Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Towards health monitoring using remote heart rate measurement using
digital camera: A feasibility study

M.A. Hassan, A.S. Malik, D. Fofi, B. Karasfi, F. Meriaudeau

Electrical and Computer Engineering, The University of Alabama, SERC 3060, Tuscaloosa, AL 35487, USA
Centre for Intelligent Signal and Imaging Research (CISIR), Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak, Malaysia
Le2i UMR 6306, CNRS, Arts et Métiers, Univ. Bourgogne Franche-Comté 12, rue de la fonderie, 71200 Le Creusot, France

ABSTRACT

The paper presents a feasibility study for heart rate measurement using a digital camera to perform health monitoring. The feasibility study investigates the reliability of the state of the art heart rate measuring methods in realistic situations. Therefore, an experiment was designed and carried out on 45 subjects to investigate the effects caused by illumination, motion, skin tone, and distance variance. The experiment was conducted for two main scenarios; human-computer interaction scenario and health monitoring scenario. The human-computer scenario investigated the effects caused by illumination variance, motion variance, and skin tone variance. The health monitoring scenario investigates the feasibility of health monitoring at public spaces (i.e. airports, subways, malls). Five state of the art heart rate measuring methods were re-implemented and tested with the feasibility study database. The results were compared with ground truth to estimate the heart rate measurement error. The heart rate measurement error was analyzed using mean error, standard deviation; root means square error and Pearson correlation coefficient. The findings of this experiment inferred promising results for health monitoring of subjects standing at a distance of 500 cm.

1. Introduction

Heart rate is an important parameter to estimate the physiological state of an individual/person. Traditionally electrocardiogram (ECG) and pulse oximeter attached to the fingertips or earlobes were used to measure the heart rate. These methods often caused discomfort to the patient [1,2]. The ability to monitor heart rate remotely by non-contact means is a growing interest in healthcare [3,4]. The Research on remote health sensing is more focus towards personal health monitoring. The applications mainly spread through, monitoring of neonatal patients, health monitoring for ergonomics, for driving assistance, monitoring while training/exercising and monitoring during human-computer interaction [5]. The main aim is to monitor the physiological signals from patient/individual. The heart rate, respiratory rate, breathing rate, and oxygen saturation are some of the commonly analyzed physiological parameters.

The outbreak of infectious diseases at crowd gathering has been a threat for local and global health care systems. The World health organization (WHO) defines crowd gathering as “sufficient to strain the planning and response resources of the community, state or nation hosting the event” [6]. Crowd gathering is part and parcel of the very existence of our civilization. However, overcrowding brings the risk of importing and exporting or locally spreading infectious diseases. The overcrowding phenomenon is common in world’s largest and popular crowd gatherings; Hajj, FIFA world Cup, and Olympics [6].

At present infectious disease monitoring or health monitoring of crowd is performed in the form of contingency approach. Specialist utilizes technology driven modules to perform monitoring. The specialist has used technologies based on Internet-based systems, wireless sensor network, and mobile phone application and syndromic surveillance system. Fever monitoring using a thermometer or thermal camera are example of some of the common devices used to detect individuals with the infectious disease (see Fig. 1). Change in temperature of the human body is a pathological variation that occurs when the body is affected by a foreign agent (i.e. infectious disease). However, most of the infectious diseases are respiratory or cardiovascular diseases that affects the heart rate and breathing rate [7,8], such as; SARS, H1N1 influenza, swine flu, coronavirus and Ebola.
Therefore the ability to monitor the heart rate of an individual remotely would be ideal to monitor potentially infected persons from mixing with large gatherings. Recently many methods have been proposed to estimate the heart rate remotely. Kranjec et al. [9], in 2014 presented a feasibility study reporting the reliability of different remote heart rate measuring methods such as; Capacitively Coupled ECG (CCECG) [10], Microwave Distance Measurement Method (MDM) [11], and Ultrasound Distance Measurement method (UDM) [12]. The feasibility study was for a variable distance between 5 cm and 60 cm. Remote heart rate measuring methods have grown rapidly, and many methods using different mediums have emerged. The technological maturity along with the reliability of each of the methods are discussed in detail in [9,13].

RGB camera based health monitoring has become one of the interesting avenues for remote health monitoring due to its inexpensive sensor and, low computational and implementation complexity. The proficiency to perform multiple sensing using single sensor has motivated researchers to make this technology sustainable [14,1]. The application and maturity of the technology are still in progress. Digital camera based heart rate measurement method operates based on two broad principles, photoplethysmography (PPG), and ballistocardiography (BCG). The digital camera based heart rate measurement method using photoplethysmography extracts physiological signals from the heart by capturing microscopic color variations of the skin, mainly the facial area [15–17]. Ballistocardiography based methods rely on the mechanical motion of the heart. The mechanical motion contributes to a microscopic displacement of the head/face or facial skin [18,19]. Many digital camera based heart rate measuring methods have been proposed using these two principles; a comprehensive survey and review on the details of the methods can be found in [20,13].

RGB camera based heart rate measurement methods are mostly developed for human computer interaction (HCI) applications and fitness monitoring applications. Most of the existing public databases data are collected under controlled environment with simplistic scenarios which are not as challenging as in real world scenarios. Furthermore, to the best of our knowledge the application of health monitoring for a person from distance is an area that has not been investigated. Therefore, we performed this feasibility study to investigate real world factors that could challenge the reliability of RGB camera based heart rate measuring methods. This study investigates the reliability of the methods for challenges such as; illumination variance and motion variance, the effects caused by different skin tones and finally, we investigate of distances variance which would be one of the major concerns on deploying RGB camera based methods for health monitoring. The remaining of the paper is discussed as follows; Section 2 describes the methodology involved in collecting the data for the feasibility study. Section 3 provides an overview of the experiment setup and Section 4 discusses the results from the experiments, and finally, the findings of this study are concluded in Section 5.

2. Methodology of the feasibility study

This section mainly focuses on the methodology of the feasibility study experiment of heart rate extraction using a digital camera. Previous databases such as MANHOB-HCI database [21], have presented data which exhibit physical motion regarding moving the head while the body remained static. Also, previous database provided emotion variations which also exhibits motion variations only with respect to the muscles movements in the face. Most of the existing databases are compiled under controlled environment. The aim of our study is to generate heart rate measurement database under realistic environment. Our experiment involved various scenarios that were similar to real-world situations emulating a realistic situation.

The database resulting from our experiment intended to investigate the effect of four real-world situations that would help us understand improve reliability of the methods for heath monitoring. The data was collected for two scenarios mainly; human-computer interaction scenario (i.e. a person sitting in front of a computer and interacting with the computer), and health monitoring scenario. The human-computer interaction scenario investigate the feasibility of issues related to real world situations; effect of illumination variance, motion variance and the effect of different skin tone of the subject. The health monitoring scenario investigate the feasibility of heart rate measurement for subjects from various distances from the camera sensor. The health monitoring scenario was conducted by using a surveillance cameras setup.

During the experiment, 45 healthy subjects were used for data collection. The subjects were separated into three groups based on the skin tones (i.e. Fair, Brown and Black). The subject’s age varied from 21 to 63 with a mean of 29.77 years and standard deviation of 7.88 years. Two camera sensors were used to collect data in RGB color space. Camera sensor one denoted as ‘cam1’ is a Bayer mosaic 24-bit RGB camera which records data at 30 fps with a resolution of 1080 × 1920. The camera was mounted at 120 cm from the ground with 90° degree angle. Camera sensor two denoted as ‘cam2’ is also a Bayer mosaic 24-bit RGB camera which records data at 30 fps with a resolution of 1080 × 1920. The camera was mounted at 240 cm from the ground with 59° degree angle. The camera sensor information is tabulated in Table 1.

The ground truth heart rate was captured using a pulse oximeter. The WristOx2 model 3150 wrist-worn pulse oximeter by Nonin Medical [22], was used to collect the heart rate (see Fig. 2). The device was highly accurate during clinical tests and during lab tests with a MDA (Medical Device Authority) approved DATENSTT DT-HW5 ECG device. The oximeter was strapped to the wrist of the subject, and the heart rate measuring sensor was attached to the index finger of the subject. The subject was asked to wear the pulse oximeter throughout the experiment, and the heart rate was monitored individually for each session in units of beats per minute (bpm).

Table 1

| Specification                  | Cam1 | Cam2 |
|-------------------------------|------|------|
| Color filter                  | Bayer Mosaic | Bayer Mosaic |
| Color model                   | 24 bit RGB | 24 bit RGB |
| Frame Processing Speed (FPS)  | 30   | 30   |
| Resolution                    | 1080 × 1920 | 1080 × 1920 |
| Focal length                  | 3.67 mm | 3.8 mm |
| Auto focus                    | Disabled | Disabled |
| Mounting height               | 120 cm | 240 cm |
| Mounting angle                | 90°   | 59°   |
due to muscle movements within the face is called non-rigid motion variance and often caused by talking, smiling, yawning and expression of other facial emotions. Motion variance caused by body movement is due to changing the head orientation or by changing the posture of the body is called rigid motion variance. Either of these scenarios, results in changing the illumination level on the face and fluctuation of the ROI, ROI drift and loss of ROI. During this session, the data were collected from 45 subjects for two scenarios; During the first scenario, the subject is asked to be idle with minimum physical movement facing the camera. The second scenario, the subject was requested to use the computer and interact with the computer by talking and expressing emotions. During the second scenario, the subject was also asked to posture them self freely so that they emulate a realistic situation while interacting with the computer.

2.1. Human-computer interaction scenario

The human-computer interaction scenario intended to investigate the feasibility of heart rate extraction at realistic situations. During this scenario we use cam 1 to record the data at 80 cm from the subject (see Fig. 3). This experiment investigated the effects caused by illumination variance, motion variance and the effect of different skin tones of the subject.

2.1.1. Illumination variance session

The change of light intensity of the environment correlates with the quality of the video. A high light intensity would smear an image while at the same time a low light intensity would dim an image. The facility used for data collection was illuminated by fluorescent light and light from the surrounding environment. The fluorescent light consisted of a light spectrum of 400 nm to 700 nm wave length [23]. During this session, we collected data considering the effect of global illumination variance, that is about three illumination intensity levels of a room. The luminance level was measured using a luminance meter T-10 [24] the luminance level was noted as Lux and denoted as lx.

Here we estimated the heart rate for three different illumination levels ranging from 310 lx to 560 lx [25]. The lighting condition was categorized into high lighting intensity between i.e. 510 and 550 lx, medium lighting intensity 430–470 lx and low lighting intensity between 340 and 380 lx. During this session, we used 36 subjects separating 12 subjects for each of the three sessions. The subjects were asked to sit idle with minimum physical movement facing the camera since the objective of the session is to monitor the effects caused by illumination variance. The data collection involved a period of five minutes.

2.1.2. Motion variance session

Motion variance from facial images related to the muscle movements within the face and change of body posture. Motion variance

2.1.3. Skin tone variance session

Variation of skin tone could play a pivotal role on the reliability of a digital camera based heart rate measuring methods. The majority of the methods proposed are dependent on measuring the microscopic color change in the skin of the individual/subject to estimate the heart rate. The dermis consists the capillaries which circulate hemoglobin that generates the microscopic color change on the surface of the skin. This microscopic color change is transmitted as scattering of light through the epidermis which consists of melanin that induces skin pigmentation. The melanin present in the epidermis is prone to attenuate and weaken the scattered signal.

Therefore, the amount of melanin present in the skin pigmentation would relate to the conspicuous/ambiguous detection of the microscopic color change. During this session, 45 subjects were used to collect the data. The subjects were divided into three groups of 15 subjects each by visually inspecting their skin tone. The three groups pertained to three major skin tones: Fair/Pale, Brown, and Black see (Fig. 4). Since the object is to determine the effect caused by the variance of skin tone, the subjects were asked to sit idle with minimum physical movement, facing the camera.

2.2. Health monitoring scenario

The health monitoring scenario was conducted to investigate the feasibility of health monitoring for individual/crowd at public spaces. The objective of this scenario is to estimate the reliability of heart rate measurement of individual/subject standing at different distances from the camera. As the distance of the subject from the camera increases the clarity of the facial features also would reduce (see Fig. 5). Therefore, the feasibility of extracting microscopic variations of the features also would reduce, which results in an uncertainty of reliability of the heart rate measuring method. During this session, the cam2 was used to record the data of 45 individuals/subjects from three different distances. The subject was asked to ascend to three distance points from the camera and was asked to stand idle with minimum physical movement, facing the camera. The subject stood idle at three points from the camera at a distance of 500 cm on point A, a distance of 700 cm on point B and a distance of 1000 cm on point C (See Fig. 6).
3. Experiment setup

We re-implemented five state of the art digital camera based heart rate measuring methods and tested them on our experimental data. The experiment consisted of two parts; Benchmarking and Feasibility study. The sensor parameters and the experimental parameters of all the session and datasets are tabulated in Table 2 for convenient data visualization. The re-implemented methods include some of the popular methods from Poh et al. 2010 [2], Poh et al. 2011 [14], Balakrishnan et al. 2013 [18], Li et al. 2014 [26], and Feng et al. 2015 [27]. The overall experiment setup is described using a block diagram in Fig. 7.

3.1. Benchmarking

The benchmarking of the methods were performed on two data sets; 1. MANHOB-HCI database [21], and 2. Neutral session of feasibility study database. The purpose of checking the methods on MANHOB-HCI database is to demonstrate that we have correctly re-implemented the state of the art methods. Therefore, we used simplistic data of 27 subjects under a neutral condition with minimum physical movements. MANHOB-HCI consisted 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.

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 40 cm. 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 \( HR_{gt} \). The neutral session of feasibility study database also consisted of simplistic data under realistic situation for 45 subjects. Here we used the data of human-computer interaction scenario for the session where the subject stays idle with minimum physical movement, facing the camera. The purpose of checking the methods on a neutral session of feasibility study database was to investigate the reliability of the methods operating on the controlled environment to realistic environment.

3.2. Feasibility study

The feasibility study experiment was conducted to investigate the main objective of this study, that is, the feasibility of the state of the art methods to operate under realistic environments. During this experiment session, the reliability of the heart rate measuring methods for illumination variance, motion variance, and skin tone variance and distance variance are investigated. The data of the 45 subjects were separated into different groups and sessions as described in the experiment design section. The effects of motion variance sessions were expressed in comparison to the reliability of neutral session of the feasibility study database for the human

| Database                | Dataset                                  | Session                          | Camera Specification                                         | No. of subjects | Table No. |
|-------------------------|------------------------------------------|----------------------------------|-------------------------------------------------------------|-----------------|-----------|
| MANHOB-HCI              | Emotion elicitation                       | Neutral                          | Frame speed 61 fps Resolution 780 × 580 24-bit              | 27              | 3         |
| Feasibility study       | Human-computer interaction scenario       | Neutral                          | Frame speed 30 fps Resolution 1080 × 1920 24-bit            | 45              | 4         |
|                         | Illumination variance – high              |                                  | RGB distance from subject – 40 cm                           | 12              | 5         |
|                         | Illumination variance – medium            |                                  |                                                              | 12              | 6         |
|                         | Illumination variance – low               |                                  |                                                              | 12              | 7         |
|                         | Motion variance                           |                                  |                                                              | 45              | 8         |
|                         | Skin tone variance – Fair                 |                                  |                                                              | 15              | 9         |
|                         | Skin tone variance – Brown                |                                  |                                                              | 15              | 10        |
|                         | Skin tone variance Black                  |                                  |                                                              | 15              | 11        |
| Health monitoring       | Distance point A                          |                                  | Frame speed 30 fps Resolution 1080 × 1920 24-bit            | 45              | 12        |
| scenario                | Distance point B                          |                                  | RGB distance from subject – 500 cm                          | 45              | 13        |
|                         | Distance point C                          |                                  | Frame speed 30 fps Resolution 1080 × 1920 24-bit            | 45              | 14        |

Fig. 5. Description on the effect of distance for the health monitoring experiment, (A) when subject at point A, (B) when subject at point B and (C) when subject at point C.

Fig. 6. Experiment setup for Health monitoring scenario.
3.3. Data preprocessing

Data preprocessing was used to reduce the false positive/negative detection of the face detection algorithms used by the state of the art methods, mainly Viola-Jones face detector and Discriminative Response Map Fitting (DRMF) based face detector. The data collected for the feasibility study emulated a realistic situation and environment. Therefore, the face detector methods which operated reliably at controlled environment failed at realistic environments. As elaborated in Fig. 8 the diversity of the databases are shown by comparing our feasibility study database to MANHOB-HCI database. Here it is evident that the diversity of the surrounding environment plays a pivotal role towards emulating methods to operate in realistic situations.

The state of the art methods operated by detecting the face/ROI in the initial frame using face detecting methods such as; Viola-Jones face detector and Discriminative Response Map Fitting method. The Viola-Jones face detector detected an average of 3 false faces/ROI for each subject accounting to a total of 163 false positive faces (see Fig. 9(e)) throughout the 45 subjects. The Discriminative Response Map Fitting method also failed to obtain facial landmark fitting for 39 subjects (see Fig. 9(f)). The false detection disabled the heart rate measuring methods to operate under realistic situations. Therefore, a foreground segmentation based preprocessing step was included to all the heart rate measuring methods to narrow down the region of interest to the foreground of the scene (see Fig. 9(g)). The foreground detection method, was implemented by extracting the initial frame and segmenting the foreground by using active contour model [28] to extract a binary silhouette. The color image was cropped using the binary silhouette of the detected foreground object to narrow down the region of interest. The face detection algorithm was applied to the segmented image to extract the feature points. After that, the feature points were used for the face detection method that operated on the remaining of the incoming frames to reduce the face detection.

3.4. Evaluation metric

The heart rate measured by the state of the art methods were derived in terms of absolute heart rate measurement error $HR_e$. Heart rate measurement error is a function of the absolute
difference (See Eq. (1)) between the heart rate measured $HR_{\text{video}}$ and corresponding ground truth heart rate $HR_{gt}$. The ground truth was extracted from a device that is clinically approved. For MANHOB-HCI database we used the heart rate computed from the ECG signal and for the feasibility study database the heart rate was obtained from the pulse oximeter. The heart rate measurement error was validated using four statistical operators; Mean of the heart rate error $Me_{HR}$ (see Eq. (2)), standard deviation of the heart rate error $SDe_{HR}$ (see Eq. (3)), the root is mean square error $RMSE$ (see Eq. (4)) and Pearson correlation coefficient (see Eq. (5)) of the heart rate error $r$.

The results were computed from the difference between the actual heart rate and measured heart rate. The Mean heart rate error signifies the overall/average accuracy limit of the heart rate measurement model regarding bpm. The standard deviation gives an overview of the fluctuation of the heart rate measurement from the mean heart rate error regarding bpm. The $RMSE$ measures the standard deviation of the residual (i.e. the spread of the points about the fitted regression line) regarding bpm. Therefore, providing an estimate of the noise that is generated by the methods. Finally, the Pearson correlation coefficient provides the linear correlation between the actual heart rate and measured heart rate by a measure of 0–1, where 1 been the highest correlation.

$HR_e(i) = |HR_{\text{video}}(i) - HR_{gt}(i)|$  \hfill (1)

$Me_{HR} = \frac{1}{N} \sum_{i=1}^{N} HR_e(i)$  \hfill (2)

$SDe_{HR} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (HR_e(i) - Me_{HR})^2}$  \hfill (3)

$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (HR_e(i))^2}{N}}$  \hfill (4)

$r = \frac{N \sum HR_{\text{video}}HR_{gt} - \sum HR_{\text{video}} \sum HR_{gt}}{\sqrt{N \sum HR_{\text{video}}^2 - (\sum HR_{\text{video}})^2} \cdot \sqrt{N \sum HR_{gt}^2 - (\sum HR_{gt})^2}}$  \hfill (5)

where $N$ is the total number of samples and $i$ is the sample at an instance. $\sum HR_{\text{video}}$ is the sum of the measured heart rate and $\sum HR_{gt}$ is the sum of the actual heart rate and $\sum HR_{\text{video}}HR_{gt}$ is the sum of the products of the paired scores.

### 4. Results and discussion

Reliability is an important aspect of the development of a health monitoring method. Heart rate measurement error within five bpm is an acceptable error margin [29]. Feasibility of a method to operate at realistic situations would enable the method to be sustainable. The experiment was conducted to investigate the effect of four real-world situations: effect of illumination variance, motion variance, and the effect of different skin tones of the subject and finally to verify the feasibility to perform health monitoring for surveillance purpose. The MANHOB-HCI database was used to benchmark the methods and to demonstrate that the methods were correctly re-implemented by us. The neutral session of the feasibility study database was used to investigate the reliability of the methods, operating from a controlled environment to realistic environment.

#### 4.1. Benchmarking

The benchmarking of the methods using MANHOB-HCI database shows that the results reported in our experiment (see Table 3) mainly the RMSE were similar to the results reported in [26,30,31], therefore, demonstrating that we have correctly re-implemented the state of the art method. Measured heart rate of each of the methods was validated with the actual heart rate. The measured heart rate was estimated by extracting the power spectral density (PSD) of the PPG or BCG signal. The first harmonic of the PSD was extracted as the heart rate frequency. The heart rate frequency was multiplied by 60 to estimate the heart rate for beats per minute (bpm).

Feng et al. [27] reported the lowest error by a mean error of 6.64 bpm with a standard deviation of 8.01 bpm. Feng et al. [27], also reported the highest correlation coefficient between actual heart rate and the measured heart rate. Balakrishnan et al. [18], reported the highest error rate by a mean error of 17 bpm with a standard deviation of 14.66 bpm. Poh et al. [14] heart rate measuring method based on BSS did not perform well as to the other PPG methods [26,27]. However, [14], reported a correlation coefficient of 0.44, while conceding a high mean error. The correlation coefficient shows that [14] extracted an accurate heart rate for 44% of the experiments while conceding higher error rate for the remaining experiments.

The accuracy of the heart rate measuring method reduced under the realistic environment (see Table 4) of the feasibility study database. Feng et al. [27], proposed method using adaptive green, red differentiation approach reported a mean error of 10.14 bpm and standard deviation of 11.51 bpm. The method also reported person correlation coefficient of 0.41, which shows that the 41% of the measurements were accurate. Li et al. [26] conceded a higher error rate during the feasibility study experiment. The illumination compensation method by using adaptive filter method based on the illumination of the background could be the cause of the high error rate.

The background of the controlled environment used by the MAHNB database varies with the backdrop of the realistic environment. Therefore the assumption made by [26] for illumination compensation may not be valid for a realistic environment. The performance of the PPG methods reduced at the realistic environment. The approach of extracting the spatial mean from the ROI to generate the raw PPG signal could have contributed to the change.

![Fig. 9. Illustration on ROI detection before and after preprocessing, (a) & (d) are the original images (g) is the preprocessed image, (b) & (c) are the face detection results of Viola-Jones & DRMF face detectors for HCI database. (e) & (f) are the face detection results of Viola-Jones & DRMF face detectors for Feasibility database. (h) & (i) are the face detection results of Viola-Jones & DRMF face detectors after preprocessing.](image-url)
in accuracy. Under the realistic situation, illumination is not controlled. Therefore, it can be seen that there was a variance of the illumination intensity present at different parts of the face/ROI (see Fig. 10(a)), compared to the illumination intensity distribution of the controlled environment (see Fig. 10(b)).

The BCG based approach of Balakrishnan et al. [18], method based on motion features performed better at the feasibility study database. The improvement in the correlation coefficient indicates that [18] method performed better under realistic conditions. The usage of motion feature compared to the gradient features may have improved the accuracy under illumination varying conditions. However [18] reported a high error rate. The PCA based blind source separation of the tracked features of the KLT may have contributed to the uncertainty of the accuracy. Since [18] extracts the BCG signal from the first five principal components and the prospective BCG signal may result in any principle component of the number of tracked features.

### 4.2. Illumination variance

The results of the illumination variance session showed that the PPG-based heart rate measuring method performed better under high lighting intensity (i.e. 510–550 lx) and the performance reduced as the lighting intensity reduced. Feng et al. [27] performed well under high lighting intensity. The adaptive green and red differencing method reported the lowest mean error of 9.61 bpm (see Table 5) with a standard deviation of 8.65 bpm. Whereas, the error rate increased as the lighting intensity decreased. The method reported a mean error of 10.86 with a standard deviation of 11.59 (see Table 6) for the medium light intensity (i.e. 430–470 lx), increasing the mean error by ≈ 1 bpm with a standard deviation of ≈ 4 bpm.

For the low lighting (i.e. 340–380 lx) session, the method reported the highest error rate by a mean error of 13.13 bpm with a standard deviation of 12.74 bpm (see Table 7). The results showed an increment of the error rate by a mean error of ≈ 3 bpm with a standard deviation of ≈ 3bpm, compared to the error rate reported in the high lighting intensity session. As mentioned earlier the method [18] suffers in the BCG signal extraction approach after BSS. However, the Pearson correlation coefficient shows that the motion feature based method [18] did not exhibit any major change in accuracy during the global illumination variance experiment.

### 4.3. Motion variance

The motion variance experiment was highly challenging for the ROI tracking of the heart rate measuring methods. The free rigid and non-rigid motion of the person increased the error rate of the heart rate measurement. The adaptive green and red differentiation approach to compensate motion variance performed better compared to other PPG methods (see Table 8). However, the overall accuracy of the heart rate measurement did reduce as indicated by the Pearson correlation coefficients of Table 4 and Table 8. The BSS based methods [2,14] suffered largely due to failure in ROI tracking. The Viola-Jones face detector based face tracking method failed to track the face at 17 occasions. Similarly the green spectrum based PPG method [26] also failed to track the ROI at nine sessions.

The motion feature based method of [18] failed the most by not being able to detect the ROI at 23 sessions. The ROI loss was mainly due to rigid motion variance caused by posture variation. The motion feature based method mainly failed due to loss of feature points. The rise of the error rate for Feng et al. [27], and Li et al. [26], was due to ROI drift (See Fig. 11). These methods used KLT feature tracking algorithm to track the detected ROI, however as the subject moved the ROI shift and the tracking algorithm failed to track the ROI to the actual location, causing ROI drift.

![Fig. 10. Illustration on the distribution on the illumination intensity values across the face, (a) MANHOB-HCI database & (b) Feasibility study database.](image-url)
Also, the results showed that Li et al. [26] performed better under motion variance session compare to the neutral session by reporting a lower error rate. The temporal pruning based motion artifact removal method operated efficiently during motion variance. However, the uncertainty in the accuracy may have raised due to the PPG signal reconstruction approach used after the temporal pruning function. In general, the high error rate during the motion variance experiment was due to the ROI loss and ROI drift issues caused by poor ROI tracking.

4.4. Skin tone variance

The effect of skin tone variance may have a direct relationship to the nature of the feature that is used to compute the PPG or BCG signals. The adaptive green and red differencing approach of [27], performed similarly to fair (see Table 9) and dark skin tones (see Table 11). The Error rate and the Pearson correlation coefficients did not vary largely for different skin tones. However the green spectrum based PPG method [26] showed an increment in error rate with a reducing Pearson correlation coefficient for the skin tones from fair to dark. Also, it was noticeable that Poh et al. [14], showed a higher correlation between the actual heart rate and the measured heart rate for fair skin (0.48), brown skin (0.26) and black skin (0.31).

The BSS based approach [14] using all three RGB spectral information was better compared to the signal spectral method of [26] while measuring PPG signal from different skin tones. The change in skin tone surprisingly affects the motion feature based BCG method [18]. The BCG based method reported a Pearson correlation coefficient of 0.22 for fair skin, 0.25 for brown skin and 0.18 for dark skin. The usage of KLT to track the feature points may have resulted in the change in accuracy for dark skin tone since KLT tracks the motion of feature points based on the pixel intensity variation of the neighboring pixels of the feature point.

4.5. Distance variance

The feasibility of health monitoring experiment for health monitoring was performed by measuring the heart rate, while the subject stood at different distances from the camera. The results of this experiment showed that the reliability of the heart rate measuring methods reduced as the distance of the subject from the camera increased. The lowest error rate was measured by Feng et al. [27], the method reported a mean error of 15.32 bpm with a standard deviation of 10.11 for point A (i.e. 500 cm) (see Table 12) and mean error of 18.23 bpm with a standard deviation of 14.24 (see Table 13) for point B (i.e. 700 cm) and mean error of 24.06 bpm.
with a standard deviation of 12.84 (see Table 14) for point C (i.e. 1000 cm).

Similarly, the Pearson correlation coefficient also reduced from 0.39 for point A (i.e. 500 cm) (see Table 12) and 0.12 (see Table 13) for point B (i.e. 700 cm) and 0.09 (see Table 14) for point C (i.e. 1000 cm). The drop in the correlation coefficient can relative to average face size of the subjects varied from 25/C2 for point A, 25/C2 for point B (see Fig. 12) and 20/C2 for point C (see Fig. 12(c)). The methods were also affected significantly by the issues of poor ROI tracking.

The methods failed to detect the ROI at many sessions causing ROI loss, and ROI drift. The number of ROI loss and ROI drift for each session is tabulated in Table 15. As the distance from the camera increased the credibility of the face detection methods reduced

Table 9
Performance validation of the methods for the skin tone variance (fair) session of Feasibility study database.

| Method          | \( M_{eHR} \) (bpm) | SD_{eHR} (bpm) | RMSE (bpm) | r  |
|-----------------|---------------------|----------------|-------------|----|
| Poh et al. 2010 [2] | 72.83              | 54.32          | 89.69       | 0.11 |
| Poh et al. 2011 [14] | 59.13              | 54.02          | 78.78       | 0.48 |
| Balakrishnan et al. 2013 [18] | 92.47              | 56.38          | 107.25      | 0.02 |
| Li et al. 2014 [26] | 60.96              | 52.40          | 79.16       | 0.29 |
| Feng et al. 2015 [27] | 14.37              | 10.80          | 17.75       | 0.49 |

Table 10
Performance validation of the methods for the skin tone variance (brown) session of Feasibility study database.

| Method          | \( M_{eHR} \) (bpm) | SD_{eHR} (bpm) | RMSE (bpm) | r  |
|-----------------|---------------------|----------------|-------------|----|
| Poh et al. 2010 [2] | 29.45              | 25.72          | 37.99       | 0.14 |
| Poh et al. 2011 [14] | 24.18              | 28.63          | 36.29       | 0.26 |
| Balakrishnan et al. 2013 [18] | 36.76              | 31.45          | 54.14       | 0.25 |
| Li et al. 2014 [26] | 26.80              | 22.36          | 34.33       | 0.24 |
| Feng et al. 2015 [27] | 11.00              | 13.65          | 17.10       | 0.41 |

Table 11
Performance validation of the methods for the skin tone variance (black) session of Feasibility study database.

| Method          | \( M_{eHR} \) (bpm) | SD_{eHR} (bpm) | RMSE (bpm) | r  |
|-----------------|---------------------|----------------|-------------|----|
| Poh et al. 2010 [2] | 26.86              | 20.96          | 36.50       | 0.17 |
| Poh et al. 2011 [14] | 21.32              | 20.33          | 32.65       | 0.31 |
| Balakrishnan et al. 2013 [18] | 32.92              | 25.65          | 48.71       | 0.18 |
| Li et al. 2014 [26] | 25.28              | 27.51          | 37.36       | 0.19 |
| Feng et al. 2015 [27] | 12.95              | 9.37           | 16.49       | 0.44 |

Table 12
Performance validation of the methods for the distance variance (point A) session of Feasibility study database.

| Method          | \( M_{eHR} \) (bpm) | SD_{eHR} (bpm) | RMSE (bpm) | r  |
|-----------------|---------------------|----------------|-------------|----|
| Poh et al. 2010 [2] | 29.83              | 20.29          | 38.33       | 0.16 |
| Poh et al. 2011 [14] | 31.27              | 19.81          | 35.90       | 0.22 |
| Balakrishnan et al. 2013 [18] | 32.11              | 28.11          | 53.23       | 0.18 |
| Li et al. 2014 [26] | 27.71              | 22.32          | 39.95       | 0.24 |
| Feng et al. 2015 [27] | 15.32              | 10.11          | 19.51       | 0.39 |

Table 13
Performance validation of the methods for the distance variance (point B) session of Feasibility study database.

| Method          | \( M_{eHR} \) (bpm) | SD_{eHR} (bpm) | RMSE (bpm) | r  |
|-----------------|---------------------|----------------|-------------|----|
| Poh et al. 2010 [2] | 29.25              | 29.23          | 50.72       | 0.07 |
| Poh et al. 2011 [14] | 28.78              | 24.21          | 49.11       | 0.09 |
| Balakrishnan et al. 2013 [18] | 33.47              | 27.40          | 54.97       | 0.04 |
| Li et al. 2014 [26] | 30.64              | 23.25          | 46.70       | 0.09 |
| Feng et al. 2015 [27] | 18.23              | 14.24          | 23.14       | 0.12 |

Table 14
Performance validation of the methods for the distance variance (point C) session of Feasibility study database.

| Method          | \( M_{eHR} \) (bpm) | SD_{eHR} (bpm) | RMSE (bpm) | r  |
|-----------------|---------------------|----------------|-------------|----|
| Poh et al. 2010 [2] | 35.41              | 29.11          | 59.25       | 0.05 |
| Poh et al. 2011 [14] | 32.32              | 28.55          | 53.07       | 0.09 |
| Balakrishnan et al. 2013 [18] | 40.19              | 26.39          | 58.81       | 0.02 |
| Li et al. 2014 [26] | 32.19              | 24.74          | 44.67       | 0.09 |
| Feng et al. 2015 [27] | 24.06              | 12.84          | 27.19       | 0.09 |
in detecting facial features, therefore, causing ROI loss and ROI drift. Also, the distance, reduce the clarity of the pixel values, therefore, baffling the microscopic color changes that amount to the PPG signal generation. Similarly the motion feature based method [18], also conceded a high error rate and reported a low Pearson correlation coefficient. The BCG based approach could have also been affected by the size of the ROI. Since the KLT feature tracking algorithm would not be able to track the microscopic motion of the face while object is standing at a distance (a) detected face at point A, (b) detected face at point B & (c) detected face at point C.

5. Conclusion

This paper presented a feasibility study on health monitoring using a digital camera. The study was mainly carried out to understand the effects of realistic conditions for the application of health monitoring which deals with health monitoring at public spaces (i.e. airports, subways, malls). The experiment was conducted on 45 subjects to assess four issues that could affect the reliability of the heart rate measuring methods at realistic situations. The feasibility was evaluated for illumination variance, motion variance, and skin tone variance and distance variance. The results were analyzed and discussed for each experiment. Five heart rate measuring methods based on five different feature selection methods were re-implemented and tested for these scenarios. We tested the methods on the neutral session of the MANHOB-HCI database to verify that the state of the art methods were re-implemented correctly.

The methods were then tested on the data of the feasibility study. The results of the first experiment (i.e. neutral session), showed that the reliability of the methods was highly affected by the diversity of the realistic environment. A preprocessing foreground selection step was included to all the heart rate measuring methods to narrow the ROI to operate at the realistic environment. The reliability was mainly affected by motion artifact caused by the spatial illumination variance of the face. The experiments of the illumination variance session showed that the reliability of heart rate measuring methods based on PPG is directly proportional to the illumination level of the room. However, the motion feature based BCG method showed that the BCG based approach is invariant to illumination variance and can measure heart rate under low illumination levels.

The motion variance session was mainly affected by ROI loss and ROI drift, the face detection methods used by the heart rate measuring methods failed to detect the face during rigid motion variation and failed to track the face consistently while the body was moving. The effect of skin tone variance was significant for the PPG-based method. The BSS approach of using RGB spectrum was able to perform better in measuring the heart rate from all three skin tones. However, the results derived from the green spectrum based method did not perform well as the other PPG method. Therefore the idea of using a function of two spectra or all three spectra may be useful in measuring the heart rate across different skin tones.

The results of the health monitoring experiment showed that the reliability of the heart rate is inversely proportional to the distance of the subject from the camera. As the subject was further apart from the camera; the size of the ROI reduced greatly. This affected the heart rate measuring methods since most of the methods generated the raw signal by estimating the spatial mean of the ROI. Thus smaller ROI leading to poor spatial mean estimation. Also the methods were affected by the clarity of the facial features of the face while standing at a distance. This is mainly due to the camera hardware specifications, specifically the focal length. The methods were not able to operate at many sessions due to ROI loss and ROI drift issues. However, the heart rate measurement by Feng et al. [27], at 500 cm point was promising. The re-implemented methods that were tested in this study were developed for controlled environments. By incorporating the ideas and the feature selection methods used by each of the methods; one can further improve the reliability of heart rate measurement for uncontrolled realistic environments.

In conclusion, the computer vision aspect of selection of ROI, ROI detection, ROI loss and ROI drift and rigid motion variation and spatial illumination variance appeared to be the main problem for heart rate measuring methods to operate at the realistic environment. Future research should be conducted to address these issue. An interesting direction would be to incorporate advanced computer vision and deep learning based method to perform face/landmark detection and tracking in realistic environments. Remote health monitoring using RGB camera would provide a cheap yet highly efficient solution for health monitoring of crowd. One could perform health monitoring at bottleneck locations of public spaces (e.g. exit and entrance of train stations, malls, and airport) or by using multiple RGB cameras at checkpoints (e.g. immigration checkpoint or ticket counters) to perform individual health monitoring.

The main idea of this study is to understand the effect of the realistic situation while measuring heart rate remotely. The heart rate measurement is performed by remotely extracting the physiological signals (BCG/PPG) from the human body. Understanding heart rate variability of these physiological signals can further derive respiratory and autonomic nervous system functions of the human body. Therefore the application of RGB camera based health monitoring would be encouraging as, developing a remote health monitoring method to monitor the physiological state of people in public spaces would prevent fatalities, the spread of communicable disease.
Acknowledgment

This experiment is supported by the graduate student scheme of University Teknologi PETRONAS and supported by the ministry of education malaysia under the higher institutional centre of excellence (HICoE) award to Centre for Intelligent Signal and Imaging Research Centre (CISIR).

References

[1] M. Hassan, A.S. Malik, N. Saad, B. Karasfi, D. Fofi, W. Sohail, Towards health monitoring in visual surveillance, in: Intelligent and Advanced Systems (I-CASIS), 2016 6th International Conference on, IEEE, 2016.
[2] M.-Z. Poh, D.J. McDuff, R.W. Picard, Non-contact automated cardiac pulse measurements using video imaging and blind source separation, Optics Exp. 18 (10) (2010) 10762–10774.
[3] A. Bauer, M. Malik, G. Schmidt, P. Barthel, H. Bonnemeier, I. Cygankiewicz, P. Guzik, F. Lombardi, A. Müller, A. Otto, 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.
[4] Q. Zhang, Q. Wu, Y. Zhou, X. Wu, Y. Ou, H. Zhou, Webcam-based, non-contact, real-time measurement for the physiological parameters of drivers, Measurement 100 (2017) 311–321, https://doi.org/10.1016/j.measurement.2017.01.007, URL: www.sciencedirect.com/science/article/pii/S0261926816300076.
[5] K.-Y. Lin, D.-Y. Chen, W.-J. Tsai, Face-based heart rate signal decomposition and evaluation using multiple linear regression, IEEE Sens. J. 16 (5) (2016) 1351–1360, cited by 0. doi:10.1109/JSEN.2015.2500032.
[6] E.O. Nsoesie, S.A. Kluberg, S.R. Mekaru, M.S. Majumder, K. Khan, S.I. Hay, J.S. Brownstein, New digital technologies for the surveillance of infectious diseases at mass gathering events, Clin. Microbiol. Infection 21 (2) (2015) 134–140.
[7] J.A. Hirsch, B. Bishop, Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate, Am. J. Physiol. – Heart Circulatory Physiol. 241 (4) (1981) H620–H629, arXiv: http://ajpheart.physiology.org/content/241/4/H620.full.pdf.
[8] S. Benkouiten, R. Charrel, K. Belhouchat, T. Drali, N. Salez, A. Nougaiere, C. Zandotti, Z.A. Memish, M. al Masri, C. Gaillard, C. Gaillard, et al., Circulation of respiratory viruses among pilgrims during the 2012 hajj pilgrimage, Clinical Infect. Dis. 57 (7) (2013) 992–1000.
[9] J. Kranjec, S. Begus, J. Drnovsek, G. Gersak, Novel methods for noncontact heart rate measurement: a feasibility study, IEEE Trans. Instrum. Meas. 63 (4) (2014) 838–847.
[10] A.E. Mahdi, L. Faggion, Non-contact biopotential sensor for remote human detection, Journal of Physics: Conference Series, vol. 307, IOP Publishing, 2011, p. 012056.
[11] G. Lu, F. Yang, Y. Tian, X. Jing, J. Wang, Contact-free measurement of heart rate variability via a microwave sensor, Sensors (Basel, Switzerland) 9 (12) (2009) 9572–9581, https://doi.org/10.3390/s91209572.
[12] D. Oberd, S. Sadek, G. Zaharia, G. El Zen, Multitunable microwave system for touchless heartbeat detection and heart rate variability extraction, Microwave Opt. Technol. Lett. 52 (1) (2010) 192–198.
[13] M. Hassan, A. Malik, D. Fofi, N. Saad, B. Karasfi, Y. Ali, F. Meriaudeau, Heart rate estimation using facial video: a review, Biomed. Signal Process. Control 38 (2017) 346–360.
[14] 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.
[15] W. Verkruysse, L.O. Svaasand, J.S. Nelson, Remote plethysmographic imaging using ambient light, Optics Exp. 16 (26) (2008) 21434–21445.
[16] W. Wang, B. den Brinker, S. Stuijk, G. de Haan, Algorithmic principles of remote-ppg, IEEE Trans. Biomed. Eng.
[17] M. Hassan, A. Malik, D. Fofi, N. Saad, F. Meriaudeau, Novel health monitoring method an using RGB camera, Biomed. Opt. Exp. 8 (11) (2017) 4838–4854.
[18] G. Balakrishnan, F. Durand, J. Guttig, Detecting pulse from head motions in video, in: proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 3430–3437.
[19] M.A. Hassan, A.S. Malik, D. Fofi, N.M. Saad, Y.S. Ali, F. Meriaudeau, Video-based heartbeat rate measuring method using ballistocardiography, IEEE Sens. J. 17 (14) (2017) 4544–4557.
[20] D.J. McDuff, J.R. Estrepp, A.M. Piasecki, E.B. Blackford, A survey of remote optical photoplethysmography imaging methods, in: Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE, IEEE, 2015, pp. 6398–6404.
[21] M.P. Tarvainen, P.O. Ranta-Aho, P.A. Karjalainen, An advanced detrending method with application to HRV analysis, IEEE Trans. Biomed. Eng. 49 (2) (2002) 172–175.
[22] Nonin, WristO2 (jan 2016). http://www.nonin.com/OEMSolutions/WrisO23150-OEM.
[23] R.J. Bielski, J. Mayor, J. Rice, Phototherapy with broad spectrum white fluorescent light: a comparative study, Psychiatry Res. 43 (2) (1992) 167–175.
[24] K. Minolta, T-10ma illuminance meters (jan 2016). http://www.sensing.konicaminolta.us/products/t-10a-illuminance-meter/.
[25] A. Pears, Strategic study of household energy and greenhouse issues a report for environment australi, Strategic study, Sustainable Solutions Pty Ltd, 2/78 William Street Brighton 3186 (jun 1998).
[26] X. Li, J. Chen, C. 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. 4246–4251.
[27] 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.
[28] T.F. Chan, L.A. Vese, Active contours without edges, IEEE Trans. Image Process. 10 (2) (2001) 266–277.
[29] M. Cecconi, J. Favrett, R.M. Grounds, A. Rhodes, A prospective study to evaluate the accuracy of pulse power analysis to monitor cardiac output in critically ill patients, BMC Anesthesiology 8 (1) (2008) 1.
[30] S. Tufikakov, X. Alameda-Pineda, E. Ricci, L. Yin, J.F. Cohn, N. Sebe, Self-adaptive matrix completion for heart rate estimation from face videos under realistic conditions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2396–2404.
[31] 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.