

Obtaining physiological parameters of biological objects from video stream methods analysis

Maxim S. Beloyvanov*, and Issam Samara

Don State Technical University, 346500 Shakhty, Russia

Abstract. The paper presents the analysis of existing methods for the non-contact measurement of the pulse of the person from the video stream received from the RGB camera. The proposed methods improve the accuracy of their work and decrease the negative impact of human movement on measurement results.

1 Introduction

Currently there are a great interest in studies about measuring biological parameters, for example, breathing rate, blood pressure, micro-expression of the face, micro-motion of the body, the heart beat rate. from the above parameters of particular interest is the measurement of pulse, since it is an important life indicator, reflecting the physical and psychological state of the person, and one of the decisive factors in the diagnosis of heart disease.

Currently, there is a huge amount of work devoted to methods of non-contact determination of human pulse, each of which has its advantages and disadvantages. The carried out analysis of the given works allows to understand in detail the negative sides of each of the methods, and to suggest ways of solving or eliminating the noted flaws.

2 Review

The existing methods of non-contact pulse detection can be divided into two fundamentally different approaches to its measurement by video stream.

The first is based on photoplethysmography — a method of recording blood flow using an infrared or light source and a photoresistor, phototransistor or RGB camera. For the first time such an approach using an RGB camera was used by Costa et al [1].

The second approach is based on the strengthening the differences between the frames which are invisible to the human eye, usually the head movements are strengthened [2].

The analysis showed that most of the methods based on the first approach, since the second method is subject to a large influence of the interfering component of the oscillations of the head in the frame. Thereby, we will consider methods based on the first method.

* Corresponding author: wintermakc@gmail.com

The detection of the pulse on the video stream is usually made over a certain regions of interest (ROI) of the frame, contains pixels that refer to the person's face. The way ROI is chosen differs in different works, as in the work of Muhammad Waqar et al [3]. In the work [4] a very reliable approach is used, which is effective for minor movements of the person and ensures high accuracy of the selection of the area. Using the resulting ROI, cheeks are allocated, because it is least likely that they will be hidden by the hair, beard or glasses.

In the work of H. Rahman et al [5], the Viola—Jones method is used for face detection [6, 7], the rectangular region obtained as a result of using this method is used for further analysis.

The Discriminative Response Map Fitting (DRMF) [8] is used to detect 66 coordinates of facial reference points in the work of Xiaobai Li et al [9]. This method is a discriminative method based on regression for structuring constrained local models (CLM) which can find the exact reference points of the face. ROI determination for these points is made taking into account the following features: it is necessary to exclude the eye area, since blinking can interfere with the evaluation of heart rate, and exclude areas at the borders of the face, which may adversely affect the evaluation of heart rate, so ROI is selected with a small paddings. Figure 1 shows the variants of ROI allocation.

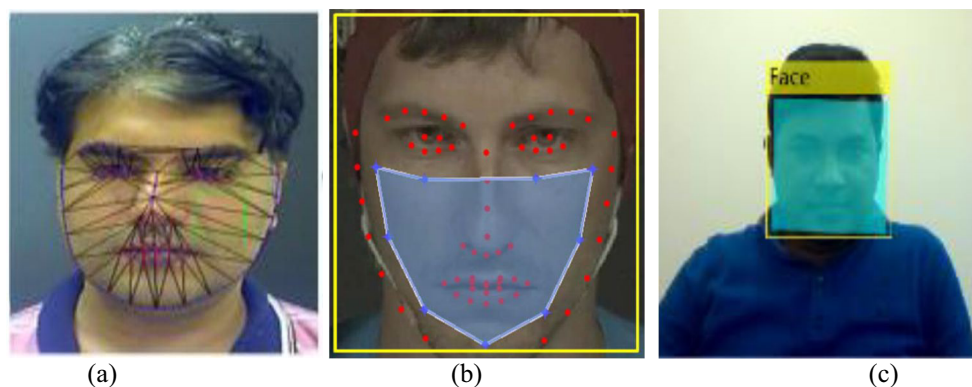


Fig. 1. Ways of detecting the face in the image (a) Face Tracking Saragih [3] (b) 66 points DRMF [9] (c) Viola-Jones method [5].

After determining the areas of interest for processing, in the first group of methods, the image is preprocessed, since the raw signal taken from the ROI is affected by lighting and head movements.

In all studies, in order to obtain the signal, an image is decomposed into three color components R, G and B.

Thus, Muhammad Waqar et. al. [3] receive three raw signals $x_1(t)$, $x_2(t)$ and $x_3(t)$, which are normalized as follows:

$$x'i(t) = x_i(t) - \mu_i \sigma_i \quad (1)$$

for each $i = 1, 2, 3$, where μ_i и σ_i - the mean and standard deviation of each color component (x_i) respectively. The raw signals x_i is converted to a normalized raw signal $x'i$, which have zero mean and deviation equals one. In their study, they used the joint approximate diagonalization of eigenmatrices (JADE) and ICA (Independent Component Analysis) [10], which is able to remove the noise generated by the movement of participants. The independent components are returned randomly, and the component with the highest peak in its power spectrum is selected as the desired pulse signal. A fast Fourier transform is applied to the selected signal to obtain the power spectrum. The researchers selected the heart rate as the frequency with the highest peak in the signal power spectrum

within the working frequency band. The operating range is between 0.7-4 Hz, which corresponds to 40-240 beats per minute and covers almost all permissible heart rate measurements.

During the experiments, the researchers found that the method of detecting a peak value inside a fixed window is not reliable so, during the optimization of the moss filter [11] they developed a selection function that determines the peak of the heart rate and measures the heart rate.

In the work of H. Rahman et al [5], isolating the signal was done using ICA and Principal component analysis (PCA). Before applying PCA, ICA and fast Fourier transform (FFT), red, green and blue signals generated from the color values of each frame which are filtered with a Hamming window (128 points, 0.6-2 Hz), which corresponds to the heart rate normal range 36-120 beats per minute.

It follows from [12] that information on the photoplethysmographic signal is contained in the red, green, and blue channels. However, green is the most informative, this is due to the fact that hemoglobin absorbs the green part of the spectrum stronger than the red and blue [13].

On this basis, in Xiaobai Li et al [9], the signal for analysis is calculated as the average value of the green color within the ROI of each frame.

The value of the signal is affected by two factors: the change in blood volume in the vessels and the change in lighting during video analysis, while the illumination brings noise into the signal being processed. Assuming also that the signal changes caused by these two factors are adaptive, background regions can be used to reduce noise in the face area. Figure 2 shows the result of this approach.

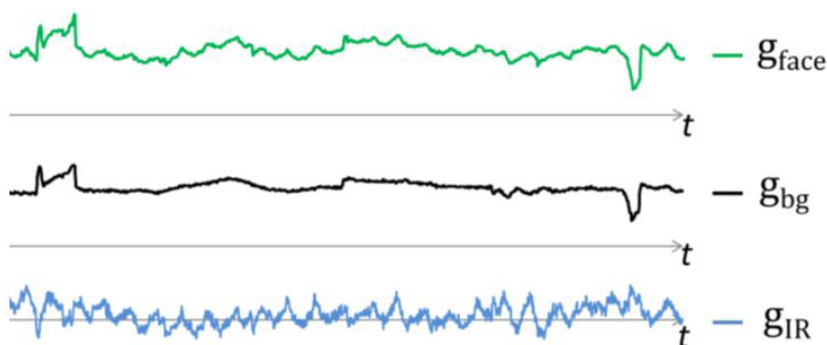


Fig. 2. Filtering noise caused by a light source [9] (g_{face} – the raw signal ROI, g_{bg} – average value of the green channel on the background, g_{IR} – filtered signal).

The next step in the work of Xiaobai Li et al [9] is a temporal filtering of the signal. They apply several time filters to exclude frequencies that are not within the range of normal heartbeat values. The frequency range is set within the range 0.7-4 Hz, which corresponds to a normal heart rate of 42-240 beats per minute.

In [14], several time filters are also used to measure the heart rate: the first one is a detrending filter based on a smoothness priors approach [15]. The second is a moving average filter that reduces random noise using the average value of adjacent frames. The third is a filter based on the Hamming window with a cutoff frequency of 0.7-4 Hz.

After filtering, the signal is converted from the time domain to the frequency domain by the Welch method [16]. The frequency with maximum power is used as the heart rate.

As a result of the above researches, the authors of all works came to the conclusion that such methods can achieve high accuracy of measurements, but the quality of measurements is greatly influenced by several factors:

- the variability of lighting;
- the noise caused by the drift of the coordinates of facial landmarks;
- the large amplitude of displacement of the ROI region under investigation;
- the quality of used filters.

3 Conclusion

Measurement of heart rate using the method of photoplethysmography has some disadvantages. The results of the studies have shown that it is possible to improve the accuracy of measurements by improving the quality of used filters, this can be done using the Filtering method proposed in [17]. To reduce the negative impact of the drift of the coordinates of the facial landmarks can be due to adaptive tracking of several areas on the facial landmarks and switching between them.

References

1. G. Da Costa, Optical remote sensing of heartbeats, *Opt. Comm.*, **117**(5-6), 395 (1995)
2. H.Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, W. Freeman, *ACM Transactions on Graphics* 31(4), 1 (2012) <http://dx.doi.org/10.1145/2185520.2185561>
3. M. Waqar, B. Tiddeman, *Proc. 7th UK Br. Mach. Vis. Work.*, 7, 1-9 (2015)
4. S. Jason FaceTracker Available: <https://github.com/kylemcdonald/FaceTracker> (2015)
5. H. Rahman, M.U. Ahmed, S. Begum, P. Funk, In *The 29th Annual Workshop of the Swedish Artificial Intelligence Society (SAIS)*, Malmö, Sweden, 129, 15 (2016)
6. P. Viola, M. Jones, *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2001) <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.10.6807>
7. P. Viola, M. Jones. In *Workshop on statistical and computational theories of vision*, 266, 56 (2001) <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.110.4868>
8. A. Asthana, S. Zafeiriou, S. Cheng, and M. Pantic. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3444 (2013)
9. X. Li, J. Chen, G. Zhao, M. Pietikainen, In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4264 (2014)
10. J. Cardoso, *Neural Comput.* 11(1), 157 (1999)
11. D.S. Bolme, J.R. Beveridge, B.A. Draper, Y.M. Lui, In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 2544 (2010)
12. W. Verkruyse, L. O. Svaasand, J. S. Nelson, *Optics express*, 16, 21434 (2008)
13. S. Prahl. Optical absorption of hemoglobin <http://omlc.ogi.edu/spectra/hemoglobin/> (1999)
14. M.-Z. Poh, D. J. McDuff, R. W. Picard, *IEEE Trans. on Biomedical Engineering*, 58(1), 7-11 (2011)
15. M.P. Tarvainen, P.O. Ranta-Aho, P.A. Karjalainen, *IEEE Transactions on Biomedical Engineering*, 49(2), 172 (2002)
16. P. Welch, *IEEE Trans. on Audio and Electroacoustics*, 15(2), 70 (1967)
17. V.I. Marchuk, I.S. Shraifel, *Methods for isolating the useful component with a priori uncertainty and a limited amount of measurement results*. Monograph (Publishing House of the SPSU, Shakhty, 2008)