



Review

Heart rate estimation using facial video: A review

M.A. Hassan^{a,b}, A.S. Malik^{a,*}, D. Fofi^b, N. Saad^a, B. Karasfi^a, Y.S. Ali^c, F. Meriaudeau^a^a Centre for Intelligent Signal and Imaging Research (CISIR), Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia^b Le2i UMR 6306, CNRS, Arts et Métiers, Univ. Bourgogne Franche-Comté 12, Rue de la fonderie, 71200 Le Creusot, France^c Science and Technology Unit, Umm Al-Qura University, 21955 Makkah, Saudi Arabia

ARTICLE INFO

Article history:

Received 1 August 2016

Received in revised form 10 April 2017

Accepted 11 July 2017

Keywords:

Heart rate measurement

Photoplethysmography

Ballistocardiography

Non-contact

Facial imaging

ABSTRACT

Photoplethysmography and Ballistocardiography are two concepts that are used to measure heart rate from human, by using facial videos. Heart rate estimation is essential to determine the physiological and pathological state of a person. This paper presents a critical review of digital camera based heart rate estimating method on facial skin. This review extends the investigation on to the principles and theory behind photoplethysmography and ballistocardiography. The article contains reviews on the significance of the methods and contributions to overcome challenges such as; poor signal strength, illumination variance, and motion variance. The experiments were conducted to validate the state of the art methods on a challenging database that is available publicly. The implemented methods were validated using the database, on 27 subjects for a range of skin tones from pearl white, fair, olive to black. The results were computed using statistical methods such as: mean error, standard deviation, the root mean square error, Pearson correlation coefficient, and Bland-Altman analysis. The results derived from the experiments showed the reliability of the state of the art methods and provided direction to improve for situations involving illumination variance and motion variance.

© 2017 Elsevier Ltd. All rights reserved.

Contents

1. Introduction	346
2. Ballistocardiography	347
2.1. Theory	347
2.2. Related work	347
3. Photoplethysmography	348
3.1. Theory	348
3.2. Related work	349
3.2.1. Signal strength	350
3.2.2. Illumination variance	351
3.2.3. Motion artifacts	352
4. Experiment setup	353
5. Results and discussion	355
6. Conclusion	358
Acknowledgements	359
References	359

1. Introduction

Heart rate estimation is an essential component to determine the physiological and pathological state of a person/individual. The heart is one of the most important organs in the human body. The function of the heart is nothing more than to pump blood through-

* Corresponding author at: Centre for Intelligent Signal and Imaging Research (CISIR), Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia.

out the body by beating approximately at 60 to 100 beats per minute (bpm) [1]. Heart rate above 100 bpm which is known as Tachycardia and heart rate below 60 bpm is known as Bradycardia are considered as abnormalities [2]. Abnormalities in the heart rate are caused by physiological and pathological factors that disrupt the normal electrical impulse which controls the heart's pumping action. Tachycardia is usually a cause of high blood pressure, smoking, fever, sudden stress, side effects of medications and damage to heart tissues from heart disease. Bradycardia is usually a cause of hypertension, rheumatic fever, obstructive sleep apnea, myocarditis and heart tissue damage related to aging. The risk factor to either of these abnormalities would result fatally.

Estimation and monitoring of heart rate have been essential in the fields of medicine and biomedical engineering. Researchers have developed various methods to estimate the heart rate from contact to non-contact sensors. Reviews have been written on contact and noncontact heart rate measuring methods (for recent reviews, see [3,4]). The standard and established heart rate measuring methods are in contact with the body such as: electrocardiogram and photoplethysmography which are based on conductive electrodes and phototransistors. Non-contact, experimental methods are microwave/ultrasound distance measurement, optical vibrio cardiography, thermal imaging, and RGB imaging (see Table 1). The common object of either of these methods is the extraction of physiological parameters to estimate and monitor the heart rate in a reliable manner.

The progress on non-contact heart rate measurement have attained much focus since remote sensing has emerged. Remote sensing has minimized the amount of cabling and cluttering associated with contact based heart rate measurement methods. Heart rate estimation using digital camera sensor is a rapidly growing research area due to its low cost and non-invasive nature for measuring the heart rate. The heart rate estimation using digital camera sensor is mainly operated based on the concepts of photoplethysmography and ballistocardiography. The digital camera based heart rate extraction methods have evolved from using high definition charge coupled device (CCD) and complementary metal-oxide-semiconductors (CMOS) camera sensor to cellphone and webcam camera sensors [4]. Therefore, heart rate measurement from facial videos using digital camera sensors has become one of the motive directions to extract physiological signals in a non-invasive approach.

The progress of heart rate measurement using facial videos has given a new course towards health monitoring in fields of health care, telemedicine, rehabilitation, sports, and ergonomics [5]. This paper will discuss the development and present the state of the art on heart rate measurement using facial videos. The remaining of the paper will be discussed based on the two main concepts: photoplethysmography and ballistocardiography, that are used to measure heart rate from facial videos. Sections 2 and 3 will discuss on methods related to ballistocardiography and photoplethysmography, the principles and theory behind concepts; the recent developments, significance and challenges on heart rate estimation. Section 4 will discuss the experiment design to test state-of-the-art methods on a common platform to determine the reliability of the methods. Section 5 will cross-validate the derived results, qualitatively and quantitatively and Section 6 will report the conclusions of the study and future direction in the research area.

2. Ballistocardiography

2.1. Theory

Ballistocardiography is a method for measuring the heart rate by estimating the motion generated by pumping of blood from the heart at each cardiac cycle. It is one of many methods that rely on



Fig. 1. Illustration on involuntary head movement due to the blood circulation from heart to head.

the mechanical motion of the cardiovascular systems, such as apex cardiography, kinetocardiograph, phonocardiograph, and seismocardiography. Ballistocardiography (BCG) was initially discovered in the 19th century and became the focus of interest in research from the 1940s to 1980s, after which the method faded away. BCG is defined by Newton's Third law of force (acceleration, velocity, or displacement), whereas the whole body is subjected to a rigid motion due to the cardiac ejection of blood. Therefore, the physiological signals obtained from BCG are more related to describing the blood flow within the heart, inside the arteries (mainly the aorta), and the movement of the heart itself. The most typical BCG extraction methods involve the subject to lay on a plane bed or sitting on a chair with minimal voluntary and involuntary movement [6].

The main motivation to extract BCG signals from the facial video was due to the involuntary head movement caused by blood flow from the heart to the head. At each cardiac cycle, when the heart beats, the left ventricle contracts and ejects blood at a high pressure to the aortic arch. This flow of blood at each cycle pass through the carotid arteries on either side of the neck, generating a force on the head (see Fig. 1). Therefore by Newton's 3rd law of force, this force created by the blood flow on the head, equals the force of the head acting on the blood flow causing reactionary cyclical head movement. White arrows illustrate this in Fig. 1. This head movement is too small to be noticed by the naked eye. However, the work presented in [7], by using video amplification, proved that the head moves periodically to the motion of the heart rate at a smaller amplitude.

2.2. Related work

In 2013, Balakrishnan et al. [7], proposed a heart rate estimation method using facial video based on BCG. This method focused on extracting the involuntary head motion in the video. The heart rate from the facial video was extracted by tracking velocities of feature points on the face region. The velocity of the feature points was extracted to track the microscopic movement of the head. The feature points of the head were extracted from the region of interest (ROI) specified by the authors. The authors used a combination of regions above the eye line (i.e. forehead) and the region below the eye line which includes the cheeks and the upper portion of the lips. The justification for the selection of ROI is that the area selected contained much of the capillaries which branch out from the carotid arteries. The eye line was excluded because blinking of the eye would result in motion artifacts in the extracted BCG signal.

The authors used the Viola-Jones face detector [8], to detect the face and extract the ROI. The feature point of the ROI was tracked by using Kanade Lucas Tomasi (KLT) face tracking algorithm [9]. The tracked points were temporally filtered using a bandpass filter

Table 1
Classification of contact and non-contact heart rate measurement methods based on technology specification.

														Acoustic	Noise Immunity
														Electrical	
														Radio frequency	
														Computational Efficiency	Robustness
														Reliability	
														Multi heartrate estimation	
														Commercial availability	Technology maturity
														Development phase	
														Low	Cost
														Fair	
														High	
														Health care	Application
														Sports	
														Ergonomics	
														Mobile	
✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Electrocardiogram	Contact
✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Photoplethysmography	
														Headphones	
														Microphone	
✓														CCECG	Non-Contact
✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Optical Vibrocardiography	
✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	MDM	
✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	UDM	
✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Thermal Camera	
✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	RGB Camera	
✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		

which operated at 0.75 Hz to 2 Hz. The operational frequency was chosen in order to resemble the heart rates within 45 to 120 bpm. Principal component analysis (PCA) was used to eliminate artifacts and extract the BCG like periodic signal. Consequently, the heart rate of the subject was determined by a simple peak detection algorithm.

Shan et al. [10] proposed a BCG extraction method based on [7]. Here the authors used a single ROI, instead of a combination of ROI. The forehead region was only used to extract the feature points since the area below the eye line (i.e. cheeks and upper lip) was much more prone to motion artifacts, caused by non-rigid movements of the facial muscle due to talking or expressing emotions. Also, independent component analyses (ICA) was used instead of PCA to extract the BCG signal. Finally, heart rate is measured by extracting the power spectral density (PSD) of the BCG signal. Although there was not much difference between the results obtained using the methods ICA or PCA, the main significance of BCG based approach was that the methodology itself was invariant to illumination variance and variations of skin complexion.

Haque et al. [11] proposed a BCG based heart rate estimation approach by using fusing corner feature points of good features to track [9] and 49 predefined facial landmarks by using deformable face fitting algorithm [12]. The authors used the feature fusion based approach to overcome the limitations of [7,10] regarding loss of features due to movement or occlusion of the face. The feature points were tracked using the KLT tracking algorithm and the vertical components of the trajectories were extracted as raw BCG traces. The authors further refined the raw traces by performing a process of bandpass filtering using an eight order Butterworth band-pass filter at operational frequencies of 0.75 Hz to 5 Hz.

The filtered BCG signals were further refined by applying a moving average filter to eliminate motion artifacts. Here the authors used a moving average filter with an window size of 300 frames/samples. The authors applied PCA to estimate the linearly uncorrelated basis from the raw BCG signals. The heart rate was estimated by applying periodicity estimation algorithm proposed by Irani et al. [13]. The periodicity estimation algorithm operated by estimating the most periodic PCA component by performing a stochastic search by using the frequency indexes derived from the discrete cosine transform (DCT). Therefore, the authors [11], estimated the heart rate by selecting first harmonic frequency by applying fast Fourier transform (FFT) on the most periodic component.

Also, BCG based heart rate estimation method is reliable in a circumstance where the face is occluded (see Fig. 2). Since the BCG based approach is mainly based on estimating the cyclical motion of feature points on surface of the head. The BCG based approach is highly advantageous, since the method is invariant to illumination variance. However, the dependence of motion features to estimate the BCG signal brings a major drawback for this approach. Where the method would be highly vulnerable to voluntary motion variations of the subject. Therefore, measuring the heart rate during voluntary head motion could degrade the reliability of the method.

3. Photoplethysmography

3.1. Theory

Hertzman and colleagues [14] first introduced Photoplethysmography in the 1940s, derived from plethysmography. Plethys-



Fig. 2. Reference frames of the videos for which BCG signals were extracted for the circumstance face occlusion (modified from [7]).

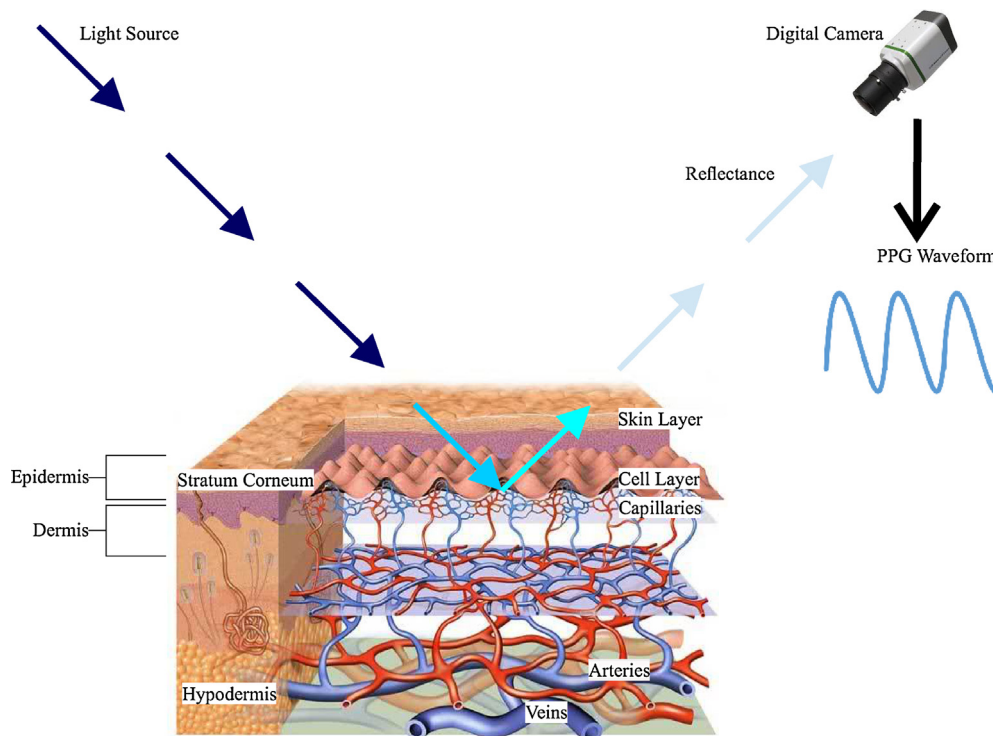


Fig. 3. Simplified description of PPG extraction from facial tissue (cross-section view).

mography is a concept that relies on the principle that blood absorbs light more than the surrounding tissues [15]. Therefore, photoplethysmography (PPG) is measured by the reflectance of the variation in light absorption due to volumetric changes of blood volume pressure. Ideal PPG signal contains vital physiological information, such as oxygen saturation (pulse oximetry), respiration rate, blood pressure, and heart rate for the assessment of automatic functions and detection of peripheral vascular diseases. PPG signal is extracted from the finger, arm or face of a subject. The PPG like pulsating signal is generated by the reflectance of ambient light that scattered through the skin layer, cell layer and the capillaries beneath the cell layer (see Fig. 3).

Capillaries are the smallest blood vessels in the human body circulating only about 2 to 5% of the total blood. Accounting to a very small volumetric change compared to the overall change of blood volume pressure in the entire body. Light propagation generates PPG signal through the skin (see Fig. 3), mainly through three layers of skin, stratum corneum, epidermis, and dermis. According to Lambert's law of light intensity [16], light reflected through the skin

can be regarded as diffusion and scattering. The stratum corneum of skin diffuses light while the epidermis and dermis scatter light. Epidermis layer mostly consists of melanin, whereas dermis consists of capillaries that circulate hemoglobin through their microvascular network; therefore, making dermis accountable to stronger scattering of light. The firm scattering of light constitutes to the microscopic color change in the subsurface of the skin. Therefore resulting in a small change in light intensity values that are recorded by the camera sensor (Fig. 4).

3.2. Related work

Verkrusse et al. [17] in 2008 proposed, the first remote PPG imaging method using ambient light. Here the authors measured PPG signals from subjects remotely (>1 m) using a digital camera. The method evolved on three main steps; extracting red (R), green (G) and blue (B) channels from the selected region of interest (ROI) and extracting the spatial average value of the pixels in the ROI per each frame of the video. Generating a raw PPG signal based on

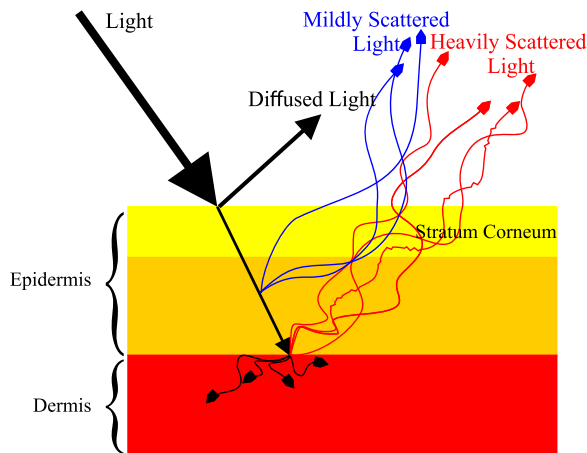


Fig. 4. Illustration of light propagation through simplified skin model.

the spatial average pixel values over time. Moreover, the raw PPG signal was filtered using a bandpass filter to extract a refined PPG signal.

Poh et al. [18] in 2010 proposed the first facial videos based PPG signal extraction method using blind source separation. The authors used a simple webcam camera to extract PPG signals from the facial region of the subject. The whole face was used as the ROI and was extracted using Viola-Jones face detector (VJFD) [8]. The spatial mean of each of R, G, B spectrums of the ROI was extracted. After which R, G, B raw signal traces were generated using the spatial mean of the pixel value of each frame over a period. The raw traces were then normalized and decomposed into three independent source signals using ICA [19]. After that, the best resembling independent source signal to a PPG like signal was used to measure the heart rate.

In 2011, Poh et al. [20], extended their work to refining the PPG signal by performing a series of temporal filtering. The filtering consisted of a passage of de-trending and smoothening the signals using moving average filter. The process generated a much-refined set of signals for blind source separation (BSS) using ICA [18]. Similar to the previous work [18], the best resembling independent source signal to a PPG like signal was selected as the PPG signal. The selected PPG signal was further refined by bandpass filtering at the operational frequency of 0.7 to 4 Hz corresponding to the heart-beat of 40 to 240 bpm. The heart rate was measured by extracting the PSD of the selected PPG signal.

De Haan et al. [21] proposed a PPG estimation method by deriving chrominance features from the RGB spectrum of the digital camera. The proposed method involved an approach of estimating the ratio of two orthogonal chrominance signals. The method operated by extracting two orthogonal chrominance signals from the RGB traces of the RGB spectrum. The RGB colors were transformed into chrominance colors by using an experimentally extracted fixed coefficients. The PPG signal were estimated by measuring the ratio between the orthogonal chrominance colors. De Haan et al. [22] also proposed a normalization function based on the different absorption spectra. The authors proposed a better PPG extracting algorithm by determining the light spectrum and the transfer characteristic of the optical filter in the camera. Here a normalized blood volume pulse vector was developed based on Hulsbusch noise-free spectrum model [23] to normalize the RGB color spectrum.

In 2014, McDuff et al. [24,25], proposed an improvement to the optical properties of the digital camera using a five band lens. The five-band camera was used to extract alternative frequency bands, in particular, orange and cyan in addition to the regular bands such as RGB. The approach was carried out to generate a higher overlap between the absorption spectra of hemoglobin and oxyhemoglobin

(HbO₂), which is about 520 to 580 nm. Hemoglobin (Hb) and Oxy-hemoglobin (HbO₂) are the two largest constituent chromophores in blood. Therefore, these chromophores make a significant influence on the rate of reflectance variation in light absorption that stimulates the PPG signal.

The absorption spectra of a digital camera range approximately from 425 to 675 nm, red, green and blue individually having an absorption spectra of 560 to 675 nm, 520 to 620 nm and 425 to 525 nm respectively. Therefore, the authors proposed a camera sensor model with overlapping color spectra combination of cyan (470 to 570 nm), orange (530 to 630 nm) and green (C, O, G) to maximize the accuracy of heart rate measurement. The heart rate extraction process from the C, O, G frames followed a similar process to [20]. The raw traces were extracted from the ROI by taking the spatial mean of the pixel intensity value of each spectrum. The signals were decomposed using ICA. The best signal to a signal resembling PPG like waveform was selected. The signal was further refined by band-pass filtering at operation frequency at 0.7 to 4 Hz. Finally, the heart rate was measured by taking the PSD of the signal.

Yan et al. [26], proposed a signal weighted analysis method to extract PPG signal. Here the authors used a strategy of taking the weighted average of the raw traces of the RGB color spectrum. The main idea was to maximize the signal and minimize the noise. Upon performing weighted signal analysis, the signal was processed for baseline drift removal and then de-noised using wavelet transform to extract the PPG signal. In 2015, Kumar et al. [5], proposed PPG estimation method by using a monochrome camera with a green spectrum filter. The authors extracted the PPG signal by using a weighted average on a combination of different ROIs in the face that are tracked separately.

The method comprised of two parts, maximizing the signal to noise ratio (SNR) by employing maximum ratio combination to combine the average pixel intensity of different ROIs. Extracting and tracking different parts of the face across a video. The authors used the deformable face fitting algorithm [12], to detect the face and extract landmark locations to track using KLT tracking algorithm [9]. The tracked feature points were stabilized using random sample consensus (RANSAC) [27] and these regions were sent to the maximum ratio combination to maximize the SNR.

The application of PPG extraction from a digital camera has evolved since 2010, from extracting physiological signals while sitting in front of the computer, monitoring patients and exercising in a gym [28]. The main focus area has been to improve the signal to noise ratio (i.e. signal strength), overcoming the effects of illumination variance and motion variance. The remaining of this section will be discussed with regards to the contributions made to overcome these issues.

3.2.1. Signal strength

PPG signal is extracted from the face as a reflectance of ambient light that is scattered through the skin. This scattering causes small changes of color in the subsurface of the skin. These minor changes are recorded as a microscopic change in light intensity values in the camera. The PPG signal is generated by extracting the spatial mean of a selected ROI over a period; this is the most common approach used by researchers [29]. The spatial mean of a ROI for a frame would give an average pixel value of all the pixels present in the ROI. Therefore, the microscopic changes in the ROI are extracted as a mean; this makes the PPG signal strength from facial videos weak in nature. Researchers have analyzed this issue and have come up with different solutions involving the selection of ROI.

Poh et al. [20], initially proposed modification on the selection of ROI from a selection of full face to an ROI of 60% of the width of the full face and full height. Lewandowska et al. [30], and Scalise et al. [31], proposed to select the ROI on the forehead region of the face (i.e. 40% width and 30% height of the full face). Aarts et al. [32], on a

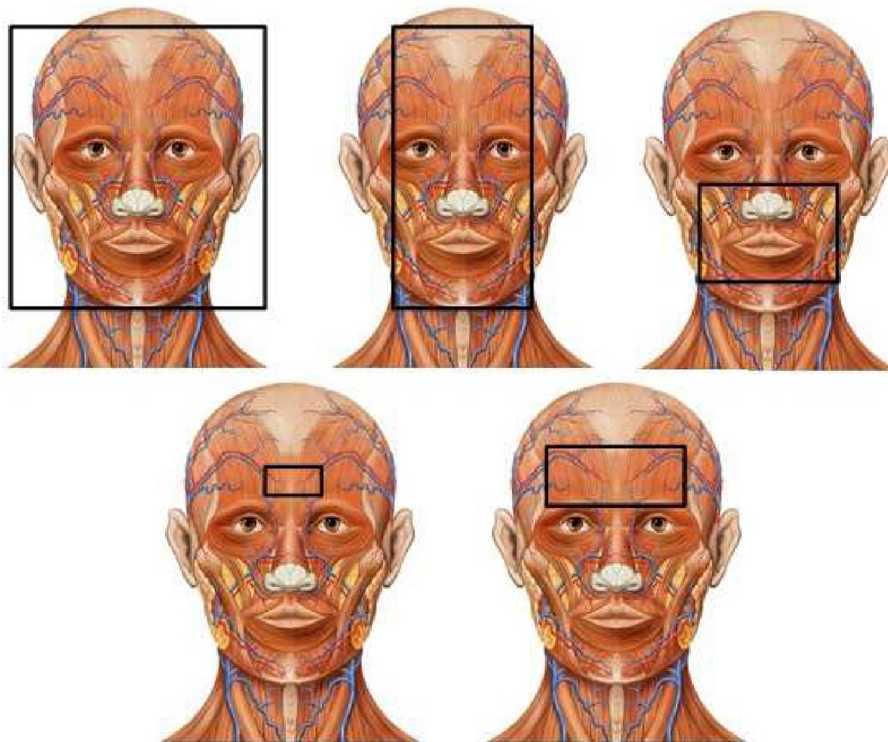


Fig. 5. ROI selection on the human face anatomy highlighting with respect to the microvascular bed of tissue.

pilot study, to monitor the heart rate of infants, used an ROI on the cheek region of the infant's face. Mesta et al. [33], also conducted a study for neonatal patients; here the authors used facial landmark fitting algorithm [12] to extract the ROI. Yong-Poh et al. [34], used an ROI below the eye line with a height of 20% and width of 55%, that cover skin region within the nose and cheeks. McDuff et al. [24], proposed an ROI selection similar to [20]. However, the eye region was excluded using facial landmarks. Li et al. [35], proposed an ROI using discriminative response map fitting (DRMF) and selected the entire area of the face below the eyes.

Kumar et al. [5], proposed ROI selection from the entire face using 20×20 pixel blocks and the goodness metrics to estimate the best combination of ROI's. Lam et al. [36] also, proposed an ROI selections criteria using facial landmarks fitting. Here the authors used random patches from the landmarks in the facial region within below the eyes and the lips. Feng et al. [37], proposed a trapezoidal ROI with 50% and 40% width, and 58% height, that covers facial skin region within the low forehead, cheeks, and upper lips. The ROI selection from Poh et al. [18], was modified to exclude non-facial regions that would affect the strength of the PPG signal. The selection of ROI within the facial parts of the cheek, lips and chin, were based on physiological facts of the human anatomy (see Fig. 5). These skin regions comprise a higher proportion of capillaries that would generate greater signal strength due to the higher absorption of light compared to others in regions of the face.

However, many researchers have selected the forehead region that consists of a lesser proportion of capillaries underneath the skin. This is because the forehead is less prone to muscle movements compared to other areas of the face. Muscle movements within the face would constitute to weaken the PPG signal strength by adding noise in the form of motion artifacts. In general, it is noticed that the eye region was excluded from the ROI since the eye does not consist pulsatile component related to blood volume change and also, blinking of the eye would result in motion artifacts that would weaken the PPG signal. Extracting PPG signal from

multiple ROI has been shown to improve the overall signal strength using comparative [35,5] and stochastic approaches [36].

3.2.2. Illumination variance

The human face is non-symmetric in nature, which leads to disproportional scattering of light through the surface of the skin. The positioning of the light source and the orientation of the face cause illumination variance. The non-symmetric nature of the face leads to spatial illumination variance across the skin surface (see Fig. 6). Diffusion of light on the non-symmetric skin plane results on illuminating the skin with different color intensities. Therefore, faulty pixel values are allocated (as seen in the zoomed image of the ROI in Fig. 6). These pixel values were recorded in contrary to the actual skin tone in the real world. Temporal illumination variance is caused by a change in the intensity/position of the lighting source over a period (see Fig. 7). Similar to spatial illumination variance, temporal illumination variance also results in the faulty allocation of pixel values of the skin. Spatial illumination variance is a challenge for the methods which depends on spatial averaging of pixels to extract the raw PPG signal sample. Whereas, temporal illumination variance challenges PPG signal extraction over a period.

Researchers have come up with solutions to overcome the effect of illumination variance. However, most of the solutions at present, address the effect caused by temporal illumination variance. Li et al. [35] proposed an illumination rectification method by taking the background region color as a reference to estimate the illumination variance. Here illumination is modeled in the form of noise that affects the PPG signal over a period, as an additive operator. Therefore, the illumination is rectified by subtracting the noise from the PPG signal. Lam et al. [36] proposed an illumination rectification method by using two points on the skin of the face in the same frame to extract the green spectrum of each point. After that, applying independent component analysis to extract the PPG signal from the two green spectrum signals. Therefore, treating the effect of illumination variance as a blind source separation problem.

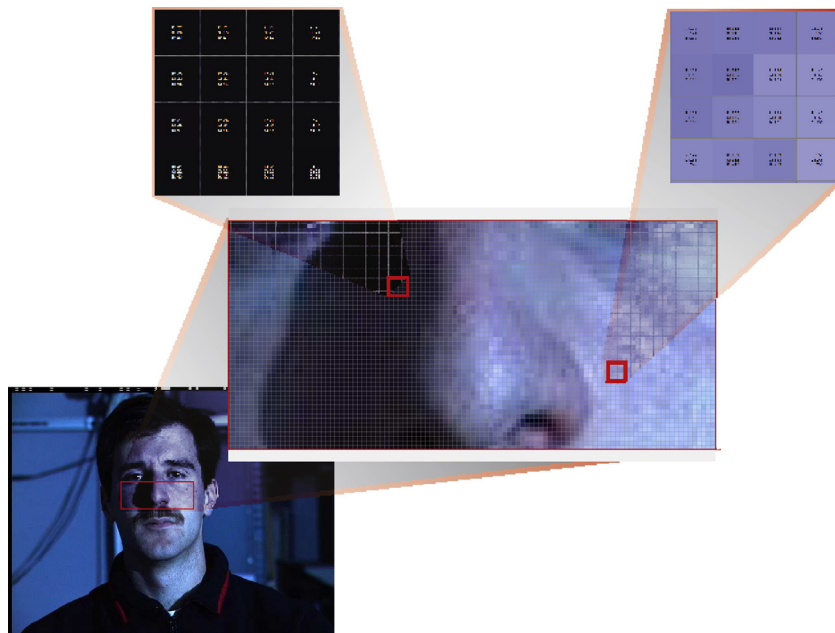


Fig. 6. Illustration of spatial illumination variance across the facial skin surface.



Fig. 7. Illustration of temporal illumination variance on the face across a period of time.

Most recently, Moco et al. [38], highlighted the effect of ballistocardiographic artifacts in PPG signals due to the position of the lighting source. Here the authors reported that the use of a homogeneous and orthogonal illuminating source could minimize ballistocardiographic artifacts in PPG signals, which is treated as noise caused by illumination variance. Researchers have treated the problem of illumination variance as a generalized motion artifact removal process. However, the effect of spatial and temporal illumination variance pose as a great challenge to PPG signal estimation.

3.2.3. Motion artifacts

PPG signals extracted from facial features are affected by motion artifacts caused by rigid and non-rigid movement of the face and body. Motion artifacts resulting from the rigid movement are due to the change of head orientation or by a change in posture of the body. Motion artifacts caused by the non-rigid movement are due to movements of the face, often caused by expressing facial emotions such as: smiling, yawning and talking. Rigid and non-rigid movement of the face and body creates a change in the illumination on the skin, which results in fluctuation of color intensity

Table 2

Operational and sensor parameters of the experiments conducted in previous studies Part-1.

Citation	Sensor parameter	Distance from sensor (meters)	Experiment parameter	No. of subjects	Availability
Poh et al. 2010 [18]	Computer Webcam, RGB color, 640 × 480 resolution, 25 fps, saved as AVI	0.5	Controlled Illumination, Indoor	12 participants (10 males, 2 females), Skin color: Asian, African and Caucasian	No
Lewandowska et al. 2011 [30]	Computer Webcam, RGB color, 640 × 480 resolution, 20 fps, Saved as AVI	1	Natural lighting mid-day, Indoor	12 participants (10 males, 2 females) Skin color: Caucasian	No
Poh et al. 2011 [20]	Computer Webcam, RGB color, 640 × 480 resolution, 15 fps, Saved as AVI	0.5	Controlled Illumination, Indoor	12 participants (8 males, 4 females) Skin color: Caucasian	No
Scalise et al. 2012 [31]	Digital web camera WeC, (Microsoft LifeCam VX-3000). RGB color, 640 × 480 resolution, 30 fps	0.2	Controlled Illumination, Indoor	7 subjects (3 males, 4 females)	No
Kwon et al. 2012 [48]	I phone front camera, RGB color, 640 × 480 resolution, 30 fps	0.3	Controlled Illumination, Indoor	10 subjects (8 males, 2 females)	No
Aarts et al. 2013 [32]	Standard color digital camera, RGB color, 300 × 300 face detector, 15 fps, Saved as AVI	1	Controlled Illumination, Indoor	19 subjects (infants age: 25 weeks to 42 weeks)	No

values in the ROI. Therefore, these movements, generate motion artifacts within PPG signal. This causes the signal to drift from the baseline and attenuate the information present in the signal. Until now the majority of the research has focused on contributing to eliminating the effect caused by motion artifacts.

Researchers have treated the issue of motion artifact removal as a blind source separation problem. Initially, Poh et al. [18], applied independent component analysis to separate the source signal from the raw signals which were contaminated by motion artifacts. Lewandowska et al. [30], also treated the motion artifacts removal as a blind source separation problem and applied PCA to extract the PPG signal. Following this direction, many researchers have proposed methods using ICA [25,39–43], PCA [31], fast ICA [32,44] and robust ICA [45]. Each of the methods mentioned above used either ICA or PCA to extract the source signal from the raw signals. However, neither of these methods selected a certain component (i.e. component 1, component 2, ..., etc.), rather the component with the best match to a PPG like signal was selected manually.

Monkarezi et al. [46] proposed to improve the motion artifact removal method involving ICA and machine learning approach. The authors used linear regression and K-nearest neighbor classifiers to classify the heart rate value extracted of the PSD. Here the idea was to automate the component selection processes by selecting the best corresponding heart rate. Li et al. [35], proposed a two-step approach to nullify the effect caused by the motion artifacts. The authors employed the KLT tracking algorithm to track through the ROI within consecutive frames to minimize the loss of ROI, which causes loss of samples in the signal. Further, the motion artifacts were eliminated by separating the PPG signal into segments and eliminating the segments with the highest standard deviation/fluctuation.

Lam et al. [36], proposed an approach using facial landmarks. Seventeen facial landmarks that were less prone to non-rigid motion variance were selected. The authors used 41 × 41 pixel patches on each side of the face and extracted PPG signal and decomposed using ICA. The ICA component was selected using a majority voting scheme to eliminate the motion artifacts and to estimate an accurate heart rate. Feng et al. [37], proposed an adaptive green and red differentiation operation to remove motion artifacts of the PPG signal. The adaptive green and red differentiation operated based raw traces that were extracted from the ROI.

The difference between the green and the red signals were extracted because the green signal contains information related to the variation of blood volume pressure of the capillaries. Whereas, the red signal contains information related to muscle movements

in the face. Wang et al. [47], proposed a motion robust PPG estimation module, by compensating the rigid and non-rigid movement of the ROI. The approach was based on the idea that the camera can capture multiple skin sample in an image and consider each pixel region as an independent source. Therefore, using multiple samples of the same image to eliminate the motion artifacts by spatial pruning and temporal filtering pixel to samples of the PPG signal.

In general, researchers have proposed to overcome the motion artifact problem caused by motion variance as a BSS problem solving or a adaptive filtering/pruning approach. The BSS based motion artifact compensation approach rise the issue of component selection and autonomous heart rate estimation; whereas, the filtering/pruning approaches raise the issue of over compensation/under compensation. However, the recent direction of motion artifact compensation has evolved in the direction of adaptive matrix computation approach and blind source separation approach.

4. Experiment setup

Many contributions have been presented by researchers to improve heart rate measurement from the facial image using BCG and PPG. Upon each proposal researchers have shown improvement in the accuracy of heart rate measurement. These claims can be subjectable to the data which was used during the experiment conducted. The majority of the data are not publicly available, which constrains researchers from cross-validating their methods. We have outlined a detailed summary (see Tables 2–4) containing operational and sensor parameter of the experiments conducted in previous studies. During our experiment, we will cross-validate some of the significant literature from BCG and PPG. The cross-validation will be carried out on the publicly available human-computer interaction (HCI) database.

MAHNOB-HCI database is considered as one of the most challenging HCI databases since the data is recorded under realistic human-computer interaction scenarios. The database is a publicly available multi-modal database recorded by [49]. MAHNOB-HCI consists of two main experiments, 'emotion elicitation experiment', and 'implicit tagging experiment'. During this study, we used the emotion elicitation experiment since this experiment consists of sessions of longer duration. The database provides challenges such as illumination variance (i.e. spatial and temporal) and motion artifacts (i.e. rigid and non-rigid movement). The database also provides a range of skin tones from pearl white, fair, olive to black across 30 subjects.

Table 3
Operational and sensor parameters of the experiments conducted in previous studies Part-2.

Citation	Sensor parameter	Distance from sensor (meters)	Experiment parameter	No. of subjects	Availability
Balakrishnan et al. 2013 [7]	Standard color digital camera, RGB color, 1280 × 720 resolution, 30 fps, Saved as AVI	1	Controlled Illumination, Indoor	18 subjects (7 males, 11 females)	No
Yu et al. 2013 [34]	Standard color digital camera, RGB color, 720 × 576 face detector, 25 fps	0.6	Natural lighting mid-day, Indoor	1 subject (3 phases/videos)	No
Mestha et al. 2014 [33]	Standard color digital camera, RGB color, 640 × 480 resolution, 30 fps, Saved as.mp4	0.9	Controlled Illumination, Indoor	9 subjects (4 males, 5 females)	No
Monkaresi et al. 2014 [46]	Logitech Webcam Pro 9000, RGB color, 640 × 480 resolution, 30 fps, Saved as AVI	–	Controlled Illumination, Indoor	10 subjects (8 males, 2 females)	No
Wen Jun et al. 2014 [43]	Smartphone camera, RGB color, 50 × 50 face detector, 30 fps	0.3	Controlled Illumination, Indoor	15 subjects (12 males, 3 females)	No
Li et al. 2014 [35]	Standard color digital camera, RGB color, 780 × 580 resolution, 61 fps	0.4	Controlled Illumination light source and Illumination from LCD screen, Indoor	27 subjects (15 females and 12 males)	Yes, MAHNOB-HCI database

Table 4
Operational and sensor parameters of the experiments conducted in previous studies Part-3.

Citation	Sensor parameter	Distance from sensor (meters)	Experiment parameter	No. of subjects	Availability
Lam et al. 2015 [36]	Standard color digital camera, RGB color, 780 × 580 resolution, 61 fps	0.4	Controlled Illumination light source and Illumination from LCD screen, Indoor	27 subjects (15 females and 12 males)	Yes, MAHNOB-HCI database
Yonggang et al. 2015 [26]	CCD camera, 24-bit RGB 1280 × 720 resolution pixels) at 15 fps	0.5	Controlled Illumination, Indoor	15 subjects (7 males, 8 females)	No
Feng et al. 2015 [37]	Computer Webcam, 24-bit RGB color, 640 × 480 resolution, 30 fps, Saved as AVI	0.75	Controlled Illumination, Indoor	10 subjects (7 males, 8 females)	No
Haque et al. [11]	Standard color digital camera, RGB color, 780 × 580 resolution, 61 fps	0.4	Controlled Illumination light source and Illumination from LCD screen, Indoor	27 subjects (15 females and 12 males)	Yes, MAHNOB-HCI database

During our experiment, we used 27 subjects of which 15 were male and 12 were female of age 19 to 40 years old (mean of 26.06 and standard deviation of 4.39). We used seven sessions for each subject about emotions such as neutral, anger, amusement, joy, fear, sadness and disgust. Each session, we used 90 s of the recording excluding the initial 10 s. We used the data, recorded by the color camera of 61 fps (frames processed per second) at a resolution of 780 × 580. The camera was positioned in front of the subject at a distance of 0.4 m. The ECG signal was recorded using three sensors attached to the body. We used the ECG signal of channel 34 from the sensor attached to the upper left corner of the chest, under clavicle bone. This ECG signal was used as ground truth to measure the heart rate HR_{gt} .

We implemented some significant contributions that were selected from Tables 5 and 6 based on their overall novelty and contribution towards improving the accuracy of heart rate estimation. The table mainly classifies the methods based on the methodology used, signal selection criteria, heart rate (HR) estimation method and proposed enhancement. The adopted methodology was mainly discussed based on the methods used to detect the face, (i.e. ROI), usage of blind source separation and operational parameters of the bandpass filter. The proposed strategies and methodology were classified to corresponding literature to highlight the overall significance of each method.

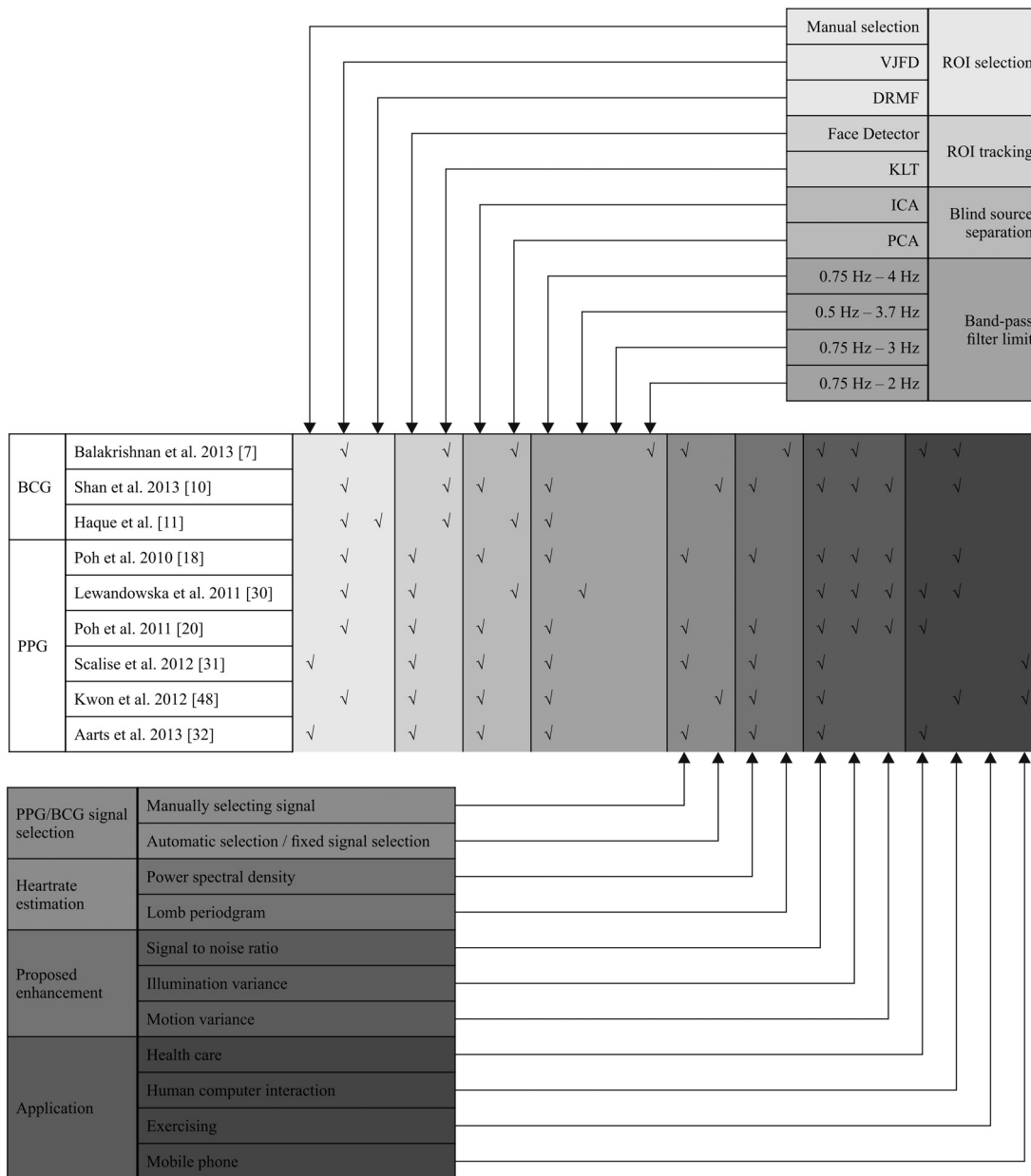
The experiment was conducted in Matlab environment. We have implemented method from the initially proposed method of Poh et al. [18] 2010 based on applying BSS using ICA to the temporal traces of the RGB spectrum and the proposed method of

Lewandowska et al. [30] 2011, based on applying BSS using PCA. We also implemented the proposed method of Poh et al. [20] 2011 based on applying BSS using ICA to the temporal traces of the RGB spectrum and using temporal filtering to refine the PPG signal. We implemented these three methods since majority of the contributions [25,31,32,40–46], proposed during 2010 to 2013 were based on similar methodology. The heart rate estimation method based BCG were also implemented, here we tested the proposed methods of Balakrishnan et al. [7] 2013, Shan et al. [10] 2013.

The proposed methods of Li et al. [35] 2014, and Lam et al. [36] 2015, based green spectrum were also implemented due to their overall novelty in estimating the PPG signal and measuring the heart rate. The proposed method of Feng et al. [37] 2015, was also implemented to validate the performance of using the green and red spectrum of the RGB camera. Finally a recent method by Haque et al. [11] 2016, based on BCG was also implemented due its novelty in feature selection and heart rate estimation approach.

The experiment was carried under two phases; 1. Simplistic scenario and 2. Dynamic scenario. The purpose of the experiment on the simplistic scenario is to demonstrate that the methods were implemented correctly. Therefore, we used the data of the neutral session of the 27 subjects, since this session exhibited least amount of illumination and motion variance. Whereas, the dynamic scenario pertaining rigid and non-rigid movement and illumination variance exhibited greater diversity and challenge to the experimented method. During the dynamic scenario, we experiment the methods on 27 subjects for all six sessions about anger, amusement, joy, fear, sadness and disgust.

Classification of heart rate estimation using facial videos methods based on specification and application part-1.


$$HR_{error} = |HR_{video} - HR_{gt}| \quad (1)$$

The initial methods based on PPG (i.e. Poh et al. [18] and Lewandowska et al. [30]) using BSS to estimate HR suffered greatly during our experiment. The methods reported a mean error rate about 20 bpm (see Table 8) and the lowest correlation coefficient between HR_{gt} and HR_{video} . Poh et al. 2011 [20], improved these results by using enhanced ROI selection along with a series of temporal filtering. The enhancement resulted in reduced the mean error rate to 11.87 bpm and reported a higher correlation coefficient of 0.52 between HR_{gt} and HR_{video} . However, [20] was highly challenged by illumination variance and motion variance; this was shown by the high standard deviation of 10.26 bpm and root mean

Table 6
Classification of heart rate estimation using facial videos methods based on specification and application part-2.

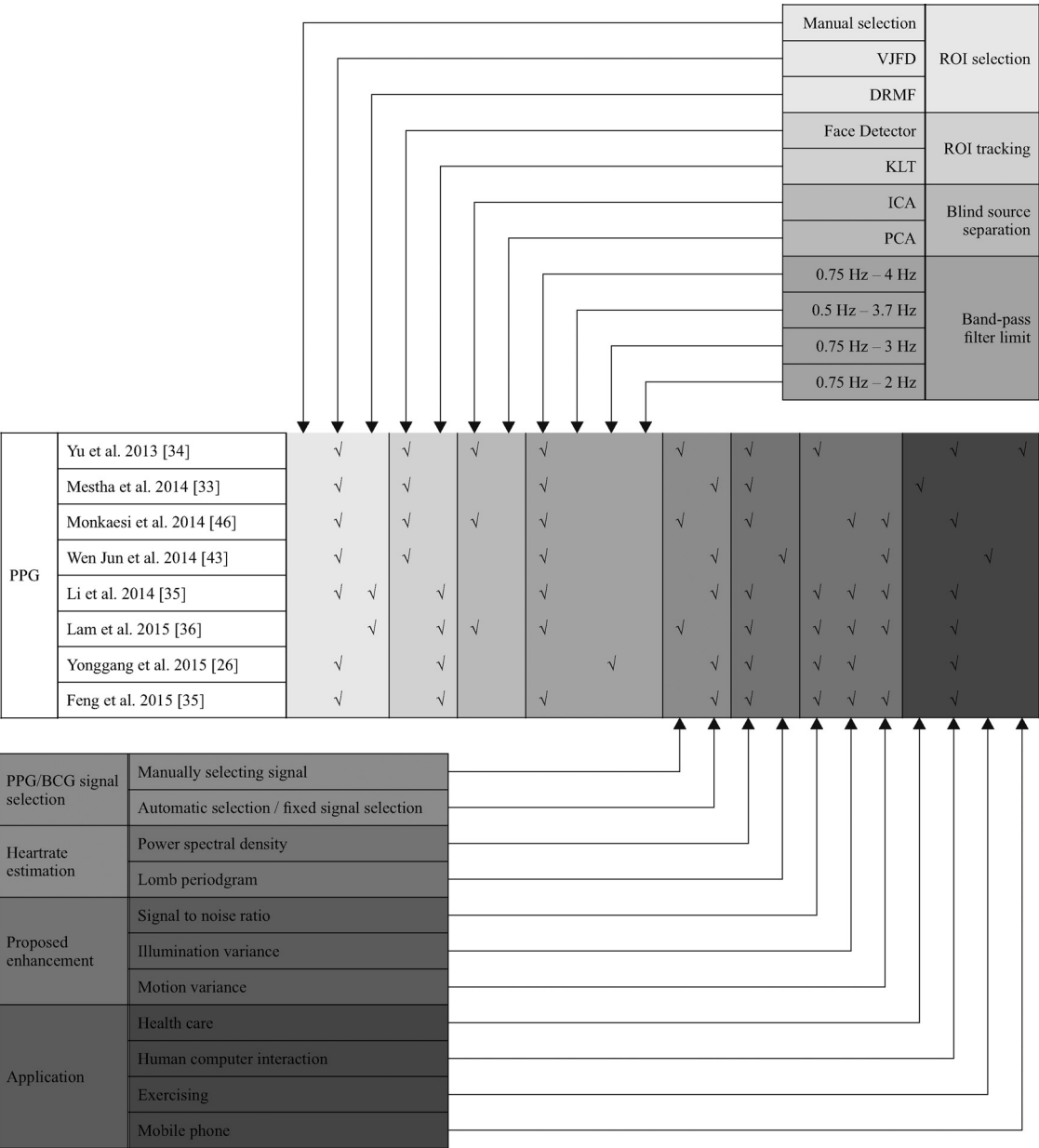


Table 7
Performance validation of the methods for the simplistic scenario experiment.

Method	Experiment results				Reported results	
	Me (bpm)	SD (bpm)	RMSE (bpm)	<i>r</i>	RMSE (bpm)	<i>r</i>
Poh et al. 2010 [18]	15.24	12.82	21.08	0.34	25.9	0.08
Lewandowska et al. 2011 [30]	16.47	13.21	22.47	0.31	–	–
Poh et al. 2011 [20]	10.00	12.66	14.01	0.44	13.60	0.36
Balakrishnan et al. 2013 [7]	17.00	14.66	22.01	0.14	21.0	0.11
Shan et al. 2013 [10]	7.48	4.76	9.11	0.55	–	–
Li et al. 2014 [35]	7.14	9.53	12.47	0.53	15	–
Lam et al. 2015 [36]	6.49	8.54	10.34	0.66	8.9	–
Feng et al. 2015 [37]	6.64	8.01	10.12	0.64	–	–
Haque et al. 2016 [11]	4.69	3.43	5.96	0.86	3.85	–

square error of 15.67 bpm, which highlights the unstable nature of the method under diverse situations.

Balakrishnan et al. [7] HR measurement method based on BCG also suffered greatly during the experiment. The method reported

the highest mean error rate of 21.68 bpm and the lowest correlation coefficient of 0.1. The method also reported the higher standard deviation and root mean square error (see Table 8). This approach was clearly unstable under the diversity of the dynamic scenario

Table 8

Performance validation of the methods for the dynamic scenario experiment.

Method	Me (bpm)	SD (bpm)	RMSE (bpm)	<i>r</i>
Poh et al. 2010 [18]	19.54	9.58	21.75	0.18
Lewandowska et al. 2011 [30]	20.44	12.92	24.16	0.03
Poh et al. 2011 [20]	11.87	10.26	15.67	0.52
Balakrishnan et al. 2013 [7]	21.68	11.91	24.72	0.10
Shan et al. 2013 [10]	44.99	34.15	56.42	0.12
Li et al. 2014 [35]	9.78	7.45	11.71	0.69
Lam et al. 2015 [36]	8.73	4.94	10.02	0.81
Feng et al. 2015 [37]	8.05	7.04	10.36	0.62
Haque et al. 2016 [11]	6.47	3.92	7.56	0.84

and also failed in 20% of the sessions to extract features consistently. This resulted on restricting the method to measure heart rate. Shan et al. [10], reported a high mean error rate of 44.99 bpm with a standard deviation of 34.15 (see Table 8). These results were vastly different to the results of the simplistic scenario (see Table 7). The modified ROI on the forehead along with the motion artifact compensation using ICA was able to reduce the error rate compare to [7].

However, the method performed poorly on the dynamic scenario since Shan et al. [10] failed to determine the heart rate at 56% of the sessions due to inconsistent feature extraction. This raises the issue of reliability since the method only operated on merely a half of the session of this experiment. Therefore, the error rate raised for the dynamic scenario. In general, BCG based methods failed to determine heart rate under the presence of motion variation in the ROI/subject. This resulted in a high error rate or failed to estimate heart rate due to the inconsistent feature extraction.

The proposed method of Li et al. [35], based on PPG extraction from the green spectrum was able to improve the reliability for heart rate estimation from the blind source separation based approach. The illumination rectification and non-rigid motion compensation methods were also estimate the PPG signal compared to the BSS approach. The improvement shown by method were established with, a mean error rate of 9.78 bpm with a higher correlation coefficient of 0.69 (see Table 8). The method was able to provide reliable heart rate measurement for most of the sessions. However, the error rate increased at sessions with greater diversity. This was highlighted by the high root mean square error of 11.36 bpm. The error rate may have increased due to the illumination rectification method, since the authors assumed that the illumination variance of the background of the scene as a reference to the illumination present in the face. The assumption may have hindered the reliability of Li et al. [35] as the illumination variance of the background is not always the same as the illumination variance of the face.

Lam et al. [36] overcame the limitations of Li et al. [35] by applying BSS to the green spectrum of the RGB camera. The method reported a higher Pearson correlation coefficient of 0.81, with a low root mean square error. The method reported a high mean error of 8.73 bpm with a low standard deviation below 5 bpm. The results showed that [36] operated reliably for most cases as the method reported a low standard deviation and high Pearson correlation coefficient. The application of the majority voting scheme could have contributed to estimating the hear rate reliably. However the high mean error could have occurred due to the random patch selection processes, since the method dose not always select the patches of same illumination spectrum and selects the patches with the highest similarity. This may cause the BSS approach to under perform while estimating the PPG signal.

Feng et al. [37], proposed a HR measurement method using the red and green spectrum of the RGB camera. Here the authors used a adaptive green and red differentiation operation to estimate the PPG signal and measure the heart rate. The method reported a

similar RMSE compare to [36], (see Table 8) inferring that both method performed similarly during our experiment. However, the high standard deviation and the low Pearson correlation coefficient suggest that the method was less stable compare to [36]. The under-performance could have occurred due green and red differentiation operation as the temporal signal of the red spectrum may cancel the pulsatile components of the resultant PPG signal.

The motion feature based approach proposed by Haque et al. [11] operated robustly compared to rest of the method that were discussed during the experiment. The method reported the highest Pearson correlation coefficient for both experiments. The method operated much reliably during the simplistic scenario experiment (see Table 7) by reporting a low mean error with standard deviation and root mean square error, within limits of an acceptable error margin. The feature fusion along with the periodicity based BCG signal estimation approach may have contributed to improving feature stability and the reliability of the method. However, mean error and root mean square error of the method increased during the dynamic scenario experiment (see Table 8). The increase of the error rate may have occurred due window size selection for the moving average smoothing approach. The authors used a window size of 300 frames/samples compared to the 5 frames/samples used by [20], this may over smoothen the BCG signal and bluer physiological information.

The Bland-Altman analysis was derived from the reading of HR_{gt} and HR_{video} for all the sessions of the dynamic scenario experiment. Here the Bland-Altman graph was plotted in Fig. 8 to provide an overview of the experiment. It can be seen that the Bland-Altman analysis estimated a poorly grouped plot (see Fig. 8(a) and (b)) for [18] and [30] accordingly. The analysis reported a 95% agreement between 50 bpm to −14 bpm with a bias of 18 bpm, this show that the methods measured heart rates with an error rate up to 50 bpm at certain sessions. Poh et al. 2011, maintained a better reliability by showing a densely grouped plot for most of the sessions (see Fig. 8(c)). However, the analysis reported a 95% agreement between 22.16 bpm to −35.55 bpm with a bias of −6.5 bpm, since the method reported a high error rate for a session of greater diversity.

The poorly grouped and populated graphs for BCG based methods [7,10] clearly shows (see Fig. 8(d) and (e)) the poor functionality and reliability. The Bland-Altman analysis reported a 95% agreement between 47.55 bpm to −6.89 bpm with a bias of 21.33 bpm. This clearly showed that the [7] method recorded an error rate of 21.33 bpm for the most of the sessions and recorded a high error rate up to 47.55 bpm for certain sessions. Li et al. 2014 [35] reported a densely grouped plot with a 95% agreement between 15.69 bpm to −25.83 bpm with a bias of −5.07 bpm (see Fig. 8(f)). This method recorded a lower bias point compare to the PPG methods using blind source separation. However, the method reported a poor grouping towered the negative direction of the difference of the heart rate axis. This is a result due to the strategy of segment based motion artifact removal approach. The segment based motion artifact removal approach eliminates motion artifact by removing signal segments with the highest standard deviation.

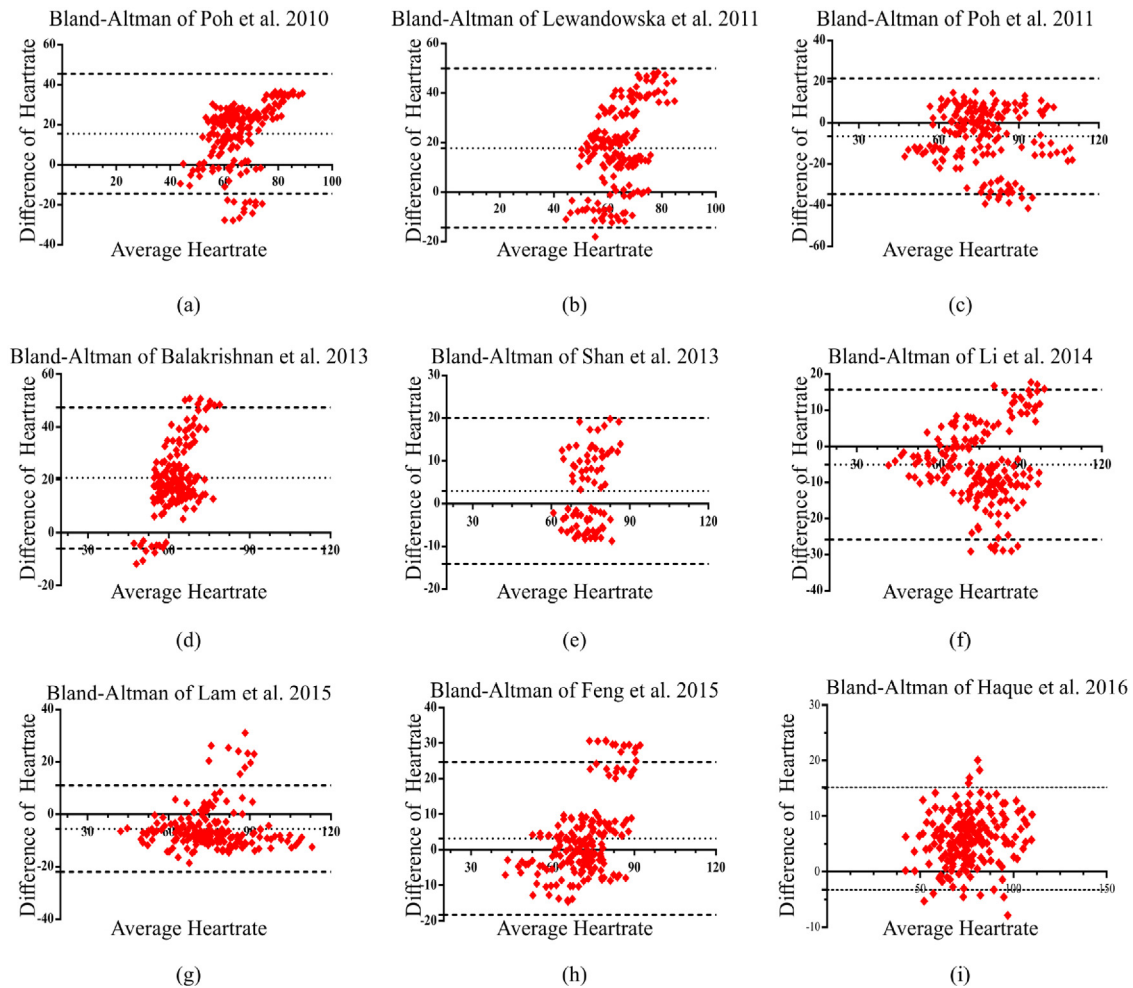


Fig. 8. Bland-Altman analysis for Heart rate measurement of dynamic scenario experiment.

Therefore, attenuating the information and the synchronous nature of the signal.

Proposed method of Lam et al. 2015 [36] reported a densely grouped plot with 95% agreement between 11.05 bpm to -21.99 bpm with a bias of -5.46 bpm (see Fig. 8(g)). The dense grouping showed that the method was able to compensate most of the challenges compared to rest of the methods. The application of blind source separation for artifacts removal improved the robustness of this method compared to other methods. Feng et al. 2015 [37] also reported a densely grouped plot (see Fig. 8(h)) with 95% agreement between 24.64 bpm to -18.25 bpm with a bias of 3.19 bpm. The method reported the lowest bias limit compare to all another method. Finally, the Bland-Altman analysis of Haque et al. 2016 [11], also reported a dense plot with 95% agreement between -3.3 bpm to -15 bpm with a bias of 5.9 bpm (see Fig. 8(i)).

The proposed methods by Lam et al. 2015 [36] based on estimating the PPG signal and method of Haque et al. 2016 [11] based on estimating the BCG signal operated robustly compared to rest of the methods. The application of majority voting scheme based heart rate estimation and periodicity based BCG signal estimation improved the robustness of the heart rate estimation methods and automated the heart rate measurement process. However the reliability of the motion artifact compensation method are subjective as the motion artifact compensation methods often either over compensated or under compensated during our experiment sessions. The summary of the discussion for this experiment is outlined in Table 9. The table will discuss multiple aspects of the proposed

methods about the effectiveness of the ROI selection, feature stability and motion artifact removal method/strategy and the reliability of the methods for both scenarios.

6. Conclusion

This paper has critically reviewed the progress and presented the state of the art in heart rate monitoring using facial videos. We investigated two mainstream heart rate measurement concepts, PPG and BCG. The principles and theory behind each of these methods were discussed on heart rate measurement using facial videos. The significance and challenges were discussed along with the contributions proposed to improve the reliability. We experimented certain methods based on their overall significance on a common experimental platform. We cross-validated and evaluated the reliability of the methods for heart rate measurement under realistic situations.

The findings of the experiment for heart rate measurement under the realistic situation, suggests that heart rate measurement by facial videos is promising under simplistic situations (i.e. idle person and controlled lighting). BCG based approach showed promising results. However, the inconsistent feature extraction was a major drawback to the reliability of the method. Invariance to illumination was one of the great advantages of BCG based methods. Recent contributions of [11] have addressed the issues related to feature stability. However the motion artifact removal method cause the method to under-perform. Further research should be

Table 9

Summary on the effectiveness of the facial imaging based heart rate measuring methods, the effectiveness of the method is ranked with star symbol *** where * is the minimum value and ***** is the maximum value.

Method	ROI selection	Feature stability	Motion artifact removal	Reliability	
				Simplistic	Dynamic
Poh et al. 2010 [18]	*	**	**	*	*
Lewandowska et al. 2011 [30]	**	**	*	*	*
Poh et al. 2011 [20]	***	***	***	***	**
Balakrishnan et al. 2013 [7]	**	*	**	*	*
Shan et al. 2013 [10]	***	*	*	*	*
Li et al. 2014 [35]	***	*****	***	****	****
Lam et al. 2015 [36]	*****	****	****	****	****
Feng et al. 2015 [37]	****	*****	****	****	****
Haque et al. 2016 [11]	****	****	***	****	****

conducted to compensate the effects of motion artifacts during motion variance sessions for reliable heart rate measurement.

Photoplethysmography based heart rate measurement is the major driving trend. The majority of the research has been presented to enhance the reliability. The direction for extracting PPG signal from multiple ROI has improved the overall signal strength. However, many improvements should be made for heart rate measurement under realistic situations. Much more attention should be given towards rectifying spatial and temporal illumination variance and motion variance caused by the change of body posture. Combining the features of the BCG based approach of [11] and PPG based approach of [36,37] could overcome the limitations of the methods related to motion variance and illumination variance. Deriving a solution based on Lambert's law of light intensity could be a promising direction towards achieving a reliable heart rate measurement method. Also, considering the digital camera as a sensor and its accountability towards quantization noise would help to minimize the effects caused by noise.

This paper has strictly discussed on heart rate measurement. However, extraction of physiological signals can lead to determining oxygen saturation (pulse oximetry), respiration rate, blood pressure, and heart rate, assessment of automatic functions and detection of peripheral vascular diseases. The non-invasive nature of this application opens possibilities for health monitoring towards various fields such as health care, telemedicine, rehabilitation, sports, ergonomics and crowd analytics.

Acknowledgements

This research work has been supported by the HiCoE grant for CISIR (Ref No. 0153CA-002), Ministry of Education (MOE), Malaysia.

References

- [1] A. Bauer, M. Malik, G. Schmidt, P. Barthel, H. Bonnemeier, I. Cygankiewicz, P. Guzik, F. Lombardi, A. Müller, A. Oto, et al., Heart rate turbulence: standards of measurement, physiological interpretation, and clinical use: international society for holter and noninvasive electrophysiology consensus, *J. Am. Coll. Cardiol.* 52 (17) (2008) 1353–1365.
- [2] S. Kazemi, A. Ghorbani, H. Amindavar, C. Li, Cyclostationary approach to doppler radar heart and respiration rates monitoring with body motion cancellation using radar doppler system, *Biomed. Signal Process. Control* 13 (2014) 79–88.
- [3] J. Kranjec, S. Beguš, G. Geršak, J. Drnovšek, Non-contact heart rate and heart rate variability measurements: a review, *Biomed. Signal Process. Control* 13 (2014) 102–112.
- [4] F. Bousefsaf, C. Maaoui, A. Pruski, Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate, *Biomed. Signal Process. Control* 8 (6) (2013) 568–574.
- [5] M. Kumar, A. Veeraraghavan, A. Sabharwal, Distanceppg: Robust non-contact vital signs monitoring using a camera, *Biomed. Opt. Express* 6 (5) (2015) 1565–1588.
- [6] O.T. Inan, Recent advances in cardiovascular monitoring using ballistocardiography, in: *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, IEEE, 2012*, pp. 5038–5041.
- [7] G. Balakrishnan, F. Durand, J. Guttag, Detecting pulse from head motions in video, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 3430–3437.
- [8] P. Viola, M.J. Jones, Robust real-time face detection, *Int. J. Comput. Vis.* 57 (2) (2004) 137–154.
- [9] C. Tomasi, T. Kanade, *Detection and Tracking of Point Features*, School of Computer Science, Carnegie Mellon Univ., Pittsburgh, 1991.
- [10] L. Shan, M. Yu, Video-based heart rate measurement using head motion tracking and ICA, in: *2013 6th International Congress on Image and Signal Processing (CISP)*, vol. 1, IEEE, 2013, pp. 160–164.
- [11] M.A. Haque, R. Irani, K. Nasrollahi, T.B. Moeslund, Heartbeat rate measurement from facial video, *IEEE Intell. Syst.* 31 (3) (2016) 40–48.
- [12] J.M. Saragih, S. Lucey, J.F. Cohn, Deformable model fitting by regularized landmark mean-shift, *Int. J. Comput. Vis.* 91 (2) (2011) 200–215.
- [13] R. Irani, K. Nasrollahi, T.B. Moeslund, Improved pulse detection from head motions using DCT, in: *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*, vol. 3, IEEE, 2014, pp. 118–124.
- [14] A.B. Hertzman, J.B. Dillon, Applications of photoelectric plethysmography in peripheral vascular disease, *Am. Heart J.* 20 (6) (1940) 750–761.
- [15] D.S. Seidman, J. Moise, Z. Ergaz, A. Laor, H.J. Vreman, D.K. Stevenson, R. Gale, A prospective randomized controlled study of phototherapy using blue and blue-green light-emitting devices, and conventional halogen-quartz phototherapy, *J. Perinatol.* 23 (2) (2003) 123–127.
- [16] W.R. McCluney, *Introduction to Radiometry and Photometry*, Artech House, 2014, pp. 33–57.
- [17] W. Verkrusysse, L.O. Svaasand, J.S. Nelson, Remote plethysmographic imaging using ambient light, *Opt. Express* 16 (26) (2008) 21434–21445.
- [18] M.-Z. Poh, D.J. McDuff, R.W. Picard, Non-contact automated cardiac pulse measurements using video imaging and blind source separation, *Opt. Express* 18 (10) (2010) 10762–10774.
- [19] P. Comon, Independent component analysis, a new concept? *Signal Process.* 36 (3) (1994) 287–314.
- [20] M.-Z. Poh, D.J. McDuff, R.W. Picard, Advancements in noncontact, multiparameter physiological measurements using a webcam, *IEEE Trans. Biomed. Eng.* 58 (1) (2011) 7–11.
- [21] G. de Haan, V. Jeanne, Robust pulse rate from chrominance-based RPPG, *IEEE Trans. Biomed. Eng.* 60 (10) (2013) 2878–2886.
- [22] G. de Haan, A. Van Leest, Improved motion robustness of remote-PPG by using the blood volume pulse signature, *Physiol. Meas.* 35 (9) (2014) 1913.
- [23] M. Huelsbusch, *An Image-Based Functional Method for Opto-Electronic Detection of Skin-Perfusion*, Ph.D. Thesis, RWTH Aachen DEPT. of EE, 2008 (in German).
- [24] D. McDuff, S. Gontarek, R.W. Picard, Improvements in remote cardiopulmonary measurement using a five band digital camera, *IEEE Trans. Biomed. Eng.* 61 (10) (2014) 2593–2601.
- [25] D. McDuff, S. Gontarek, R. Picard, Remote measurement of cognitive stress via heart rate variability, in: *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE, IEEE, 2014*, pp. 2957–2960.
- [26] Y. Yan, X. Ma, L. Yao, J. Ouyang, Noncontact measurement of heart rate using facial video illuminated under natural light and signal weighted analysis, *Bio-Med. Mater. Eng.* 26 (s1) (2015) 903–909.
- [27] M.A. Fischler, R.C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Commun. ACM* 24 (6) (1981) 381–395.
- [28] M.A. Hassan, A.S. Malik, N. Saad, B. Karasfi, D. Fofi, W. Sohail, Towards health monitoring in visual surveillance, in: *2016 6th International Conference on Intelligent and Advanced Systems (ICIAS)*, 2016, pp. 1–6, <http://dx.doi.org/10.1109/ICIAS.2016.7824046>.
- [29] M.A. Hassan, A.S. Malik, N. Saad, B. Karasfi, Y.S. Ali, D. Fofi, Optimal source selection for image photoplethysmography, in: *2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, 2016, pp. 1–5, <http://dx.doi.org/10.1109/I2MTC.2016.7520406>.
- [30] M. Lewandowska, J. Rumiński, T. Kocejko, et al., 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)*, IEEE, 2011, pp. 405–410.

- [31] L. Scalise, N. Bernacchia, I. Ercoli, P. Marchionni, Heart rate measurement in neonatal patients using a webcam, in: 2012 IEEE International Symposium on Medical Measurements and Applications Proceedings (MeMeA), IEEE, 2012, pp. 1–4.
- [32] L.A. Aarts, V. Jeanne, J.P. Cleary, C. Lieber, J.S. Nelson, S.B. Oetomo, W. Verkruysse, Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit – a pilot study, *Early Hum. Dev.* 89 (12) (2013) 943–948.
- [33] L.K. Mestha, S. Kyal, B. Xu, L.E. Lewis, V. Kumar, Towards continuous monitoring of pulse rate in neonatal intensive care unit with a webcam, in: Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE, IEEE, 2014, pp. 3817–3820.
- [34] Y.-P. Yu, B.-H. Kwan, C.-L. Lim, S.-L. Wong, P. Raveendran, Video-based heart rate measurement using short-time Fourier transform, in: 2013 International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS), IEEE, 2013, pp. 704–707.
- [35] X. Li, J. Chen, G. Zhao, M. Pietikainen, Remote heart rate measurement from face videos under realistic situations, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 4264–4271.
- [36] A. Lam, Y. Kuno, Robust heart rate measurement from video using select random patches, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 3640–3648.
- [37] L. Feng, L.-M. Po, X. Xu, Y. Li, R. Ma, Motion-resistant remote imaging photoplethysmography based on the optical properties of skin, *IEEE Trans. Circuits Syst. Video Technol.* 25 (5) (2015) 879–891.
- [38] A.V. Moço, S. Stuijk, G. De Haan, Ballistocardiographic artifacts in PPG imaging, *IEEE Trans. Biomed. Eng.* 63 (9) (2016) 1804–1811.
- [39] K.-Z. Lee, P.-C. Hung, L.-W. Tsai, Contact-free heart rate measurement using a camera, in: 2012 Ninth Conference on Computer and Robot Vision (CRV), IEEE, 2012, pp. 147–152.
- [40] C. Lueangwattana, T. Kondo, H. Haneishi, A comparative study of video signals for non-contact heart rate measurement, in: 2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), IEEE, 2015, pp. 1–5.
- [41] K. Mannapperuma, B.D. Holton, P.J. Lesniewski, J.C. Thomas, Performance limits of ICA-based heart rate identification techniques in imaging photoplethysmography, *Physiol. Meas.* 36 (1) (2014) 67.
- [42] E. Christinaki, G. Giannakakis, F. Chiarugi, M. Pediaditis, G. Iatraki, D. Manousos, K. Marias, M. Tsiknakis, Comparison of blind source separation algorithms for optical heart rate monitoring, in: 2014 EAI 4th International Conference on Wireless Mobile Communication and Healthcare (Mobihealth), IEEE, 2014, pp. 339–342.
- [43] W.J. Jiang, S.C. Gao, P. Wittek, L. Zhao, Real-time quantifying heart beat rate from facial video recording on a smart phone using Kalman filters, in: 2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom), IEEE, 2014, pp. 393–396.
- [44] J.R. Estepp, E.B. Blackford, C.M. Meier, Recovering pulse rate during motion artifact with a multi-imager array for non-contact imaging photoplethysmography, in: 2014 IEEE International Conference on Systems, Man and Cybernetics (SMC), IEEE, 2014, pp. 1462–1469.
- [45] T. Pursche, J. Krajewski, R. Moeller, Video-based heart rate measurement from human faces, in: 2012 IEEE International Conference on Consumer Electronics (ICCE), IEEE, 2012, pp. 544–545.
- [46] H. Monkaresi, R.A. Calvo, H. Yan, A machine learning approach to improve contactless heart rate monitoring using a webcam, *IEEE J. Biomed. Health Inform.* 18 (4) (2014) 1153–1160.
- [47] W. Wang, S. Stuijk, G. de Haan, Exploiting spatial redundancy of image sensor for motion robust rppg, *IEEE Trans. Biomed. Eng.* 62 (2) (2015) 415–425.
- [48] S. Kwon, H. Kim, K.S. Park, Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone, in: Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, IEEE, 2012, pp. 2174–2177.
- [49] M. Soleymani, J. Lichtenauer, T. Pun, M. Pantic, A multimodal database for affect recognition and implicit tagging, *IEEE Trans. Affect. Comput.* 3 (1) (2012) 42–55.