

# Dynamic heart rate estimation using principal component analysis

Yong-Poh Yu,<sup>1,3</sup> P. Raveendran,<sup>1,\*</sup> Chern-Loon Lim,<sup>1,4</sup> and Ban-Hoe Kwan<sup>2,5</sup>

<sup>1</sup>Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur 50603, Malaysia

<sup>2</sup>Department of Mechatronics and Biomedical Engineering, Lee Kong Chian Faculty of Engineering and Science, University Tunku Abdul Rahman, Malaysia

<sup>3</sup>richieyp@yahoo.com

<sup>4</sup>limchernloon@gmail.com

<sup>5</sup>banhoe@gmail.com

\*ravee@um.edu.my

**Abstract:** In this paper, facial images from various video sequences are used to obtain a heart rate reading. In this study, a video camera is used to capture the facial images of eight subjects whose heart rates vary dynamically, between 81 and 153 BPM. Principal component analysis (PCA) is used to recover the blood volume pulses (BVP) which can be used for the heart rate estimation. An important consideration for accuracy of the dynamic heart rate estimation is to determine the shortest video duration that realizes it. This video duration is chosen when the six principal components (PC) are least correlated amongst them. When this is achieved, the first PC is used to obtain the heart rate. The results obtained from the proposed method are compared to the readings obtained from the Polar heart rate monitor. Experimental results show the proposed method is able to estimate the dynamic heart rate readings using less computational requirements when compared to the existing method. The mean absolute error and the standard deviation of the absolute errors between experimental readings and actual readings are 2.18 BPM and 1.71 BPM respectively.

©2015 Optical Society of America

**OCIS codes:** (100.0100) Image processing; (100.2960) Image analysis; (170.3880) Medical and biological imaging.

## References and links

1. A. A. Kamshilin, S. Miridonov, V. Teplov, R. Saarenheimo, and E. Nippolainen, "Photoplethysmographic imaging of high spatial resolution," *Biomed. Opt. Express* **2**(4), 996–1006 (2011).
2. M. Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Express* **18**(10), 10762–10774 (2010).
3. A. B. Hertzman and C. R. Spealman, "Observations on the finger volume pulse recorded photoelectrically," *Am. J. Physiol.* **119**(334), 3 (1937).
4. M. Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Trans. Biomed. Eng.* **58**(1), 7–11 (2011).
5. M. Kumar, A. Veeraraghavan, and A. Sabharwal, "DistancePPG: Robust non-contact vital signs monitoring using a camera," *Biomed. Opt. Express* **6**(5), 1565–1588 (2015).
6. S. Xu, L. Sun, and G. K. Rohde, "Robust efficient estimation of heart rate pulse from video," *Biomed. Opt. Express* **5**(4), 1124–1135 (2014).
7. Y. P. Yu, P. Raveendran, and C. L. Lim, "Dynamic heart rate measurements from video sequences," *Biomed. Opt. Express* **6**(7), 2466–2480 (2015).
8. X. Liu, D. Wang, F. Liu, and J. Bai, "Principal component analysis of dynamic fluorescence diffuse optical tomography images," *Opt. Express* **18**(6), 6300–6314 (2010).
9. J. Vargas, J. A. Quiroga, and T. Belenguer, "Phase-shifting interferometry based on principal component analysis," *Opt. Lett.* **36**(8), 1326–1328 (2011).
10. A. Krishnaswamy and G. V. Baranoski, (2004). A study on skin optics. Natural Phenomena Simulation Group, School of Computer Science, University of Waterloo, Canada, Technical Report, 1, 1–17.
11. C. A. Poynton, (1996). A technical introduction to digital video. John Wiley & Sons, Inc.
12. P. Viola and M. Jones, (2001). Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on (Vol. 1, pp. I-511)*. IEEE.

13. M. P. Tarvainen, P. O. Ranta-Aho, and P. A. Karjalainen, "An advanced detrending method with application to HRV analysis," *IEEE Trans. Biomed. Eng.* **49**(2), 172–175 (2002).
14. A. Pietila and T. Tammi, (1997). *U.S. Patent No. 5,622,180*. Washington, DC: U.S. Patent and Trademark Office.
15. I. Heikkila, (1998). *U.S. Patent No. 5,840,039*. Washington, DC: U.S. Patent and Trademark Office.

## 1. Introduction

Video based heart rate estimation is based on the photoplethysmography (PPG) technique [1,2]. PPG technique is a non-invasive optical technique used for measuring the blood volume pulses (BVP) through the intensity variations in the reflected light [3].

Poh *et al.* [4] used Independent component analysis (ICA) to recover the BVP signals from the facial images of the subjects. They showed that the color components of the facial images, namely red, green and blue (RGB), are dependent to each other. The method was tested on several 60-second videos and accurate results were obtained.

Kumar *et al.* [5] proposed a model, known as DistancePPG, to improve the signal-to-noise ratio of the camera-based PPG signal by combining the color change signals obtained from different regions of the face using a weighted average. Additionally, they introduced a method to track different regions of the face separately to extract the PPG signals under motion. The method was evaluated on people having diverse skin tones, under various lighting conditions and natural motion scenarios. Kumar *et al.* concluded that the accuracy of heart rate estimation was significantly improved using the proposed method.

On the other hand, Xu *et al.* [6] designed a simplified Mathematic model to predict the human heart rates based on the light absorbance by the skin. In this work, the change of blood concentration due to arterial pulsation was defined as a pixel quotient in log space. The method was tested on subjects who were stationary. Xu *et al.* concluded that the method gave accurate results.

Previous works did not focus on the dynamic heart rate variation. For dynamic heart rate variation, short video sequence is essential for real time implementation. Yu *et al.* [7] proposed a method that uses a combination of ICA and mutual information to compute the dynamic heart rate variation from short video sequence. For short video sequences, the challenge in using ICA method is that the ICA sources may not have sufficient independence amongst them. Hence, mutual information was used to establish the independence of the sources to obtain an accurate reading. However, this method is computationally intensive.

In this paper, we propose to use principal component analysis (PCA), which is less computationally intensive than ICA, to estimate the instantaneous heart rate that varies dynamically from short video sequences. PCA is a data reduction technique that is commonly used in image and video analysis [8,9]. Since the pixel intensities in log space for the facial images are correlated to each other, PCA is used to recover the de-correlated principal components. However, for short video sequence, the principal components (PC) may still have high correlation amongst them and that may render inaccurate reading. To determine the video duration that gives uncorrelated PCs, the Pearson correlation coefficient is used.

Section 2 discusses the relationship between RGB and  $YC_bC_r$  components and hemoglobin concentration where it can be used to determine the heart rate from the video sequences. The proposed method that uses PCA and Pearson correlation coefficient is presented in Section 3. The results obtained from the method are presented and discussed in Section 4. Section 5 concludes the study.

## 2. Relationship between RGB and $YC_bC_r$ components of facial images

Human skin is composed of different layers [10] and its color is highly related to melanin and hemoglobin concentrations. Xu *et al.* [6] derived the relationship between the RGB pixel intensities obtained from a facial image and the hemoglobin and melanin concentrations,  $c_h$  and  $c_m$ , in the skin layer as:

$$\log P_R = -\{v_m(R)c_m + v_h(R)c_h + A_0(R)\} + \log kE(R), \quad (1)$$

$$\log P_G = -\{v_m(G)c_m + v_h(G)c_h + A_0(G)\} + \log kE(G), \quad (2)$$

$$\log P_B = -\{v_m(B)c_m + v_h(B)c_h + A_0(B)\} + \log kE(B). \quad (3)$$

where  $R, G$  and  $B$  represent the red, green and blue components of the image respectively while  $v_h$  and  $v_m$  are the product of pigment extinction coefficient of hemoglobin and melanin respectively and their mean path length of photons in the skin layer.  $A_0$  denotes the baseline skin absorbance while  $k$  is the constant number for the camera gain and  $E$  is the power of incident light for each color component.

Considering the video is captured under constant background light and for a short duration, then the AC components of the RGB pixel intensities in log space consist mostly of hemoglobin concentration. Since hemoglobin concentration is related to blood concentration, the frequency of this hemoglobin concentration is considered as the BVP, i.e. the heart rate pulse.

As Eqs. (1)-(3) depend only on pigment concentration, baseline skin absorbance and the incident light, we may conclude that the RGB in log space are correlated to each other. Figure 1 shows the distribution of RGB pixel intensities in log space over a period of time used in our experiment. Table 1 shows the correlation among  $\log P_R$ ,  $\log P_G$ , and  $\log P_B$ . The values in Table 1 show the RGB pixel intensities in log space are highly correlated to each other. Therefore, PCA can be used to decorrelate these color components and recover the corresponding uncorrelated PCs.

However, for short video duration, the correlation amongst the PCs is still high which upon using them in the subsequent operation may lead to inaccurate heart rate readings. To address this issue, we add three more color components, i.e. luminance  $Y$ , chrominance  $C_b$  and  $C_r$ , in log space as the input features.  $YC_bC_r$  components are correlated to the RGB components. These components can be derived from the RGB [11] as follows:

$$Y = 16 + 65.481 \cdot R + 128.553 \cdot G + 24.966 \cdot B \quad (4)$$

$$C_b = 128 - 37.797 \cdot R - 74.203 \cdot G + 112 \cdot B \quad (5)$$

$$C_r = 128 + 112 \cdot R - 93.786 \cdot G - 18.214 \cdot B \quad (6)$$

With these six color components as input features, the corresponding PCs have much lower correlation as compared to the PCs whose input features are RGB components only. Figure 2 illustrates the correlation coefficients for six PCs and three PCs respectively. It shows that the correlation coefficient for the six PCs decreases as the video duration increases while the correlation coefficient for the three PCs is not consistent and relatively higher when compared to the six PCs. Therefore, in this paper, we use PCA to recover the PPG signals from these six color components. After applying the PCA, the first PC that has the largest possible variance is considered as the PPG signal that consists of the hemoglobin concentration. The heart rate can be computed from this PPG signal.

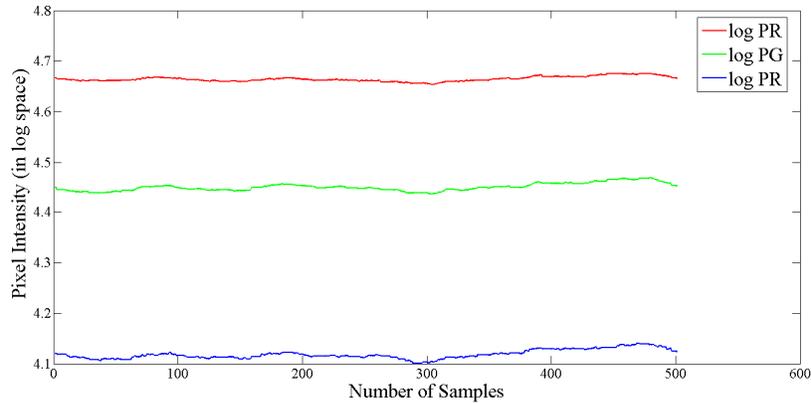


Fig. 1. The distribution of  $\log P_R$ ,  $\log P_G$  and  $\log P_B$ .

**Table 1. Correlation coefficient among  $\log P_R$ ,  $\log P_G$ , and  $\log P_B$**

	$\log P_R$	$\log P_G$	$\log P_B$
$\log P_R$	1.00	0.92	0.84
$\log P_G$	0.92	1.00	0.93
$\log P_B$	0.84	0.93	1.00

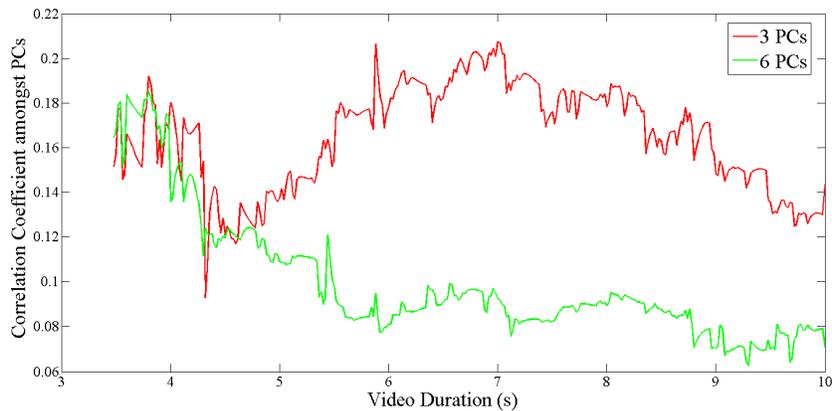


Fig. 2. The graph of correlation coefficient amongst PCs vs video duration for 3 PCs and 6 PCs respectively.

### 3. Proposed method

In this section, the proposed model to estimate the dynamic heart rate measurements using PCA is presented. The relationship between the correlation among PCs, video duration and heart rate accuracy is also discussed. As the video duration affects the accuracy of the heart rate reading, a stopping criterion is set to determine the video duration needed for dynamic heart rate estimation. The details are described in this section.

#### 3.1 Relationship between the correlation among PCs and video duration

Ideally, PCA will compute its PCs by maximizing the correlation among the input features. However, for short video duration, the PCs may still have high correlation. Hence, it is important to find out the minimum video duration that gives the least correlation. We use

Pearson correlation coefficient to determine the correlation between any two PCs. For any given two PCs  $\mathbf{x}$  and  $\mathbf{y}$ , the Pearson correlation coefficient of these two PCs,  $R$  is given as:

$$R(x, y) = \frac{C(x, y)}{\sqrt{C(x, x)C(y, y)}}. \quad (7)$$

where  $C(\mathbf{x}, \mathbf{y})$  is the covariance of PCs  $\mathbf{x}$  and  $\mathbf{y}$ ,  $C(\mathbf{x}, \mathbf{x})$  and  $C(\mathbf{y}, \mathbf{y})$  are the variances of PCs  $\mathbf{x}$  and  $\mathbf{y}$  respectively. Since six PCs are recovered from the PCA, the averaged correlation coefficient amongst PCs,  $R_{avg}$  is computed using

$$R_{avg} = \frac{1}{\binom{6}{2}} \sum_{m=2}^6 \sum_{n=1}^m R(m, n). \quad (8)$$

Figure 3 illustrates the relationship between the averaged correlation coefficient  $R_{avg}$  and video duration for a particular heart rate reading used in the experiment. A power function curve is fitted to represent the function  $R_{avg}$ . It is found that the value of the  $R_{avg}$  decreases significantly at the beginning, but remains almost constant when the video duration exceeds a specific duration. The value of the  $R_{avg}$  varies very little after this duration. Hence, the stopping criterion to determine the video duration is set as when the difference of  $R_{avg}$  for 3 continuous video frames is smaller than  $2 \times 10^{-4}$ .

To illustrate how correlated PCs affect the accuracy of computed heart rate readings, two points X and Y are selected in Fig. 3. The actual heart rate reading for this particular instant is 143 BPM. Point X represents a very short video duration where the  $R_{avg}$  doesn't meet the stopping criterion. Point Y represents the suitable video duration where the  $R_{avg}$  has met the stopping criterion. Point Y gives more accurate heart rate estimation, i.e 142.38 BPM as compared to point X that gives 63.75 BPM. When the stopping criterion is met, the corresponding video duration is used to compute the instantaneous heart rate for that particular instant.

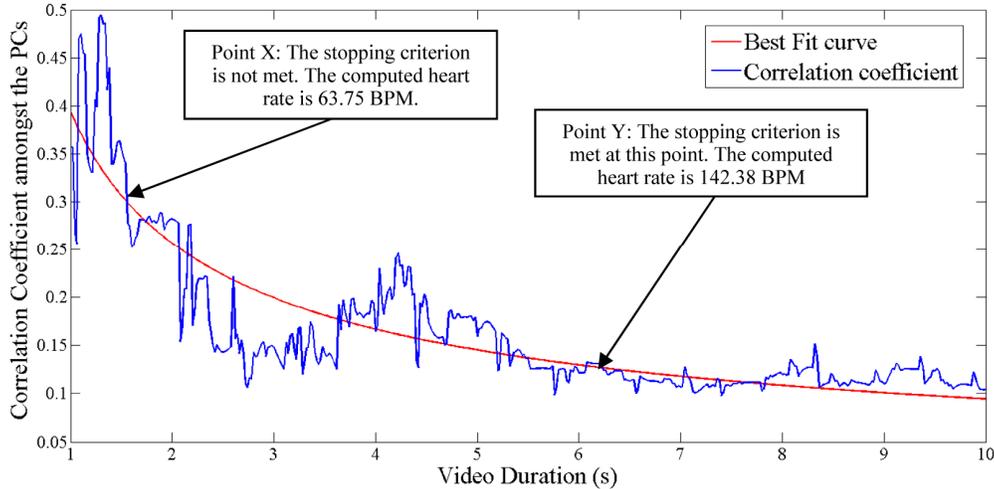


Fig. 3. The relationship between the averaged correlation coefficient  $R_{avg}$  and the video duration and the respective computed heart rates.

### 3.2 Block diagram of proposed model

The block diagram of the proposed model is illustrated in Fig. 4. The face region is identified by using the model described in [12] and the region of interest (ROI) is fixed at the area below eyes and above the upper lip of mouth. For each frame, the spatially average of the

RGB and  $YCbCr$  components, i.e.:  $\mu_R$ ,  $\mu_G$ ,  $\mu_B$ ,  $\mu_Y$ ,  $\mu_{Cb}$ , and  $\mu_{Cr}$  are computed respectively. All six color components are projected into log space. Therefore, at any time instant, a set of six input features  $\log P_R$ ,  $\log P_G$ ,  $\log P_B$ ,  $\log P_Y$ ,  $\log P_{Cb}$  and  $\log P_{Cr}$  are formed. The set of input features are then detrended using the model developed by [13]. PCA is then used to recover six PCs from these six input features. The set of PCs is bandpass filtered (128-point Hamming window, 0.8-4 Hz).

The entire process is repeated by increasing the number of previous video frames, until the stopping criterion described in *Section 3.1* is met. At this point, the corresponding number of frames is chosen as the video duration needed to compute the instantaneous heart rate reading. The first PC is then chosen as the PPG signal. The corresponding frequency of this PPG signal is considered as the instantaneous heart rate reading for that particular instant.

#### 4. Experimental study

In this section, the experimental setup and the experimental results are discussed and analysed. A comparative study between the proposed method and the method used in [7] is also presented.

##### 4.1 Experimental setup

All experiments were set up under constant office fluorescent light. A Sony camcorder (HDR-PJ260VE) was used for the video recording purposes. All videos were recorded and sampled at 50 frames per second. The camcorder was fixed at a position with a distance of about 0.60 m from the subject's face. In the experiment, eight subjects were selected and requested to carry out a cycling activity. In the first stage of the experiment, four subjects were asked to cycle at different speeds for about two minutes. Then they were asked to stop for one minute. The camcorder was used to capture their facial images during that time. In the second stage of the experiment, the remaining four subjects were asked to cycle continuously and their facial images were captured by the camcorder for one minute. An increasing heart rate trend was observed. Throughout the video recordings, all subjects were asked to remain stationary. Sixty heart rate readings (sampled at each second) were computed for every subject.

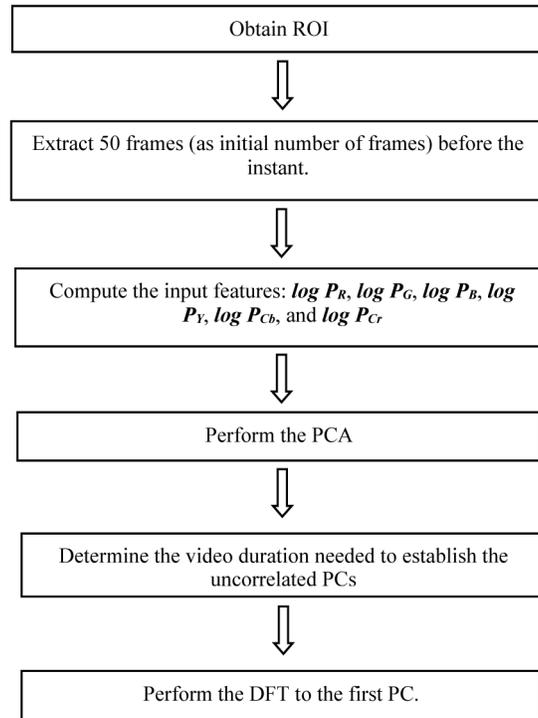


Fig. 4. Flow chart of the proposed method.

As reference, the instantaneous heart rates of each subject that obtained from the proposed method were compared to the actual heart rate readings measured from Polar Heart Rate Monitor – Polar Team<sup>2</sup> Pro. Polar Team<sup>2</sup> Pro samples and computes the instantaneous heart rate by measuring at least one ECG signal waveform, as described in the patents [14,15].

#### 4.2 Experimental results and analysis

A total of 480 instantaneous heart rate readings were obtained from this experiment. In the experiment, the subjects' heart rates were varying between 81 BPM and 153 BPM. Table 2 summarizes the details of the computed heart rate readings of all subjects. The highest and the lowest mean absolute errors are 2.99 and 1.37 BPM. Figure 5 shows the scattered plot of all computed and actual heart rate readings. It shows that the computed heart rate readings are closely correlated to the actual heart rate readings. The correlation coefficient between the computed and actual heart rate readings is 0.99. The mean absolute error for all readings is 2.18 BPM while the standard deviation of absolute errors is 1.71 BPM. The Bland Altman plot is shown in Fig. 6. It shows that only a small number of computed heart rate readings are located outside the 95% limit of agreement interval.

**Table 2. Summary of Heart Rate Readings Results Obtained from Proposed Method**

Subject	Heart Rate Readings (BPM)		Mean absolute error (BPM)	Standard deviation of absolute errors (BPM)
	Highest	Lowest		
1	141	127	1.37	1.02
2	134	122	1.91	2.09
3	105	96	1.88	1.21
4	153	119	1.42	0.99
5	108	81	2.32	1.54
6	120	104	2.99	1.86
7	153	127	2.65	2.00
8	133	105	2.85	1.77

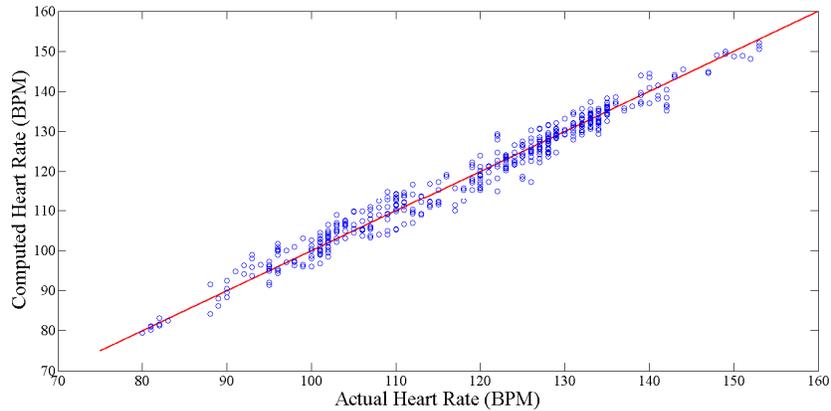


Fig. 5. Comparison of actual heart rate readings and computer heart rate readings.

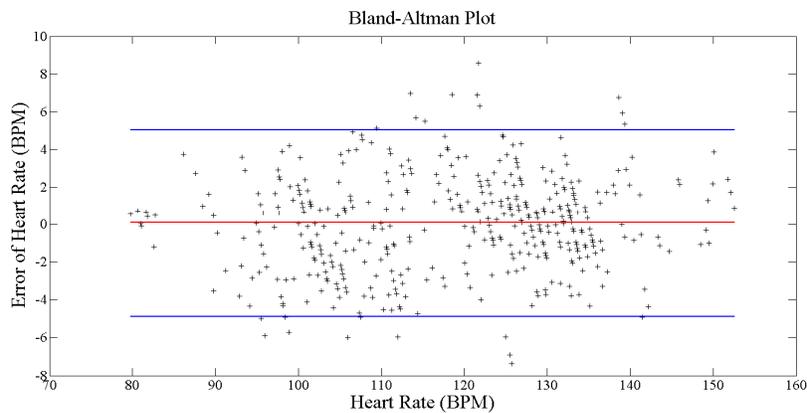


Fig. 6. Bland-Altman Plot for all computed heart rate readings.

#### 4.3 Comparative study between proposed method and existing method

A comparative study has been done to compare the accuracy (mean error and standard deviation of error), video duration for the heart rate computation, and the computational cost of using the method described in this paper and the method suggested in [7]. To calculate the computational cost, both ICA [7] and PCA computations are repeated for 1000 times and the average time taken is recorded. Table 3 summarizes the results of the comparative study. As can be seen in Table 3, both accuracy and video duration are not much different for these two methods. However, in terms of the computational cost, the proposed method is much more efficient than the method used in [7]. Additionally, the proposed method directly uses the first PCs to compute the heart rate while [7] investigated all ICA sources first and then chose the source with high peak in frequency domain as the source giving heart rate information. As low computational cost and small memory resources are important factors for the eventual implementation in mobile phones, the proposed method is more efficient than the previous method.

**Table 3. Comparison of proposed method (using PCA) and method suggested in [7] (using ICA)**

	PCA	ICA
Mean absolute error	2.18	1.64
Standard deviation of absolute errors	1.71	1.48
Video duration (mean), in second	5.33	5.49
Computational cost (mean), in second	2.25	6.62

## 5. Conclusion

In this study, it is found that heart rate readings can be obtained by applying PCA to the facial images. When the PCs are uncorrelated to each other, then an accurate reading can be obtained. An important consideration for dynamic heart rate estimation is the need for video duration. Instead of using RGB components only, three additional components,  $Y C_b C_r$  are used. In doing so, a shorter video duration is obtained. To ensure the reliability of the heart rate estimation, the PCs must have least correlation. To validate the criterion, Pearson correlation coefficient is used. Experimental results show that this method is able to estimate dynamic heart rates from short video sequences using less computational requirements when compared to [7].

## Acknowledgment

This research is supported by High Impact Research Chancellory Grant UM.C/HIR/MOHE/ENG/42 from the University of Malaya.