application used. For example, a filter with a wide pass band is often preferred for motion magnification, whereas a narrow-passband filter is preferred for color amplification, such as blood flow, because the latter results in less noise distortion. However, for real time implementation for motion magnification and color amplification, low-order IIR filters are convenient to use.

Various frequency responses of several types of temporal filters are shown in Figure below.

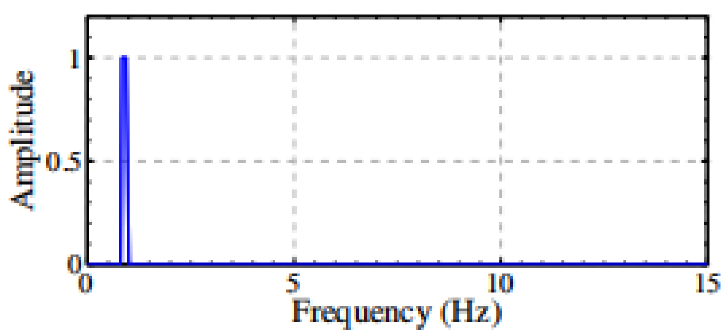


Figure 11. (a) Ideal 0.8-1 Hz (face)

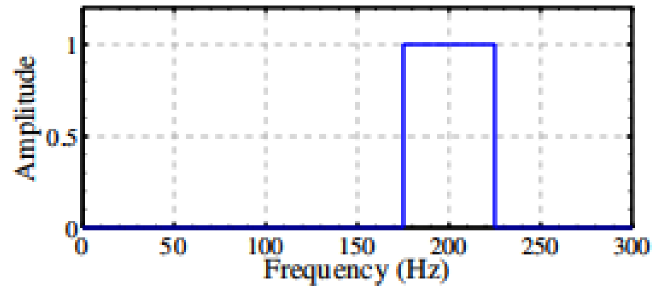


Figure 12. (b) Ideal 175-225 Hz (guitar)

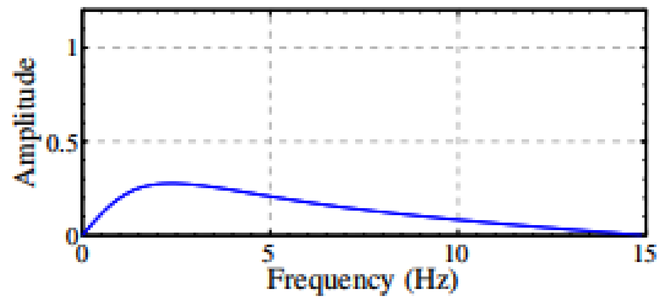


Figure 13. . (c) Butterworth 3.6-6.2 Hz (subway)

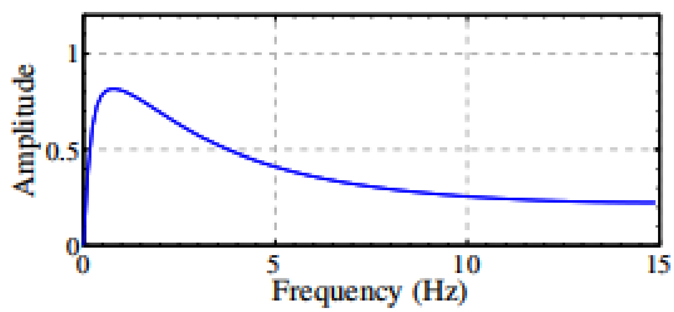


Figure 14. . (d) Second-order IIR (pulse detection)

5.5 Summary of EVM technique

starts by describing the main goal and motivation for developing a lightweight, real time Eulerian Video Magnification-based method for the Android platform. Then, an overview of the implemented algorithm’s steps is described. The algorithm begins by detecting a person’s face and magnifying it using the Eulerian Video Magnification method. Then, it extracts a possible pulse signal by averaging the green channel of a rectangle’s face, which is validated, processed by detrending and normalization, and validated again, by verifying its shape and timing. Finally, the heart rate is estimated using the power spectrum technique to obtain a

signal frequency.

5.6 ICA vs EVM

Poh et al. proposed a method for estimating HRV features and frequency domain parameters. This technique used an RGB camera recording from a short distance; however, it has a limited frequency band resolution (only red, green and blue).

On the contrary, Mcduff recommended combining the cyan, orange and green bands to achieve better results.

For that, a novel five band digital camera that has the ability to record cyan and orange along the RGB bands was used instead of the standard RGB camera.

Using this method, alternative frequency bands can be used to improve the accuracy of measuring physiological features.

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

The difference between ICA and other methods such as classical factor analysis is that ICA searches for statistically independent and non-Gaussian signals.

ICA is applied to many applications in various areas. For example, it can be used in brain imaging since the electrodes attached to the scalp linearly combine the signals generated from the various sources of the brain. Another application is in econometrics where ICA is used to break down parallel time series into independent components to reveal information about the data set structure.

ICA has also been used in image feature extraction to locate image features that are very independent.

Recovering the PPG signal using ICA is currently the most studied method. However, this method is sensitive to movement and variation in light. The EVM technique was shown to be effective when it comes to HR extraction. It is based on the Eulerian principle that states that pressure and velocity develop over time and hence the term” Eulerian video magnification”.

When using the Eulerian perspective on videos, each pixel is processed independently and treated as a time.

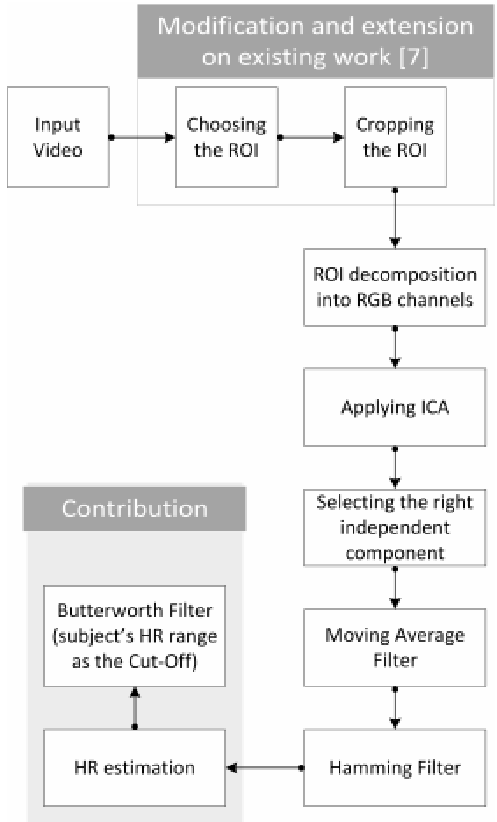


Figure 15. Block diagram of the ICA based method

EVM utilizes localized spatial pooling and temporal filtering to pull out the signal of the cardiac pulse.

This technique allows for the magnification of the subtle changes in the skin color instigated by the blood flow in facial vessels. Hence this method can be used to make this phenomenon visible to the naked eye. Nonetheless, extracting the HRV signal using this method is still a challenge and only few attempts have been made

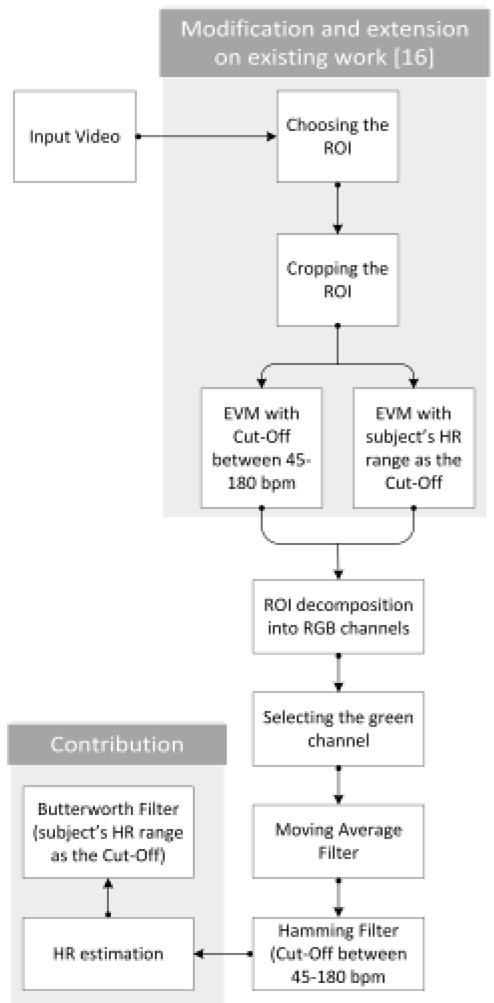


Figure 16. Block diagram of the EVM based methods

$$y_i'(t) = \frac{y_i(t) - \mu_i}{\theta_i}$$

(1)

Where μi and θi are the mean and standard deviation respectively for i = 1, 2, 3.

5.7 ICA based method

We summarize the ICA based method as follows

1) Obtain the raw signals from the ROI,

2) Extract the red, green and blue (RGB) channels from the ROI.

3) Spatially average each of the RGB channels using all pixels in the ROI to produce red, green and blue discrete signals. These signals are denoted as y1(t), y2(t) and y3(t) respectively and are referred to as the raw elements. Here, we detrend the raw elements with $\lambda = 2000$ as the smoothness parameter and normalize the detrended raw elements using equation (1).

4) Apply the ICA technique using the joint approximate diagonalization of eigen matrices (JADE) algorithm to break up the raw elements into three independent source signals. ICA helps separate the observed raw elements from the noise caused by motion artifacts. When applying ICA, the source signals are returned in a random order.

5) To select the best component, apply Fast Fourier Transform (FFT) on the output sources and choose the one with the highest peak within the range of 45 and 180bpm.

6) Apply a series of filters to the resulting signals. The first filter is a K-point moving average to smooth the signal using equation (2) with $K = 5$.

7) Apply a 3db band-pass Hamming window filter with a cut-off between 0.75 and 3HZ, which are low and high frequencies at 45 and 180bpm (normal human HR range).

8) Since our experimental setup is lit by artificial lighting, this kind of illumination produces high frequency signals which interfere with the PPG signals extracted from the video. Consequently, additional filtering is required to clean up the signal. If we were measuring the HR, it would not be

a major issue (since HR involves counting the number of PPG peaks regardless of the time interval between them).

However, in the case of HRV estimation, to get the R-R interval, the signal should be very clean. Hence, we add a narrow IIR band pass filter with cut-off frequencies based on the specific HR range of the measurement period. This technique is especially effective for short measurement periods. It is worth noting that in numerous emotional and physiological assessment applications, HRV records of 5 minutes length are typically used.

9) Determine minHR (the minimum HR) and maxHR.

10) Use the minFreq equation (3) and the maxFreq equation (4) as cut-off range to create a narrow IIR filter. Apply the IIR filter to the signal retrieved in Step 7. 11) Design a 2nd order Butterworth filter to attenuate frequencies outside the interest band. We found that α = 5 and β = 5 rendered good results.

We employ the Butterworth filter as opposed to other IIR filters because of its flat passband and stop band-band can help avoid ripples.

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

N is the number of the neighbors
Ys is the output of smoothed for the Ith values

$$minFreq = (minHRi - \alpha)/60 \qquad (3)$$

Alpha is constant number and i where to the period of measurements

$$maxFreq = (maxHRi - \beta)/60 \qquad (4)$$

beta is constant number

5.8 EVM based method

To the best of our knowledge, the proposed EVM based method shown in Figure 9 is the first successful attempt at extracting accurate HRV features from a facial video using the EVM technique.

If the goal is just to obtain the HR, the green channel extracted from the EVM magnified video is sufficient for such operation and no further processing is necessary.

However, the HRV signal is very sensitive to noise and therefore more filtering is required. The following are the steps needed to perform the proposed EVM method.

1) Apply the EVM magnification algorithm on the cropped ROI instead of the whole frame.

2) Apply the Gaussian pyramid for spatial pooling and a narrow band-pass filter, by applying the Fourier transform for every pixel, and setting to zero every frequency that does not fall between the Cut-Off (45 and 240bpm). Then apply the inverse Fourier transform to rebuild the frame.

3) Spatially average each of the RGB channels using the same process employed in Step 3 for the ICA based method.

4) From our experiments, the green channel seems to render the cleanest PPG signal. Therefore, it is the one chosen for further processing. To reduce noise, the green channel is filtered with 5dp moving average, 3dp Hamming window with cut-off frequencies between 45 and 180bpm and with a Butterworth band-pass filter with minFreq. and maxFreq. as cut-off frequencies (same filters used in the ICA based method).

5.9 HRV signal processing for ICA and EVM

Due to the limit of the camera frame rate, (we use 30f ps in our data collection) the resulting PPG signal is considered to be at low resolution one.

This would drastically affect the HRV analysis. To resolve this issue, a cubic spline interpolation is implemented at 240Hz. After the signal is interpolated, a peak detection algorithm is applied. McDuff et al. proved that a moving window of 0.25 seconds gives highly correlated results with the contact sensor measurements.

The highest point is detected inside the window. If the detected point is greater than the highest point in the previous window and greater than the highest point in the next window, it will be selected as a peak.

The same procedure is repeated for the whole signal. The RR intervals series is extracted by calculating the time between consecutive peaks.

Head motion, variation in the illumination and other artifacts can influence the accuracy of the HRV analysis by affecting the RR interval. To reduce the effect of such sources of error, the HRV series needs to be filtered. The interval correction used in our work is based on the technique

A second order Butterworth filter is used in this case. The normal HR range for humans is between 45 and 180bpm, therefore the cut-off frequency given to the filter is 0.75 – 3Hz.

After that we converted the time domain signal to the frequency by applying the Fast Fourier Transform (FFT). To estimate the HR continuously, FFT procedure, peak detection and smoothing procedures are performed every 0.5 seconds in a moving window of the last three seconds of signal.

A duration of three seconds is chosen as it gave the best results relative to other lengths that we attempted. These steps are performed every 0.5 seconds to increase the number of HR calculations which will be later smoothed to give a better estimation.

To reduce the artificial high frequencies that appear when the signal is treated as periodic by the FFT, a Hanning window was used to bring the edges to zero.

After performing the FFT, the highest peak was detected in the interest band which is the HR band between 45 and 180. The detected peak is then converted to the right frequency in the FFT vector.

The HRV analysis of the frequency domain is achieved by a power spectrum density (PSD) estimation. The main parameters we are interested in are Low Frequency (LF) which corresponds to 0.04 to 0.15Hz, High Frequency (HF) which corresponds t0 0.15to 0.4Hz and LF/HF. The values of LF and HF parameters are converted to normalized units (n.u.).

6. Photo-plethysmography (PPG)

Human vital signs like heart rate, blood oxygen saturation and related physiological measures can be measured using a technique called photo-plethysmography (PPG). This technique involves optically monitoring light absorption in tissues that are associated with blood volume changes. Typically, this is done via a contact sensor attached to the skin surface. Remote Photoplethysmography (rPPG) detects the blood volume pulse remotely by tracking changes in the skin reflectance as observed by a camera.

The process of rPPG essentially involves two steps: detecting and tracking the skin colour changes of the subject, and analysing this signal to compute measures like heart rate, heart rate variability and respiration rate. Recent advances in computer video, signal processing, and machine learning have improved the performances of rPPG techniques significantly. Current state-of-the-art methods are able to leverage image processing by deep neural networks to robustly select skin pixels within an image and perform HR estimation.

However, this reliance upon heavy machine learning (ML) processes has two primary drawbacks:

(i) it necessitates rPPG specific training of the ML model, thereby requiring collection of large training sets;

(ii) complex models can require significant computation time on CPUs and thus can potentially add a bottleneck in the pipeline and limit real-time utility.

Since rPPG analysis is originally a signal processing task, the use of an end-to-end trainable system with no domain knowledge leaves room for improvement in efficiency (e.g., we know that pulse signal is embedded in average skin colour changes, but the ML system has to learn this).

We introduce a simplified and efficient rPPG pipeline that performs the full rPPG analysis in real-time. This method achieves state-of-the-art results without needing any rPPG related training. This is achieved via extracting regions of interest robustly by 3D face modelling, and explicitly reducing the influence of head movement to filter the signal.

While heart rate is a useful output from a PPG/rPPG analysis, finer analysis of the obtained blood volume pulse (BVP) signal can yield further useful measures.

One such measure is heart rate variability: an estimate of the variations in the time-intervals between individual heart beats.

This measure has high utility in providing insights into physiological and psychological state of a person (stress levels, anxiety, etc.). While traditionally this measure is obtained based on observation over hours, short and ultrashort duration (≤ 5 mins) HRV are also being studied.

to obtain ultra-short HRV measure as a proof-of-concept/technology demonstrator for longer duration applications. The computation of heart rate variability requires temporally locating heart beats with a high degree of accuracy.

Unlike HR estimation, where errors in opposite directions average out, HRV analysis is sensitive to even small artefacts and all errors add up to strongly distort the final measurement. Thus, estimating HRV is a challenging task for rPPG and this has received relatively little focus in literature. Our method extracts a clean BVP signal from the input via a twostep wide and narrow band frequency filter to accurately time heart beats and estimate heart rate variability.

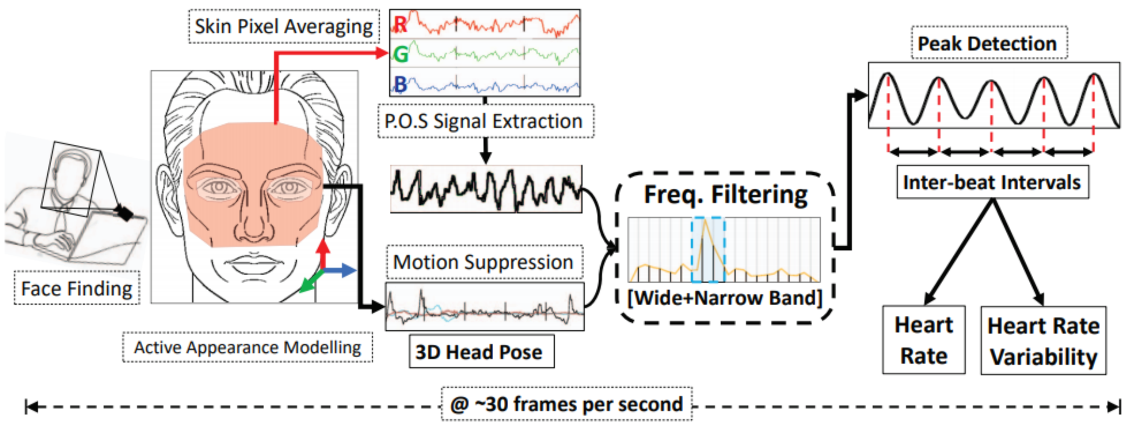


Figure 17. An overview of the proposed heart rate and heart rate variability estimation pipeline (left to right).

face in captured webcam images are detected and modelled to track the skin pixels in the region of interest. A single 1-D signal is extracted from the spatially averaged values of these pixels over time.

In parallel, 3-D head movements are tracked and used to suppress motion noise. An FFT based wide and narrow band filtering process produces a clean pulse waveform from which peaks are detected.

The inter-beat intervals obtained from these peaks are then used to compute heart rate and heart rate variability. The full analysis can be performed in real time on a CPU.

6.1 Signal processing based rPPG methods

Since the early work of Verkruysse showed that heart rate could be measured from recordings from a consumer grade camera in ambient light, a large body of research has been conducted on the topic.

Most published rPPG methods work either by applying skin detection on a certain area in each frame or by selecting one or multiple regions of interest and track their averages over time to generate color signals.

A general division can be made into methods that use blind source separation (ICA, PCA) vs those that use a 'fixed' extraction scheme for obtaining the BVP signal. The blind source separation methods require an additional selection step to extract the most informative BVP signal. To avoid this, we make use of a 'fixed' extraction scheme in our method.

Among the 'fixed' methods, multiple stand out and serve as inspiration and foundation for this work.

Tasli et al. presented the first face modelling-based signal extraction method and utilized detrending based filtering to estimate BVP and heart rate. The CHROM method uses a ratio of chrominance signals which are obtained from RGB channels followed by a skin-tone standardization step. Li et al. proposed an extra illumination rectification step using the colour of the background to counter illumination variations.

The SAMC method proposes an approach for BVP extraction in which regions of interest are dynamically chosen using self-adaptive matrix completion. The Plane-orthogonal to skin (POS) method improves on CHROME. It works by projecting RGB signals on a plane orthogonal to a normalized skin tone in normalized RGB space and combines the resulting signals into a single signal containing the pulsatile information. We take inspiration from Tasli et al. and further build upon POS. We introduce additional signal refinement steps for accurate peak detection to further improve HR and HRV analysis.

6.2 Deep learning based rPPG methods

Most recent works have applied deep learning (DL) to extract either heart rate or the BVP directly from camera images. They rely on the ability of deep networks to learn which areas in the image correspond to heart rate. This way, no prior domain knowledge is incorporated, and the system learns rPPG concepts from scratch.

DeepPhys is the first such end-to-end method to extract heart and breathing rate from videos. HR-Net uses two successive convolutional neural networks to first extract a BVP from a sequence of images and then estimate the heart rate from it. Both show state-of-the-art performance on two public datasets and a number of private datasets.

Our presented algorithm makes use of an active appearance model to select regions of interest to extract a heart rate signal from.

Due to this, no specific rPPG training is required while prior domain knowledge is more heavily relied upon.

6.3 HRV from PPG/rPPG

Some past methods have also attempted extracting heart rate variability from videos. A good overview is provided by Rodriguez et al.

Because of the way HRV is calculated, it is crucial that single beats are detected accurately with a high degree of accuracy.

Methods that otherwise show good performance in extracting HR can be unsuitable for HRV analysis, since they may not provide beat locations. Rodriguez et al. evaluate their baseline rPPG method for HRV estimation. Their method is based on bandpass filtering the green channel from regions of interest. However, their results are only reported on their own dataset, which makes direct comparison difficult.

Our method also estimates heart rate variability by obtaining precise temporal beat locations from the filtered BVP signal.

7. Our work and testing

The results obtained during this dissertation and the analysis of these reports include several performance improvements to the implemented Eulerian Video Magnification method and the developed algorithm to estimate a person's heart rate.

compares the measurements from the implemented Android application, Pulse, to a sphygmomanometer, and also, to another Android application by flutter that estimates a person's heart rate.

7.1 Algorithm

EVM steps

1. Loading the video {we have the frames of the video}

2. We do the Gaussian Pyramids and then we specify the levels of it with the functions "depth = floor(log(min_size) / log (2)) - 4"

3. We take the output frames and then use it in a temporal filter "Ideal BPF" its ranges from 0.5 to 4 and this is the normal heart rate for the humans so before it or after it gets removed

4. We do the magnification, here we multiply in a constant

5. We do the construction, here we get the last level of Laplacian and then we add it to the same opposite frame

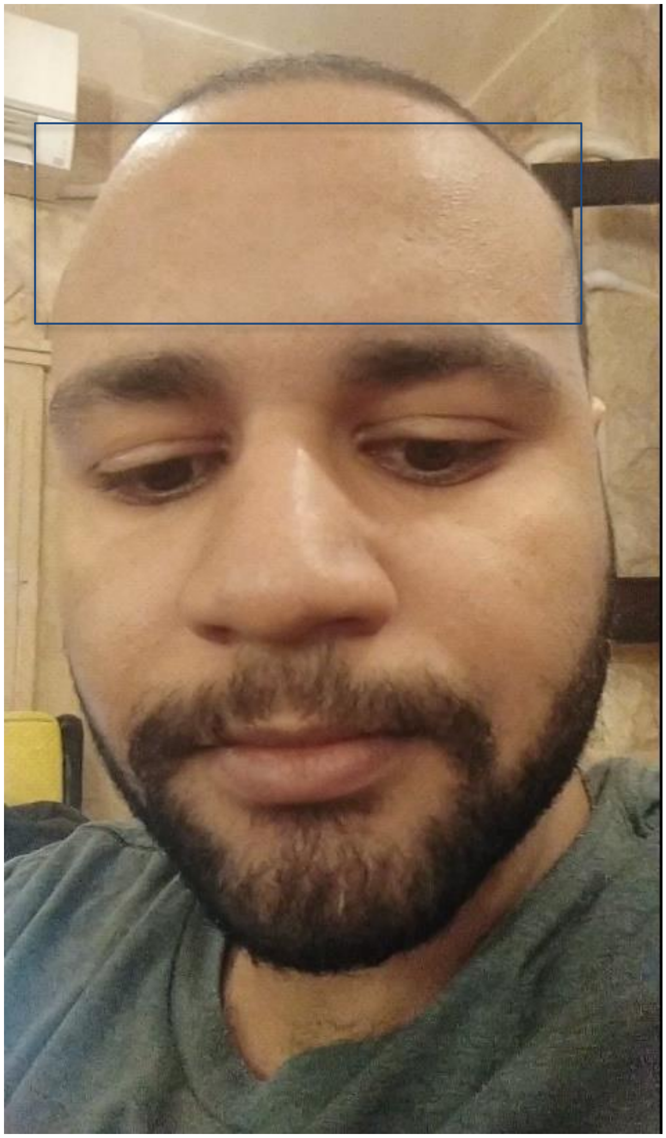


Figure 18. Video frame

Figure 19 Magnified ROI

7.2 Heart Rate calculation

We take the reconstructed frame and take from it the illumination of each frame shown in the next figure, we extract it by transforming the pixels of the frame to the YIQ plane “YIQ is the color space used by the NTSC color TV system, employed mainly in North and Central America, and Japan. I stand for in-phase, while Q stands for quadrature, referring to the components used in quadrature amplitude modulation.

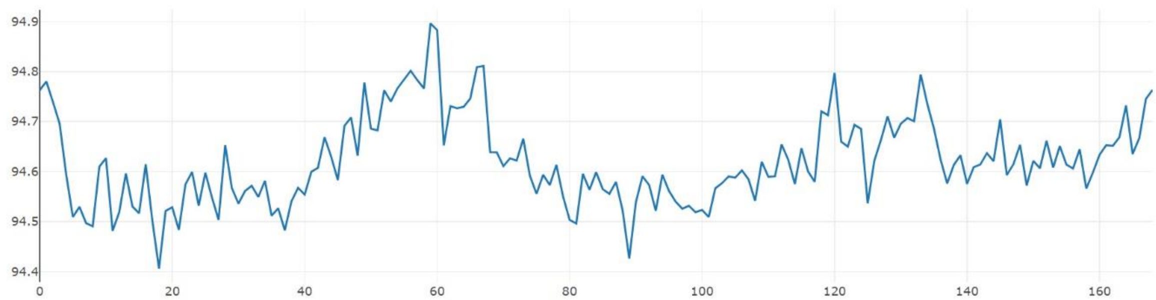


Figure 20. Illumination in time domain (magnitude vs frame number)

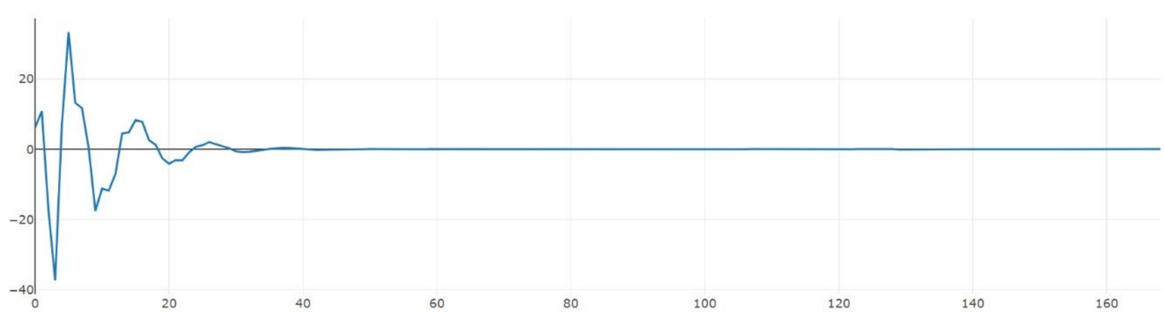


Figure 21. Filtered illumination (magnitude vs frame number)

Then we consider the YIQ plane to be our signal, so we apply to it the BPF and its cutoff frequency from 0.75 to 3 HZ as the human heart rate

Then we apply the FFT and peak detection and smoothing, and those three functions occur each 0.5 seconds with sliding window and the window size is the last 6 seconds of the signal, we take those last 6 seconds apply to them the FFT and zero padding and Hanning window

First, we apply the Hanning window by multiplying the Hanning window function by this interval then we apply the zero banding to a specific range then we apply the FFT.

Then we get the peak of the result and then smooth it to the right frequency by getting the bin of the peak and then multiply it by the sampling frequency and then divide it to the total bins in the FFT.

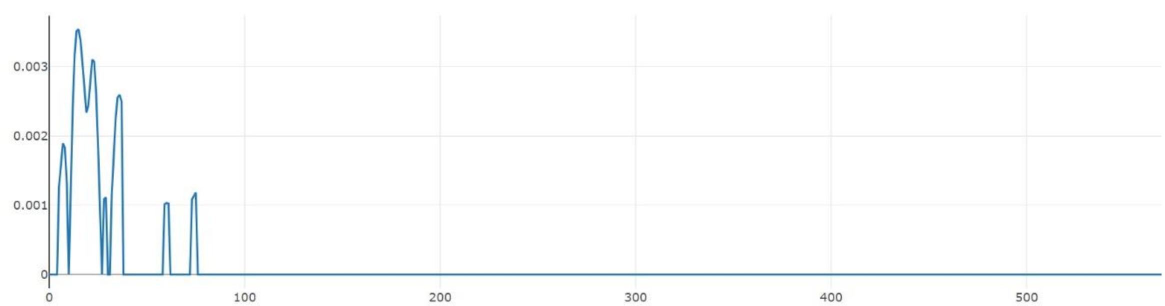


Figure 22. Illumination after FFT (magnitude vs FFT pins)

The heart rate can be calculated from the smoothed frequency multiplied by 60, and here it is the heart rate.

7.3 Testing

During working on the project for all phases we definitely passed through many problems and trials.

- We thought we should work on the green signal from the frames where actually we should have worked from the beginning on the illumination
- It was assumed that we deal with the signal as pixels while it is frames
- It took us a long time to determine the suitable sampling rate
- Heart rate calculation was a big deal since each paper was calculating in a different ways and techniques without enough details

7.4 Results

Algorithm calculation for heart got fine results with acceptable accuracy. It should have been tested much more times but due to lack of time that all we have got.

Smart watch heart rate calculation	Algorithm heart rate calculation	Accuracy
74	75	98.6%
66	67	98.5%
96	99	96.9%
82	84	97.6%
96	98	97.9%
70	72	97.2%

Table 1 Heart rate calculation smart watch vs algorithm

8. Tools

In the beginning we will be talking about the software applications and tools that we use then we will be talking about the hardware devices.

8.1 Node.JS

Figure 23. Node js

8.2 What is Node.js?

Node.js is a server-side platform built on Google Chrome's JavaScript Engine (V8 Engine). Node.js was developed by Ryan Dahl in 2009 and its latest version is v0.10.36. The definition of Node.js as supplied by its official documentation is as follows:

- Node.js is a platform built on Chrome's JavaScript runtime for easily building fast and scalable network applications. Node.js uses an event-driven, non-blocking I/O model that makes it lightweight and efficient, perfect for data-intensive real-time applications that run across distributed devices.
- Node.js is an open source, cross-platform runtime environment for developing server-side and networking applications. Node.js

applications are written in JavaScript and can be run within the Node.js runtime on OS X, Microsoft Windows, and Linux.

- Node.js also provides a rich library of various JavaScript modules which simplifies the development of web applications using Node.js to a great extent.

8.3 Features of Nodejs

Following is some of the important features that make Node.js the first choice of software architects.

1. Asynchronous and Event Driven – All APIs of Node.js library are asynchronous, that is, non-blocking. It essentially means a Node.js based server never waits for an API to return data. The server moves to the next API after calling it and a notification mechanism of Events of Node.js helps the server to get a response from the previous API call.
2. Very Fast – Being built on Google Chrome's V8 JavaScript Engine, Node.js library is very fast in code execution.
3. Single Threaded but Highly Scalable – Node.js uses a single threaded model with event looping. Event mechanism helps the server to respond in a non-blocking way and makes the server highly scalable as opposed to traditional servers which create limited threads to handle requests. Node.js uses a single threaded program and the same program can provide service to a much larger number of requests than traditional servers like Apache HTTP Server.
4. No Buffering – Node.js applications never buffer any data. These applications simply output the data in chunks.

5. License – Node.js is released under the MIT license

Most works in this filed was built using python, and by our choice to use nodejs to make the processing time less than that way by using python.

8.4 Concepts

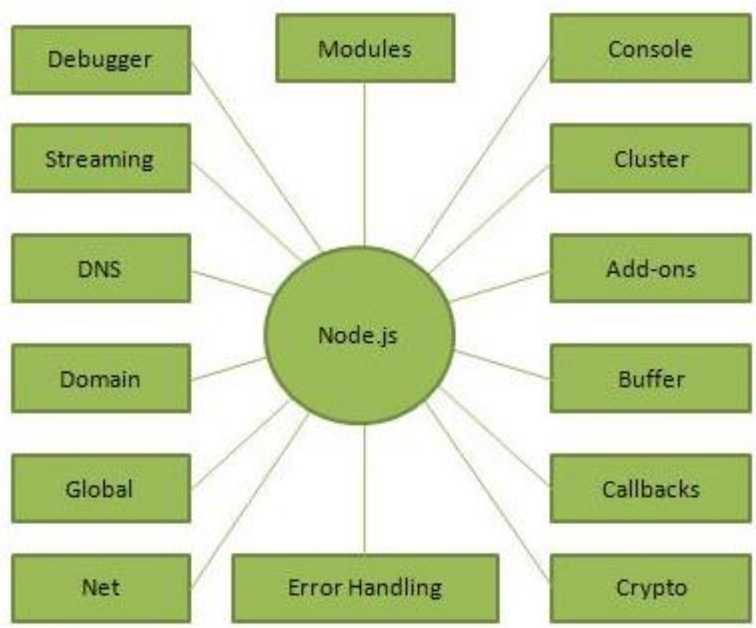


Figure 24. Concepts of node js

8.5 Where to Use Node.js?

Following are the areas where Node.js is proving itself as a perfect technology partner.

- I/O bound Applications
- Data Streaming Applications
- Data Intensive Real-time Applications (DIRT)

- JSON APIs based Applications
- Single Page Applications

8.6 Flutter

Figure 25. flutter

8.7 What is Flutter?

Flutter is Google’s portable UI toolkit for crafting beautiful, natively compiled applications for mobile, web, and desktop from a single codebase. Flutter works with existing code, is used by developers and organizations around the world, and is free and open source.

For users, Flutter makes beautiful apps come to life.

For developers, Flutter lowers the bar to entry for building apps. It speeds app development and reduces the cost and complexity of app production across platforms.

For designers, Flutter provides a canvas for high-end user experiences. Fast Company described Flutter as one of the top design ideas of the decade for its ability to turn concepts into production code without the compromises imposed by typical frameworks.

It also acts as a productive prototyping tool, with Code Pin support for sharing your ideas with others.

For engineering managers and businesses, Flutter allows the unification of app developers into a single mobile, web, and desktop app team, building branded apps for multiple platforms out of a single codebase.

Flutter speeds feature development and synchronizes release schedules across the entire customer base.

8.8 What makes Flutter unique?

Flutter is different from most other options for building mobile apps because it doesn't rely on web browser technology nor the set of widgets that ship with each device. Instead, Flutter uses its own high-performance rendering engine to draw widgets.

In addition, Flutter is different because it only has a thin layer of C/C++ code.

Flutter implements most of its system (compositing, gestures, animation, framework, widgets, etc.) in Dart (a modern, concise, object-oriented language) that developers can easily approach to read, change, replace, or remove. This gives developers tremendous control over the system, as well as significantly lowers the bar to approachability for most of the system.

9. Website

We used ReactJs to make the web site, then we make the back-end of the website and connect it with the application to make the all process on it, to prevent the long time and the more RAM for processing time.

It includes the pages which is

Home: to give a preview of what our services provide our users.

Services: to make availability to user to upload his video

Log in: to give the user the availability to log in if he has an account before

Sign up: to create a new account to the user if he used our application or website for the first time.

10.Mobile application

Also, the application has the same pages

We made the website and the application to make it easy for every user who has a smartphone with a good camera only to know his heart rate, not needing the big RAM or more space, or strong processor by specific features.

Conclusion page: to give results to the user.

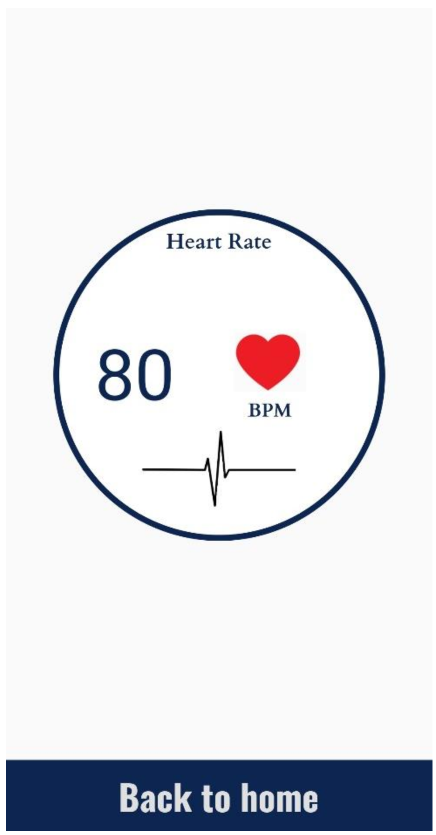


Figure 26. Application showing result

Other services: is that if the user faces ill people around him or accident happened next to him and there is somebody get injured , so he need to first aids, our application give instructions to him to follow it to help these people.

10.1 Pulse: Android application

The user interface of the implemented Android application, Pulse, is shown on figure 3.5. When the application is opened, the camera starts processing images and requests the user to place its face in front of the camera. If a face is detected the user is instructed to remain still in a bright place. Finally, if a cardiac pulse is detected, the heart rate estimated and the signal are shown on the top of the screen, while the person's face magnification is also shown. Three buttons exist on the top-right corner. From left to right, the first one, represented with a play icon, starts the record mode. The record mode averages all the heart rate values estimated in that time period. Pressing the button again finishes the record mode and displays the average of the beats per minute. This mode was used for the procedure described before the second button, represented with a camera icon, allows the application to switch between

the front and back camera. The button, represented with a wrench icon, opens a settings dialog, where the Eulerian Video Magnification may be enabled and disabled and its

amplification value, α, can be changed. Also, an option to turn on and off the number of frames per second that the application is running at is available.

About the hardware we only need a smartphone with a camera to get us the face that we will analyses.

A smartphone is a cell phone that allows you to do more than make phone calls and send text messages. Smartphones can browse the Internet and run software programs like a computer. Smartphones use a touch screen to allow users to interact with them.

11.Conclusions

presents a review of the relevant information obtained from this work and an exposition of further work and research gives an overall description of the work done, from the performance improvements of the Eulerian Video Magnification method, to the creation of the Android application capable of estimating a person's heart rate using the device's camera. Finally, section 5.2 exposes future work that could follow the development of an Android based implementation of the Eulerian Video Magnification

The main goal of this work was providing an Eulerian Video Magnification-based method capable of running on an Android device. To achieve that, various real-time implementations of the Eulerian Video Magnification method.

These were not efficient enough to execute on a smartphone in real-time.

Hence, a performance profiler was integrated into a desktop application, in order to increase the performance of the application and the Eulerian Video Magnification method.

This Eulerian Video Magnification method implementation was using a temporal bandpass filter composed by subtracting two first-order IIR lowpass filters, which is more convenient for real-time implementation than an ideal temporal filter implemented by applying the Fourier Transform to each pixel for a video segment.

Since the implemented algorithm was developed in the programming by nodejs for performance reasons, the integration into the Android platform was done through the use flutter. The application workflow started by grabbing an image from the device's camera. A person's face was detected

A region of interest (ROI) of the person's face would then be fed into the implemented Eulerian Video Magnification method to amplify color variations. The average of the ROI green channel was computed, in order to increase the signal-to-noise ratio and stored. Along The time, these stored values represent a PPG signal of the underlying blood flow variations. The signal is further processed using the detrend method to remove trends from the signal without magnitude distortion.

It is then validated as a cardiac pulse signal by detecting its peaks in order to verify its shape and timing. Finally, the heart rate estimation is computed by identifying the frequency with the higher power spectrum of the signal

11.1 Future work

Having developed a lightweight, real-time Eulerian Video Magnification-based method for the Android platform whose goal is to amplify color variations, the performance of different variants of the Eulerian Video Magnification method could be improved.

This would increase the usage of this method in other kinds of devices and in other kinds of applications. Other uses for the Eulerian Video Magnification method could be studied, such as, using it as a security camera to detect small motion by magnifying such motion, or to identify suspicious people by detecting its heart rate in a contact-free way. Another idea would be to use the Eulerian

Video Magnification method with the objective of identifying if a person is drunk or not, based on the work Nevertheless, the implemented Pulse application needs to improve its heart rate estimation accuracy, and its face detection module in order to not lose track of a person’s face. In addition, other vital signs could be monitored, such as breathing rate.

In order to support some features of the implemented Android application.

12. References

[1] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Fredo Durand, William Freeman, “Eulerian Video Magnification for Revealing Subtle Changes in the World”, ACM Transactions on Graphics, July 2012

[2] Haider Ismael Shahadi, Zaid Jabbar Al-allaq, Hayder J. Albatat, Ahmed Toman Thahab, “Eulerian video magnification”, Article in Indonesian Journal of Electrical Engineering and Computer Science, may 2020

[3] BURT, P., AND ADELSON, E. 1983. The laplacian pyramid as a compact image code. IEEE Trans. Comm. 31, 4, 532–540.

[4] C. Liu, A. Torralba, W. T. Freeman, F. Durand, and E. H. Adelson, “Motion magnification,” ACM Trans. Graph., vol. 24, no. 3, pp. 519-526, 2005.

[5] C. Ordóñez, C. Cabo, A. Menéndez, and A. Bello, “Detection of human vital signs in hazardous environments by means of video magnification,” PLoS One, vol. 13, no. 4, pp. 1-15, 2018.

[6] A. Sarra, Z. Mao, C. Niezrecki, and P. Poozesh, “Vibration-based damage detection in wind turbine blades using Phase-based Motion Estimation and motion magnification,” J. Sound Vib., vol. 421, pp. 300-318, 2018.

[7] J. G. Chen, N. Wadhwa, Y. J. Cha, F. Durand, W. T. Freeman, and O. Buyukozturk, “Modal identification of simple structures with high-speed video using motion magnification,” J. Sound Vib., vol. 345, pp. 58-71, 2015.

[8] M. Z. Poh, D. J. McDuff, and R. W. Picard, “Non-contact, automated cardiac pulse measurements using video imaging and blind source separation,” Opt. Soc. Am., vol. 18, no. 10, pp. 10762-10774, 2010.