Non-Contact Video-Based Remote Photoplethysmography for Human Stress Detection

Submitted: 20th October 2018; accepted: 2nd June 2020

Sergii Nikolaiev, Sergii Telenyk, Yury Tymoshenko

DOI: 10.14313/JAMRIS/2-2020/21

Abstract: *This paper presents the experimental results for stress index calculation using developed by the authors information technology for non-contact remote human heart rate variability (HRV) retrieval in various conditions from video stream using common wide spread web cameras with minimal frame resolution of 640x480 pixels at average frame rate of 25 frames per second. The developed system architecture based on remote photoplethysmography (rPPG) technology is overviewed including description of all its main components and processes involved in converting video stream of frames into valuable rPPG signal. Also, algorithm of RR-peaks detection and RR-intervals retrieval is described. It is capable to detect 99.3% of heart contractions from raw rPPG signal. The use-cases of measuring stress index in a wide variety of situations starting with car and tractor drivers at work research and finishing with students passing exams are presented and analyzed in detail. The results of the experiments have shown that the rPPG system is capable of retrieving stress level that is in accordance with the feelings of experiments' participants.*

Keywords: *video processing; web cameras; stress index; remote photoplethysmography; rPPG; heart rate; heart rate variability; Predictive, Preventive, Personalized and Participatory Medicine.*

1. Introduction

At present, the formation of the XXI century medicine requires a new philosophy and the platforms for more effective person's treatments to advance current healthcare systems. There are several needs that the modern and innovative healthcare systems across the planet should respond to and among them are: the rising costs of medical care and the emerging need to reduce such costs; the grand challenges facing the healthcare and biomedical industry in order to utilize many of the novel technologies; the necessity of radical improvement in wellness and disease prevention; the developing shortage of healthcare professionals; the strong desire of the individual persons to participate more in every aspects of their healthcare.

These tasks to be solved need a new paradigm of advanced healthcare in terms of predictive, preventive, personalized, and participatory (P4) medicine [1]. The core elements of that vision are widely accepted now providing physicians and patients with personalized information about everyone's health on different system levels. During the development of P4 medicine, many rapidly developing technologies such as artificial intelligence, telemedicine, smart clothes with wearable sensors, mobile apps and beyond will result in treating the causes rather than the symptoms of disease, more efficient patient management and hence a better quality of life. The healthcare industry is on the cusp of substantial changes in the coming decade as new technologies are being developed. In this paper, the authors follow the paradigm of P4 Medicine, which is a global trend in the 21st century and it involves continuous monitoring of the humans' condition even before any signs of negative changes [2].

The authors have developed robust contactless remote photoplethysmography information technology that uses video stream processing in real time. The developed system is able to obtain biological indicators like HRV through the use of widely distributed web and other video cameras. Mass adoption of such information technology (IT) will allow people to provide an appropriate level of heart health through continuous contactless monitoring of HRV without changing their life styles and could make our medicine preventive.

Considering that the relationship between stress, heart disease and sudden death has been recognized since antiquity the main focus of this paper will be dedicated to stress. The choice of studying stress is dictated by the following facts:

According to statistics, annual costs to employers because of stress related health care issues and missed work are estimated in $300 Billion [3]. Moreover 77 % of respondents of the research marked that they regularly experience physical symptoms caused by stress. And 73 % of people experience psychological symptoms because of stress. 33% feel that they are living with extreme stress. So, detecting and measuring stress may not only help to make lives better, but also to reduce heart and other diseases.

In this paper the results of measuring stress index of humans in different conditions with the help of developed system are presented.

2. Heart Rate Variability

"Heart Rate Variability" (HRV) has become the conventionally accepted term to describe variations of both instantaneous heart rate and the series of times between consequential pairs of heart contractions (so called RR-intervals). The analysis of HRV has been widely used as a non-invasive and reliable tool to evaluate cardiovascular autonomic control in health and disease. To describe oscillation in consecutive cardiac cycles, other terms have also been used in the literature: for e.g. cycle length/ heart period variability or RR interval sequences (tachograms), and they more appropriately emphasize the fact that it is the interval between consecutive beats that is being analyzed rather than the heart rate per se [4]. In this work the term HRV will be used throughout the article.

Usually for the accurate diagnosis of cardiovascular diseases the Holter device is used as the medical standard for heart activity measuring. It requires a patient to visit a doctor, install the device for couple of days, and then follow doctor's examining of obtained electrocardiograms (ECGs) manually. But many heart diseases do not require the entire ECG to be examined, and for diagnostics it is enough to have only beat-to-beat time intervals, so-called RR intervals. The phenomenon to focus on is the oscillation in the intervals between consecutive heart beats as well as the oscillations between consecutive instantaneous heart rates. Patterns in these oscillations contain enough information for unveiling not only heart pathologies but also dysfunctions of the whole organism.

The modern IT development infinitely extends the possibility of tracking various biological signals of a person with further computer processing of digital data. In recent years, there have been various alternatives to the Holter device, namely: personal pulse meters, "smart" clocks, fitness trackers that allow you to record HR, continuous monitoring of the cardiovascular system and reduce the risk of cardiovascular diseases (CVD. Modern markets of mobile soft- and hardware are filled with a different kind of applications for health monitoring and pulsometer-like gadgets that may read, store and process our biological signals.

But the only problem remains that all these approaches are contact and in some types of applications it is impossible to make contact measurement and remote technology is needed to estimate HR and heart rate variability.

3. Contactless Remote Photoplethysmography

In recent years, the possibility to extract the heart rate (HR) with the help of a remote photo detector has been established. The approach is called remote photoplethysmography (rPPG) [5–9]. The new technique offers a heart rate (HR) measurement that does not need to have contact with the studied object, which is a valuable feature for both medical and surveillance purposes [8]. rPPG contactless monitors of human heart activity detect subtle human skin color variations from reflected light observed by the camera which are induced by heart contractions and blood flow [10].

Lately, several new rPPG algorithms have been developed for pulse-signal extraction from the face with RGB-cameras as photo detectors [6, 7]. These include:

(a) Blind Source Separation (e.g., PCA-based [11] and ICA-based [12]), which use different criteria to separate temporal RGB traces into uncorrelated or independent signal sources to retrieve the pulse;
(b) CHROM [13], which linearly combines the chrominance signals by assuming a standardized skin-color to white-balance the camera;
(c) PBV [14], which uses the signature of blood volume changes in different wavelengths to explicitly distinguish the pulse-induced color changes from motion noise in RGB measurements; and
(d) 2SR [15], which measures the temporal rotation between spatial subspaces of skin-pixels for pulse extraction.

The essential difference between these rPPG algorithms is in the way of combining RGB-signals into a pulse-signal. The use of three color channels with multiple wavelengths gives the methods the possibility to be robust to motion of the subject. A better understanding of the core rPPG algorithms could benefit many systems/applications for video health monitoring, such as the monitoring of heart-rate [16–20], respiration [17], SpO2 [21], blood pressure [22], neonates [23–24], and the detection of atrial fibrillation [25] and mental stress [26].

4. Remote PPG System's Architecture Description

The developed by the authors of this work rPPG technology is based on the one-pixel camera mathematical model. This means that we in theory can replace our camera with theoretical camera that has only one single pixel. This pixel measures with infinite precision the amount of lite that falls into photosensitive sensor for all three color-channels and do not generate any noise. Thus, the mathematical model expects noise-free real valued luminance signal in three different spectral channels (red, green and blue). To be able to use such abstraction with real multi-pixel cameras that produce many types of noise and have finite precision of measurement the following modules structure is proposed:

1) Face detection module;
2) Images spatial filtering module;
3) Module for skin tints time series frequencies filtration;
4) Heart beats' time detection module

Video processing begins with sequential analysis of each video frame applying face detector, and spatial filters like: skin-detector to find binary mask of skin pixels on the frames; transformations of color tint signal spaces to compensate energy of skin luminance;

aggregating skin pixel colors to reduce camera's sensor's pixel noise. At this stage the more pixels are aggregated the higher precision is achieved and the lower noise levels are present so the processed signal satisfies criteria of one-pixel camera model. On the next stage temporal filters including frequency finite impulse response pass-band filter with frequencies of heart rate to remove all temporal noises except heart signal. Heart beats' time detection module returns sequence of heart con-traction moments in time that allow to calculate time deltas between each pair of R-peaks resulting in series of RR-intervals.

In the following sections each processing stage will be described in more details.

4.1. Face Detection

The problem of finding faces in the images and video frames can be defined as follows: in a given picture with dimensions K*M pixels, it is necessary to find the coordinates of rectangles that correspond to the bounding boxes of minimal size that fully contain faces of the image.

Different approaches are known that solve this task [27]. As a front face detector, in this work Haar cascade classifier was used because of its high speed and accuracy in real-time video processing applications.

In order to train the classifier, it is necessary to have labeled images dataset, which consists of a set of images with faces (positive samples) and a set of background images (negative samples). The resulting detector was trained on a dataset containing 35,200 frontal faces and 60,000 background images. Each image in the dataset has dimension of 20x20 pixels.

The face detector was trained using the Viola and Jones algorithm (ADAboost). For the Haar cascade quality improvement, bootstrapping procedure has been sequentially applied to modify the training set of negative samples before each training iteration.

The depth of the resulting cascade is 25 stages. The sequence of stages with increasing number of features (only two features are checked at the first stage and 218 features – at the last stage) causes low level of false detections and small average calculation time for the whole cascade. Such cascade structure allows to process 17 frames/second in resolution 640x480 pixels on AMD dual core 1.8GHz laptop and more that 30 frames/second in resolution 320x240 which is enough for our purposes. On i7-7700 HQ CPU processing more than 30 frames per second was achieved on frames with resolution 1280x720 pixels.

The quality of the trained classifier was checked on a test sample of 7000 faces and 3400 background images. The following metrics of the classifier quality were obtained:

tp=6953 – true positives – the number of correctly classified images with faces,

fp=47 – false positives – the number of images with faces classified as background,

tn=3372 – true negatives – the number of correctly classified images with background,

fn=28 – false negatives – the number of images with background classified as faces.

$$precision\frac{tp}{tp+fp} = 0.9932 \tag{1}$$

$$recall = \frac{tp}{tp+fn} = 0.9959 \tag{2}$$

$$accuracy = \frac{tp+tn}{tp+fp+tn+fn} = 0.9927 \tag{3}$$

To speed up the process of finding faces in live stream video and to lower CPU load, the refinement search principle is being used during detection phase. This principle allows reducing the amount of computation when processing the next frames of video stream after frame with already found faces. The essence of the refinement search is that after a successful detection of a face in one frame, the search for this face on the next frame is performed only in the vicinity of the already found bounding box rectangle of this given face and not throughout the whole frame.

The task of finding pulse and separate heartbeats from video stream requires low noise levels from all components of the system. In the described above algorithm detected bounding box coordinates may change on couple of pixels frame to frame and this effect add additional noise on the next levels of signal processing. So, to avoid this noise the algorithm of stabilizing the position and size of the face bounding box on the frame is used. Its main purpose is cuts off accidental changes in the position and size of the found face rectangle.

Thus, the face detection module output consists of the video stream with stabilized human face images.

4.2. Frames Spatial Filtering

In this paper one-pixel camera model is used so the purpose of this module is to create mapping $S(I) \to R$ of each video frame $I \in R^{N \times N \times 3}$ into a one-dimensional real number $s^* \in R$, which encodes face skin tint from the input image. At this stage the analysis takes place within the same frame over spatial coordinates excluding temporal factors. This operation of spatial filtration aggregates pixels colors of skin and thus excludes pixel noise with zero mean and captures mean of skin tint which can be tracked later in temporal context.

In order for the pixels of the background not to add additional noise, the first step of spatial filter module is to apply a skin detector to the input image and obtain the skin mask.

The most skin detectors are based on the comparison of pixel colors with the database of known skin colors, and if the examined color is in the database – then in the resulting mask corresponding pixel is set to 1, and 0 otherwise. Another approach is to have a model (for example based on RBF neural network) that approximates skin colors probability distribution and returns for each given input pixel or region of pixels a probability that the input is part of skin.

Applying threshold for resulting probabilities the skin bitmask is obtained.

Finding a skin binary pixel mask is done with the following algorithm:

- blur input image to reduce pixel noise artifacts;
- convert the color space of the input image into HSV (Hue-Saturation-Value) space;
- select skin mask M of pixels where model probability for each pixel of being skin pixel is higher than 80%;
- apply "dilation" and "erosion" operations to reduce noise between frames in time;
- calculate the weighted average of the pixel intensities $i_{x,y,z}$ under the mask $M:(x,y)\in M$ for each of the channels $z\in\{R,G,B\}$.

Thus, after applying this algorithm, three real values of the average intensities of the skin pixels are obtained for each input image: $I_z = E_{(x,y)\in M} i_{x,y,z}$ one intensity value for each of the color channel $z\in\{R,G,B\}$.

From the mathematical point of view the intensity of one single pixel can be represented as

$$i_{x,y,z} = i_z + s_{x,y,z} + \eta_{x,y,z},$$

where i_z is the average illumination of the observed object from external light source; $\eta_{x,y,z}$ component is the pixels noise induced by light sensing element of the camera; $s_{x,y,z}$ component is the skin tint change we want to measure.

Usually $\eta_{x,y,z} \gg s_{x,y,z}$, but given the fact that the noise has zero mean ($E_{x,y}\eta_{x,y,z} = \eta_z = 0$), the average of a sufficient number of pixels over the mask allows to reduce the noise level $\eta_{x,y,z}$ to values where the amplitude of the useful skin tint signal s_z becomes larger than the amplitude of the pixels noise $s_z \gg \eta_z$. The larger the number of pixels is used for averaging is the better.

Therefore, after applying the transform:

$$E_{(x,y)\in M}\left(i_{x,y,z} - i_z\right) = E_{x,y}s_{x,y,z} + E_{x,y}\eta_{x,y,z} =$$

$$E_{x,y}s_{x,y,z} = s_z$$

the desired signal s_z is obtained. But at this stage, the component i_z remains unknown.

Let's consider the following fact that oxyhemoglobin and hemoglobin absorb and reflect light in different parts of the spectrum differently. So, between channels $z\in\{R,G,B\}$ the values of s_R, s_G, and s_B are non-constant quantities that are nonlinearly dependent on the ratio of levels oxyhemoglobin and hemoglobin in the capillaries of the skin. Thus, the time series $s_R(t)$, $s_G(t)$ and $s_B(t)$, where t is timestamp of received frame from the camera, also show nonlinear dependence.

At the same time, when the illumination of the object changes, components $i_R(t)$, $i_G(t)$ and $i_B(t)$ change proportionally or remain unchanged with constant illumination. By selecting the constants $T_z, z\in\{R,G,B\}$ so that $T_R * i_R(t) + T_G * i_G(t) + T_B * i_B(t) = 0$ the variation of the useful signal can be preserved while the intensity component of the object's illumination in time is removed from the equation.

Introducing the following designation of the vectors $s(t) = [s_R(t), s_G(t), s_B(t)]$ and $T_{RGB} = [T_R, T_G, T_B]$, the final equation becomes:

$$\sum_{z\in\{R,G,B\}} E_{(x,y)\in M} T_z * i_{x,y,z} =$$

$$\sum_{z\in\{R,G,B\}} T_z * E_{x,y}s_{x,y,z} + T_z * i_z + T_z * E_{x,y}\eta_{x,y,z} =$$

$$\sum_{z\in\{R,G,B\}} T_z * E_{x,y}s_{x,y,z} = T_{RGB} * s = s^* \quad (4)$$

where s^* – is the desired signal.

To calculate the T_{RGB} value let's use the fact that the variances of the object illumination level time series and the signal differ by several orders of magnitude $D_t(i_z(t)) \gg D_t(s_z) \gg D_t(\eta_z)$. This allows us to apply the methods of blind source separation, which rotate and distort the space $\{I_z(t), z\in\{R,G,B\}\}$ in the way so that the useful signal s^* becomes one of the components, and everything else is distributed among other components.

The bests results were obtained by applying Independent Component Analysis (ICA) as blind source separation method that changes the joint distribution of the components and separates the desired signal s^*. The following coefficients were obtained as a result of ICA algorithm application:

$$T_{RGB} = [-0.25, 0.764, -0.285] \quad (5)$$

After this transformation, a time series of skin tints $s^*(t)$ are obtained that may also contain other noise in addition to the heartbeats signal. So filtering in frequency domain should be applied to $s^*(t)$ to leave only the frequency band that contains the heart signal, and throw away other bands as noise.

4.3. Skin Tints Time Series Frequencies Filtration

This module is needed to separate the useful heartbeat signal from the noise. Considering that frames from the video camera are obtained at different time intervals, it is necessary to process the resulting unevenly sampled series applying the following steps:

- Evenly resample the output series $s^*(t)$.
- Perform frequency filtering of the uniformly resampled series with sampling rate of 30Hz by applying bandpass filter in the range of heartbeat frequencies – from 30/60 Hz to 180/60 Hz.

For non-uniformly sampled time series it is necessary to use only those resampling methods that can work in real time applications and that do not distort the shape of the signal. Therefore, three methods were considered: linear interpolation, interpolation with cubic splines and optimal sinc interpolation. Researching the specifics of the residual noise and the form of the signal $s^*(t)$, it was found that the linear interpolation has the smallest impact on the signal distortion, which is why it was chosen as resampling method in this work. The raw input signal sampling rate varies in the range from 5 Hz to 30 Hz. After the

resampling process the signal with a constant sampling rate equal to 30Hz is obtained.

The frequency filters can be applied to the uniformly sampled time series, leaving only heart frequencies band in the range from 30 beats per minute to 180 beats /min.

Finite impulse response (FIR) filter has shown the best performance. The filters of the moving average, Bessel, Butterworth and Chebyshev were also considered. The authors did not manage to construct Bessel, or Butterworth filters that would satisfy the quality requirements, so the choice was stopped on the family of bandpass finite impulse response filters of the order greater than 60.

Optimal quality/execution time tradeoff was achieved with 61 order of the filter. Such a high filter order is caused by the necessity to obtain the width of the signal bandwidth in diapason 1.5–2 hertz with very sharp transition between passing and blocking filter mode and this requires the amplitude-frequency characteristic (AFC) of the filter to have very similar form with the ideal filter.

The higher the order of the filter, the better is the quality of the filtered signal but on the other hand, the greater time delay in the data processing pipeline is observed. Time delays are undesirable effect for real-time systems but at lower orders of the filter the signal remained too much noise because of too smooth AFC form.

4.4. Heart Beats' Time Detection

The next step after frequency filtering is to get markup of signal peaks corresponding to heart beats.

The filtered signal has the form of a quasiperiodic sinusoid ($E\left[s(t)\right]=0$), where local maxima correspond to the moments of the heart ventricles contraction.

Therefore, it is necessary to determine the moments in time of the appearance of these maxima in the filtered signal, which uses the following algorithm:

1. Find the time intervals T_i, where $s(t) > 0$,
2. For each interval from 1, the global maximum $s^i_{\max} = \max_{t\in T_i} s(t)$ should be computed,
3. Determine the time moment t_i when this global maximum $s^i_{\max}$: $t_i = argmax_{t\in T_i}\left(s^i_{\max} = s(t)\right)$, has appeared,
4. For two consecutive maxima, calculate the RR-interval values as $rr_k = t_{2k+1} - t_{2k}$,
5. If $rr_k < \frac{60}{30}$ or $rr_k > \frac{60}{180}$, where 30 and 180 are responsible for the minimum and maximum heart rate, then mark the RR interval as false,

Thus, in this section, the structure and algorithms used to retrieve human rPPG by processing facial image from the video were described.

The system using wide-spread web camera with minimal frame resolution of 640x480 pixels and average speed of 25 frames per second, detects 99.3% of heart contractions. The experiments have shown that standard deviation of delta time between heart beat contractions' time detected by the system and Holter monitor is 0.046 seconds.

As the output – the system returns time series of RR intervals and as result HRV can be calculated, same as spectrograms of retrieved cardiointervalograms [27–29].

HRV and RR intervals are very useful indicators for estimating user's body regulatory abilities and reactions to external factors.

For example, stress represents a wide range of physical responses that occur as a direct effect of a stressor causing an upset in the homeostasis of the body, and the corresponding state of the nervous system of the body (or the body as a whole). In medicine, physiology, psychology positive (eustress) and negative (distress) forms of stress are distinguished.

5. Stress Index Calculation

Having series of RR-intervals, it is possible to apply variational pulsometry that is used for stress calculation. The essence of variational pulsometry consists in learning the distribution law of cardio intervals. The distributions of cardio intervals are also called histograms. A traditional manner of grouping cardio intervals in the range from 400 to 1300 ms with the buckets' intervals of 50 ms was constituted in perennial practice. Thus, 20 fixed ranges of cardio-intervals' length are considered that allow to compare pulsograms received by different researchers. The timing capacity of pulsograms is set to 5-minute standard.

Cardio interval histogram is a bar plot with buckets' width of 50 ms. Cardio RR-intervals are distributed among these buckets and form columns. The higher the column, the more cardio intervals it includes with duration within the beginning and end time of the bucket. A healthy person with a normal energy potential has symmetrical histogram of pyramidal shape with its central column containing between 30%–50% of all cardio intervals.

According to [30] variational pulsometry is widely practiced in Russia and post-soviet countries and is called "index of regulatory systems tension" or stress index (SI).

Prof. R. Baevskiy has proposed [30] to capture the factors that are caused by stress into one single formula, namely:

$$SI = \frac{AMo}{2 * Mo * MxDMn} \tag{6}$$

where:

Mo (the mode) – is the most frequently occurring value of cardio intervals in milliseconds. Mo differs a little from mathematic expectation (*M*) in the case of normal allocation and high stationarity;

Amo (the mode amplitude) – is the proportion of the most common cardio intervals, from which the central column of the histogram was formed, of all cardio intervals.

MxDMn (the RR-intervals variation range) – the difference between cardio intervals of the minimum and maximum duration.

Stress index calculation is only one of approaches in interpretation and estimation of the variational pulsogram.

This parameter is very sensitive to amplification of sympathetic tone and in norm SI varies within the limits of 80–150. Small load (physical or emotional) increases SI up to 1,5–2 times and for significant loads it increases up to 5–10 times. SI in rest can be equal to 400–600 units for the patients with constant tension of regulatory systems illness and may reach 1000–1500 units for people with coronary heart diseases.

6. Stress Index Measurement Using Developed Remote rPPG System

With the help of developed rPPG system several experiments were conducted to measure people stress level in different situations.

6.1. Drivers' Stress

The rPPG system was used to track HRV and driver stress during many hours of operating cars during summer 2017. The camera was mounted in the car's interior on the celling to have clear view on the drivers' face (see fig. 1). The drivers were around 5-6 hours in the road covering intercity distances and measurement of SI before and immediately after the journey were compared. Also the drivers were asked to fill in small survey with questions about their level of tiredness. The results of comparisons have shown that stress index that was in range 14-32 before the journey was almost twice higher in the end of the route.

Fig. 1. Stress index of the driver after 5-hour intercity journey is twice higher compared to measurement before the journey

The system was also tested on several drivers in the field using various models of tractors. The camera was mounted near the steering wheel (see fig. 2) or on the ceiling depending on the tractor model. It was shown that drivers who were working in more comfortable models of tractors and have been fewer hours on their shift by the moment of experiment conduction had less stress index values.

Fig. 2. The system testing also involved several tractor-drivers in field conditions on tractors of different models

6.2. Students Stress

Another experiment was conducted at IASA NTUU "Igor Sikorsky KPI" involving group of 23 students. The purpose of this experiment was to detect a change in the student's internal state directly before and during the exam at session. The null hypothesis to test was that students would feel calm two days before the exam as they still have two days to prepare. During the exam, everyone will be little worried. Those who fail to pass the exam will be very worried and experience severe stress.

The experiment consisted of two phases: on the first phase students' stress index was measured in calm conditions. At this stage no coming events for the next two days have been planned that could have been treated by students as alert factor. Also, students have not performed any active physical exercises at least for half an hour before the beginning of the experiment that could have affected the measurements. During this stage experimental measurements of SI with the rPPG system have shown average SI to be in the range 12–37 and heart rate to be in range 61–73 bpm.

The second phase of the experiment was conducted during exam. It appeared that students were extremely stressed out. SI was in range 75–261 and the lowest average heart rate per student was 83 while the highest was 145 bpm.

Each experimental recording lasted from four to seven minutes. The goal was to collect at least 300 heartbeats in a row for each recoding to be able to calculate the stress index by Baevsky's method.

After processing the results, it turned out that the null hypothesis was confirmed but not for all students. On average, the stress index level during the exam has been greater than two days before the exam. But there were also students who had normal stress index level during both phases and who were also stressed all the time.

For example, let's look at the obtained data of the student whose heart rate was within the normal range

during both phases of experiment (approximately 67 beats per minute).

The figure 3 shows his RR intervals obtained during the experiment on the exam. They are in range from 630 to 1017 milliseconds. The bar height shows the duration of each of the RR intervals in milliseconds (at the bottom of the bar durations in milliseconds are presented in white). Numbers above the bars indicate the instantaneous pulse (beats per minute). Labels under the abscise axis indicate the time of each RR interval occurrence given in seconds from the beginning of the experiment recording.

Judging from the figure 3, one can assert that the student is in normal condition and has little stress. After calculating the SI, we get the figure 4, which shows that the student's stress was in the range of 20 to 40 conventional units. To calculate the stress formula, the duration of RR-intervals series was taken equal to 150 seconds.

Although most of the students had "poker faces" during this experiment and showed no external signs of their emotions and neither the lecturer nor the stuff conducting the experiment noticed any difference from other students, the rPPG system has detected very high heart rate and big SI values from some of the students. For example, the figure 5 shows RR-intervals of an extremely worried student.

Her pulse was in the range of 110 to 140 beats per minute and averaged 127 beats per minute, indicating high level of adrenaline in the blood and, accordingly, a high level of stress.

At the same time, two days before the exam, she was calm, and her average pulse rate was 65 beats per minute.

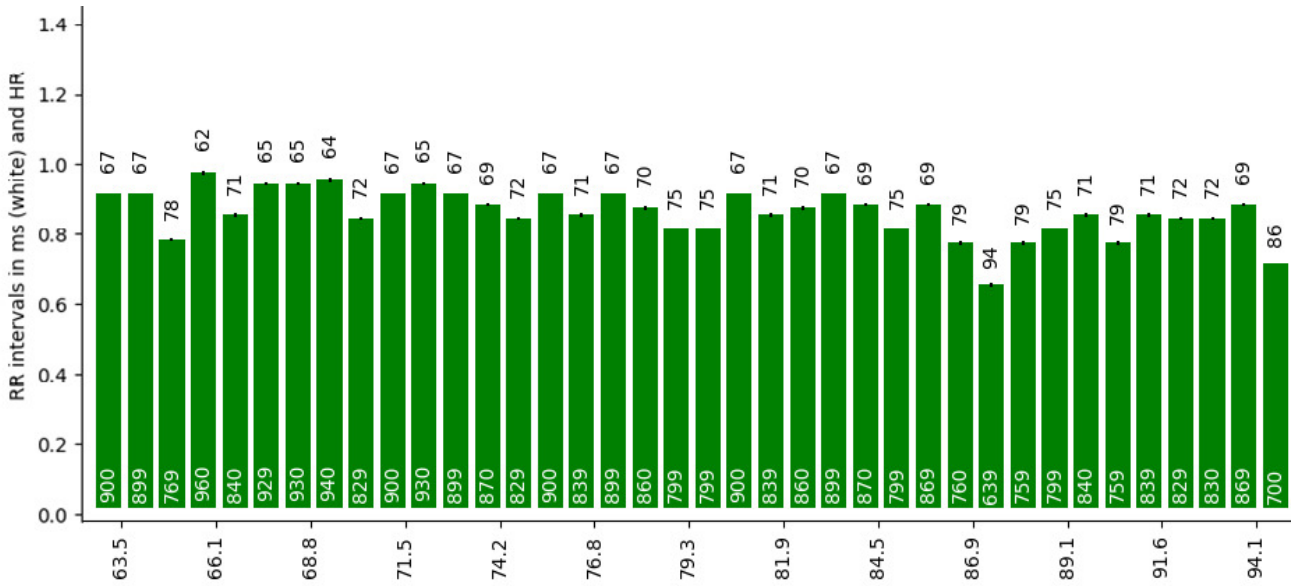

Fig. 3. Example of RR-intervals of the student in time (sec). Signatures over the bars show the instantaneous pulse (beats/minute)

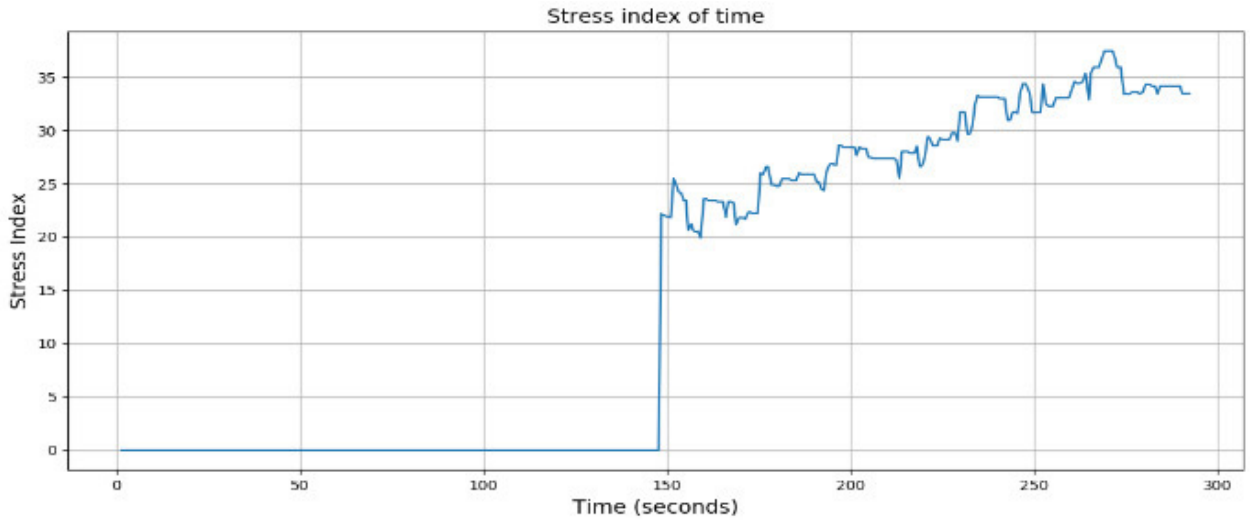

Fig. 4. Stress index from time plot of calm student (conventional units)

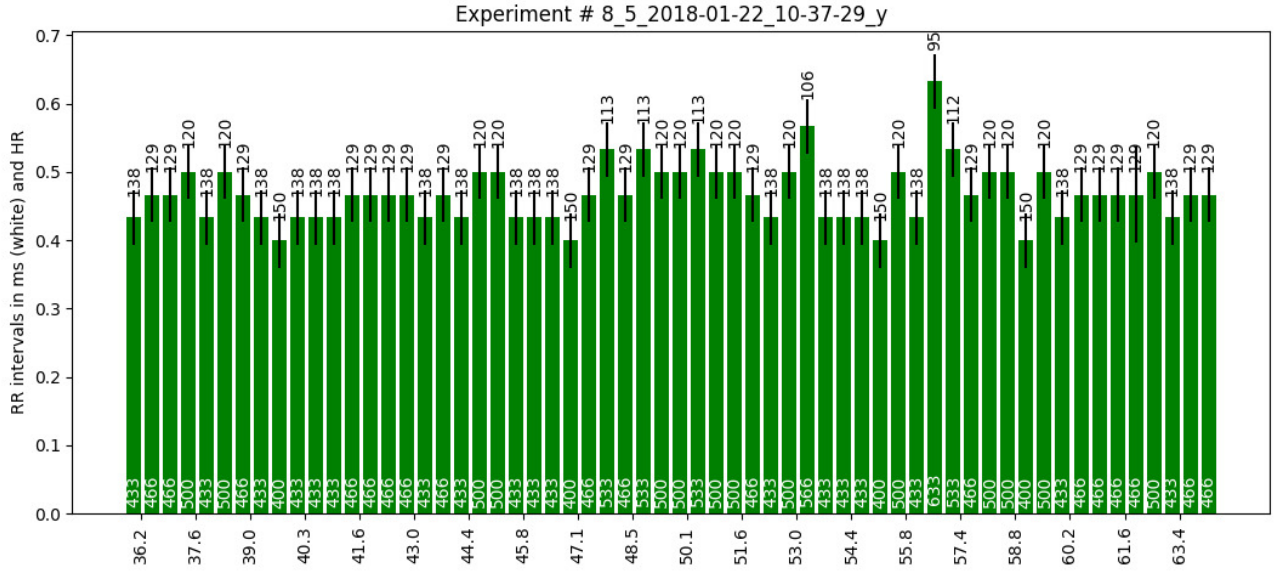

Fig. 5. Example of obtained RR-intervals from experiment during exam)

Using the proposed IT the student's photoplethysmogram raw signal (yellow-colored) and after frequency filtration (violet-colored) were obtained (figure 6). Yellow vertical lines represent moments in time of heart contractions. A good noise to signal ratio of raw signal can be noticed because the amplitude of the signal is much bigger than the amplitude of noise.

This testifies to the high quality and reliability of the obtained PPG and RR-intervals, respectively.

After calculating the stress index, we get the following figure 7, from which it can be seen, that the student's stress was in the range of 125 to 132 conventional units.

The obtained students' levels of stress during the exam were surprising even for professor with many years of experience who was conducting this exam.

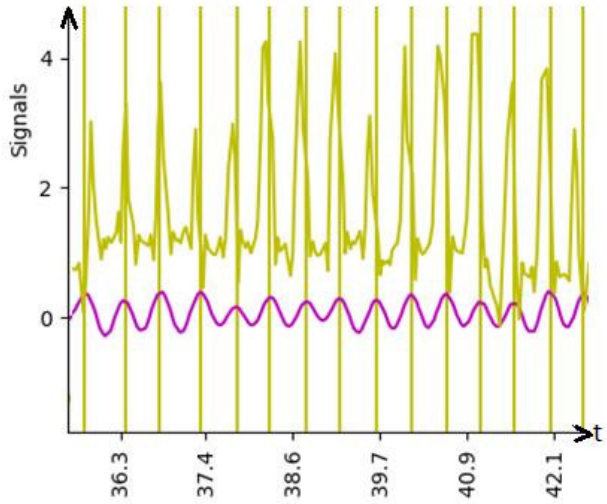

Fig. 6. Student's rPPG heartbeat signal before (yellow) and after filtering (violet) in time (sec)

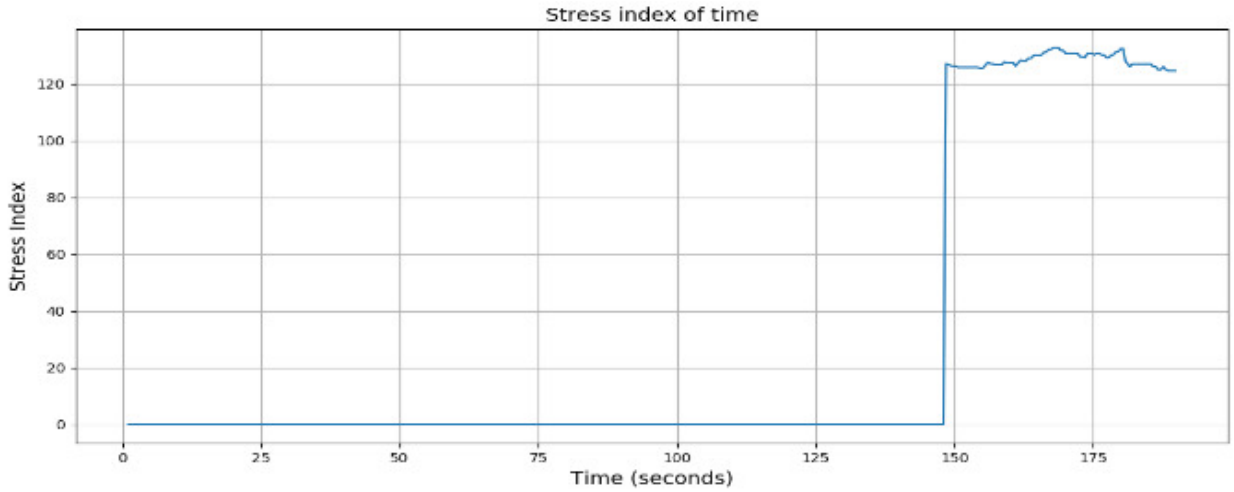

Fig. 7. Stress index from time plot of the stressed student (conventional units)

It was shown that the system determined the pulse and stress levels of the students without any problems and the rPPG system measurements results were in accordance with students' answers from questionnaires about their feelings. Also, after a detailed analysis of the rPPG signals retrieved by the system, their quality was confirmed as high and conclusions about the reliability of the results of the experiments themselves were marked as reliable.

7. Conclusion

The need of the new non-contact widely available sensors for human bio-signals constant monitoring was described within the framework of predictive, preventive, personalized, and participatory medicine. It was shown that the heart rate variability can be a good indicator for estimating internal states of the human including heart rate, stress levels and heart diseases that cause decrease in productivity and financial loses for the enterprises and to the whole economy.

Overview of recent papers and approached was made to show how using widely spread web cameras it is possible to build remote non-contact information technology that can extract precise timings of heart beats from video stream based on remote photoplethysmography. Also, the architecture and main modules of the developed by the authors contactless rPPG system were presented.

Several experiments were described for human stress index calculation in different conditions including car and tractor drivers.

It was shown that despite the absence of visual signs perceivable by other humans, the rPPG system was able to differentiate and measure internal states of the people who were participating in the experiments. The obtained measurements were in accordance with the feelings of participants and the quality of the obtained results was confirmed by in-depth examining of all stages of signal processing within the rPPG system.

In the summary, it can be stated that the developed personal non-contact automatic remote photoplethysmography system for heart beats time moments' contraction (so called RR-intervals) retrieval, can be used in different applications like:

- person's functional and emotional states detection;
- person's Identification & Authentication via remote detection of vital signs presence;
- contactless remote HRV and stress tracking.

Developed rPPG method has shown good performance using ordinary web cameras for online R-peaks detection. The technology misses heart beats at rate 0.69% and detects false positives heart beats at rate 1.16%. Root mean square time deviation between correctly classified heart beats is 0.046 seconds. Using cameras filming at higher frame rates one can greatly decrease these errors.

AUTHORS

Sergii Nikolaiev – Institute for Applied System Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine, e-mail: sergiynicolaev@gmail.com.

Sergii Telenyk* – Department of Electrical Engineering and Computer Science, Cracow University of Technology, Krakow, Poland, e-mail: stelenyk@pk.edu.pl.

Yury Tymoshenko – Institute for Applied System Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine, e-mail: yury.alex.tym@gmail.com.

* Corresponding author

REFERENCES

[1] O. Golubnitschaja, J. Kinkorova and V. Costigliola, "Predictive, Preventive and Personalised Medicine as the hardcore of 'Horizon 2020'", *EPMA Journal*, vol. 5, no. 1, 2014
DOI: 10.1186/1878-5085-5-6.

[2] S. Nikolaiev and Y. Timoshenko, "Reinvention of the cardiovascular diseases prevention and prediction due to ubiquitous convergence of mobile apps and machine learning". In: *2015 Information Technologies in Innovation Business Conference (ITIB)*, 2015, 23–26
DOI: 10.1109/ITIB.2015.7355066.

[3] "Stress and Heart Disease". The American Institute of Stress, www.stress.org/stress-and-heart-disease. Accessed on: 2020-08-11.

[4] M. Malik, "Heart Rate Variability. Standards of Measurement, Physiological Interpretation, and Clinical Use", *Annals of Noninvasive Electrocardiology*, vol. 1, no. 2, 1996, 151–181
DOI: 10.1111/j.1542-474X.1996.tb00275.x.

[5] X. Li, J. Chen, G. Zhao and M. Pietikainen, "Remote Heart Rate Measurement from Face Videos under Realistic Situations". In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, 4264–4271
DOI: 10.1109/CVPR.2014.543.

[6] W. Verkruysse, L. O. Svaasand and J. S. Nelson, "Remote plethysmographic imaging using ambient light", *Optics Express*, vol. 16, no. 26, 2008
DOI: 10.1364/OE.16.021434.

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

[8] M. van Gastel, S. Stuijk and G. de Haan, "Motion Robust Remote-PPG in Infrared", *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 5, 2015, 1425–1433
DOI: 10.1109/TBME.2015.2390261.

[9] K. Humphreys, C. Markham and T. E. Ward, "A CMOS camera-based system for clinical photoplethysmographic applications". In: F. D. Murtagh (eds.), *Proceedings Volume 5823, Opto-Ireland 2005: Imaging and Vision*, 2005
DOI: 10.1117/12.604822.

[10] W. Verkruysse, L. O. Svaasand and J. S. Nelson, "Remote plethysmographic imaging using ambient light", *Optics Express*, vol. 16, no. 26, 2008
DOI: 10.1364/OE.16.021434.

[11] M. Lewandowska, J. Rumiński, T. Kocejko and J. Nowak, "Measuring pulse rate with a webcam — A non-contact method for evaluating cardiac activity". In: *2011 Federated Conference on Computer Science and Information Systems (FedCSIS)*, 2011, 405–410.

[12] M.-Z. Poh, D. J. McDuff and R. W. Picard, "Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam", *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, 2011
DOI: 10.1109/TBME.2010.2086456.

[13] G. de Haan and V. Jeanne, "Robust Pulse Rate From Chrominance-Based rPPG", *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, 2013, 2878–2886
DOI: 10.1109/TBME.2013.2266196.

[14] G. de Haan and A. van Leest, "Improved motion robustness of remote-PPG by using the blood volume pulse signature", *Physiological Measurement*, vol. 35, no. 9, 2014, 1913–1926
DOI: 10.1088/0967-3334/35/9/1913.

[15] W. Wang, S. Stuijk and G. de Haan, "A Novel Algorithm for Remote Photoplethysmography: Spatial Subspace Rotation", *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 9, 2016, 1974–1984
DOI: 10.1109/TBME.2015.2508602.

[16] X. Li, J. Chen, G. Zhao and M. Pietikainen, "Remote Heart Rate Measurement from Face Videos under Realistic Situations". In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, 4264–4271
DOI: 10.1109/CVPR.2014.543.

[17] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. A. Clifton and C. Pugh, "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models", *Physiological Measurement*, vol. 35, no. 5, 2014, 807–831, 10.1088/0967-3334/35/5/807.

[18] W. Wang, S. Stuijk and G. de Haan, "Exploiting Spatial Redundancy of Image Sensor for Motion Robust rPPG", *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, 2015, 415–425
DOI: 10.1109/TBME.2014.2356291.

[19] M. Kumar, A. Veeraraghavan and A. Sabharwal, "DistancePPG: Robust non-contact vital signs monitoring using a camera", *Biomedical Optics Express*, vol. 6, no. 5, 2015
DOI: 10.1364/BOE.6.001565.

[20] S. Tulyakov, X. Alameda-Pineda, E. Ricci, L. Yin, J. F. Cohn and N. Sebe, "Self-Adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, 2396–2404
DOI: 10.1109/CVPR.2016.263.

[21] A. R. Guazzi, M. Villarroel, J. Jorge, J. Daly, M. C. Frise, P. A. Robbins and L. Tarassenko, "Non-contact measurement of oxygen saturation with an RGB camera", *Biomedical Optics Express*, vol. 6, no. 9, 2015
DOI: 10.1364/BOE.6.003320.

[22] I. C. Jeong and J. Finkelstein, "Introducing Contactless Blood Pressure Assessment Using a High Speed Video Camera", *Journal of Medical Systems*, vol. 40, no. 4, 2016
DOI: 10.1007/s10916-016-0439-z.

[23] L. K. Mestha, S. Kyal, Beilei Xu, L. E. Lewis and V. Kumar, "Towards continuous monitoring of pulse rate in neonatal intensive care unit with a webcam". In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, 3817–3820
DOI: 10.1109/EMBC.2014.6944455.

[24] S. Fernando, W. Wang, I. Kirenko, G. de Haan, S. Bambang Oetomo, H. Corporaal and J. van Dalfsen, "Feasibility of Contactless Pulse Rate Monitoring of Neonates using Google Glass". In: *Proceedings of the 5th EAI International Conference on Wireless Mobile Communication and Healthcare – "Transforming healthcare through innovations in mobile and wireless technologies"*, 2015
DOI: 10.4108/eai.14-10-2015.2261589.

[25] J.-P. Couderc, S. Kyal, L. K. Mestha, B. Xu, D. R. Peterson, X. Xia and B. Hall, "Detection of atrial fibrillation using contactless facial video monitoring", *Heart Rhythm*, vol. 12, no. 1, 2015, 195–201
DOI: 10.1016/j.hrthm.2014.08.035.

[26] B. Kaur, S. Moses, M. Luthra and V. N. Ikonomidou, "Remote stress detection using a visible spectrum camera". In: H. H. Szu, L. Dai and Y. Zheng (eds.), *Proceedings Volume 9496, Independent Component Analyses, Compressive Sampling, Large Data Analyses (LDA), Neural Networks, Biosystems, and Nanoengineering XIII*, 2015
DOI: 10.1117/12.2177159.

[27] S. S. Nikolaiev, Y. O. Tymoshenko and K. Y. Matviiv, "Haar Cascade Face Detector Quality Dependence on Training Dataset Variablity", *Research Bulletin of the National Technical University of Ukraine "Kyiv Politechnic Institute"*, no. 6, 2017, 38–46
DOI: 10.20535/1810-0546.2017.6.115181.

[28] S. Nikolaiev and H. Chereda, "Sampling Rate Independent Filtration Approach for Automatic ECG Delineation", *International Scientific Journal "Internauka"*, vol. 27, no. 5, 2017.

[29] S. Nikolaiev, "Metric and algorithm for similarity between two temporal event sequences cal-

culation", *System research and information technologies*, no. 3, 2017, 127–135
DOI: 10.20535/SRIT.2308-8893.2017.3.12.

[30] R. M. Bayevsky and G. G. Ivanov, "Cardiac Rhythm Variability: the Theoretical Aspects and the Opportunities of Clinical Application", *Ultrazvukovaya i funktsionalnaya diagnostika*, no. 3, 2001, 108–127 (in Russian).