



INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

WEBCAM BASED HEART RATE EXTRACTION

LARISSA CARVALHO, H. G. VIRANI

Dept. of Electronics and Telecommunication Engineering, Goa College of Engineering,
Farmagudi, Ponda, Goa.

Accepted Date: 27/02/2014 ; Published Date: 01/05/2014

Abstract: In this project I have implemented a Non Invasive Heart Rate monitoring system to monitor subjects using Digital Image Processing. Here the main focus is to reveal variations in videos that are not possible to be seen with the naked eye. A standard video of the subject is taken as input which is followed by pyramid decomposition and filtering of the frames. The resulting signal is then amplified at pixel level 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.

Keywords: Heart Rate, Pyramid Decomposition, Magnification factor, Filter, Restoration.

Corresponding Author: MS. LARISSA CARVALHO



PAPER-QR CODE

Access Online On:

www.ijpret.com

How to Cite This Article:

Larissa Carvelho, IJPRET, 2014; Volume 2 (9): 469-476

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, 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 $\lambda = 550$ nm, because it is the approximate wavelength of the human eye's greatest sensitivity and $D = 3$ mm.

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 $\Delta l \approx 5 \times 10^{-6}$ 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 $\Delta l = 0.94 \mu\text{m}$. This means that a typical detector with pixel size approximately $5 \mu\text{m}$, 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 $\Delta l \approx 10 \mu\text{m}$ 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 $\theta = \Delta l/D$.

This estimate leads to the result, $\theta \approx 6.7 \times 10^{-4}$ rad which is approximately 2' 17". This means that the Eulerian video magnification algorithm does not depend on cameras capabilities to reveal unseen motion, but on other principles.

I. 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 in front 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

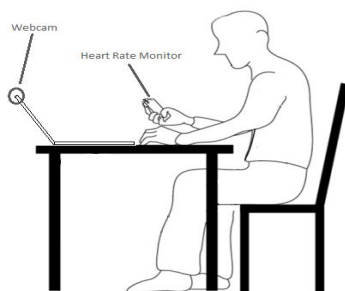


Fig. 1 Experimental Setup

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

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 data has no clean RGB signals but has a mixture of illumination data along with noise and motion artifacts.

The processing and analysis of the video data is done using MATLAB R2013a.

II. Algorithm

1) Consider the 50th frame of the video to avoid noise artifacts and perform the Gaussian pyramid decomposition.

2) Filter each level of the pyramid using a band pass filter with cut-off frequency $[(\text{hifreq} * \text{framerate} / \text{samplingrate}) \pm 1]$

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

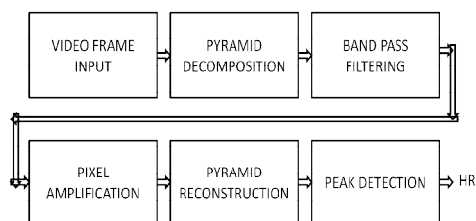


Fig 2:Block Diagram

Pyramid Decomposition

Pyramids can be used to accelerate search algorithms, i.e. in pattern finding or to perform multi-resolution blending operations. When constructing the BURT and ADELSON's pyramid, we do the following

- Blur and subsample the original image by a factor of two.
- Store this in the next level of the pyramid.

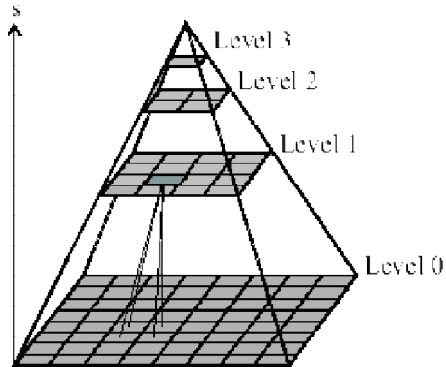


Fig 3. Pyramid Decomposition

To compute the pyramid we first interpolate the lower resolution image to obtain a reconstructed low pass version of the original image. We then subtract the low pass version from the original image.

The Gaussian pyramid is computed by convolving the image frame with a Gaussian kernel as shown in figure 6. The output image is a low pass filtered version of the original image. Its cut-off frequency can be controlled. The pyramid can be found by taking the difference between the original and the lowpass filtered image. Thus we get a set of band pass images. Thus a pyramid is a set of band pass filters.

	1	4	7	4	1
	4	16	26	16	4
$\frac{1}{273}$	7	26	41	26	7
	4	16	26	16	4
	1	4	7	4	1

Figure 4: Gaussian Kernel

Implementation:

If $I(x,y)$ is the original image, then the Gaussian pyramid on the image I is given as

$$G_0(x,y)=I$$

$$G_{i+1}(x,y)=REDUCE(G_i(x,y))$$

Where, the REDUCE is the convolution of the image with a low pass filter. The center pixel of the pixel mask has the highest weight and the sum of the weights in the mask is 1.

The Gaussian kernel is given as

$$W(r,c)=w(r) * w(c)$$

$$\text{where, } w(r) = (1/4 \quad -a/2 \quad 1/4 \quad a \quad 1/4 \quad 1/4)$$

and a is chosen in the range 0.3 to 0.6.

Filtering

This step converts the time series to the frequency domain using the Fast Fourier Transform and the band pass filter is applied to the signal.

This method is used to amplify the temporal variations in a video.

Pixel Amplification

After getting the frequency band of interest, we need to magnify and amplify it and add the result to the original image.

We then multiply the extracted band pass signal by a magnification factor whose value can be specified by the user. Then we add the magnified signal to the original and collapse the pyramid to get the final frame.

Image Restoration

After amplifying the pixels and adding back to the original image to get the output frame, collapse the pyramids into a single image per frame. We can change the amplification factor to obtain different results. We can also pass the amplified signal through a low pass filter to reduce effects of the high frequency components, such as borders

III. 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 for 5 subjects. 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 7 shows the PPG signal corresponding to the red component after reconstructing the pyramid.

After pulse detection, the number of peaks corresponding to the various subjects were found as shown in Table 1. The Heart Rate was found using the formula

$$\text{Heart Rate} = \frac{1800 \times \text{Number of peaks}}{\text{Number of frames}}$$

TABLE I.

Time	Using HRM	Calculated Heart Rate
1 st subject	74	72
2 nd subject	68	66
3 rd subject	80	75
4 th subject	60	65
5 th subject	150	144

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

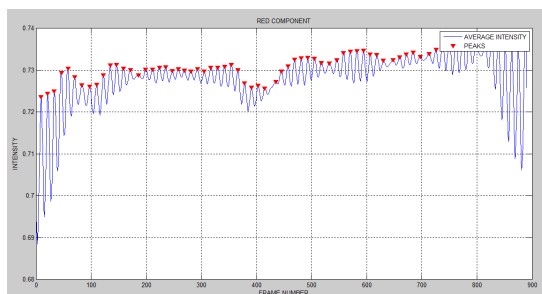


Figure 5: Plot of average intensity vs frame number of 5th subject

CONCLUSION

This Thesis presents the implementation of a non invasive, non contact system which reveals temporal variations in videos which was presented at SIGGRAPH 2012. Using this system we are able to determine the pulse rate of an individual through non contact means.

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, J. Krajewski, and Reinhard Moeller, University of Wuppertal, Germany "Video base Heart Rate Measurement From Human Faces" IEEE International Conference on Consumer Electronics 2012.
3. Magdalena Lewandowska, Jacek Ruminski, Tomasz Kocejko, Jędrzej 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.
4. Raphael C. Gonzalez, Richard E. Woods and Steven L. Eddins Digital Image Processing using Matlab 1st edition.