THE INTERNATIONAL JOURNAL OF SCIENCE & TECHNOLEDGE

A Real-Time and Non-Contact Approach to Monitor Heart Rate for Night Care Applications

Ren-You Huang
Ph.D. Candidate, SoCLab, Institute of Electrical Control Engineering,
National Chiao-Tung University, Hsinchu, Taiwan (R.O.C.)
Lan-Rong Dung
Professor, SoCLab, Department of Electrical and Computer Engineering,
National Chiao-Tung University, Hsinchu, Taiwan (R.O.C.)

Abstract:
Approaches to measure heart rate using a regular color camera has been proposed since their non-contact and low-cost properties. However, in some scenarios, e.g. monitoring the heart rate of the sleeping people at night, are not suitable for the methods using color camera since there is usually no light when people are sleeping. This paper proposed a framework to monitor the heart rate of sleeping people with a camera, an infrared (IR) source, and IR cut filter. The IR sources are attached to the camera, and put on the top of the sleeping people. We computed the intensity variance of ROI instead of mean to extract the pulse signal, followed by Discrete Fourier Transform to find the main frequency as the heart rate. We also addressed the motion interference produced by the sleeping people and alleviated it by an efficient way. We further improved the accuracy by rejecting the unreliable measurements. The experiments show that our method achieved 92% accuracy.

Keywords: *night care, heart rate, real-time, non-contact photoplethysmography.*

1. Introduction

Cardiac pulse rate (i.e. heart rate) is an important physiological parameter. A simple way to measure heart rate was described in the 1930s, which is known as photoplethysmography (PPG) (Hertzman, 1937). PPG detects the optical absorption variations of the human skin due to the blood volume variations. Although PPG is a simple mechanism to measure the cardiac pulse, it still needs to contact the skin (e.g., attached to the subjects' fingers), which is not suitable for the cases of extreme sensitivity, e.g. neonates, skin-damaged patients, or when the non-contact property is required (e.g., surveillance, night care). Peng et. al (Peng, Zhou, Lin, & Zhang, 2015) proposed an alternative method for extracting the PPG signal through the smart-phone camera followed by computing the HRV. However, this method still requires the subjects to put their finger on the smartphone camera and keep themselves static, which has similar disadvantages as the traditional PPG device.

Another works has shown that cardiac pulse rate can be measured in a non-contact way, which is also known as remote-PPG (rPPG) (Huelsbusch & Blazek, 2002; Takano & Ohta, 2007; Verkruysse, Svaasand & Nelson, 2008; Poh, McDuff & Picard, 2010; Poh, McDuff & Picard, 2011; Lewandowska, Rumiński & Kocejko, 2011; de Haan & Jeanne, 2013; Wang, Stuijk & de Haan, 2015; Wu, Rubinstein, Shih, Guttag, Durand & Freeman, 2012; Tarassenko, Villarroel, Guazzi, Jorge, Clifton & Pugh, 2014). These works measure the pulse rate with only one camera, which are low-cost, simple, and effective. The main idea of rPPG is that the blood volume variations can be captured from skin color changes during video recording. The earlier works (Huelsbusch & Blazek, 2002; Takano & Ohta, 2007) first obtain the mean intensity of skin region and then perform frequency analysis, using Fourier or wavelet transforms, to estimate the pulse rate. Recent works (Verkruysse, Svaasand & Nelson, 2008; Poh, McDuff & Picard, 2010; Poh, McDuff & Picard, 2011; Lewandowska, Rumiński & Kocejko, 2011; de Haan & Jeanne, 2013; Wang, Stuijk & de Haan, 2015; Wu, Rubinstein, Shih, Guttag, Durand & Freeman, 2012; Tarassenko, Villarroel, Guazzi, Jorge, Clifton & Pugh, 2014) estimate the pulse rate using a regular color video camera. The first step of these methods is to locate the region of interest by manual selection or automatic face detection, followed by the analysis of skin color changes to extract the pulse signals. The analysis can be done by directly calculating the RGB signals (Verkruysse, Svaasand & Nelson, 2008), source separation (Poh, McDuff & Picard, 2010; Poh, McDuff & Picard, 2011; Lewandowska, Rumiński & Kocejko, 2011), chrominance (de Haan & Jeanne, 2013; Wang, Stuijk & de Haan, 2015), motion magnification (Wu, Rubinstein, Shih, Guttag, Durand & Freeman, 2012), or autoregressive model (Tarassenko, Villarroel, Guazzi, Jorge, Clifton & Pugh, 2014).

Poh et al. (Poh, McDuff & Picard, 2010) proposed an algorithm for heart rate (HR) measurement. They first detected the face every frame and extracted the mean RGB color values to form a three-dimensional time series. Then they applied independent component analysis (ICA) (Comon, 1994; Cardoso, 1999) to separate the independent sources from these RGB signals which may contain the pulse signal, followed by FFT and select the frequency with maximum amplitude in the spectral of the component which has highest peak as the heart rate. Later on, the authors proposed a similar method in (Poh, McDuff & Picard, 2011) to extract the R-R intervals by finding the peaks of the pulse signal. The peaks of pulse signal are treated as the R wave of ECG signal, and the peak intervals are treated as R-R intervals. Alternatively, one may apply PCA, as shown in (Lewandowska, Rumiński & Kocejko, 2011), instead of ICA to separate the pulse signal from RGB time series. Wu et al. (Wu, Rubinstein, Shih, Guttag, Durand & Freeman, 2012) proposed a Eulerian-based motion magnification to magnify the subtle motions or color changes in temporal domain using Laplacian pyramid. This method is able to obtain a clean pulse signal if the subjects are almost static. Tarassenko et. al (Tarassenko, Villarroel, Guazzi, Jorge, Clifton & Pugh, 2014) proposed algorithms using autoregressive (AR) model to extract the heart rate and breath rate for the patients during dialysis. Haan and Jeanne (de Haan & Jeanne, 2013) proposed a chrominance-based remote photoplethysmography (we denote "C-rPPG" in the rest of this paper) which takes different factors into account to form the color model captured by camera. Given the pulsatility as a function of wavelength exhibits a strong peak in green and the dips in red (Crowe & Damianou, 1992; Martinez, Paez & Strojnik, 2011). To exploit this fact and to reduce the specular reflection problem mentioned in (Tominaga, 1994), they proposed a model using difference of weighted color channels to obtain chrominance signals. This method is robust to different skintone and adaptive to non-white illumination. Moreover, the authors showed the impressive results of HR estimation for the scenario with the subjects exercising on stationary bike. Wang et al. (Wang, Stuijk & de Haan, 2015) proposed an algorithm exploiting the spatial redundancy of image sensor and the idea of chrominance to improve the robustness to motions.
The previous works have addressed different problems, e.g., motion artifacts (de Haan & Jeanne, 2013; Wang, Stuijk & de Haan, 2015) and proposed corresponding solutions, which can obtain impressive results. However, most of the works are suitable for the day-time scenario or the environments with visible lighting source since those methods basically extract the pulse signal from the captured RGB color signals. Moreover, some of the works need a lot of computations to obtain accurate heart rate, which are not suitable for real-time applications. In this paper, we proposed a framework to monitor the heart rate of sleeping people, which has following features: (a) using infrared source such that the sleeping people will not be bothered by the visible light at night; (b) tolerant to the slight motions made by sleeping people; (c) very low computational complexity which is suitable for low level computer or embedded system.

2. The Proposed Method

The proposed framework is shown in Fig. 1. The camera is put on the top of the subject's face, with an IR cut filter attached to the lens, and IR sources surrounding with the camera. One may find some off-the-shelf IR cameras on the Internet and setup the same system as the proposed one. The overall processing flow is shown in Fig. 2, while the details of the processes will be described in the following sub-sections.

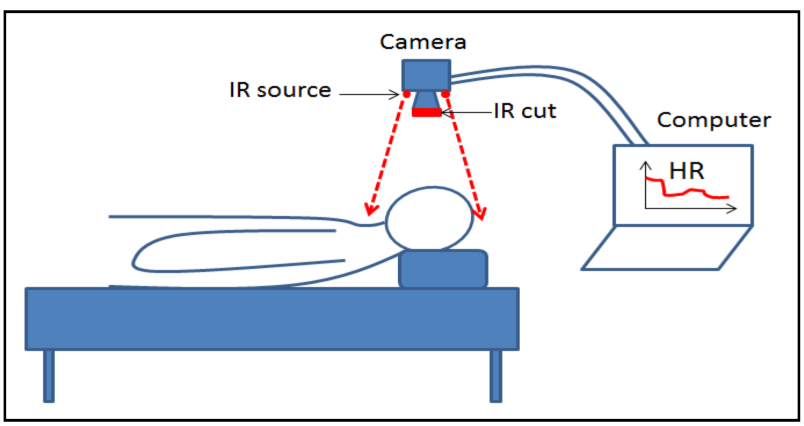

Figure 1: The proposed framework.

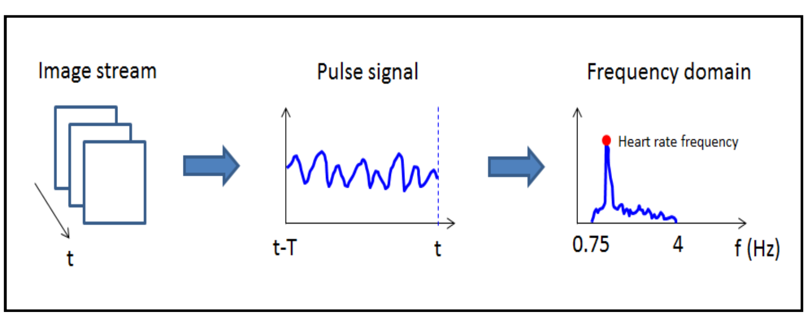

Figure 2: The processing flow of our approach.

2.1. ROI Selection
The first step of all the related works is to determine a region of interest (ROI) for the further processing. The ROI can be determined by manual selection or some automatic algorithm (e.g., face detection). Since the sleeping people are not under controlled (i.e., they might move out of the fixed ROI), the automatic algorithms such as face detection seem to be better choice for a moving subjects. However, the face detection algorithms are usually too complex to be used in real-time application. Moreover, the out-of-plan rotation will greatly degrade the performance of face detection. Considering the above issues, and assuming that the camera can cover the face of the sleeping people within the pillow, we take whole frame as the ROI in our approach.

2.2. Extract the Pulse Signal
For each frame, one may just average the intensity values of all the pixels $I_t(x)$ of the ROI at t-th frame to form the pulse signal s(t) as follows:

$$s(t) = \frac{1}{n}\sum_x I_t(x) \tag{1}$$

followed by some signal analysis to determine the heart rate frequency. However, we found that the mean values might have some small variations produced by the lighting source or the auto-exposure function of camera. Such variations will degrade the accuracy of heart rate estimation. To alleviate this problem, we compute the intensity variance of each frame rather than the mean value, as follows:

$$s(t) = \frac{1}{n}\sum_x \left(I_t(x) - \mu_t\right)^2 \tag{2}$$

where µt is the mean value of the ROI at t-th frame. It is easy to explain the idea by using the histogram of image. Fig. 3 shows an example to explain our idea with assumption that both the background and the subject are static, where g is the intensity value of image and h(g) is the histogram value of g-th bin. If the illumination of lighting source is not constant, it might shift the original distribution of a frame, as shown in Fig. 3(a). The shift changes the mean value of an image, and might become the main frequency of the temporal signal, which results in wrong estimation of heart rate. On the contrary, the variance is more robust to the lighting problem. It is clear that the mean value can be changed by both the light variations and pulsation of blood volume (as shown in Fig. 3(b)), while the variance can be changed by only the pulsation of blood volume.

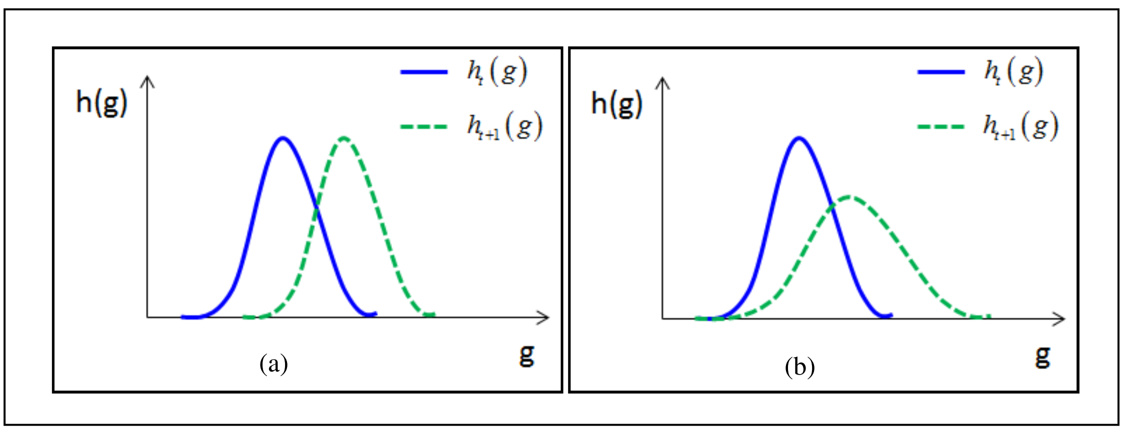

Figure 3: Example to explain the reason that using variance is better than using mean. (a) The histogram of (t + 1)-th frame which is changed by the lighting fluctuation. (b) The histogram of (t + 1)-th frame which is changed by pulsation of blood volume.

Another factor to change the histogram of a frame is the changes of the scene, e.g., the non-rigid deformation of the human face, incoming background object, and the movements of sleeping people. These changes will influence the values of both mean and variance. To alleviate this problem, we compute the absolute frame difference between current frame and previous one,

$$D_f(x) = \left|I_t(x) - I_{t+1}(x)\right| \tag{3}$$

and ignore the pixels which have the difference larger than a pre-defined threshold. To select the threshold, we have considered that the pulsation of blood volume is usually small, therefore we ignore the pixels with frame differences larger than 10 in the computation of variance.

2.3. Determine the Heart Rate
The heart rate is estimated by performing Fast Fourier Transform (FFT) to the pulse signal. We select the frequency with maximum amplitude in frequency domain and multiply the frequency by 60 to obtain the heart rate (bpm), i.e.,

$$H'(t) = \left(\arg\max_{\omega\in\Omega} \left|S(\omega)\right|\right)\cdot 60 \tag{4}$$

where Ω = [0.75, 4] Hz is the observed frequency band, $S(\omega)$ is the Fourier transformed version of the pulse signal s(t) obtained by (2). However, there may be spectral leakage since the non-integer periods of signals sampled in the limited time interval (rectangular window). The spectral leakage may lead to false peak in frequency domain, which decrease the quality of signal analysis. In the theory of digital signal processing (Oppenheim, Schafer & Buck, 1989), multiplying the observed data in the time interval by some window functions (e.g. Gaussian window, BlackmanHarris window) can greatly reduce the spectral leakage and benefit finding the true peak frequency. For simplicity, we selected the Blackman-Harris window in our approach, which is defined as follows:

$$w_B(n) = \sum_{i=0}^{3} (-1)^i a_i \cos\left(\frac{2\pi ni}{N-1}\right) \quad (5)$$

where a_0 = 0.35875, a_1 = 0.48829, a_2 = 0.14128, and a_3 = 0.01168 in (Harris, 1978).

2.4. Unreliable Heart Rate Rejection

Although we have alleviated some interferences mentioned above, there might be still unreliable estimations due to other unknown interference. To reject the unreliable pulse rate, we can simply use the idea of incremental learning as follows:

$$H(t) = \begin{cases} H'(t) & \text{,if } |\Delta H| \le \eta \\ (1-\beta)\mu_H(t) + \beta H'(t) & \text{,otherwise.} \end{cases} \quad (6)$$

where η is a predefined threshold, $\Delta H = |H'(t) - H'(t-1)|$, and $\beta = \exp(-|\Delta H - \eta|)$ is the weighting of estimated pulse rate. If the difference between current pulse rate and previous one is below the threshold η, then we accept the estimated one; otherwise, we reject the value by an exponential factor. By applying this "soft" rejection strategy, the algorithm is more stable and non-sensitive to the value of threshold η. Our system will update the heart rate every one second, thus we can simply assume the variation of heart rate is under 10 bpm, i.e., we set η = 10 in this work. The mean of heart rate μ_H (t) is updated by:

$$\mu_H(t) = (1-\alpha)\mu_H(t-1) + \alpha\mu_H(t) \quad (7)$$

where α is learning rate. To make μ_H (t) capable to adapt the heart rate variation, we suggest α should be in the range of 0.05 to 0.2. Note that we update μ_H (t) only once a second, which is the same as the heart rate estimation. For simplicity, the first few measurements are directly accepted without the rejection strategy and simply averaged to obtain μ_H (t). In other word, we only apply the rejection strategy after the first few measurements. For convenience, we denote the unreliable pulse rate rejection as "UPRR" in the rest of this work.

2.5. Details for Real-Time Implementation

In order to implement the algorithm for the real-time system, we have to clarify some details. We set $N = T \cdot F_s$ where T = 30 is the length of temporal window, and F_s = 30 is the frame rate of the camera. In other word, the temporal window we used to analysis the heart rate is of size N = 900. Our system will start to report the heart rate after t = 30 (seconds).

3. Experiments

We had taped one video clip with about 5 minutes long which has a subject lied on the bed. The subject was asked to wear a chest band which is able to report the heart rate immediately to the mobile or tablet applications. The heart rate measured by the chest band is considered to be the ground truth and will be used to verify the accuracy of our approach.

3.1. Quantitative Indexes

For objective comparisons, we use two quantitative indexes in our experiments, which are mean of absolute error (MAE), and success rate (SR), respectively. The MAE represents the overall average of difference between estimated heart rate and the ground truth, which is in unit of beat-per minute (bpm). The SR represents the accuracy of the overall estimation compared with the ground truth. We consider an estimation is one success if the absolute error is under an acceptable tolerance. These quantitative indexes are defined as follows:

$$MAE = \frac{1}{N_s} \sum_{t \ge 30s} |H(t) - G(t)| \quad (8)$$

$$SR = \left(\frac{1}{N_s} \sum_{t \ge 30s} \delta_\theta(|H(t) - G(t)|)\right) \cdot 100\% \quad (9)$$

where H(t) is the heart rate estimated by different methods, G(t) is the ground truth measured by chest band, N_s is the number of samples from t = 30 to the end of video, $\delta_\theta(\cdot)$ is equal to 1 if the value is smaller than the predefined threshold θ otherwise 0. Here we θ = 5 in (9), which means the errors under 5 bpm are acceptable.

3.2. Results and Discussions

Fig. 4 shows the heart rate of the subject in the video estimated by different approaches. The heart rate estimated by the pulse signal using Eq. (1) is shown in Fig. 4(a). There are three obvious drops of heart rate which are wrong estimations. By replacing the mean value with variance in Eq. (2), there is only one drop of heart rate left, as shown in Fig. 4(b). This error is produced by the large motion of subject. By using the frame difference and ignoring the pixels with large motions, we can alleviate this problem and improve the accuracy slightly, as shown in Fig. 4(c). The drop of heart rate becomes "narrow", which means the accuracy is improved.

However, the interference produced by large motions are still difficult to perfectly be alleviated. Finally, Fig. 4(d) shows the heart rate estimated by the complete proposed method. It is clear that the proposed UPRR is able to reject the unreliable heart rate and significantly improve the accuracy. The quantitative evaluations are listed in the Table 1. It is clear that the complete proposed method is able to achieve 92.30% accuracy with the averaged estimated error 1.80 bpm. The main factor of the error is the "time delay" effect produced by the temporal we used. We can observe that there is about 15 seconds delay from our estimation to the ground truth. This is because of the fact that the temporal is centered at t-N/2, where N = 30 (seconds) in this paper.

	Mean	Var	Var+FD	Var+FD+UPRR
MAE (bpm)	4.17	2.92	2.62	1.80
SR (%)	83.76	88.04	89.20	92.30

Table 1: The quantitative evaluations of different approaches

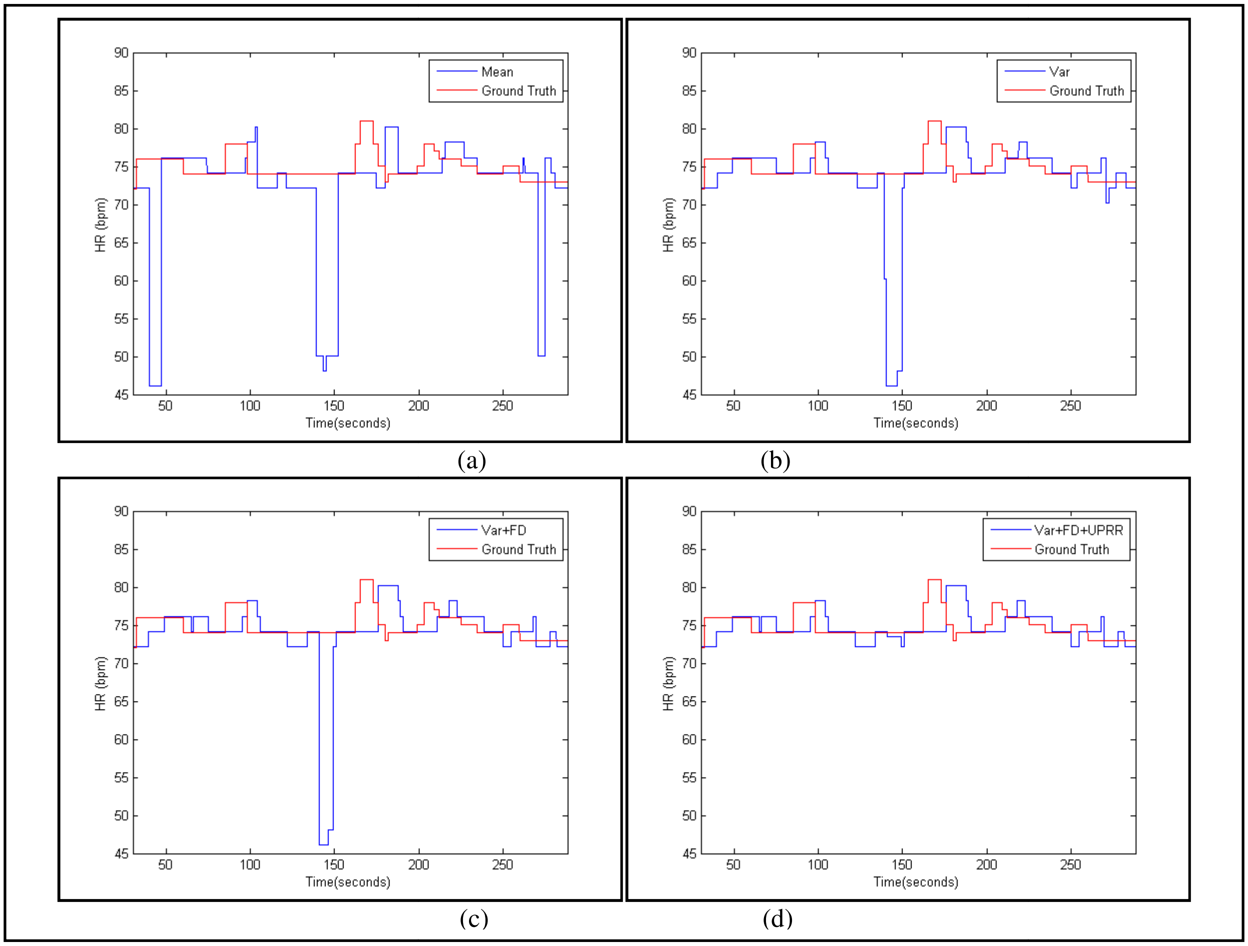

Figure 4: The results of different methods compared with the ground truth. The heart rate estimated by (a) using mean of ROI (b) using variance of ROI (c) using variance and frame difference (d) the complete proposed algorithm.

4. Conclusions

This paper proposed a real-time and non-contact approach to monitor the heart rate of sleeping people with a camera, an IR source and an IR cut filter. The proposed approach has the following features: (a) using infrared source such that the sleeping people will not be bothered by the visible light at night; (b) tolerant to the slight motions made by sleeping people; (c) very low computational complexity which is suitable for low level computer or embedded system. We computed the intensity variance of ROI instead of mean to extract the pulse signal, followed by Fourier Transform with a Blackman-Harris window to find the main frequency as the heart rate. To alleviate the motion interferences, we ignoring the pixels with large frame differences computed by two successive frames. To improve the accuracy, we proposed an UPRR method to reject the unreliable heart rate. The experiments have shown that our method has 92.30% accuracy and the mean of absolute error is only 1.80 (bpm). The remaining problem is that the "time delay" problem produced by the temporal we used, which would be one of our future works.

5. References

i. Hertzman, A. B. (1937). Photoelectric plethysmography of the fingers and toes in man. Experimental Biology and Medicine, 37(3), 529-534.
ii. Peng, R. C., Zhou, X. L., Lin, W. H., & Zhang, Y. T. (2015). Extraction of Heart Rate Variability from Smartphone Photoplethysmograms. Computational and mathematical methods in medicine, 2015.
iii. Huelsbusch, M., & Blazek, V. (2002, April). Contactless mapping of rhythmical phenomena in tissue perfusion using PPGI. In Medical Imaging 2002 (pp. 110-117). International Society for Optics and Photonics.
iv. Takano, C., & Ohta, Y. (2007). Heart rate measurement based on a time-lapse image. Medical engineering & physics, 29(8), 853-857.
v. Verkruysse, W., Svaasand, L. O., & Nelson, J. S. (2008). Remote plethysmographic imaging using ambient light. Optics express, 16(26), 21434-21445.
vi. Poh, M. Z., McDuff, D. J., & Picard, R. W. (2010). Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Optics express, 18(10), 10762-10774.
vii. Poh, M. Z., McDuff, D. J., & Picard, R. W. (2011). Advancements in noncontact, multiparameter physiological measurements using a webcam. Biomedical Engineering, IEEE Transactions on, 58(1), 7-11.
viii. Lewandowska, M., Rumiński, J., & Kocejko, T. (2011, September). Measuring pulse rate with a webcam—a non-contact method for evaluating cardiac activity. In Computer Science and Information Systems (FedCSIS), 2011 Federated Conference on (pp. 405-410). IEEE.
ix. de Haan, G., & Jeanne, V. (2013). Robust pulse rate from chrominance-based rPPG. Biomedical Engineering, IEEE Transactions on, 60(10), 2878-2886.
x. Wang, W., Stuijk, S., & de Haan, G. (2015). Exploiting Spatial Redundancy of Image Sensor for Motion Robust rPPG. Biomedical Engineering, IEEE Transactions on, 62(2), 415-425.
xi. Wu, H. Y., Rubinstein, M., Shih, E., Guttag, J. V., Durand, F., & Freeman, W. T. (2012). Eulerian video magnification for revealing subtle changes in the world. ACM Trans. Graph., 31(4), 65.
xii. Tarassenko, L., Villarroel, M., Guazzi, A., Jorge, J., Clifton, D. A., & Pugh, C. (2014). Non-contact video-based vital sign monitoring using ambient light and auto-regressive models. Physiological measurement, 35(5), 807.
xiii. Comon, P. (1994). Independent component analysis, a new concept?. Signal processing, 36(3), 287-314.
xiv. Cardoso, J. F. (1999). High-order contrasts for independent component analysis. Neural computation, 11(1), 157-192.
xv. Crowe, J. A., & Damianou, D. (1992, October). The Wavelength Dependence of the Photoplethysmogram and its implication to Pulse Oximetry. In Engineering in Medicine and Biology Society, 1992 14th Annual International Conference of the IEEE (Vol. 6, pp. 2423-2424). IEEE.
xvi. Martinez, L. F. C., Paez, G., & Strojnik, M. (2011, August). Optimal wavelength selection for noncontact reflection photoplethysmography. In International Commission for Optics (ICO 22) (pp. 801191-801191). International Society for Optics and Photonics.
xvii. Tominaga, S. (1994). Dichromatic reflection models for a variety of materials. Color Research & Application, 19(4), 277-285.
xviii. Oppenheim, A. V., Schafer, R. W., & Buck, J. R. (1989). Discrete-time signal processing (Vol. 2). Englewood Cliffs: Prentice-hall.
xix. Harris, F. J. (1978). On the use of windows for harmonic analysis with the discrete Fourier transform. Proceedings of the IEEE, 66(1), 51-83.