UNIVERSITÄT AUGSBURG
Universitätsbibliothek

OPEN ACCESS
OPUS AUGSBURG

Assessment of blind source separation techniques for video-based cardiac pulse extraction

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

Angaben zur Veröffentlichung / Publication details:

Wedekind, Daniel, Alexander Trumpp, Frederik Gaetjen, Stefan Rasche, Klaus Matschke, Hagen Malberg, and Sebastian Zaunseder. 2017. "Assessment of blind source separation techniques for video-based cardiac pulse extraction." *Journal of Biomedical Optics* 22 (3): 035002. https://doi.org/10.1117/1.jbo.22.3.035002.

Nutzungsbedingungen / Terms of use:

licgercopyright

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:
Deutsches Urheberrecht
Weitere Informationen finden Sie unter: / For more information see:
https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/

Assessment of blind source separation techniques for video-based cardiac pulse extraction

Daniel Wedekind,[a,*] Alexander Trumpp,[a] Frederik Gaetjen,[b] Stefan Rasche,[b] Klaus Matschke,[b] Hagen Malberg,[a] and Sebastian Zaunseder[a]
[a]TU Dresden, Institute of Biomedical Engineering, Dresden, Germany
[b]TU Dresden, University Hospital Carl Gustav Carus Dresden, Herzzentrum Dresdem GmbH, Dresden, Germany

Abstract. Blind source separation (BSS) aims at separating useful signal content from distortions. In the contactless acquisition of vital signs by means of the camera-based photoplethysmogram (cbPPG), BSS has evolved the most widely used approach to extract the cardiac pulse. Despite its frequent application, there is no consensus about the optimal usage of BSS and its general benefit. This contribution investigates the performance of BSS to enhance the cardiac pulse from cbPPGs in dependency to varying input data characteristics. The BSS input conditions are controlled by an automated spatial preselection routine of regions of interest. Input data of different characteristics (wavelength, dominant frequency, and signal quality) from 18 postoperative cardiovascular patients are processed with standard BSS techniques, namely principal component analysis (PCA) and independent component analysis (ICA). The effect of BSS is assessed by the spectral signal-to-noise ratio (SNR) of the cardiac pulse. The preselection of cbPPGs, appears beneficial providing higher SNR compared to standard cbPPGs. Both, PCA and ICA yielded better outcomes by using monochrome inputs (green wavelength) instead of inputs of different wavelengths. PCA outperforms ICA for more homogeneous input signals. Moreover, for high input SNR, the application of ICA using standard contrast is likely to decrease the SNR. © *2017 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JBO.22.3.035002]

Keywords: blind source separation; independent component analysis; principal component analysis; video-based vital signs monitoring; camera-based photoplethysmogram; signal-to-noise ratio.

1 Introduction

The contactless acquisition of vital signs allows a convenient medical assessment and enables clinical and out-of-hospital applications. Various systems and principles for contactless measurements have been introduced in recent years.[1,2] Among such approaches, the usage of cameras, referred to as camera-based photoplethysmography or photoplethysmography imaging, is one promising solution to assess the cardiac pulse in a very user-friendly setting.

The acquisition of the cardiac pulse using cameras was first demonstrated by Huelsbusch and Blazek.[3] Meanwhile, many researchers have addressed the camera-based photoplethysmogram (cbPPG), most often to assess the heart rate.[4–10] The most important drawback of the technique is its susceptibility to artifacts induced by movements and changes in illumination. Sophisticated image and signal processing techniques are required to cope with such factors and facilitate the camera-based assessment of cardiac pulse even under real-world conditions.

Poh et al.[6] were the first to use blind source separation (BSS) algorithms in the context of the cbPPG. Since then BSS algorithms became a core part of signal processing schemes to extract the heart rate from cbPPG recordings. BSS aims at separating the desired signal content (i.e., cardiac pulse) from noise and artifacts by means of decorrelation and utilizing the concept of statistical independence. Principal component analysis (PCA) and independent component analysis (ICA)[11] realized as Joint approximation diagonalization of eigen-matrices (JADE)[12] or FastICA[13] are typical BSS techniques that have been widely applied to cbPPGs.

Common approaches[6,14–19] use different color channels [typically red, green, and blue (RGB)] extracted from regions of interest (ROI), typically the face, as input to PCA or JADE ICA. FastICA has also been applied to RGB signals[8,19,20] and achieved a slightly better performance in comparison to other ICA algorithms.[8] Tsouri et al.[21] proposed a constrained ICA for RGB information of a face ROI. Other researchers have further developed the idea of applying multispectral cbPPG to PCA/ICA but have used alternatives to RGB, namely combinations of RGB with orange and cyan channels or chrominance as well as hue and infrared-based signals, respectively.[9,15,22–25]

In addition to wavelength-based considerations, more selective ROI choices, such as reducing the face ROI to a more concise area, have been outlined in the context of PCA/ICA.[8,10,16,19,22,26] These approaches seek to exclude regions that are not supposed to contribute with useful signals but can introduce distortions, e.g., by mouth movements during speaking/smiling or blinking eyes.[19,27] Approaches, which are described in literature, typically rely on a spatial preselection and use multispectral information (RGB) as input to BSS techniques. Moreover, a monochrome cbPPG, extracted from the forehead, was used as input for spatio-temporal ICA.[26] Wang et al.[23] alternatively addressed spatial selection without using explicit face detection. The authors utilized the temporal behavior of pixel traces to distinguish skin-like areas showing temporally periodic content from motion-like content. Even Guazzi et al.[28] pursued

*Address all correspondence to: Daniel Wedekind, daniel.wedekind@mailbox.tu-dresden.de

the idea of spatially selecting ROIs according to the local distribution of signal quality.

Despite the frequent multispectral BSS use, there is no consensus on performance improvements by using BSS techniques with multispectral inputs. In particular, Kwon et al.[17] described a blurred spectral peak after applying RGB ICA as well as an increased heart rate error. Christinaki et al.[8] identified only subtle improvements but similar heart rate errors with/without using RGB ICA. Feng et al.[10] showed a lack of robustness by applying standard approaches that use ICA with RGB channels.

A possible reason for the inadequacy of applied BSS techniques can be the assumption of a linear mixing process[13] of available sources in standard PCA/ICA. In particular, the wavelength-dependent penetration depth into human skin[3] could introduce nonlinear mixing behavior, which may degrade the performance of BSS algorithms. Other factors may also impact the success of BSS techniques in extracting pulsatile signals. For example, an effect of ROI size was addressed by Mannapperuma et al.[29] in the context of multispectral BSS.

Based on BSS's multiple use and high capacity on the one hand, conflicting findings and oppositional statements on the other hand, this contribution investigates the performance of BSS to enhance the cardiac pulse from cbPPGs in dependency to varying input data characteristics. We use standard BSS techniques to compare BSS's application on multispectral as well as monochrome inputs to identify beneficial conditions for efficient usage of these techniques. To facilitate an appropriate comparison, we further introduce a spatial ROI selection based on cbPPG's signal quality. Finally, this contribution shall allow for a targeted application of BSS techniques to cbPPG recordings to achieve a higher signal quality and make the technique applicable under real-world conditions.

2 Materials and Methods

2.1 *Patients and Clinical Setting*

The data for this study were gathered within the scientific project "CardioVisio—Contactless aquisition of vital parameters." Measurements were carried out at the cardiac surgical intensive care unit at the Heart Center Dresden, University Hospital Carl Gustav Carus, TU Dresden, Dresden, Germany. The project was approved by the Institutional Review Board of the TU Dresden (IRB00001473, EK168052013). Patients after elective cardiac surgery were included if they gave written, informed consent prior to surgery. Video data were recorded during the immediate recovery from surgery after admission at the intensive care unit.[30] Postoperative care followed clinical standards including mechanical ventilation and external cardiac pacing by temporal atrial and ventricular wires adjusted to intrinsic cardiac rhythm and haemodynamic needs. A four-lead electrocardiogram (ECG) and finger photoplethysmogram (PPG) were simultaneously recorded at 300 Hz and used to derive the reference heart rate.

2.2 *Video Recording and Selection*

Video data were recorded using an industrial camera (IDS UI-3370CP-C-HQ, IDS Imaging Development Systems GmbH, Obersulm, Germany, 100 fps, 420×320 pixels, RGB 3×12 bit). The camera was placed at a distance of ~60 to 100 cm to patients' faces. Clinical ceiling fluorescent lamps served as the primary illumination source. However, the luminous color, intensity, and homogeneity of the illumination varied across the measurements due to varying patient positions with respect to the illumination, varying room geometries, and entering daylight. Therefore, a broad range of illumination characteristics is covered by the used data.

To use only suitable data for further analysis, we restricted our analysis to data segments that showed high-quality reference PPGs to correctly identify the true heart rate. Furthermore, data segments with severe cardiac disorders were excluded. Only continuous segments with a minimum length of 500 s (one per patient) were considered for further processing. Based on such criteria, overall recordings of 18 patients (13 males and 5 females; 30 min per recording) were selected from a larger collective of 76 patients. The selected material included a total of about 6 h video data (average length 1200 ± 400 s per patient). The selection did not consider video quality, i.e., slight patient motion as well as illumination inadequacies (changes or insufficient lightning) persisted in the dataset.

2.3 *Signal Extraction*

The selected video data were processed in windows of 10 s resulting in 2197 windows (106 ± 37 per patient). The cbPPGs were extracted in three ways as depicted in Fig. 1. To allow a spatial selection of desired ROIs, every video frame was covered by 25×19 overlapping square ROIs (50% overlap at each direction) of 32×32 pixels. The ROI placement is indicated in Fig. 1(a). The ROI size was chosen since own prior investigations addressing the relationship between ROI size and signal quality [signal-to-noise ratio (SNR)] using a comparable technical setup (i.e., camera sensor, image resolution, and camera distance to subject) showed that no higher signal quality can be obtained by further increasing the ROI size. [Other investigations,[31] which showed an appropriate ROI size for pulse wave extraction to be larger (100 to 150 pixels ROI side length), used a higher image resolution.] To compare the spatial ROI selection to standard ROI selection prior to BSS,[6,8,22] manually annotated ROIs were used to extract the cbPPG of the complete face ($\mathrm{ROI_F}$) and the forehead-cheeks region ($\mathrm{ROI_{FhC}}$), respectively. See Figs. 1(b) and 1(c) for exemplary ROI annotations. cbPPG was extracted from each ROI_n (with $n = 1,2,\ldots,475$) as well as from $\mathrm{ROI_F}$ and $\mathrm{ROI_{FhC}}$ at every wavelength by averaging its pixels values[4] for each frame. See Fig. 2 for exemplary signals together with the reference PPG.

2.4 *Signal Processing*

2.4.1 *Preprocessing*

Each 10-s cbPPG signal (from ROI_n, $\mathrm{ROI_F}$, and $\mathrm{ROI_{FhC}}$) was normalized by a three-step procedure. The signal was linearly detrended followed by 0.5 Hz high-pass filtering (fifth-order Butterworth) to limit low-frequency content below an expected heart rate.[32] Furthermore, the signal amplitude was normalized by subtracting its mean and dividing the result by signal's standard deviation. Suchlike preprocessed $\mathrm{cbPPG}_{n,\mathrm{color}}$ (and $\mathrm{cbPPG_{F,color}}$, $\mathrm{cbPPG_{FhC,color}}$) were used for further processing.

2.4.2 *Definition of Inputs to Blind Source Separation*

To evaluate BSSs' benefit in consideration of varying inputs, different input sets S were defined. Each input set contained three input signals to reflect the common number of input channels when RGB videos are used. The input sets differed

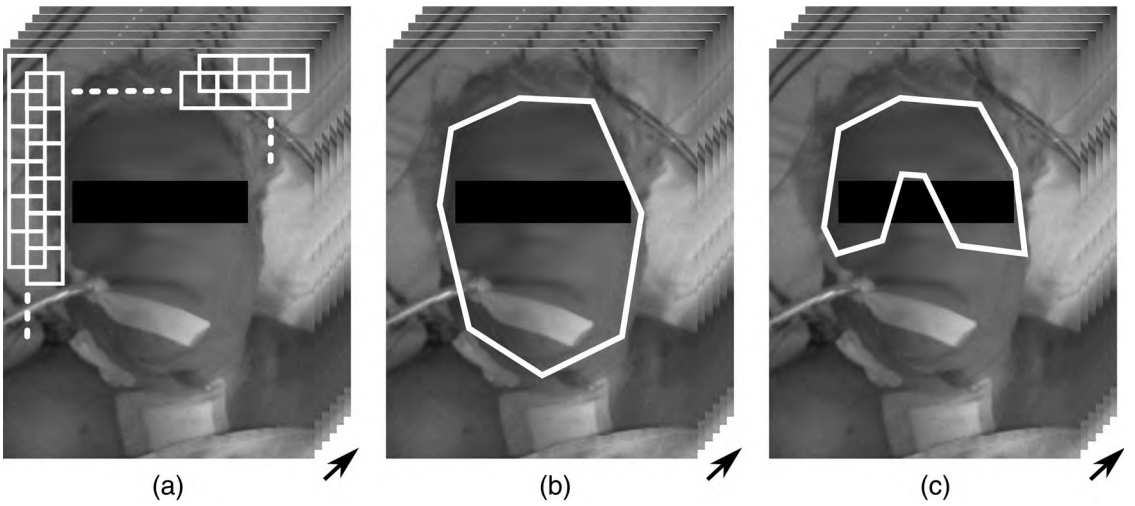

Fig. 1 ROI selection on video frames. White lines indicate ROI borders. (a) Evenly distributed 32×32 pixel ROIs with 50% overlap, (b) manually annotated ROI including the face $\mathrm{ROI}_{\mathrm{F}}$, and (c) manually annotated ROI including forehead and cheeks $\mathrm{ROI}_{\mathrm{FhC}}$.

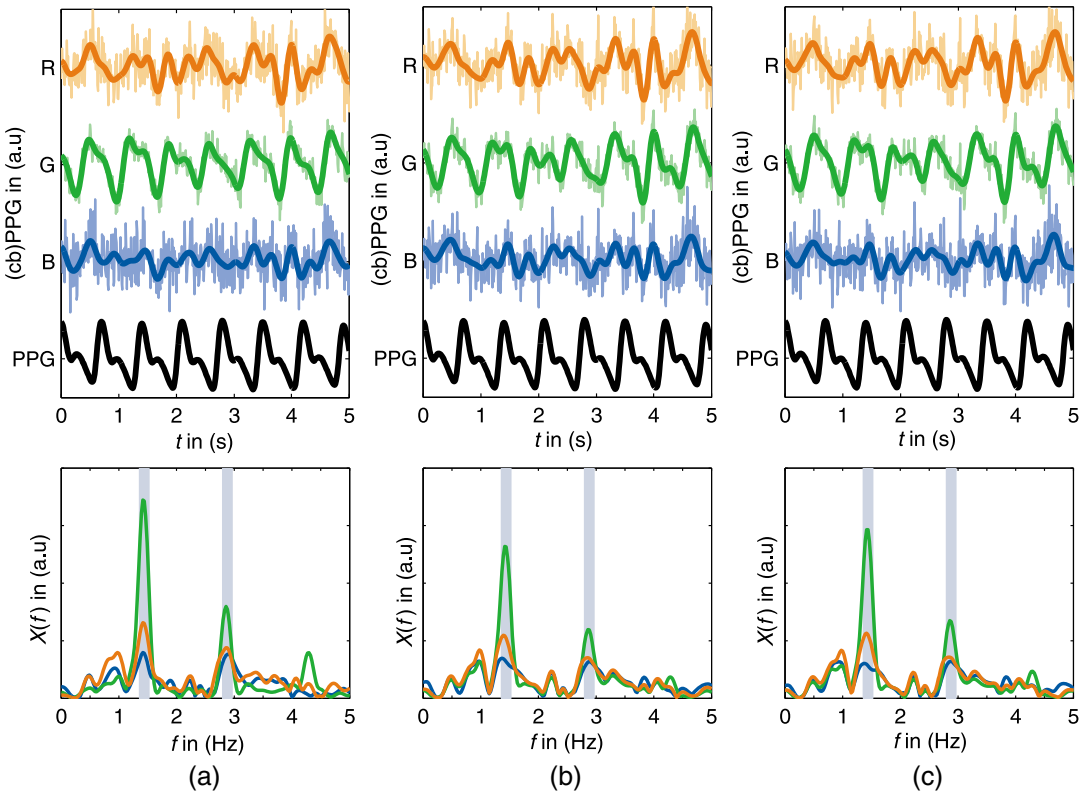

Fig. 2 Exemplary signal extraction. Sample signal excerpts [normalized and normalized +4 Hz low-pass filtered (bold signal) versions] and amplitude spectra from different ROIs. (a) 32×32 pixel ROI, (b) manually annotated ROI including the face, and (c) manually annotated ROI including forehead and cheeks. The true heart rate and its harmonic (± 5 bpm) are indicated by the colored areas in the back of the spectra. Colors of the respective spectra are according to the time signals.

regarding the wavelength(s) to be used and the frequency content of chosen ROIs. Regarding the wavelength, we distinguished using cbPPGs from the green channel (monochrome approach) or using RGB channels (multispectral approach). Regarding the frequency content, we distinguished using ROIs, which showed equal dominant frequencies and using ROIs, which showed differing dominant frequencies ("dominant frequency" refers to the location of the global maximum in the fast Fourier transform of the cbPPG signal from a ROI after applying a Hanning window and zero padding to 4096 points). To select three ROIs from ROI_n, we further distinguished between a deterministic choice and a random choice. The

deterministic choice selects three ROIs that showed the highest signal quality. For the random selection, three ROIs, which possess the desired frequency content, were chosen randomly (independently of its signal quality). To give an example, the selection "equal dominant frequency + deterministic choice" means that ROIs were first ordered according to their dominant frequency. Afterward, the ROIs that showed the highest SNRs within the desired dominant frequencies were selected. Note that it has to be defined which dominant frequency is the desired one. As there is no unambiguous answer to this question, we decided to create input sets for the three most often occurring dominant frequencies (i.e., for "equal dominant frequency" always three different input sets were used). A detailed mathematical definition of equal/differing dominant frequency and the definition of signal quality are depicted in Appendix A (a graphical overview on the selection process is also provided in Fig. 8). Table 1 summarizes the resulting input sets.

2.4.3 *Blind Source Separation Processing*

Every set S was further processed with PCA and ICA, respectively.

The PCA was computed without dimension reduction by conducting the singular value composition of the covariance matrix of each set.[11] Let us consider a set

$$S = [\mathrm{cbPPG}_{\#,\mathrm{color}}, \mathrm{cbPPG}_{\#,\mathrm{color}}, \mathrm{cbPPG}_{\#,\mathrm{color}}]^T = \mathbf{x}, \tag{1}$$

where $\mathrm{cbPPG}_{\#,\mathrm{color}}$ denotes a preprocessed cbPPG of the dimension $1 \times m$ (window length m) extracted using $\mathrm{ROI}_{\#}$. A transformation $\mathbf{B}$ was computed that satisfies

$$\Lambda = \mathbf{B}^T E[\mathbf{x}\mathbf{x}^T]\mathbf{B}, \tag{2}$$

Table 1 Definition of input sets S for BSS.

Set ID	Wavelength	Frequency content	Selection	Overall size
MC1	Green	Equal dominant frequency*	Highest SNR	3× (32 × 32)
MC2	Green	Differing dominant frequencies	Highest SNR	3× (32 × 32)
MS1	RGB	Equal dominant frequency*	Highest SNR	3× (32 × 32)
MS2	RGB	Equal dominant frequency*	Highest SNR	3× [3 × (32 × 32)]
MCR	Green	Equal dominant frequency*	Random choice	3× (32 × 32)
MSR	RGB	Equal dominant frequency*	Random choice	3× (32 × 32)
F	RGB	n/a	$\mathrm{ROI_F}$	Whole face
FhC	RGB	n/a	$\mathrm{ROI_{FhC}}$	Suitable regions

Note: See Appendix A for a detailed mathematical description.
*Note that in case of equal dominant frequency, three different input sets were evaluated for the three most occurring dominant frequencies.
Set IDs refer to MC: monochrome, MS: multispectral, R: random SNR and standard ROI sets from F: face and FhC: forehead and cheeks.

with Λ as a 3×3 diagonal matrix with squared singular values (equivalent to the eigenvalues of the covariance matrix) as diagonal entries and $E[\mathbf{x}\mathbf{x}^T]$ as covariance matrix of $\mathbf{x}$. Accordingly, the orthogonal transformation matrix $\mathbf{B}$ consists of the covariance matrix's eigenvectors of S and diagonalizes $E[\mathbf{x}\mathbf{x}^T]$. The PCA output $\mathbf{y}_{\mathrm{PCA}}$ was obtained by

$$\mathbf{y}_{\mathrm{PCA}} = \mathbf{B}^T\mathbf{x}. \tag{3}$$

The output $\mathbf{y}_{\mathrm{PCA}}$ was further normalized (according to the preprocessing) prior to evaluation.

As representative ICA method, the FastICA algorithm[13] was chosen to compute a transformation $\mathbf{W}$, which aims at statistical independence of the output $\mathbf{y}_{\mathrm{ICA}}$

$$\mathbf{y}_{\mathrm{ICA}} = \mathbf{W}\mathbf{x}, \tag{4}$$

because Christinaki et al.[8] have shown a superior performance of FastICA compared to JADE for processing the cbPPG to extract the heart rate. Prewhitening of $\mathbf{x}$ was used. Prewhitening works according to the PCA while ensuring $E[\mathbf{x}\mathbf{x}^T] = \mathbf{I}$ the identity matrix, which addresses an orthonormal output. The FastICA was initialized with a fixed random demixing matrix $\mathbf{W}$, which was used as starting point for every processed segment. The FastICA was symmetrically conducted for dimension preservation between $\mathbf{x}$ and $\mathbf{y}$. The standard tanh-nonlinearity was applied as contrast function, which supports super-Gaussian source extraction,[13] as indicated for the PPG signal by Tsouri et al.[21] Simultaneously, it does not aim at highly super-Gaussian signals, which is a consequence of Morris et al.[33] selecting the PPG component after ICA by using the lowest kurtosis of the components.

2.5 *Evaluation Metrics*

The signal quality of inputs, i.e., of each single cbPPG, and outputs, i.e., each independent component/principal component, was assessed by a spectral SNR, which was proposed by de Haan and Jeanne.[14] The SNR considers the true heart rate $f = f_{\mathrm{PPG}}$, which is gained from reference recordings. Based on the true heart rate, a binary mask $\mathrm{BM}^{f_{\mathrm{PPG}}}$ was defined according to

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

$\mathrm{BM}^{f_{\mathrm{PPG}}}$ sustains the spectral indices of the heart rate as well as its first harmonic. The precision ± 5 bpm refers to the accuracy demanded for heart rate meters specified in ANSI/AAMI EC13:2002.[32] The SNR was calculated from a given amplitude spectrum $X(f)$ by

$$\mathrm{SNR}^{f_{\mathrm{PPG}}} = 10 \log_{10}\left(\frac{\sum_{f=30\ \mathrm{bpm}}^{240\ \mathrm{bpm}} \mathrm{BM}^{f_{\mathrm{PPG}}}(f) \cdot X(f)^2}{\sum_{f=30\ \mathrm{bpm}}^{240\ \mathrm{bpm}} [1 - \mathrm{BM}^{f_{\mathrm{PPG}}}(f)] \cdot X(f)^2}\right). \tag{6}$$

For this contribution, the true heart rate f_{PPG} was estimated by first averaging manually annotated beat-to-beat intervals from the reference ECG to obtain f_{ECG}. Afterward the closest frequency peak in the amplitude spectrum of the reference PPG was searched ($f_{\mathrm{PPG}} \approx f_{\mathrm{ECG}}$).

Input sets S were typically formed of cbPPGs with three different dominant frequencies, respectively, signal qualities (see Appendix A). Accordingly, one obtains 3×3 input signals as well as 3×3 output signals after BSS with respective $\mathrm{SNR}^{f_{\mathrm{PPG}}}$ values (S^{FhC} and S^{F} only contain 1×3 input and output signals). Since the automated selection of an appropriate output component is not in the scope of this contribution (but has been considered by our group),[34] we always selected the best possible SNR of inputs or BSS outputs for the further evaluation. This procedure comprises assessing only the highest input and output SNR out of three dominant frequencies.

2.6 *Statistical Assessment*

This contribution investigates the performance of BSS to enhance the cardiac pulse from cbPPG in dependency to varying input data characteristics. To that end, the results and respective statistical analyses are broken down into three aspects:

1. Do varying input sets provide differing input SNR?
2. Which benefit can generally be expected from applying BSS to input sets of varying constitution and quality?
3. Given an input, which of the applied BSS techniques is to be preferred?

The first aspect was addressed by a comparison of input SNRs using one-way analysis of variance (ANOVA). As posthoc tests, a selection of 13 pairwise t-tests with Bonferroni–Holm correction was applied (see Table 2 for the test selection). We selected pairwise comparisons of the deterministic small-sized ROIs (S^{mc1}, S^{mc2}, S^{ms1}, and S^{ms2}) plus the tests of the random small-sized ROI selection matched with S^{mc1} and S^{ms1} (i.e., S^{mcR} and S^{msR}). Moreover, pairwise comparisons of all deterministic multispectral sets were conducted (S^{ms1}, S^{ms2}, S^{FhC}, and S^{F}). To avoid the large sample size to determine the statistical results,[35] we calculated subjects' means and applied the statistical analysis to the mean values (i.e., $n = 18$ for ANOVA and posthoc analysis). Furthermore, the effect size measure Hedges g,[36] including the 95% confidence interval (CI) of g,[37] is used as standardized mean difference between groups. The interpretation of g is straight forward: given a comparable CI, the larger the effect size, the bigger the impact of an experimental variable.[36,38] As contextual information is required to interpret effect sizes in terms of absolute values,[38] we abstain from interpreting g's absolute value. To define which effects are relevant for further discussion, we instead introduce the concept of CI consistency: an effect is regarded as consistent, if the CI of a given g is completely positive or negative, respectively (denoted as consistent effect).

Table 2 Results of the comparison of input sets.

Set ID 1	Set ID 2	p	g
MS1	F	<0.001	2.01 [1.23, 2.78]
MS2	F	<0.001	1.86 [1.15, 2.56]
MS1	MSR	<0.001	1.57 [0.95, 2.19]
MS1	FhC	<0.001	1.34 [0.77, 1.91]
MS2	FhC	<0.001	1.21 [0.71, 1.72]
MC1	MCR	<0.001	1.09 [0.60, 1.59]
MC1	MS2	0.06	0.14 [0.03, 0.25]
MC2	MS2	0.06	0.14 [0.03, 0.25]
MS1	MS2	0.06	0.12 [0.01, 0.22]
MC1	MS1	0.01	0.02 [0.01, 0.04]
MC2	MS1	0.01	0.02 [0.01, 0.04]
MC1	MC2	n/a	n/a
F	FhC	<0.001	−0.51 [−0.74, −0.27]

Note: Pairwise results: Bonferroni–Holm corrected p values from posthoc pairwise t-tests, effect size g and 95% CIs of g[37] in brackets.

For the second aspect, the comparison of output SNRs does not suffice the needs for a statement on the BSS performance, because the BSS performance ΔSNR (difference between output and input SNR) and the output SNR must be assumed to be heavily dependent on the input SNR. A statement, which bases solely on output SNRs, thus, might favor the output featuring the highest input SNR and will not provide a meaningful statement on BSSs' performance. A dependence on the input SNR implies using the input SNR as covariate, i.e., analysis of covariance (ANCOVA). A suchlike analysis provides a meaningful statement on the benefit of applying BSS. However, a poor input SNR will be favored by this analysis as a large improvement can be obtained while the outcome still might be worse than using another input. For high input SNRs, on the other hand, the potential improvement, which can be gained by BSS, is limited as the output SNR is bounded. For such reasons, we decided to combine both analyses, ANOVA and ANCOVA. An ANOVA to the output SNRs and respective posthoc tests could be used as described before. For ANCOVA, the input SNR served as covariate. As posthoc tests for ANCOVA, t-tests with centered mean[39] were applied and Hedges g was calculated (again subjects means were used to avoid that large sample sizes determine the statistical results). The selection of posthoc tests applied for the first aspect was also used for the second aspect. For the pairwise comparison of two settings, one of them was regarded as superior and relevant for further discussion if ANOVAs and ANCOVAs posthoc test show consistent effects with the same sign (both CI entirely positive or negative, for a detailed example, see the description of results in Sec. 3.2). Additionally, the question if a significant ΔSNR could be achieved by applying a BSS algorithm to a single input was answered by pairwise t-tests of the differences between output and input SNRs.

The third aspect was assessed by pairwise t-tests of ICA and PCA outputs. Again, subject means were used and additionally Hedges g was considered.

3 Results

3.1 *Signal Quality of Input Sets*

The ANOVA yields a highly significant difference between inputs ($p < 0.001$). Table 2 shows the results of the pairwise comparison between input sets (i.e., significance from posthoc test and results on Hedge's g, the input SNR values can be found

Table 3 Pairwise comparison between ICA and PCA outputs.

Set ID	Input SNR	PCA	ICA	p	g
MC1	2.56 ± 2.50	3.01 ± 2.21	2.69 ± 2.37	<0.001	0.14 [0.07,0.20]
MC2	2.56 ± 2.50	2.10 ± 2.31	2.92 ± 2.48	<0.001	−0.33 [−0.53, −0.13]
MS1	2.50 ± 2.52	2.45 ± 2.06	2.36 ± 2.33	0.53	0.04 [−0.08, 0.16]
MS2	2.18 ± 2.76	2.68 ± 2.24	2.74 ± 2.56	0.67	−0.03 [−0.15, 0.09]
MCR	-0.74 ± 3.34	-0.86 ± 2.88	-0.63 ± 2.94	<0.01	−0.08 [−0.14, −0.02]
MSR	-2.33 ± 3.43	-1.09 ± 3.13	-1.60 ± 3.02	<0.001	0.16 [0.09,0.24]
F	-4.05 ± 3.74	1.70 ± 2.99	2.18 ± 3.25	<0.01	−0.15 [−0.25, −0.05]
FhC	-2.05 ± 3.97	2.85 ± 3.01	2.98 ± 3.16	0.32	−0.04 [−0.12, 0.04]

Note: SNR in dB shown as mean ± standard deviation, p values from pairwise t-tests and effect size g and 95% CIs of g[37] in brackets.

in Table 3). g shows consistent effects between various inputs. posthoc t-tests confirm significant differences between various inputs.

The deterministic automated selection of 32×32 pixels input ROI_n and respective $\mathrm{cbPPG}_{n,\mathrm{color}}$ according to Appendix A provides higher quality cbPPGs compared to random selection of ROI_n of the same size or standard ROI_F and $\mathrm{ROI}_\mathrm{FhC}$, respectively. Figure 3 confirms this finding graphically. Omitting the random selection, an increased ROI size comes along with a decreased input SNR [for example, on a multispectral input ROI (MS1 < MS2 < FhC < F)]. Regarding the random selection of $\mathrm{cbPPG}_{\mathrm{n,color}}$, the input signal quality in both cases, the monochrome and multispectral random case, is significantly worse than the one achieved by the deterministic selection.

3.2 *Blind Source Separation Performance on Different Input Sets*

Figure 3 shows boxplots of individually taken BSS performances, namely the patient-wise averaged output and input SNRs including the statistical measure of the pairwise difference. Especially for a low input SNR, statistically significant SNR improvements are obtained for both PCA and ICA. PCA moreover significantly improves the SNR of the monochrome set MC1 with equal dominant frequencies whereas ICA significantly improves the SNR of the monochrome set MC2 with different dominant frequencies. ICA also shows a highly significant SNR improvement on MS2. Figure 4 illustrates a distinct dependence of the BSS performance, i.e., the obtained SNR difference ΔSNR, on the input SNR. ANCOVA proves that there are no significant differences in the strength of that dependence (i.e., no differences in the slope of separate regression lines with $p = 0.27$ for PCA and $p = 0.15$ for ICA). ANCOVA further proves highly significant differences in terms of adjusted means (i.e., significant differences in the intercepts of parallel regression lines with $p < 0.001$ for PCA and ICA). ANCOVA posthoc tests confirm significant differences between various outputs (see Fig. 5). ANOVA for the output SNR yields a highly significant difference between outputs ($p < 0.001$ for PCA and ICA). ANOVA posthoc tests confirm significant differences between various outputs (see also Fig. 5). Figure 5 gives a comprehensive overview on the posthoc results of ANCOVA and ANOVA together with Hedges g including its 95% CIs. As stated before, BSS's application can be considered as superior in a pairwise comparison if g and its CI are consistent and show the same direction (sign of g) for both ANOVA and ANCOVA posthoc tests.

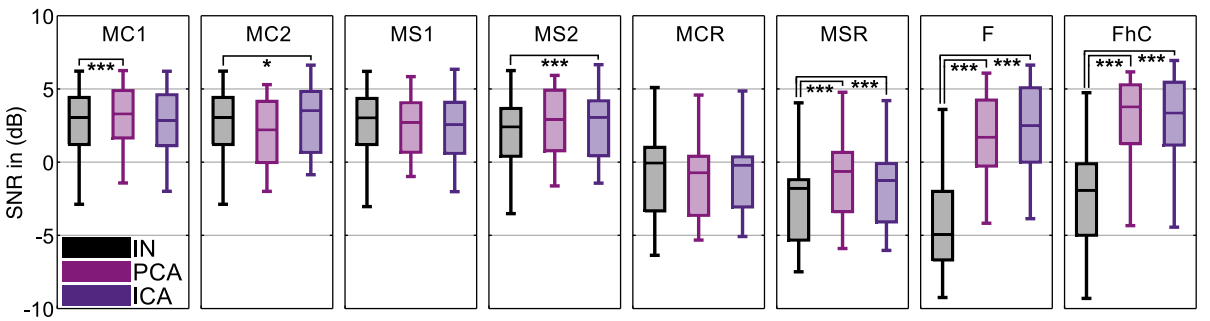

Fig. 3 Boxplots showing input and output SNRs for BSS processing. Patient-wise averaged input and output SNRs ($n = 18$) of all sets S according to "selection of inputs to BSS." Innerbox lines indicate the median. The maximum whisker length is set to 150% the interquartile range outside the interquartile borders. No outliers were found by applying this criterion. Significance of differences ΔSNR between output and input SNRs by pairwise t-tests is indicated between the boxes (in case of significance) denoting p values as: * denotes $p \leq 0.05$, ** denotes $p < 0.01$, and *** denotes $p < 0.001$.

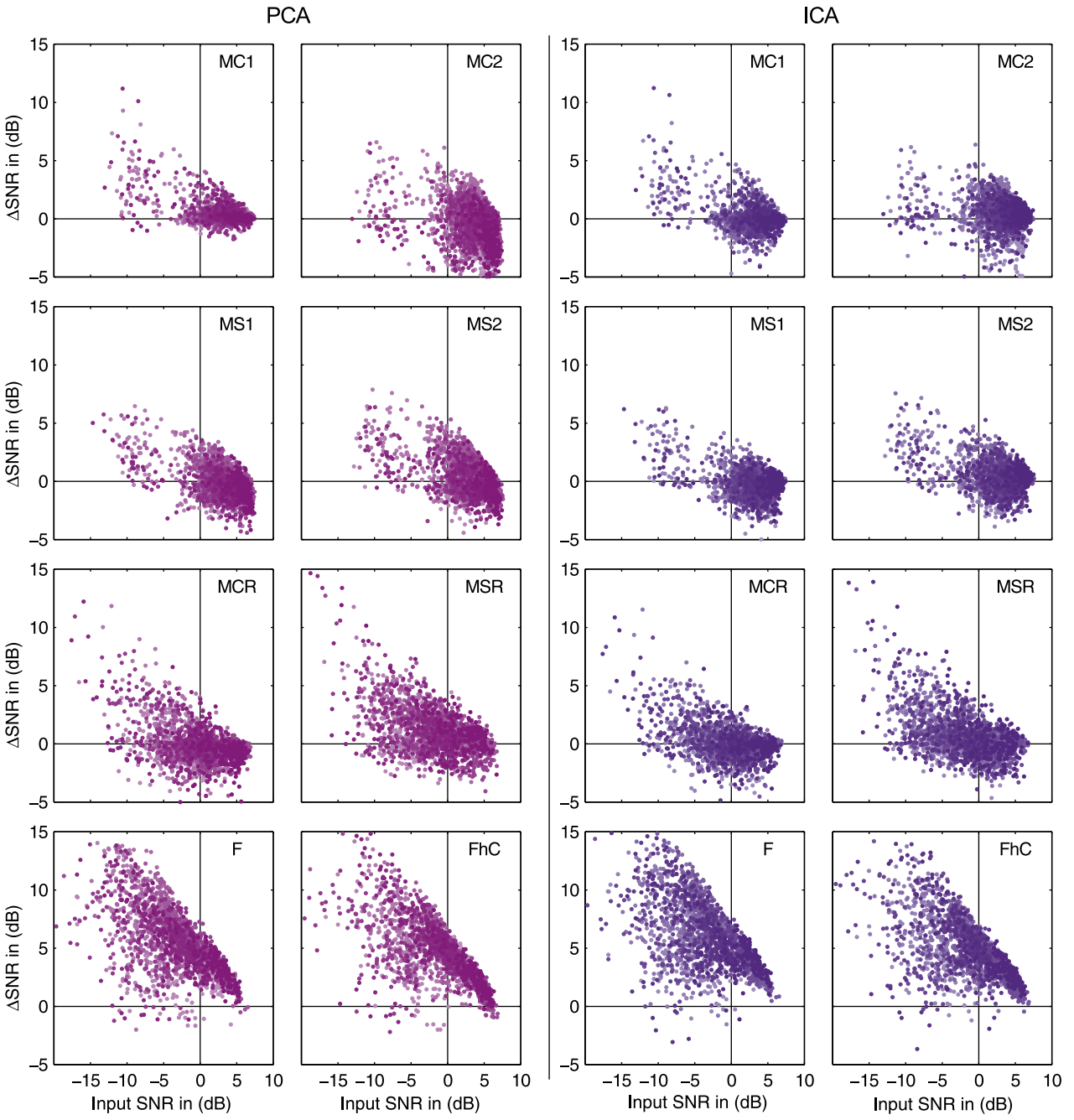

Fig. 4 BSS performance regarding the SNR. SNR changes (output–input) by BSS processing as a function of the input SNR. Every point depicts the performance of a single 10-s window. Color gradation indicates single patients.

As can be seen in Fig. 5, not every pairwise comparison shows relevant differences. The comparison of ICA results for ANCOVA and ANOVA for the standard approaches of sets FhC and F may serve as an example for the interpretation of Fig. 5. Both sets show significant differences in their input SNR (y-axis) with a negative effect size $g = -0.51$ (input SNR of S^{F} is smaller than input SNR of S^{FhC}). Considering the comparison of adjusted means, i.e., ANCOVA's posthoc test, a consistent effect in favor of S^{F} is found (indicated by the entirely positive CI of g). Concerning the comparison of the output SNR without adjustment, i.e., ANOVA's posthoc test, also a consistent effect, is obtained. However, this time the CI is entirely negative, which indicates that S^{FhC} provides a significantly higher output SNR. Apparently, the lower input SNR of S^{F}, together with a bounded SNR, favors S^{F} within ANCOVA's posthoc test. As example for a pairwise comparison, which shows a relevant difference, readers may be referred to MC1 versus MCR using ICA: besides significantly differing input SNRs (with $g = 1.09$, input SNR of S^{mc1} is higher than input SNR of S^{mcR}), both ANOVA and ANCOVA show consistent effects, either of them is positive. It indicates the better performance of ICA on homogeneous (frequency) inputs of best available SNR compared to homogeneous (frequency) input of random SNR regardless of the input SNR.

The results depicted in Fig. 5 can be summarized as follows. Consistent effect size measures, according to the definition

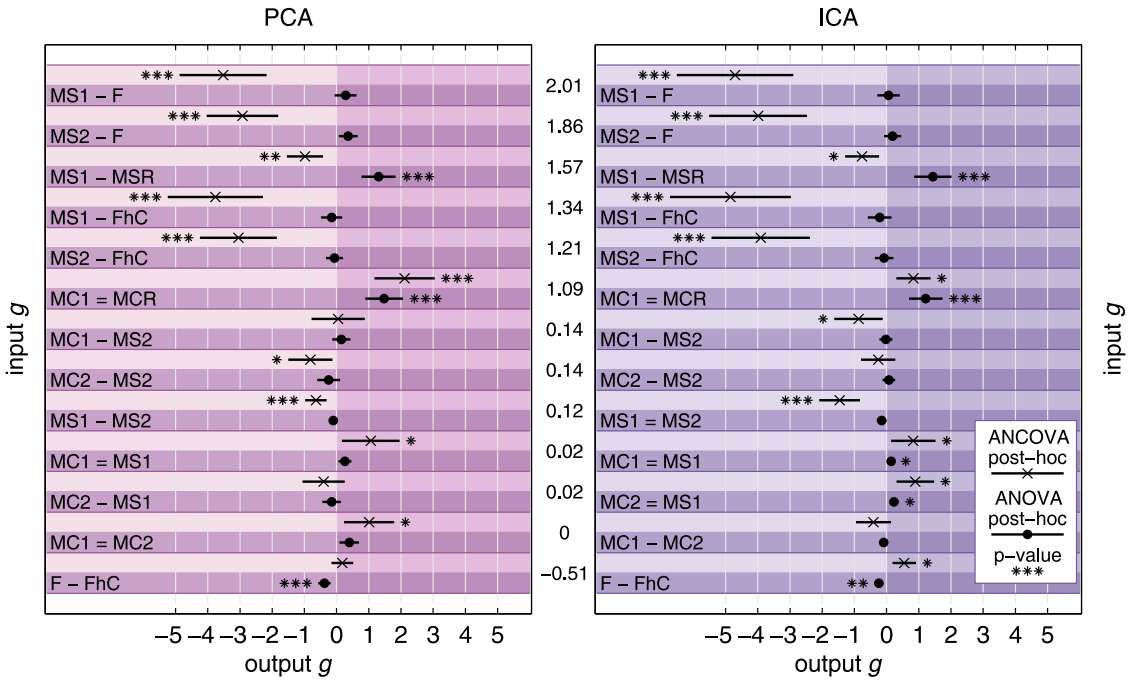

Fig. 5 ANOVA and ANCOVA posthoc statistics for PCA and ICA output sets. The figure shows pairwise *t*-test results (ANOVA) and pairwise *t*-test results using mean-centered-independent variable (ANCOVA) of patient-wise averaged ($n = 18$) SNR outputs. The results are characterized by its effect size g (• for ANOVA and x for ANCOVA) as well as the 95% CI of g[37] stated as line length. Sign of g indicates the effect direction according to the mean difference obtained by always setting the second set as subtrahend. The set comparisons are vertically (y-axis) ordered according to the effect size g of the input comparison of the same sets. Significance according to the Bonferroni–Holm corrected p value is denoted by * $p \leq 0.05$, ** $p < 0.01$, and *** $p < 0.001$. Consistency of an effect is denoted by = between the set names, - otherwise.

(i.e., fully positive, respective negative CI of g), can be found for both PCA and ICA, for example, assessing the adjusted means (ANCOVA) of S^{FhC} when compared to high SNR inputs S^{ms1} and S^{ms2}. Thus, in case of poor input SNRs, PCA and ICA both can be efficiently applied. However, relevant effects, according to the definition, i.e., considering ANCOVA and ANOVA at the same time, are found for PCA with $S^{\mathrm{mc1}} > S^{\mathrm{mc2}}$ and $S^{\mathrm{mc1}} > S^{\mathrm{ms1}}$ as well as for ICA with $S^{\mathrm{mc1}} > S^{\mathrm{ms1}}$, $S^{\mathrm{mc2}} > S^{\mathrm{ms1}}$, and $S^{\mathrm{ms2}} > S^{\mathrm{ms1}}$. Accordingly, PCA is considered performing worse on inhomogeneous frequency inputs compared to homogeneous frequencies. Moreover, PCA and ICA are considered performing better on monochrome inputs compared to multispectral inputs of the same ROI size. Increasing the ROI size (from S^{ms1} to S^{ms2}) also favors ICA performance.

3.3 *Blind Source Separation Performance Given an Input Set*

Table 3 shows the input and output SNR for all input sets and BSS techniques (PCA and ICA). Note that for equal dominant frequencies, originally three different input sets were available but only the highest SNR is shown (independently from the used dominant frequency; in 44%/36%/20% the first/second/third dominant frequency yielded the highest SNR). Additionally, Fig. 3 shows the according boxplots of input and output SNRs including pairwise t-test results in case of significance ($p \leq 0.05$). Segment-wise ΔSNR is also shown as a function of the respective input SNR in Fig. 4.

Comparing the performance of PCA and ICA on a given input (see Table 3), PCA works significantly better on homogeneous inputs S^{mc1} (dominant frequency and wavelength). ICA works significantly better on inhomogeneous inputs S^{mc2} (dominant frequency) as well as for inhomogeneous ROIs (S^{F}).

4 Discussion

4.1 *Spatial Selection versus Standard Approaches*

In real applications, ROIs for cbPPG extraction have to be selected by using automated video processing algorithms. Typically, rectangular ROIs using frontal face classifiers for detection and (stabilized) tracking of the complete face[6] or selective rectangular ROIs of smaller face parts further using facial landmarks[10,22] are used. Also, nonrectangular and, thus, more specific ROIs based on facial landmarks[27] have been used. To evaluate our spatial ROI selection, we simulate a perfect functioning face/face part detection by manually selecting standard ROIs (see Fig. 1).

As can be seen from the results (see Tables 2 and 3), our automated spatial selection among small-sized ROI_n (S^{mc1} and S^{ms1}) does not require any face detection and outperforms the input SNR of selecting ROIs of the whole face or parts of it (forehead and cheeks) including consistent effect sizes g. Together with BSS, the output SNR (Fig. 5) shows inconsistent effects. Thereby, ANCOVA posthoc tests are affected by the highly differing input SNR while ANOVA posthoc tests show similar output SNRs between small-sized ROI_n and $\mathrm{ROI}_{\mathrm{F}}$, and

ROI_{FhC}. Given comparable output SNRs, an automated estimation of an ROI_{FhC} and ROI_F would require a higher effort for stable automated landmark detection in different subjects. Our automated spatial selection shows to compete successfully against even manually annotated ROIs achieving higher input SNRs and comparable output SNRs.

However, the comparison of deterministic and random selection of ROI_n clearly motivates that assembling homogeneous input characteristics (wavelength and dominant frequency) for BSS processing is not sufficient to yield the best BSS outcome. Both, the deterministic as well as the random selection of ROIs make use of periodicity in terms of frequently occurring dominant frequencies. However, the random selection does not force the highest possible SNR of a periodic component, like it is conducted in the deterministic selection. Accordingly, this comparison serves as a test, if always the best available signal quality needs to be selected in context of BSS. According to that, input and output SNR of the sets MCR and MSR of both BSS algorithms show absolute values below 0 dB (see Tables 2, 3, and Fig. 3). In addition, the BSS performance as a function of the input SNR (Fig. 4) shows an inferior performance compared to deterministic selection of equally sized ROIs in terms of a broadened area of negative performance (performance below x-axis), which widely lasts into the range of negative input SNRs. This finding underlines the necessity to process the best available input. Consequently, BSS is not necessarily able to compensate for lower input SNR under comparable conditions (ROI size and dominant frequency).

4.2 *Blind Source Separation Performance*

Figures 6 and 7 show examples of selecting $cbPPG_{n,color}$ according to the proposed BSS input selection and the further processing of these signals with PCA and ICA for the monochrome (MC1) and multispectral (MS1) case, respectively. In both cases, PCA and ICA are able to extract a distinct pulsatile component in the time domain. Focusing on the markedness of the spectral peak related to the heart rate in the spectra $X(f)$, it is worth noting that in the monochrome set (Fig. 6) PCA and ICA are performing similar. On the contrary, in the multispectral set (Fig. 7), PCA shows a decrease in the spectral power of the cardiac pulse compared to ICA. However, both examples and both BSS algorithms show at least one output component of proper quality regarding common postprocessing tasks (e.g., heart rate estimation). Accordingly, even a decrease in signal quality by application of BSS not necessarily renders postprocessing impossible. Nevertheless, morphology retention through BSS should be considered carefully.[40]

Several researchers have reported that BSS not necessarily improves the cbPPG quality and outcome. ICA was found to, if any, only subtly decrease the heart rate error for a small-sized cheek ROI[8] and even (slightly) increase for rectangular face ROI[16,17] compared to the BSS inputs. Moreover, PCA was found to perform worse on multispectral inputs compared to FastICA.[16] Since movements affect the input signal quality, it is worth relating these results with the movement conditions during recording. Christinaki et al.[8] allowed for small movements (facial expression) while extracting cheek ROIs, whereas

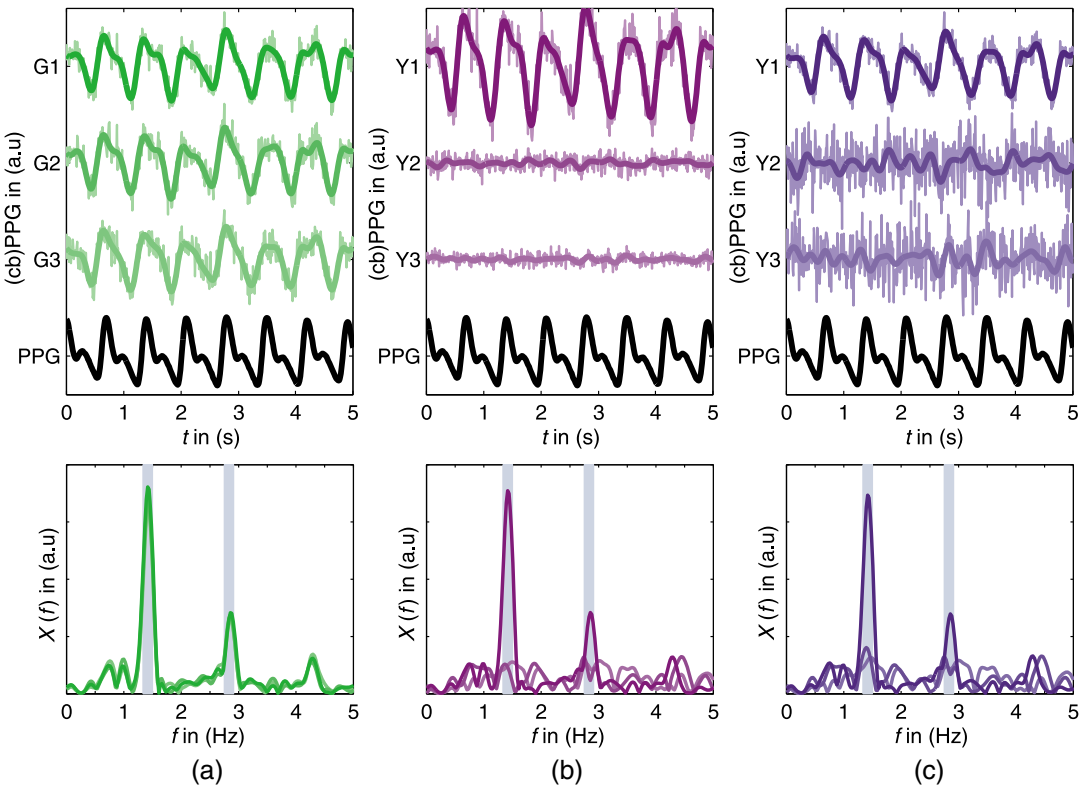

Fig. 6 Exemplary BSS performance on monochrome input. Sample signal excerpts [+4 Hz low-pass filtered (bold signal) versions] and amplitude spectra according to an automatically selected set S^{mc1}. (a) BSS input, (b) PCA output, and (c) ICA output. The true heart rate from the reference and its harmonic (±5 bpm) are indicated by the colored areas in the back of the spectra. Colors of the respective spectra are according to the time signals.

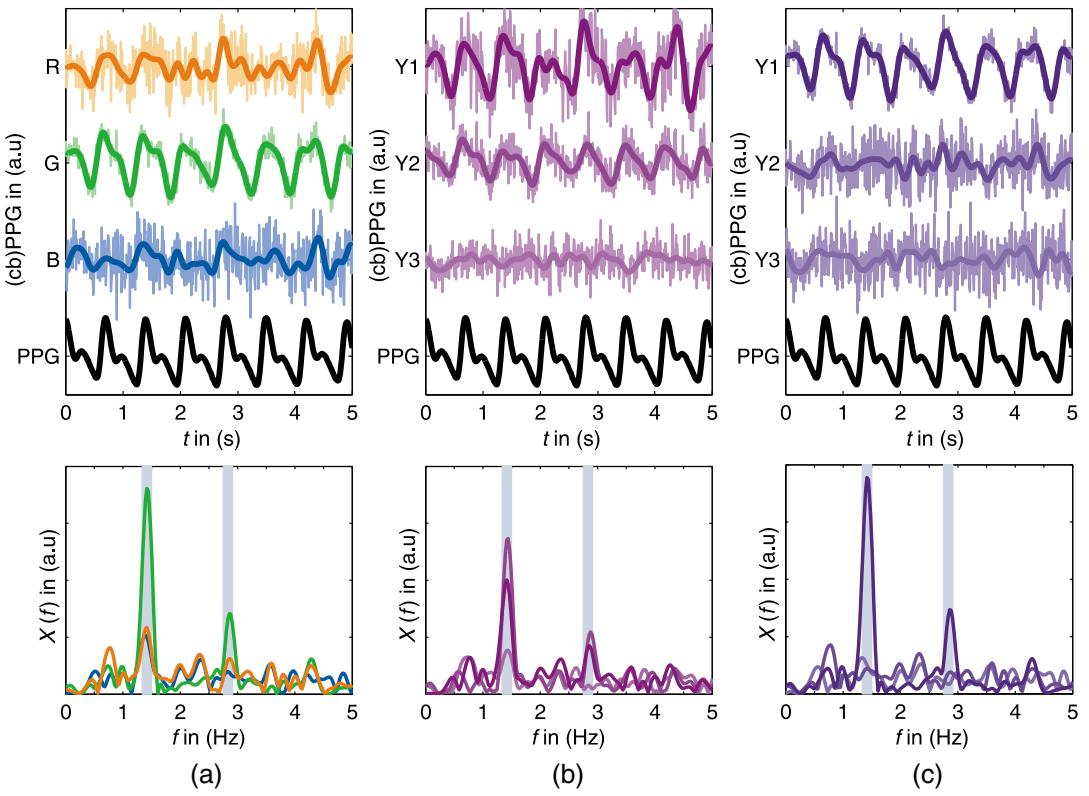

Fig. 7 Exemplary BSS performance on multispectral input. Sample signal excerpts [+4 Hz low-pass filtered (bold signal) versions] and amplitude spectra according to an automatically selected set S^{ms1}. (a) BSS input, (b) PCA output, and (c) ICA output. The true heart rate from the reference and its harmonic (±5 bpm) are indicated by the colored areas in the back of the spectra. Colors of the respective spectra are according to the time signals.

face ROIs were extracted from subjects who were asked not to move.[16,17] On the other hand, improvements of Bland–Altman heart rate measures were shown for facial ROIs,[6] while these improvements turned out to be higher for movement phases compared to no movement. Accordingly, the beneficial usage of BSS for cbPPG is rather found in conditions of less signal quality.

Such findings are in accordance with the results obtained by our investigations. Considering Fig. 4, which indicates an inverse relationship between input SNR and BSS performance (negative ΔSNR), the usage of BSS can even decrease the signal quality; its application, thus, should be considered with care. Particularly for small-sized ROIs after deterministic spatial selection (MC1, MC2, MS1, and MS2), an SNR decrease mainly appears in case of high-quality inputs (to the right of the *y*-axis) and is differently pronounced for different sets. On the contrary, the standard approaches using the face ROI and the forehead-cheeks ROI (Fig. 4) are not showing this marked negative BSS performance, whereas the increase of SNR for good quality inputs is also limited.

Accordingly, the performance of the compared methods is limited for high-quality inputs. This behavior might be attributed to the used contrast. Focusing for instance on MC1 that consists of inputs, which are as homogeneous as possible, the decorrelative transformation conducted by PCA is mostly able to preserve the SNR and shows the lowest number of segments with negative ΔSNR for high (positive) input SNRs (see Fig. 4). In comparison, the additional rotation introduced by ICA decreases the SNR for this set. However, the exclusive usage of PCA for high-quality inputs is not sufficient if the input is not as homogeneous as assembled by MC1 as the results of MC2, MS1, and MS2 show. So far, we used a standard tanh-contrast for FastICA as well as symmetric optimization for uniformity of the amount of output signals. One should further test optimized contrasts for example rather abstaining from super-Gaussian source optimization as indicated by the cbPPG component selection of Morris et al.[33] or alternatively let the demixing be guided by an expected cardiac pulse composition.[15] Furthermore, deflationary ICA could be applied for avoiding a model order violation. Another possibility to avoid undesired SNR decrease, if no SNR preserving contrast for high input quality is available, could be an adaptive decision, whether a BSS algorithm should be applied or not. This decision could be based on the prior SNR estimate based on peak frequency detection, which is done during the selection process of inputs.

4.3 *Homogeneity of the Blind Source Separation Input*

In this work, different factors of (in-)homogeneity of input signal sets to BSS are assessed in a controlled fashion.

First, the sensitivity of BSS algorithms regarding input of different dominant frequencies is analyzed. The results show that the ICA can take advantage of input signals comprising content with different dominant frequencies (MC2) while PCA shows a significantly worse performance (see Table 3). One might infer that the concept of statistical independence as applied for ICA is rather suited to such content compared to the

concept of decorrelation utilized for PCA. However, in case of a perfect homogeneous BSS input (MC1) comprising only one uniform wavelength and dominant frequency, PCA performs significantly better than ICA. As we used prewhitening prior to ICA that is similar to PCA, again one might deduce that the contrast applied for ICA, which is used to additionally transform the prewhithened data, is not well chosen to extract the cardiac pulse component.

Second, the question of sensitivity of BSS algorithms regarding wavelength homogeneity could clearly be answered for same ROI sizes and equal dominant frequencies (MC1 and MS2) in support of the monochrome approach. Both, PCA and ICA showed a significantly higher output SNR (see Table 3) by using the monochrome input. Also, consistent effects are found for both, ANCOVA and ANCOVA on the outputs (Fig. 5). Such findings support the idea that wavelength-dependent penetration depth into human skin imposes a nonlinear problem,[3] which BSS cannot consistently handle properly. However, our results also confirm the better suitability of ICA compared to PCA for multispectral face ROI input.[16]

Another result giving insight into BSS input homogeneity is stated by the comparison between the multispectral sets MS1, MS2, FhC, and F. While ROI size and input SNR show an inverse relationship, the output SNR of multispectral PCA and ICA shows proportionality to the ROI size except for the rectangular face ROI (see Tables 2 and 3). Up to $\mathrm{ROI}_{\mathrm{FhC}}$, one might assume a homogeneous ROI augmentation since mostly homogeneous skin regions without marked edges and regions, which do not necessarily contribute to a distinct cardiac pulse, are consolidated with $\mathrm{ROI}_{\mathrm{FhC}}$. The same skin regions are principally addressed by the sets MS1 and MS2. The only exception gives $\mathrm{ROI}_{\mathrm{F}}$ where also less-suited regions, such as mouth and nose, are included in the ROI, thus, serving an inhomogeneous ROI augmentation. Consequently, homogeneous ROI augmentations seem to be beneficial for multispectral BSS, whereas inhomogeneous areas inside ROIs should be omitted to optimize the extraction of the cardiac pulse. This behavior could also be found regarding heart rate error measures after FastICA on 15-s cbPPGs of different ROIs.[29] Despite that investigation neglected the input quality, the FastICA output showed decreasing heart rate error measures while step-wise excluding the face surrounding and face borders from the ROI. On the other hand, heart rate errors of outputs increased again while assessing ROIs with highly edged face regions mostly containing nose and mouth structures. Nevertheless, the positive effect of homogeneity of the input seems to be limited especially considering monochrome inputs of high signal quality, which links to the previous discussion on appropriate BSS contrasts.

A future application of the result that only PCA is able to preserve the SNR in very homogeneous (wavelength and dominant frequency) high-quality inputs of small-sized ROIs could be the evaluation of ΔSNR of spatially distributed ROIs to address the spatial homogeneity of the cutaneous microcirculation. Two-dimensional statements of the microcirculation may provide clinical significance in critical care patients.[30]

5 Conclusion

In conclusion, we investigated the performance of BSS to enhance the cardiac pulse from the cbPPG in dependency to varying input data characteristics. To that end, we developed an automated spatial selection of small-sized ROIs to locate the cardiac pulse in video recordings and control the input characteristics. The cbPPGs obtained from this spatial ROI selection significantly increased the SNR of the cardiac pulse compared to standard approaches, which select the whole face or anatomically defined subregions as ROIs. Subsequent BSS application did not show an unambiguous effect, rather input characteristics and particular BSS techniques had to be considered for future BSS usage on cbPPGs. While PCA developed the better performance compared to ICA using a very homogeneous input of same (green) wavelength and dominant frequency, ICA showed a significantly better performance compared to PCA on inhomogeneous inputs. Both, PCA and ICA performed better on monochrome inputs compared to multispectral inputs for the same ROI size. Algorithms, such as the proposed automated spatial ROI selection, can help to ensure passing appropriate inputs to respective BSS routines. Our results indicate, that, regardless of a subsequent BSS application, the usage of signal characteristics, such as simple frequency-domain features, can help identifying beneficial ROI locations to obtain superior cbPPGs compared to using classical ROI definitions based on face features. Furthermore, it turned out that BSS application might suffer from an inverse relationship between input signal quality and BSS performance, which even can cause a decrease of signal quality compared to the input data. Future research should address adaptive cbPPG processing schemes, which invoke BSS optionally or search for BSS contrasts capable of separating cbPPGs from noise independently from the input quality. Notwithstanding this, unsupervised signal processing techniques, such as BSS, should be carefully characterized in context with the measurement techniques and respective signal characteristics to which they are applied. The introduction of g in the context of quality assessment and our results allow comparisons of other algorithms and datasets in the future.

Appendix: Algorithm for Selection of BSS Input

Algorithm Input

All available $\mathrm{cbPPG}_{\mathrm{n,color}}$ signals with $n = 1,2,\ldots,475$ and color $\in \{R, G, B\}$ serve as input to the input set selection algorithm. The selection is mainly based on evaluating peak frequencies $\hat{f}_{\mathrm{G}}$ and its grouping clusters $\tilde{f}_{i,\mathrm{G}}$ considered as the dominant frequency, respectively, periodic component of the amplitude spectrum $X(f)$ of a $\mathrm{cbPPG}_{n,\mathrm{G}}$.

1. The maximum peak frequency $\hat{f}_{n,\mathrm{G}}$ of the amplitude spectra $X(f) = \mathcal{F}\{\mathrm{cbPPG}_{n,\mathrm{G}}\}$ between [30, 240] bpm is estimated for every ROI. The green wavelength is chosen for this selection according to the suitability for detecting the cardiac pulse inside this channel.[4]
2. The histogram $H(\hat{f}_{\mathrm{G}})$ of peak frequencies $\hat{f}_{\mathrm{G}}$ for all 475 ROIs is estimated (see Fig. 8 for an example). Clusters of $\mathrm{cbPPG}_{n,\mathrm{G}}$ are formed according to peak frequencies $\hat{f}_{n,\mathrm{G}}$ of maximum spread of 10 bpm ($\pm$5 bpm). The cluster width adapts to the signal quality measure used for evaluation. In case of multiple possibilities for forming $\mathrm{cbPPG}_{n,\mathrm{G}}$ clusters due to a continuous range of peak frequencies $\hat{f}_{n,\mathrm{G}}$ with spread >10 bpm bpm, always the cluster with maximum amount of $\mathrm{cbPPG}_{n,\mathrm{G}}$ is formed and the cluster limits to surrounding clusters are adjusted accordingly. The clustering according to peak frequencies in general

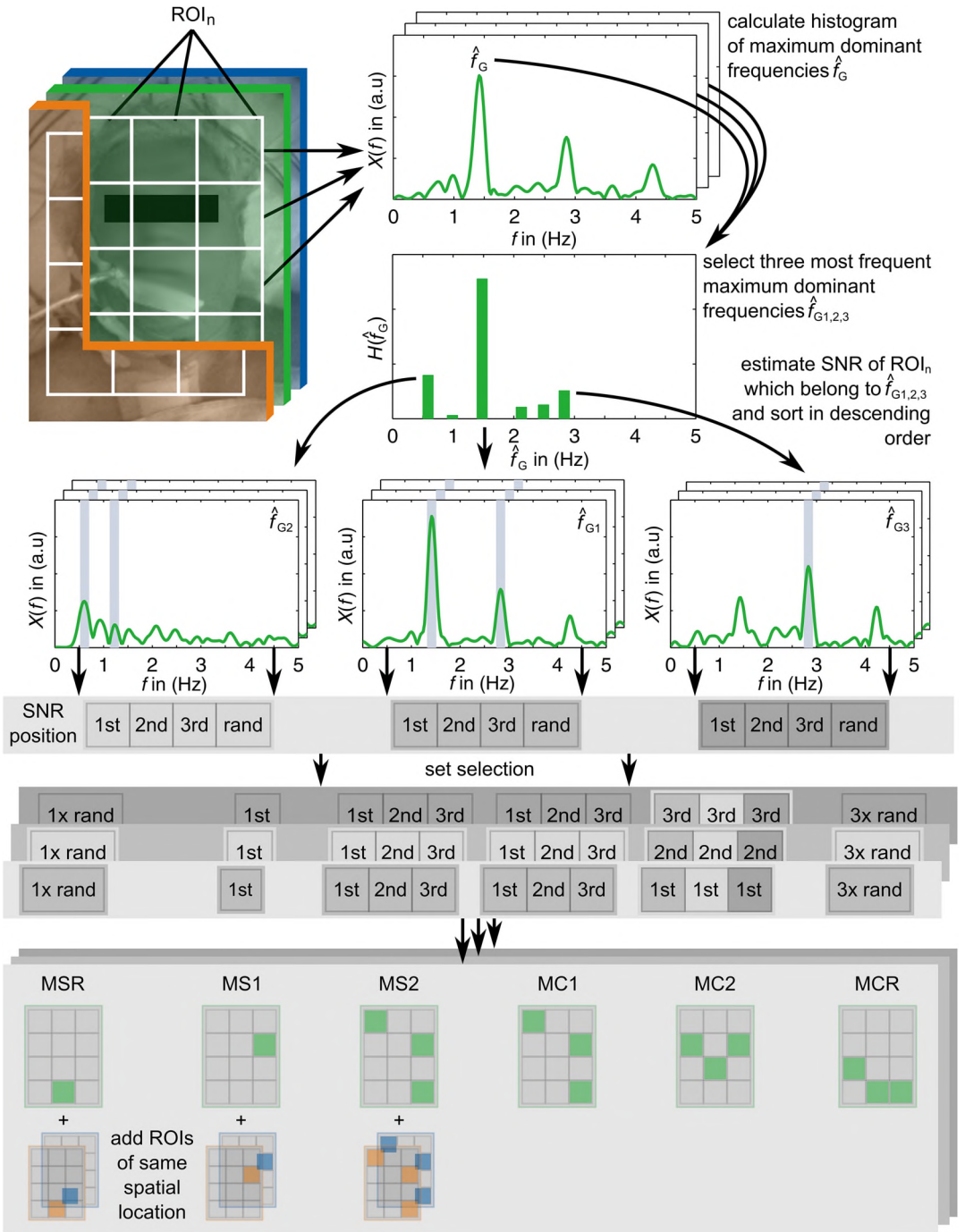

Fig. 8 Selection of ROIs for BSS input. ROI selection based on the three most frequent dominant frequencies of the green wavelength. ROI assembly (MC# and MS#) based on signal quality inside dominant frequencies.

addresses the search for highly periodic components as we expect the nature of the cardiac pulse.

3. The three largest clusters of $\text{cbPPG}_{n,\text{G}}$ of $H(\hat{f}_\text{G}) \rightarrow \tilde{f}_{i,\text{G}}$ with $i \in \{1,2,3\}$ according to its central peak frequency $\tilde{f}_{i,\text{G}}$ (see Fig. 8) are located.

4. To assess the strength of the periodic component defined by its peak frequency $\hat{f}_{n,\text{G}}$ inside a $\text{cbPPG}_{n,\text{G}} \in \tilde{f}_{i,\text{G}}$ with $i \in \{1,2,3\}$, the $\text{SNR}^{\tilde{f}_i}_{n,\text{G}}$ ($i \in \{1,2,3\}$) of the $\text{cbPPG}_{n,\text{G}}$ is calculated using Eqs. (5) and (6) while considering the peak frequency $\hat{f}_{n,\text{G}}$ as usable signal frequency estimate f_PPG. Accordingly, $\text{SNR}^{\tilde{f}_i}_{n,\text{G}} = \text{SNR}^{f_\text{PPG}}$ with $f_\text{PPG} = \hat{f}_{n,\text{G}}$ and $\hat{f}_{n,\text{G}} \in \tilde{f}_{i,\text{G}}$ ($i \in \{1,2,3\}$).

5. Input sets S are principally formed of ROIs of identical size so approaches could equally benefit from spatial averaging. Exceptions are formed by the standard multispectral approaches using a face or a forehead-cheek ROI, respectively. Moreover, one multispectral set is build with larger ROI area to adapt to the area intrinsically formed by the monochrome approach.

Monochrome Approach (Homogenous Frequency Content)

Assemble the green channel from appropriate ROIs inside a cluster $\tilde{f}_{i,\mathrm{G}}$ to three input sets according to ($i \in \{1,2,3\}$)

$$\left\{\begin{array}{ll} \mathrm{cbPPG}_{n,\mathrm{G}} \text{with} & \text{highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{\mathrm{n,G}} \stackrel{!}{=} \tilde{f}_i, \\ \mathrm{cbPPG}_{n,\mathrm{G}} \text{with} & \text{second highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i, \\ \mathrm{cbPPG}_{n,\mathrm{G}} \text{with} & \text{third highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i \end{array}\right\} \rightarrow S_{\mathrm{i}}^{\mathrm{mc1}}.$$

Monochrome Approach (Heterogeneous Frequency Content)

Assemble the green channel from appropriate ROIs inside an SNR positioning (of absolute SNRs) j for all three clusters $\tilde{f}_{i,\mathrm{G}}$ with $i \in \{1,2,3\}$ to three input sets according to ($j \in \{\text{first}, \text{second}, \text{third}\}$)

$$\left\{\begin{array}{ll} \mathrm{cbPPG}_{n,\mathrm{G}} \text{ with} & \text{j highest SNR}_{n,\mathrm{G}}^{\tilde{f}_1} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_1, \\ \mathrm{cbPPG}_{n,\mathrm{G}} \text{ with} & \text{j highest SNR}_{n,\mathrm{G}}^{\tilde{f}_2} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_2, \\ \mathrm{cbPPG}_{n,\mathrm{G}} \text{ with} & \text{j highest SNR}_{n,\mathrm{G}}^{\tilde{f}_3} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_3 \end{array}\right\} \rightarrow S_{\mathrm{j}}^{\mathrm{mc2}}.$$

Multispectral Approach (Homogenous Frequency Content)

Assemble the color channels from appropriate ROIs inside a cluster $\tilde{f}_{i,\mathrm{G}}$ to three input sets according to ($i \in \{1,2,3\}$)

$$\left\{\begin{array}{ll} \mathrm{cbPPG}_{n,\mathrm{R}} \text{ with} & \text{highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i, \\ \mathrm{cbPPG}_{n,\mathrm{G}} \text{ with} & \text{highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i, \\ \mathrm{cbPPG}_{n,\mathrm{B}} \text{ with} & \text{highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{\mathrm{n,G}} \stackrel{!}{=} \tilde{f}_i \end{array}\right\} \rightarrow S_{\mathrm{i}}^{\mathrm{ms1}}.$$

Multispectral Approach (ROI Area Adaption to Monochrome Approach)

Assemble the color channels from appropriate ROIs inside a cluster $\tilde{f}_{i,\mathrm{G}}$ containing the area (ROI_n) of the three best SNR ROIs obtained by frame-wise averaging the respective ROIs located as in S_i^{mc1} for three wavelengths each to three input sets according to ($i \in \{1,2,3\}$)

$$\left\{\begin{array}{ll} \mathrm{cbPPG}_{\mathrm{n,R}} \text{ with} & \text{highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i, \\ & \text{second highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i, \\ & \text{third highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i \\ \mathrm{cbPPG}_{\mathrm{n,G}} \text{with} & \text{highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i, \\ & \text{second highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i, \\ & \text{third highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i \\ \mathrm{cbPPG}_{\mathrm{n,B}} \text{ with} & \text{highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i, \\ & \text{second highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i, \\ & \text{third highest SNR}_{n,\mathrm{G}}^{\tilde{f}_i} \text{ subject to } \hat{f}_{n,\mathrm{G}} \stackrel{!}{=} \tilde{f}_i \end{array}\right\} \rightarrow S_{\mathrm{i}}^{\mathrm{ms2}}.$$

Random Approaches

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

Standard Approaches

For testing against standard multispectral BSS processing for cbPPG, form sets S^{F} and S^{FhC} from $\mathrm{cbPPG}_{\mathrm{F,color}}$ and $\mathrm{cbPPG}_{\mathrm{FhC,color}}$, respectively.

Algorithm Output

Input sets $S_{\mathrm{i}}^{\mathrm{mc1}}$, $S_{\mathrm{j}}^{\mathrm{mc2}}$, $S_{\mathrm{i}}^{\mathrm{ms1}}$, $S_{\mathrm{i}}^{\mathrm{ms2}}$, $S_{\mathrm{i}}^{\mathrm{mcR}}$, and $S_{\mathrm{i}}^{\mathrm{msR}}$ with $i \in \{1,2,3\}$ and $j \in \{\text{first}, \text{second}, \text{third}\}$ (i.e., three input sets, containing three channels each for the multispectral and monochrome approach) as well as one set S^{F} and S^{FhC}, respectively.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

Acknowledgments

The authors thank the Saxon State Ministry of Science and Culture (SMWK) for funding the project CardioVisio—Contactless acquisition of vital parameters. All authors report grants from Saxon State Ministry of Science and Culture (SMWK) during the conduct of the study. D. Wedekind and A. Trumpp report grants from Steinbeis Innovation Center Applied Medical Technology outside the submitted work.

References

1. Y. L. Zheng et al., "Unobtrusive sensing and wearable devices for health informatics," *IEEE Trans. Biomed. Eng.* **61**(5), 1538–1554 (2014).
2. C. Brueser et al., "Ambient and unobtrusive cardiorespiratory monitoring techniques," *IEEE Rev. Biomed. Eng.* **8**, 30–43 (2015).
3. M. Huelsbusch and V. Blazek, "Contactless mapping of rhythmical phenomena in tissue perfusion using PPGI," *Proc. SPIE* **4683**, 110–117 (2002).
4. W. Verkruysse, L. Svaasand, and J. Stuart Nelson, "Remote plethysmographic imaging using ambient light," *Opt. Express* **16**(26), 21434–21445 (2008).
5. G. Balakrishnan, F. Durand, and J. Guttag, "Detecting pulse from head motions in video," in *Proc. of the 26th IEEE Computer Vision and Pattern Recognition Conf. (CVPR)*, Portland, Oregon, pp. 3430–3437 (2013).
6. M. Poh, D. McDuff, and R. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Express* **18**(10), 10762–10774 (2010).
7. H. Wu et al., "Eulerian video magnification for revealing subtle changes in the world," *ACM Trans. Graphics* **31**(4), 1–8 (2012).
8. E. Christinaki et al., "Comparison of blind source separation algorithms for optical heart rate monitoring," in *Proc. of the 4th Mobihealth*, Athens, Greece, pp. 339–342 (2014).
9. Y. Hsu, Y. Lin, and W. Hsu, "Learning-based heart rate detection from remote photoplethysmography features," in *Proc. of the 39th Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Florence, Italy, pp. 4433–4437 (2014).
10. L. Feng et al., "Motion artifacts suppression for remote imaging photoplethysmography," in *Proc. of the 19th Int. Conf. on Digital Signal Processing (DSP)*, Hong Kong, China, pp. 18–23 (2014).
11. P. Comon, "Independent component analysis, a new concept?" *Signal Process.* **36**(3), 287–314 (1994).
12. J. Cardoso, "High-order contrasts for independent component analysis," *Neural Comput.* **11**(1), 157–192 (1999).
13. A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural. Networks* **10**(3), 626–634 (1999).
14. G. de Haan and V. Jeanne, "Robust pulse-rate from chrominance-based rPPG," *IEEE Trans. Biomed. Eng.* **60**(10), 2878–2886 (2013).
15. G. de Haan and A. van Leest, "Improved motion robustness of remote-PPG by using the blood volume pulse signature," *Physiol. Meas.* **35**(9), 1913–1926 (2014).
16. B. Holton et al., "Signal recovery in imaging photoplethysmography," *Physiol. Meas.* **34**(11), 1499–1511 (2013).
17. 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 *Proc. of the 34th Annual Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*, San Diego, California, pp. 2174–2177 (2012).
18. M. Poh, D. McDuff, and R. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Trans. Biomed. Eng.* **58**(1), 7–11 (2011).
19. M. Lewandowska et al., "Measuring pulse rate with a webcam: a non-contact method for evaluating cardiac activity," in *Proc. of the Federate Conf. on Computer Science and Information Systems (FedCSIS)*, Szczecin, Poland, pp. 405–410 (2011).
20. F. Zhao et al., "Remote measurements of heart and respiration rates for telemedicine," *PLoS One* **8**(10), e71384 (2013).
21. G. Tsouri et al., "Constrained independent component analysis approach to nonobtrusive pulse rate measurements," *J. Biomed. Opt.* **17**(7), 077011 (2012).
22. D. McDuff, S. Gontarek, and R. Picard, "Remote detection of photoplethysmographic systolic and diastolic peaks using a digital camera," *IEEE Trans. Biomed. Eng.* **61**(12), 2948–2954 (2014).
23. W. Wang, S. Stuijk, and G. de Haan, "Exploiting spatial-redundancy of image sensor for motion robust rPPG," *IEEE Trans. Biomed. Eng.* **62**(2), 415–425 (2015).
24. C. Lueangwattana, T. Kondo, and H. Haneishi, "A comparative study of video signals for non- contact heart rate measurement," in *Proc. of the 12th Int. Conf. on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, Hua Hin, China, pp. 1–5 (2015).
25. L. Yang et al., "Motion-compensated non-contact detection of heart rate," *Opt. Commun.* **357**, 161–168 (2015).
26. Y. Sun et al., "Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise," *J. Biomed. Opt.* **16**(7), 077010 (2011).
27. H. Tasli, A. Gudi, and M. Uyl, "Remote PPG based vital sign measurement using adaptive facial regions," in *Proc. of the 16th Int. Conf. on Image Processing (ICIP)*, Paris, France, pp. 1–5 (2014).
28. A. R. Guazzi et al., "Non-contact measurement of oxygen saturation with an RGB camera," *Biomed. Opt. Express.* **6**(9), 3320–3338 (2015).
29. K. Mannapperuma et al., "Performance limits of ICA-based heart rate identification techniques in imaging photoplethysmography," *Physiol. Meas.* **36**(1), 67–83 (2015).
30. S. Rasche et al., "Camera-based photoplethysmography in critical care patients," *Clin. Hemorheol. Microcirc.* **64**(1), 77–90 (2016).
31. L. Tarassenko et al., "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models," *Physiol. Meas.* **35**(5), 807–831 (2014).
32. ANSI/AAMI, "Cardiac monitors, heart rate meters, and alarms," EC13:2002, American National Standard (2002).
33. D. Morris et al., "Determining pulse transit time non-invasively using handheld devices," Patent application WO2014137768A1 (2014).
34. D. Wedekind et al., "Automated identification of cardiac signals after blind source separation for camera-based photoplethysmography," in *Proc. of the 35th Int. Conf. on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine, pp. 422–427 (2015).
35. B. Lantz, "The large sample size fallacy," *Scand. J. Caring Sci.* **27**(2), 487–492 (2013).
36. C. O. Fritz, P. E. Morris, and J. J. Richler, "Effect size estimates: current use, calculations, and interpretation," *J. Exp. Psychol. Gen.* **141**(1), 2–18 (2012).
37. H. Hentschke and M. Stüttgen, "Computation of measures of effect size for neuroscience data sets," *Eur. J. Neurosci.* **34**(12), 1887–1894 (2011).
38. S. Nakagawa and I. C. Cuthill, "Effect size, confidence interval and statistical significance: a practical guide for biologists," *Biol. Rev.* **82**(4), 591–605 (2007).
39. E. C. Hedberg and S. Ayers, "The power of a paired t-test with a covariate," *Social Sci. Res.* **50**, 277–291 (2015).
40. F. Andreotti et al., "An open-source framework for stress-testing non-invasive foetal ECG extraction algorithms," *Physiol. Meas.* **37**(5), 627–648 (2016).

Daniel Wedekind is a PhD student at the Institute of Biomedical Engineering, TU Dresden. He received his diploma degree in electrical engineering from the TU Dresden in 2013. His current research interests include practically relevant questions of the application of blind source separation for contactless vital signs monitoring.

Alexander Trumpp is a PhD student at the Institute of Biomedical Engineering, TU Dresden. He received his diploma degree in mechatronics engineering from the TU Dresden in 2013. His current studies focus on camera-based vital sign monitoring, both from a research and application point of view.

Sebastian Zaunseder received his PhD in Electrical Engineering from TU Dresden, Germany in 2011. Subsequently, he joined the Institute of Biomedical Engineering of TU Dresden, where he is currently head of the group Biosignals. His research interests include contact-free measurement systems, processing of biomedical signals and images to acquire robust information on vital signs, investigations on the cardio-respiratory autonomic modulation and research related to sleep.

Biographies for the other authors are not available.