Pulse Rate Estimation

Computer Vision Project
Aneri Sheth (1401072)
Janki Desai (1401103)
Rupande Shastri (1401102)

***Abstract*—Human Pulse Rate can be estimated using several techniques ranging from direct measurements (contact-based) to non-contact based approaches. Over the past few years, a computer vision assisted non-contact based approach has evolved significantly for two reasons: limitations of contact based approaches (for example, the patient should be physically present, et cetera) and use of pulse rate information for non-clinical purposes. However, existing computer-vision guided techniques have not been benchmarked and the improvements are still required. A technique called Photoplethysmography (PPG) can be carried out using facial video recording from a laptop/smart phone camera. We explore this technique (PPG) and describe an algorithm that can estimate heart beat of a person from facial video recording.**

***Index Terms*—Photoplethysmography (PPG), Blood Volume Pulse (BVP) Independent Component Analysis (ICA), Frequency domain signal processing**

I. Introduction

Non - contact pulse rate estimation is important in today's scenario to avoid difficulties faced in contact based estimation (separate device needed, inconvenience/discomfort to patients)According to camera guided techniques, non-contact pulse rate estimation can be broadly classified into two classes:

1) Image-guided Techniques - They mostly use the intensity levels of different color channels over time to build the feature trajectory to be fed into some statistical model for PR estimation.
2) Motion-guided Techniques - They make use of tracking key points over time to estimate the feature for pulse estimation.

In this project, we are using image-based technique - extracting PPG signal by analyzing periodic changes in skin color. PPG is a non-invasive optical technique of determining pulse rate by sensing the cardiovascular pulse wave or the blood volume pulse (BVP) through variations in transmitted or reflected light from the skin. According to the Beer-Lambert law, the change of blood volume in the tissue is reversely proportional to the intensity of the light reflected or transmitted from skin. This is hardly seen by the naked eye, but it can be captured by the commercial camera. We take a video with a patient sitting in front of the laptop and the frame rate is 15fps. From the given video, face (region of interest)is detected using Viola Jones face detection method and then we extract RGB components from the face and perform Independent Component Analysis to de-correlate the three vector components.Then, we pass the features extracted after statistical analysis through thresholds, to remove outliers. The filtered features undergo Fourier Transform. In the frequency domain we construct a Power Spectral Density (PSD) plot of the signal and the peak of the PSD is estimated to be the BVP value.

II. Literature Review

The algorithm that we have used for our project is primarily based on two papers, both of which describe the techniques for non-contact based PR Estimation.
Non contact, automated cardiac pulse measurements using video imaging and blind source separation, by M. Poh, D. McDuff, and R. Picard, in 2010. This novel approach can be applied to color video recordings of the human face and is based on automatic face tracking along with blind source separation of the color channels into independent components. Using Bland-Altman and correlation analysis, we compared the cardiac pulse rate extracted from videos recorded by a basic webcam to an FDA-approved finger blood volume pulse (BVP) sensor and achieved high accuracy and correlation even in the presence of movement artifacts. This is the first demonstration of a low-cost accurate video-based method for contact-free heart rate measurements that is automated, motion-tolerant and capable of performing concomitant measurements on more than one person at a time.
Real-time Quantifying Heart Beat Rate from Facial Video Recording on a Smartphone using Kalman Filters by Wen Jun Jiang,, Shi Chao Gao, Peter Wittek and Li Zhao, 2014. Thy extract the green channel from the video. Then they normalize it and use a Kalman filter with a particular structure to eliminate ambient noise. This filter also enhances the heart pulse component in the signal distorted by Gaussian noise and white noise. After that they employ a band-pass FIR filter to remove the remaining motion artifacts. This is followed by peak detection or Lomb periodogram to estimate heart rate. The algorithm has low computational overhead, low delay and high robustness, making it suitable for real-time interaction on a smart phone.
Both the papers make use of Video-based monitoring methods which are largely suitable for nonclinical purposes. The common method is extracting PPG signal and performing statistical analysis. We make use of some techniques and designed the algorithm for pulse rate estimation.

III. Algorithm

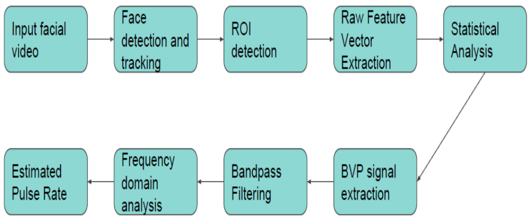


Figure 1 - Basic Flow of the Algorithm

The algorithm has three main parts:

1) Face Detection and Region of Interest(ROI) extraction - Viola Jones is implemented for face detection over a given video sequence and for pulse rate estimation, the region of interest extracted is the entire face (ideally from forehead to chin). The face is tracked oin each subsequence frame using KLT Algorithm
2) Feature Extraction and Independent Component Analysis (ICA) - The underlying source signal of interest is the cardiovascular pulse wave that propagates throughout the body. Volumetric changes in the facial blood vessels during the cardiac cycle modify the path length of the incident ambient light such that the subsequent changes in amount of reflected light indicate the timing of cardiovascular events.By recording a video of the facial region with a webcam, the RGB color sensors pick up a mixture of the reflected plethysmographic signal along with other sources of fluctuations in light due to artifacts such as motion and changes in ambient lighting conditions. The algorithm ICA defines a generative model for the observed multivariate data. In the model, the data variables are assumed to be linear mixtures of some unknown latent variables, and the mixing system is also unknown. The latent variables are assumed non-gaussian and mutually independent, and they are called the independent components of the observed data.The video frame is processed and the R,G and B components are extracted. They are assumed to be the linear combination of observed signal. The aim is to find a demixing matrix that can de-correlate the three signals. Out of the three signals, we use the green trace for the BVP Value.
3) Frequency domain signal processing - For further processing, we perform Fourier Transform on the filtered features and construct Power Spectral Density Plot and we show that the peak of the PSD is the estimated BVP Value.

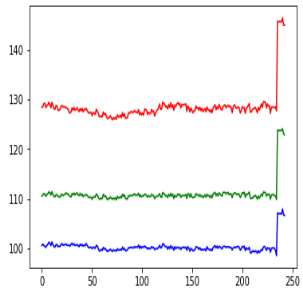
IV. Implementation and Results

The algorithm was implemented on different videos of 16 seconds for three different subjects.The resolution was 848x480 pixels. We choose forehead, cheeks and the entire face as region of interest and compared the results in the table below.

Original bpm	Full face	Cheek	Forehead
102	101	98	104
78	82	82	74
75 (beard)	76	95	76

Figure 2 - Comparing BMP obtained using different approach

In the approach where the entire face is taken as the region of interest we've taken 60 % as ROI from the detected face. Next step is to extract RGB traces of the frames which are displayed in Figure 3. We pre-process the video by applying moving average to smoothen the frames. We obtain three signals after statistical analysis using ICA as shown in Figure 4.



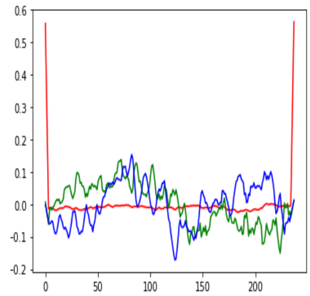


Figure 3 Figure 4

Lastly, we filter the individual frames using bandpass filters and analyze them in frequency domain. After obtaining the Power Spectrum Density we get the strongest frequency as our BPM measure.

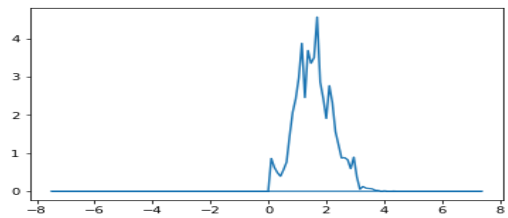


Figure 5 - Power Spectral Density

V. Future Work

1) We can use video tracking instead of frame by frame processing as it would computationally less expensive.
2) Better segmentation techniques can be used to detect Region of Interest such as forehead or cheeks.
3) Ambient Noise Removal in the range of 0.75 to 4 Hz

VI. Conclusion

Here is a non-contact video-based technique to measure heart rates. On comparing different parts of the face as ROI, we found that the full face works best. However, it depends on the amount of skin in the image as well, so people with facial hair (beards) and prominent foreheads will give better/similar results on the forehead. Also, people with prominent cheeks and small foreheads will give better results in cheek. The idea to define an ROI which has more skin was to remove motion artefacts caused by blinking or movement of the lips. But we see that the movement has mostly been taken care of by smoothing the signal and bandpass filtering.

References

[1] M. Poh, D. McDuff, and R. Picard, Non contact, automated cardiac pulse measurements using video imaging and blind source separation, 2010

[2] Real-time Quantifying Heart Beat Rate from Facial Video Recording on a Smartphone using Kalman Filters, Wen Jun Jiang,, Shi Chao Gao, Peter Wittek and Li Zhao, 2014

[3] Viola Jones https://en.wikipedia.org/wiki/Viola

[4] ICA https://www.cs.helsinki.fi/u/ahyvarin/whatisica.shtml

[5] http://iopscience.iop.org/article/10.1088/0967-3334/28/3/R01/meta

[6] http://neurosky.com/2015/01/ecg-vs-ppg-for-heart-rate-monitoring-which-is-best/