Motion and Illumination Resistant Facial Video based Heart Rate Estimation Method using Levenberg-Marquardt Algorithm Optimized Undercomplete Independent Component Analysis

Ankit Gupta, Antonio G. Ravelo-García, and Fernando Morgado Dias *Member, IEEE*

Abstract—Heart Rate (HR) estimation is of utmost need due to its applicability in diverse fields. Conventional methods for HR estimation require skin contact and are not suitable for scenarios such as sensitive skin or prolonged unobtrusive HR monitoring. Therefore remote photoplethysmography (rPPG) methods have been an active area of research. These methods utilize the facial videos acquired using a camera followed by extracting the Blood Volume Pulse (BVP) signal for heart rate calculation. The existing rPPG methods either used a single color channel or weighted color differences, which has limitations dealing with motion and illumination artifacts. This study considers BVP extraction as an undercomplete problem and proposed a method U-LMA. First, a non-linear Cumulative Density Function (CDF) approximated by a hyperbolic tangent (tanh) was used to deal with the non-linearity associated with rigid and non-rigid motions and illumination variations. Then, the entropy of the proposed non-linear CDF was optimized using a customized LMA for BVP signal extraction, followed by maximum peak estimation for HR calculation. The performance of the proposed method was tested under three scenarios: constrained, motion, and illumination variations scenarios. High Pearson correlation coefficient values and smaller lower-upper statistical limits of bland-altman plots, justified the good performance of the U-LMA. Comparative analysis of U-LMA with undercomplete ICA with negentropy (U-neg) and other state-of-the-art methods demonstrated its best performance of U-LMA by achieving the lowest error and highest correlation values (0.01 significance level). Additionally, higher accuracy satisfying the clinically accepted error differences also justified its clinical relevance.

Index Terms—Blind Source Separation, Blood volume pulse extraction, Dimensionality Reduction, Heart rate estimation, Levenberg-Marquardt Optimization algorithm, Independent Component Analysis.

I. INTRODUCTION

The cardiovascular disease growth rate has been inclining faster worldwide in recent years [1]. Therefore, Heart Rate (HR) is a vital physiological parameter. It reflects the physiological, physical, and emotional state of an individual. Heart Rate monitoring has numerous applications in diverse fields like criminals' false statements

This manuscript is submitted to the journal on June 23, 2021. This work is supported by the LARSyS Project—UIDB/50009/2020, co-financed by Regional Development European Funds for the "Operational Programme Madeira 14–20"—Priority Axis 1 of the Autonomous Region of Madeira, number M1420-01-0145-FEDER-000002.

Ankit Gupta is working as a researcher at Madeira Interactive Technologies Institute (ITI/LARSyS/M-ITI), 9020-105 Funchal and a PhD student of Informatics Engineering, University of Madeira Caminho da Achada 10, Funchal 9000-208 Portugal email(ankit.gupta@iti.larsys.pt)

Antonio G Ravelo-García is with Madeira Interactive Technologies Institute (ITI/Larsys/M-ITI), 9020-105 Funchal, Portugal and also associate professor at Institute for Technological Development and Innovation in Communications, Universidad de Las Palmas de Gran Canaria, 35001 Las Palmas de Gran Canaria, Spain email(antonio.ravelo@ulpgc.es).

Fernando Morgado-Dias is with Madeira Interactive Technologies Institute (ITI/LARSyS/M-ITI), 9020-105 Funchal, Portugal and also associate professor at Faculty of Exact Sciences and Engineering, University of Madeira, Penteada Campus, 9020-105 Funchal, Portugal (email:morgado@uma.pt).

detection, neonatal vital signs monitoring, etc. Photoplethysmography (PPG) and electrocardiogram (ECG) are gold standards for measuring HR. However, both methods follow a contact-based approach which causes discomfort to individuals and is also unsuitable for prolonged and unobtrusive HR monitoring. Therefore, remote photoplethysmography (rPPG) has been an active area of research, which does not require any physical contact with the skin [2]–[11]. rPPG based methods is of utmost need where contact must be prevented in scenarios like sensitive skin, neonatal vital signs monitoring, or during sleep [10], [11]. The principle behind rPPG is the measurement of periodic variations due to the absorbance of hemoglobin in the blood and pulsation because of cardiac activity [12]–[14]. HR estimation using rPPG method is a three-step process: Region Of Interest (ROI) selection; Blood Volume Pulse (BVP) extraction; and average HR calculation. Most of the methods used the facial region as ROI, which has been further used for BVP extraction [15]. Capturing the facial region is predominantly done using an RGB camera because it allows less constrained conditions, unlike other methods like NIR, radar, or ultrasound systems [11]. The rPPG based methods use reflected light which is acquired through a photodetector, after light absorption by the skin tissues, arteries, veins, bones, and blood [16], [17]. This reflected light contains the blood volume variations along with various undesirable noise interferences [18] due to rigid and non-rigid motions [10] and illumination variations [14] which degrades the performance of HR estimation methods due to noisy BVP signal [6]. Furthermore, noise due to these artifacts easily dominates the relatively weaker strength of the resultant BVP signal [19]. The few frequently used BVP extraction methods used in the literature are Wavelet transforms [20], and Independent Component Analysis (ICA) [21], Ensemble Empirical Mode Decomposition (EEMD) [22]. Wavelet transform needs the selection of appropriate filtering coefficients at different decomposition levels [23], whereas EEMD requires selecting amplitude and noise frequency [24]. However, ICA starts with random initialization of unmixing matrix with just a single prerequisite of unmixing matrix dimension, depending on the number of independent components, which is comparatively trivial than other two. Additionally, among these methods, ICA is a common method for BVP signal extraction [25]. It considers BVP extraction as Blind Source Separation (BSS) problem, which deals with extracting the desired signal with no or limited information, from mixture signals. Moreover, Joint diagonalization approximation of matrices (JADE) a variant of ICA proposed by Poh et al. [2] has shown motion tolerance up to certain extent. In addition, Multichannel ICA proposed by Zhang et al. [26] conducted their experiments under low illumination as well. To the best of our knowledge, none of the ICA method based studies analysed the impact of motion and illumination variations effect simultaneously under constrained or natural conditions.

A general assumption about ICA based methods is that the number of independent signals is equal to the number of mixed signals. In other words, the signal constructed by each color channel (mixed-signal) results in an Independent Component (IC). This assumption requires analyzing each IC as a potential candidate for the BVP signal and requires apriori knowledge. Moreover, there is no de-

defined criterion for selecting the BVP signal from the independent components from different color channels [27]. Conventionally, BVP signal extraction includes selecting the component with the highest periodicity, which may result in selecting the wrong IC as BVP signal in case of periodic motions by the subjects [11]. Although, the majority of the studies selected the second IC for BVP signal extraction by discarding the 1st and 3rd IC corresponding to the red and blue color channels [6], [28], [29]. This results in loss of information present in the red and blue channels [28], which may be vital for HR estimation. Color subspace transformation methods like CHROM [11] and POS [13] were proposed to overcome this information loss, which employed orthonormal vector transformations to construct a raw signal for BVP extraction. The main drawback of these methods is the weights assigned to color channels which may degrade the BVP information [28]. Considering the above mentioned limitations, the current study proposed the BVP signal extraction as an under-complete problem. In other words, given three mixture signals corresponding to R, G, and B color channels, the task is to extract one IC which corresponds to the motion and illumination resistant BVP signal. A novel method combining undercomplete ICA [30] with customized Levenberg-Marquardt algorithm (LMA) [31], [32] was proposed for optimizing the unmixing matrix for BVP signal extraction without losing information from any color channel. Additionally, the proposed method eradicates the need for IC selection since the output is a single IC. The mean heart rate calculation was performed using power spectral density analysis using Fast Fourier Transform (FFT), post bandpass filtering. This study contributes to the extant literature with:

- A novel non-linear optimization function using a cumulative density function approximated by the hyperbolic tangent to deal with the non-linearity due to rigid and non-rigid motions and illumination variation artifacts.
- A customized LMA for optimizing the entropy of the proposed non-linear least square function, ensuring the statistical independence of the resultant BVP signal.
- A novel method constituting Undercomplete ICA with customized LMA (U-LMA), for artifacts free BVP signal extraction, followed by its performance analysis under three scenarios (database used): Constrained (VIPL-HR [33]), motion constituting rigid and non-rigid motions (UBFC-rPPG [34]) and illumination variations (COHFACE [35]).
- Testing the performance of U-LMA with negentropy based undercomplete ICA (U-neg) and other state-of-the-art methods to analyze the impact of non-linear function along with optimization using LMA under scenarios mentioned above.

II. RELATED WORK

The first attempt of heart rate estimation under normal light conditions was performed by Verkruysse et al. [4]. The study used the PPG signal extracted using the green color channel of the ROI selected from face videos acquired by a digital camera. PPG signal was then processed using filtering techniques and, subsequently, heart rate calculation. Poh et al. [29] extracted the PPG signal using JADE (ICA) from the R, G, and B signal traces acquired from the facial ROI captured using a webcam. Consequently, three ICs were extracted from each color channel, followed by selecting an appropriate IC as a PPG signal for heart rate estimation. This study was further extended by adding a temporal filtering component, which consisted of de-trending and signal smoothing using a moving average filter for better PPG signal extraction [2]. The above methods mainly used the green component of the ROI since it is considered to have maximum PPG information. The method given by Poh et al. [29]

used kurtosis optimization, which does not have descent statistical properties to support statistical independence among components. Gill et al. [36] addressed the problem of unsorted ICs of ICA, which is a challenging problem to select the appropriate independent component as BVP signal. They proposed constrained-ICA which uses negentropy as an optimization function, avoiding local minima convergence. It is important to note that negentropy possesses better statistical properties and symmetric decorrelation than kurtosis to ensure statistical independence. Considering the periodicity of the PPG signal, Macwan et al. proposed multi-objective optimization using Autocorrelation and ICA (MAICA), which constitutes negentropy and signal autocorrelation at different time lags, for BVP signal extraction. Kalman filter was also utilized to address motion and illumination artifacts [37]. A different approach was presented by De Haan et al. [10], utilizing the chrominance features of R, G, and B spectra. The approach extracted the two chrominance vectors, orthogonal to each other, from the RGB color spectra. RGB to chrominance vector transformations was done by empirically known coefficients. Finally, the ratio for the two vectors was used for PPG signal extraction. Furthermore, De Haan et al. [11], further improved this method by employing the absorption spectra changes of the RGB spectra for BVP signal extraction, where the Hulsbusch noise-free spectrum model was used to develop a normalized BVP vector. Combining the advantages of chrominance based signals and ICA, Song et al. introduced a semi-blind source separation method named Kernel Density ICA, which takes chrominance signals as input and extracts the PPG signal [21]. The kernel density ICA was used to address the problem of similar magnitude among illumination variation and PPG signal. In addition, the authors have also tested the effect of different shooting distances and image resolution for PPG signal extraction. Realizing the need to add more channels for accurate BVP extraction, McDuff et al. [38] used a five-band lens camera to extract the orange and cyan spectra along with three traditional color spectra. This enables monitoring the absorption of light differences between Hb and HbO₂ by creating a bigger overlap between cyan, orange, and green spectra, for accurate heart rate estimation, using the approach presented by Poh et al. [2]. A similar approach was presented by Gupta et al. [9] in which magenta color filter, the thermal camera was utilized with RGB camera, addressing the illumination variations and proposed red and green channel with thermal imaging can better estimate HR. FastICA was used for BVP signal separation, which uses negentropy to maximize statistical independence. Yan et al. [5] proposed an approach of using a weighted average of RGB spectra of the selected ROI for improving the Signal-to-Noise Ratio (SNR) followed by denoising the signal using wavelet transform for PPG signal extraction. Kumar et al. [7] used a monochrome camera for extracting the green spectra followed by its weighted average using varied ROI combinations. Purcuhe et al. [3] analyzed the effect of using different facial region parts on HR estimation and divided the facial region into three ROIs: forehead, eye, and nose surrounding area, and mouth area. The heart rate was computed using each ROIs, applying ICA for PPG signal extraction, followed by Fourier transform. The limitations of the existing state-of-art methods are multi-fold. First, due to the unordering of independent components, selecting the appropriate IC containing BVP information is challenging. Second, existing statistical dependence metrics do not consider the non-linearity associated with the PPG extraction problem. Third, adding further channels for HR estimation enhances the complexity of the problem by increasing its dimensionality with an added effect due to different types of motions and illumination variations. Fourth, color difference equations proposed in color subspace transformation methods have associated coefficients with the color channels, which may affect PPG information. Finally, the

semi-blind source separation may need additional information about PPG signal statistical properties for accurate signal extraction. The problem of unordering of independent components is resolved by assuming the BVP extraction problem as undercomplete, which deals with taking raw RGB traces as input and a single BVP signal as output. To ensure the consideration of non-linearity with statistical independence, a non-linear optimization function (CDF approximated by tanh) is proposed. The presented method is capable of dealing with the associated non-linearity due to artifacts with three channels of the RGB color space. The proposed method does not assign the weights to color channels which ensures considering each color channel independently to contribute to the BVP signal. Finally, U-LMA does not require apriori information to extract the BVP signal from raw signals. Hence the proposed method manages to overcome all the limitations pointed out by the existing literature.

III. PROPOSED METHODOLOGY

The proposed method takes a face video recorded under ambient light conditions as input and estimates the mean heart rate. It calculates the heart rate via a three-step procedure: ROI selection, BVP signal extraction, and Heart Rate Estimation. This section discusses the detailed information of the proposed U-LMA for heart rate estimation, explaining all the constituting steps in the following subsections. A detailed flow diagram for the proposed approach is shown in Fig. 1.

A. ROI Selection and Signal Construction

ROI selection deals with identifying the face using the Viola-Jones face detector [39], followed by skin segmentation. The skin was segmented using Cb and Cr components of the YCbCr color model using the parameters proposed by Mahmoud [40]. Subsequently, a spatial averaging operation for each channel was performed on each image frame of the video. A detrending process was also applied to remove slow non-stationary drifts in the signal using the approach by Tarvainen *et al.* [41]. Finally, an overlapping moving window operation was applied to each channel for constructing raw signals. Fig. 2 depicts the workflow for ROI selection and raw signal construction.

B. BVP signal extraction

The raw signals are further refined to extract the BVP signal for HR estimation. Following the standard ICA annotations, the raw signals are considered as mixture signals containing BVP and other information along with noise interferences due to motion and illumination artifacts. The goal is to extract the BVP signal as one of the ICs from them [42]. Ideally, the 2nd IC from the green channel is selected as a BVP signal, while other ICs are discarded, which may contain BVP information. Therefore, this study defines a BVP extraction as an undercomplete problem that takes three mixture signals and extracts a single IC, consisting of BVP information from all three channels [43]. This problem is solved using the proposed U-LMA, which uses a CDF of the raw signals approximated by tanh, followed by its optimization using the customized LMA proposed in this study. The proposed approach is motivated by the work of Porri *et al.* [30] for signal separation and dimensionality reduction. The difference lies in the context, optimization algorithm, and termination condition. The present work uses a customized version of LMA for optimizing the unmixing matrix W and the number of iterations as the only termination condition. The reason behind choosing the number of iterations as a termination condition is to consider the absence of BVP signal information. Since there is no reference BVP signal, it

is not possible to check the correlation of the resultant signal with the original BVP signal, so the correlation criterion is not taken into consideration. LMA is chosen because initial values of W due to random initialization may or may not lie near the desired solution. Both conditions need separate ways of approaching the solution. This optimization algorithm will allow the solution to efficiently converge to the desired values of W in both conditions [44]. The details of the customized version of LMA will be discussed in the following subsection.

Mathematically, $x(t) \in R^{3 \times t}$ and $y(t) \in R^{3 \times t}$ are mixed signals and IC matrices, respectively. $x(t)$ consist of three mixed signals $x_1(t)$, $x_2(t)$, and $x_3(t)$, corresponding to the color channels, whereas $y(t)$ comprises 3 ICs $y_1(t)$, $y_2(t)$ and $y_3(t)$. ICA assumes that mixture signals are linear combinations of ICs:

$$x(t) = Ay(t) \quad (1)$$

where A is called the mixing matrix, and the goal is to estimate unmixing matrix W , which, when multiplied by mixed signals, gives the independent components as:

$$y(t) = Wx(t) \quad (2)$$

From (1), the matrix W is an approximation of A^{-1} . For U-LMA, W will be a rectangular matrix since the number of independent components is less than the number of mixtures, unlike the standard ICA, where W is a square matrix. As statistically independent signals having CDF denoted by σ , have maximum entropy [30], W can be determined by maximizing the entropy of the CDF, ensuring the statistical independence of ICs [45]. The entropy $H(y)$ of CDF for the independent component y is mathematically defined as:

$$H(y) = H(x) + E[\log \sigma'] \quad (3)$$

Where $H(y)$ is the entropy of the IC, and σ' defines the derivative of σ . It is important to note that in (3), $x(t)$ and $y(t)$ is written as x and y for brevity. The approximation of $H(x)$ to multidimensional Gaussian for a single IC can be written as:

$$H(x) = 0.5 * \log(c) + 0.5 * (1 + \log(2 * \pi)) \quad (4)$$

For $H(y)$ maximization, $H(x)$ can be reduced to $0.5 * \log(c)$ where c is written as $E[xx^T] = WSW^T$ and S is the covariance of x . Considering the reduced form of $H(x)$ and approximating $\sigma = \tanh$, a new function can be deduced as:

$$h(W) = 0.5 * \log|WSW^T| + E[\log(\sec^2(y))] \quad (5)$$

(5) is used as a criterion for extracting the IC from the mixed signals. Differentiating (5) with respect to W_{ij} yields the ∇ vector for updating the unmixing matrix W , given by:

$$\nabla_{(w^h)} = (SW^T / WSW^T)^T - 2E[y^T x] \quad (6)$$

where $(SW^T / WSW^T)^T$ is the pseudo inverse of W with respect to S , positive definite matrix.

The proposed algorithm approximates the unmixing matrix W using the customised LMA algorithm by maximizing (5) and updating the matrix W using (6).

C. Customised Levenberg Marquardt Algorithm (LMA)

LMA is a widely used optimization algorithm that is used to find the global minima for non-linear least square functions with faster convergence property [46] and dual algorithmic adaptability depending on the current solution. In other words, it can be considered as a combination of gradient descent and the gauss-newton method depending on the proximity of the current solution to the global minima. The present study customized the conventional LMA

Fig. 1. The workflow of the proposed method for Heart Rate Estimation.

Fig. 2. ROI and raw signal construction.

by introducing the entropy of cumulative density function (CDF) approximated by a hyperbolic tangent $\tanh$ defined in [5] as an optimization function, followed by its maximization for statistical independence. The advantage of approximation using $\tanh$ lies in the fact that it introduces processing with higher-order statistics to deal with the non-linearity associated with the optimization problem [47]. The workflow of updating W using the proposed method for BVP signal extraction is presented in fig. 3. The process starts with the random initialization of W followed by calculating the entropy of CDF and subsequently validating the convergence condition. If the convergence condition is reached, then W_{curr} , the W at the current iteration is returned as output; otherwise, the Jacobian, Hessian, and a diagonal matrix are calculated for updating W . The diagonal matrix consists of the highest value of the Jacobian achieved till the last iteration performed. The cost function, i.e., entropy is calculated before and after updating W and is compared. The damping parameter λ is increased if entropy decreases, followed by calculating the cost function again after updating W using W_{prev} from the previous iteration until the entropy is increased. If there is a rise in the entropy value after updating W , is decreased until the convergence condition is reached. The raw signal is multiplied with W to extract the BVP signal for heart rate estimation as:

$$BVP(t) = W * x(t) \quad (7)$$

D. Heart Rate Estimation

The BVP signal extracted through U-LMA is processed using a bandpass filter with cut-off frequencies 0.7-4.0 Hz, respectively, corresponding to 42-240 beats per minute (bpm). Finally, the Fast Fourier Transform (FFT) is applied for analyzing the power spectral density for maximum peak estimation, which is then used for heart rate calculation by taking its $\log_{10}$ and multiplied by 60.

IV. RESULT ANALYSIS

The proposed method was tested under constrained and natural conditions using three benchmark databases: VIPL-HR, UBFC-rPPG, and COHFACE. VIPL-HR database was used for performance validation under constrained conditions, UBFC-rPPG tested the method's performance for rigid and non-rigid motions, and illumination variations effect on the proposed method was tested using COHFACE database. It is necessary to analyze the method's performance under constrained and unconstrained conditions since testing a method under constrained conditions gives insight into its steps and their precision, whereas unconstrained conditions test its robustness. The detailed description of the databases used for this study is presented in the following subsections, with a summary in Table I.

A. Databases

1) **VIPL-HR**: The database consists of 2378 videos with visible light spectra and 752 videos with Near Infrared (NIR) spectra from 107 subjects (79 males and 28 females), aged between 22 and 41 years. Nine variable scenarios were considered for sample collection. For each scenario, the samples were collected using digital cameras of different frame rates and NIR cameras. Each database sample comprises a 30 second (s) subject video, BVP signal, HR, and SpO2 values [33]. This study has used the subset of videos that corresponds to a frame rate of 30 frames per second (fps) with 1920×1080 pixels resolution, covering the HR range between 47 and 100 beats per minute (bpm). The ground truth heart rate was acquired using a CMS60C pulse oximeter synchronized with the subject's video. This resulted in 107 samples (one from each individual) that were analyzed for testing U-LMA. One sample (p41) was left out due to insufficient video length (18 s) smaller than the processing window (25s).

2) **UBFC-rPPG**: UBFC-rPPG is a publicly available database consisting of 50 video samples, synchronized with a CMS50E pulse oximeter (sampling rate 60 Hz). The videos are available in uncompressed form with a resolution of 640×480 pixels at 30 fps, covering the HR range between 63 and 112 bpm. Each video is 2 minutes (min) long in which participants were asked to sit facing the camera and play a mathematical game that causes abrupt rise and fall in HR value promoting rigid and non-rigid movements [21]. The

Fig. 3. Customized LMA for entropy maximization.

Fig. 4. Face detection and skin segmentation.

database did not provide age-specific information. All videos were used to test the performance of the proposed method [34].

3) **COHFACE Database:** COHFACE dataset is a collection of 160 videos with physiological recordings for HR and respiration rate from 40 healthy subjects with a mean age of 35.6 years. The dataset constitutes 60 seconds (s) videos from 12 females and 28 males covering the HR range between 54 to 97 bpm. The videos were recorded with a resolution of 640x480 pixels at 20 fps with the synchronized BVP measurements using BVP model SA9308M, belt model SA9311M (sampling rate 256 Hz) [35]. The dataset offers constrained and challenging natural conditions, especially in terms of illumination variations over the facial region. Therefore, this study tests the performance of the proposed method using natural conditions video samples.

B. Performance Metrics

For each video sample, the estimated heart rate was compared to the corresponding heart rate value acquired by the ground truth sensor (ECG or PPG). The degree of differences between the actual and the estimated readings were analyzed and summarized using the Bland-Altman plot [48] with upper and lower statistical limits as $\pm 1.96 \times \text{standard deviation (STD)}$, mean error, standard deviation,

TABLE I
DATABASE SUMMARY USED FOR THIS STUDY.

Database Features	VIPL-HR	UBFC-rPPG	COHFACE
No. of subjects	107	50	40
Video Resolution	1920 × 1080	640 × 480	640 × 480
Frame rate	30	30	20
Video Duration	30 seconds	90 seconds	60 seconds
Ground truth sensor	CMS60C	CMS50E	SA9308, SA9311M
Shooting Distance	1 meter	1 meter	–
HR range	47-100 bpm	63-112 bpm	54-97 bpm
Considered Artifacts	Constrained*	Rigid and Non-rigid motion	Illumination

*Constrained: Conditions with minimum permissible motion and illumination.

Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE), Pearson correlation at 0.01 significance level [49], regression analysis at 95% confidence intervals and accuracy defined by clinically accepted error differences between ground truth and estimated HR values, i.e., $\pm 5 \text{ bpm}$.

C. ROI selection and Signal construction

Before this step, the RGB image frames of the video were pre-processed for adjusting the pixel intensities, using gamma correction. ROI selection deals with face detection followed by segmenting the skin in the YCbCr color space in which Y represents the luminance with pixel intensity ranges between 16 and 235, while for chrominance blue (Cb) and chrominance red (Cr) components, the pixel values lie between 16 and 240. The thresholds used for Cb and Cr components are in the range of 77 to 127 and 133 to 173, respectively, with no thresholding for the luminance component [40]. Finally, the ROI is selected as 70% height and 60% width of the segmented skin region. Fig. 4 depicts the results of the face detection and skin segmentation process. For de-trending the temporal RGB traces, the regularization parameter was set to an empirically defined value, i.e., 10. The raw signal was constructed using a moving window operation with a 96% overlap (1-sec increment) for each color channel.

D. BVP Signal Extraction and HR estimation

BVP signal extraction was performed using U-LMA. The unmixing matrix W was first initialized randomly, and the values of damping parameter λ were set empirically as 5 and 2.5, respectively, as a part of standard LMA initialization. Subsequently, the customized LMA was employed to maximize the entropy of the proposed non-linear CDF optimization function using 1000 iterative steps because none of the video samples took this many iterations for convergence to global maxima. Finally, the optimized unmixing matrix W was used to extract the BVP signal. Finally, FFT was applied to the resultant signal, followed by calculating the $\log_{10}$ value of peak maxima and multiplying it with 60 for mean heart rate estimation.

E. Performance Analysis

As mentioned before, the performance of the proposed U-LMA method is analyzed considering three scenarios: constrained, Rigid and non-rigid motions, and illumination variations. Table I specifies that the VIPL-HR database has been used for performance testing under the constrained or stable scenario, UBFC-rPPG for testing its robustness in rigid and non-rigid motions and, COHFACE in illumination variations scenarios. For each scenario, Bland-Altman and regression plots will be presented and analyzed, taking into account the respective measured parameters for the plots.

Fig. 5. Bland-Altman Plot for the constrained scenario.

Fig. 6. Regression plot for the constrained scenario.

1) Constrained Scenario: For the constrained (VIPL-HR database) scenario, the subjects were asked to sit in the still position at a distance of one meter away from the camera with the ceiling lamp on. The Bland-Altman and regression plot for the constrained scenario are shown in Fig. 6. The mean bias for the proposed method is -0.35 bpm which is near to zero error difference between ground truth and estimated values. In other words, on average, the heart rate estimated by the algorithm measures 0.35 bpm less than the conventional BVP sensor used. Furthermore, the majority of the differences lie within upper (7.3651) and lower (-8.0632) level statistical limits, which justifies the good performance of the method. Additionally, the Pearson correlation value denoted by r for this scenario is 0.92, which confirms a higher correlation among ground truth and estimated values. Therefore, Bland-Altman's analysis and high correlation value justify the good performance of the proposed method under the constrained scenario.

2) Rigid and Non-rigid motions Scenario: Performance analysis under this scenario was performed using all video samples of the UBFC-rPPG database. The videos were collected while subjects were playing a time-sensitive mathematical game which causes an abrupt increase or decrease in HR values along with involuntary head movements due to the subject's action. The samples have a certain amount of illumination variations, too, since the video samples were collected considering natural conditions. Fig. 7 depicts the Bland-Altman, and regression plots. As expected, the mean bias for this scenario is 1.84 bpm due to the presence of motion artifacts which means, the U-LMA predicts 1.84 bpm more than the traditional BVP. Consequently, the

Fig. 7. Bland-Altman Plot for Rigid and Non-rigid motion scenario.

upper statistical limit of the Bland-Altman plot is slightly greater than 10 bpm; however, the method achieved a far lower statistical limit -6.6005. All of the data points lie between the upper and lower statistical limits. Interestingly, the ground truth and estimated HR values showed a very high correlation (0.94), despite having a higher overall mean difference. Hence, the Bland-Altman analysis and regression plot confirm the proposed method's effectiveness under challenging motion conditions and handling abrupt rise and fall of HR values, considering the mean of the HR values during the interval.

3) Illumination Variations Scenario: COHFACE database is utilized to assess the ability of U-LMA under different illumination scenarios. It is worth noting that the samples for the database have motion artifacts too, but with the predominance of uneven illumination distribution over the face due to ambient light. The performance analysis using Bland-Altman and regression plot is presented in Fig. 8. The mean bias achieved with the illumination scenario is -0.85 with lower and upper statistical limits -9.7546 and 8.0546, respectively. Similar to the other scenarios, almost all the differences lie within the statistical limits. Furthermore, the Pearson correlation achieved under this scenario is 0.92, which confirms a good correlation between estimated and ground truth values. Both plots and their measured parameters prove the efficiency of U-LMA for HR estimation using illumination variant facial videos.

F. Comparative analysis

The proposed method works on the principle of rPPG; hence its performance is compared to GREEN [4], ICA-Poh [29], CHROM [10], POS [13] as state-of-art methods, which are rPPG based methods. GREEN, ICA-Poh works on utilizing the green channel of the RGB image model, whereas CHROM and POS work on the principle of orthonormal color subspace transformations. However, BCG [50] is also included in the comparative analysis since it works on tracking the periodic movements of the head. As the effect of both types of motion on the proposed method is also tested in this study, it is worth including BCG as one of the state-of-art methods for performance comparison. To assess the performance of the proposed non-linear cumulative density function approximated tanh and customized LMA, another variant of the proposed undercomplete analysis (U-neg) is also introduced, which utilizes the differential entropy or negentropy as an objective function. This objective function is optimized using the standard ICA procedure given by Hyvärinen et al. [42]. The state-of-the-art methods were implemented using standard implementation using iPhys toolbox by Mc Duff et al. [51]. All methods were

Fig. 8. Regression plot for Rigid and Non-rigid motion scenario.

Fig. 9. Bland-Altman plot for Illumination variations scenario.

Fig. 10. Regression plot for Illumination variations scenario..

tested under three scenarios, as explained in previous subsections. In other words, the video samples from all three databases were tested for all the methods used for comparative analysis by calculating RMSE, MAPE, mean error, standard deviation, accuracy, and Pearson correlation values.

1) *Constrained scenario*: Similar to section 4.5.1, the videos from the VIPL-HR database were used for comparative analysis. The performance metrics for the comparative analysis are presented in Table II. Among all methods, BCG and POS were the worst

TABLE II
PERFORMANCE METRICS FOR THE METHODS UNDER CONSTRAINED SCENARIO.

Methods	RMSE (bpm)	MAPE (%)	SD* (bpm)	Mean (bpm)	Accuracy (%)	r*
Green	23.74	27.24	13.85	19.33	4.72	0.10
ICA - Poh	21.23	24.23	13.03	16.81	3.77	0.23
CHROM	17.54	20.47	17.63	-0.15	16.98	0.18
POS	24.69	27.13	21.79	11.82	3.77	-0.28
BCG	24.89	27.82	12.84	21.36	7.55	-0.04
UICA without LMA (U-neg)	17.03	17.21	13.84	10.01	23.58	0.47
UICA with LMA(U-LMA)	3.85	4.07	3.86	-0.35	84.91	0.92

SD*: Standard Deviation; r*: Pearson correlation is calculated at the 0.01 significance level; Accuracy is defined as the percentage of achieving the error difference with ± 5 bpm.

performing methods for this scenario, despite minimal motion and illumination variation artifacts. On the other hand, CHROM, GREEN, and ICA-Poh methods exhibited almost the same performance. One key observation during this analysis is the lowest mean error achieved by CHROM, showing its effectiveness under constrained conditions. Nevertheless, all state of art methods failed to achieve decent performance. This is because most of the methods have created their self-created databases, which were not shared publicly. Additionally, the original papers lack information about the conditions considered during data collection, which would potentially be the reason for their poor performance. U-neg achieved relatively better performance than other state-of-the-art methods depicting its better accuracy and reliability. However, the method did not achieve convincing results. On the other hand, the proposed U-LMA achieved the best results than any other method used in the study, justifying its performance under minimal motion and illumination artifacts. Moreover, the highest accuracy with clinically accepted error difference is achieved by U-LMA.

2) *Rigid and Non-rigid motions Scenario*: The video samples from the UBFC-rPPG database were used to assess the effect of rigid and non-rigid motions on HR estimation. Table III presents the performance metrics for all the compared methods. Similar to the constrained scenario, POS performed worst for the motion scenario too. Although the original study tested its performance under a fitness scenario in which subjects were asked to do fitness exercises during video recording, a fluorescent lamp was also used to avoid illumination variations. In other words, a single condition was tested at a time, whereas the video samples used for this study have both illumination and motion artifacts at the same time, which could be a potential reason for its worst performance. Additionally, the number of subjects for the task was just 5, which is 10 times lesser than this study. BCG was expected to perform well in these settings but did not perform well due to higher dominance of non-periodic movements. ICA-Poh and GREEN have shown almost the same performance in terms of all metrics. Interestingly, GREEN achieved relatively better accuracy than constrained scenarios, thereby showing its resistance for motion artifacts. Finally, U-neg has performed far better than any other state-of-art method (except CHROM). Although RMSE, MAPE, mean error, and error standard deviation of U-neg was reduced, the accuracy was degraded in the motion scenario, as expected. Due to the ability to handle motion artifacts, CHROM was the second best method for HR estimation. Nevertheless, the proposed U-LMA outperformed all the methods, reporting the minimum value of errors and highest accuracy and Pearson correlation, justifying its best performance than other methods.

TABLE III
PERFORMANCE METRICS OF THE METHODS UNDER RIGID AND NON-RIGID MOTION SCENARIO

Methods	RMSE (bpm)	MAPE (%)	SD* (bpm)	Mean (bpm)	Accuracy (%)	r*
Green	37.61	34.88	19.14	32.49	10	0.34
ICA – Poh	34.87	31.79	19.64	28.94	10	0.27
CHROM	14.43	13.57	11.03	9.43	18	0.73
POS	44.06	38.15	33.43	29.08	2	-0.08
BCG	36.08	33.75	16.90	31.97	8	0.03
UICA without	14.97	12.86	11.27	9.98	22	0.59
LMA(U-neg)						
UICA with LMA(U-LMA)	4.57	4.00	4.22	1.84	78	0.94

SD*: Standard Deviation; r*: Pearson correlation is calculated at the 0.01 significance level; Accuracy is defined as the percentage of achieving the error difference with ± 5 bpm.

TABLE IV
PERFORMANCE METRICS OF THE METHODS UNDER ILLUMINATION VARIATIONS SCENARIO.

Methods	RMSE (bpm)	MAPE (%)	SD* (bpm)	Mean (bpm)	Accuracy (%)	r*
Green	22.76	26.93	.39	1 0.61	2.5	-0.18
ICA – Poh	24.88	28.52	17.84	17.56	7.5	-0.28
CHROM	19.49	22.48	10.54	16.47	15	0.18
POS	26.28	30.71	12.63	23.14	5	-0.18
BCG	25.53	27.00	15.72	20.27	22.50	0.19
UICA without	11.13	13.03	10.76	3.33	25	0.60
LMA(U-neg)						
UICA with LMA(U-LMA)	4.48	5.16	4.45	-0.85	80	0.92

SD*: Standard Deviation; r*: Pearson correlation is calculated at the 0.01 significance level; Accuracy is defined as the percentage of achieving the error difference with ± 5 bpm.

3) Illumination Variation Scenario: The effect of illumination variations on the methods used for this study was evaluated using COHFACE database. State-of-the-art methods GREEN, ICA-Poh and, POS, have shown a negative correlation for this scenario, indicating their susceptibility for uneven illumination distribution. CHROM and BCG, on the other hand, performed better than these three methods in terms of accuracy. CHROM achieved better RMSE, MAPE, and standard deviation, whereas BCG achieved relatively better accuracy than other state-of-the-art methods. The lower performance of state-of-the-art methods could be due to the constrained conditions maintained during video acquisition and the relatively lower number of subjects for the analysis. For instance, the GREEN method used ambient light, and fluorescent light combination and POS uses at most 15 subjects with different kinds of fluorescent lamps near the face for checking the illumination effect, rather than using natural light illumination conditions. The proposed U-neg and U-LMA have performed better than the state-of-the-art methods achieving better performance metrics. Furthermore, U-LMA achieved the lowest error values and highest accuracy and correlation values showing its superiority for HR estimation under illumination variations scenario.

4) RMSE analysis: The RMSE for the HR estimation has been predominantly analyzed in most of the studies conducted so far. It is calculated as the square root of the averaged squared error differences among different samples which provides the overall distribution of errors. Fig. 9 depicts the box and whisker plot of RMSE for analyzing the RMSE distribution among methods based on databases and vice-versa. These RMSE plots provide a deep insight into the performance of state-of-the-art and proposed methods for all the databases used in the study. UBFC-rPPG dataset was challenging for all methods used

Fig. 11. RMSE Box and whisker plot for the databases used.

Fig. 12. RMSE Box and whisker plot for the databases used.

in the study due to its realistic conditions considered during video acquisition, whereas the performance with VIPL-HR is the best due to constrained conditions. The COHFACE dataset is also challenging in terms of illumination variations throughout the samples. Among all, the worst performing method is POS, which can be easily identified using RMSE plots. The reason is that the original POS study considered a fairly small number of samples for the study and tackled one condition at a time and constant spectrum with a single light source [13], [52]. Moreover, except, VIPL-HR all other databases have collected the data maintaining realistic conditions than POS original study. Consequently, POS could not perform better in any of the scenarios analyzed in this study. It is clear from the RMSE plots that GREEN, ICA-Poh, and BCG performed equally. However, CHROM has shown far better performance than any of them in all scenarios due to its ability to handle constrained and natural conditions in a better way. U-neg performed relatively better than any state-of-the-art method, except rigid and non-rigid motions Scenario scenario, in which CHROM achieved better performance than U-neg, although the difference in performance was not significant. U-LMA performed far better than U-neg and other state-of-the-art methods either in terms of databases or state-of-the-art methods comparison. The proposed methods performed relatively better due to the proposed undercomplete ICA, which ensures better PPG information extraction from all three channels of RGB color space. Furthermore, U-neg used negentropy (differential entropy) for optimizing W with standard ICA implementation, but the experiments conducted during the study revealed that the entropy of CDF approximated by tanh yielded better statistical independence than negentropy. Additionally, the lowest RMSE ranges by U-LMA when compared to U-neg in all scenarios were due to better optimization of unmixing matrix W using customized LMA proposed in this work.

V. DISCUSSIONS

U-LMA outperformed all methods in the various scenarios considered for the study due to its following components: Non-linear objective function (entropy of the CDF approximated by tanh), efficient optimization algorithm (customized LMA). The non-linear function provided an advantage to counter the non-linearity associated with different types of motions and illumination variations. Customized LMA was able to find the global maxima for the entropy of the non-linear function in all samples with empirically evaluated damping parameter values. The study introduced U-neg which included undercomplete ICA with negentropy as an optimization function which allowed us to test the effectiveness of the U-LMA components. Kurtosis was not used because it is previously known that negentropy exhibits better statistical properties than kurtosis [42]. Moreover, JADE used kurtosis, whose performance was surpassed by U-neg, as shown in the previous section. The performance comparison between both methods confirmed that processing using higher-order statistics enhanced the performance of the proposed method. On the other hand, undercomplete ICA prevented the loss of BVP information by considering all the color channels, justifying the vitality of BVP information in red and blue channels for accurate HR estimation.

Following Bland-Altman and regression analysis, the mean bias of the illumination scenario is slightly far away from zero error difference when compared to the constrained scenario, depicting the effect of illumination artifacts on HR estimation. As expected, the mean bias for the motion scenario is comparatively larger than the illumination and constrained scenario, respectively, due to the presence of rigid and non-rigid motions of the subjects in the video samples. Furthermore, the Pearson correlation value for the UBFC-rPPG database is higher by 0.2 bpm than other databases, which could be due to the utilization of uncompressed videos for HR estimation.

The samples used under motion and constrained scenarios have the same sampling rate (30 fps), whereas the samples tested under illumination scenarios have a relatively lower frame rate (20 fps). It is worth pointing that low frame rate results in missing blood volume variations, which can be captured using a higher sampling rate. The proposed U-LMA worked equally well for the videos captured with different framerates. Hence, blood volume variations information loss due to compressed or low sampling rate videos did not significantly affect the proposed method for HR estimation.

Although the proposed method managed to confine most of the differences between ground truth and estimated HR values within the statistical limits, this did not justify its clinical relevance since none of the statistical limits were found to match with the clinically accepted error differences. Therefore, a new metric accuracy (error difference; 5 bpm) was defined to test the clinical relevance. The proposed U-LMA achieved sufficiently better accuracy justifying its clinical relevance as compared to other methods.

Nevertheless, U-LMA also possesses certain limitations which need to be addressed in the near future. First, although the effect of video compression and frame rates were tested, the effect of different camera-subject distances on the HR estimation was not addressed. The effect of different shooting distances could be tested on HR estimations by creating such databases followed by analyzing them. Second, although U-LMA was tested for the rigid and non-rigid motions, scenarios with periodic motions such as walking running, treadmill exercises, etc., or unconscious motions during sleep conditions were not addressed, which is a future direction of the study. Third, the proposed method works well with various illumination variation conditions, but the effect of zero luminance or dark conditions was not addressed in this work, due to non-availability of reliable databases. The majority of the limitations for

the proposed work are due to the limited or unavailability of sufficient benchmark databases. Finally, the proposed method only estimates HR, whereas other physiological parameters estimations like blood oxygen saturation, blood pressure, respiratory rate using BVP signals will be an interesting approach for future studies.

VI. CONCLUSION

This work addressed the BVP extraction as an undercomplete problem and proposed U-LMA. Taking into account the non-linearity due to motion and illumination artifacts, a novel entropy based non-linear function was proposed. The proposed non-linear function proved its effectiveness by addressing both type of artifacts. Furthermore, the non-linear function was optimized using the proposed customized LMA for entropy maximization, maximum statistical dependence due to better optimization of unmixing matrix W . The optimization using customized LMA, also aimed at reduced the effect of motion illumination artifacts. Additionally, the proposed method eradicates the need for IC selection and preserves the maximum possible BVP information from all channels of the RGB color space. Performance analysis for U-LMA was performed by comparing it under three scenarios: constrained, motion (rigid and non-rigid) and, illumination variations scenario. The Bland-Altman analysis and regression plots proved the efficacy of the proposed method under all scenarios. Additionally, to check the influence of customized LMA and the entropy based CDF function, U-neg with negentropy optimization function using standard ICA implementation was also tested. The results demonstrated the effectiveness of the proposed non-linear function and LMA by achieving comparatively better results than U-neg. In addition, U-LMA and U-neg, with other state-of-the-art methods, were also compared under the aforementioned scenarios using RMSE, MAPE, mean error, standard deviation, and accuracy (error difference ± 5 bpm) Pearson correlation values. The results depicted the dominance of U-LMA in all scenarios by reporting comparatively lowest error values and high accuracy and Pearson correlation values under 0.01 significance level. Higher accuracy values also suggest the applicability of the proposed method under clinical conditions.

VII. ACKNOWLEDGMENT

This work is supported by the LARSyS Project–UIDB/50009/2020, co-financed by Regional Development European Funds for the "Operational Programme Madeira 14–20"—Priority Axis 1 of the Autonomous Region of Madeira, number M1420-01-0145-FEDER-000002. The authors are also thankful to anonymous reviewers for their suggestions and recommendations to improve this manuscript for journal publication.

REFERENCES

- [1] S. Cook, M. Togni, M. C. Schaub, P. Wenaweser, and O. M. Hess, "High heart rate: a cardiovascular risk factor?," *European heart journal*, vol. 27, no. 20, pp. 2387–2393, 2006.
- [2] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in non-contact, multiparameter physiological measurements using a webcam," *IEEE transactions on biomedical engineering*, vol. 58, no. 1, pp. 7–11, 2010.
- [3] T. Pursche, J. Krajewski, and R. Moeller, "Video-based heart rate measurement from human faces," in *2012 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 544–545, IEEE, 2012.
- [4] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Optics express*, vol. 16, no. 26, pp. 21434–21445, 2008.
- [5] Y. Yan, X. Ma, L. Yao, and J. Ouyang, "Noncontact measurement of heart rate using facial video illuminated under natural light and signal weighted analysis," *Bio-medical materials and engineering*, vol. 26, no. s1, pp. S903–S909, 2015.

- [6] X. Li, J. Chen, G. Zhao, and M. Pietikainen, "Remote heart rate measurement from face videos under realistic situations," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4264–4271, 2014.
- [7] M. Kumar, A. Veeraraghavan, and A. Sabharwal, "Distanceppg: Robust non-contact vital signs monitoring using a camera," *Biomedical optics express*, vol. 6, no. 5, pp. 1565–1588, 2015.
- [8] J. Kranjec, S. Beguš, G. Geršak, and J. Drnovšek, "Non-contact heart rate and heart rate variability measurements: A review," *Biomedical signal processing and control*, vol. 13, pp. 102–112, 2014.
- [9] O. Gupta, D. McDuff, and R. Raskar, "Real-time physiological measurement and visualization using a synchronized multi-camera system," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 46–53, 2016.
- [10] G. De Haan and V. Jeanne, "Robust pulse rate from chrominance-based rppg," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, pp. 2878–2886, 2013.
- [11] G. De Haan and A. Van Leest, "Improved motion robustness of remote-ppg by using the blood volume pulse signature," *Physiological measurement*, vol. 35, no. 9, p. 1913, 2014.
- [12] X. Chen, J. Cheng, R. Song, Y. Liu, R. Ward, and Z. J. Wang, "Video-based heart rate measurement: Recent advances and future prospects," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 10, pp. 3600–3615, 2018.
- [13] W. Wang, A. C. den Brinker, S. Stuijk, and G. De Haan, "Algorithmic principles of remote ppg," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1479–1491, 2016.
- [14] D.-Y. Chen, J.-J. Wang, K.-Y. Lin, H.-H. Chang, H.-K. Wu, Y.-S. Chen, and S.-Y. Lee, "Image sensor-based heart rate evaluation from face reflectance using hilbert–huang transform," *IEEE Sensors Journal*, vol. 15, no. 1, pp. 618–627, 2014.
- [15] M. A. Hassan, G. S. Malik, N. Saad, B. Karasfi, Y. S. Ali, and D. Fofi, "Optimal source selection for image photoplethysmography," in *2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings*, pp. 1–5, IEEE, 2016.
- [16] T. Y. Abay and P. A. Kyriacou, "Reflectance photoplethysmography as noninvasive monitoring of tissue blood perfusion," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 9, pp. 2187–2195, 2015.
- [17] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological measurement*, vol. 28, no. 3, p. R1, 2007.
- [18] S. Kado, Y. Monno, K. Yoshizaki, M. Tanaka, and M. Okutomi, "Spatial-spectral-temporal fusion for remote heart rate estimation," *IEEE Sensors Journal*, vol. 20, no. 19, pp. 11688–11697, 2020.
- [19] R. Song, S. Zhang, C. Li, Y. Zhang, J. Cheng, and X. Chen, "Heart rate estimation from facial videos using a spatiotemporal representation with convolutional neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 7411–7421, 2020.
- [20] F. Bouesfah, C. Maaoui, and A. Pruski, "Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 568–574, 2013.
- [21] R. Song, S. Zhang, J. Cheng, C. Li, and X. Chen, "New insights on super-high resolution for video-based heart rate estimation with a semi-blind source separation method," *Computers in biology and medicine*, vol. 116, p. 103535, 2020.
- [22] J. Cheng, X. Chen, L. Xu, and Z. J. Wang, "Illumination variation-resistant video-based heart rate measurement using joint blind source separation and ensemble empirical mode decomposition," *IEEE journal of biomedical and health informatics*, vol. 21, no. 5, pp. 1422–1433, 2016.
- [23] U. Bal, "Non-contact estimation of heart rate and oxygen saturation using ambient light," *Biomedical optics express*, vol. 6, no. 1, pp. 86–97, 2015.
- [24] M. Kedadouché, M. Thomas, and A. Tahan, "A comparative study between empirical wavelet transforms and empirical mode decomposition methods: Application to bearing defect diagnosis," *Mechanical Systems and Signal Processing*, vol. 81, pp. 88–107, 2016.
- [25] C. Wang, T. Pun, and G. Chaneil, "A comparative survey of methods for remote heart rate detection from frontal face videos," *Frontiers in bioengineering and biotechnology*, vol. 6, p. 33, 2018.
- [26] C. Zhang, X. Wu, L. Zhang, X. He, and Z. Lv, "Simultaneous detection of blink and heart rate using multi-channel ica from smart phone videos," *Biomedical Signal Processing and Control*, vol. 33, pp. 189–200, 2017.
- [27] I. R. Tayibnapis, Y.-M. Yang, and K. M. Lim, "Blood volume pulse extraction for non-contact heart rate measurement by digital camera using singular value decomposition and burg algorithm," *Energies*, vol. 11, no. 5, p. 1076, 2018.
- [28] P. Sahindrakar, G. de Haan, and I. Kirenko, "Improving motion robustness of contact-less monitoring of heart rate using video analysis," *Technische Universiteit Eindhoven, Department of Mathematics and Computer Science*, 2011.
- [29] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Optics express*, vol. 18, no. 10, pp. 10762–10774, 2010.
- [30] J. Porri and J. V. Stone, "Undercomplete independent component analysis for signal separation and dimension reduction," report, Citeseer, 1998.
- [31] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of applied mathematics*, vol. 2, no. 2, pp. 164–168, 1944.
- [32] D. W. Marquardt, "An algorithm for least-squares estimation of non-linear parameters," *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.
- [33] X. Niu, H. Han, S. Shan, and X. Chen, "Vipl-hr: A multi-modal database for pulse estimation from less-constrained face video," in *Asian Conference on Computer Vision*, pp. 562–576, Springer, 2018.
- [34] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, and J. Dubois, "Un-supervised skin tissue segmentation for remote photoplethysmography," *Pattern Recognition Letters*, vol. 124, pp. 82–90, 2019.
- [35] G. Heusch, A. Anjos, and S. Marcel, "A reproducible study on remote heart rate measurement," *arXiv preprint arXiv:1709.00962*, 2017.
- [36] G. R. Tsouri, S. Kyal, S. A. Dianat, and L. K. Mestha, "Constrained independent component analysis approach to nonobtrusive pulse rate measurements," *Journal of biomedical optics*, vol. 17, no. 7, p. 077011, 2012.
- [37] R. Macwan, Y. Benezeth, and A. Mansouri, "Heart rate estimation using remote photoplethysmography with multi-objective optimization," *Biomedical Signal Processing and Control*, vol. 49, pp. 24–33, 2019.
- [38] D. McDuff, S. Gontarek, and R. W. Picard, "Improvements in remote cardiopulmonary measurement using a five band digital camera," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 10, pp. 2593–2601, 2014.
- [39] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, vol. 1, pp. I–I, IEEE, 2001.
- [40] T. M. Mahmoud, "A new fast skin color detection technique," *World Academy of Science, Engineering and Technology*, vol. 43, pp. 501–505, 2008.
- [41] M. P. Tarvainen, P. O. Ranta-Aho, and P. A. Karjalainen, "An advanced detrending method with application to hrv analysis," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 2, pp. 172–175, 2002.
- [42] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks*, vol. 13, no. 4-5, pp. 411–430, 2000.
- [43] A. Tharwat, "Independent component analysis: An introduction," *Applied Computing and Informatics*, 2020.
- [44] M. K. Transtrum and J. P. Sethna, "Improvements to the levenberg-marquardt algorithm for nonlinear least-squares minimization," *arXiv preprint arXiv:1201.5885*, 2012.
- [45] S. Amarai, A. Cichoki, and T. Chen, "A new learning algorithm for blind source separation," *Advances in Neural Information Processing Systems*, vol. 8, pp. 757–763, 1996.
- [46] J. J. Moré, *The Levenberg-Marquardt algorithm: implementation and theory*, pp. 105–116. Springer, 1978.
- [47] E. Fan, "Extended tanh-function method and its applications to nonlinear equations," *Physics Letters A*, vol. 277, no. 4-5, pp. 212–218, 2000.
- [48] D. Giavarina, "Understanding bland altman analysis," *Biochemia medica: Biochemia medica*, vol. 25, no. 2, pp. 141–151, 2015.
- [49] R. Artusi, P. Verderio, and E. Marubini, "Bravais-pearson and spearman correlation coefficients: meaning, test of hypothesis and confidence interval," *The international journal of biological markers*, vol. 17, no. 2, pp. 148–151, 2002.
- [50] G. Balakrishnan, F. Durand, and J. Guttag, "Detecting pulse from head motions in video," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3430–3437, 2013.
- [51] D. McDuff and E. Blackford, "iphys: An open non-contact imaging-based physiological measurement toolbox," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 6521–6524, IEEE, 2019.
- [52] P. Gupta, B. Bhowmick, and A. Pal, "Mombat: Heart rate monitoring from face video using pulse modeling and bayesian tracking," *Computers in biology and medicine*, vol. 121, p. 103813, 2020.