

Non-contact Heart Rate Monitoring Analysis from Various Distances with different Face Regions

Norwahidah Ibrahim¹, Razali Tomari², Wan Nurshazwani Wan Zakaria³, Nurmiza Othman⁴

^{1,2,3}Department of Robotic and Mechatronics, Faculty of Electric & Electronics, University Tun Hussein Onn Malaysia, Malaysia

⁴Department of Electronics, Faculty of Electric & Electronics, University Tun Hussein Onn Malaysia, Malaysia

Article Info

Article history:

Received Aug 1, 2017

Revised Oct 28, 2017

Accepted Nov 12, 2017

Keyword:

Facial tracking

Independent component

analysis

Non-contact heart rate

photoplethysmograph

ABSTRACT

Heart rate (HR) is one of vital biomedical signals for medical diagnosis. Previously, conventional camera is proven to be able to detect small changes in the skin due to the cardiac activity and can be used to measure the HR. However, most of the previous systems operate on near distance mode with a single face patch, thus the feasibility of the remote heart rate for various distances remains vague. This paper tackles this issue by analyzing an optimal framework that capable to works under the mentioned issues. Initially, plausible face landmarks are estimated by employing cascaded of regression mechanism. Next, the region of interest (ROI) was constructed from the landmarks in a face location where non rigid motion is minimal. From the ROI, temporal photoplethysmograph (PPG) signal is calculated based on the average green pixels intensity and environmental illumination is separated using Independent Component Analysis (ICA) filter. Eventually, the PPG signal is further processed using series of temporal filter to exclude frequencies outside the range of interest prior to estimate the HR. As a conclusion, the HR can be detected up to 5 meters range with 94% accuracy using lower part of face region.

Copyright © 2017 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Razali Tomari,

Department of Robotic and Mechatronics,

University Tun Hussein Onn Malaysia,

Parit Raja, 86400 Batu Pahat, Johor, Malaysia.

Email: mdrazali@uthm.edu.my

1. INTRODUCTION

There is growing interest in measuring HR without contact, particularly for populations such as an elderly with fragile skin and premature infant. One of the prominent non-contact HR methods is by measuring the variation of reflected light intensity from the human skin that known as plethysmograph (PPG) signal. Basically PPG make use of a pulse oximeter to measure the light absorption rate as a result of light illuminates the skin. Since the amount of light absorption is proportional to the blood volume, the PPG signal can be used to measure the blood volume pulse. Some improvement of this concept has been done, however a special lighting condition is necessary such as from specific narrowband wavelength or from a special sensor [1].

Recent studies have shown a possibility to infer the PPG signal from a color-based method via commercial web cameras [2], [3]. For instance, works in [3] measure a HR from recorded face videos by detecting head location using Haar cascade method [8], computed the mean value of the three color channels inside the head region, run the Independent Component Analysis (ICA) to separate the PPG from the color channels, and eventually transform the signal to the frequency domain for finding maximum power strength with the range of 0.7Hz to 4Hz that indicate the HR frequency.

Instead relying on three channel of color traces, it has been reported in [4] that the green color channel provide strongest PPG signal since hemoglobin light absorption is most sensitive to oxygenation changes for green light compare to blue and red. However recent work in [3], claimed that single green trace, provide more accurate ICA outcome to separate between source and noise which in contrast with [5], which found that ICA slightly dropped the performance when using green trace alone. A head motion based method for pulse detection is proposed by [6] which show a promising outcome for translating a subtle motion into the HR estimator.

Most of the mentioned methods work well in controlled laboratory environment. However, their test subjects remain in static with near distance with the acquisition device. In real scenario it is common that the subject would move and also the distance would vary from the camera. The recent work in [7-10] addressing the motion issue by incorporating combination of facial landmark extractions, tracker and various illumination rectification. However, still they did not test the system various acquisition distances scenario in which the strength of remote HR reside since the task is daunting for the contact based HR system. In this paper, we investigate the capability of remote HR for acquiring data from various distances and provide an optimal framework that capable to satisfy the constraint.

2. SYSTEM OVERVIEW

Figure 1 shows overview of the proposed framework that consist of five main steps namely, facial tracker, signal extraction, Blind Source Separation (BSS) using ICA, filtering and lastly, histogram analysis.

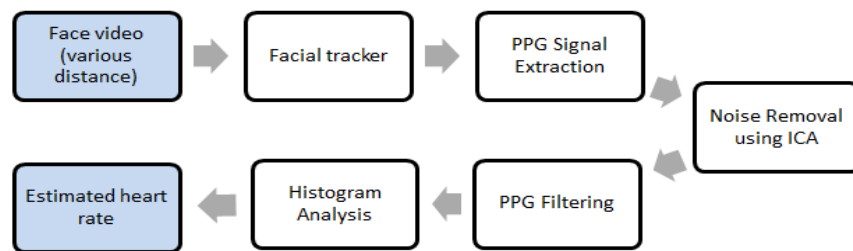


Figure 1. Overall system flow for the proposed framework

The input for this system is face video that was recorded by using a video camera and the output is the HR estimation reading that is measured by using beat per minute (bpm) unit. For the transformation process from input to the output, initially the facial tracker is applied to the image sequences. This facial tracker would track rigid and non-rigid facial landmark, and 49 points of important facial point would be displayed. The 49 points of human facial landmarks produced will be used for constructing region of interest (ROI). Next PPG signal is extracted from the constructed ROI based on temporal random trace information of green color channel. Since the signal acquired is usually contain noise generated from the environment, an ICA based BSS strategy is used to separate the primer PPG signal and the unwanted information. Following that, the primer PPG signal is undergo sequence of filtering step to obtain the optimal state of signal power spectral density (PSD) information and consequently the HR. However, relying on single sequence of HR reading is still subject to the measurement variation. To overcome this, a histogram analysis of repetitive HR reading is constructed based on the same ROI with different random traces. Eventually, the HR is estimated from the average of histogram with lowest variation reading.

2.1. Facial Detection

Facial tracker is used in this system to detect the face and extract facial landmarks information. To satisfy the face detection requirement, a set of validated face regions is extracted using AdaBoost-based cascade with Haar like feature [11]. This classifier works by constructing a strong classifier (positive images) as linear combination weak classifiers (negative images). During detection, a series of classifiers are applied to every image sub-window with different scaling factor. Regions are considered valid if they pass through all the classifier stages.

From the detected face region, facial landmarks are detected by using cascaded of pose regression model [12]. The method works perform a raw initial guess of facial landmark positions and uses a cascade of regressors to infer the shape as whole and explicitly minimizes the alignment error over the training data. Let

$S = (x_1, x_2, \dots, x_p)$ denotes the coordinate of all p facial landmarks in a bounding box I and $r_t(\dots)$ be the regressor cascade, the current shape estimate of the facials location can be estimated by using following formulation :

$$S^{(t+1)} = S^{(t)} + r_t(I, S^{(t)}) \quad (1)$$

The critical point of the cascade is that the regressor makes its predictions based on features, such as pixel intensity values and indexed relative to the current shape estimate. Each regressor in the cascade returns a vector which is used to update the current shape estimate in an additive manner. The method generally does not build any parametric model; instead merely study the correlations between image features to infer a facial shape. To train the regressor, each ground truth shape in the training set is transformed to the mean shape which is rescaled and centered at the origin. Sample of the detected 49 points facial landmarks is shown in Figure 2(a).

From the annotated facial landmarks location, a ROI of face region was constructed. In this paper, four ROIs were investigated to determine the most suitable patch that can be used in various distances requirement as shown in Figure 2(b). The first patch was on the right cheek of the face. Right cheek was chosen since less non-rigid motion is generated in this area compared to other regions. Next, the second patch was selected on the center of the face that includes eyes and nose region but excluding the forehead area since the forehead tends to be covered by hair. This area was chosen because previous study [13] reported that the center of face to be the most suitable ROI for video based HR. The third patch was the whole face since basically the larger the region, the possibility to extract the PPG information from a far distance is high.

However, for the whole face region a reduction of 10% horizontal dimension and 20% of vertical dimension with respect to center location is performed to exclude plausible region that constitute to the background. Finally, the fourth patch was selected at the lower part of the face that includes nose and mouth but excluding eyes and chin area. This region was chosen because of there are less non-rigid motion in it and wider ROI dimension compared to the patch placed on the right cheek. From the selected ROI, random pair temporal green color channel values that indicate the PPG signal is extracted.

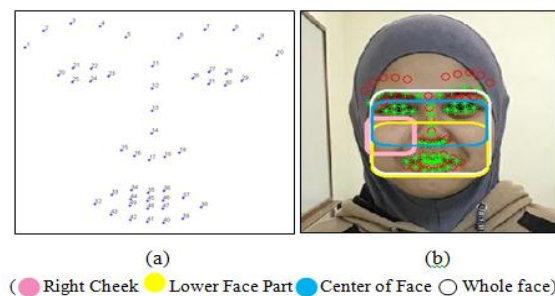


Figure 2. Sample of facial tracker result, (a) 49 facial landmarks, (b) ROI for extracting the PPG signal

2.2. PPG Signal Extraction and Noise Removal

In each of the constructed ROI the PPG signal is obtained randomly pickup two different locations inside the ROI, and consequently taps the reading of green color channel in the selected location over the predefined video sampling time. The Green channel was selected for PPG extraction because it provides the strongest PPG signal information compared to other color channel [4]. Sample of the extracted two traces of PPG is shown in Figure 3. It can be seen from the figure that the obtained signals contaminated with noise due to the effect from the environment and therefore need to undergo some filtering step prior to estimate the HR value.

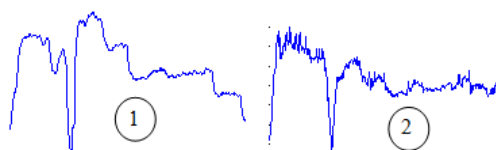


Figure 3. Sample of extracted two raw PPG signal from average of green trace in the face region

2.3. Noise Removal using ICA

The raw PPG traces calculated from the green color channel of the tracked patch incorporated the illumination variation information and hence produced a noise traces. In this section, a mechanism on how to isolate the illumination effect that hinders an accurate PG signal extraction is described. Fundamentally, two factors that affect the reading of the raw PPG signal which are the blood volume variation cause by cardiac pulse and the illumination changes. The combination of both parameters can be represented by a linear correlation define by Equation (2) in which s is the green channel signal and y is the variation of illumination. Ideally if the y parameter can be estimated, then the pure cardiac pulse signal can be obtained. However, in practice the signal of y cannot be measured directly.

$$PPG_{\text{raw}} = s + y \quad (2)$$

In this paper ICA is used separate the components of the green channel signal from the illumination variation effect. Basically ICA is one of the BSS techniques that capable to recover unobserved signal from a set of observed mixture in which the mixing process is unknown and assume to be linearly correlated. The raw PPG traces obtained previously is fed to the ICA by assuming the rectified PPG and the illumination components are independent. After solving for the components, the ICA will produce two separated signals as depicted in Figure 4. From the figure, it can be clearly seen that the fed two traces produce two components in which one of the ICA components signal pattern is quite similar with one of the raw traces. To explicitly determine the generated ICA signals belong to which group, i.e., PPG or illumination, nearest average Eulclidean distance measure is performed to find the similarity score between the two components and the traces, and the one with the lowest value will be labeled as the rectified PPG signal while the other one will be denoted as illumination variation .

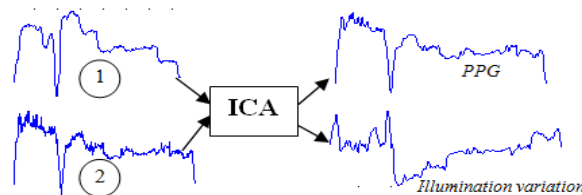


Figure 4. Sample of ICA implementation of BSS, original signal in the left side and two ICA components on the right side

2.4. PPG Filtering

Once the pure PPG signal is in hand, the final step is to determine corresponding heart rate value. However, the obtained signal still contains noise and some filtering step is necessary. In this paper, two temporal filter which are detrending [14] and moving average is implemented for reducing slow and non-stationary trend of signal and polishing the random noise to make the signal smooth prior to frequency domain conversion. The filtered PPG signal is converted to frequency domain for determine the power spectrum density (PSD) using Welch method [15] with the constrained frequency spectrum within the range of 0.7Hz to 4Hz that represent the HR value range from 42bpm to 240bpm. Eventually, the HR is calculated by multiplying the maximal PSD response with 60. Illustration of the generated signal as a result of the elaborated process can be seen in Figure 5 which consist of detrending result in (a), moving average outcome in (b) and the corresponding PPG frequency that is converted using Welch method.

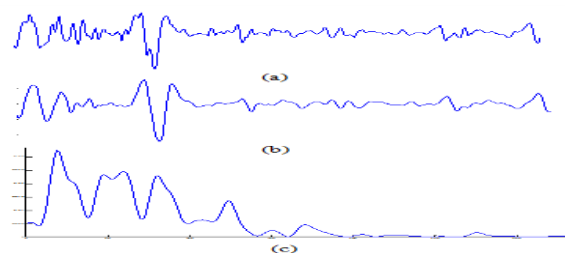


Figure 5. Filtering step, (a) Detrending, (b) Moving Average, (c) Pwelch with frequency range from 0.7Hz – 4Hz

2.5. Histogram Analysis

The HR value obtained from a single sequence calculation is still subject to the variation and hence a histogram based analysis is performed to determine the consistent reading against repetition. In this paper 10 repetition of HR reading from the same image sequence is performed. Any HR reading that is significant with the majority in the bin will be eliminated while the remaining will be used to determine an average of HR which shows consistency over the sampled time period.

3. RESULTS AND ANALYSIS

To analyze the performance of HR reading over various distances and different face regions we conduct an experiment in an indoor environment under controlled light condition. In the experiment, a recorded video of 1440*1080 pixels resolution with 60 Fps is taken from participant with different distance setting of 1 meter, 2 meters and 5 meters respectively. During recording, the participant exhibit normal static motion along the process. Sample snapshots of experiment setting are shown in Figure 6 which indicates pictures taken with facial landmarks detected.

The analysis is conducted to determine an optimal cross correlation setting between distances and face ROI that yield highest accuracy of the estimated HR value. As mentioned previously, there were four ROI selected for this experiment which are right cheek, center face, whole face and lower face. All the patch will be utilize in each of the video taken at four different distances as explained above and the result is showcase in Table 1. From the table, for 1 meter reading (Figure 6 (top left)) the most accurate HR was found to be generated from right cheek region with 98.77% accuracy. Clearly this happen because of such region generate less non-rigid movement and hence would reduce the motion artifact in the obtained PPG signal, and consequently would results in more accurate HR reading.

Table 1. Accuracy table for each patch corresponding to the distances

Patch (ROI)	Right Cheek Accuracy (%)	Center of Face Accuracy (%)	Lower Part of Face Accuracy (%)	Whole face Accuracy (%)
Range (m)				
(HR accuracy results for near distance)				
1	98.77	77.78	87.65	87.65
2	79.27	80.49	87.80	87.65
(HR accuracy results for farther distance)				
5	93.99	91.57	97.60	91.62
Average	92.27	89.22	93.61	91.62



Figure 6. Experiment setup to determine the HR value with detected facial landmarks from distance of 1m (top left), 2m (top right) and 5m (bottom)

Moving to a 2 meter case (Figure 6 (top right)), it is found that the most accurate HR was obtained from the lower part of the face area with 87.80% accuracy. Apparently, same as the right chick region, the

lower part yield less non-rigid movement and since the ROI dimension is bigger; it can generate an accurate HR reading even from a far distance. For a 5 meter case which consider as the farthest distance from the camera as indicate in Figure 6 bottom, the most accurate HR was also obtained from the lower part of the face area with 97.60% accuracy. Again, this was due to the less non-rigid movement in the area, so there was less motion artifact and bigger dimension as mentioned previously that make the obtained PPG to be more accurate.

Since the analyses were focusing on heart rate for various distances, the distance for both experiments ranges up until five meters. Quantitative assessment such as accuracy and percent of error are calculated using Equation below.

$$\text{Percentage of Error} = \left(\frac{|\text{measured value} - \text{actual value}|}{\text{actual value}} \right) * 100\% \quad (3)$$

$$\text{Accuracy (\%)} = 100\% - \text{Percent of Error} \quad (4)$$

From the obtained data, for a single person HR reading over a 5 meters range, it can be deduce that the ROI that produce the most inaccurate reading was the patch that placed in the center of the face that with an average of 88.62% of accuracy while the optimal ROI is found to be from the lower part of the face with an average of 93.65% of accurate HR estimation. One of the reason that contribute to the inaccuracy of the HR reading is the eye movement is included in the selected ROI. However, generally the entire selected path is capable to produce around 80% of accurate HR estimation in each of the distance variation.

4. CONCLUSION AND FUTURE WORKS

This paper investigates effect of ROI variation over a different distance for non contact HR estimation framework. There were total of five steps in order to accomplish the framework which is video acquisition, facial detection, PPG signal extraction using, PPG signal filtering and histogram. From the experiments, we can conclude that the system capable to work properly with various distances variation ranging from 1 to 5 meters with an average accuracy of 92%. Among the evaluated ROI, we found that the lower part of face region capable to produce the highest average of accuracy over the distance variation with 93.61%.

For future, the system is expected to be able to detect the HR reading from various ranging of skin tone and simultaneously extract the HR reading from a for multiple persons. We would also interested to study the effect of using different types of signal analysis method to obtain more accurate the HR reading. Apart from that, the system is also can be improved by reducing the processing time to make it applicable for real time data HR measurement scenario.

ACKNOWLEDGEMENTS

The authors would like to thank to Ministry of Education (MOE) and Universiti Tun Hussein onn Malaysia (UTHM) for supporting this research under Fundamental Research Grant Scheme (Vot. no. 1582).

REFERENCES

- [1]. G. Cennini, J. Arguel, K. Aksit and A. Van Leest, "Heart Rate Monitoring via Remote Photoplethysmography with Motion Artifact Reduction", *Optic Express*, 18(5) : 4867-4875, 2010.
- [2]. C. Li, C. Xu, C. Gui, M. D. Fox, "Distance regularized level set evolution and its application to image segmentation", *IEEE Trans. on Image Processing*, 2010.
- [3]. M.-Z. Poh, D.J. McDuff, and R.W. Picard, "Advancements in noncontacts, multiparameter hysiological measurements using a webcam", *IEEE Trans. on Biomedical Engineering*, 2011.
- [4]. W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light", *Optics Express*, 16(26) : 21434-21445, Dec 2008.
- [5]. S. Kwon, H. Kim, and K. S. Park, "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone", In IEEE Engineering in Medicine and Biology Society (EMBC), pages 2174-2177, Aug 2012.
- [6]. G. Balakrishnan, F. Durand, and J. Guttag, "Detecting pulse from head motions in video", In IEEE Computer Vision and Pattern Recognition (CVPR), pages 3430-3437, June 2013.
- [7]. X. Li, J. Chen, G. Zhao, and M. Pietikainen, "Remote Heart Rate Measurement from Face Video under Realistic Situations", In IEEE Computer Vision and Pattern Recognition (CVPR), pp. : 4264-4271, 2014.
- [8]. M. Kumar, A. Veeraraghavan and A. Sabharwal, "Distance cePPG: Robust non-contact Vital Sign Monitoring using a Camera", *Biomedical Optic Express*, 6(5):1565-1288, 2015.

- [9]. A. Lam and Y. Kuno, "Robust Heart Rate Measurement from Video using Select Random Patches", In IEEE International Conference on Computer Vision (ICCV), 2015.
- [10]. R.-Y. Huang and L.-R. Dung, "Measurement of Heart Rate Variability using off the shelf smart phones", *Biomedical Engineering online*, 15:11, pp. 1-16, 2016.
- [11]. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", In Proc. of Int. Conf. on Comp. Vision and Pattern Recognition, 2001, pp. 511-518.
- [12]. P. Dollar, P. Welinder, P. Perona, "Cascaded Pose Regression", In Proceeding on Computer Vision and Pattern Recognition, 2010.
- [13]. T. Pursche, J. Krajewski, and R. Moeller, "Video-based heart rate measurement from human faces", In Proceedings of IEEE International Conference on Consumer Electronics (IEEE, 2012), 2012, pp. 544-545
- [14]. P. Tarvainen, P. O. Ranta-aho, and P. A. Karjalainen, "An advanced detrending method with application to hrv analysis", *IEEE Trans. on Biomed. Eng.*, 2002
- [15]. P. Welch, "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging overshoot, modified periodograms", *IEEE Trans. on Audio and Electroacoustics*, 1967

BIOGRAPHIES OF AUTHORS



Norwahidah Ibrahim was born in Johor, Malaysia in 1993. She received the B.Eng degree in Electronics (Mechatronics) from University Tun Hussein Onn Malaysia in 2016. Currently, she is doing her M.Eng degree in Electrical at the same university. Her current research interest is computer vision, image processing and signal processing.



Razali Tomari was born in Johor, Malaysia, in 1980. He received the B.Eng degree in Mechatronics from the Universiti Teknologi Malaysia, in 2003, M.Sc in intelligent system from the Universiti Putra Malaysia, in 2006. And PhD degree in computer vision and robotic from the Saitama University, Japan in 2013. In 2003, he joined Faculty of Electrical Engineering, University Tun Hussein Onn as a tutor and later on become a Senior Lecturer in 2013. His current research interest includes computer vision, pattern recognition, smart wheelchair and sensing technology.



WN Wan Zakaria received B.Eng (2007) in Electronics and Mechanical Engineering from Chiba University and MSc(2008) and PhD (2012) from Newcastle University. She is currently a lecturer in Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia. Her current interests include Medical Robotics System specifically on development of robot force control, Image Processing and Computer Aided Diagnosis, and development of Wearable Device. She is author and co-author of several journal papers and conference proceedings.



Nurmiza Binti Othman was born in Terengganu, Malaysia, in 1983. She received the M. Sc. (Electrical and Electronic Engineering) from Utsunomiya University, Japan in 2009 and the Ph.D (Electrical and Electronic Engineering) from Kyushu University, Japan in 2014. In 2007, she joined the Department of Electronic Engineering, Universiti Tun Hussein Onn Malaysia as a tutor. From 2014 to present, she has appointed as a lecturer with the same institution. Currently, she is also a Senior Researcher of UTHM Research Center For Applied Electromagnetics, Universiti Tun Hussein Onn Malaysia. She is a member of professional membership IEEE under Engineering in Medicine and Biology Society. Her research interests include Superconductor Engineering, Microfabrication, Nanomagnetic Materials, Magnetic Particle Imaging, and Magnetic Tomography.