Remote Heart Rate Extraction Using Microsoft Kinect™ v2.0

Lakmini Malasinghe
University of the West of Scotland
High St, Paisley, United Kingdom, PA1 2BE
+44 (0) 141 849 4143
Lakmini.Arachchige@uws.ac.uk

Stamos Katsigiannis
University of the West of Scotland
High St, Paisley, United Kingdom, PA1 2BE
+44 (0) 141 849 4143
Stamos.Katsigiannis@uws.ac.uk

Naeem Ramzan
University of the West of Scotland
High St, Paisley, United Kingdom, PA1 2BE
+44 (0) 141 848 3648
Naeem.Ramzan@uws.ac.uk

Keshav Dahal
University of the West of Scotland
High St, Paisley, United Kingdom, PA1 2BE
+44 (0) 141 848 3305
Keshav.Dahal@uws.ac.uk

ABSTRACT
Remote and contactless heart rate detection is still an open research issue of great clinical importance. Available approaches lack the necessary accuracy and reliability for acceptance by medical experts. In this study, we propose a new method for remote heart rate extraction using the Microsoft Kinect™ v2.0 image sensor. The proposed approach relies on signal processing and machine learning methods in order to create a model for accurate estimation of the heart rate via RGB and infrared face videos. Electrocardiography (ECG) recordings and RGB and infrared face videos, captured using the Kinect™ v2.0 image sensor, were acquired from 17 subjects and used to create a machine learning model for remote heart rate detection. Experimental evaluation through supervised regression experiments showed that the proposed approach achieved a mean absolute error of 6.972 bpm, demonstrating the capabilities of the underlying technology.

CCS Concepts
•Applied computing → Consumer health; •Applied computing → Health informatics.

Keywords
Heart rate; Remote heart rate extraction; Remote patient monitoring; Kinect v2.0; RGB sensor; Infrared sensor; Fast fourier transform.

1. INTRODUCTION
The Remote patient monitoring is an emerging research field in the area of engineering applications in healthcare. The world's population is aging fast, i.e. the ratio of the number of aged people to the number of young people keeps increasing [1], [2]. The nature of illnesses has become more complex and the number of people with chronic health conditions in every country keeps increasing [1]. Most chronic illnesses require continuous monitoring, but hospitals cannot cope with this demand as they need to allocate their limited resources to emergencies and patients with more critical conditions. Therefore, to address the ever-increasing demand for better healthcare, the number and the resources of healthcare facilities (hospitals, health clinics, ambulances, health visitors, etc.) must be increased. In order to meet this demand, a huge cost must be borne by patients and states and as a result, less costly alternatives, such as remote patient monitoring, are preferred.

ICBBT '18, May 16–18, 2018, Amsterdam, Netherlands

ACM ISBN 978-1-4503-6366-2/18/05…$15.00
DOI: https://doi.org/10.1145/3232059.3232060

The main purpose of a remote patient monitoring (RPM) system is to obtain vital signs from a patient for prolonged periods of time, process and analyse them for early detection of various health conditions, and/or monitor existing health conditions such as chronic illnesses [3]. This information is then sent to a doctor/medical practitioner as per the requirement. Conventional RPM systems suffer from a major disadvantage, i.e. the requirement of attaching devices to the patient, which limits his/her mobility and freedom for daily activities. To address this limitation, research is moving towards contactless monitoring [4], i.e. the use of sensors that do not need to be attached on the patient. Contactless monitoring has two main subcategories: a) radar-based methods, and b) image-based methods. Out of these two, image-based methods are more popular due to advancements in their underlying technology and the wide usage and availability of camera technology [5].

Out of the many vital signs that can be monitored using the existing technology, heart rate is dominant as it can reveal abnormalities in the human body caused by many different types of illnesses. Furthermore, heart related chronic illnesses are the main cause of mortality [3]. In this work, a system for obtaining the heart rate using a Microsoft Kinect™ v2.0 sensor is presented. The proposed system uses the camera's colour and infrared (IR) channels in combination with a low complexity custom algorithm in order to extract an accurate estimate of the heart rate of a person. The proposed algorithm utilises the in-built face detection and tracking features of the Kinect™ sensor in order to limit the search area into specific regions of interest (ROIs) inside the face area. These methods have been able to mitigate effects of small head motions, thereby solving the issue of motion artefacts to a certain extent.

The rest of this paper is organised in four sections. In Section 2, a review of some existing recent systems and technologies for contactless image-based heart rate monitoring is given. Section 3 provides a detailed description of the proposed system and methodology, whereas Section 4 presents the results and

discussion of the experimental evaluation. Finally, conclusions and feature research directions are drawn in Section 5.

2. BACKGROUND

The main idea behind contactless monitoring using a camera is to capture and observe visual changes that are related to heart rate. Some of the physical changes that occur with a heartbeat are bobbing of the head, small movements in veins, pupillary fluctuations in eyes and skin colour changes [6]-[8]. The camera captures some of these changes and by analysing these changes over time it is possible to estimate the cause of the change, i.e. the heart rate.

The concept of contactless heart rate monitoring was demonstrated as early as 1995 by Da Costa [9]. In this paper, two optical methods to detect heart rate are shown. One method is an invasive method, whereas the other method is contactless. The invasive method uses a mirror attached to the body which is illuminated by a laser beam. The mirror's movement from vein displacement by dilatational waves under the skin (movement of blood with every heart beat) causes changes in the reflected light back to the camera. Analysis of speckle pattern gives the heart rate. In the contactless method, the skin is directly illuminated from the laser beam and a similar analysis is done. This method however suffered from large processing times and cumbersome digitising procedure due to the technology available at that time. Since then, there have been many efforts to detect heart rate and heart rate variability. Most approaches use image-based methods.

McDuff et al. [5] reviewed photoplethysmographic imaging methods for non-contact remote health monitoring targeting various illnesses, including heart related ones. The systems discussed use photoplethysmography (PPG) based imaging techniques. Heart related systems used colour signals, captured via a camera, which are then processed to recover the blood volume pulse signal out of which physiological data such as pulse rate, pulse rate variability, respiration rate and blood oxygen levels can be obtained. A recent system for remotely assessing heart rate by skin colour processing is described in [10]. In that work, principal component analysis (PCA), empirical mode decomposition, and the Hilbert Huang Transform have been used in order to estimate the heart rate. Monkaresi et al. [11] presented a method for heart rate measurement using a web-camera, while Lukáč et al. [12] used a web camera in order to estimate respiration rate. In another work, Gupta et al. [13] proposed a system utilising a web camera that has been tested on 20 subjects and achieved very good results (a mean error of 1.8 bpm). Their proposed approach uses a fusion of Eulerian and Lagrangian methods in order to obtain a good PPG signal.

Similarly, a Kinect™ camera has been used in [14] for measurement of heart rate and respiration rate, while in [15], Mishra et al. presented a blood pulse measurement system using linearly polarized light. The Kinect™ v2.0 camera has also been used in [16] for remote detection of heart rate. After filtering the RGB data, blind source separation (BSS) is performed using PCA. Then, the Fast Fourier Transform (FFT) is performed on the PCA output. The study demonstrated promising results; however, occasionally the system showed very large errors. Procházka et al. [17] used the Kinect™ v2.0 colour, infrared, and depth sensors to analyse breathing and heart rate. The chosen ROI was the mouth area, which is an uncommon choice compared to other works. The infrared and depth sensors were used for heart rate analysis, with three subsections on the lip considered for comparison of the results. This work achieved a very low error rate, but seemingly no information has been given to establish how many subjects participated in the study. Bakhtiyari et al. [18] also used the Kinect™ sensor in order to measure heart rate variability (HRV) via colour and depth data. They used independent component analysis (ICA) on all colour and depth signals in order to find the optimal signal/data that corresponds to HRV. The system's results are good; however, it was tested only on five male subjects.

While lots of works have been proposed for the problem of remote heart rate monitoring, it is evident that it still remains an open issue since accuracy and reliability are not sufficient for use in a clinical environment.

3. METHODOLOGY

3.1 Signal Acquisition

The data acquisition device used in this work was the Kinect™ for Windows v2.0 sensor/camera, which is equipped with a 1920x1080 - 1080p (30 Hz) colour camera, a 512x424 (30 Hz) infrared (IR) sensor, and a 512x424 (30 Hz) depth sensor, and is capable of body and face detection and tracking. The camera was placed on a still surface while subjects sat in front of the camera at a distance of 1 m. The subjects were asked to remain as still as possible (without large head or facial movements excluding non-voluntary muscle movements and blinking of eyes). Using the in-built face detection and tracking mechanism of the Kinect™ sensor, still images of the face region of the frame were recorded for one minute using both the colour (1920x1080) and IR (524x424) sensors, at a rate of approximately 30 samples per second. The coordinates for the eyes, nose, and mouth, provided by the Kinect™ sensor, were also recorded for each captured image, along with their corresponding timestamps. For the off-line analysis, the raw images were encoded (without loss) as TIFF images so that no information would be lost due to compression. It must be noted that for the real-time version of the proposed system, the storing of all the captured images is not necessary since only the regions of interest need to be stored in buffer memory until they get analysed.

A Shimmer™ [19] wireless 4-lead electrocardiography (ECG) sensor was used in order to obtain an ECG signal with a sampling frequency of 256 Hz, concurrently with the image acquisition. The real heart rate of the participants during the experiment was then computed from the ECG recordings in order to establish the ground-truth for the validation of the proposed method. Timestamps were also recorded for all the captured ECG samples. Both the imaging and the ECG data were recorded on a single PC (Intel Core i5-4590 CPU, 3.3 GHz, 8 GB RAM, Windows 8 64-bit OS) and synchronisation was achieved by means of timestamps with millisecond precision. An application in C# using the Kinect™ v2.0 SDK was developed in order to acquire the signals using the Kinect™ sensor, while MATLAB 2016b [20] and the Shimmer™ MATLAB Instrument Driver was used for capturing the ECG signals.

3.2 Subjects

A total of 17 healthy subjects (11 male, 6 female) between the age of 23-47 (μ = 31.625, σ = 5.349) of Asian, African, and Caucasian origin participated in the study. The subjects had varying facial features (beards, spectacles, cosmetics etc.), further increasing the diversity of the dataset and better resembling a real operating scenario. Furthermore, data acquisition took place indoors, in the same environment for all subjects. Nevertheless, since lighting consisted of both indoor lighting (fluorescent light) as well as a smaller portion of sunlight (through a roof-top window), lighting

conditions for each subject differed slightly depending on the time and the weather. As a result, the acquired data are closer to real-world conditions than to a completely controlled environment.

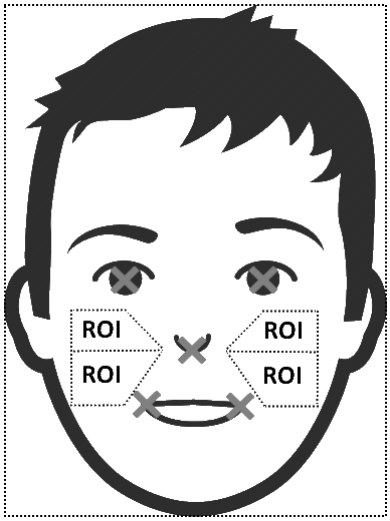

Figure 1. The regions of interest (ROIs) selected from the face region and the eye, nose, and mouth positions returned by the Kinect™ sensor.

Ethical approval from the *Ethics Committee* of the *University of the West of Scotland* was obtained in order to conduct these experiments. Written consent for the participation in the study and for the anonymous use of the collected data was also obtained from each participant prior to the experiment. Participation was voluntarily and participants were informed that they may stop the experiment at any time.

3.3 Data Processing

Data processing involves three sub tasks; ROI computation, blood volume pulse (BVP) signal extraction, and deriving actual heart rate from ECG data.

3.3.1. Region of Interest Computation

After acquiring the facial images, the coordinates of the eyes, nose, and mouth (as shown in Figure 1) are used in order to define the ROIs that would be subsequently analysed. In order to reduce the impact of hair, facial hair, glasses, and headscarves, the ROIs were limited to four regions, two in each cheek, as shown in Figure 1. Computation of the ROIs within the right cheek is performed as follows:

Since face detection is performed using rectangular detection areas, the acquired images contain pixels that do not belong to the actual face, especially at the corners and at the borders of the image. A number of pixels d is selected to be used as an offset between the ROIs and the boundaries of the image, the eye sack below the eyes, the corners of the mouth and the sides of nose. The offset d is then added to or subtracted from the coordinates of the facial landmarks in order to denote the ROIs. Let the size of the image be M x N (*rows* x *columns*) and (x_i, y_i) be the coordinates of facial landmark i. The ROIs within each cheek are defined by two trapezoidal regions with their corner pixels defined as shown in Table 1.

ROI computation is performed twice for each captured frame; once for the colour image and once for the infrared image. It must be noted that the offset d and the coordinates of the facial landmarks are different between the colour and the infrared images due to the difference in resolution. After ROI computation, the brightness values of the pixels within the ROIs are then jointly stored in one data array for each of the Red, Green, Blue, and IR channels.

Table 1. Coordinates of corner pixels of ROIs

ROI	Corner pixels
Right cheek - Upper	$(x_{right\_eye} + d\ ,\ y_{right\_eye})$ $(x_{right\_eye} + d\ ,\ N - d)$ $(x_{nose}\ ,\ y_{nose} + d)$ $(x_{nose}\ ,\ N - d)$
Right cheek - Lower	$(x_{nose}\ ,\ y_{nose} + d)$ $(x_{nose}\ ,\ N - d)$ $(x_{right\_mouth}\ ,\ y_{right\_mouth} + d)$ $(x_{right\_mouth}\ ,\ N - d)$
Left cheek - Upper	$(x_{left\_eye} + d\ ,\ y_{left\_eye})$ $(x_{left\_eye} + d\ ,\ d)$ $(x_{nose}\ ,\ y_{nose} - d)$ $(x_{nose}\ ,\ d)$
Left cheek - Lower	$(x_{nose}\ ,\ y_{nose} - d)$ $(x_{nose}\ ,\ d)$ $(x_{left\_mouth}\ ,\ y_{left\_mouth} - d)$ $(x_{left\_mouth}\ ,\ d)$

Note: Left and right refer to the position in relation to the image and not to the actual position of the real object.

3.3.2. Blood Volume Pulse Signal Extraction

After acquiring the brightness values of the pixels from each channel, a 256-bin histogram is created for each of them. The histograms from each frame are then normalised (each bin value divided by the total number of pixels) and the difference between the histograms of each frame and the previous is then computed. A set of four difference signals (D_R, D_G, D_B, D_{IR}) is now available for each frame, except for the first. The Fourier Transform of each difference signal is then computed using the Fast Fourier Transform in order to decompose each signal to its frequencies. The highest frequency from each of the four channels is then selected as the representation of the processed frame. Using these representations and the timestamps referring to each frame, four time series I_R, I_G, I_B, I_{IR}, are created, one for each of the four original image channels.

A peak detection algorithm is then used in order to count the local maxima within the four time series. A threshold for accepting a local maximum as information pertaining the heart rate was set to the mean amplitude of the time series. Consequently, any peak with an amplitude lower than the mean amplitude was not taken

into consideration when counting the local maxima. The number of local maxima per minute is then returned as a feature to be used in a regression model for the prediction of the actual heart rate.

Table 2. Error rates for the single channel approach and various regression methods

Regression method	Red		Green		Blue		Infrared	
	MAE	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**
SVM (Linear)	7.550	10.453	**7.279**	**9.839**	**6.972**	10.852	7.474	**9.973**
K-Nearest Neighbour	12.366	16.247	9.418	11.746	8.542	12.111	10.82	14.424
Linear Fitting	7.379	10.530	7.476	10.080	7.283	11.193	7.415	10.093
Stepwise Linear regression	**7.295**	**10.373**	7.636	11.170	7.295	**10.373**	**7.295**	10.373

Note: Bold values refer to the best values for each channel and metric.

3.3.3. Regression

The obtained peak counts from the I_R, I_G, I_B, and I_{IR}, signals, as well as the real heart rate computed from the ECG recording from each of the subjects participating in this study were then used in order to create and evaluate a machine learning model that could accurately predict the heart rate. The performance for each image channel using various regression techniques (linear support vector machines, *k*-nearest neighbour, linear regression, step-wise linear regression) was evaluated using leave-one-out cross validation in order to avoid overfitting the proposed models. Results are reported only for the test samples of each fold.

An outline of the proposed methodology is given on Figure 2 in the form of a flowchart.

4. RESULTS AND DISCUSSION

The predicted heart rate from the proposed regression models was evaluated in terms of the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE) compared to the actual heart rate of the subjects. Results for each channel and each regression method used are reported in Table 2. The best performance in terms of the MAE (6.972 bpm) was achieved using the signal acquired from the blue channel (I_B) and using linear SVM regression. It is worth pointing out that in terms of the MAE, the best results for the I_R, I_G, I_{IR} signals are similar (σ = 0.0092) and their difference from the best result for I_B is 0.323, 0.307, and 0.323 bpm respectively. The best performance in terms of the RMSE (9.839) was achieved using the signal acquired from the green channel (I_G) and using linear SVM regression. Considering that the MAE achieved for I_G and I_{IR} is marginally worse (+0.307 bpm and +0.323 bpm) than the one achieved for I_B and that the RMSE achieved for I_G and I_{IR} is better than the one achieved for I_B (-1.013 and -0.879 respectively), it seems that using the blue channel leads to larger prediction errors than when the green or the infrared channel are used. Furthermore, the green channel provides slightly better performance compared to the infrared channel, as shown in Table 2.

Apart from only using single channel information for creating the regression models, we evaluated the performance of models using all the different combinations of channels. The multichannel approach underperformed in terms of the MAE compared to the single channel approach, with the best results reaching a MAE of 7.413 bpm for the combination of I_G and I_{IR} using linear multivariate regression. The combination of I_G, I_B, and I_{IR} provided the second best MAE (7.471 bpm). In terms of the RMSE, the best result (8.898) was obtained for the combination of all channels. Detailed results for the multichannel approach are given in Table 3. From this table, it is evident that the (I_G, I_{IR}), (I_R, I_G, I_{IR}), (I_G, I_B, I_{IR}), and (I_R, I_G, I_B, I_{IR}) combinations provide very similar performance (σ_{MAE} = 0.0992, σ_{RMSE} = 0.055) that is marginally worse than the single channel approach in terms of the MAE, but better than the single channel approach in terms of the RMSE. This shows that the multichannel approach is less prone to large errors, thus making it more suitable for the proposed application.

Table 3. Error rates for the single channel approach and various regression methods

#	Channels				MAE	RMSE
2	R	G	-	-	9.543	11.019
	R	-	B	-	13.727	15.68
	R	-	-	IR	9.083	10.957
	-	G	B	-	9.782	11.613
	-	G	-	IR	**7.413**	9.030
	-	-	B	IR	9.064	10.98
3	R	G	B	-	9.310	10.935
	R	G	-	IR	7.616	8.968
	R	-	B	IR	8.166	10.316
	-	G	B	IR	7.471	8.987
4 (all)	R	G	B	IR	7.601	**8.898**

R: Red, G: Green, B: Blue, IR: Infrared

An interesting observation is that the acquired results agree with previous works which demonstrated that the green or the infrared channel provide the best results for heart rate extraction, while earlier work suggested that red and blue channels may have complementary information [8]. The results of the proposed single channel approach are consistent with these findings, while results for the multichannel approach showed that although the combination of the green and the infrared channels provides one of the best performances, the addition of the red and/or blue channel provides a minor improvement (in terms of the RMSE).

The proposed heart rate extraction method achieved satisfactory results. Table 4 presents the evaluation results for other image-based heart rate extraction methods proposed in the literature. Since some of these works report the mean error instead of the MAE or the RMSE, we computed the mean error for the configuration of the proposed method that achieved the lowest MAE (I_B, linear SVM), which was -1.45. Gupta et al. [13] reported a mean error of 1.80 for their proposed system, while Li et al. [21] achieved a mean error of -3.30. Furthermore, Bosi et al. [16] reported a MAE equal to 10.054. It is evident that the proposed approach performed better than the examined approaches.

Figure 2. Flowchart of the proposed heart rate extraction method.

Table 4. Performance compared to other methods

Method	# of subjects	Mean error	MAE
Bosi et al. [16]	30	n/a	10.054
Gupta et al. [13]	20	1.80	n/a
Li et al. [21]	10	-3.30	n/a
Proposed	17	-1.45	6.972

5. CONCLUSION AND FUTURE WORK

In this work, the authors proposed and evaluated a heart rate extraction method based on RGB and infrared facial video input from the Microsoft Kinect™ v2.0 sensor. Different regression models were trained and tested using signals acquired from the red, green, blue, and infrared channel of the sensors. The experimental evaluation showed that the green and the infrared channels provided the best results for the single channel approach, with the green channel offering marginally better performance than the infrared. This finding shows that for a single channel approach, an infrared sensor is not required for achieving good performance. Nevertheless, the infrared channel was used in all the channel combinations that provided the best results for the multichannel approach. Since the multichannel approach provided better results than the single channel approach, it is evident that the infrared channel benefits the examined application. Furthermore, while the combination of the green and the infrared channels achieved good performance, the marginal improvement by the addition of the red and/or blue channel showed that that the red and blue channels may also contain heart rate information, as proposed in other works. As a result, we further conclude that all the colour channels may contain heart rate information.

The proposed method exhibited resilience in relation to the variability of the input data, which included multiple facial/skin variations, including natural features such as various skin tones, beards, imperfect/uneven skin textures and additional features such as cosmetics and spectacles. Furthermore, the proposed method worked well with the naturally occurring small and involuntary facial movements such as blinking of eyes and soft bobbing of head, or vibration of blood vessels and muscles due to cardiac activity.

Future work will include filtering of the input data in order to remove noise and increase the accuracy of the prediction, as well as an updated mechanism for selecting the regions of interest within the face area. Furthermore, the use of signal processing approaches instead of regression will also be explored in order to further reduce the computational complexity of the proposed method.

6. ACKNOWLEDGMENTS

We would like to thank all the participants for their support and voluntary participation in our experiments. Funding for this research was provided by the Erasmus Mundus Action 2 - SmartLink project.

7. REFERENCES

[1] United Nations. 2002. World Population Ageing: 1950-2050. (2002). http://www.un.org/esa/population/publications/worldageing19502050/pdf/80chapterii.pdf

[2] Lutz, W., Sanderson, W., and Scherbov, S. 2008. The coming acceleration of global population ageing. *Nature* 451, 7179 (Feb 2008), 716–719. DOI= https://doi.org/10.1038/nature06516

[3] Silva, B. M.C., Rodrigues, J. J.P.C., de la Torre Díez, I., López-Coronado, M., and Saleem, K. 2015. Mobile-health: A review of current state in 2015. *Journal of Biomedical Informatics* 56, Supplement C (2015), 265–272. 1532-0464 DOI= https://doi.org/10.1016/j.jbi.2015.06.003

[4] Malasinghe, L. P., Ramzan, N., and Dahal, K. 2017. Remote patient monitoring: a comprehensive study. *Journal of Ambient Intelligence and Humanized Computing* (Oct 2017). 1868-5145 DOI= https://doi.org/10.1007/s12652-017-0598-x

[5] McDuff, D. J., Estepp, J. R., Piasecki, A. M., and Blackford, E. B. 2015. A survey of remote optical photoplethysmographic imaging methods. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 6398–6404. 1094-687X DOI= https://doi.org/10.1109/EMBC.2015.7319857

[6] Balakrishnan, G., Durand, F., and Guttag, J. 2013. Detecting Pulse from Head Motions in Video. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*. 3430–3437. 1063-6919 DOI= https://doi.org/10.1109/CVPR.2013.440

[7] Verkruysse, W., Svaasand, L. O., and Nelson, J. S. 2008. Remote plethysmographic imaging using ambient light. Opt Express 16, 26 (Dec 2008), 21434–21445.

[8] Giovangrandi, L., Inan, O. T., Wiard, R. M., Etemadi, M., and Kovacs, G. T. A. 2011. Ballistocardiography: A method worth revisiting. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 4279–4282. 1094-687X DOI= https://doi.org/10.1109/IEMBS.2011.6091062

[9] Da Costa, G. 1995. Optical remote sensing of heartbeats. *Optics Communications* 117, 5 (1995), 395 – 398. 0030-4018 DOI= https://doi.org/10.1016/0030-4018(95)00181-7

[10] Bogdan, G., Radu, V., Octavian, F., Alin, B., Constantin, M., and Cristian, C. 2015. Remote assessment of heart rate by skin color processing. In *2015 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)*. 112–116. DOI= https://doi.org/10.1109/BlackSeaCom.2015.7185097

[11] Monkaresi, H., Calvo, R. A., and Yan, H. 2014. A Machine Learning Approach to Improve Contactless Heart Rate Monitoring Using a Webcam. *IEEE Journal of Biomedical and Health Informatics* 18, 4 (July 2014), 1153–1160. 2168-2194 DOI= https://doi.org/10.1109/JBHI.2013.2291900

[12] Lukáč, T., Púcik, J., and Chrenko, L. 2014. Contactless recognition of respiration phases using web camera. In 2014 *24th International Conference Radioelektronika*. 1–4. DOI= https://doi.org/10.1109/Radioelek.2014.6828427

[13] Gupta, P., Bhowmick, B., and Pal, A. 2017. Serial fusion of Eulerian and Lagrangian approaches for accurate heart-rate estimation using face videos. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2834–2837. 1557-170X DOI= https://doi.org/10.1109/EMBC.2017.8037447

[14] Bernacchia, N., Scalise, L., Casacanditella, L., Ercoli, I., Marchionni, P., and Tomasini, E. P. 2014. Non contact measurement of heart and respiration rates based on Kinect . In *2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*. 1–5. DOI= https://doi.org/10.1109/MeMeA.2014.6860065

[15] Mishra, D., Gogna, G., Barsaiyan, A., and Sarkar, M. 2015. Blood Pulsation Measurement Using Linearly Polarized Light. *IEEE Sensors Journal* 15, 8 (Aug 2015), 4488–4495. 1530-437X DOI= https://doi.org/10.1109/JSEN.2015.2421553

[16] Bosi, I., Cogerino, C., and Bazzani, M. 2016. Real-time monitoring of heart rate by processing of Microsoft Kinect 2.0 generated streams. In *2016 International Multidisciplinary Conference on Computer and Energy Science (SpliTech)*. 1–6. DOI= https://doi.org/10.1109/SpliTech.2016.7555944

[17] Procházka, A., Schätz, M., Vyšata, O., and Vališ, M. 2016. Microsoft Kinect Visual and Depth Sensors for Breathing and Heart Rate Analysis. *Sensors* 16, 7 (June 2016), 996. 1424-8220 DOI= https://doi.org/10.3390/s16070996

[18] Bakhtiyari, K., Beckmann, N., and Ziegler, J. 2017. Contactless heart rate variability measurement by IR and 3D depth sensors with respiratory sinus arrhythmia. *Procedia Computer Science* 109, Supplement C (2017), 498 – 505. 1877-0509 DOI= https://doi.org/10.1016/j.procs.2017.05.319 8th International Conference on Ambient Systems, Networks and Technologies, ANT-2017 and the 7th International Conference on Sustainable Energy Information Technology, SEIT 2017, 16-19 May 2017, Madeira, Portugal.

[19] Burns, A., Greene, B. R., McGrath, M. J., O'Shea, T. J., Kuris, B., Ayer, S. M., Stroiescu, F., and Cionca, V. 2010. SHIMMER - A Wireless Sensor Platform for Noninvasive Biomedical Research. *IEEE Sensors Journal* 10, 9 (Sept 2010), 1527–1534. 1530-437X DOI= https://doi.org/10.1109/JSEN.2010.2045498

[20] Mathworks. 2016. Matlab 2016b. (2016). https://uk.mathworks.com/products/matlab.html

[21] Li, X., Chen, J., Zhao, G., and Pietikäinen, M. 2014. Remote Heart Rate Measurement from Face Videos under Realistic Situations. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 4264–4271. 1063-6919 DOI= https://doi.org/10.1109/CVPR.2014.543