Image Analysis on Fingertip Video To Obtain PPG

Ayan Chatterjee and Andreas Prinz

Department of ICT - eHealth, The University OF Agder, Grimstad, Norway.
*Corresponding author E-mail: ayan1.c2@gmail.com

http://dx.doi.org/10.13005/bpj/1554

(Received: 06 November 2018; accepted: 14 December 2018)

Cardiovascular problems are evolving as the chief cause of death worldwide. Heart Rate, Blood Pressure, Respiratory Rate, Oxygen Saturation, Systolic Upstroke Time, Heart Beat duration, Diastolic time, RR interval are some important physiological parameters that help to monitor our daily health condition. Those parameters are very useful to determine if a person is suffering from any cardiovascular problems or not based on daily data collection and monitoring over a certain period of time and in this context machine learning algorithms will be very helpful for developing a smart cardiovascular tele-monitoring & recommendation system for better lifestyle. Irregularities in the heart signal can pop up a serious indication for upcoming cardiac problem. Here, we have concentrated on intensity variation based heart rate calculation process from PPG with major analysis on captured contact video. Here we have used normal handy smart phone camera which is available to everyone and able to capture fingertip videos of flowing blood in the vessels with visible light wavelength. In this paper, we have performed analysis on captured videos for accurate health parameter capturing and compared it with standard devices (FDA approved).

Keywords: PPG; Sampling; Intensity; Histogram; Filter; AC&DC component.

PPG signal can be described as non-invasive, inexpensive, portable and convenient tool for calculating Oxygen saturation (SPO2), blood pressure (systole/diastole), Heart Rate, Arterial age, vascular disease etc. During the capture, red contact video of fingertips, miss touch errors can produce significant variation in real time results due to noise. Similarly, high pressure on camera and low pressure on camera can produce incorrect pixel intensity and therefore, produces incorrect result. They can be treated as noise and needs to be removed up to a level to keep up the originality of a signal to give correct heart rate.

PPG signal can be measured with conventional pulse oximetry devices or smart phone camera. Here in our experiment we have used smart phone camera to capture fingertip red image video for 30 sec. on average in visible wavelength (400-700 nm) under normal physical & environmental condition for obtaining brightness signal (based on intensity variation) to plot and analyze the PPG signal. Infrared is a special form of invisible electromagnetic wave that can be felt by heat. For better picture, IR radiation is blocked by film or filter (camera glass). If no IR filter, pictures will not be as good as digital camera. So, it is a challenge to determine SPO2 using mobile camera video as for this both the Red & IR light is needed.

Here, in our study we have performed analysis on captured video to obtain good quality PPG signal. Using this analysis we can tell if the captured image is a suitable choice for plotting the brightness signal or not! We have
also compared the result with FDA approved oximetry devices to draw the conclusion. There are few challenges with PPG like a. different smartphones lead to different color saturation in the image b. motion artifacts due to shifting of fingers c. different fingertip pressure on the lens d. different environmental condition (dark / light) e. ROI (region of interest selection) in a frame f. Total frame number to be processed out of total count f. LED/Flash is used or not!

**Video Capturing & PPG**

The captured reflected light at camera sensor has two components a. AC or pulsatile component of artery b. DC or non-pulsatile component due to tissue background, venous blood and constant part of arterial blood flow. This AC & DC componentsof both Red & IR are used to measure oxygen saturation in the blood. Oxygen saturation is nothing but a ratio in between oxyhemoglobin to the amount of hemoglobin present in the blood. Oxygen combined with hemoglobin to form the compound to carry oxygen in blood. It’s normal range is in between 95-100.

The amplified waveform of brightness signalis used to measure heart rate, systolic point, diastolic point etc.

**PPG Capturing with Mobile Camera**

PPG is a form of cardiac pulse which is measured optically. It reflects the blood movement in the vessels that travels from heart to fingertips and toes through the vessels in the wave like motion. Here, in our experiment it is measured with average intensity variation method. There is another method namely curve-line movement based method. In intensity variation, we have measured average intensity variation of every frame in red channel. For systolic peak the corresponding image intensity is high & for diastolic peak it is low. In this experiment, we have shown why not to choose green or blue channel for measuring average pixel intensity!

Every smartphone camera has an optical sensor. Introduction of smartphone & tabs to healthcare segment has a potential to improve healthcare & thereby reduce healthcare costs. Latest smartphones have advanced powerful hardware, software, better memory & storage, well connectivity to cloud storage and backup option, advanced OS, embedded sensors got diverse, they are fast and do have long running batteries. The quality of good PPG signals depends on the
properties of the skin area, camera characteristics, light emitting diode (LED) usage, and light conditions. A good PPG signal can help to predict accurate Heart Rate, BP, SPO2 & Respiratory rate.

In our experiment, left & right hand fingertips are held on smartphone camera lens with flash for 30 sec & without flash for 30 sec without creation of additional pressure and captured video is stored in the server for further analysis.
in MATLAB. PPG data can be considered of two types - a. Reflective Mode (Video with flash) b. Transmission mode (Video without flash). Data has been collected with two different smartphones Moto-G Turbo & Samsung J7 Next Gen. in this project. FPS of mobile camera is also an important parameter for capturing video and experiment has shown its range in between 25-30 FPS and we have used same data as sampling frequency for corresponding video.

The samples as described in X axis of above pictures, are PPG cases captured under different circumstances (with camera flash on) like normal Light, normal Light with external

Fig. 6.1. Red Channel Brightness Signal (Average Pixel Intensity vs Length) [Chatterjee et al.]

Fig. 6.2. Blue Channel Brightness Signal (Average Pixel Intensity vs Length) [Chatterjee et al.]

Fig. 6.3. Green Channel Brightness Signal (Average Pixel Intensity vs Length) [Chatterjee et al.]
Sound, dark condition, dark & mobile heat, normal Walking. In all cases with respect to both the cameras, the fps lies in between 25-30 range on average. The captured image is in RGB format so every pixel consists of R (0-255), G (0-255), B (0-255) values. For proper touch measurement, we have used a. average intensity count of frame in red channel b. histogram analysis for red, green and blue channel c. Rules and Values for different parameters d. Gray scale image conversion and analysis (binary method). Under histogram analysis, if LED is turned on, then G and B channel intensity values are lower half of the histogram means less than 128. If LED is turned off, then G and B channel intensity values are close to zero in histogram. Hence, method of use, whether the LED is used or not can be detected with above rules.

If more contact pressure is produced on camera:

a. Deforming arterial wall leads to wrong reading
b. Block micro circulation in the capillary
 c. Sudden up & down of pixel intensity.
d. After FFT analysis of brightness signal flat line (Y=constant) is produced

The process of heart rate measurement can be shortened as:

Video Signal à Calibration Check à Average Pixel Intensity of red channel for first 660 frame à Analysis of image in the video à Brightness Signal Obtaining à Set Filter stabilization time = 2 & initial signal cutoff = 0 à Band Pass Butterworth Filter à Select WINDOW_START = 6, WINDOW_END = 12 out of 128 frames (FFT data point) from total frame count & select FFT window as 128 à For each set perform FFT à Leakage reduction à FFT Analysis à For each set calculate local maxima à For each set calculate HR à Calculate average HR

Assumption
Here, Experimental purpose: 40 – 220 bpm = .667 to 3.66 Hz
Sampling Frequency = 2* 3.66 Hz = 7.32 Hz (Nyquist)< 8 Hz
N.B: DC component operates at 0Hz and low frequency noise up to 0.5 Hz so filtering at above range will help PPG to get rid of DC & Low frequency noise. Higher the BP, shorter will be the duration of every heartbeat.

Choice of Channel And Histogram analysis [6, 17, 18]

We have picked up one fingertip video when body is at rest and selected its first 90 frames for analysis. We have obtained there brightness...
Fig. 8. Initial Bad Frame Capture (Average Pixel Intensity vs Length) [Chatterjee et al.]

Fig. 9.1. Histogram analysis & raw brightness signal for signal-1 [Chatterjee et al.]

Fig. 9.2. Histogram analysis of green & blue intensity for signal-1 [Chatterjee et al.]
signal of their corresponding average pixel intensity calculation / frame on red, green & blue channel. Here we have retained original frame size 1920 * 1080.

The result is showing that only red channel brightness signal is matching to the heart pulsatile signal properties due to uniform distribution of pixel intensity across red channel. Basic convention is that for average pixel intensity >= 200 is a good touch and we have done histogram analysis for R, G & B of the frames for further image analysis. Blue channel pixel intensity can be used for measuring AC & DC component for IR wave due to certain similarity so we can use it to measure oxygen saturation along with Red AC & DC component.

Normal smartphone has few problems based on the categories:
1. While camera is on for long time, heat is produced and that can be felt by touching for few headsets.
2. Camera stabilization takes some time, thereby sometimes for initial 1-3 seconds, bad images of contact fingertips result in and for red channel those images have average pixel intensity of >250-255 as depicted below:

**Observation on Moto-G Turbo & Samsung J7 Next Gen and noisy image**

Bad Frame with average pixel intensity > 254-255 on red channel:

So, it shows that images of captured video depends on the quality of the camera and circumstances. In case of Moto-G Turbo, systole and diastole can be captured correctly for those frames where average pixel intensity > 200 – 250 on red channel. Here frames with average pixel intensity > 250 – 255 on red channel, shows incursion of noise. So, we have to remove initial few of captured frames from full video (28-32 sec) that lies in the range of > 250-255. They are basically non-proper red color picture. We can’t remove a frame with mistouch in the middle of frame sequence as it might cause missing of some

![Fig. 10.1. Histogram analysis & raw brightness signal for signal-2[Chatterjee et al.]](image1)

![Fig. 10.2. Histogram analysis of green & blue intensity for signal-2[Chatterjee et al.]](image2)
important point so may lead to incorrect reading. It’s better to take reading again if any mistouch observed. For Samsung J7 Next Gen, systole and diastole can be captured correctly for those frames where average pixel intensity > 200 – 254.6 on red channel. Sometimes, noise can be observed in the range of >254.6 to 255.

The above diagram (left diagram: captured with Moto-G Turbo & right diagram: captured with Samsung J7 Next Gen.) shows that instability in the initial few frames (1-3 seconds). So, in the final algorithm we have removed initial few of the captured frames from full video (28-32 sec) and restricted the number to first 660 frames for all the videos (around > 20 sec. of capture).
Average / mean intensity in red channel of image captured by Samsung J7 Next Gen is higher than image captured by Moto-G Turbo. So, creation of a threshold seems to be very difficult as it depends on video capture. More detail study has been described in the calibration section of this paper.

Histogram analysis is used represent frequency of occurring of pixel intensity in the

Fig. 12.2. Histogram analysis of green & blue intensity for signal-4 [Chatterjee et al.]

Fig. 13.1. Histogram analysis & raw brightness signal for signal-5 [Chatterjee et al.]

Fig. 13.2. Histogram analysis of green & blue intensity for signal-5 [Chatterjee et al.]
whole data set. It is very useful for image analysis (correct region identification, outlier detection, calibration detection, identifying PPG image pattern & noisy pattern) as depicted below:

In the below diagrams, the green line is showing standard deviation & red line is showing the mean. The left side diagram is showing histogram analysis of red channel average intensity for the complete captured raw video before preprocessing. The right side diagram is showing corresponding brightness signal obtained. The standard deviation and mean can be used for outlier detection.

For captured fingertip video flash in Moto-G Turbo set, intensity variation is close to 1.5 - 5. It is observed that, continuous flash produced some heat and felt while touching the camera during multiple capture. In the next paper we will be discussing if any noise incorporated in the image for this heat generation and sweating.

For captured fingertip video flash in Samsung J7 Next Gen set, intensity variation is < 5.

For captured fingertip video without flash, intensity variation is abrupt (more than 5) and not fit for calculating heart rate. For contact fingertip videos without flash, histogram analysis shows blue and green channel intensity per frame shifts left and for blue channel the range becomes > 0 &< 10 and for blue channel it ranges in between > 0 &< 1. From this pattern and comparing same to the

### Table 1. Signal Analysis Report [Chatterjee et al.]

<table>
<thead>
<tr>
<th>Mean</th>
<th>Standard dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>237.5845</td>
<td>3.3690</td>
<td>229.2114</td>
<td>255</td>
</tr>
<tr>
<td>243.4559</td>
<td>4.8610</td>
<td>236.4386</td>
<td>255</td>
</tr>
<tr>
<td>246.6277</td>
<td>1.9523</td>
<td>241.8376</td>
<td>255</td>
</tr>
<tr>
<td>246.4072</td>
<td>1.9284</td>
<td>237.2134</td>
<td>254.8942</td>
</tr>
<tr>
<td>246.2110</td>
<td>1.7492</td>
<td>241.6756</td>
<td>255</td>
</tr>
</tbody>
</table>

Captured by Samsung J7 Next Gen for a 31 years of healthy man with flash:

**Fig. 14.1.** Histogram analysis & raw brightness signal for signal-1 [Chatterjee et al.]

**Fig. 14.2.** Histogram analysis of green & blue intensity for signal-1 [Chatterjee et al.]
frames captured with flash on, we can predict if an image is captured with flash or without flash. When flash is on, from the above histogram analysis, blue and green intensity values for good frames are less than 128 or values are lower half of the histogram.

We have applied below rules to the captured videos and received true outcome in Boolean:

<table>
<thead>
<tr>
<th>When LED is Used</th>
<th>When LED is not used</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{mean}(G) - \sigma_G &lt; R_{\text{LED}}\text{min})</td>
<td>(\text{mean}(G) + \sigma_G &lt; G_{\text{NOLED}}\text{max})</td>
</tr>
<tr>
<td>(\text{mean}(R) - \sigma_R &gt; R_{\text{LED}}\text{min})</td>
<td>(\text{mean}(B) + \sigma_B &lt; B_{\text{max}})</td>
</tr>
<tr>
<td>(\sigma_R, \sigma_G, \sigma_B &lt; \sigma_{\text{max}})</td>
<td>(\sigma_R, \sigma_G, \sigma_B &lt; \sigma_{\text{max}})</td>
</tr>
<tr>
<td>(\text{mean}(B) + \sigma_B &lt; B_{\text{max}})</td>
<td>(\text{mean}(G) + \sigma_G &lt; G_{\text{NOLED}}\text{max})</td>
</tr>
</tbody>
</table>

Where, \(G_{\text{LED}}\text{min}= 13\) instead of 10 (used in other research paper) & other fixed values are as depicted above. [3,4,5]

**Calibration**

This is the step for measuring threshold for each video frame. Here we have considered 80%, 85% & 90% of first 200 red channel frames pixel intensity and thereby its mean to measure the threshold. This methodology can be used for all type of smartphone videos where \(H*W\) is not fixed.

\[ T = \text{mean}(T_i) \text{ where } T_i \text{ is the measured threshold for frame } i. \]

\(\text{“every pixel intensity in red channel } \times n\% = \text{Total average Intensity in red channel } \times n\%, \text{ Where } n = 80 / 85 / 90.\)

From the above table, result shows that for thresh-holding point (Tp) selection, we can choose first 200 frames & obtain 90% of the average intensity variation in the red channel for every frame and sum them up or obtain 90% of the
average intensity variation in the red channel for first 200 frames. The post condition is that it must be more than 200& less than minimum average intensity in red channel. It is independent of camera / selection.

**A rule can be established as below to improve the algorithm**

90% of the average intensity variation in the red channel > 200 &&< (mean(red Channel intensity) – 5 * standard deviation).

From histogram analysis, it can be concluded that good touch frames are generally in between (mean(red Channel intensity) ± 5 * standard deviation in red channel intensity).

**testing [6,7,11]**

For maintaining originality, we have manually captured red image video of fingertips
using mobile camera then used that raw *.mp4 file for analysis in MATLAB software. We have captured all videos in average voltage of light and at normal room temperature and at rest physical condition. Then the result is compared with two devices like Dr. Trust’s Pulse Oximetry (FDA Approved) as shown in Figure 4 [Ayan Chatterjee et al.] & portable wearable BP Machine by MEDISANA as shown in Figure 5 [Ayan Chatterjee et al.].

We have used below devices, software & hardware tools:

<table>
<thead>
<tr>
<th>Mean</th>
<th>Standard dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>252.8043</td>
<td>0.8014</td>
<td>250.7790</td>
<td>255</td>
</tr>
<tr>
<td>253.0314</td>
<td>0.7809</td>
<td>250.8493</td>
<td>255</td>
</tr>
<tr>
<td>254.5383</td>
<td>0.3534</td>
<td>252.9994</td>
<td>255</td>
</tr>
<tr>
<td>237.2475</td>
<td>3.6375</td>
<td>231.8594</td>
<td>255</td>
</tr>
<tr>
<td>254.8057</td>
<td>0.1369</td>
<td>254.3104</td>
<td>255</td>
</tr>
</tbody>
</table>

**Table 2. Signal Analysis Report**

[Chatterjee et al.]

Smart Phone: Moto G Turbo Edition
Android: 6.0.1
Camera: Rear
Software: MATLAB Signal & Image Processing Toolbox
Version: R2018b
Hardware: Acer Laptop, Intel Celeron Processor
Platform: Windows 10

We have taken below HR readings from Dr. Trust for 2 min in normal rest condition:

**Frame selection (ROI) [3,4,5,6]**

Here, in our study we have used frame size of 1920*1080 and used the total area to find average pixel intensity in red channel for calculating the brightness signal. As depicted below, a close look can find a dark red and less red region of a good touch fingertip image. The boundary is oval in shape and we can assume it is nothing but top part of our finger. So, dark part seems to be unnecessary and noise can be there. We are working on less red region to find a “ROI” of fingertip image which is actual and without potential noise. If we are able to...
find out proper “ROI” with area < 1920*1080 then it will rather improve time & space complexity. In the second picture, we have put a black line to distinguish dark and less dark region of a fingertip image with good touch.

We are trying to use different clustering algorithms, edge detection algorithms with different operators like Prewitt, Sobel, Robinson Compass Masks, Kirsch Compass Masks, Laplacian

**Fig. 18.2.** Histogram analysis of green & blue intensity for signal-5 [Chatterjee et al.]

**Fig. 19.1.** Histogram analysis & raw brightness signal for signal-1 [Chatterjee et al.]

**Fig. 19.2.** Histogram analysis of green & blue intensity for signal-1 [Chatterjee et al.]
The paper has given a detail analysis on fingertip red image and popped up new questions. Operator etc. with sharpening strategy to obtain “ROI”. So, it can be concluded as a part of future research paper.

**CONCLUSIONS**

The paper has given a detail analysis on fingertip red image and popped up new questions.

**Table 3.** Signal Analysis Report [Chatterjee et al.]

<table>
<thead>
<tr>
<th>Mean</th>
<th>Standard dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>30.8882</td>
<td>1.8652</td>
<td>25.5499</td>
<td>33.7479</td>
</tr>
<tr>
<td>37.9662</td>
<td>9.1182</td>
<td>23.4509</td>
<td>52.8612</td>
</tr>
</tbody>
</table>

**Fig. 20.1.** Histogram analysis & raw brightness signal for signal-1 [Chatterjee et al.]

**Fig. 20.2.** Histogram analysis of green & blue intensity for signal-1 [Chatterjee et al.]

**Fig. 21.** Pulse Oximetry Dr. Trust
and area where research progress can be made like ROI selection from fingertip image and finding effect of external circumstances on PPG video. A good capture will result in average intensity

![Heart Rate Reading with Dr. Trust](image1)

**Fig. 21.** Heart Rate Reading with Dr. Trust

![Heart Rate Calculation with Video Pre-Processing](image2)

**Fig. 22.** Heart Rate Calculation with Video Pre-Processing (660 frames & skipping initial 2 fps)

![Fingertip Image & Region Selection](image3)

**Fig. 23.** Fingertip Image & Region Selection

<table>
<thead>
<tr>
<th>Signal</th>
<th>80%</th>
<th>85%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moto-1</td>
<td>190.1648</td>
<td>202.0501</td>
<td>213.9354</td>
</tr>
<tr>
<td>Moto-2</td>
<td>200.0661</td>
<td>212.5702</td>
<td>225.0744</td>
</tr>
<tr>
<td>Moto-3</td>
<td>197.3136</td>
<td>209.6458</td>
<td>221.9779</td>
</tr>
<tr>
<td>Moto-4</td>
<td>197.3930</td>
<td>209.7300</td>
<td>222.0671</td>
</tr>
<tr>
<td>Moto-5</td>
<td>197.4792</td>
<td>209.8216</td>
<td>222.1640</td>
</tr>
<tr>
<td>Samsung-1</td>
<td>202.5047</td>
<td>215.1613</td>
<td>227.8178</td>
</tr>
<tr>
<td>Samsung-2</td>
<td>202.7144</td>
<td>215.3840</td>
<td>228.0537</td>
</tr>
<tr>
<td>Samsung-3</td>
<td>203.4036</td>
<td>216.1163</td>
<td>228.8290</td>
</tr>
<tr>
<td>Samsung-4</td>
<td>191.9096</td>
<td>203.9039</td>
<td>215.8983</td>
</tr>
<tr>
<td>Samsung-5</td>
<td>203.8867</td>
<td>216.6296</td>
<td>229.3725</td>
</tr>
</tbody>
</table>

**Table 4.** % of Intensity Analysis Report [Chatterjee et al.]
variation for every frame more than 200 in red channel. Average pixel intensity of 254-255 in red channel happens at the start of the video & lasts for 1-4 sec based on the situation and different set. So, it is better to avoid them. Average pixel intensity for red channel is used to consider for brightness signal due to uniform intensity distribution. Heart Rate Calculation with Video Pre-Processing (processing on initial 660 frames by skipping initial 2 fps) is a new addition to the algorithm. If any average pixel intensity for red channel is below 200 in the range of considered 660 frames then it’s better to recalculate after taking video again.

REFERENCES


