Comparative Analysis of Face ROI Outline Algorithms for Contactless Heart Rate Measurement Under Different Registration Conditions

Ivan Semchuk, Konstantin Zlobin, Natalia Muravskaya, Andrey Samorodov

Bauman Moscow State Technical University
Moscow, Russian Federation
semvan@bmstu.ru

Abstract—**Contactless heart rate measurement techniques can be applied in medical and biometrical tasks such as vital signs measurement and vitality detection. Incorrect measurement result can cause serious consequences. The paper presents the results of experimental studies aimed at comparing algorithms for face images processing for contactless heart rate measurement using video analysis. Algorithms based on color segmentation and facial landmarks are performed. A method for evaluating the effectiveness of algorithms and data acquisition process are described. The least heart rate measurement error is shown by algorithm based on a grid with the nodes at facial landmarks formation and independent signal calculation for each grid sector.**

I. Introduction

Videoplethysmography (VPG) is a contactless pulse curve recording technology based on processing the video stream of a face image followed by time-frequency analysis of the signal in order to calculate physiological parameters [1], [2].

Various modifications of methods for video images processing and VPG signals are presented in literature aimed at increasing heart rate (HR) measurements stability while reducing the influence of image registration parameters [3] and taking into account the physiological features of facial blood flow [4], [5].

VPG signal acquisition model can be described using the following equation [5]:

$$VPG(t) = \iint_{x,y} A(x,y) \times P(t, \omega, \varphi(x,y)) \, dxdy \quad (3)$$

Where A(x, y) is amplitude of reflected light, P – pulse wave function, t – time, ω – heart rate frequency, $\varphi(x, y)$ – pulse wave at (x, y) point.

Expression (1) allows to select several research areas at once:

- Evaluating the effectiveness of different approaches to processing facial images under different registration conditions
- Splitting the integration area into smaller areas with independent analysis of the received signals

will allow you to ignore areas of the image with a low signal-to-noise ratio (SNR)

The purpose of this work is a comparative analysis of various algorithms for ROI outline in a face image under different signal registration conditions.

II. Description of the algorithms under study

Before calculating the VPG signal it is necessary to outline image sections containing useful information about the blood flow dynamics [6].

To detect the patient's face widely used computer vision methods such as Haar cascades or LBP are applied [7]. However, the result of detection is coordinates of the bounding box which fits the desired object best while the solution to the segmentation problem is a set of pixels belonging to the desired object. In VPG task segmentation is used to detect pixels that belong to skin areas of the face surface. The segmentation methods studied in this paper are described below.

A. Color-based method

The area of color space which contains the distribution of human skin colors is quite limited. Thus segmentation of the required face area can be reduced to sifting pixels belonging to a cluster defined by a parallelepiped with borders (0,133,77), (255,173,127) in the YCrCb color space [8]. The value of the VPG amplitude corresponding to the i-th frame of the video image of the face is calculated using the formula:

$$VPG_i = \frac{\sum_{j,k} G_{j,k}}{\sum_{j,k} R_{j,k} + \sum_{j,k} B_{j,k}} \quad (2)$$

Where i is the frame number of the video image; j, k are the coordinates of pixels belonging to the selected area; R, G, B are the color coordinates of the corresponding pixels. An example of the result of an algorithm based on the color segmentation method is shown in Fig. 1 (b) (the original image is shown in Fig. 1 (a)).

B. Facial landmarks-based method

Since the selected area of skin in the image has a characteristic geometry it makes sense to rely not only on its shape in the color space but also on the standard metric representation. The dLib open library provides methods for segmenting 68 points of face contours using a histogram of directed gradients. By cropping the image along the contours one can get an image with pixels containing information only about the human skin. The contour that bounded the desired area was based on landmarks that described the eyebrows and the lower part of the face. The eyes and lips that fall inside this contour were blackened before calculating the VPG signal using expression (2). Fig. 1 (c) shows an example of how this algorithm performance.

C. Combined method

Combining operations in both metric and color spaces allows to use areas of the skin that do not fall inside the contour drawn facial landmarks and at the same time has the property of adaptability to the lighting parameters of the biological object [9]. Segmentation of the forehead and exclusion of hair are possible if the distribution of pixels sifted in the color space is correctly analyzed during the operation of the algorithm based on facial landmarks detection. Calculations have shown that the RGB coordinates of points belonging to the areas selected by this method are approximated by a linear function (with an average value of $R^2=0.97$, calculated from 5000 random frames from the dataset used). Thus the combined algorithm consists of the following steps:

1) Facial landmarks detection

2) Selecting a contour based on detected points

3) Calculation of coefficients of a linear function f(R, G, B) that approximates a point cloud in RGB space based on pixels inside the selected area

4) Search for pixels outside the selected contour where the following expression is valid:

$$r(\mathrm{p},\mathrm{f}) < \mu_r \qquad (3)$$

where r(p, f) is the Euclidean distance from point p to the approximating function f and μ_r is the average distance between the color coordinates of pixels belonging to the selected contour and the approximating function.

An example of the algorithm result is shown in Fig. 1 (d).

D. Landmarks-net-based method

Expressions (1) and (2) imply that each frame of the video corresponds to a single count of the VPG signal which takes into account all pixels belonging to the selected area. It can potentially worsen the signal-to-noise ratio if there are uninformative areas in the selected area of the image. Building the grid with the nodes at facial landmarks allows to calculate VPG for each sector of the grid independently. This permits to

exclude noisy skin regions and take into account phase distribution of the pulse wave along face surface. In this work 68 facial landmarks were detected and 56 points were used for creating the grid. The grid was made up of rows and columns with lines parallel to image boarders. An example of an image divided into sectors is shown in the Fig. 3. For each sector the value of the VPG was calculated using the expression (2).

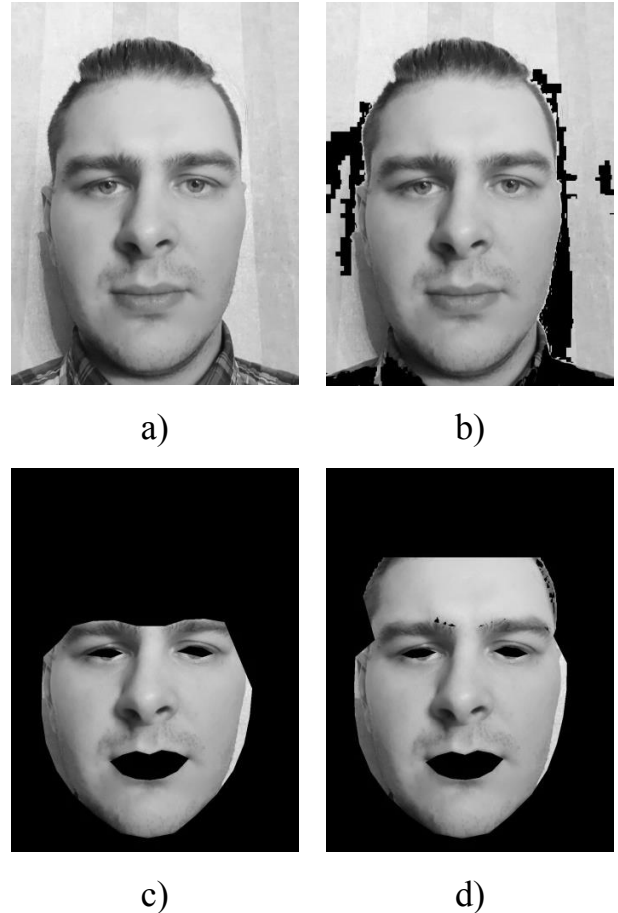a)   b)

c)   d)

Fig. 1. Example of the results of algorithms ROI outline performance: a) the original image, b) color-based method, c) facial landmarks-based method, d) combined method
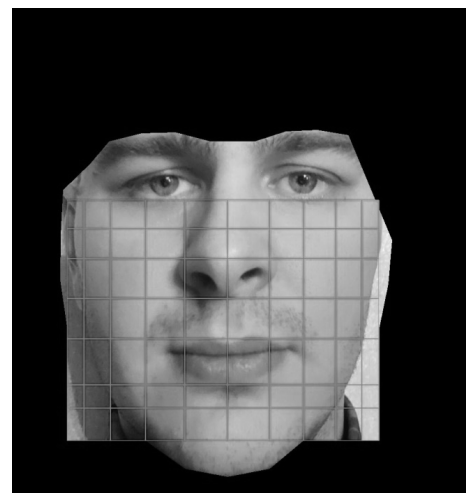

Fig. 2. An example of landmarks-net-based method performance

IV. ALGORITHMS EFFECTIVENESS EVALUATION METHOD

According to the purpose of the technology under study it is obvious that the main metric of the effectiveness of the image processing algorithms should be based on the error of heart rate measurement using the contactless method. Let HR^{PPG} be the heart rate value calculated using contact photoplethysmography (PPG) signal and HR^{VPG} be the heart rate value calculated using VPG signal. Then the result of comparing the two signals data:

$$q = \begin{cases} 1, \left| HR^{PPG} - HR^{VPG} \right| \leq 1 \\ 0, otherwise \end{cases} \tag{3}$$

Acceptable error rate (±1 bpm) is chosen according to the pulse oximeter test recommendations [10].If the signals have a sufficiently long duration (N points), it is advisable to divide them into smaller sections with the size of k points. Then:

$$q_i = \begin{cases} 1, \left| HR_i^{PPG} - HR_i^{VPG} \right| \leq 1 \\ 0, otherwise \end{cases} \tag{4}$$

Where i takes the values [k/2; N-k / 2] and describes the position of the center of the window of size k, within which the sections of contact and non-contact photoplethysmograms are compared.

The final metric for the quality of a contactless photoplethysmogram signal is the result of the expression:

$$Q = \frac{\sum_{\frac{k}{2}}^{N-\frac{k}{2}} q_i}{N - \frac{k}{2}} \tag{5}$$

When calculating this metric for the landmarks-net-based algorithm expression (5) is converted as:

$$q_i = \begin{cases} 1, \exists j, k \left| HR_i^{PPG} - HR_{i,j,k}^{VPG} \right| \leq 1 \\ 0, otherwise \end{cases} \tag{6}$$

Where j,k are the numbers of the row and column of the grid where the sector used for calculating the videoplethysmogram is located.

The heart rate calculation is based on the localization of systolic peaks in the signal with a peak detector using wavelet transform followed by the calculation of the average time interval between peaks.

V. MATERIALS AND METHODS

The illumination of a biological object and the distance between the registration tool and the biological object were used as parameters that simulate video recording conditions. The illumination was changed by means of dimmed lamps and controlled by a luxmeter. The distance between the camera and the face of the subject was measured using a laser rangefinder.

Two series of measurements were performed in each of which one of the parameters was changed and the second was maintained at a constant value. Thus when changing the distance between the camera and the subject's face illumination was maintained at the level of 1000 lx. When the light intensity was changed the distance between the bioobject and the registration tool was equal to 60 cm.

The measurements involved 4 subjects each of whom underwent parallel registration of the video of the face and contact PPG three times at a given value of the controlled environmental parameter. The duration of each measurement was 30 seconds. Thus for each value of the controlled parameter, the signals were registered 12 times.

For each of the VPG signal the value of the signal quality metric was calculated using the formulas (2, 3, 4). To check whether Q-vectors are well-modeled by normal distribution Shapiro - Wilk test was performed. Checking the statistical significance of the difference between Q-vectors for different algorithms with the same parameter values was performed using the Student's t-test.

VI. RESULTS

Fig. 3 and 4 show the dependence of the average value of the metric calculated using the formulas (2, 3, 4) on the illumination of the subject's face and the distance between the face and the registration tool respectively. Analysis of the measurement results allows us to conclude that the algorithm based on an independent analysis of individual sections of the face image showed the best result in terms of the value of the efficiency metric. There was no statistically significant difference between the values of the performance metric for the other three algorithms.
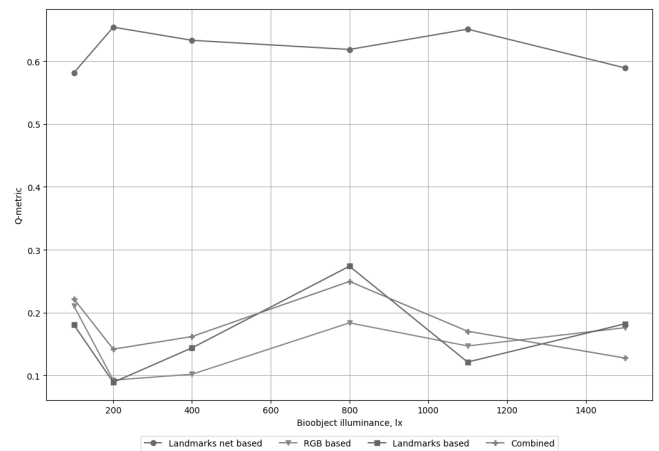

Fig. 3. The dependence of the efficiency metric of ROI outline algorithms on the illumination of a biological object
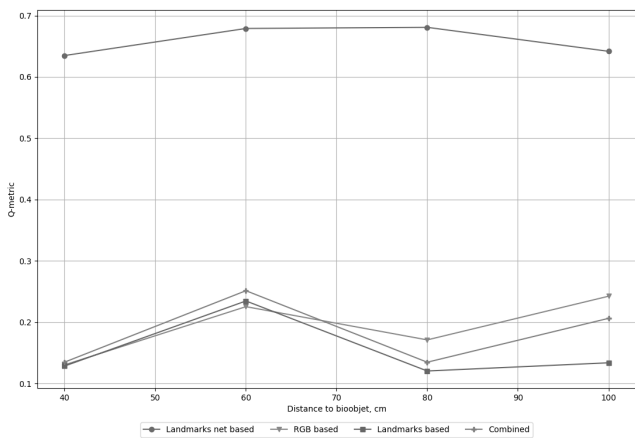
Fig. 4. Dependence of the efficiency metric of ROI outline algorithms on the distance between the camera and the bioobject

VII. CONCLUSION

The paper is devoted to the comparative analysis of algorithms for ROI outline on face images in the task of contactless heart rate measurement. The operation of four algorithms based on color segmentation and the use of facial landmarks was studied under different illumination conditions of the subject's face and also at different distances between the camera and the subject's face. The results suggest that the algorithm based on the grid with the nodes at facial landmarks with independent videoplethysmogram calculation for each sector shows best measurement accuracy and stability.

REFERENCES

[1] Cennini G., Arguel J., Akşit K., van Leest A., "Heart rate monitoring via remote photoplethysmography with motion artifacts reduction". *Opt Express.*, Vol. 18, 2010, № 5, pp.4867-4875.

[2] Semchuk, I.P., Zmievskoy, G.N., Muravskaya, N.P., Volkov A.K., Murashko M.A., Samorodov A.V., "An Experimental Study of Contactless Photoplethysmography Techniques", *Biomed Eng.*, № 53, 2019, pp.1-5.

[3] Przybylo J, Kantoch E, Jablonski M, Augustyniak P., "Distant measurement of plethysmographic signal in various lighting conditions using configurable frame-rate camera", *Metrology and Measurement Systems*, Vol. 23, 2016, № 4, pp 579–92.

[4] Yoshizawa M., Sugita N., Tanaka A., Ichiji K., Homma N., Yambe T., "An Optimization Technique to Extract Video Pulse Wave for Non-Contact Remote Monitoring of Autonomic Nervous System and Blood Pressure Variability", *IEEE 7th Global Conference on Consumer Electronics(GCCE)*, 2018, pp 425–28.

[5] Semchuk I., Muravskaya N., Samorodov A., "The study of facial ski surface bloodflow dynamics", *AIP Conference Proceedings*, 2019, 2140:1.

[6] S. Thakur, S. Paul, A. Mondal, S. Das, A. Abraham, "Face detection using skin tone segmentation", *2011 World Congress on Information and Communication Technologies, Mumbai*, 2011, pp. 53-60.

[7] A. Adouani, W. M. Ben Henia and Z. Lachiri, "Comparison of Haar-like, HOG and LBP approaches for face detection in video sequences", *2019 16th International Multi-Conference on Systems, Signals & Devices (SSD), Istanbul*, 2019, pp. 266-271.

[8] Kaur A., Kranthi B. V., "Comparison between YCbCr color space and CIELab color space for skin color segmentation" *International Journal of Applied Information Systems,* Vol. 3, 2012, №. 4. pp. 30-33.

[9] J. Fritsch, S. Lang, A. Kleinehagenbrock, G. A. Fink and G. Sagerer, "Improving adaptive skin color segmentation by incorporating results from face detection", *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication, Berlin,* 2002, pp. 337-343.

[10] МИ 3280-2010 Пульсовые оксиметры и пульсоксиметрические каналы медицинских мониторов. Методика поверкиб 2010 – с.12