Extraction of Heart Rate from a Facial Video

Larissa Carvalho[1] H.G. Virani[2] Shajahan Kutty[3]
[1,2,3]Dept. of Electronics and Telecommunication
[1,2,3]Goa University,Ponda, Goa, India

Abstract—*In this project I have implemented a noninvasive heart rate monitoring system to monitor subjects using Digital Image Processing. Here the main focus is on revealing the variations in videos that are difficult or impossible to see with the naked eye. This method, takes a standard video of the subject as input, performs face tracking and applies pyramid decomposition, followed by filtering of the frames. The resulting signal is then amplified to reveal hidden information. We are thus able to visualize the flow of blood as it fills the face.This method is based on the Eulerian Video magnification algorithm presented at SIGGRAPH 2012.*

I. INTRODUCTION

Heart rate describes how many times a heart, beats per minute.The body gets stressed out during physical exercises or illness or mental stress, therefore the heart rate is a parameter of high importance to medicine, physics, and psychology and to many other fields.

Cardiac pulse measurements play an important role in diagnosing heart diseases such as Tachycardia and Bradycardia. It also helps in monitoring during physical exercises to avoid health hazards and to estimate the extent of one's physical training. Nowadays Electrocardiograms (ECG), Heart Rate Monitors (HRM) and Pulse Oximeters are used to measure heart rate however these devices are not feasible to operate during physical exercises or there is minimal contact required with the patient. Thus by using a non contact method, we are able to minimize the amount of cabling and clutter associated with neonatal ICU monitoring, burn or trauma patient monitoring, sleep studies and long term epilepsy monitoring thus facilitating patient comfort and use of minimal hardware.

Why do we get the idea that something cannot be seen with a naked eye, but can be seen by a camera?To understand this, let us compare their spatial resolutions, which describes the ability of an imaging device or organ, to distinguish between two points, that are located at a small angular distance - the smaller is the angular distance, the higher is the spatial resolution and more details can be seen.To do this let us use the Rayleigh's Criterion. When light enters a lens, a diffraction pattern is produced. The Rayleigh's Criterion defines the minimal distance between diffraction patterns of two distinguishable objects as the distance, when the diffraction minimum of one source point coincides with the diffraction maximum of the other source point. According to Rayleigh's Criterion, we can calculate the minimal angular distance θ, which must separate two distinguishable points of an object as

$$\theta = 1.22\frac{\lambda}{D}$$

Where λ is the wavelength of the light
D is the diameter of the eye pupil aperture.

In our case λ = 550 nm, because it is the approximate wavelength of the human eye's greatest sensitivity and D= 3 mm.

By inserting these two parameters in the equation we get θ as 2.24 x 10^{-4} rad.

The spatial resolution on the human light perceptive cells can be expressed as

$$\Delta l = \theta f$$

Where f is the focal length of the eye and is approximately 22 mm. Inserting the angular resolution calculated above and the focal length, we get Δl ≈ 5 x 10^{-6} m.

Let us now take a look at cameras spatial resolution which is given as

$$\Delta l = 1.22\frac{f\lambda}{D}$$

Where f/D is called the focal ratio which is about 1.4 to 22.In our case let us consider it to be 1.4. Then Δl = 0.94μm. This means that a typical detector with pixel size approximately 5 μm, would not be able to distinguish between these two points and that middle class cameras' resolutions are not determined by the diameter of their aperture, but by the pixel sizes of their detectors.

Let us assume that the distance between two photons on the detector, Δl, to be perceived as photons from different sources, must be two pixel sizes, so Δl ≈ 10μm for an average middle class camera. This way, they can be detected by two pixels that lay one pixel apart. Knowing the value of a typical focal length, f, of a middle class camera, which is about 15 mm, we can calculate the spatial resolution of an average camera, determined by the pixel size as θ=Δl/D . This estimate leads to the result ≈ 6.7 x 10^{-4} rad which is approximately 2' 17". This means that the Eulerian video magnification algorithm does not depend on cameras capabilities to revealunseen motion, but on other principles.

II. EXPERIMENTAL SETUP

The test setup was designed in a way to minimize external environmental influences in order to extract the heart rate from the facial video.The subject was placed infront of a webcam(1.3MP, 24 bit RGB, 8 bit per channel), and colour videos were recorded for 25 sec at 30 frames per sec with a resolution of 640x480 pixels.

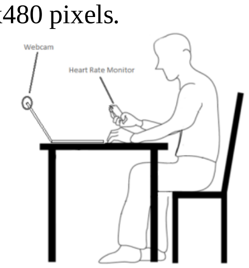

Fig. 1: Experimental Setup

The subject was placed at a distance of 1 meter away from the webcam with bright light as the source of illumination.As the video was being recorded,the heart rate of the subject was noted down manually using a heart rate monitor for every 10 sec interval.
The pulse wave initiated by the heart beat travels through the whole arterial system and reaches the face causing short term volume changes in the blood.
The intensity of the absorbed light depends on the volume.However this recorded video date has no clean RGB signals but has a mixture of illumination data along with noise and motion artifacts.

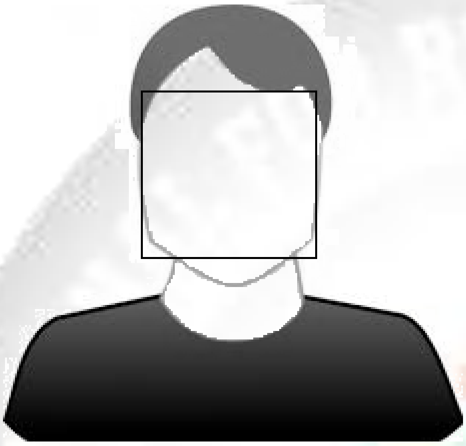

Fig. 3: Face detection

The processing and analysis of the video data is done using MATLAB R2013a.Firstly,the ROI is selected using the face detection algorithm published by MATLAB.

III. Algorithm

1) Consider the 50th frame of the video to avoid noise artefacts and perform the Laplacian pyramid decomposition.
2) Filter each level of the pyramid using a band pass filter with cut-off frequency [(hifreq*framerate/samplingrate)±1]
3) Magnify the intensities of the frequency band of interest by a magnification factor.
4) Collapse the pyramid into a single frame.
5) Repeat the above steps for the remaining frames.

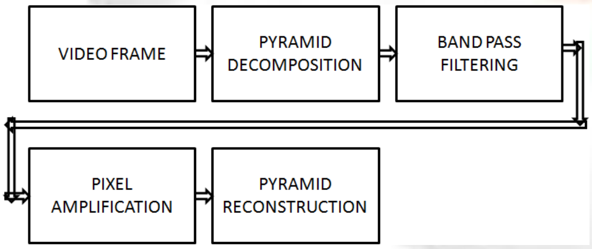

Fig. 4: Block diagram

IV. Results

The results were generated using MATLAB R2013a.The computation time per video was in the order of a few minutes. The Face Detection algorithm was implemented. The Heart rate generated by the code was compared to that of the HRM and the results were obtained as shown in Table 1.Figure 3 shows the PPG signal corresponding to the red, green and the blue component after reconstructing the Pyramid.
After pulse detection,the numbers of peaks corresponding to the various 10 second windows were found as shown in Table 1.
The Heart Rate was found using the formula
Heart Rate=6 * Number of peaks in the 10second window.

Time	Heart rate Using a HRM	Calculated Heart Rate	Percentage Error
10^{th} second	74	72	2.7%
20^{th} second	68	66	2.9%

After reconstructing the pyramid,the graph of the average intensity of each frame versus time/frame number was plotted and a peak detection algorithm was used to note the position of the r peaks.

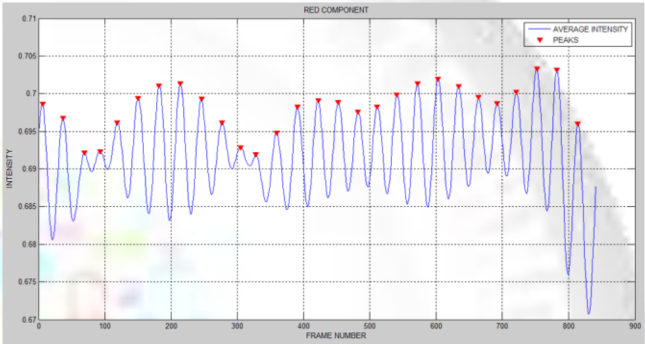

Fig. 4: Plot of average intensity vs frame number

The number of peaks was found in each 10 second window which corresponded to 34 frames and the result was compared to that found by the Heart Rate Monitor (HRM).

References

[1] Hao-Yu Wu, Michael Rubinstein, Eugene Shih, John Guttag, Fredo Durand, and William Freeman, "Eulerian Video Magnification for Revealing Subtle changes in the world."ACM Trans. Graph.(Proceedings SIGGRAPH 2012).
[2] T.Pursche,JKrajewski, and Reinhard Moeller, University ofWuppertal,Germany "Video base Heart Rate Measurement From Human Faces" IEEE International Conference on Consumer Electronics 2012.
[3] Magdalena Lewandowska,JacekRuminski,Tomasz KocejkoJedrzej Nowak "Measuring Pulse Rate with a webcam - a non-contact method for evaluating cardiac activity". Proceedings of the Federated Conference on Computer Science and Information systems pp.405-410. IEEE 2011.M.Z. Poh, D.J. McDuff, and R.W. Picard. Advancements in noncontact,
[4] Multiparameter physiological measurements using a webcam. Biomedical Engineering, IEEE Transactions on, 58(1):7–11, 2011.
[5] JawadNagi,SyedKhaleel Ahmed and FarrukhNagi "A MATLAB based Face recognition system using Image Processing and Neural Networks".4th International Colloquium on signal processing and its applications, march 7-9 2008,Kuala Lumpur,Malaysia.
[6] SungjunKwon,Hyunseok Kim and Kwang Suk Park, "Validation of heart rate extraction using video imaging

on a built in camera system of a smart phone" 34th Annual International Conference of the IEEE EMBS San Diego, California USA 1 Sept 2012