Assessment of source separation techniques to extract vital parameters from videos

Daniel Wedekind, Alexander Trumpp, Fernando Andreotti, Frederik Gaetjen, Stefan Rasche, Klaus Matschke, Hagen Malberg, Sebastian Zaunseder

Angaben zur Veröffentlichung / Publication details:

Wedekind, Daniel, Alexander Trumpp, Fernando Andreotti, Frederik Gaetjen, Stefan Rasche, Klaus Matschke, Hagen Malberg, and Sebastian Zaunseder. 2015. "Assessment of source separation techniques to extract vital parameters from videos." In *23rd European Signal Processing Conference (EUSIPCO), 31 August 2015 - 4 September 2015, Nice, France*, edited by Jean-Luc Dugelay, Dirk Slock, Patrizio Campisi, and Claude Delpha, 434–38. Piscataway, NJ: IEEE. https://doi.org/10.1109/eusipco.2015.7362420.

ASSESSMENT OF SOURCE SEPARATION TECHNIQUES TO EXTRACT VITAL PARAMETERS FROM VIDEOS

Daniel Wedekind[*], *Alexander Trumpp*[*], *Fernando Andreotti*[*], *Frederik Gaetjen*[†], *Stefan Rasche*[†], *Klaus Matschke*[†], *Hagen Malberg*[*], *Sebastian Zaunseder*[*]

[*] Institute of Biomedical Engineering, Faculty of Electrical and Computer Engineering
TU Dresden, Dresden, Germany
[†] Herzzentrum Dresden GmbH, University Hospital Carl Gustav Carus Dresden
TU Dresden, Dresden, Germany

ABSTRACT

Camera-based photoplethysmography is a contactless mean to assess vital parameters, such as heart rate and respiratory rate. In the field of camera-based photoplethysmography, blind source separation (BSS) techniques have been extensively applied to cope with artifacts and noise. Despite their wide usage, there is no consensus that common BSS approaches contribute to an improved analysis of camera-based photoplethysmograms (cbPPG). This contribution compares previously proposed multispectral BSS techniques to a novel spatial BSS approach for heart rate extraction from cbPPG. Our analysis indicates that the application of BSS techniques not necessarily improves cbPPG's analysis but signal properties like the signal-to-noise-ratio should be considered before applying BSS techniques.

Index Terms— Camera-based Photoplethysmography, Blind Source Separation, Independent Component Analysis, Principal Component Analysis, Heart Rate

1. INTRODUCTION

Contact-less acquisition of vital signs allows the implementation of novel clinical and out-of-hospital applications. Various systems and techniques for contactless measurements have been introduced in the last years. Amongst such approaches, camera-based monitoring is one promising solution to assess the cardiac pulse.

The acquisition of the cardiac pulse, using near infrared cameras, was firstly demonstrated by Huelsbusch *et al.* 2002 [1]. Since then, many researchers have worked with camera-based photoplethysmography, most often to assess the heart rate [2–8]. Since the technique is highly sensitive to artifacts, induced by movements and lightening changes, elaborated image and signal processing techniques are required to make use of the camera-based photoplethysmogram (cbPPG) under real world conditions.

The application of blind source separation (BSS) algorithms has been highlighted to improve heart rate extraction from cbPPG recordings. BSS aims at separating the desired signal content (i.e. cardiac pulse) from noise and artifacts by means of decorrelation and exploiting statistical independence. Principal Component Analysis (PCA) and Independent Component Analysis (ICA) are typical BSS techniques that have been widely applied to cbPPG.

Proposed techniques use segments of different color channels (typically RGB) extracted from regions of interest (ROI), typically the face, as input to PCA or JADE ICA [4, 9–14]. FastICA [15] has also been applied to RGB signals [6, 14] and achieved a slightly better performance in comparison to other ICA algorithms [6]. Tsouri *et al.* [16], on the other hand, used RGB information of a face ROI as input for a constrained ICA. They successively embedded sines at assumed heart rates as constraint to select the most probable heart rate. Other researchers have further developed the idea of multispectral cbPPG with PCA/ICA but have utilized alternatives to RGB, namely combinations of RGB with orange and cyan channels or chrominance based signals, respectively [7, 10, 17, 18].

Besides investigating the impact of using different wavelengths, spatial approaches like reducing the face ROI to a more concise area have been recently addressed in the context of PCA/ICA [6, 8, 11, 14, 17, 19]. Spatial reduction seeks to exclude regions that are not supposed to contribute with useful signals but can introduce distortions like mouth movements during speaking/smiling or blinking eyes [14, 20]. Approaches described in literature typically rely on a spatial preselection and use multispectral information (RGB) as input to BSS techniques. However, a monochrome cbPPG, extracted from the forehead, is used as input for spatio-temporal ICA in [19]. Wang *et al.* alternatively addressed spatial reduction without using face detection. The authors utilized the tem-

The authors would like to thank the Saxon state ministry of science and culture (SMWK) and the European Union for funding the project CardioVisio - Contactless acquistion of vital parameters. FA is financially supported by the Conselho Nacional de Desenvolvimento Tecnologico (CNPq - Brazil).

poral behavior of pixel traces to distinguish skin-like areas showing temporally periodic content from motion-like content [18].

Despite the frequent multispectral BSS use, there is no consensus on the benefit of using BSS techniques with multispectral inputs. Kwon *et al.* described a blurred spectral peak after applying RGB ICA as well as an increased heart rate error [12]. Christinaki *et al.* identified only subtle improvements but similar heart rate errors with/without using RGB ICA [6]. Feng *et al.* showed a lack of robustness applying a standard RGB ICA [8].

One possible reason for BSS techniques' limitations is the assumption of a linear mixing process [15] of available sources in standard PCA/ICA. In particular, the wavelength-dependent penetration depth into human skin [1] may introduce nonlinear mixing behavior, which may affect the performance of BSS algorithm negatively. However, other factors may also impact the success of BSS techniques in extracting pulsatile signals. Thus, deepened investigations on the performance of BSS techniques are required. This contribution compares a novel spatially applied monochrome BSS approach to an accordingly adjusted multispectral application of BSS, in order to characterize the potential efficacy of both approaches.

2. MATERIALS AND METHODS

2.1. Data Recording and Selection

Video data was recorded using an industrial camera (IDS UI-3370CP-C-HQ, 100 fps, 420x320 pixels, RGB 3x12 bit). The camera was placed at a distance of approximately 60 cm to patients' faces. For illumination, a fluorescent light source and natural light was used respectively. Overall 18 recordings (13 male, 5 female; 30 minutes per recording) of resting cardiovascular patients in a supine position during recovery after heart surgery were selected from a larger collective. Electrocardiogram (ECG) and finger PPG were simultaneously recorded at 100 Hz as reference. Written informed consent was obtained from all patients. The study was approved by the Ethics Committees of Technische Universität Dresden.

In order to use only suitable data for further analysis, we restricted our analysis to data segments which showed good quality on the reference PPG and did not contain severe cardiac disorders. A total of about 6 h video data was selected (average length 1200 ± 400 s per patient). Since the selection did not consider video quality, slight patient motion as well as lightning inadequacies (changes or insufficient lightning) persisted on the dataset.

2.2. cbPPG Extraction

The preselected video excerpts were processed in windows of 10 s resulting in an overall 2197 windows. Every video frame was covered by 25 x 19 overlapping ROIs (50% overlap at each direction) of 32 x 32 pixels (see figure 1 for a video frame example demonstrating ROI placement). The ROI size was chosen according to own investigations regarding the camera sensor chip which revealed that no significant signal quality improvement can be obtained with larger ROIs in comparable recording setups (distance to subject, resolution). The cbPPG was extracted from each ROI_n (with n = 1,2,. . . ,475) at every wavelength (color = R,G,B) by averaging its pixel values for consecutive frames (see figure 2 for an example) [1].

Fig. 1. Exemplary video frame demonstrating overlapping 32x32 pixels ROIs for cbPPG extraction.

Each 10 s cbPPG segment was normalized by subtracting its mean and dividing it by its standard deviation. Furthermore, 0.5 Hz highpass filtering (5th order Butterworth) was applied to limit low frequency content below an expected heart rate [21]. Suchlike preprocessed $\mathrm{cbPPG}_{n,color}$ were used for further processing.

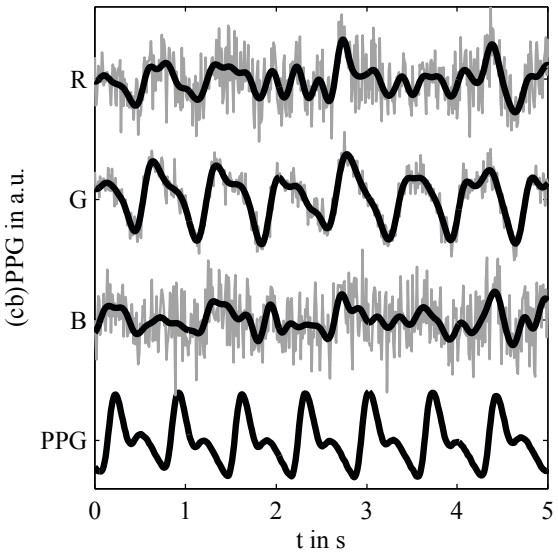

Fig. 2. Sample cbPPG excerpt showing signals extracted from one ROI on patients face for three color channels R,G,B (grey: $\mathrm{cbPPG}_{n,color}$ signals, black: lowpass (5 Hz) filtered signals) with corresponding reference PPG.

Algorithm 1 Procedure to form input sets for applied BSS techniques

Input: Signals $\text{cbPPG}_{n,color}$ with $n = 1, 2, \ldots, 475$ and $color \in \{R, G, B\}$

1: Localize maxima in amplitude spectra $X(f) = |\mathcal{F}\{\text{cbPPG}_{n,G}\}|$ between $[30, 240]$ bpm for every ROI $\rightarrow \hat{f}_{n,G}$

2: Calculate histogram of $\hat{f}_{n,G} \rightarrow H_{\hat{f}_{n,G}}$ (see figure 3)

3: Identify most often occurring frequencies from $H_{\hat{f}_{n,G}} \rightarrow \tilde{f}_{i,G}$ with $i \in \{1, 2, 3\}$ (see figure 3)

4: Calculate SNRs from $\text{cbPPG}_{n,G}$ (see section 2.4) for $\tilde{f}_i$ with $i \in \{1, 2, 3\} \rightarrow \text{SNR}_{n,G}^{\tilde{f}_i}$ with $i \in \{1, 2, 3\}$

5: Form input sets

Monochrome approach: assemble the green channel from appropriate ROIs to three input sets according to

$$\left.\begin{array}{ll} \text{cbPPG}_{n,G} \text{ with} & \text{highest SNR}_{n,G}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,G} \stackrel{!}{=} \tilde{f}_i, \\ \text{cbPPG}_{n,G} \text{ with} & \text{second highest SNR}_{n,G}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,G} \stackrel{!}{=} \tilde{f}_i, \\ \text{cbPPG}_{n,G} \text{ with} & \text{third highest SNR}_{n,G}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,G} \stackrel{!}{=} \tilde{f}_i \end{array}\right\} \rightarrow S_i^{mc} \text{ with } i \in \{1, 2, 3\}$$

Multispectral approach: assemble color channels from one appropriate ROIs to three input sets according to

$$\left.\begin{array}{ll} \text{cbPPG}_{n,R} \text{ with} & \text{highest SNR}_{n,G}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,G} \stackrel{!}{=} \tilde{f}_i, \\ \text{cbPPG}_{n,G} \text{ with} & \text{highest SNR}_{n,G}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,G} \stackrel{!}{=} \tilde{f}_i, \\ \text{cbPPG}_{n,B} \text{ with} & \text{highest SNR}_{n,G}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,G} \stackrel{!}{=} \tilde{f}_i \end{array}\right\} \rightarrow S_i^{ms} \text{ with } i \in \{1, 2, 3\}$$

Random selection: For testing against choosing only the highest SNRs, assemble analogous monochrome S_i^{mcR} and multispectral S_i^{msR} sets with respective random selection out of available $\text{cbPPG}_{n,G}$ subject to $\hat{f}_{n,G} \stackrel{!}{=} \tilde{f}_i$ and $i \in \{1, 2, 3\}$

Output: Input sets S_i^{mc}, S_i^{ms}, S_i^{mcR} and S_i^{msR} with $i \in \{1, 2, 3\}$ (i.e. three input sets, containing three channels each for the multispectral and monochrome approach)

2.3. Definition of input sets for applied BSS techniques

The multispectral approach and the monochrome approach are characterized by their respective input signals. We used identical ROI sizes and defined a number of three cbPPG as input to the applied BSS techniques for both approaches (further referred to as *input set*). Identical ROI sizes ensure that both approaches can equally benefit from spatial averaging. The choice of three input channels reflects the common number of input channels when RGB videos are used. The definition of input sets S_i^{mc}, S_i^{ms}, S_i^{mcR} and S_i^{msR} i.e. the selection of proper $\text{cbPPG}_{n,color}$, for the monochrome and the multispectral approach, respectively, are detailed in algorithm 1.

Note that for each approach (multispectral and monochrome) we defined three input sets consisting of three input signals each. Such signals served as input to the applied BSS techniques, namely PCA and FastICA (symmetric use, general-purpose tanh-nonlinearity [15]). For each technique we obtained three output components per input set.

2.4. Definition of signal quality

We adopted the spectral measure proposed in [9] as indicator of signal quality. The measure bases on a frequency f_{si} which is considered as usable signal. Given f_{si} a binary mask (BM) is defined by

$$\text{BM}^{f_{si}}(f) = \begin{cases} 1 \text{ if } f \in [f_{si} \pm 5 \text{ bpm}] \\ 1 \text{ if } f \in [2 \cdot f_{si} \pm 5 \text{ bpm}] \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

$\text{BM}^{f_{si}}$ points the spectral indices of the usable signal as well as the first harmonic. According to DeHaan *et al.* [9] the signal-to-noise-ratio (SNR) is calculated from a given amplitude spectrum $X(f)$ by

$$\text{SNR}^{f_{si}} = 10 \log_{10} \left(\frac{\sum_{f=30\text{bpm}}^{240\text{bpm}} \text{BM}^{f_{si}}(f) \cdot X(f)^2}{\sum_{f=30\text{bpm}}^{240\text{bpm}} ((1 - \text{BM}^{f_{si}}(f)) \cdot X(f))^2} \right) \tag{2}$$

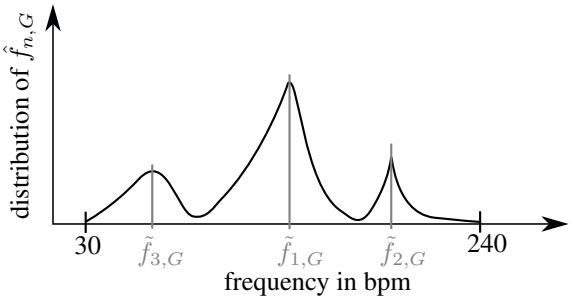

Fig. 3. Schematic view on the distribution of maximum frequencies $\hat{f}_{n,G}$ and derivation of $\tilde{f}_{i,G}$. See algorithm 1 for details.

2.5. Performance evaluation

The efficacy of BSS techniques is evaluated in terms of output components' SNR. The SNR is calculated using the reference heart rate $\hat{f}_{\mathrm{REF}}$. $\hat{f}_{\mathrm{REF}}$ is obtained from the reference PPGs' spectra by detecting the highest peak neighboring the mean heart rate. This mean heart rate is derived from manually annotated beat-to-beat intervals from the reference ECG. The binary mask $\mathrm{BM}^{\hat{f}_{\mathrm{REF}}}$ is generated according to (1). The SNR of every output component for each input subset is estimated by using (2). The highest SNR from a single output component for both approaches (multispectral, monochrome) serves as evaluation metric. Using only the highest SNR reflects our aim to characterize the compared approaches in terms of their potential efficacy.

3. RESULTS

Table 1 shows the results for the multispectral and monochrome approach in terms of absolute SNR and SNR improvement, respectively. SNR improvement is calculated by pairwise subtraction of highest output SNR and highest input SNR within input sets. Note that positive values correspond to an improvement of SNR. Figure 4 shows the SNR improvement as a function of input SNRs. In general, a decreasing BSS performance with increasing input SNR can be deduced. In each configuration, BSS techniques can cause an absolute decrease of SNR. Figure 4 further confirms table 1 in that sense that the monochrome approach outperforms the multispectral approach on average. The respective left clusters of low input SNR can be assigned to cases, where the actual heart frequency was not part of the three most often occurring maximum frequencies during input set construction.

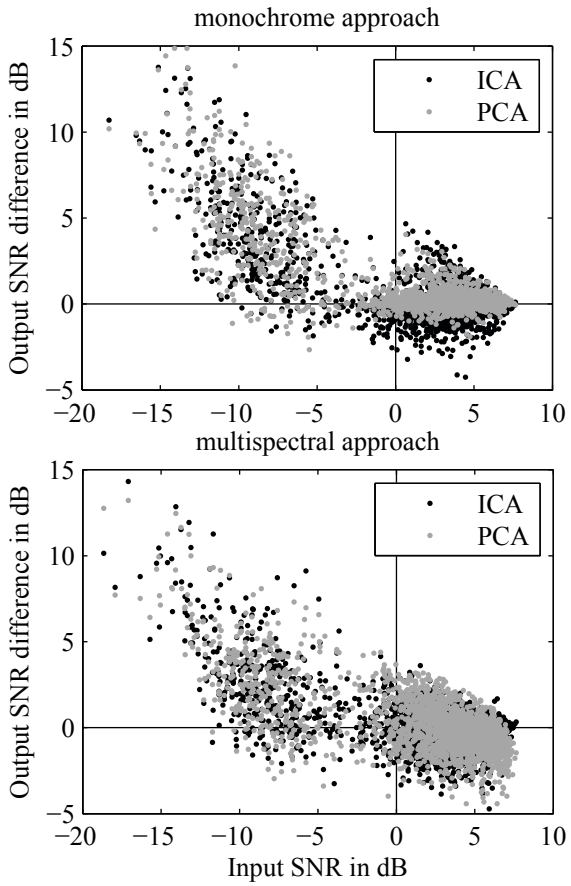

Fig. 4. SNR changes after BSS application (highest input SNR selection) as a function of the input SNR.

4. DISCUSSION AND CONCLUSION

The presented work demonstrates that the application of standard linear BSS techniques on cbPPG not necessarily contributes to a SNR improvement. This holds particularly in case of high input SNR. Moreover, the monochrome approach outperforms common multispectral approaches, especially for PCA. This result may be interpreted as a hint that nonlinear mixing, introduced by using different wavelengths, in fact can degrade the performance of BSS techniques. The average input SNR decrease during random input selection together with a partially better BSS performance is not able to compensate the lower input SNR in terms of a higher output SNR (see table 1). This could be interpreted as justification of exclusively selecting the highest input SNRs. Similar behavior was discovered by applying monochrome approaches using the blue or the red color channel in comparison to the shown performance generated by the green color channel.

Some limitations of our analysis should kept in mind. Firstly, our examination only considers the highest possible SNR. This handling is intended to characterize the potential efficacy of BSS techniques. However, it neglects the permutation indeterminacy of BSS which is essential to the practical application of BSS techniques and was recently addressed by our group [22]. Secondly, any similar comparison has to cope

Table 1. SNR and SNR improvement (pairwise SNR differences from all windows) after applying BSS techniques (in dB).

	cbPPG (before BSS) SNR		PCA SNR improvement		ICA SNR improvement	
	mean (median)	std	mean (median)	std	mean (median)	std
multispectral approach	1.46 (3.12)	5.25	+ 0.33 (– 0.06)	2.03	+ 0.31 (+ 0.05)	1.92
monochrome approach	1.43 (3.14)	5.38	+ 0.96 (+ 0.18)	2.20	+ 0.64 (+ 0.02)	2.23
multispectral approach (random selection)	– 1.38 (– 0.93)	4.04	+ 0.94 (+ 0.76)	1.82	+ 0.55 (+ 0.25)	1.80
monochrome approach (random selection)	0.02 (0.82)	4.14	+ 0.46 (– 0.06)	2.09	+ 0.62 (+ 0.20)	1.88

with the problem of defining a well balanced input, in order to not favor any of the compared methods. We addressed this issue by using equal areas and three input channels. However, other choices are reasonable. For example, multiple ROIs could be combined to larger areas by averaging their pixel intensities in order to cover the same image area for the multispectral and monochrome approach. Future works should vary the conditions of both approaches to further deepen the gained insights.

REFERENCES

[1] M. Huelsbusch and V. Blazek, "Contactless Mapping of Rhythmical Phenomena in Tissue Perfusion Using PPGI," in *Proceedings of the Medical Imaging 2002 (SPIE 4683)*, San Diego, USA, 2002, pp. 110–117.

[2] W. Verkruysse, L.O. Svaasand, and J. Stuart Nelson, "Remote plethysmographic imaging using ambient light," *Optics Express*, vol. 16, no. 26, pp. 21434–21445, 2008.

[3] G. Balakrishnan, F. Durand, and J. Guttag, "Detecting Pulse from Head Motions in Video," in *Proceedings of the IEEE CVPR*, Portland, USA, 2013, pp. 3430–3437.

[4] 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.

[5] H.Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, "Eulerian video magnification for revealing subtle changes in the world," *ACM Transactions on Graphics*, vol. 31, no. 4, pp. 1–8, 2012.

[6] E. Christinaki, G. Giannakakis, F. Chiarugi, M. Pediaditis, and G. Iatraki, "Comparison of Blind Source Separation Algorithms for Optical Heart Rate Monitoring," in *Proceedings of the 4th Mobihealth*, Athens, Greece, 2014, pp. 339–342.

[7] Y. Hsu, Y.L. Lin, and W. Hsu, "Learning-based heart rate detection from remote photoplethysmography features," in *Proceedings of the 39th ICASSP*, Florence, Italy, 2014, pp. 4433–4437.

[8] L. Feng, L. Po, X. Xu, and Y. Li, "Motion Artifacts Suppression for Remote Imaging Photoplethysmography," in *Proceedings of the 19th DSP*, Hong Kong, China, 2014, pp. 18–23.

[9] 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.

[10] 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, pp. 1913–1926, 2014.

[11] B.D. Holton, K. Mannapperuma, P.J. Lesniewski, and J.C. Thomas, "Signal recovery in imaging photoplethysmography," *Physiological measurement*, vol. 34, no. 11, pp. 1499–1511, 2013.

[12] S. Kwon, H. Kim, and K. Suk Park, "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone," in *Proceedings of 34th IEEE EMBS*, San Diego, USA, 2012, pp. 2174–2177.

[13] 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, 2011.

[14] M. Lewandowska, J. Ruminski, T. Kocejko, and J. Nowak, "Measuring pulse rate with a webcam: A non-contact method for evaluating cardiac activity," in *Proceedings of the FedCSIS*, Szczecin, Poland, 2011, pp. 405–410.

[15] A. Hyvärinen, "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis," *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

[16] G.R. Tsouri, S. Kyal, S. Dianat, and L.K. Mestha, "Constrained independent component analysis approach to nonobtrusive pulse rate measurements," *Journal of Biomedical Optics*, vol. 17, no. 7, pp. 077011, 2012.

[17] D. McDuff, S. Gontarek, and R.W. Picard, "Remote Detection of Photoplethysmographic Systolic and Diastolic Peaks Using a Digital Camera," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 12, pp. 2948–2954, 2014.

[18] W. Wang, S. Stuijk, and G. de Haan, "Exploiting Spatial-redundancy of Image Sensor for Motion Robust rPPG," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 415–425, 2014.

[19] Y. Sun, S. Hu, V. Azorin-Peris, S. Greenwald, J. Chambers, and Y. Zhu, "Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise," *Journal of Biomedical Optics*, vol. 16, no. 7, pp. 077010, 2011.

[20] H.E. Tasli, A. Gudi, and M.D. Uyl, "Remote PPG based vital sign measurement using adaptive facial regions," in *Proceedings of ICIP*, Paris, France, 2014, pp. 1–5.

[21] ANSI/AAMI, "American National Standard - Cardiac monitors, heart rate meters, and alarms EC13:2002," .

[22] D. Wedekind, F. Gaetjen, S. Rasche, K. Matschke, H. Malberg, and S. Zaunseder, "Automated Identification of Cardiac Signals after Blind Source Separation for Camera-Based Photoplethysmography," in *Proceedings of the 35th ELNANO*, Kyiv, Ukraine, 2015, pp. 422–427.