Frame Adaptive ROI for Photoplethysmography Signal Extraction from Fingertip Video Captured by Smartphone

Lai-Man Po[a], Xuyuan Xu[a], Litong Feng[a], Yuming Li[a], Kwok-Wai Cheung[b] , Chun-Ho Cheung[c]

[a] Department of Electronic Engineering, City University of Hong Kong, Hong Kong, China
[b] Department of Computer Science, Chu Hai College of Higher Education, Hong Kong, China
[c] Department of Information Systems, City University of Hong Kong, Hong Kong, China

ABSTRACT

Photoplethysmography (PPG) has been widely used in clinical applications for monitoring vital signs especially heart rate by pulse oximeter. Recent researches have demonstrated the possibility of using fingertip video based PPG approach to estimate heart rate by smartphones. However, due to the variation of camera sensor characteristics in difference smartphones, the conventional fixed region-of-interest (ROI) for PPG signal extraction technique is not reliable. In this paper, a novel frame adaptive ROI method is proposed to detour the color saturation or cut-off distortion in the fingertip video capturing process for improving the reliability due to variation and limited dynamic range of the camera sensors in different smartphone models. Experimental results demonstrate that the proposed method can produce good pulsatile waveform and achieve high heart rate estimation accuracy using different smartphone models as compared with a FDA (U.S. Food and Drug Administration) approved commercial pulse oximeter.

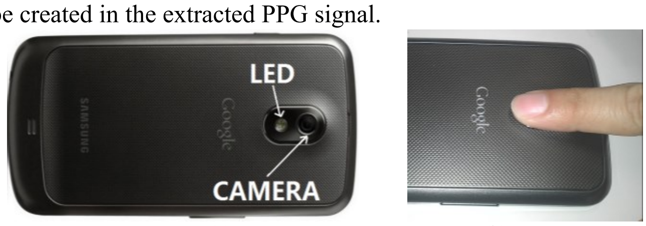
***Index Terms*—** Photoplethysmography (PPG); Pulse oximeter; Heart rate; Color saturation; Color cut-off.

I. INTRODUCTION

Recent report [1] has shown that self-measurement of vital signs is crucial for patients with chronic diseases, myocardial infarction or stroke. It is because vital signs such as heart rate, respiration rate, and blood oxygen saturation are important for timely detection of health conditions [2]. Photoplethysmography (PPG) has been widely used in clinical applications of monitoring heart rate by pulse oximeter. Basically, PPG is a noninvasive technique as human skin is illuminated by the light source and a photodetector is used to detect cardiovascular pulse wave that propagates through the body [3].

For most of people, the pulse oximeter is not the best device for performing timely self-monitoring of vital signs since they may not be willing to purchase and carry this special equipment [4]. Nowadays, smartphone has become the most portable, mobile and affordable communication device and people bring smartphones with them everywhere all the time. Besides its portability, mobility, connectivity and high computational power, smartphone usually has a built-in digital camera together with a light emitting diode (LED) or so called flashlight, as shown in Fig. 1(a). This hardware configuration is very suitable for reflection mode PPG to assess physiological signal using the smartphone camera. In which, video is captured by placing the fingertip over both camera and flashlight as indicated in Fig. 1(b).

A common approach of PPG signal extraction in fingertip video is to use a predefined region of interest (ROI) in each video frame for a particular color channel and then average the pixel values within the ROI to convert the 2-dimensional video signal into an 1-dimensional signal. In [5], Jonathan used a 10×10-pixel region located at the center of each frame as ROI and the green channel is supposed to provide a stronger signal than red and blue channels. Nokia E63 smartphone model was used in their study and they demonstrated the feasibility of using smartphone for heart rate measurement by frequency-domain analysis. Similarly, Blokhovsky [6] and Scully [7] both used central 50×50-pixel region as ROI to generate the PPG signal from fingertip videos. Bolkhovsky used the green channel for iPhone4S and red channel for Motorola Droid, while Scully used the green channel for Motorola Droid. However, Grimaldi [8] reported that the distribution of the pixels in the green channel is not uniform for different smartphone models such as Samsung, HTC HD2, iPhone4 and Nokia. This raised a main concern of the reliability and accuracy of PPG signal extraction from fingertip video with the use of different smartphone models due to the variation of camera characteristics. To improve the reliability, color saturation distortion is considered in [9] and PPG signal is calculated as a radius of the circle fitting the binarized image of red channel by an adaptive threshold. Unfortunately, the use of radius as the PPG signal sometimes may not be accurate in some occasions. It is because the change of radius could not be obvious and noise may be created in the extracted PPG signal.



(a) (b)

Fig. 1**:** (a) Camera configuration of smartphone (Model: Galaxy Nexus), and (b) Fingertip video capturing for reflection mode PPG using smartphone.

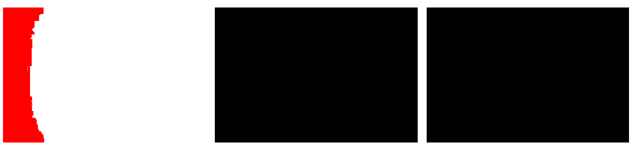
To tackle the problem of camera characteristic variation in different smartphone models, a new frame adaptive ROI technique is proposed in this paper to detour the color saturation and cut-off distortion in the fingertip video capturing process. Thus, signal distortion caused by the limited dynamic range of camera sensors can be minimized. This paper is organized as follows. In section 2, the problem raised from the color saturation and cut-off distortions in fingertip video is discussed. The new frame adaptive ROI method for PPG signal extraction is proposed in section 3. Experimental results are given in section 4 and followed by our conclusion in section 5.

978-1-4799-8391-9/15/$31.00 ©2015 IEEE

II. Colour Saturation and Cut-off Distortions of Fingertip Video

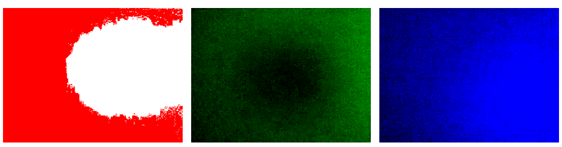
To improve the PPG signal extraction from fingertip video, we have to first understand the main characteristics of color saturation and cut-off distortions. Similar to the findings of color distortion in [10], due to the limited dynamic range of the camera sensors, color saturation and cut-off distortions are often found in the fingertip video. It is probably because the fingertip closely contact with the smartphone camera with the flashlight turned on, as illustrated in Fig. 1(b). It results in some regions may be too bright or too dark. Examples of red, green and blue channels of fingertip video frames that captured by smartphone cameras of Galaxy Nexus and LG Optimus P920 are shown in Fig. 2 and Fig. 3, respectively. From these figures, we can easily realize some regions of the video frames are too bright that over the maximum brightness of the camera sensors and saturation distortion is caused, i.e. the white regions in Fig. 2(a) and Fig. 3(a) are equivalent to the maximum intensity of 255 in red channel. The main feature of the saturation regions of video frames is that they are capped to the maximum level or 255 for an 8-bit dynamic range. Similarly, some regions of the video frames may be too dark, for example, the blue and green channels as shown in Fig. 2(b)-(c), and Fig. 3(b)-(c). This may be due to that human tissue in vivo of a fingertip almost covers the whole camera and the reflected light other than red is hardly detected and sensed by the camera sensors. These too dark regions are considered as cut-off distortions since their brightness are lower than the minimum sensitivity of the camera sensors. The main feature of the cut-off regions is that they are quantized to the minimum level of zero.

In order to save the computation, the conventional PPG signal extraction methods [5-7] use a small region of the video frame, e.g. 50×50-pixel ROI at center of the video frame. However, this fixed ROI is not reliable for different fingertip position and different smartphone models. First column of graphs in Fig. 4 shows the signals generated by averaging the central 50×50-pixel ROI of frames using the RGB color channels of Fig. 2. The extracted signal remains constant values of 255 or 0, and manifests the problem due to the color saturation or cut-off distortion. Second column of graphs in Fig. 4 shows the second example of PPG signal extracted from the fingertip video of Fig. 3. Since not all the central regions are in the color saturation or cut-off regions, PPG signals are possible to be generated in green and blue channel. However, these extracted signals are very noisy, and it is difficult for peak detection in heart rate estimation.



(a) (b) (c)

Fig. 2: Fingertip video frames (640×480) captured by Galaxy Nexus smartphone: (a) red channel, (b) green channel, and (c) blue channel.



(a) (b) (c)

Fig. 3: Fingertip video frames (320×240) captured by LG Optimus P920 smartphone (a) red channel, (b) green channel (magnify by 5 times), and (c) blue channel (magnify by 5 times).

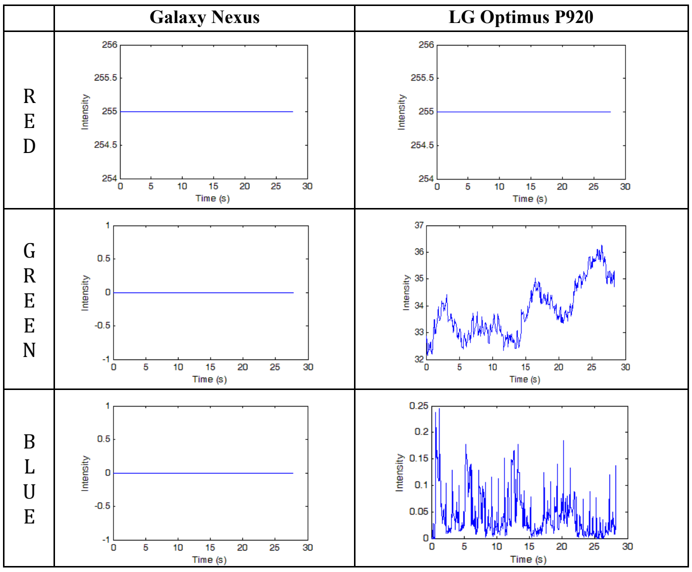


Fig. 4: Extracted PPG signals using Galaxy Nexus and LG Optimus P920 smartphone from central 50x50-pixel region of red channel, green channel, and blue channel.

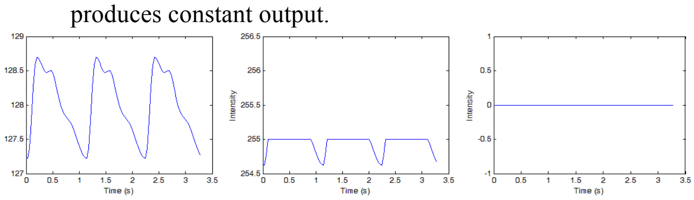
III. Frame Adaptive ROI for PPG Signal Extraction

Based on the color saturation and cut-off distortion phenomena in section 2, pixels of a specified color channel in fingertip video can be classified into three types:

(1) Full-PPG pixel: The temporal pixel intensity variation of the PPG signal lies within the dynamic range of the camera sensors. Thus, high quality PPG signal can be extracted from full-PPG pixel. An example is shown in Fig. 5(a), whereas neither saturation nor cut-off distortion are occurred.

(2) Partial-PPG pixel: The temporal pixel intensity variation of the PPG signal does not always lie within the dynamic range of the camera sensors. Sometimes the pixel intensities are saturated due to too bright, as depicted in an example of Fig. 5(b), or being cut off. This type of signals can significantly degrade the quality of PPG signal since the saturated and cut-off timeslots of partial-PPG signals are hardly to be same and may be varied largely.

(3) None-PPG pixel: The temporal pixel intensity is in either saturated or cut-off region, as shown in Fig. 5(c), and always produces constant output.



(a) (b) (c)

Fig. 5: The signals of (a) a full-PPG pixel, (b) a partial-PPG pixel with saturation distortion in some time sessions due to a large off-set, and (c) a none-PPG signal with cut-off distortion.

Based on this model, a PPG signal extracted from a color channel of a fingertip video using a specified ROI can be expressed as a combination of these three types of pixels:

$$p_c(t) = \frac{1}{D}\left(\sum_{(x,y)\in E_f} I_c(x,y,t) + \sum_{(x,y)\in E_p} I_c(x,y,t) + \sum_{(x,y)\in E_n} I_c(x,y,t)\right) \quad (1)$$

where c represents the color channel and it can be R (red), G (green), or B (blue) channel. $p_c(t)$ is the extracted PPG signal from color channel c by averaging the pixel intensity in the ROI, and D is the total number of pixels within the ROI. $I_c(x,y,t)$ is the pixel intensity of color channel c at pixel position (x,y) for frame t. E_f, E_p and E_n are the regions with full-PPG pixels, partial-PPG pixels, and none-PPG pixels, respectively.

Based on Eq. (1), high quality PPG signal can be extracted by using full-PPG pixels as ROI only. To determine full-PPG pixels as ROI, however, is not efficient and effective, as it needs to scan each pixel in temporal domain with sufficient duration for identifying full-PPG pixels. Moreover, the determined ROI may not be temporally stable due to slightly fingertip movement and change of lighting conditions. To tackle these problems, frame adaptive ROI is proposed, which uses full-PPG pixels and partial-PPG pixels that are not in saturation or cut-off levels to define the adaptive ROI in each frame. In the proposed method, red channel is used for PPG signal extraction, and the adaptive ROI only consist of full-PPG and partial-PPG pixels in the red video frame, which can be expressed as:

$$ROI(t) = \{(x,y) | 0 < I_R(x,y,t) < 255\} \quad (2)$$

where $ROI(t)$ denotes the ROI at frame t. Thus, the PPG signal for R channel at frame t can be calculated as

$$P(t) = \frac{1}{|ROI(t)|} \sum_{(x,y)\in ROI(t)} I_R(x,y,t) \quad (3)$$

where $|ROI(t)|$ is number of pixels within the $ROI(t)$. Fig. 6(a) shows the results of extracted PPG signals using Eq. (3) by averaging the full-PPG signal of Fig. 5(a) and the partial-PPG signal of Fig. 5(b) with exclusion of the saturation sessions. As the partial-PPG signal has a large offset, the generated PPG signal is greatly deformed due to offset in the unsaturated time sessions of the partial-PPG signal as Fig. 6(a).

To avoid the waveform deformation due to different offsets, the first order difference of the full-PPG signals and partial-PPG signals can be used to generate high quality PPG signal. It is because the first order differences of both full-PPG and partial-PPG signals are very similar in unsaturation or non-cut-off time sessions. To obtain the correct PPG signal waveform, frame adaptive ROI for PPG extraction is used and described as:

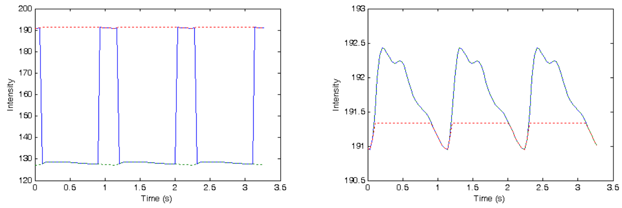
$$P(t) = P(t-1) + \frac{1}{|ROI(t) \cap ROI(t-1)|} \cdot \sum_{(x,y)\in ROI(t)\cap ROI(t-1)} \big(I_s(x,y,t) - I_s(x,y,t-1)\big) \quad (4)$$

where $t \geq 1$ and $P(0)$ is the first sample of the extracted PPG signal from the first frame of fingertip video, which is given by

$$P(0) = \frac{1}{|ROI(0)|} \sum_{(x,y)\in ROI(0)} I_R(x,y,0) \quad (5)$$

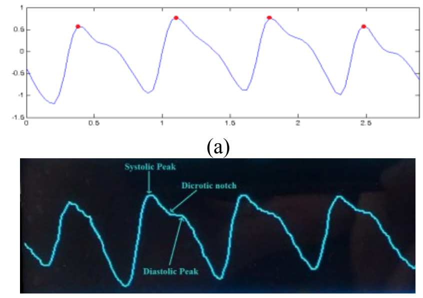
Basically, the sample of the extracted PPG signal is obtained by adding previous extracted PPG sample $P(t\text{-}1)$ with the average pixel intensity differences of the current frame pixels and the previous frame pixels within the intersection of the adaptive ROIs. Fig. 6(b) shows the extracted PPG signal using Eq. (4) and Eq. (5) from the full-PPG signal of Fig. 5(a) and partial-PPG signal of Fig. 5(b). In Fig. 6(b), the red dotted line is the partial-PPG signal and green dotted line is the full-PPG signal. The waveform of the extracted PPG is not affected by the offset from the partial-PPG signal. With this frame adaptive ROI, high quality PPG signal can be extracted. Its main advantage is that it can adapt to the change of ROI due to fingertip movement or other conditions such as background lighting or flashlight variations.

The extracted PPG signal from the fingertip video comprises a pulsatile physiological waveform, and a lower frequency component [11]. To extract the pulsatile waveform that related to heart rate, a 88-order equiripple finite impulse response (FIR) filter with bandpass frequencies ranging from 0.7 to 3.5 Hz is used to obtain the pulsatile signal. Fig. 7(a) shows an example of pulsatile PPG signal extracted by proposed method, which is shown for comparing with PPG signal in Fig. 7(b) that obtained at the same time by a FDA approved pulse oximeter EDAN M3 [12]. From Fig. 7(a), we can observe that the pulsatile PPG extracted by the proposed methods can maintain the characteristics of pulsatile waveform, systolic peak, dicrotic notch and diastolic peak, which is very similar to Fig. 7(b). These characteristics are very important for diagnosing the heart related disease based on the PPG signal [1].



(a) (b)

Fig. 6: PPG signals generated by (a) averaging the full-PPG signal of Fig. 8(a) and the partial-PPG signal of Fig. 8(b) with exclusion of the saturation sessions, and (b) the proposed frame adaptive ROI with these two signals.



(a)

(b)

Fig. 7: PPG signal extraction by (a) proposed method, (b) PPG signal from EDAN M3.

It is difficult to objectively measure the performance of the proposed frame adaptive ROI PPG signal extraction using waveforms comparison. However, the major application of fingertip video based PPG is for heart rate estimation. The accuracy of heart rate estimation is good for objective evaluation by comparing the results from pulse oximeter at the same time during the fingertip video capturing. To obtain the heart rate from the extracted PPG signal after the bandpass FIR filter, local peak detection method is used, which aims at finding the systolic peaks in the PPG signal as the red dots in Fig. 7(a). The systolic peak is the local maximum point for one cardiac cycle, which can be detected by finding the maximum value during one cardiac cycle of the PPG signal. After the peak detection, average heart rate can be calculated by average time differences of two systolic peaks. Using the systolic peaks for the calculation of heart rate can provide continuous and instant heart rate, as it requires a short period of the PPG signal.

IV. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed methods, the FDA approved vital signs monitor EDAN M3 [12] with PPG based heart rate monitoring function is used to obtain the reference measures. Left-hand index finger of volunteer is used for fingertip video capture by smartphone camera, while the right-hand index finger is used for EDAN M3 PPG measurement. Five smartphone models are used for our evaluation. They are Samsung Galaxy Nexus, LG Optimus P920, Samsung Galaxy S2, Samsung Galaxy Tablet 7.0, and Motorola Atrix.

PPG pulsatile signal is extracted for heart rate estimation using 20-second fingertip videos that captured from ten volunteers by different smartphones and analyzed by three well-known methods of Jonathan's method [5], Scully's method [7], Lamonaca's method [8], and the proposed methods. As Jonathan's method uses the frequency domain for heart rate estimation, the duration of captured video is kept at 20 seconds for maintaining consistence of all methods.

Table 1 shows the average heart rate and standard deviation (SD), i.e. (mean, SD), of different methods compared with the results from EDAN M3. For Galaxy Nexus smartphone, the green and blue channels are almost cut-off at the center of fingertip video frame. That is why the heart rate cannot be estimated by Jonathan's and Scully's methods as in Table 1. Table 2 shows the root-mean-square distortion (RMSD) of the estimated heart rates between different methods and the pulse oximeter. Jonathan's method performs well for Galaxy S2 with relatively lower RMSD compared with other smartphones. Inside the ROI, Scully has higher percentage of full-PPG pixels, and lower RMSD compared with Jonathan's method. Lamonaca's method has similar average RMSD to Scully's method, and can handle the saturation problem for red channel. However, its performance highly depends on the accuracy of the changing of radius for representation of the PPG signal. For some smartphone models, this change is not obvious and does not have an acceptable RMSD performance. From the last row of Table 2, we can find that the proposed methods can achieve lowest RMSD among all the compared methods. Moreover, the average RMSD is around 1.9, which is just around one-third of the Scully's method.

	Nexus	LG Optimums	Galaxy S2
EDAN M3	(79.21, 12.33)	(79.48, 13.92)	(75.01, 13.12)
Jonathan	NA	(76.84, 8.91)	(72.12, 13.43)
Scully	NA	(77.84, 9.52)	(73.69, 13.09)
Lamonaca	(79.03, 12.48)	(79.62, 11.91)	(72.81, 11.02)
Proposed	(77.60, 12.04)	(78.42, 13.61)	(73.11, 12.69)
	Galaxy Tablet	**Atrix**	**Average of all smartphone**
EDAN M3	(79.96, 12.30)	(72.46, 9.40)	(76.15, 11.69)
Jonathan	(82.35, 16.03)	(71.85, 10.96)	(75.31, 11.56)
Scully	(83.05, 12.52)	(74.90, 9.52)	(77.66, 10.47)
Lamonaca	(80.37, 12.81)	(73.17, 8.86)	(75.74, 11.17)
Proposed	(78.30, 12.53)	(70.81, 8.92)	(74.75, 11.52)

Table 1: (Mean, SD) of heart rate for different smartphones.

	Nexus	Optimums	Galaxy S2	Galaxy Tablet	Atrix	Average
Jonathan	NA	12.42	4.69	6.08	6.55	7.81
Scully	NA	7.94	3.48	4.99	4.46	5.77
Lamonaca	8.40	5.75	4.28	3.09	4.83	5.61
Proposed	2.31	1.39	2.02	2.19	2.07	1.92

Table 2: RMSD of heart rate for different smartphones compared with the pulse oximeter EDAN M3.

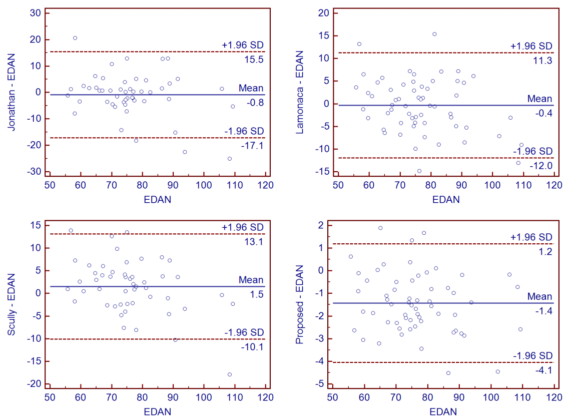


Fig. 8: Bland-Altman plots comparing agreement between the average estimated heart rate and the pulse oximeter EDAN M3.

Fig. 8 shows the Bland-Altman plots [13] of average estimated heart rate for different methods compared with measured results by the pulse oximeter EDAN M3. It compares two clinical measurements to show the errors, with the solid line representing the bias (difference between means) between the pulse oximeter and the compared methods. The two dash lines stand for 95% limits of agreement. Jonathan's method has the largest 95% limits of agreement -17.1 bpm and 15.5 bpm. While Lamonaca's and Scully's methods have similar 95% limits of agreement, which are -10.1 bpm and 13.1 bpm, and -12.0 bpm and 11.3 bpm. For the proposed methods, the 95% limited of agreement is between -4.1 bpm and 1.2 bpm, which shows that the proposed methods can achieve higher reliability in heart rate estimation among all the methods using different devices against the pulse oximeter EDAN M3, since the heterogeneous characteristics of cameras in different smartphone model are considered.

V. Conclusion

In smartphone fingertip video based PPG application, quality of PPG signal is strongly influenced by the heterogeneous characteristics of cameras in different smartphone models. To improve the reliability and accuracy of PPG signal extraction, new frame adaptive ROI is proposed, which can detour the color saturation and cut-off distortion in the fingertip video. In which, the first order difference of the full-PPG signals and partial-PPG signals are used to generate high quality PPG signal for avoiding waveform deformation due to different offsets. Experimental results demonstrate that the proposed methods outperform the three well-known fingertip video heart rate estimation methods, and have a higher reliability for heart rate estimation for different smartphones.

References

[1] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," Physiological measurement, 2007.

[2] A. Smith, T. Ferriss, "Home health monitoring may significantly improve blood pressure control," Kaiser Permanente Study Finds, http://xnet.kp.org/newscenter/pressreleases/co/2010/052110telehealthstudy.html, 2010.

[3] M. Elgendi, "On the analysis of fingertip photoplethysmogram signals", Current cardiology reviews, vol. 8(1), pp. 14-25, 2012.

[4] M. Steffen, A. Aleksandrowicz; S. Leonhardt, "Mobile Noncontact Monitoring of Heart and Lung Activity," IEEE Trans. Biomedical Circuits and Systems, vol.1(4), pp.250-257, 2007.

[5] E. Jonathan, M. Leahy, "Investigating a smartphone imaging unit for photoplethysmography," Clinical Physics and Physiological Measurement, vol. 31(11), 2012.

[6] J. B. Bolkhovsky, C. G. Scully, K. H. Chon, "Statistical Analysis of Heart Rate and Heart Rate Variability Monitoring Through the Use of Smart Phone Cameras", IEEE Conf. Engineering in Medicine and Biology Society, pp. 1610-1613, 2012.

[7] C. G. Scully, J. Lee, J. Myer, A. M. GOrbach, Y. Mendelson, K. H. Chon, "Physiological Parameter Monitoring from Optical Recording With a Mobile Phone", IEEE. Trans. on Biomedical Engineering, vol. 59, 2012.

[8] D. Grimaldi, Y. Kurylyak, F. Lamonaca, A. Nastro, "Photoplethysmography Detection by Smartphone's Videocamera", IEEE. Conf. on Intelligent Data Acquisition and Advanced Computing System: Technology and Application, pp. 488-491, 2011.

[9] F. Lamonaca, D. Grimaldi, V. Spagnuolo, "Reliable Pulse Rate Evaluation by Smartphone", IEEE Symposium Medical Measurements and Applications Proceedings, pp.1(4), pp. 18-19, 2012.

[10] I. Omer, M. Werman, "Color Lines: Image Specific Color Representation," IEEE Conf. Computer Vision and Pattern Recognition, vol.2, pp.II946-953, 2004.

[11] J. Allen and A. Murray, "Variability of photoplethysmography peripheral pulse measurements at the ears, thumbs, and toes," IEEE. Proc. Science Measurement and Technology, pp.147 403–7, 2000.

[12] EDAN M3, http://www.edan.com.cn/html/EN/products/patientmonitoring/VitalSigns/201203/20380.html

[13] D.G. Altmam, J.M. Bland, "Measurement in medicine: the analysis of method comparison studies," Statistician, vol.32, pp307-317, 1983.