A Machine Learning Approach to Improve Contactless Heart Rate Monitoring Using a Webcam

Hamed Monkaresi, *Student Member, IEEE*, Rafael A. Calvo, *Senior Member, IEEE*, and Hong Yan, *Fellow, IEEE*

***Abstract*—Unobtrusive, contactless recordings of physiological signals are very important for many health and human–computer interaction applications. Most current systems require sensors which intrusively touch the user's skin. Recent advances in contact-free physiological signals open the door to many new types of applications. This technology promises to measure heart rate (HR) and respiration using video only. The effectiveness of this technology, its limitations, and ways of overcoming them deserves particular attention. In this paper, we evaluate this technique for measuring HR in a controlled situation, in a naturalistic computer interaction session, and in an exercise situation. For comparison, HR was measured simultaneously using an electrocardiography device during all sessions. The results replicated the published results in controlled situations, but show that they cannot yet be considered as a valid measure of HR in naturalistic human–computer interaction. We propose a machine learning approach to improve the accuracy of HR detection in naturalistic measurements. The results demonstrate that the root mean squared error is reduced from 43.76 to 3.64 beats/min using the proposed method.**

***Index Terms*—Blood volume pulse (BVP), computer vision, non-contact, remote sensing.**

I. Introduction

THE recent work of Poh *et al*. [1] for producing heart rate (HR) signals from a webcam has opened the opportunity for numerous new applications. This technique uses video of the face and independent component analysis (ICA) [2] to detect core physiological signals such as HR and respiration. This technology can be used to help monitor the health parameters of both individuals with cardiovascular disease and healthy individuals in nonlaboratory conditions.

Poh *et al*. [1] improved previous attempts [3], [4] to develop a low-cost accurate video-based method for contact free HR measurement. The algorithm is based on a blind source separation (BSS) technique that recovers unobserved signals or sources from a set of observed mixtures with no prior information about the mixing process. There are several techniques for BSS, including ICA [2] used by Poh *et al*. ICA is a technique for uncovering the independent source signals from a set of observations that are composed of linear mixtures of the underlying sources. The source signal in this study is the blood volume pulse (BVP) that propagates throughout the body. High degrees of agreement (correlation coefficient $r = 0.99$) were achieved for measures of HR using the proposed method and HR extracted from the BVP sensor.

In this paper, we evaluate Poh *et al*.'s technique under two new conditions; naturalistic human computer interaction, exercising scenarios which contain more user motion and different lighting conditions. We analyze the current limitations of this method and provide our own improvements using machine learning (ML) techniques.

II. Background

Remote, contactless monitoring of vital signs can be divided into three categories: microwave Doppler radar [5]–[7], thermal imaging [8], [9], and video-based imaging methods [1], [3], [4]. One of the earliest applications of remote contactless monitoring was reported in the 1980s [7], in which a microwave life-detection system for sensing the heartbeat and breathing of human subjects lying on the ground at a distance of about 30 m or located behind a cinder block wall. To our knowledge, this was the first effort to remotely monitor vital signs using microwave Doppler radar.

In thermal imaging, remote HR detection is performed through the analysis of skin temperature modulation. Pulsative blood flow modulates tissue temperature because of the heat exchange by convection and conduction between vessels and surrounding tissue. Such modulation is more pronounced in the vicinity of major superficial blood vessels [8]. A superficial blood vessel should be selected as the region of interest (ROI) for image analysis. Detecting and tracking these areas on human bodies is a challenging task. On the other hand, remote breathing measurements take advantage of the fact that the expired air has higher temperature than the inspired air due to heat exchange in the lungs and respiratory passageways. This thermic nature of breath around the nostril area creates an opportunity for thermal measurement [9]. Sprager and Zazula [10] proposed a new method to measure HR and respiration rate from optical interferometric signals under two separate conditions: at rest and during cycling.

Among the aforementioned methods, video-based remote monitoring methods are considered cheaper and easier to adopt [11]. Most of them use the photoplethysmography (PPG) methodology to detect HR variability. PPG is a low-cost and noninvasive means of sensing the cardiovascular BVP through

Manuscript received February 28, 2013; revised June 12, 2013, August 25, 2013, and October 8, 2013; accepted November 17, 2013. Date of publication November 20, 2013; date of current version June 30, 2014.

H. Monkaresi and R. A. Calvo are with the School of Electrical and Information Engineering, The University of Sydney, N.S.W. 2006, Australia (e-mail: hamed.monkaresi@sydney.edu.au; rafael.calvo@sydney.edu.au).

H. Yan is with the Department of Electronic Engineering, City University of Hong Kong, Kowloon, Hong Kong, and also with the School of Electrical and Information Engineering, University of Sydney, Sydney, N.S.W. 2006, Australia (e-mail: h.yan@cityu.edu.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JBHI.2013.2291900

variations in transmitted or reflected light [12]. PPG is based on the principle that blood absorbs light more than surrounding tissue, so variations in blood volume affect transmission or reflectance correspondingly.

PPG is measured through reflection of dedicated light sources such as infrared with a shallower penetration depth in skin. Recent works [3], [4] have demonstrated that pulse measurement could be obtained by analyzing the skin color with ambient light as the illumination source. Poh *et al*. proposed the method for recovering BVP signals by tracking the face skin color changes using an ordinary camera. It has shown its robustness under a nonlaboratory condition with little user motion. Automatic face tracking ability enhances the method to work well in a wide range of everyday activities.

III. Methodology

Our evaluation consists of three studies. First, the implementation of Poh *et al*. [1] method was replicated using a new dataset. In the second study, the accuracy of this method was evaluated on a long naturalistic HCI. Third, the robustness of the method was also assessed under new conditions which included motion artifacts and a wide range of HR changes.

A. Studies, Participants, and Recording Methods

1) First Study—Recording at Rest: Ten volunteer participants (mean age = 26.7 years, eight males, two females, 80% Caucasian, and 20% Asian) from The University of Sydney participated in the first and second study. All participants signed an informed consent form prior to data collection. This study was approved by the University of Sydney's Human Ethics Research Committee prior to data collection.

All participants were seated in front of the same computer running Windows XP in a normal indoor environment. Video recording was carried out using an ordinary webcam (Logitech Webcam Pro 9000) mounted on the screen. All videos were recorded in color (24-bit RGB with three channels, 8 bits/channel) at 30 frames/s (fps) with pixel resolution of 640 $\times$ 480 pixels and saved in AVI format. In order to record physiological changes, electrocardiogram (ECG), respiration, and galvanic skin response (GSR) sensors were placed on participants' bodies. BIOPAC MP150 system with AcqKnowledge (v. 3.8.2) software was used to acquire the physiological signals at a sampling rate of 250 Hz. Two electrodes were placed on the wrists for collecting ECG. The ECG-100C amplifier was used for ECG recording. GSR was recorded from the index and middle finger of the left hand and a respiration band was strapped around the chest to collect respiration rate. GSR and respiration rate data were not used in this study. The participants were asked to keep their movement to a minimum for 1 min to record the baseline physiological signals. This part of the experiment was used in the first study. The study was conducted in a normal room with normal artificial fluorescent ceiling light in combination with a varying amount of sunlight coming through windows from the left side of the participant.

2) Second Study—Naturalistic HCI: The second study was conducted using the same materials used in the first study. Immediately after completing the first study, participants were asked to create their personal websites in Google sites. They were allowed to explore the Internet and use external resources for completing the tasks. During the interaction, video and ECG signals were recorded simultaneously. Each interaction lasted about 30 min ($\pm$10 min).

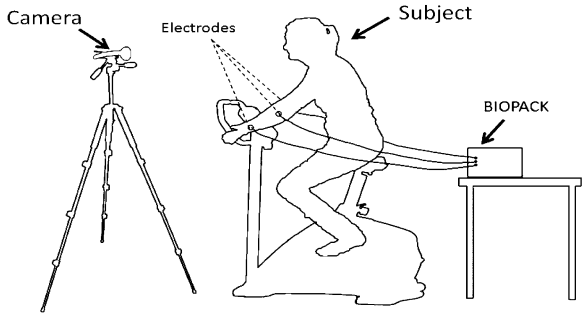

Fig. 1. Experiment setup for the third study.

TABLE I
Levels of Activity in the Indoor Cycling Study

Part	Level	Duration
1	Rest	2 min
2	Easy cycling	2 min
3	Normal cycling	3 min
4	Hard cycling	3 min
5	Normal cycling	3 min
6	Easy cycling	2 min
7	Rest	2 min

3) Third Study—Indoor Cycling: This study was conducted in an indoor gym environment with the participation of a female participant. The only illumination source was the ambient artificial fluorescent light. The same camera was mounted in front of a cycling machine for recording the participant's upper body. Two electrodes were placed on the participant's wrists and the earth electrode was placed on her left arm to record ECG. The participant's physiological signals were recorded while the participant was cycling (see Fig. 1).

The experiment consisted of seven levels. In the first and last parts, the participant was seated on the cycling machine at rest position (no cycling) for 2 min. The other parts of the experiment were captured while the participant was cycling at different resistance levels (as shown in Table I). The resistance of the cycling machine was gradually increased up to level 4 and returned to the easiest level. The participant was cycling for 2 min in the "Easy" levels and for 3 min in "Normal" and "Hard" cycling levels.

B. HR Measurement Methods

In this section, a brief explanation of the methodology used by Poh *et al*. [1] is presented. Then, our proposed extension to their method is explained. Fig. 2(a) shows the flowchart of recovering HR signals as described by Poh *et al*. Our variation of their method is illustrated in Fig. 2(b).

1) HR Recovering Using ICA: All the video and physiological recordings were analyzed offline using MATLAB (v. 7.12.0) and Open Computer Vision library (OpenCV) [13].

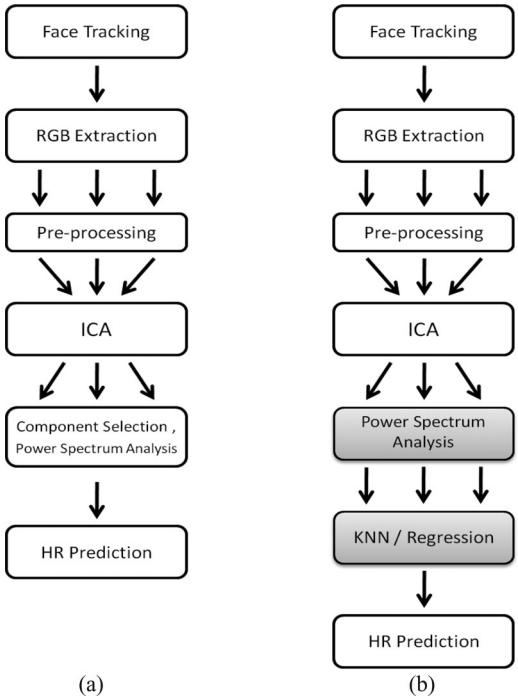

Fig. 2. Flowchart of HR extraction from video recording. (a) Method proposed by Poh *et al*. (b) Extended method improved by ML techniques.

First, for each video, we detected the face area using OpenCV and an extended boosted cascade classifier (Haar classifier [14]). In order to increase the calculation speed and reduce the false face detection rate due to background artifacts (false positives), a robust face tracking algorithm was implemented and applied to the dataset. Therefore, in each frame, a dynamic ROI was selected based on the extended face area which had been detected in a previous frame (extended by 30% in width and height). The algorithm looked for the face in this tracking ROI. If the face was not found, the ROI was expanded to the whole image. False positive and false negative rates were reported in each study. A false positive rate indicates the number of frames in which the algorithm detected more than one face divided by the total number of analyzed frames. A false negative rate indicates the number of failures of the algorithm in detecting the face over the total number of analyzed frames.

Through tracking ROI, a rectangular region in the middle with 60% of the width and the full height of the detected face region was selected as the new ROI and then separated into the RGB channels. In order to compose the raw signal for the red, green, and blue channels, the average of each color amplitude values of all pixels in the ROI was calculated for each frame.

Afterward, for subsequent processing, the 30-s moving window was defined with 1-s increments. Three raw traces in selected windows were preprocessed (detrended and normalized [1]) and finally decomposed into three independent source signals using ICA based on the Joint Approximate Diagonalization of Eigenmatrices algorithm [15].

a) Component selection: The outputs of the ICA were three row vectors [16]. However, it is still not clear which component contains HR signals. Hence, selection of the component which contains the HR signals is a challenging issue. For example, Poh *et al*. selected the second component for measuring HR in their first study [1]. They argued that the cardiovascular signals were clearly visible in the second component. However, in their latest study [11], they looked for HR signals in all three components. To find the cardiovascular pulse frequency, the power spectrum of the selected component was calculated and the spectrum with the maximum peak across all spectra was selected. The frequency of that spectrum was considered as the frequency of heart beats. In this paper, we have evaluated the spectrums between 0.75 and 4.0 Hz (corresponding to [45, 240] bpm) inside each component separately and compared the performance of each of them. Besides those estimations, the spectrum which contains the highest peak among all three components was considered as the fourth estimation. We refer to this method as MPA (Maximum Peak among All) in the rest of this paper. This module (component selection) has been replaced by ML techniques in our implementation.

b) Noise reduction: Motion artifacts or changes in illumination source can influence the ICA analysis and accuracy of frequency computation. To address this issue, Poh *et al*. [1] utilized the historical estimations of the pulse frequency to reject artifacts by fixing a threshold for maximum change in the pulse rate between successive measurements (taken one second apart). If the difference between the current pulse rate estimation and the last computed value exceeded the threshold (they used a threshold of 12 bpm in their experiments), the algorithm rejected it and searched the operational frequency range for the frequency corresponding to the next highest power that met this constraint. If no frequency peaks that met the criteria were located, then the algorithm retained the last computed pulse frequency estimation.

2) ML Techniques: We have evaluated the performance of the three output components of ICA in estimating actual HR. Each component could estimate the HR better than others on different occasions. We could not rely on one component to estimate the HR even for the same participant. By getting some information obtained from the ICA and training a model, we can make a better estimation. In our proposed model, nine features extracted from power spectral density (PSD) analysis were applied on output components of the ICA. Six of them included the most powerful frequencies in each component before and after applying the noise reduction method. The index of the spectral peak in the PSD of each component, selected in the noise reduction method, formed the rest of the features. Two ML techniques, linear regression and k-nearest neighbor (kNN) classifier, have been used to estimate HR from these computed values. Linear regression is an approach that tries to model a linear relationship between a dependent variable and one or more explanatory variables. kNN is a type of instance-based learning and uses a simple distance measure to find the training instance closest to the known test instance, and considers that the test instance belongs to the same class as this training instance [17]. When the target parameter (y) has a numeric value, the kNN would be considered to be a regression problem. In this case, the k neighbors nearest to the test instance are selected first, and then, the average of their

predicted values is assigned to the test instance. Suppose we have training data $(x_1, y_1), \ldots, (x_n, y_n)$, where $x_1, \ldots, x_n \in \mathbb{R}^p, y_1, \ldots, y_n \in \mathbb{R}$. If x^* is a test vector where $x^* \in \mathbb{R}^p$, we can predict y^* by finding the kNN of x^* and then computing the mean of the k nearest training values $(y_{r_1}, \ldots, y_{r_k})$ as shown in (1). The training and testing processes were performed separately with a tenfold cross-validation approach. The Waikato environment for knowledge analysis (Weka [18]) was used for executing these techniques:

$$y^* = \frac{1}{k} \sum_{i=1}^{k} y_{r_i}. \quad (1)$$

The mean absolute error (MSE), root mean squared error (RMSE), and Pearson's correlation coefficient were calculated for the estimated HR and actual HR extracted from reference ECG. Bland–Altman plots [19] were used for comparing proposed methods and actual HR values. The mean differences with 95% limits of agreement (LoA) ($\pm$1.96 SD) were also reported for each method. The LoA specify a range that most of the measurement errors lie within it.

3) Extracting Actual HR From ECG: The reference HR values were also extracted from the 30-s moving window (1-s incremental) from recorded ECG signals. We chose these windows to compare the results with our proposed method. Data preprocessing and feature extraction were done using the Augsburg Biosignal Toolbox [20].

IV. Results

A. First Study: Validation

This study aimed to replicate the results of [1] with our own implementation of the algorithm, and our own data. The first 1 min of each participant's video recording in which their movements were kept to a minimum was analyzed at this stage. Thirty-one estimations were performed for each participant. The first 30-s window was considered to estimate the HR at $t = 30$. The standard deviations of the x- and y-coordinates of the face tracker of all participants were 4.56 and 2.57 pixels, respectively. The average actual HR of 78.80 bpm was extracted from ECG signals among all participants. The standard deviation of extracted actual HR was 8.84 bpm (see Table II).

Table III presents a comparison of our implemented algorithms applied in this experiment and Poh *et al*.'s experiment results. It should be mentioned that after ICA analysis, the third component produced the best results among all three components. The MPA method produced a similar performance. The results achieved from the third component are presented in Table III. A correlation coefficient of 0.99 between estimated HR and the actual HR was achieved in this study. The RMSE of measurement at rest in [1] was 2.29, while in this study, the RMSE was reduced to 1.69. Furthermore, using this experimental data, the mean bias was 0.86 bpm and the LoA span was 5.7 bpm, slightly better than Poh *et al*.'s [1] results in the sitting still experiment (mean bias = 0.05 bpm, LoA span = 8.99 bpm). Our results also showed reasonable accuracy compared to a new method for HR measurements at rest using Fiber Bragg grating-based sensor [21] (mean bias = $-$ 0.01 bpm, LoA span = 3.64 bpm). Overall, the results showed that the accuracy of our implementation and this new dataset was comparable with Poh *et al*.'s report.

TABLE II
First Study Description: Validation of the Implemented Method for Estimating HR

	Poh et al.'s Study	First Study
Parameters		
Window size (seconds)	30	30
Experiment length (seconds)	60	60
# of participants	12 (2 F, 10 M)	10 (2 F, 8 M)
Recording rate (fps)	15	30
# of frames	10800	18000
# of measurements	372	310
SD of x- coordination (pixels)	-	4.56
SD of y- coordination (pixels)	-	2.57
Range of x- movement (pixels)	-	23
Range of y- movement (pixels)	-	18
Face Detection Acc	FN=0% FP=0%	FN=0% FP=0%
Mean actual HR	-	78.80
SD of actual HR	-	8.84

F=Females, M=Males, FN = false negative, FP = false positive.

TABLE III
Comparison of the Results Published by Poh *et al*. and Our Study Using the Same Algorithm

	Poh et al.'s Study	First Study
Method	**ICA**	**ICA**
Selected component	2nd	3rd
Mean bias (bpm)	-0.05	0.86
SD of bias (bpm)	2.29	1.46
Upper limit (bpm)	4.44	3.71
Lower limit (bpm)	-4.55	-1.99
RMSE	2.29	1.69
Corr. coefficient	0.98^{***}	0.99^{***}

***: ($p<0.001$).

B. Second Study: Evaluate the Algorithm in HCI

The goal of this study was to evaluate the robustness and reliability of the video based, noncontact HR measurement in a naturalistic HCI scenario. For this reason, participants were free to move and no advice was provided regarding positions. The interaction with the computer was immediately started after 1 min of a no-movement situation which was also used in the first study. We considered the whole session as one dataset. On average, each session lasted about 30 min, but in order to compare across the same length for all participants, the first 20 min of each session were analyzed here. This analysis has been done based on 30-sec moving window. Table IV provides more information about this study.

The face tracking algorithm was robust enough in detecting the face during head movement. However, sometimes because of the large degree of head tilting and turning (more than 45°) or the occlusion of the face, the algorithm could not find the frontal face in the recorded images. In order to have a fair evaluation of the ICA analysis, the time windows which included false negative image frames were ignored for further analysis. Thus,

TABLE IV
SECOND STUDY DESCRIPTION: HR ESTIMATION DURING HUMAN–COMPUTER INTERACTION

	Second Study
Parameters	
Window size (seconds)	30
Experiment length (seconds)	1200
# of participants	10 (2 female, 8 male)
Recording rate (fps)	30
# of frames	360000
# of measurements	9329 (of 11710)
SD of x- coordination (pixels)	12.16
SD of y- coordination (pixels)	7.24
Range of x- movement	64.00
Range of y- movement	35.50
Face Detection Acc	FN=0.04%, FP=0%
Minimum actual HR	61.27
Maximum actual HR	111.79
Mean actual HR	80.55
SD of actual HR	9.52

among 11 710 potential 30-s windows, only 9329 were selected for ICA analysis. The average range of the x- and y-coordinates of the detected faces was 64.00 and 35.5 pixels, respectively.

1) User-Dependent Analysis: In this part, we have evaluated different components to estimate HR for each participant. The third component achieved the best results among eight out of ten participants by comparing RMSEs. The MPA method improved the HR estimation for the remained two participants. The best results obtained by ICA analysis for each participant are presented in Table V. In addition, the results achieved by applying kNN and linear regression are also presented in this table. The figures show the significant improvement of ICA with the kNN method over the former method (only ICA) for all participants. The regression did not improve the performance very much. These results imply the feasibility of building personal models for HR estimation during HCI in naturalistic scenarios.

2) Combined-Participants Analysis: Data from individual participants were combined to yield one large dataset. This dataset was used for training and testing the models with the tenfold cross-validation approach. On the other hand, the HR was also estimated by the ICA methods. Here, the third component also provided the best performance among all three components. The descriptive statistics of HR estimation through different approaches are presented in Table VI.

The large values for the mean bias –25.38 bpm (SD = 35.65 bpm) would also indicate the weakness of the ICA method in detecting HR in the natural HCI environments. Applying ML techniques and the ICA method significantly reduced the absolute error of the HR estimation. The MSE reduced from 27.17 to 1.32 bpm and 9.10 bpm by applying kNN ($k = 1$) and linear regression techniques, respectively. The ICA and kNN technique also obtained a strong correlation with the reference sensor by reaching the Pearson's correlation coefficient of 0.89 ($p < 0.001$).

The results of the Bland–Altman analysis for the three methods of HR prediction are presented in Fig. 3. The LoA range was very wide for the ICA and for ICA and regression predictions. In contrast, the ICA and kNN method achieved the best accuracy with the LoA from -7.09 to 7.21 bpm [see Fig. 3(c)]. In Fig. 3(c), the difference values have more variations for mid-range average values (around 80–90) and smaller variations for small and large average values. One possibility is that there are more samples in the mid-range. That is, there are more people in this range and fewer people at two extremes. When there are more people, you would naturally have a wider range of difference values. However, there were no systematic or proportional errors observed for the ICA and kNN method as shown in Fig. 3(c).

TABLE V
RESULTS FOR HR ESTIMATION USING THE ICA METHOD AND THE IMPROVED METHOD BY ML TECHNIQUES (PARTICIPANT-DEPENDENT ANALYSIS)

ID	Method	r	MAE	RMSE
1	ICA (3rd Comp.)	-0.38	8.67	21.16
	ICA + kNN	**0.85**	**0.63**	**1.20**
	ICA + Reg.	0.44	1.49	1.92
2	ICA (3rd Comp.)	0.10	2.79	4.34
	ICA + kNN	**0.81**	**0.63**	**1.38**
	ICA + Reg.	0.16	1.68	2.20
3	ICA (3rd Comp.)	0.29	46.83	59.13
	ICA + kNN	**0.92**	**0.44**	**1.09**
	ICA + Reg.	0.34	1.77	2.54
4	ICA (3rd Comp.)	0.41	45.34	57.39
	ICA + kNN	**0.93**	**0.43**	**0.92**
	ICA + Reg.	0.48	1.39	2.12
5	ICA (MPA)	0.18	33.92	41.98
	ICA + kNN	**0.90**	**0.69**	**1.62**
	ICA + Reg.	0.47	2.48	3.20
6	ICA (3rd Comp.)	0.29	3.96	8.20
	ICA + kNN	**0.92**	**0.48**	**0.95**
	ICA + Reg.	0.49	1.54	2.06
7	ICA (3rd Comp.)	-0.34	32.41	44.31
	ICA + kNN	**0.83**	**0.72**	**1.19**
	ICA + Reg.	0.50	1.37	1.77
8	ICA (3rd Comp.)	-0.02*	14.27	28.35
	ICA + kNN	**0.86**	**0.58**	**1.18**
	ICA + Reg.	0.46	1.51	1.98
9	ICA (MPA)	0.15	56.85	63.62
	ICA + kNN	**0.81**	**0.84**	**1.75**
	ICA + Reg.	0.18	2.27	2.81
10	ICA (3rd Comp.)	0.25	16.92	28.67
	ICA + kNN	**0.91**	**1.38**	**3.10**
	ICA + Reg.	0.49	5.15	6.37

*: p<0.05; p<0.001 for the rest; kNN(k=1); Reg.: Linear Regression; r: Pearson's correlation coefficient; MAE: Mean Absolute Error; RMSE: Root Mean Squared Error.

TABLE VI
DESCRIPTIVE STATISTICS OF HR ESTIMATION USING THE ICA METHOD AND IMPROVED METHOD BY ML TECHNIQUES (COMBINED-PARTICIPANTS ANALYSIS)

	Second Study		
Method	**ICA [1]**	**ICA + kNN**	**ICA + Regression**
Selected component	3rd	-	-
Mean bias (bpm)	-25.38	0.06	0.00
SD of bias (bpm)	35.65	3.65	9.10
RMSE	**43.76**	**3.64**	**7.31**
Corr. Coefficient	**-0.10*****	**0.93*****	**0.29*****
Mean absolute error	27.17	1.32	9.10
Relative absolute error	3.43	0.16	0.92
Root relative squared error	21.14	0.38	0.95

***: (p<0.001).

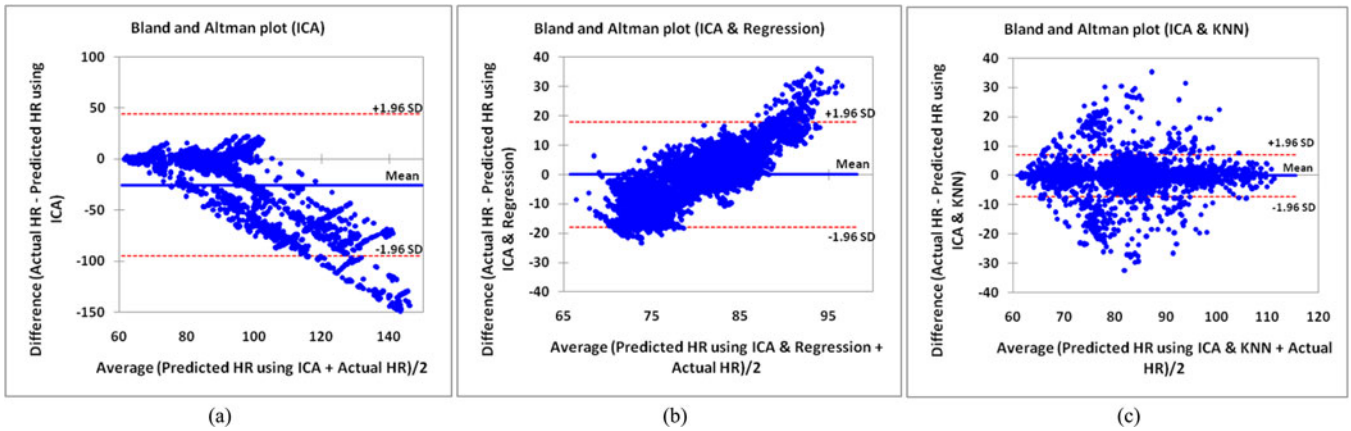

Fig. 3. Bland–Altman plots analyzing the agreement between measured actual HR and measured HR using (a) ICA, (b) ICA with regression, and (c) ICA with kNN in the second study.

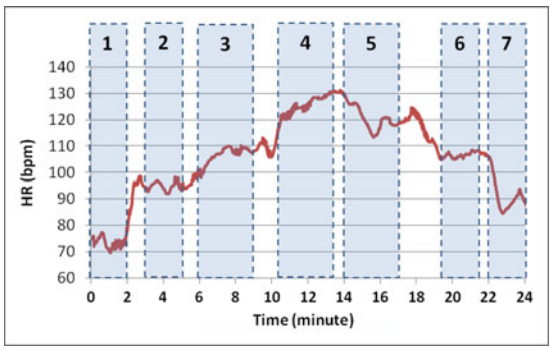

Fig. 4. Actual HR changes during the third study (indoor exercising) extracted from ECG.

According to Fig. 3(a), even though the actual HR range was between 61.27 and 111.79 bpm, the ICA method predicted very large values most of the time. This might be because other strong frequencies appeared in the third independent component. Fig. 3(b) shows a case of proportional error for the ICA and regression method. It suggests that the relation between the nine features and actual HR might not be linear.

C. Third Study—Indoor Exercising

In this study, the participant was asked to keep her movement to a minimum during the experiment. Obviously, some unwanted movement during the cycling was acceptable. The false negative ratio was 0.0.

We explored here the robustness of the proposed algorithm for a wide dynamic range of HR. In the second study, the average of the range and standard deviation of actual HRs among all participants were 16.50 and 3.00 bpm, respectively. In this study, the HR of the participant started from 75 bpm, reached 130 bpm in the middle of the experiment, and went down to 85 bpm. Fig. 4 presents the actual HR changes during the experiment extracted from the ECG signals. Seven parts of the experiment are shown in the corresponding columns in Fig. 4. The range and standard deviation of actual HR in the first, second, and sixth parts were smaller than in the other parts.

TABLE VII
SUMMARY OF THE THIRD STUDY RESULTS: HR ESTIMATION DURING INDOOR EXERCISING USING THE ICA METHOD

Parts	1	2	3	4	5	6	7
Moving window size	30 sec						
Recording rate	15 fps						
Length (sec)	120	120	180	180	180	120	120
# of measurements	91	91	151	151	151	91	91
SD of x-coordination	4.92	5.06	6.54	9.97	8.65	8.06	4.05
SD of y-coordination	5.14	4.10	1.68	3.39	3.9	2.53	1.75
Range of x-coordination	36	30	28	50	46	39	18
Range of y-coordination	27	35	11	24	22	14	11
Range of Actual HR (bpm)	7.85	6.32	8.34	10.37	13.08	3.62	9.24
SD of Actual HR (bpm)	2.13	1.55	2.02	2.63	3.74	1.00	2.68
Results Using ICA							
Selected Component	1st	1st	1st	3rd	2nd	1st	1st
Mean bias (bpm)	2.82	-4.07	11.95	-1.59	2.55	-10.30	10.42
SD of bias (bpm)	4.47	1.94	6.07	20.27	5.70	1.50	9.04
Max bias (bpm)	14.12	0.12	27.22	25.40	14.17	-7.93	31.06
Min bias (bpm)	-12.95	-8.89	-0.38	-49.13	-15.47	-18.43	-2.08
RMSE	5.30	4.52	13.44	20.34	6.25	10.47	13.84

The results of applying the proposed algorithm on these parts are presented in Table VII. The strategy for component selection in each part was to choose the component which contains the spectrum that minimizes the RMSE value. Consequently, the third component was selected in part 4, the second component was selected in part 5 and the first component was selected in the other parts. Our evaluations also stated that the HR spectrum could not be estimated accurately by considering the spectrum containing the highest peak among all three components.

As shown in Table VII, the ICA method achieves acceptable results in the first and second parts with RMSE values of 5.30 and 4.32. In these two parts, the range of actual HR changes was smaller than other parts except part 6 which did not achieve a good RMSE. On the other hand, the algorithm showed the worst result in part 4 which contained a gradually increased HR.

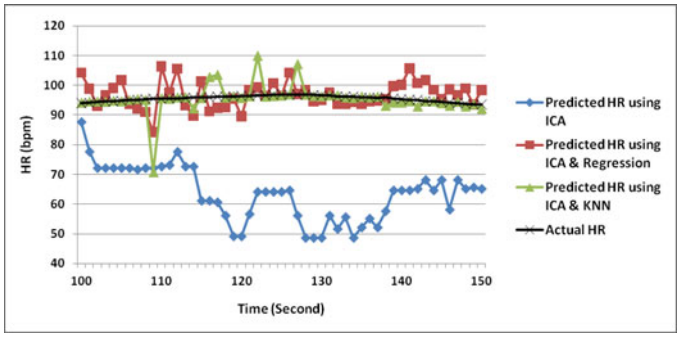

Fig. 5. Actual HR changes during the third study (indoor exercising) extracted from ECG.

TABLE VIII
DESCRIPTIVE STATISTICS OF HR ESTIMATION USING ICA METHOD AND ML TECHNIQUES DURING INDOOR EXERCISING (THIRD STUDY)

	Third Study		
	ICA	**ICA + kNN**	**ICA + Regression**
Selected component	3rd	-	-
Mean bias (bpm)	-24.37	-0.28	0.05
SD of bias (bpm)	25.54	4.33	13.70
RMSE	**35.31**	**4.33**	**13.69**
Corr. Coefficient	0.53^{***}	0.97^{***}	0.58^{***}
Mean absolute error	28.53	1.41	11.24
Relative absolute error	2.09	0.10	0.82
Root relative squared error	4.37	0.25	0.81

***: ($p<0.001$).

However, exploring the whole experiment as one dataset showed that the third component produced the RMSE of 35.23 which was the best value among the components and the MPA method. On the other hand, the kNN technique dramatically increased the performance of the estimation. Fig. 5 illustrates an example of the improvement of ML techniques over the ICA in measuring HR.

The Bland–Altman analysis showed that the best agreement between actual HR measured by ECG and the proposed methods was achieved by the ICA and kNN method. This method reduced the mean bias from –24.37 to 0.05 bpm with 95% LoA −8.20 to 8.76. Table VIII summarizes the results of HR detection in the third study. These results represented the robustness of the proposed method in a new condition which includes lots of user motions and a large range of HR changes (from 69.5 to 130.79 bpm).

V. CONCLUSION

We have evaluated a method for remote HR measuring in three applications: a controlled laboratory task, a naturalistic HCI, and an indoor cycling exercise. This study evaluated Poh *et al.*'s method and showed the feasibility of their methodology to measure HR at rest. Their seminal work is one of the successful attempts at remote physiological sensing. However, their proposed method did not show positive results in naturalistic HCI and indoor exercise situations. The study analyzes the problems caused by unwanted movements and the wide dynamic range of HR which are common during real world measurement. We have addressed these issues by building and training specific models for each participant using ML techniques. The results suggest that the kNN-based technique outperforms other approaches (manual or computational) that try to select the best independent component for HR estimation. On average, the mean of absolute errors in HR estimation is reduced to 0.68 bpm by applying the kNN technique to the ICA outputs among ten participants in the HCI scenario. The kNN technique also improved the accuracy of the HR estimation over the former method in indoor exercising conditions.

Although the accuracy of estimation using the proposed approach is increased, in some applications (e.g., emergency severity index triage), the achieved accuracy is not acceptable and more improvements are needed to implement a more accurate estimation. Another limitation might be the objection to drawing major generalizations from the small number of subjects studied in the third study. It should be noted that the purpose of the third study was to measure the impact of a larger dynamic range on the measurements. The lighting conditions and other environmental variables were kept the same. Recruiting more participants for generalizing our findings from the third study is considered for possible future work. However, building a user-independent model to yield reliable results when presented with new users without the need for retraining is one of our concerns for future work.

ACKNOWLEDGMENT

The authors would like to thank The New South Wales Institute of Sport for their help in the third study.

REFERENCES

[1] M. Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Exp.*, vol. 18, no. 10, pp. 10762–10774, May 2010.
[2] P. Comon, "Independent component analysis, a new concept?," *Signal Process.*, vol. 36, no. 3, pp. 287–314, 1994.
[3] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Opt. Exp.*, vol. 16, no. 26, pp. 21434–21445, Dec. 2008.
[4] C. Takano and Y. Ohta, "Heart rate measurement based on a time-lapse image," *Med. Eng. Phys.*, vol. 29, no. 8, pp. 853–857, Oct. 2007.
[5] C. Li, J. Cummings, and J. Lam, "Radar remote monitoring of vital signs," *Microw. Mag.*, vol. 10, pp. 47–56, Feb. 2009.
[6] E. F. Greneker, "Radar sensing of heartbeat and respiration at a distance with applications of the technology," in *Proc. RADAR*, 1997, pp. 150–154.
[7] K. M. Chen, D. Misra, H. Wang, H. R. Chuang, and E. Postow, "An X-band microwave life-detection system," *IEEE Trans. Biomed. Eng.*, vol. BME-33, no. 7, pp. 697–701, Jul. 1986.
[8] M. Garbey, N. Sun, A. Merla, and I. Pavlidis, "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 8, pp. 1418–1426, Aug. 2007.
[9] J. Fei and I. Pavlidis, "Thermistor at a distance: Unobtrusive measurement of breathing," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 4, pp. 988–998, Apr. 2010.
[10] S. Sprager and D. Zazula, "Detection of heartbeat and respiration from optical interferometric signal by using wavelet transform," *Comput. Methods Programs Biomed.*, vol. 111, no. 1, pp. 41–51, Jul. 2013.
[11] M. Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7–11, Jan. 2011.
[12] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological Meas.*, vol. 28, no. 3, pp. R1–R39, Mar. 2007.

[13] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*. Sebastopol, CA, USA: O'Reilly Media, 2008.
[14] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2001, pp. 511–518.
[15] J. F. Cardoso, "High-order contrasts for independent component analysis," *Neural Comput.*, vol. 11, no. 1, pp. 157–192, Jan. 1999.
[16] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Netw.*, vol. 13, no. 4/5, pp. 411–430, 2000.
[17] D. W. Aha, D. Kibler, and M. K. Albert, "Instance-based learning algorithms," *Mach. Learning*, vol. 6, no. 1, pp. 37–66, Jan. 1991.
[18] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*. San Mateo, CA, USA: Morgan Kaufmann, 2005, p. 525.
[19] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 1, no. 8476, pp. 307–310, Feb. 1986.
[20] J. Wagner, J. Kim, and E. Andre, "From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification," in *Proc. IEEE Int. Conf. Multimedia Expo.*, 2005, pp. 940–943.
[21] Ł. Dziuda, F. W. Skibniewski, M. Krej, and P. M. Baran, "Fiber Bragg grating-based sensor for monitoring respiration and heart activity during magnetic resonance imaging examinations," *J. Biomed. Opt.*, vol. 18, no. 5, p. 057006, 2013.

Hamed Monkaresi (S'12) is currently working toward the Ph.D. degree at the School of Electrical Engineering and Information Engineering, The University of Sydney, N.S.W., Australia.

He has been a member of the Learning and Affect Technologies Engineering (Latte) Research Group since July 2010. His research interests are in affective computing and computer vision. More specific interests include facial expression recognition, and using machine learning techniques for understanding complex mental states.

Rafael A. Calvo (SM'09) received the Ph.D. degree in artificial intelligence applied to automatic document classification from the Universidad Nacional de Rosario (UNR), Argentina, in 2000.

He has also worked at Carnegie Mellon University and the Universidad Nacional de Rosario, and as a Consultant for projects worldwide. He is currently an Associate Professor in the School of Electrical and Information Engineering, The University of Sydney, N.S.W., Australia, and the Director of the Learning and Affect Technologies Engineering (Latte) Research Group. He has authored numerous publications in the areas of affective computing, learning systems, and web engineering.

Dr. Calvo is a member of the IEEE Computer Society. He is an Associate Editor of the IEEE TRANSACTIONS ON AFFECTIVE COMPUTING and the IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES.

Hong Yan (F'06) received the Ph.D. degree from Yale University, New Haven, CT, USA.

He was a Professor in the School of Electrical and Information Engineering, The University of Sydney, N.S.W., Australia. He is currently a Professor in the Department of Electronic Engineering, City University of Hong Kong, Hong Kong. His research interests include image processing, pattern recognition, and bioinformatics. He has more than 300 journal and conference publications in these areas.

Dr. Yan is a Fellow of the International Association of Pattern Recognition.