ISSN 2581 – 3781

AJCSE
ASIAN JOURNAL OF COMPUTER SCIENCE ENGINEERING

Available Online at www.ajcse.info
Asian Journal of Computer Science Engineering 2021;6(1):1-15

REVIEW ARTICLE

A Survey on Non-Contact Heart Rate Estimation from Facial Video

Duaa H. Ali, Mazin H. Aziz

Department of Computer Engineering, University of Mosul, Mosul, Iraq

Received on: 24-02-2021; Revised on: 25-02-2021; Acceptance on: 10-03-2021

ABSTRACT

One of the most important health parameters and a critical indicator of the physiological condition of people is the heart rate (HR); thus, it is important to monitor. HR estimation from facial videos is a low-cost, simple, and comfortable method; therefore, it has enjoyed great interest in recent years. In this paper, we give a review of some public datasets used in evaluating non-contact HR estimation methods, the parameters of HR estimating from facial video system with a set of recent studies that dealt with them, review some of recent related reviews and surveys, and finally, the applications of non-contact HR monitoring system by a camera. This review shows that the non-contact photoplethysmography method attracted more interest since 2013. We can deduce that there was no general-purpose system that can fit all types of HR-monitoring, instead, the designer can first select the usage environment, then sets the suitable system parameters that are most fulfills the application demands.

Key words: Facial video, heart rate, non-contact, photoplethysmography, regions of interest

INTRODUCTION

While heart rate (HR) can be monitored through some body attached devices,[1] many studies have dealt with methods for estimating HR from facial videos in recent years; however, their accuracy and computational complexity have not yet reached the desired standards due to the presence of many challenges such as spontaneous facial expressions, movement, and changing illumination.

Extracting HR from facial videos is mainly divided into two categories: Photoplethysmography (PPG) and ballistocardiography (BCG). PPG means the extraction of the HR from skin color variation,[2] while BCG means the extraction of the HR from head motion.[3,4] Both the skin color variation and the head motion are due to the internal functioning of the heart. HR extraction based on PPG gained more research interest than BCG's HR extraction. The early methods succeeded in estimating HR under controlled conditions that required constant and appropriate illumination, the person remaining steady without movement, and with minimum facial expressions. Hence, it is not suitable for real-world applications such as drivers' monitoring or during exercise. On the other hand, there were successful studies in estimating HR under realistic conditions and achieved promising results,[5,6] but using more complex ways than the early methods.

The basic system is composed of an illuminator, one camera or more, and a laptop as a processing unit with a display as shown in Figure 1a. The illuminated person's face-video is acquired by the camera and send to the processing unit to extract the PPG signal and then estimate the HR. The video is processed in three main stages to extract the HR: Face detection and tracking, PPG signal extraction and processing, and HR estimation as shown in Figure 1b. In the first stage, various algorithms such as Haar cascade (Viola and Jones method),[7,8] discriminative response map fitting (DRMF),[9] and Kanade-Lucas-Tomasi (KLT)[10] were used for face detection and tracking. In the second stage, the raw-signal is extracted from the regions of interest (ROIs) and then processed using different methods and filters to recover the PPG signal. The final stage job is the HR estimation using peak-detection or power spectral density (PSD).

Some studies were used other body parts such as a finger instead of the face to find the PPG.[11,12]

Researchers may use their local videos or public dataset videos, or both of them for a system test. In

Address for correspondence:
Duaa H. Ali,
E-mail: duaa.hussein.ali94@gmail.com

both cases, there must be a ground-truth source for results' comparison, where some traditional medical devices are used to measure the HR throughout the experiments such as an electrocardiogram (ECG) and a pulse oximetry device.

To characterize the extracted PPG signal from a video, there are several used terminologies. The most common are: Remote PPG (rPPG),[13-16] imaging PPG (iPPG),[17] and PPG imaging (PPGi/PPGI).[18] In this paper, the term iPPG will be used for non-contact methods.

The researches and studies covered in this paper were done in the interval between 1967 and 2020, as shown in Figure 2.

PUBLIC DATASETS FOR NON-CONTACT HR ESTIMATION

The available public databases for evaluating non-contact HR estimation methods still limited. These datasets are used, but not limited to, in the evaluation of the iPPG signal taken from the facial video. Here are some of the most used dataset.

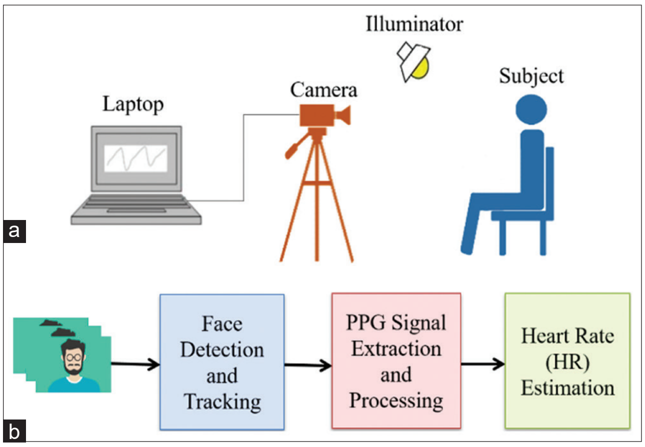

Figure 1: (a) General block diagram for non-contact heart rate (HR) estimation system. (b) The processing stages for HR extraction

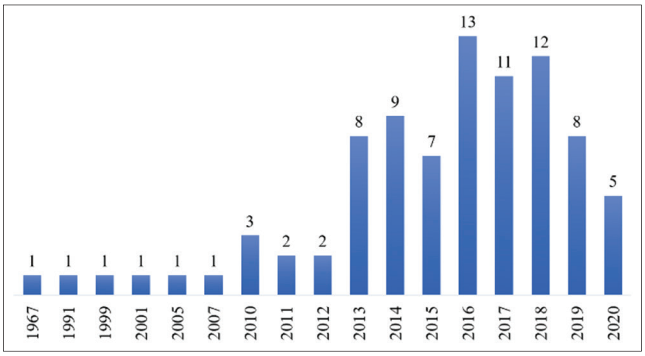

Figure 2: Histogram for the covered studies in this paper per year

MAHNOB-HCI dataset

MAHNOB-HCI dataset[19] is an emotional dataset and the most used in the researches,[5,6,20-22] which comprises 527 videos for 27 subjects (11 males and 16 females), recorded under laboratory illumination with limited head movement and facial expression. The videos were recorded using six cameras with 60 frames per second (fps).

MMSE-HR dataset

MMSE-HR dataset[6] is a part of the MMSE dataset.[23] MMSE-HR dataset is also an emotional dataset, comprising 102 videos for 40 subjects, recorded under similar circumstances to MAHNOB-HCI dataset.

The ground-truth HR calculations of both MAHNOB-HCI and MMSE-HR databases can be found using the OSET ECG Toolbox from ECG signals included within the datasets.

VIPL-HR dataset

VIPL-HR dataset[20] is a multi-modality database that contains 752 near-infrared videos and 2378 visible light videos of 107 subjects. Nine different factors were taken into account, including different head movements and lighting conditions. Both videos were captured using the Logitech C310, RealSense F200, and the front camera of the HUAWEI P9 smartphone, and aCONTEC CMS60C blood volume pulse (BVP) sensor was used to capture the ground-truth HR.

PURE dataset

PURE database[24] consists of 60 video clips acquired for ten persons (eight males and two females) captured by an eco274CVGE camera. The illumination source was the daylight that was allowed to enter in various quantities through a large window. Different head motions were applied, which were steady, talking, slow translation, fast translation, small rotation (about 20°), and medium rotation (about 35°).

Pulse from face (PFF) dataset

PFF dataset[21] was made up of 59 video clips for 13 individuals captured by a Nikon D5300 camera. Videos were recorded under dim and variation illumination conditions, for different body movements.

BioVid heat pain dataset

BioVid heat pain database[25,26] consists of physiological data and videos for 90 persons. The videos were recorded with 25 fps without imposing restrictions on the movement of the participants, as the videos included facial expressions and head movements.

DEAP dataset

DEAP database[27] is an emotional database in terms of stages of valence, arousal, like/dislike, familiarity, and dominance. It consists of physiological data for 32 participants, but the recorded facial videos are only for 22 persons. The front facial videos were recorded by a SONY DCR-HC27E camera with 50 fps. The videos were recorded in relatively dark surroundings with a variation illumination on subjects' faces, as they watched music videos during the experiment. Subjects did not ask to remain still; they could move their heads naturally, in addition to facial expressions.

UBFC-rPPG dataset

UBFC-rPPG dataset[28] is a public dataset for analysis of iPPG signal, which consists of 43 videos recorded by a webcam (Logitech C920 HD pro) with 30 fps. These videos were taken for the persons while playing a mathematical game.

SYSTEM PARAMETERS

There are different features for each of the reviewed works and can be distinguished through some of the design parameters. This section explains these system parameters. Table 1 gives a comparison summary for the covered studies regarding these parameters.

Video Camera

The choice of the camera for filming facial video clips depends on the system used to estimate the HR; some systems used visible light while others used infrared light. It is possible to use more than one camera of the same type or of different types at the same time. Red Green Blue (RGB), IP, infrared, and smartphone cameras were all used in the previous works. The most typically used in the researches was the webcam because it has low cost and available at the hand.

Webcams of different types were used for facial video by several researches.[5,29-33]

Other types of cameras were utilized such as RGB charged coupled device camera of type USB UI-2230SE-C of IDS Gmbh,[34] Panasonic Lumix GF2 camera,[35] Sony camera type HDRPJ260VE,[36] Sony camera type DCR-HC27E,[37] Nikon D610 camera,[38] and Nikon D5300 camera.[21]

FLIR Cricket IP camera[39] and Microsoft Kinect-IR camera[40] were employed also.

A single smartphone camera,[41] four identical Logitech C920 webcam camera,[42] and three different types; webcam, smartphone, and RealSense cameras in the same system[20] were used in other researches.

Color spaces and channels

Most of the videos that were used for HR estimation are standard videos,[43] while some of them are near infrared videos.[40] Many HR estimation methods have used the iPPG signal in RGB color space,[5,6,18,21,22,29-31,33,34,37,38,42,44,45] on the other hand, other methods were used alternative color spaces.[20,32,41] There were researches that have used more than one color space.[35,36]

In the studies,[6,18,30,31,33,34,37,42] all the RGB's channels were used to estimate HR, while only red and green were analyzed by another research.[22] The red channel alone was used sometimes,[29] but the green channel gained more interest.[5,21,38,44,45]

Channels of other color spaces were adopted such as the Hue channel of the hue saturation value color space,[41] the Q channel of the in-phase quadrature (YIQ) color space,[32] all the channels (Y, U, and V) of the YUV color space,[20] and the Y channel of the YCbCr color space.[39]

Multi-color-spaces may be utilized such as the use of all the channels of the RGB color space in addition to the L channel of the hue saturation lightness color space,[35] or full channels of the RGB and the YCbCr color spaces.[36]

Niu *et al*.,[20] were proved in their study that using different color spaces gives different accuracy

Table 1: A brief summary of the noncontact heart rate estimation studies

Paper	Year	Camera	Color space	Channels	Illumination	Motion	Methods	Database and results (RMSE in bpm)
[31]	2010	Webcam (iSight camera)	RGB	R + G + B	V	S	ICA	Local (1.24)
[34]	2013	RGB CCD (USB UI-2230SE-C of IDS Gmbh)	RGB	R + G + B	F	S	RoverG, XoverY, Fixed, XsminαYs	Local (0.5/0.4/0.5/0.5)
[35]	2013	Panasonic Lumix GF2	RGB + HSL	R + G + B + L	V	S	Feature tracking + PCA	Local
[5]	2014	Webcam (iSight IPAD camera)	RGB	G	V F	E/SM S	FFT	MAHNOB-HCI/VideoHR (7.62/1.27)
[29]	2014	Webcam	RGB	R	F	-	EVM	Local
[36]	2015	Sony (HDRPJ260VE)	RGB + YCbCr	R + G + B + Y + Cb + Cr	F	S	PCA	Local
[44]	2015	-	RGB	G	V	E/SM	ICA	MAHNOB-HCI (8.9)
[6]	2016	-	RGB	R + G + B	V -	E/SM E/SM	SAMC	MAHNOB-HCI/MMSE-HR (6.23/11.37)
[30]	2016	Webcam	RGB	R + G + B	V	S	FFT, ICA, PCA	Local
[45]	2016	-	RGB	G	-	E/LM	Spectral, IBI	BioVid Heat Pain (<3)
[37]	2017	Sony DCR-HC27E	RGB	R + G + B	V	E/LM	IVA, M-CCA	DEAP (5.0017)
[18]	2017	-	RGB	R + G + B	V	LM	ICA + JADE	Local (4.8)
[41]	2018	Smartphone	HSV	Hue (0-0.1)	V	SM	FFT	Local (4.1617)
[32]	2018	Webcam (Logitech C920)	YIQ	Q	V	E/SM	MVSGC	Local
[42]	2019	Four Webcam (Logitech C920)	RGB	R + G + B	V	SM/T/ LM	ICA + machine learning	Local (1.43, 0.96)
[39]	2019	FLIR Cricket IP	YCbCr	Y	V	S	Spatial filtering + temporal filtering + FFT	Local (2.07)
[20]	2019	Webcam, Phone, RealSense	YUV	Y + U + V	V	E/S/ LM/T	RhythmNet (CNN)	MAHNOB-HCI/ MMSE-HR/VIPL-HR (4.00/5.03/8.14)
[38]	2019	Nikon (D610)	RGB	G	V	S	EVM+ FFT + peak detection	Local
[40]	2019	Microsoft Kinect (IR)	-	-	D	S	Matrix decomposition + STFT	Local (5.39)
[21]	2020	Nikon (D5300)	RGB	G	V	E/SM/S/ LM/T	STFT + TFR + CNN	MAHNOB-HCI/VIPL-HR (3.08/7.82)
[22]	2020	-	RGB	R + B	V	E/SM	CNN	MAHNOB-HCI (7.45)
[33]	2020	Webcam (Logitech C920)	RGB	R + G + B	V	S/LM/T	LMS + FFT	VIPL-HR/Local (26.56/4.67)

bpm: Beat per minute, F: Fixed illumination, V: Variance illumination, D: Dim environment, E: Expression, S: Stable, T: Talk, SM: Slight motion, L: Large motion, -: Not mentioned in the research, RGB: Red Green Blue, ICA: Independent component analysis, CCD: Charged coupled device, HSL: Hue saturation lightness, PCA: Principal components analysis, FFT: Fast Fourier transform, EVM: Eulerian video magnification, SAMC: Self-adaptive matrix completion, LMS: Least medium square, IBI: Inter-beat-interval, MVSGC: Multiscale variable-weight SavitzkyGolay combination, M-CCA: Multiset canonical correlation analysis, IVA: Independent vector analysis, JADE: Joint approximate diagonalization eigen-matrices, YIQ: Q channel of the in-phase quadrature, CNN: Convolutional neural network, STFT: Short-time Fourier transform, TFR: Time-frequency representation, HSV: Hue saturation value

for the same HR estimation method, as shown in Figure 3.

It was found that the head motion has much greater influence on the intensity than on the chromaticity of the recorded video, since the chromaticity represents the inherent optical properties of blood hemoglobin.[46] Therefore, it is beneficial to choose a color space that distinguishes chromaticity from intensity such as YUV and YCrCb, to reduce the artifacts caused by motion.

Illumination and motion

The human-skin when illuminated gives two types of reflection: Spectral reflection and diffused reflection. The spectral reflected rays come directly from the surface of the skin, showing a light source color and do not contain an iPPG signal. While the diffused reflection, transmitted through the skin and shows discoloration of the skin, including differences that occur due to the cardiac cycle. The illumination affects the accuracy of HR estimation. Good illumination gives better accuracy than dimed or frequent illumination, as shown in Figure 4.

Subjects' head movement, facial expressions, and eye blinking are also considered as challenges on the iPPG signal extraction.

The reviewed papers show that the illumination sources and methods are various in addition to persons' states. Varying quantities of sunlight entered through windows,[31] professional,[34] natural varying,[35] fixed source,[5] a bright light,[29] office fluorescent,[36] constant sunlight,[30] and low[40] illumination were used and the persons were asked to stay still during the experiments. Various illumination was used in Blöcher *et al*. and Sanyal and Nundy,[18,41] with subjects' movement in Blöcher *et al*.,[18] and minimal movement in Sanyal and Nundy.[41] Various illumination and the participants were allowed to do reading, smiling, and yawning in a range (±10).[32] Normal

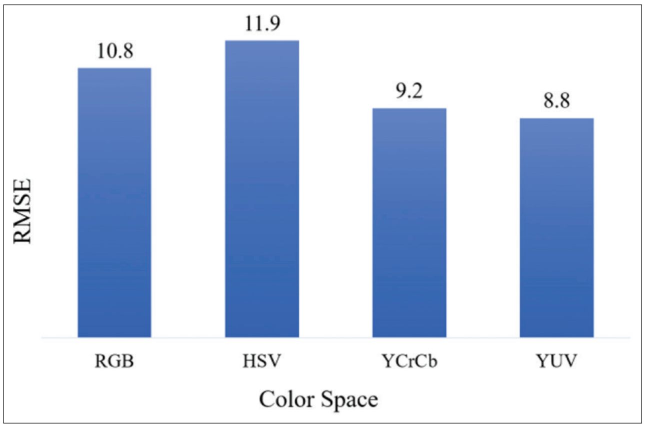

Figure 3: The root mean squared error (RMSE) of heart rate estimation using different color spaces by RhythmNet method [20]

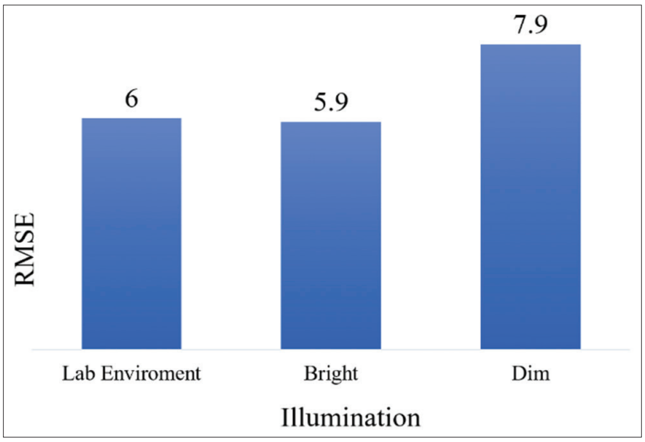

Figure 4: The heart rate estimation errors (in RMSE) by RhythmNet method under different illumination conditions [20]

fluorescent to dim illumination was used with infants allowing any covering or sleeping position for the purposes of research.[38] Mixed fluorescent and natural sunlight were applied and the subjects were asked to talking, walking, or lying down in the room.[42] Bonomi *et al*.[38] were made the participants watched video clips containing neutral emotional triggers and were asked them to stand still.

The illumination and subjects' movement for the studies that used public datasets explained in the second section of this paper.

Processing methods and its efficiency

The processing stage is composed of two main phases; the image processing phase that extracts the iPPG signal from the facial video, and the signal processing phase which is used to calculate the HR from the iPPG signal, as shown in Figure 5a. Many techniques and procedures were adopted in the covered studies here, as shown in Figure 5b and c. The image processing consists mainly of the face detection and tracking, ROI-tracking, and the iPPG signal extraction from the facial video either from color magnification or from small movement magnification. The aim of the signal processing phase is to find the best way that gives the most accurate HR from iPPG.

Blind source/signal separation (BSS) is one of the most conventional methods used to extract the iPPG signal from the raw-signal. BSS refers to the separation of a collection of signals from the mixed-signals when both the source signal and mixing process are unknown.[47] The most techniques of BSS used in the extraction of the iPPG signal are independent component analysis (ICA), principal components analysis (PCA), dependent component analysis, and joint approximate diagonalization eigen-matrices (JADE). However, the efficiency of conventional iPPG methods is readily degenerated by noise interference. Recently, several deep-learning iPPG-based approaches have been introduced, which have shown good efficiency against noise.

Poh *et al*.[31] were reached high degrees of agreement between measurements by applying ICA based on the JADE algorithm.[48]

A proposed algorithm by De Haan and Jeanne that used a robust chrominance-based HR estimation

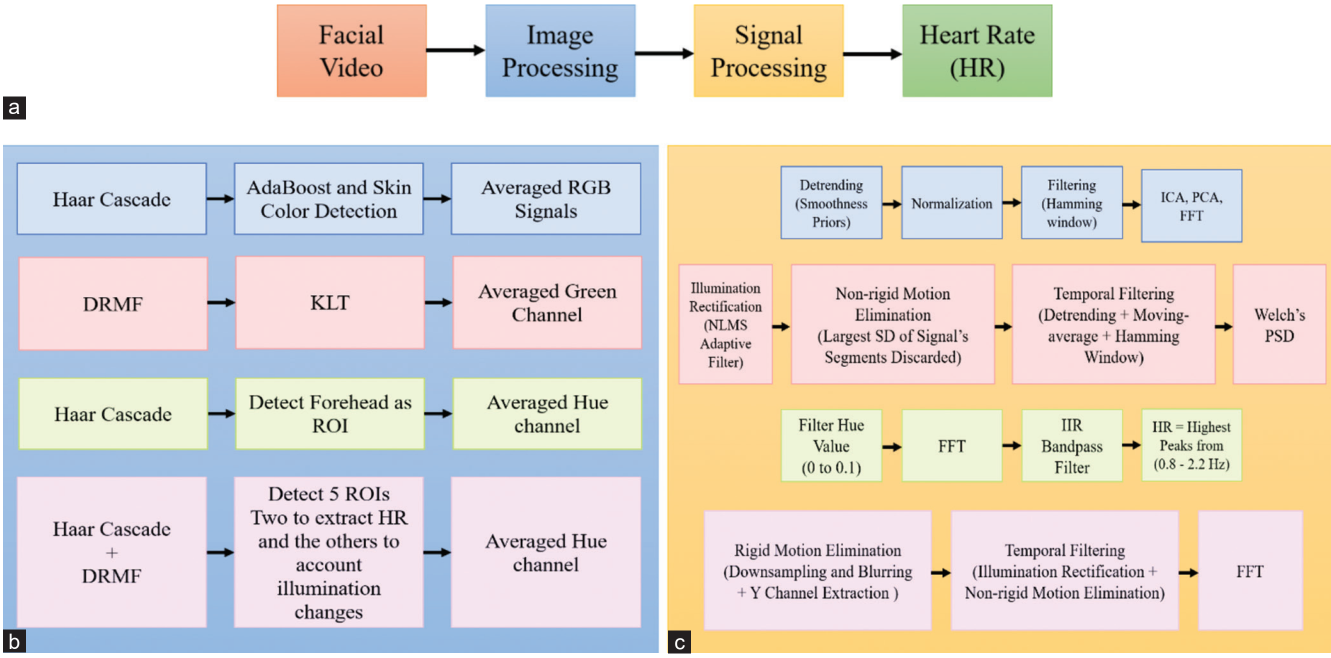

Figure 5: (a) Video processing stages to get the heart rate. (b) Image processing stage in Li *et al*., Rahman *et al*., Bonomi *et al*., Sanyal and Nundy.[5,30,39,41] (c) Signal Processing Stage in Li *et al*., Rahman *et al*., Bonomi *et al*., Sanyal and Nundy.[5,30,39,41]

method has shown better results than ICA and PCA.[34]

HR was extracted through measuring delicate head motion induced at any beat by the Newtonian reaction to the flow of blood. The suggested method tracked features, on the head and applied PCA to decompose their paths into a set of movements of components. Then, the component that most corresponds to HRs based on its spectrum of temporal frequency was chosen. Finally, the movement forecasted to this component was analyzed and peaks of the paths were identified, which correspond to Hrs. The measured HRs was very near to the ECG readings.[35]

Li *et al*.[5] were proposed a processing scenario that can reduce the noise due to subjects' motion and illumination's fluctuation. The proposed framework was composed of, DRMF method to seek for good ROI, KLT algorithm to track that ROI, raw-signal was found by calculating the average of the ROI's green channel value for each frame, Normalized Least Mean Squares adaptive filter to correct the interference of lighting variations, discarding the noisy-signal's segments due to abrupt non-rigid motions, excluding the powers of frequencies outside the HR range using several time filters, and finally applying Welch's method[49] to estimate the HR frequency. The results showed that the proposed framework achieves promising results under realistic conditions.

Based on Eulerian Video Magnification (EVM) algorithm proposed at SIGGRAPH in 2012, Carvalho *et al*. have applied face tracking, decomposition of the pyramid followed by frame filtering, then signal amplification to expose hidden information and monitor HR.[29]

Yu *et al*.[36] were estimated HR depending on the extracted BVP using PCA algorithm. The results showed that the proposed approach was able to predict dynamic HR from sequences of short videos using fewer computational criteria compared to Yu *et al*.[50]

Lam and Kuno[44] were extracted the iPPG signal from two ROIs, selected by voting from multi local ROIs, in the green channel then applied ICA algorithm. Their method was active against illumination fluctuation and subject's motions and showed better results than some previous works,[5,31] but it cannot be applied in real time.

Tulyakov *et al*.[6] were implemented an approach, named it as self-adaptive matrix completion (SAMC), to dynamically select ROIs for strong HR estimation. While discovering the best ROI of the face to be used for prediction, SAMC allows simultaneous prediction of HR. The method used in this research has surpassed modern HR estimation methods under realistic conditions.

Rahman *et al*.,[30] in their study, three different methods of signal processing have been experimented: Fast Fourier Transform (FFT), PCA, and ICA. The results showed a high degree of agreement between the presented method and reference measurements.

A low latency technique for continuous frame-based HR prediction was proposed by Rapczynski *et al*. to process and analyze the extracted iPPG signals in real-time. They used two different techniques to find the right BVP signal: An inter-beat-interval and a spectral-based technique to detect the dominant frequency, which was assumed to be the BVP. A combination of skin detection and face tracking in short time-windows (10 s) was adopted by the researchers.[45]

Qi *et al*.[37] performed two popular joint blind source separation algorithms: Multiset canonical correlation analysis and independent vector analysis. The results showed that the proposed approach was outperformed ICA-based approaches.

A developed algorithm by Blocher *et al*. included several signal and image processing techniques such as ICA, spatial filtering, and Hilbert transform for the signal which used peak detection algorithm to estimate the desired parameters. The findings of the evaluation showed high accuracy for the detection of pulse peak and HR estimation relative to the ECG system.[18]

Sanyal and Nundy[41] were utilized Haar cascade to detect face and eyes. The raw-iPPG signal was extracted by finding the average Hue channel of the forehead, and then frequency transformation was done on the average pixel. The IIR bandpass filter was used to remove unwanted frequencies located out of the HR frequencies range (0.8–2.2 Hz). The results showed that using Hue channel was more accurate than green channel usage.

The YIQ color space was used to extract iPPG signal along with the multiscale variable-weight SavitzkyGolay combination, which is a mathematical model for signal extraction in the time domain with minimized noise disruption due to illumination changes, motion, and facial expression. Good results were obtained with comparison to the gold-standard way of recording signals from a finger-tip through a pulse oximeter.[32]

The proposed method, by Ghanadian *et al*., to find HR for persons in movement state using a single camera[51] was further developed using 4-cameras.[42] ICA was used to separate the RGB signals into three components, and then the machine learning approach to choose the best iPPG signal. Compared to, the state-of-the-art status of the subject in movement, the proposed method reduces the RMSE by 18%.

Bonomi *et al*.[39] utilized an approach of three phases: spatial filtering, temporal filtering, and frequency analysis. In the first phase, the Haar cascade, DRMF, and KLT algorithms were used for face detection and ROIs tracking. Down sampling and blurring filters were adopted as processes for spatial filtering to enhance the most important ROIs' color components. Then, the Y-component of the YCbCr color space was extracted. In the second phase, the mathematical average of every ROIs Y-component was computed. The signals contained the pulse were added, but that classified as the source of noise were subtracted, and then the result has normalized. The approach presented in Li *et al*.[5] was applied to rectify the time-domain signal from non-rigid motions, where each temporal window was divided into samples, and the standard deviation (SD) was computed for each window. The highest SDs for 30% of the samples have been discarded, where the others were concatenated. In the last phase, a bandpass filter (0.67 and 1.67 Hz) was used to remove the signal elements out of the HR band. Then, the Welch's PSD[49] was applied. Their approach showed promising results in the quality of experience analysis.

Niu *et al*.[20] named a method RhythmNet for HR estimation from the facial videos. RyhthmNet was a learning approach that used a spatial-temporal representation to encode the HR signals from the ROIs as inputs. For HR estimation, the spatiotemporal representations were then fed into a convolutional neural network (CNN). Promising HR estimation precision was reached with this method.

A research on HR estimation for infants was conducted by Gibson *et al*. using a modified EVM technique to magnify the captured videos. After that, the pixels' average of the R, G, and B channels of the selected ROIs was computed, and the averaged signal of the G channel was chosen to estimate the HR signal. FFT was then used, followed by a band-pass filter with frequencies 0.5–3 Hz. Then, the inverse FFT, and peak detection was performed.[38]

Martinez *et al*.[40] presented a simple method to estimate, instantaneous HR (iHR) based on matrix decomposition using an IR camera to overcome the illumination problems. The experiments were showed that the recovered signal matched the ground truth iHR with minor proportional errors.

A method consists of four steps was presented by Hsu *et al*.[21] In Step 1, the face was detected and tracked and the required ROIs were identified. In Step 2, the average of the green channel was computed from the ROIs, and three-stage sequential filtering was applied: Light rectification, trend elimination, and signal amplification. The Short-Time Fourier Transform was used in Step 3 to translate the 1D filtered signal to 2D Time-Frequency Representation (TFR) to describe frequencies through short periods of time. In the last step, a deep CNN was explored to resolve the learning problem of the HR estimation, which was formulated by the 2D TFR. Their results ranked at the top of the list in comparison to eight other related approaches.

Song *et al*.[22] proposed an iPPG approach with CNN. The approach was used to map the image of the HR feature to the equivalent HR value through a ResNet-18 network using a feature decoder method. Using pulse signals derived from traditional iPPG techniques, the images of the spatiotemporal feature were built in a time-delayed way. The CNN model was trained first with images of synthetic features extracted from BVP or ECG signals. Then, it was further refined with images of real features produced by noise-contaminated iPPG pulses. The findings indicate that, overall, the best performance was obtained by the presented approach relative to several other typical-iPPG methods.

Zhao *et al*.[33] were presented a method of four phases.

In the first phase, the KLT algorithm was used to track the manually selected ROI in the first frame. Then, the average of each channel was computed for the ROI. In the second phase, the preliminary pulse was extracted using a statistical learning approach to determine the direction of projection. In the third phase, motion artifact was removed using least medium square. In the last phase, the PSD was computed using FFT and modulated using a weight function created by the Gaussian distribution, determined by the previous estimates of the HR. The proposed algorithm exceeded the state-of-the-art algorithms by a wide margin in terms of signal-to-noise ratio and accuracy, particularly when large motion occurs.

RELATED REVIEW AND SURVEYPAPERS

In 2014, Kranjec *et al*.[52] presented a concise summary of the two conventional approaches, ECG and PPG, and later on focused on the approaches of non-contact HR measuring with capacitive coupled ECG, optical vibrocardiography, Doppler radar, thermal imaging, HR from speech, and RGB imaging. The research represented a comparative analysis of these approaches, thus outlining their benefits and drawbacks. All of the approaches mentioned have been effective and capable of delivering the requested information in a specific environment and with certain accuracy and reliability. Although all of the approaches mentioned relied on the extraction of heart activity parameters from one individual, the RGB imaging approach has been effective in HR measurement from at least three individuals at the same time. The complexity of the discussed methods has differed; the process was most challenging in the approaches: HR from speech, thermal imaging, and RBG imaging. The advantages of RGB imaging discussed methods were multiple subjects and a low-cost sensor, while the disadvantages were complex algorithms and low temporal resolution. The advantage of thermal imaging reviewed papers was a passive non-contact approach, while the disadvantages were an expensive sensor and low temporal resolution. The studies covered in this paper were performed over the years from 1981–2014 as depicted in Figure 6a.

In 2016, Sikdar *et al*.[53] did a comprehensive review of video and image-based HR estimation approaches. Five state-of-the-art approaches on a homogeneous platform were implemented and evaluated with a large dataset in different environmental conditions. Through evaluation, they observed that none of these approaches can estimate accurate HR at all kinds of conditions, therefore, may not be appropriate for clinical environments unless optimized further, however, home-based non-clinical applications may be used. Among the five approaches that were implemented and benchmarked, LE by Wei *et al*.[54] was the strongest followed by the optimized Balakrishnan's approach[35] (PCA+FFT) presented by Irani *et al*.[3] (PCA+DCT). The authors of this paper believed that the contributions made by the discussed studies (2007–2015) in their paper are

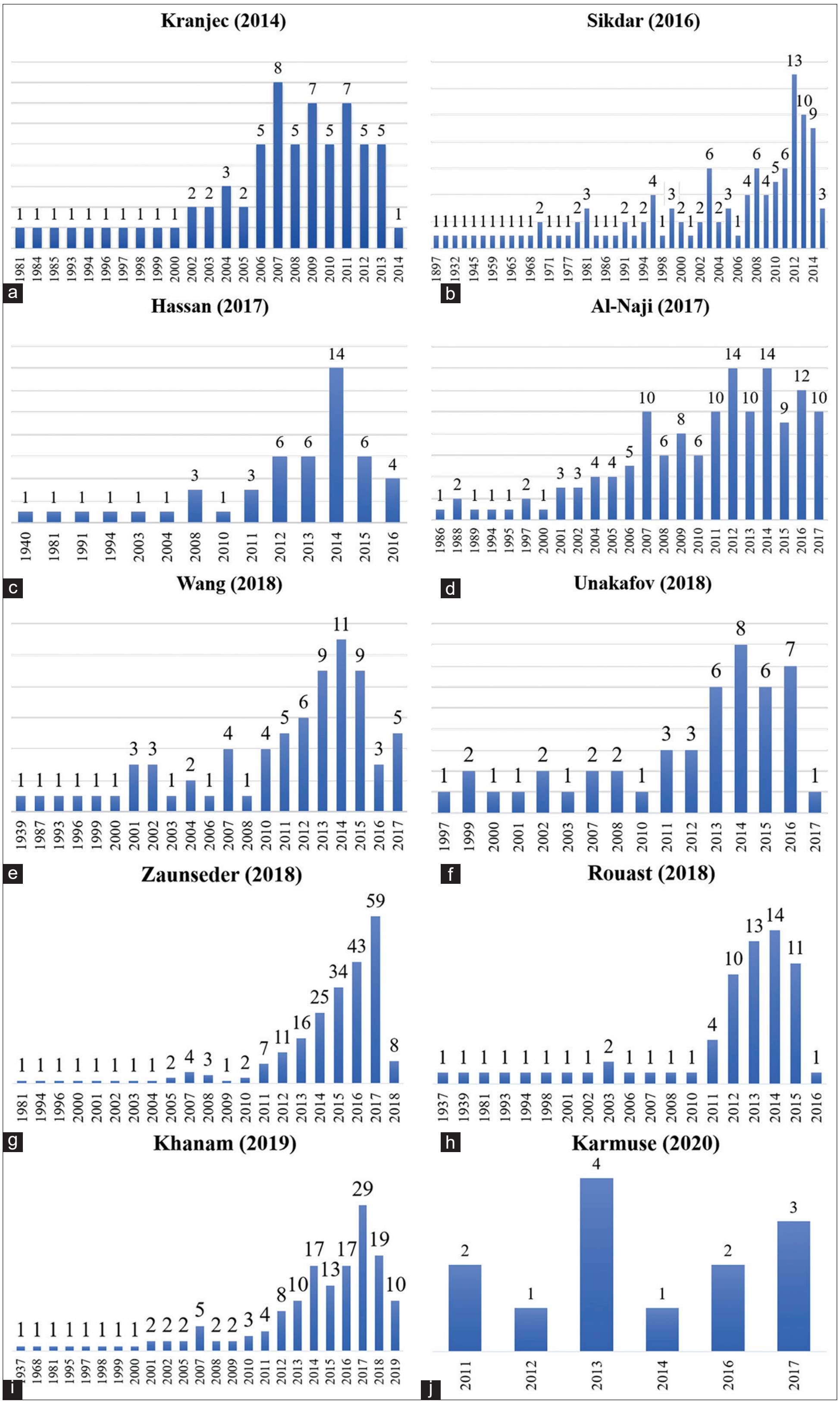

Figure 6: (a-j) Studies covered in the review and survey papers per year

important and they may open up fresh research dimensions. They also stated that with significant progress in optical camera technology, researchers could obtain extra accurate information from head tracking or skin color. They covered many studies over the years from 1897 to 2015, as shown in Figure 6b.

The published paper at 2017 by Hassan *et al*.[55] critically reviewed the progress of the state-of-the-art in HR estimation using facial video. They have studied the two predominant concepts of HR measurement, PPG and BCG. They also discussed the principles and theories underlying each of these methods on measuring HR from video. Certain methods were experimented, with cross-validated, and evaluated their reliability under realistic-situations. They concluded that BCG-based techniques showed encouraging outcomes, but inconsistent extraction of features was a major downside to the reliability of the approach, whereas lighting stability was one of the major advantages of BCG-based approaches. They stated that iPPG signal extraction from multi-ROI had increased the overall signal power. The authors think that, many improvements must be made to measure HR under realistic-situations; more attention should be paid to correcting spatial and temporal lighting contrast and motion contrast caused by changing body position. They proposed the following solutions to increase the reliability of HR estimation from facial video: Merging the features of both BCG-based method used in the paper[56] and the iPPG-based methods presented in the studies,[44,57] might overcome the limitations of the approaches related to motion contrast and illumination contrast. The published papers in the interval 1940–2016 were covered in this review, as depicted in Figure 6c.

In 2017 too, Al-Naji *et al*.[58] reviewed and compared the newest, most promising remote HR measuring methods of Doppler radar, video imaging, and thermal imaging. They discussed their features and limitations under various circumstances. In addition, the study summarized the performance of these approaches in a table with respect to the subject movement, noise artifacts, the number of ROIs, detection range, biological effects, multiple subjects, and cost. The study explained that all the discussed approaches could be effective alternative methods to traditional contact measuring ways used in clinical and biomedical environments. The authors believed that a computer vision-based system using a video camera with visible light offers advantages where the system cost, subjects' safety, detection of multiple subjects, long-term monitoring, and long-distance are the main considerations. This review covered the related works through 1986–2017 and they are shown as a histogram in Figure 6d.

The year 2018 attracted many researchers to review the published activities in this field, among them was Wang *et al*.[59] whom described and evaluated several works about HR-monitoring through videoing along the period 2008–2018. The framework used in the article consisted of three stages: Processing the face video, extract the BVP signal from the face, and HR estimation. According to every stage, the presented methods were classified and grouped. MAHNOB-HCI dataset was used to evaluate and compare the performance of the presented methods. Results were shown that skin area extracting, BSS, and peak detection are the most robust methods for HR estimation with head motions. From the point of view of the authors of this article, the state-of-the-art methods are not strong enough under naturalist situations and still cannot estimate HR in real-time; machine learning approaches could be powerful and accurate for remote HR estimation even under naturalist situations, with the deep learning develop. The studies covered in this paper were distributed over the years from 1939 to 2017, as shown in Figure 6e.

Another survey paper was prepared by Unakafov in 2018.[60] It contains comparisons of HR estimation methods using iPPG extracted from the DEAP dataset videos. They were mentioned that the DEAP database has many features as its videos are of good enough quality and allow more accurate HR estimation than the MMSE-HR and MAHNOB datasets, and also this dataset can be used to verify the estimation of HR variability and respiratory rate from iPPG. They also mentioned two drawbacks of the DEAP dataset for iPPG extraction, which are: First, as a reference, just contact-photoplethysmogram is available, although more accurate ECG would be preferred; second, a limited number of motions in DEAP videos lowers the utility of this dataset for examining face-tracking methods and suppressing motion defects. A framework consisted from five steps for iPPG-based HR estimation were proposed, which was used to compare various approaches for iPPG analysis; best HR estimation was achieved when using the POS method[61] to extract iPPG and continuous wavelet transform for HR estimation. The studies over 1997–2017 were covered in this paper as swept in Figure 6f.

In 2018, Zaunseder *et al*.[62] provided another review paper to overview the background of the acquired cardiovascular parameters using contactless cameras, and realizations. They also gave extensive information on the extremely popular application of this technique, outline additional concepts, and critically discussed the state at that time. Figure 6g depicts the reviewed studies throughout the years from 1981 to 2018.

The last survey paper of 2018 that we present here, is produced by Rouast *et al*.[63] whom reviewed the growth of the iPPG studies since its appearance in 2008 until 2016 and described them based on head motion and skin color. They also classified iPPG methods and derived a framework that offered a summary of modular steps to enable practitioners to design iPPG algorithms that satisfy their special requirements. Discussed literature was tabulated on the basis of a contribution to the domain and the used algorithm, for the first time. The two key problems discussed in the iPPG were; the robustness of the algorithm in relation to subject noise and poor signal strength processing due to lighting and skin types. According to the viewpoint of this review authors, future iPPG algorithms should trade-off between the quantity of information processed and the complexity of the algorithm, since real-time applications would constraint computation time. The covered studies in this paper, from 1937 to 2016, are shown as a histogram in Figure 6h.

In 2019, Khanam *et al*.[64] published a review on camera-imaging-based techniques of recent works, where various studies focused on motion and color-based approaches have been discussed. Then, various aspects of color-based techniques were monitored, for example, illumination variations, motion artifacts, alternate sensors, various vital-signs, different subjects, multiple subjects, multiple ROIs, and long-distance. In addition, possible applications of iPPG were described in both environments, clinical and non-clinical environments. The gaps and difficulties of the discussed studies were also established and the authors provided some pointers for future works. It emerged from this study that simultaneous multi-subject monitoring, automatic multi-ROIs selection, noise artifact removal induced by both motion and illumination variations, multi-camera fusion, publicly available datasets, and long-distance detection are topics that need a study yet to develop many real-world applications. According to the viewpoint of authors of this paper, computer vision-based approaches could be hygienic, robust, reliable, cost-effective, safe, and suitable for long-term and long-distance monitoring methods for non-contact physiological assessments. This paper could be a road for new researchers to discover and understand the gaps and challenges in recent studies. The studies covered in this paper were performed over the years from 1937 to 2019 and sketched as a statistical form in Figure 6i.

The last review of the literature included in this paper is the one presented by Karmuse and Kakhandki in 2020.[65] They described a review of real-time methods to measure HR using a laptop computer’s USB webcam. The researchers pointed to the possibility of using this technology to develop personal healthcare and telemedicine. The histogram of the covered studies over the years from 2011 to 2017 is shown in Figure 6j.

Table 2: Noncontact HR monitoring applications

Paper	Application
[66-70]	Monitoring infants in the NICU
[71]	Prevention of sudden infants’ death syndrome and detection of heart attack and stroke in the elderly at home (telemedicine)
[72]	Screening of infective people with seasonal influenza or COVID-19 in mass gathering places
[73,74]	Sleep monitoring
[75,76]	Sleep apnea monitoring and detection
[77]	A rapid examination of patients suspected of having infectious diseases
[78,79]	Monitoring during fitness exercise
[80-85]	Driver monitoring
[86]	Emotional and stress states

NICU: Neonatal intensive care unit, HR: Heart rate

Applications

Non-contact HR monitoring by a camera has many applications in clinical and non-clinical environments, such as sleep monitoring, driver monitoring, infants and elderly monitoring, telemedicine, and stress detection as summarized in Table 2.

CONCLUSION

While HR estimation from a face video is a low-cost, simple, and comfortable method, it is a difficult problem in less-constrained scenarios due to head movement, variations in illumination, and sensor variety. The conventional methods succeeded in estimating the HR but under controlled conditions. Researchers are continuing attempts to find methods for estimating HR in real conditions for this method to become reliable and approved. Recent methods used deep-learning iPPG-based approaches and showed clear improvement in estimating the HR under realistic conditions, but with more complex calculations than the conventional methods.

The available public databases that can be used to test the efficiency of HR estimation methods are limited and do not contain major movements to be used as a real test representing realistic conditions. MAHNOB-HCI dataset is the most database used in researches.

Non-contact HR monitoring by a camera has many applications in clinical and non-clinical environments, such as sleep monitoring, sleep apnea monitoring, driver monitoring, telemedicine, emotional and stress detection, screening of infective people with seasonal influenza or COVID-19 in mass gathering places, prevention of sudden infants' death syndrome at home, and detection of heart attack and stroke in the elderly.

We think that to reach practical systems for HR monitoring through videoing, the application environment should be specified first, then the appropriate system parameters and processing algorithms in addition to displaying methods. This strategy could give the most suitable system for a certain application; however, a given system may fit to more than one application. For example, real-time operation may be critical in some applications, while person's movement cannot be avoided in others and so on.

REFERENCES

1. Phan D, Siong LY, Pathirana PN, Seneviratne A, editors. Smartwatch: Performance Evaluation for Long-term Heart Rate Monitoring. 2015 International Symposium on Bioelectronics and Bioinformatics; 2015.
2. Poh MZ, McDuff DJ, Picard RW. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Opt Exp 2010;18:10762-74.
3. Irani R, Nasrollahi K, Moeslund TB, editors. Improved Pulse Detection from Head Motions using DCT. 2014 International Conference on Computer Vision Theory and Applications; 2014.
4. Lomaliza JP, Park H, editors. Detecting Pulse from Head Motions Using Smartphone Camera. International Conference on Advanced Engineering Theory and Applications. Berlin: Springer; 2016.
5. Li X, Chen J, Zhao G, Pietikainen M, editors. Remote Heart Rate Measurement from Face Videos under Realistic Situations. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2014.
6. Tulyakov S, Alameda-Pineda X, Ricci E, Yin L, Cohn JF, Sebe N, editors. Self-adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016.
7. Viola P, Jones M, editors. Rapid Object Detection Using a Boosted Cascade of Simple Features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2001.
8. Kasinski A, Schmidt A. The architecture of the face and eyes detection system based on cascade classifiers. Comput Recogn Syst 2007;2:124-31.
9. Asthana A, Zafeiriou S, Cheng S, Pantic M, editors. Robust Discriminative Response Map Fitting with Constrained Local Models. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2013.
10. Tomasi C, Kanade T. Detection and Tracking of Pointfeatures. Technical Report CMU-CS-91-132. Pennsylvania, PA: Carnegie Mellon University; 1991.
11. Hoan NV, Park JH, Lee SH, Kwon KR. Real-time heart rate measurement based on photoplethysmography using android smartphone camera. J Korea Multimed Soc 2017;20:234-43.
12. Pelegris P, Banitsas K, Orbach T, Marias K, editors. A Novel Method to Detect Heart Beat rate Using a Mobile Phone. 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology; 2010.
13. Bhelkar V, Shedge D. Design and implementation of health monitoring device using FPGA. Int J Comput Appl 2017;158:0975-8887.
14. Lindelöw M, Lindqvist A. Remote Heart Rate Extraction from Near Infrared Videos an Approach to Heart Rate Measurements for the Smart Eye Head Tracking System; 2016.
15. Kwon S, Kim J, Lee D, Park K, editors. ROI Analysis for Remote Photoplethysmography on Facial Video. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2015.
16. Hernandez-Ortega J, Fierrez J, Morales A, Tome P, editors. Time Analysis of Pulse-based Face Anti-spoofing in Visible and NIR. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops; 2018.
17. Ren L, Kong L, Foroughian F, Wang H, Theilmann P, Fathy AE. Comparison study of noncontact vital signs detection using a Doppler stepped-frequency continuous-wave radar and camera-based imaging photoplethysmography. IEEE Trans Microw Theor Tech 2017;65:3519-29.
18. Blöcher T, Schneider J, Schinle M, Stork W, editors. An Online PPGI Approach for Camera Based Heart Rate monitoring Using Beat-to-beat Detection. 2017 IEEE Sensors Applications Symposium; 2017.
19. Soleymani M, Lichtenauer J, Pun T, Pantic M. A multimodal database for affect recognition and implicit

tagging. IEEE Trans Affect Comput 2011;3:42-55.
20. Niu X, Shan S, Han H, Chen X. RhythmNet: End-to-end heart rate estimation from face via spatial-temporal representation. IEEE Trans Image Proc 2019;29:2409-23.
21. Hsu GS, Xie RC, Ambikapathi A, Chou KJ. A deep learning framework for heart rate estimation from facial videos. Neurocomputing 2020;417:155-66.
22. Song R, Zhang S, Li C, Zhang Y, Cheng J, Chen X. Heart rate estimation from facial videos using a spatiotemporal representation with convolutional neural networks. IEEE Trans Instrument Measur 2020;69:7411-21.
23. Zhang Z, Girard JM, Wu Y, Zhang X, Liu P, Ciftci U, *et al*., editors. Multimodal Spontaneous Emotion Corpus for Human Behavior Analysis. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016.
24. Stricker R, Müller S, Gross HM, editors. Non-contact Video-based Pulse Rate Measurement on a Mobile Service Robot. The 23rd IEEE International Symposium on Robot and Human Interactive Communication; 2014.
25. Walter S, Gruss S, Ehleiter H, Tan J, Traue HC, Werner P, *et al*., editors. The BioVid Heat Pain Database: Data for the Advancement and Systematic Validation of an Automated Pain Recognition System. 2013 IEEE International Conference on Cybernetics; 2013.
26. Werner P, Al-Hamadi A, Niese R, Walter S, Gruss S, Traue HC, editors. Towards Pain Monitoring: Facial Expression, Head Pose, a new Database, an Automatic System and Remaining Challenges. Proceedings of the British Machine Vision Conference; 2013.
27. Koelstra S, Muhl C, Soleymani M, Lee JS, Yazdani A, Ebrahimi T, *et al*. DEAP: A database for emotion analysis using physiological signals. IEEE Trans Affect Comput 2011;3:18-31.
28. Bobbia S, Macwan R, Benezeth Y, Mansouri A, Dubois J. Unsupervised skin tissue segmentation for remote photoplethysmography. Pattern Recogn Lett 2019;124:82-90.
29. Carvalho L, Virani M, Kutty MS. Analysis of heart rate monitoring using a webcam. Int J Adv Res Comput Commun Eng 2014;3:6593-5.
30. Rahman H, Ahmed MU, Begum S, Funk P, editors. Real Time Heart Rate Monitoring from Facial RGB Color Video Using Webcam. The 29th Annual Workshop of the Swedish Artificial Intelligence Society (SAIS), 2-3 June 2016. Malmö, Sweden: Linköping University Electronic Press; 2016.
31. Poh MZ, McDuff DJ, Picard RW. Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Trans Biomed Eng 2010;58:7-11.
32. Bai G, Huang J, Liu H. Real-time robust noncontact heart rate monitoring with a camera. IEEE Access 2018;6:33682-91.
33. Zhao C, Lin CL, Chen W, Chen MK, Wang J. Visual heart rate estimation and negative feedback control for fitness exercise. Biomed Sig Proc Control 2020;56:101680.
34. De Haan G, Jeanne V. Robust pulse-rate from chrominance-based rPPG. IEEE Trans Biomed Eng 2013;60:2878-86.
35. Balakrishnan G, Durand F, Guttag J, editors. Detecting Pulse from Head Motions in Video. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2013.
36. Yu YP, Raveendran P, Lim CL, Kwan BH. Dynamic heart rate estimation using principal component analysis. Biomed Opt Exp 2015;6:4610-8.
37. Qi H, Guo Z, Chen X, Shen Z, Wang ZJ. Video-based human heart rate measurement using joint blind source separation. Biomed Sig Proc Control 2017;31:309-20.
38. Gibson K, Al-Naji A, Fleet JA, Steen M, Chahl J, Huynh J, *et al*. Noncontact heart and respiratory rate monitoring of preterm infants based on a computer vision system: Protocol for a method comparison study. JMIR Res Prot 2019;8:e13400.
39. Bonomi M, Battisti F, Boato G, Barreda-Angeles M, Carli M, Le Callet P. Contactless approach for heart rate estimation for QoE assessment. Sig Proc 2019;78:223-35.
40. Martinez N, Bertran M, Sapiro G, Wu HT, editors. Non-contact Photoplethysmogram and Instantaneous Heart Rate Estimation from Infrared Face Video. 2019 IEEE International Conference on Image Processing; 2019.
41. Sanyal S, Nundy KK. Algorithms for monitoring heart rate and respiratory rate from the video of a user's face. IEEE J Trans Eng Health Med 2018;6:1-11.
42. Ghanadian H, Al Osman H, editors. Non-contact Heart Rate Monitoring Using Multiple RGB Cameras. International Conference on Computer Analysis of Images and Patterns. Berlin: Springer; 2019.
43. Niu X, Han H, Shan S, Chen X, editors. Continuous Heart Rate Measurement from Face: A Robust rPPG Approach with Distribution Learning. 2017 IEEE International Joint Conference on Biometrics; 2017.
44. Lam A, Kuno Y, editors. Robust Heart Rate Measurement from Video Using Select Random Patches. Proceedings of the IEEE International Conference on Computer Vision; 2015.
45. Rapczynski M, Werner P, Al-Hamadi A, editors. Continuous Low Latency Heart Rate Estimation from Painful Faces in Real Time. 23rd International Conference on Pattern Recognition; 2016.
46. Yang Y, Liu C, Yu H, Shao D, Tsow F, Tao N. Motion robust remote photoplethysmography in CIELab color space. J Biomed Opt 2016;21:117001.
47. Pal M, Roy R, Basu J, Bepari MS, editors. Blind Source Separation: A Review and Analysis. 2013 International Conference Oriental COCOSDA Held Jointly with 2013 Conference on Asian Spoken Language Research and Evaluation; 2013.
48. Cardoso JF. High-order contrasts for independent component analysis. Neural Comput 1999;11:157-92.
49. Welch P. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Trans Audio Electroacoustics 1967;15:70-3.
50. Yu YP, Raveendran P, Lim CL. Dynamic heart rate measurements from video sequences. Biomed Opt Exp 2015;6:2466-80.
51. Ghanadian H, Ghodratigohar M, Al Osman H. A machine learning method to improve non-contact heart rate monitoring using an RGB camera. IEEE Access

2018;6:57085-94.
52. Kranjec J, Beguš S, Geršak G, Drnovšek J. Non-contact heart rate and heart rate variability measurements: A review. Biomed Sig Proc Control 2014;13:102-12.
53. Sikdar A, Behera SK, Dogra DP. Computer vision guided human pulse rate estimation: A review. IEEE Rev Biom Eng 2016;9:91-105.
54. Wei L, Tian Y, Wang Y, Ebrahimi T, Huang T, editors. Automatic Webcam-Based Human Heart Rate Measurements Using Laplacian Eigenmap. Asian Conference on Computer Vision. Berlin: Springer; 2012.
55. Hassan MA, Malik AS, Fofi D, Saad N, Karasfi B, Ali YS, *et al*. Heart rate estimation using facial video: A review. Biomed Sig Proc Control 2017;38:346-60.
56. Haque MA, Irani R, Nasrollahi K, Moeslund TB. Heartbeat rate measurement from facial video. IEEE Intell Syst 2016;31:40-8.
57. Feng L, Po LM, Xu X, Li Y, Ma R. Motion-resistant remote imaging photoplethysmography based on the optical properties of skin. IEEE Trans Circ Syst Video Technol 2014;25:879-91.
58. Al-Naji A, Gibson K, Lee SH, Chahl J. Monitoring of cardiorespiratory signal: Principles of remote measurements and review of methods. IEEE Access 2017;5:15776-90.
59. Wang C, Pun T, Chanel G. A comparative survey of methods for romote heart rate detection from frontal face videos frontiers in bioengineering and biotechnology. Front Bioeng Biotechnol 2018;6:33.
60. Unakafov AM. Pulse rate estimation using imaging photoplethysmography: Generic framework and comparison of methods on a publicly available dataset. Biomed Phys Eng Express 2018;4:045001.
61. Wang W, den Brinker AC, Stuijk S, De Haan G. Algorithmic principles of remote-PPG. IEEE Trans Biomed Eng 2016;64:1479-91.
62. Zaunseder S, Trumpp A, Wedekind D, Malberg H. Cardiovascular assessment by imaging photoplethysmography a review. Biomed Eng Biomed Technik 2018;63:617-34.
63. Rouast PV, Adam MT, Chiong R, Cornforth D, Lux E. Remote heart rate measurement using low-cost RGB face video: A technical literature review. Front Comput Sci 2018;12:858-72.
64. Khanam FT, Al-Naji A, Chahl J. Remote monitoring of vital signs in diverse non-clinical and clinical scenarios using computer vision systems: A review. Appl Sci 2019;9:4474.
65. Karmuse MS, Kakhandki AL. A review on real time heart rate monitoring system using USB camera. Int J Eng Comput Sci 2020;9:24934-9.
66. Aarts LA, Jeanne V, Cleary JP, Lieber C, Nelson JS, Oetomo SB, *et al*. Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit a pilot study. Early Hum Dev 2013;89:943-8.
67. Mestha LK, Kyal S, Xu B, Lewis LE, Kumar V, editors. Towards Continuous Monitoring of Pulse Rate in Neonatal Intensive Care Unit with a Webcam. 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2014.
68. Scalise L, Bernacchia N, Ercoli I, Marchionni P, editors. Heart Rate Measurement in Neonatal Patients Using a Webcamera. 2012 IEEE International Symposium on Medical Measurements and Applications Proceedings; 2012.
69. Villarroel M, Guazzi A, Jorge J, Davis S, Watkinson P, Green G, *et al*. Continuous non-contact vital sign monitoring in neonatal intensive care unit. Healthc Technol Lett 2014;1:87-91.
70. van Gastel M, Balmaekers B, Oetomo SB, Verkruysse W, editors. Near-continuous Non-contact Cardiac Pulse Monitoring in a Neonatal Intensive Care Unit in Near Darkness. Optical Diagnostics and Sensing 18th Toward Point-of-care Diagnostics, International Society for Optics and Photonics; 2018.
71. Zhao F, Li M, Qian Y, Tsien JZ. Remote measurements of heart and respiration rates for telemedicine. PLoS One 2013;8:e71384.
72. Negishi T, Abe S, Matsui T, Liu H, Kurosawa M, Kirimoto T, *et al*. Contactless vital signs measurement system using RGB-thermal image sensors and its clinical screening test on patients with seasonal influenza. Sensors 2020;20:2171.
73. Vogels T, Van Gastel M, Wang W, De Haan G, editors. Fully-automatic Camera-based Pulse-Oximetry during Sleep. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops; 2018.
74. Hu M, Zhai G, Li D, Fan Y, Duan H, Zhu W, *et al*. Combination of near-infrared and thermal imaging techniques for the remote and simultaneous measurements of breathing and heart rates under sleep situation. PLoS One 2018;13:e0190466.
75. Li MH, Yadollahi A, Taati B. Non-contact vision-based cardiopulmonary monitoring in different sleeping positions. IEEE J Biomed Health Inf 2016; 21:1367-75.
76. Zhu K, Li M, Akbarian S, Hafezi M, Yadollahi A, Taati B. Vision-based heart and respiratory rate monitoring during sleep a validation study for the population at risk of sleep apnea. IEEE J Transl Eng Health Med 2019;7:1-8.
77. Sun G, Nakayama Y, Dagdanpurev S, Abe S, Nishimura H, Kirimoto T, *et al*. Remote sensing of multiple vital signs using a CMOS camera-equipped infrared thermography system and its clinical application in rapidly screening patients with suspected infectious diseases. Int J Infect Dis 2017;55:113-7.
78. Wang W, den Brinker AC, Stuijk S, de Haan G. Robust heart rate from fitness videos. Physiol Measur 2017;38:1023.
79. Wang W, Balmaekers B, De Haan G, editors. Quality Metric for Camera-based Pulse Rate Monitoring in Fitness Exercise. 2016 IEEE International Conference on Image Processing; 2016.
80. Lee K, Han DK, Ko H. Video analytic based health monitoring for driver in moving vehicle by extracting effective heart rate inducing features. J Adv Transport 2018;2018:8513487.
81. Kuo J, Koppel S, Charlton JL, Rudin-Brown CM. Evaluation of a video-based measure of driver heart rate. J Saf Res 2015;54:55.e29-59.
82. Guo Z, Wang ZJ, Shen Z, editors. Physiological

Parameter Monitoring of Drivers Based on Video Data and Independent Vector Analysis. 2014 IEEE International Conference on Acoustics, Speech and Signal Processing; 2014.
83. Gücüyener İ. A novel design of heartbeat monitoring system for the motor vehicle. Int J Injury Control Saf Promot 2016;23:395-9.
84. Qi H, Wang ZJ, Miao C, editors. Non-contact driver cardiac physiological monitoring using video data. 2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP); 2015.
85. Zhang Q, Wu Q, Zhou Y, Wu X, Ou Y, Zhou H. Webcam-based, non-contact, real-time measurement for the physiological parameters of drivers. Measurement 2017;100:311-21.
86. Puri C, Olson L, Pavlidis I, Levine J, Starren J, editors. StressCam: Non-contact Measurement of Users' Emotional States through Thermal Imaging. CHI'05 Extended Abstracts on Human Factors in Computing Systems; 2005.