Research article
DOI: https://doi.org/10.18721/JCSTCS.17308
UDC 004.89

NON-INVASIVE HEART RATE MEASUREMENT SYSTEM BASED ON VIDEO STREAM ANALYSIS

M.S. Antonenko[1] ✉ , *D.O. Budanov*[1], *A.Yu. Zaitceva*[2]

[1] Peter the Great St. Petersburg Polytechnic University,
St. Petersburg, Russian Federation;

[2] Institute for Analytical Instrumentation of the Russian Academy of Sciences,
St. Petersburg, Russian Federation

✉ antonenko.ms@edu.spbstu.ru

Abstract. The paper is devoted to the development and testing of a remote biomonitoring system based on the phenomenon of plethysmography. This phenomenon allows not only to measure a person's pulse rate non-invasively, but also to assess physiological state of the person. At the first stage of the system operation, it is necessary to detect regions of interest. This operation can be effectively implemented using neural networks. The task of face recognition was performed by the YOLOv7-tiny architecture, due to its speed and the ability to run on embedded systems. For the detected face, a rectangle was created, whose coordinates indicated the boundaries of the face. Next, the average brightness of the selected areas is calculated and stored in the dataset. By performing fast Fourier transform (FFT) for a given set, it is possible to obtain a signal spectrum. Using methods of digital signal processing, it is possible to filter the signal and select the part of the spectrum of interest in the region of 0.7–3 Hz. The maximum amplitude of the harmonic will correspond to the current pulse.

Keywords: remote biomonitoring, telehealth, heart rate, photoplethysmography, Fourier transform, neural network

Citation: Antonenko M.S., Budanov D.O., Zaitceva A.Yu. Non-invasive heart rate measurement system based on video stream analysis. Computing, Telecommunications and Control, 2024, Vol. 17, No. 3, Pp. 84–92. DOI: 10.18721/JCSTCS.17308

Научная статья
DOI: https://doi.org/10.18721/JCSTCS.17308
УДК 004.89

СИСТЕМА НЕИНВАЗИВНОГО ИЗМЕРЕНИЯ ЧАСТОТЫ СЕРДЕЧНЫХ СОКРАЩЕНИЙ НА ОСНОВЕ АНАЛИЗА ВИДЕОПОТОКА

М.С. Антоненко[1] ✉ , *Д.О. Буданов*[1], *А.Ю. Зайцева*[2]

[1] Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Российская Федерация;

[2] Институт аналитического приборостроения РАН, Санкт-Петербург, Российская Федерация

✉ antonenko.ms@edu.spbstu.ru

Аннотация. Статья посвящена разработке и тестированию системы дистанционного биомониторинга на основе явления плетизмографии. Данное явление позволяет не только неинвазивно измерить пульс человека, но и оценить его физиологическое состояние. На первом этапе работы системы необходимо детектировать области интереса. Данная операция может быть эффективно реализована с использованием нейронных сетей. Задачу распознавания лица выполняла архитектура YOLOv7-tiny, за счет быстродействия и возможности запуска на встраиваемых системах. Для детектированного лица создавался прямоугольник, координаты которого обозначали границы лица. Далее осуществляется вычисление средней яркости выбранных областей и сохранение в наборе данных. Выполняя Быстрое преобразование Фурье для заданного набора, можно получить спектр сигнала. Используя методы цифровой обработки сигнала можно отфильтровать сигнал и выделить интересующий нас участок спектра в районе 0.7-3 Гц. Максимальная амплитуда гармоники и будет соответствовать текущему пульсу.

Ключевые слова: дистанционный биомониторинг, телемедицина, частота сердечных сокращений, фотоплетизмография, преобразование Фурье, нейронная сеть

Для цитирования: Antonenko M.S., Budanov D.O., Zaitceva A.Yu. Non-invasive heart rate measurement system based on video stream analysis // Computing, Telecommunications and Control. 2024. Т. 17, № 3. С. 84–92. DOI: 10.18721/JCSTCS.17308

Introduction

Nowadays, telehealth systems capable for monitoring human physiological parameters are a promising area of research. As a rule, these systems are used for non-invasive express diagnosis of pathological conditions. Non-invasive method for measuring heart rate (HR) is based on the analysis of photoplethysmogram that is the result of recording changes that occur when small vessels are filled with blood, depending on the phase of the cardiac cycle [1]. This dependence is periodic and indicates the current pulse of a person. Moreover, it can be obtained using a conventional camera. By performing frame-by-frame image processing, it is possible to evaluate changes in skin tone and estimate current HR of a person [3]. In this case, of particular interest in the measurement are forehead and under-eye areas, where change in skin tone is most noticeable. Then, using methods of digital signal processing and computer vision, it is possible to obtain information about the current HR from the brightness curve [4].

In this paper, the algorithm for determining regions of interest (ROIs) on a person's face in a video stream is proposed. To solve this task, it is necessary to detect a face in the frame and then select the boundaries of the ROIs and outline them with geometric primitives (rectangles). At this stage, high-performance neural network models should be used. Next, the HR is calculated using classical signal analysis methods. Based on the results of comparison of existing models for face detection, a neural network

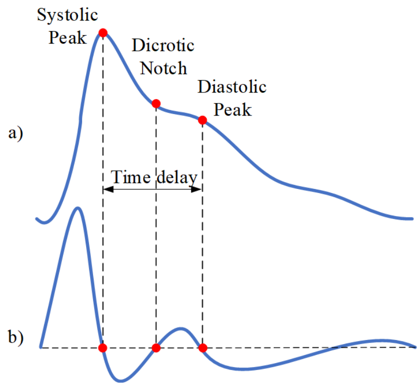

Fig. 1. Signal measurements: (a) original fingertip photoplethysmogram;
(b) first derivative wave of photoplethysmogram

with the YOLOv7-tiny architecture was chosen. The algorithm for finding ROIs on a person's face is based on coordinate regression, which is used to find the coordinates of a fixed number of points. To carry out coordinate regression it is possible to use a convolutional neural network based on the MobileNetV2 architecture [5].

Photoplethysmography

Photoplethysmography is the method that evaluates changes of blood volume in blood vessels each time a heart beats [1] (Fig. 1).

The arterial pulse waveform can be separated into three distinct components:

- The systolic phase, characterized by a rapid increase in pressure to a peak, followed by a rapid decline. This phase begins with the opening of the aortic valve and corresponds to the left ventricular ejection.
- The dicrotic notch, which is widely believed to represent the closure of the aortic valve.
- The diastolic phase, which represents the run-off of blood into the peripheral circulation [2].

The principle of photoplethysmography is based on determining the optical density of tissue. The ROI is illuminated from one side, after which the scattered light reflected and transmitted through the tissue area is received at the photodetector. The magnitude of its intensity is proportional to the change in blood supply to the tissue during contraction and relaxation of the heart muscle. The more blood in the vessel lead to an increase red blood cells that scatter light, the more light is reflected from them [3].

Optical methods of microcirculation analysis of biological tissues are based on the total spectral optical parameters of the medium (reflectance, scattering, absorption). For different biological environments, it depends on the functional, physiological and pathophysiological state of tissues, on their anatomy as well as on the percentage concentration of certain endogenous tissue chromophores in them, different forms of hemoglobin, connective tissue collagen, fat, water, melanin, etc. Each tissue chromophore has its own and specific spectral characteristic, which makes it possible to identify these molecular compounds by optical methods and distinguish them from other chromophores contained in the biological tissue. On this basis, methods have been developed that are now actively and quite successfully used in clinical practice. It can be single out the well-known and widespread method of pulse oximetry, which is accurate enough in assessing microcirculation and oxygen saturation. In addition, this method is fast and accessible, but has some significant drawbacks and limitations in interpreting the result [4].

The disadvantages of such methods and systems are as follows:

- The patient's movements can greatly affect the measurement result.
- Poor tissue perfusion distorts the measurement result and, as a consequence, this method depends on the pulse component.

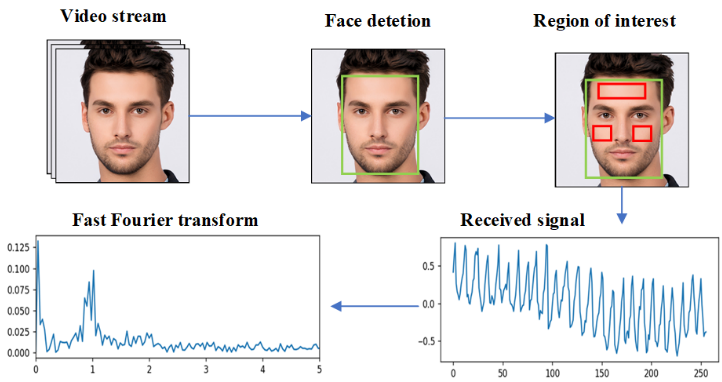

Fig. 2. Steps taken to extract the heart rate from facial video

- If there is abnormal hemoglobins in the blood, such as methemoglobin, the result may be unreliable.
- Microcirculation can be assessed only on certain skin areas: finger/nose/earlobe.
- The technique does not show the status of tissue microcirculation in the periphery, does not analyze the integral result and requires subjective interpretation of the result by medical personnel.

Changes in the parameters of backscatter radiation, in conjunction with a certain sequence of changes in the color of tissues on the face, can be recorded by wearable sensors and video camera, processed by machine learning methods to determine the presence and degree of microcirculatory pathology [5].

Heart rate determination algorithm

This section describes the basic steps taken to obtain the HR based on changes of skin tone. The entire algorithm is presented in Fig. 2.

At the first stage it is necessary to obtain an image from the camera, then detect a face and select ROIs (in this case, forehead and under-eye areas). This step is performed using the OpenCV library.

At the next stage, the average brightness of each selected area is calculated. By determining this value frame by frame for each moment of time, it is possible to obtain the dependence of changes of skin tone over time (photoplethysmogram). To obtain the numerical value of the HR, it is necessary to process the received data [7]. This can be achieved by the use of Fourier transform to convert a function from the time domain into the frequency domain. Then, the frequency with the largest amplitude in the range of 0.75–3 Hz is selected from the obtained spectrum. This value corresponds to the numerical value of the HR.

The raw heartbeat signal contains other extraneous high and low frequency components due to ambient color and motion noise induced from the data capturing environment. Therefore, to increase accuracy approximation algorithms, filters, and signal division into modes are applied [7, 8].

Using methods of NumPy statistical data processing libraries and Matplotlib data visualization libraries, the following results were obtained (Fig. 3, 4).

To process the data, first, the average value of the entire set is subtracted to eliminate the constant component of the trend. Next, appropriate filters are applied to eliminate high and low frequency components. After the Fourier transform and obtaining the spectrum, the maximum is checked against the average value to detect a static image (Fig. 5).

Results of heart rate calculation

The next step is to obtain, detect and outline a face with a primitive rectangle (bounding box). To solve this problem, a neural network for real-time object detection YOLOv7 was used. YOLOv7-tiny is the most compact and fastest model, suitable for use on devices with limited resources. Memory consumption and a recognition time were taken into account when choosing the architecture of the neural

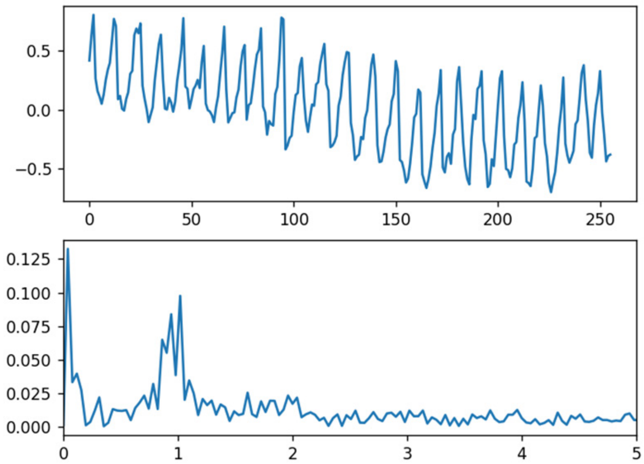

Fig. 3. Heartbeat signal in the state of calm (above) and its spectrum (below)

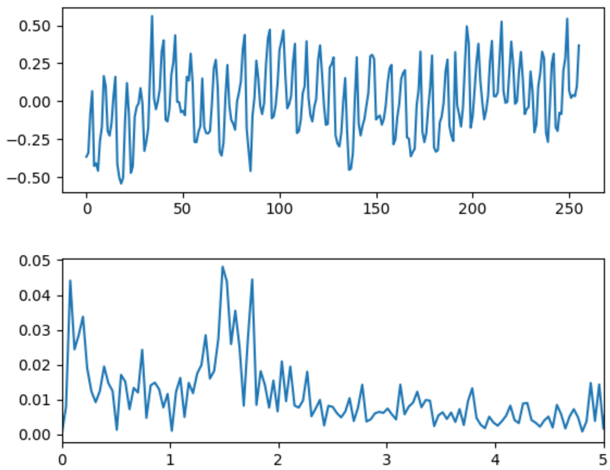

Fig. 4. Heartbeat signal (above) and its spectrum (below) under a physical activity

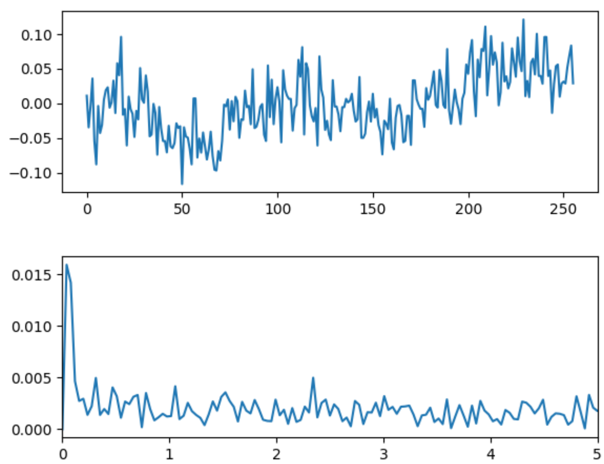

Fig. 5. Heartbeat signal of static video (above) and its spectrum (below)

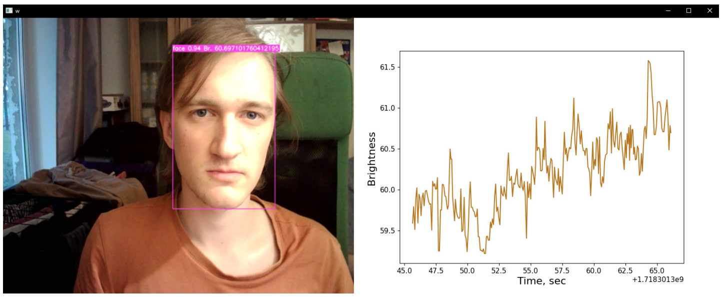

Fig. 6. Interface of an application for heart rate detection

network used in this paper. YOLOv5 and YOLOv6 offer different model sizes as well, however, on average they are more resource intensive. Developers provide a comparison of different versions of the YOLO model in [9]. In [5], for this neural network, recall, accuracy, and average accuracy were used as tracked metrics, which reached 68.7%, 62.2%, and 86.2%, respectively. All operations are performed during the single iteration unlike other detection algorithms that need to repeat the same image processing many times, such as algorithms based on region proposal networks [9]. To train the model the Wider Face dataset[1] was used, which contains 32203 images with 393703 labeled faces. Each image has the size of 640x640 pixels. During model training, the computational TPU cores were used in Google Colab with the following parameters: number of epochs – 17, batch size – 16, workers – 4.

The software part of the proposed system, containing a trained model with a HR detection algorithm and a user interface, was executed on two hardware platforms: CPU-based and GPU-based. The first platform consisted of AMD Ryzen 7 3700U CPU and 6 GB of RAM. The second platform consisted of GeForce RTX 3070 GPU with 32 GB of RAM. The GPU contained 5888 CUDA cores and 8 GB VRAM, which makes the performance of the platform several orders of magnitude better. Then a video stream from a web camera was sent to the input of the trained model and a face was detected. The resulting coordinates of bounding box were used in the HR calculation algorithm, for which the average brightness was calculated in the whole detected area. The image processing window is shown in Fig. 6. As can be seen from the left side of the figure, the camera detects a face and outlines it. The probability of detection and average brightness are shown on the top of the bounding box. The right side of the figure contains a graph of brightness (absolute value) over time (sec). This curve is polyharmonic due to the constant change in frame size and parasitic facial movements.

According to the obtained results, the influence of the camera matrix resolution prevails over the influence of the illumination level. Therefore, the influence of the illumination level will not be considered further. The frame rate of the video stream must be at least 30 fps (supported by any modern web-camera). The minimum image resolution is 640x640, since a dataset with images with this resolution was used to train the model.

Fig. 7 shows the signal spectrum before and after filtering with a high-pass filter with a cutoff frequency of 0.7 Hz. The pulse in the state of calm is 58–65 beats per minute, which corresponds to a peak around 1 Hz.

To identify the dependence of measurement accuracy from the performance of the prototype of the proposed system, measurements were carried out on two platforms: GPU (CUDA) and CPU. The results of executing the proposed application when varying the spectrum counting points are presented in Table.

Table 1 shows that calculations on the CPU have unacceptable accuracy and require more time to measure, due to the fact that video processing on the GPU occurs in real time (30 fps), while on the CPU it is 2–3 fps. As a consequence, peaks are often missed and a lot of useful information is lost. The number

[1] WIDER FACE: A Face Detection Benchmark, Available: http://shuoyang1213.me/WIDERFACE/ (Accessed 13.09.2024)

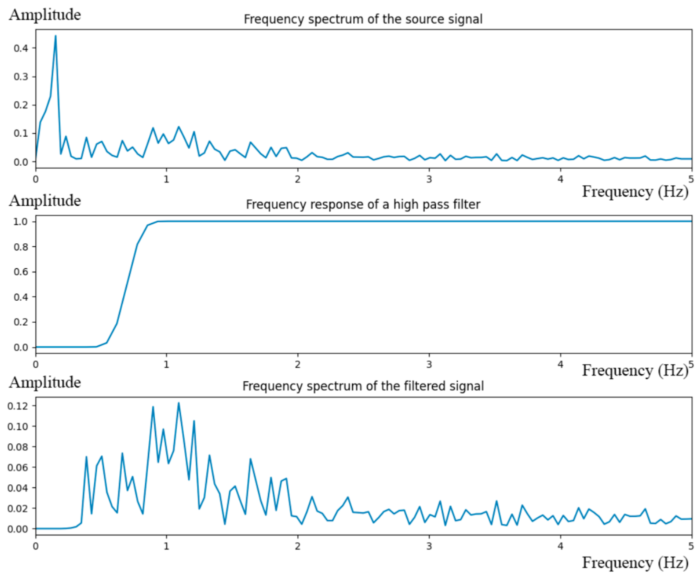

Fig. 7. Signal spectrum

of points for fast Fourier transform does not greatly affect the accuracy of measurements. However, the more points used, the more reliable results are, as there is less influence of unnecessary accidental facial movements. At the same time, the speed and complexity of calculations increase proportionally. From this point of view, it is optimal to use 128 points. For embedded systems, it is possible to use the NVIDIA Jetson platform. This single board computer has a small size and low power consumption. In addition, it contains GPU with high-performance CUDA cores suitable for machine learning tasks.

Table

Comparison of actual and received values from the platform

№	Obtained	Real	Number of points	Computing platform	Error
1	60.7	69	256	GPU	8.4
2	62	74	256	GPU	12
3	59.4	69	256	GPU	9.4
4	54.3	70	128	GPU	15.7
5	56	66	128	GPU	10
6	66	69	128	GPU	3
7	39	65	256	CPU	26
8	51	64	256	CPU	13
9	33	68	128	CPU	35
10	42	65	128	CPU	23

Comparison with other implementations

The article [7] uses a similar method for measuring HR based on changes in skin tone, as well as micromovements of the face. To remove/reduce the extraneous frequency components and trends from the signal it was decomposed using Hodrick-Prescott filter. The accuracy of the proposed method was compared with state of the art color and the motion-based methods of [11–12]. The overall error rates are less than 10% for HR estimation by counting the number of peaks for both motion and color signals.

In the description of how the application “webcam-pulse-detector” works[2], it says, that data was collected by “measuring average optical intensity in the forehead location, in the subimage's green channel alone (a better color mixing ratio may exist, but the blue channel tends to be very noisy)”. Measurement accuracy is not given, however, with a good lighting and minimal noise due to motion, a stable heartbeat should be isolated in about 15 sec. In this paper, it takes 16 seconds to measure heartbeat rate using 128 points.

Several methods are considered in [10] to achieve an error of 5 beats per minute.

An illumination rectification method by using two points on the skin of the face in the same frame to extract the green spectrum of each point was proposed in [15]. After that, independent component analysis was applied to extract the photoplethysmography signal from the two green spectrum signals. Therefore, treating the effect of illumination variance as a blind source separation problem. ROI selections criteria using facial landmarks fitting was proposed by the authors of this article as well.

Conclusion

Remote biomonitoring plays an important role in modern telehealth systems. It allows to measure non-invasively the basic parameters of a person’s physiological state without taking samples in a few seconds using only a camera and specialized software. In this work, a system for measuring HR based on analysis of video stream was proposed. The results of the experiment showed that for the correct and affordable work of the system it is necessary to use a GPU in a hardware part of the system. Therefore, currently proposed system has low accuracy since the entire area of the rectangle that contains face is taken into account, however this expands the measurement capabilities. In the future work, to increase accuracy, a neural network model to detect ROIs on the human face will be developed and applied.

REFERENCES

1. **Lazim M.R.M.L.M., Aminuddin A., Chellappan K., Ugusman A., Hamid A.A., Ahmad W.A.N.W., Mohamad M.S.F.** Is heart rate a confounding factor for photoplethysmography markers? A systematic review. International Journal of Environmental Research and Public Health, 2020, vol. 17, no. 7, article no. 2591. DOI: 10.3390/ijerph17072591

2. **Yartsev A.** Normal arterial line waveforms. 2015. Available: https://derangedphysiology.com/main/cicm-primary-exam/required-reading/cardiovascular-system/Chapter%20760/normal-arterial-line-waveforms (Accessed 10.05.2024)

3. **Guryleva A., Fomina D., Volkov M., Danilycheva I., Serdoteckova S.** Digital image processing for the diagnosis of the cold inducible urticaria by photoplethysmographybased methods. 2023 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), 2023, pp. 234–245. DOI: 10.1109/WECONF57201.2023.10147961

4. **Danishevskiy N.S., Budanov D.O., Zaytseva A.Y.** Facial Landmark Detection Approaches for Remote Medical Applications. IEEE International Conference on Control in Technical Systems, 2023, pp. 177–180. DOI: 10.1109/CTS59431.2023.10288892

5. **Danishevskiy N.S., Budanov D.O., Zaytseva A.Yu.** Development of an intelligent system for remote biomonitoring of heart rate. Technical Physics Letters, 2023, vol. 49, no. 12, pp. 23–25. DOI: 10.61011/TPL.2023.12.57575.204A

6. **Huang S.-H.** Accurate pulse transit time acquisition method for multi-point photoplethysmography. International Conference on Consumer Electronics – Taiwan (ICCE–Taiwan), 2023, pp. 551–552. DOI: 10.1109/ICCE-Taiwan58799.2023.10226655

7. **Haque M.A., Nasrollahi K., Moeslund T.B.** Estimation of heartbeat peak locations and heartbeat rate from facial video. Image Analysis. SCIA 2017. Lecture Notes in Computer Science, 2017, vol. 10270, pp. 269–281. DOI: 10.1007/978-3-319-59129-2_23

[2] Webcam-pulse-detector, Available: https://github.com/thearn/webcam-pulse-detector (Accessed 13.09.2024)

8. **Suriani N.S., Jumain N.A., Ali A.A., Mohd N.H.** Facial video based heart rate estimation for physical exercise. 2021 IEEE Symposium on Industrial Electronics & Applications (ISIEA). 2021, pp. 1–5. DOI: 10.1109/ISIEA51897.2021.9509986

9. **Wang C.-Y., Bochkovskiy A., Liao H.-Y.M.** YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023, pp. 7464–7475. DOI: 10.48550/arXiv.2207.02696

10. **Hassan M.A., Malik A.S., Fofi D., Saad N., Karasfi B., Ali Y.S., Meriaudeau F.** Heart rate estimation using facial video: A review. Biomedical Signal Processing and Control, 2017, vol. 38, pp. 346–360. DOI: 10.1016/j.bspc.2017.07.004

11. **Poh M.-Z., McDuff D.J., Picard R.W.** Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Optics Express, 2010, vol. 18, no. 10, pp. 10762–10774. DOI: 10.1364/OE.18.010762

12. **Kwon S., Kim H., Park K.S.** Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone. 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2012, pp. 2174–2177. DOI: 10.1109/EMBC.2012.6346392

13. **Balakrishnan G., Durand F., Guttag J.** Detecting Pulse from Head Motions in Video. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013, pp. 3430–3437. DOI: 10.1109/CVPR.2013.440

14. **Poh M.-Z., McDuff D.J., Picard R.W.** Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam. IEEE Transactions on Biomedical Engineering, 2011, vol. 58, no. 1, pp. 7–11. DOI: 10.1109/TBME.2010.2086456

15. **Lam A., Kuno Y.** Robust heart rate measurement from video using select random patches. Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 3640–3648.

INFORMATION ABOUT AUTHORS / СВЕДЕНИЯ ОБ АВТОРАХ

Antonenko Maksim S.
Антоненко Максим Сергеевич
E-mail: price.lt@yandex.ru

Budanov Dmitry O.
Буданов Дмитрий Олегович
E-mail: budanov_do@spbstu.ru

Zaitceva Anna Yu.
Зайцева Анна Юрьевна
E-mail: anna@da-24.ru
ORCID: https://orcid.org/0000-0003-0334-2770

Submitted: 16.06.2024; Approved: 29.08.2024; Accepted: 05.09.2024.
Поступила: 16.06.2024; Одобрена: 29.08.2024; Принята: 05.09.2024.