PAPER

Improved motion robustness of remote-PPG by using the blood volume pulse signature

IOP Publishing | Institute of Physics and Engineering in Medicine Physiological Measurement

Physiol. Meas. **35** (2014) 1913–1926 doi:10.1088/0967-3334/35/9/1913

Improved motion robustness of remote-PPG by using the blood volume pulse signature

G de Haan[1] **and A van Leest**[2]

[1] Philips Group Innovation, Research, High Tech Campus 36, 5656AE Eindhoven, The Netherlands
[2] ASML, Flight Forum 1900, 5657EZ Eindhoven, The Netherlands

E-mail: G.de.Haan@Philips.com and Arno.van.Leest@ASML.com

Received 23 January 2014, revised 2 June 2014
Accepted for publication 6 June 2014
Published 27 August 2014

Abstract

Remote photoplethysmography (rPPG) enables contact-free monitoring of the blood volume pulse using a color camera. Essentially, it detects the minute optical absorption changes caused by blood volume variations in the skin. In this paper, we show that the different absorption spectra of arterial blood and bloodless skin cause the variations to occur along a very specific vector in a normalized RGB-space. The exact vector can be determined for a given light spectrum and for given transfer characteristics of the optical filters in the camera. We show that this 'signature' can be used to design an rPPG algorithm with a much better motion robustness than the recent methods based on blind source separation, and even better than the chrominance-based methods we published earlier. Using six videos recorded in a gym, with four subjects exercising on a range of fitness devices, we confirm the superior motion robustness of our newly proposed rPPG methods. A simple peak detector in the frequency domain returns the correct pulse-rate for 68% of total measurements compared to 60% for the best previous method, while the SNR of the pulse-signal improves from −5 dB to −4 dB. For a large population of 117 *stationary* subjects we prove that the accuracy is comparable to the best previous method, although the SNR of the pulse-signal drops from +8.4 dB to +7.6 dB. We expect the improved motion robustness to significantly widen the application scope of the rPPG-technique.

Keywords: biomedical monitoring, photoplethysmography, remote sensing, image analysis

(Some figures may appear in colour only in the online journal.)

1. Introduction

Photoplethysmography (PPG) is a simple and low-cost optical technique that can be used to detect blood volume changes in the micro-vascular bed of tissue. Typically, an optical sensor placed at the skin non-invasively measures the amount of light reflected from, or transmitted through, the skin. The sensor outputs the PPG waveform, which comprises a pulsatile physiological waveform attributed to cardiac synchronous changes in the blood volume with each heartbeat, and is superimposed on a slowly varying baseline with various lower frequency components attributed to respiration, sympathetic nervous system activity and thermoregulation (Allen 2007). The technique was first described in the 1930s (Hertzman 1937) and is now ubiquitously applied in hospitals, taking the form of a contact device clipped to a finger/toe/ear, or patched to the skin.

Although this technique is non-invasive, it requires contact to the skin. For completely unobtrusive measurement, it has been attempted to measure the (cardiac) pulse-signal at a distance, using radar (Greneker 1997), a thermal camera (Garbey *et al* 2007), or a RGB-video camera (Hülsbusch and Blazek 2002; Takano and Ohta 2007; Verkruysse *et al* 2008; Hülsbusch 2008; Poh *et al* 2010; Lewandowska *et al* 2011; Sun *et al* 2011; Balakrishnan *et al* 2013). Such *remote PPG* (rPPG) measurement is interesting in cases where contact has to be prevented because of extreme sensitivity (e.g. neonates, subjects with skin damage), or when unobtrusiveness is essential/desirable (e.g. surveillance, fitness).

The main concern with rPPG is robustness to subject motion. This is where the use of an RGB-video camera is advantageous, as it simultaneously provides multiple (color) channels with different PPG and noise mixtures. Blind Source Separation (BSS) techniques, like the Principal Component Analysis (PCA) and the Independent Component Analysis (ICA), have been used for decades to retrieve a clean signal from multi-channel mixtures with physiological measurement applications, e.g. in multi-channel electroencephalogram and electrocardiogram (Glass *et al* 2004). Indeed, motion robust methods to find the pulse-signal, i.e. the pulsatile component of the PPG, from video using BSS have been described (Hülsbusch 2008; Poh *et al* 2010; Lewandowska *et al* 2011). These BSS-techniques return independent signals (different definitions of 'independent'), which are linear combinations of the normalized mean skin-color channels, but no identification of which of these is the pulse-signal. To select the proper independent component all proposals used the periodic nature of the pulse, implicitly assuming the distortions (the other independent signals) to be non-periodic.

A different, chrominance-based, approach was used by de Haan and Jeanne (2013), who assumed non-local intensity-variation and specular reflection to cause the main distortions. Furthermore, they assumed a constant standardized skin-color, which enabled them to construct a linear combination of the normalized mean skin-color signals orthogonal to the assumed distortions, regardless the color of the illumination. This eliminated the periodicity-based component-selection of BSS-based methods, which performed poorly on fitness videos with their strong periodic subject motion.

In the current paper we aim at further improved motion robustness of the rPPG-technique without making assumptions about distortions or periodicity of signals. To this end, in section 2, we derive, from physiology and optics, the unique 'signature' of the blood volume pulse available in the mean pixel-values of a skin-region. We further propose a new motion-robust method designed to suppress the signals that deviate from this 'signature', computing another linear combination of mean RGB skin-pixel values. In a further attempt to relax the influence of possibly inaccurate algorithmic assumptions on the quality of the pulse-signal, we also propose methods that use the chrominance-based or the signature-based solutions just to guide the component selection in BSS-based methods. In section 3, we evaluate the motion

robustness using six videos that registered subjects exercising on different fitness devices. Also we assess the accuracy and SNR of the methods on a data-set used and described in de Haan and Jeanne (2013) and created by Thomas (2011), comprising a population of 117 stationary subjects in a broad range of skin-pigmentation levels. In section 4 we discuss the results from these experiments and, finally, we draw our main conclusions in section 5.

2. Analysis and methods

Before analyzing the pulse-signal in RGB-space, we briefly summarize existing methods.

The pulse-signal $\overrightarrow{S}$ obtained with a BSS-based method can be written as a linear combination of the individual mean-centered normalized color channels $\mathbf{C}_\mathrm{n}$:

$$\overrightarrow{S} = \overrightarrow{W}\mathbf{C}_\mathrm{n}, \tag{1}$$

where the linear weighting $\overrightarrow{W}$, with $\overrightarrow{W}\mathbf{W}^T = 1$, is obtained using BSS, and where each of the three rows of the $3 \times N$ matrix $\mathbf{C}_\mathrm{n}$ contains N samples of the mean-centered normalized color signals $\overrightarrow{R}_\mathrm{n}$, $\overrightarrow{G}_\mathrm{n}$ and $\overrightarrow{B}_\mathrm{n}$, respectively, i.e.:

$$\overrightarrow{R}_\mathrm{n} = \frac{1}{\mu(\overrightarrow{R})}\overrightarrow{R} - 1, \quad \overrightarrow{G}_\mathrm{n} = \frac{1}{\mu(\overrightarrow{G})}\overrightarrow{G} - 1, \quad \overrightarrow{B}_\mathrm{n} = \frac{1}{\mu(\overrightarrow{B})}\overrightarrow{B} - 1, \tag{2}$$

where the operator μ corresponds to the (temporal) mean (over N frames), and the vectors $\overrightarrow{R}$, $\overrightarrow{G}$, and $\overrightarrow{B}$ contain the (spatial) mean of the red, green and blue pixel values of the skin-region (typically, and also in this paper, the face) of every frame in the analysis window with length N. We shall use 'ICA' and 'PCA' to refer to the BSS-based methods used by Poh *et al* (2010) and Lewandowska *et al* (2011), respectively.

Also the Chrominance-based method (further referred to as 'CHRO') of de Haan and Jeanne (2013) can be rewritten as a linear combination of $\boldsymbol{C}_\mathrm{n}$. They used a standardized skin-color (a good fit regardless of the skin-pigmentation level), selected as $[Rs, Gs, Bs] = [\,0.7682, 0.5121, 0.3841]$. This choice allowed them, without assuming a white-balanced camera, to derive the—specular reflection free—color difference signals $\overrightarrow{X}_\mathrm{s}$ and $\overrightarrow{Y}_\mathrm{s}$ from the normalized color channels using $\overrightarrow{X}_\mathrm{s} = 0.77\overrightarrow{R}_\mathrm{n} - 0.51\overrightarrow{G}_\mathrm{n}$ and $\overrightarrow{Y}_\mathrm{s} = 0.77\overrightarrow{R}_\mathrm{n} + 0.51\overrightarrow{G}_\mathrm{n} - 0.77\overrightarrow{B}_\mathrm{n}$, and to compute the pulse-signal as:

$$\overrightarrow{S} = \overrightarrow{X}_\mathrm{s} - \alpha\overrightarrow{Y}_\mathrm{s}, \text{ with } \alpha = \frac{\sigma\left(\overrightarrow{X}_\mathrm{s}\right)}{\sigma(\overrightarrow{Y}_\mathrm{s})}, \tag{3}$$

where the operator σ corresponds to the standard deviation. This expression can be re-written as[3]:

$$\overrightarrow{S} = \overrightarrow{W}_\mathrm{CHRO}\mathbf{C}_\mathrm{n}, \tag{4}$$

with

$$\overrightarrow{W}_\mathrm{CHRO} = \frac{1}{\sqrt{6\alpha^2 - 20\alpha + 20}}[2 - \alpha,\ 2\alpha - 4,\ \alpha]\,. \tag{5}$$

[3] For uniformity with other algorithms we scaled $\overrightarrow{S}$ such that the weight vector $\overrightarrow{W}_\mathrm{CHRO}$ is normalized. Also for simplicity, we use $\overrightarrow{S}$ to symbolize the output pulse-signal of all methods introduced in this or earlier papers, although clearly these pulse-signals are not identical.

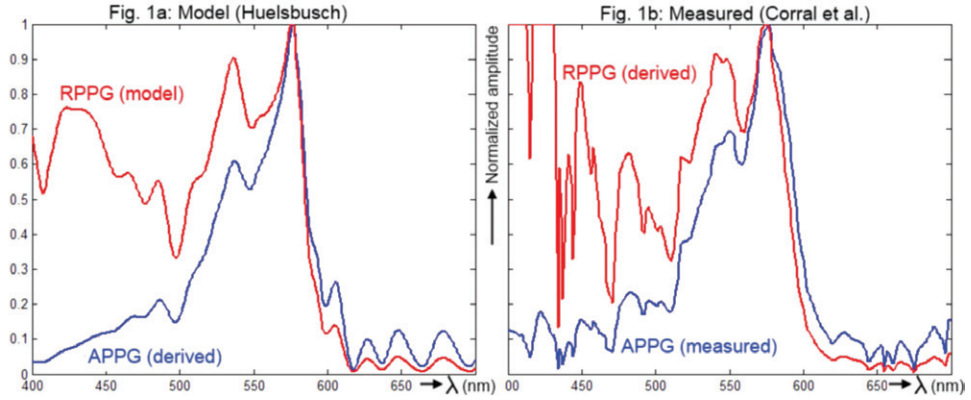

Figure 1. (*a*) Predicted spectrum from a model by Hülsbusch of the *relative* PPG-amplitude (Hülsbusch 2008) and (*b*) measurement by Corral *et al* (2011) of the *absolute* PPG-amplitude under halogen illumination. The figures also show a derived absolute/relative PPG amplitude. All spectra have been scaled to 1 for their peak at 577 nm.

The best results are obtained if band-passed filtered versions of $\overrightarrow{R_n}$, $\overrightarrow{G_n}$ and $\overrightarrow{B_n}$ are used. To this end, (in the Fourier domain) all frequencies outside the pulse-rate band, i.e. 0.5–3.5 Hz, are set to zero. The computation of α and the Fourier transform for band-pass filtering are all performed on the analysis window with length N. The optimal length N shall be discussed in the results section, but $N > 32$ to allow for the band-pass filtering.

The new approach that we propose in this paper results from an analysis of the pulse-signal in the normalized color-space. We shall first present our analysis and then introduce the methods.

2.1. Analysis of the pulse-signal in RGB-space

The beating of the heart causes pressure variations in the arteries as the heart pumps blood against the resistance of the vascular bed. Since the arteries are elastic, their diameter changes in sync with the pressure variations. These diameter changes occur even in the smaller vessels of the skin, where the blood volume variations cause absorption variations of the light that can be measured with a PPG-sensor.

Hülsbusch explained that the resulting *relative* PPG-amplitude, as a function of the wavelength λ, of the signal from this sensor, $RPPG(\lambda) = \sigma(PPG(\lambda))/\mu(PPG(\lambda))$, is determined by the contrast between the blood and the blood-free tissue (Hülsbusch 2008). Given the absorption spectrum of the oxygenated arterial blood and the absorption spectrum of the dermis and epidermis, and assuming a concentration of 3% melanin, his simulations led to the relative PPG-amplitude shown in figure 1(*a*). In this spectrum, the absorption peaks of oxygenated blood at $\lambda = 542$ nm and $\lambda = 577$ nm can be clearly recognized.

Corral *et al* (2011) *measured* $\sigma(PPG(\lambda))$ using a spectrometer analyzing the light reflected from a skin-region illuminated by a tungsten-halogen incandescent light source. Their measurement did not include the normalization, i.e. the division of the pulsatile fraction by the baseline portion. We shall refer to it as the *absolute* PPG-amplitude to clearly distinguish it from the predicted *relative* PPG-amplitude of Hülsbusch. Corral's spectrum shows a strong amplitude peak around $\lambda = 550$ nm, as shown in figure 1(*b*). Since the baseline is determined by the light-source and the skin, this absolute PPG can be related to the relative PPG

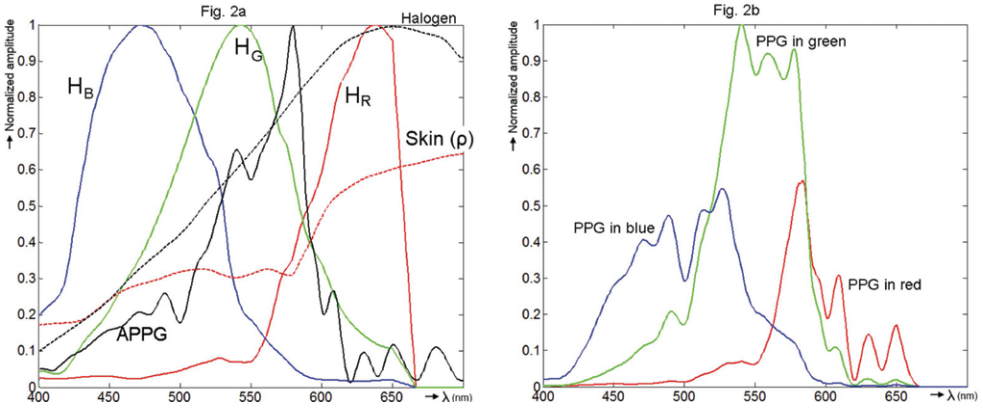

Figure 2. (*a*) Spectral responses of Hülsbusch' derived absolute PPG-amplitude, the emission spectrum of a halogen lamp, an example skin-reflection, and the optical RGB-filters of a UI-2220SE-C camera. (*b*) The PPG contributions to the RGB color channels predicted from the responses in figure 2(*a*), which after integration over the spectrum give the components of the blood volume vector $\overrightarrow{P}_{\mathrm{bv}}$.

via the continuous emission spectrum of the tungsten-halogen illumination, $I_h(\lambda)$, and the skin-reflection, $\rho_s(\lambda)$:

$$PPG(\lambda) = \rho_s(\lambda)\ I_h(\lambda)\ RPPG(\lambda)\,. \tag{6}$$

In the figure, we included a *derived* relative PPG using the emission spectrum of a tungsten-halogen incandescent lamp. Apart from the noise for shorter wavelengths, we find this to correspond fairly well with the prediction from the model of Hülsbusch. The increased noise-level for shorter wavelengths is plausible given the weaker emission of the halogen lamp and the higher absorption of the skin for shorter wavelengths.

We used the noise-free spectrum from the model of Hülsbusch to predict the *normalized blood-volume pulse vector* $\overrightarrow{P}_{\mathrm{bv}}$, which we define as the relative PPG-amplitude in the normalized RGB-color channels of the video camera registering a stationary skin-region:

$$\overrightarrow{P}_{\mathrm{bv}} = \frac{[\sigma(\overrightarrow{R}_{\mathrm{n}}),\ \sigma(\overrightarrow{G}_{\mathrm{n}}),\ \sigma(\overrightarrow{B}_{\mathrm{n}})]}{\sqrt{\sigma^2(\overrightarrow{R}_{\mathrm{n}}) + \sigma^2(\overrightarrow{G}_{\mathrm{n}}) + \sigma^2(\overrightarrow{B}_{\mathrm{n}})}}. \tag{7}$$

The camera samples the spectrum rather coarsely with the RGB color channels modelling the primary colors of the eye. The RGB-channels are typically centered around 450 nm, 550 nm and 650 nm, and have a relative wide bandwidth of some 100 nm and are consequently partially overlapping. Given $PPG\ (\lambda)$ derived from the model of Hülsbusch, we can predict the blood volume pulse vector $\overrightarrow{P}_{\mathrm{bv}}$ for different skin-reflection spectra, $\rho_s(\lambda)$, and spectral compositions, $I(\lambda)$, of the light source.

To quantify our expectations, we measured the responses, $H_{\mathrm{R}}(\lambda)$, $H_{\mathrm{G}}(\lambda)$ and $H_{\mathrm{B}}(\lambda)$ of the RGB color-channels of a global-shutter CCD camera (model USB UI-2220SE-C of IDS GmbH, white-balanced on tungsten-halogen incandescent illumination). We also measured the skin reflectance $\rho_s(\lambda)$ of an example subject, and used Hülsbusch's derived *absolute* PPG-amplitude, $PPG\ (\lambda)$, according to equation (6) (For simplicity we used the same notation for Corral's and Hülsbusch's $PPG\ (\lambda)$ and $RPPG\ (\lambda)$, although their spectra are not identical). From these spectra, shown in figure 2(*a*), the blood volume pulse vector $\overrightarrow{P}_{\mathrm{bv}}$ can be predicted:

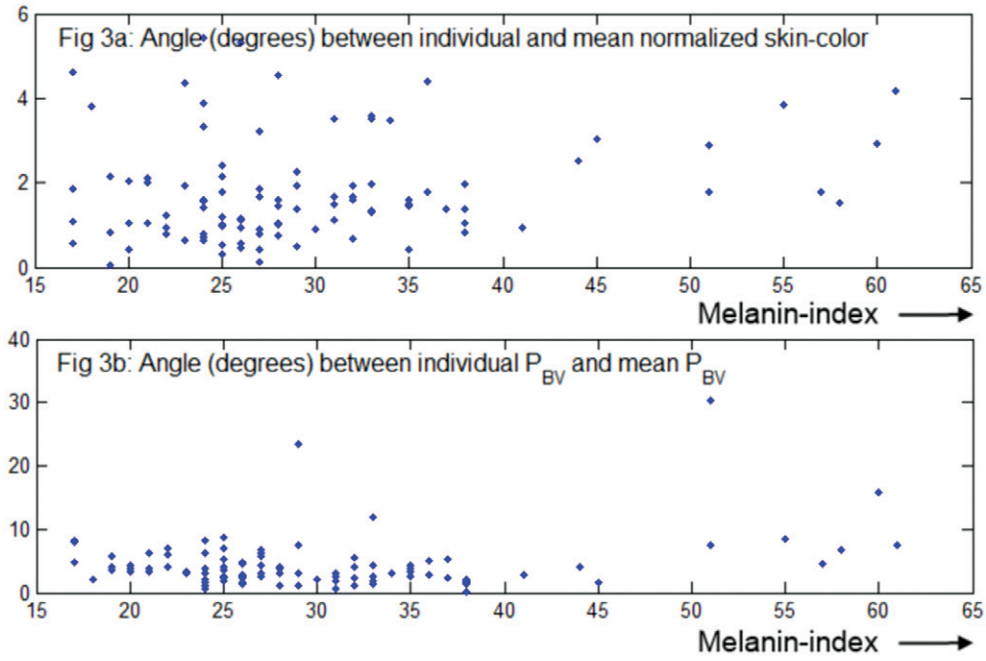

Figure 3. (*a*) The RMS-value of the angle between the (105) individual and the mean normalized skin-color vector is only 2°, i.e. the vector is quite stable over the skin-pigmentation range (melanin index 17–61). (*b*) The signal representing angle between individual $\overrightarrow{P}_{\mathrm{bv}}$ -vectors and the mean $\overrightarrow{P}_{\mathrm{bv}}$ is more noisy. The angle between the individual and the mean $\overrightarrow{P}_{\mathrm{bv}}$ has an RMS-value of 6°, but also does not show a clear influence of the melanin-index. We consider the higher noise level is caused by short videos, low blue-levels and subject movements.

$$\overrightarrow{P}_{\mathrm{bv}} = \left[\frac{\int_{400}^{700} H_{\mathrm{R}}(\lambda)\frac{I(\lambda)}{I_h(\lambda)}PPG(\lambda)\,\mathrm{d}\lambda}{\int_{400}^{700} H_{\mathrm{R}}(\lambda)\frac{I(\lambda)}{I_h(\lambda)}\rho_s(\lambda)\,\mathrm{d}\lambda}, \frac{\int_{400}^{700} H_{\mathrm{G}}(\lambda)\frac{I(\lambda)}{I_h(\lambda)}PPG(\lambda)\,\mathrm{d}\lambda}{\int_{400}^{700} H_{\mathrm{G}}(\lambda)\frac{I(\lambda)}{I_h(\lambda)}\rho_s(\lambda)\,\mathrm{d}\lambda}, \frac{\int_{400}^{700} H_{\mathrm{B}}(\lambda)\frac{I(\lambda)}{I_h(\lambda)}PPG(\lambda)\,\mathrm{d}\lambda}{\int_{400}^{700} H_{\mathrm{B}}(\lambda)\frac{I(\lambda)}{I_h(\lambda)}\rho_s(\lambda)\,\mathrm{d}\lambda} \right]. \tag{8}$$

We compared the predicted $\overrightarrow{P}_{\mathrm{bv}}$, i.e. the normalized result of equation (8), with the measured $\overrightarrow{P}_{\mathrm{bv}}$ according to equation (7), using the relative pulsatilities in the mean pixel value over a skin area of a stationary subject. The measured and the predicted blood volume pulse vector using Hülsbusch's spectrum were only 4° apart, while the model prediction using Corral's more noisy spectrum was 7° off.

We note that the pulse-signal observed in all camera color channels largely originates from wavelengths between 500 nm and 600 nm. Figure 2(*b*) shows how this interval dominates each of the camera color channel responses. As a result the blue and red channel of the camera clearly show a PPG signal, though smaller than the green channel.

Our model also shows that the blood volume pulse vector depends on the optical filters of the camera, the illumination spectrum and the skin-reflectance. Interestingly, the relative reflectance of the skin, in the RGB-channels under white illumination, is remarkably stable through the complete range of skin-pigmentation, as shown in figure 3(*a*). This is likely because the absorption spectra of the blood-free skin is dominated by the melanin absorption (Hülsbusch 2008). Although a higher melanin concentration can increase the absolute absorption considerably, the relative absorption in the different wavelengths remains substantially

the same. This implies an increase of melanin darkens the skin but hardly changes the normalized color of the skin. Consequently, also the blood volume pulse, $\overrightarrow{P_{\rm bv}}$, is quite stable under white illumination. This is confirmed by figure 3(*b*) for the same data set of 105 subjects with different skin pigmentation. This data set is a large subset of the data set built by Thomas (2011). We removed 12 subjects from the original set as they moved too much for measuring $\overrightarrow{P_{\rm bv}}$. The melanin index of the subjects was measured at the forehead using a skin-pigmentation analyzer, SPA 99, from Courage and Khazaka.

2.2. New methods

The stable $\overrightarrow{P_{\rm bv}}$ can be used as the 'signature' of blood volume change to distinguish their color variations from those due to alternative causes. So, we propose to use the known relative pulsatilities of the color channels $\overrightarrow{P_{\rm bv}}$, to discriminate between the pulse-signal and distortions. This requires no assumption that the pulse is the only periodic component in the video, as necessary for independent component selection in BSS-based methods. Also assumptions about the distortion signals, like their relative strength in the normalized color channels as in the chrominance-based method, are not necessary.

We do assume, as before, that the pulse-signal can be built as a linear combination of normalized color signals as in equation (1). Since we know that the relative amplitudes of the pulse-signal in the RGB color-channels are given by $\overrightarrow{P_{\rm bv}}$, we are searching for the weights, $\overrightarrow{W_{\rm PBV}}$, that give a pulse-signal $\overrightarrow{S}$, for which the correlation with the color channels $\overrightarrow{R_{\rm n}}$, $\overrightarrow{G_{\rm n}}$, and $\overrightarrow{B_{\rm n}}$ equals $\overrightarrow{P_{\rm bv}}$:

$$\overrightarrow{S}\,\mathbf{C}_{\rm n}^{T} = k\overrightarrow{P_{\rm bv}} \Leftrightarrow \overrightarrow{W_{\rm PBV}}\mathbf{C}_{\rm n}\mathbf{C}_{\rm n}^{T} = k\overrightarrow{P_{\rm bv}}, \tag{9}$$

and consequently $\overrightarrow{W_{\rm PBV}}$ can be calculated:

$$\overrightarrow{W_{\rm PBV}} = k\overrightarrow{P_{\rm bv}}\mathbf{Q}^{-1} \text{ with } \mathbf{Q} = \mathbf{C}_{\rm n}\mathbf{C}_{\rm n}^{T}, \tag{10}$$

with scalar k chosen such that $\overrightarrow{W_{\rm PBV}}$ has unit length. We conclude that the characteristic wavelength dependency of the PPG-signal, as reflected in the blood volume pulse, $\overrightarrow{P_{\rm bv}}$, can be used to estimate the pulse-signal from the time-sequential RGB pixel data averaged over the skin area. We shall refer to this new algorithm as the PBV-method.

In hindsight, a strong point of the BSS-based methods is that they need no assumptions regarding main distortions, a standardized skin, or blood volume pulse vectors. Their weakness is they have to assume the pulse is the only periodical signal to identify the proper independent component. In a further attempt to relax constraints, we propose to use the weighting $\overrightarrow{W}$ of equation (3) or (10), obtained from the CHRO or PBV-method just to *guide* the component selection of a BSS-based method. A good criterion is to choose the component that has the 'most similar' weighting coefficients. Relevant algorithms result in weights $\overrightarrow{W}$ that more accurately minimize distortions, since the underlying assumptions — known distortions for CHRO, and stable $\overrightarrow{P_{\rm bv}}$ for PBV—serve mainly as guiding rules where small departures play no role.

To elaborate we assume ICA or PCA generates three component signals, $\overrightarrow{S_1}$, $\overrightarrow{S_2}$, $\overrightarrow{S_3}$, each obtained as a linear combination of the normalized RGB signals, $\mathbf{C}_{\rm n}$:

$$\overrightarrow{S_i} = \overrightarrow{W_i}\mathbf{C}_{\rm n}, \text{ with } \quad i = 1,\ 2,\ 3. \tag{11}$$

The pulse-signal, $\overrightarrow{S}$, is selected from the component signals using a reference weight vector, $\overrightarrow{W_{\rm REF}} = \overrightarrow{W_{\rm CHRO}}$, or $\overrightarrow{W_{\rm REF}} = \overrightarrow{W_{\rm PBV}}$, obtained from the CHRO- or PBV-algorithm, equal to $\overrightarrow{S} = \overrightarrow{S_p}$ with $p \in \{1, 2, 3\}$ such that

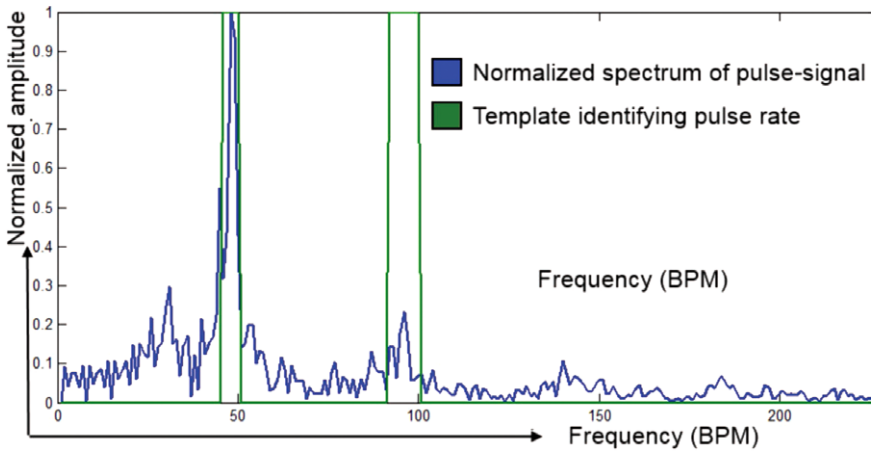

Figure 4. The SNR is calculated with a template indicating the pulse-rate and its second harmonic from the reference sensor. The SNR measures the energy ratio of the spectral segments inside and outside the template. The second harmonic was included as it seems unfair to count it as noise, even though it does not change the SNR very much. The third harmonic was small enough to be ignored.

$$\forall_{i\in\{1,2,3\}} \left| \overrightarrow{W_p} \overrightarrow{W}_{\mathrm{REF}}^{T} \right| \geq \left| \overrightarrow{W_i} \overrightarrow{W}_{\mathrm{REF}}^{T} \right|. \tag{12}$$

This gives us 4 additional algorithms: ICA or PCA guided by CHRO, or by PBV. All methods may use overlapped signal windowing as in de Haan and Jeanne (2013), and can successfully use short window lengths. Since BSS cannot guarantee the polarity of the pulse-signal, we have to keep the sign of the pulse-signal fragment of each window the same. To that end we modify $\overrightarrow{S_M}$, i.e. the pulse-signal fragment of window M:

$$\overrightarrow{S_M} = \mathrm{sign}\left(\overrightarrow{W}_{p,M} \overrightarrow{W}_{\mathrm{REF},M}^{T} \right) \overrightarrow{S}_{p,M}, \tag{13}$$

where p may vary from one window to the other. We note that the original BSS-based algorithms performed poorly with short window lengths as the independent component selection has to recognize the pulse-signal from its periodic nature, which suffers from the low frequency resolution that comes with a short window.

2.3. Evaluation metrics

For the evaluation, the pulse-signals of all methods and that of the reference (a finger-oximeter of Contec Medical Systems, model CMS50E, or Polar H3 heart-rate sensor for fitness) were pre-processed using a Butterworth (order 6) band-pass filter with cutoff frequencies 0.5–3.5 Hz. We define the first performance metric as the Percentage of total measurements 'PERC' that the pulse-rate from the reference sensor and the rate from the method under test was in good agreement. The rate was measured every frame using a peak-detector in the frequency domain using a sliding Fourier window of only 256 frames to allow for the varying pulse-rate typical for exercising subjects. A good agreement was considered as such where the pulse-rate differed no more than a single frequency bin (4.7 Hz) from the rate obtained from the reference sensor.

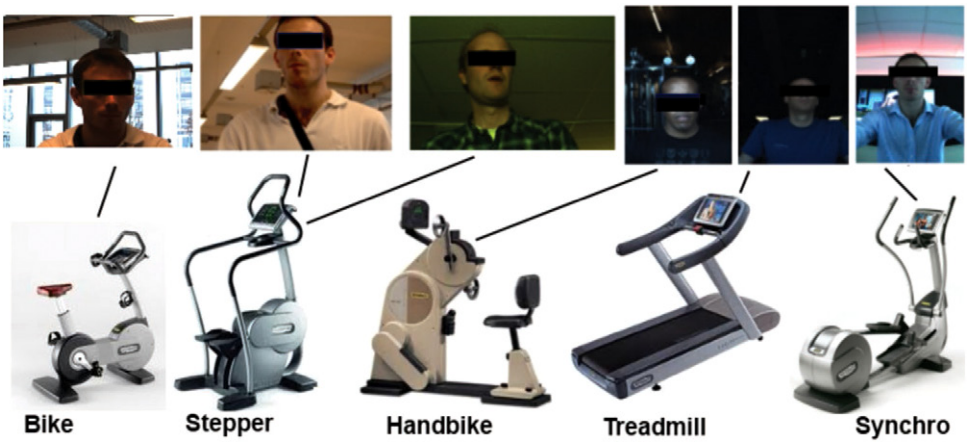

Figure 5. Fitness devices and corresponding camera view to test motion robustness.

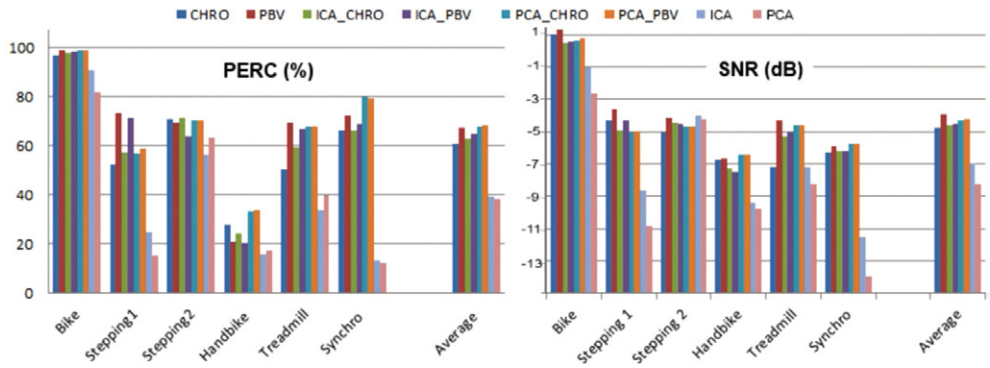

Figure 6. Our two performance metrics, PERC and SNR, resulting from the fitness recordings show that the outputs of all proposed rPPG methods do not differ very much in performance. The earlier BSS-based methods, ICA and PCA, are included for reference to show their motion robustness is very limited.

As our second metric, we compute the signal-to-noise-ratio (SNR). More specifically, we compute the ratio of the energy around the fundamental pulse-frequency (of the reference sensor) plus its second harmonic and the remaining energy contained in the spectrum, as illustrated in figure 4 and defined by:

$$\mathrm{SNR} = 10\,\log_{10}\left(\frac{\sum_{f=0.5}^{3.5}\left(\overrightarrow{U_t}(f)\,\overrightarrow{\hat{S}}(f)\right)^2}{\sum_{f=0.5}^{3.5}\left(\left(1-\overrightarrow{U_t}(f)\right)\overrightarrow{\hat{S}}(f)\right)^2}\right), \tag{14}$$

where $\overrightarrow{\hat{S}}(f)$ is the spectrum of the pulse-signal, $\overrightarrow{S}, f$ is the frequency (Hz), and $\overrightarrow{U_t}(f)$ is the binary template as shown in figure 4.

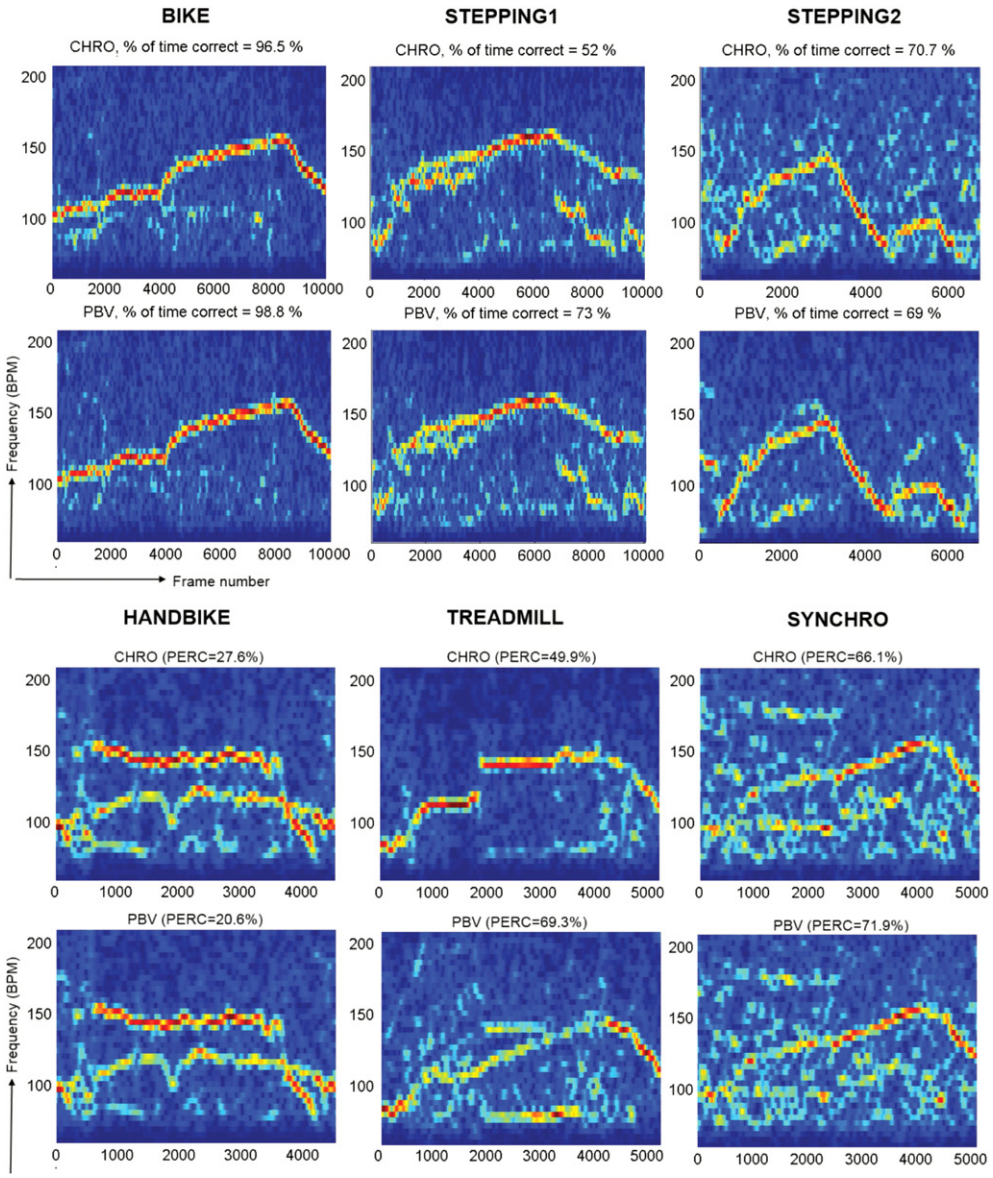

Figure 7. Spectrograms for the PBV and CHRO algorithms on the different fitness sequences. With the spectrograms the score on our PERC metric is shown. Peak detection is too simple for good results, yet it demonstrates that the pulse-signal is more often correctly found using PBV-method.

3. Results

We first tested the motion-robustness of our proposed methods with six videos recorded in a gym. We had four different subjects exercising on a bike, a stepping device, a handbike, a treadmill and a synchro-device, respectively, including moderate and strong motion.

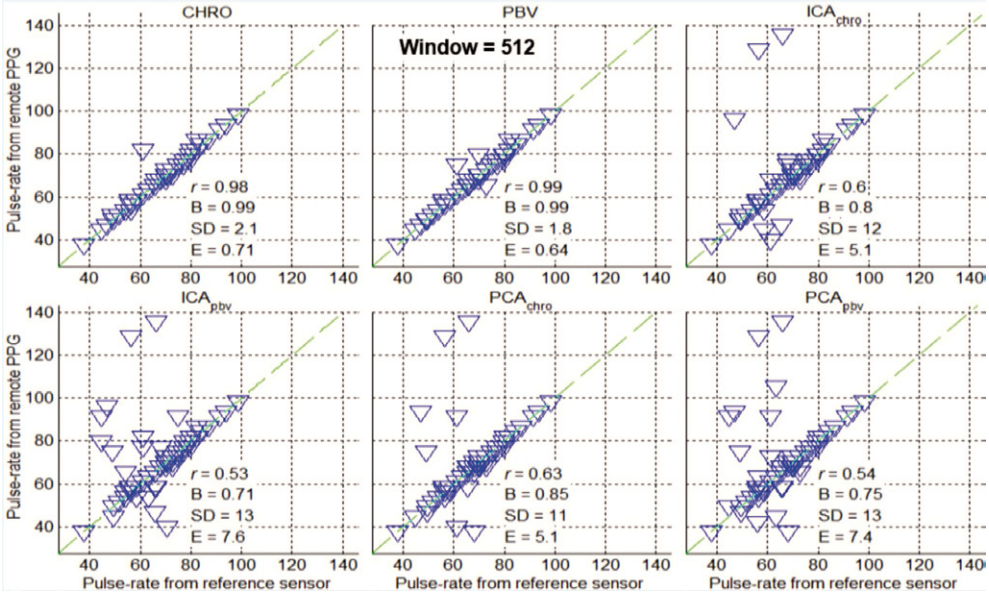

Figure 8. Correlation plots from the large scale experiment including the Pearson correlation, r, the slope of the linear fit, B, the standard deviation, SD, and the RMSE, E, for all evaluated methods for window length 512. The PBV and CHRO methods show the strongest correlation with the finger-oximeter.

Table 1. SNR (dB) with long (512) and short (64) overlap-add window.

	CHRO	PBV	ICA_{CHRO}	ICA_{PBV}	PCA_{CHRO}	PCA_{PBV}
512:	8.4	7.6	6.2	5.3	4.3	3.6
64:	8.2	5.6	3.1	1.0	1.8	0.1

Photographs of the devices and snapshots from the camera are shown in figure 5. The mean face-color was input to all methods as in de Haan and Jeanne (2013). Clearly, the illumination conditions varied greatly in brightness and in spectral composition, but methods were not modified for this. The subjects were healthy, male, age between 28 and 35, and had a light skin, except for the person on the handbike who had a dark skin. Informed consent was obtained from each subject and the study was approved by the Internal Committee Biomedical Experiments of Philips Research.

Figure 6 shows that for the two sequences with the lowest SNR, i.e. the handbike and the synchro, the PERC was highest with the two guided-PCA methods. For the three sequences with the higher SNR, i.e. bike, stepping and treadmill, the PBV-method scored the highest PERC. PBV and the two PCA-guided methods improved the average PERC to about 68% from 60% for CHROM (and less than 40% for the older BSS-based methods). PBV also improved the SNR on the exercising subjects by almost 1dB compared to CHROM and by more than 3 dB compared to the BSS-based methods.

To further illustrate the progress from this paper, we show the spectrograms along with the PERC-score of the PBV-method and the best competitor, i.e. the CHRO-method in figure 7. The spectrograms of the guided BSS-methods are very similar to that of PBV and have therefore been omitted.

To show the impact of the improved motion robustness on the signal quality for stationary subjects, we compared the pulse-rate from the various methods with the pulse-rate measured by the finger oximeter for a large population of 117 healthy individuals with various skin-pigmentation in a controlled environment with professional studio illumination. This is the same data-set used and described in de Haan and Jeanne (2013) and created by Thomas (2011). As shown by the correlation plots in figure 8, the pulse-rates obtained with the CHRO-method and that of our newly proposed PBV-method are closest to that of the finger-oximeter and perform comparably. Although the signal quality remains good, the SNR on these stationary subjects is 1 dB lower for a long overlap-add window and about 3.5 dB lower for the shorter window, as shown in table 1. The new proposed guided BSS-based methods show a greater signal-quality loss on stationary subjects.

4. Discussion

All our new methods show a very good improvement in motion robustness compared to earlier methods. Our comparison on exercising subjects shows that BSS-methods guided by CHRO are on average better than CHRO. Also a small improvement on average for PCA guided by PBV, compared to PBV, can be observed, but this difference is very small. ICA guided by PBV was found to be worse than PBV.

The PBV-method shows a very modest loss in signal quality for the stationary subjects where motion robustness cannot bring any advantage. This loss becomes significant for our guided-BSS methods, even though the guided-PCA proved best on our most challenging fitness-videos. We assume that with fewer distortions the reference weight vector, $\overrightarrow{W}_{\mathrm{REF}}$, may become poorly defined, which leads to this drop in signal quality for simple stationary cases. Similarly, as shown in table 1, for stationary subjects the SNR with a short overlap-add window—which allows algorithms to more quickly adapt to changing statistics—cannot be expected to improve the performance for stationary subjects. Nonetheless, the PBV-method still performs well with the short window. With significant motion, all methods performed best with a relatively short overlap-add window. The optimal length—used in the experiments—is around 64 frames (3 s) for all methods. (Note that the peak-detector still uses a 256 frame FFT.)

This modest SNR-loss of the PBV-method compared to CHRO is plausible, since for stationary subjects the color channels mainly contain a pulse-signal and some inevitable noise. Hence, the correctness of the blood volume pulse vector in equation (10) becomes important; a slight deviation may result in a vector $\overrightarrow{W}_{\mathrm{PBV}}$ that has a large angle with respect to the inaccurate blood volume pulse vector $\overrightarrow{P}_{\mathrm{bv}}$. As a direct result, the SNR decreases. However in case of strong distortions, these *distortions* mainly determine the direction of the vector $\overrightarrow{W}_{\mathrm{PBV}}$ and an inaccurate blood volume pulse vector cannot harm much. In the stationary case the CHRO algorithm profits from the pulse-signal having an opposite sign and similar amplitude in X_s and Y_s. Consequently, matching their standard deviation and subtraction, equation (3), will lead to a good SNR.

5. Conclusions

In this paper, we have derived the relative pulsatilities in the RGB-channels of a video camera due to blood volume variations in the skin. Our model shows the effect of the transfer of the

optical filters in the camera, the skin-reflection, and the emission spectrum of the light source. We further considered predictions of the *relative* PPG-amplitude by Hülsbusch (2008) and measurements of the *absolute* amplitude by Corral *et al* (2011), and considered them in good agreement with our results.

Measurements on 105 subjects show that the unit length vector describing the relative pulsatilities in the RGB-channels of a color camera is substantially independent of skin-pigmentation. We experimentally verified this blood volume pulse 'signature' and designed a new algorithm (the PBV-method) for robust rPPG pulse-detection that minimizes the contribution to the pulse-signal of color variations with other signatures.

To reduce the effect of small deviations from the predicted signature, we used the RGB-weighting resulting from our PBV-method to select between the independent components returned by methods based on blind source separation (BSS) leading to two additional algorithms, ICA and PCA *guided by PBV*. Similarly, we used an earlier chrominance-based method (CHRO) for BSS component selection to relax the influence of the presumed constant relative skin-reflection and known main distortions. This led to two more rPPG methods, ICA and PCA *guided by CHRO*.

We compared these methods with state-of-the-art CHRO (de Haan and Jeanne 2013). We tested the motion robustness on subjects exercising on five different fitness-devices. This experiment showed clearly improved motion-robustness of all new methods, with PBV and PCA guided by CHRO or PBV increasing the PERC-score to 68% from 60% for CHRO. In a second experiment designed to check the signal quality on stationary subjects, the PBV-method performed comparably to the state-of-the-art (CHRO) on a large data set with 117 subjects with different skin-pigmentation, but the guided BSS-methods sacrificed signal quality for the superior motion-robustness in case of stationary subjects.

We conclude that for stationary subjects the best signal quality is obtained with the earlier CHRO and new PBV method. The motion robustness of all proposed methods, using an overlap-add window of about 3 s, is clearly superior.

Acknowledgments

The authors are grateful for the help and contributions of their Philips colleagues Ingmar van Dijk, Vincent Jeanne, Ihor Kirenko, Herman Kloosterman and Wim Verkruijsse.

References

Allen J 2007 Photoplethysmography and its application in clinical physiological measurement *Physiol. Meas.* **28** R1–39

Hertzman A B 1937 Photoelectric plethysmography of the fingers and toes in man *Exp. Biol. Med.* **37** 529–34

Greneker E F 1997 Radar sensing of heartbeat, respiration at a distance with application of the technology *RADAR* **97** 150–4

Garbey M, Sun N, Merla A and Pavlidis I 2007 Contact-free measurement of cardiac pulse based on the analysis of thermal imagery *IEEE Trans. Biomed. Eng.* **54** 1418–26

Hülsbusch M and Blazek V 2002 Contactless mapping of rhythmical phenomena in tissue perfusion using PPGI *Proc. SPIE* **4683** 110–7

Takano C and Ohta Y 2007 Heart rate measurement based on a time-lapse image *Med. Eng. Phys.* **29** 853–7

Verkruysse W, Svaasand L O and Nelson J S 2008 Remote plethysmographic imaging using ambient light *Opt. Express* **16** 21434–45

Hülsbusch M 2008 An image-based functional method for opto-electronic detection of skin-perfusion *PhD Thesis* RWTH Aachen department of EE (in German) http://books.google.nl/books?id=JtLnPgAACAAJ

Poh M Z, McDuff D J and Picard R W 2010 Non-contact, automated cardiac pulse measurements using video imaging and blind source separation *Opt. Express* **18** 10762–74

Lewandowska M, Ruminski J, Kocejko T and Nowak J 2011 Measuring pulse rate with a webcam: a non-contact method for evaluating cardiac activity Proc. FedCSIS (Szczecin, Sept. 2011) **18–21** 405–10

Sun Y, Hu S, Azorin-Peris V, Greenwald S, Chambers J and Zhu Y 2011 Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise *J. Biomed. Opt.* **16** 077010

Balakrishnan G, Durand F and Guttag J 2013 Detecting pulse from head motions in video *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR) (Portland, Jun. 2013)* pp 3430–7

Glass K A, Frishkoff G A, Frank R M, Davey C, Dien J, Malony A D and Tucker D M 2004 A framework for evaluating ICA methods of artefact removal from multichannel EEG *Proc. 5th Int. Conf. ICA (Granada, Sept. 2004)*

de Haan G and Jeanne V 2013 Robust pulse-rate from chrominance-based rPPG *IEEE Trans. Biomed. Eng.* **60** 2878–86

Thomas M 2011 *Large Scale Study on Remote Photoplethysmography Graduation Report Resulting from Internship* Philips Research Eindhoven ISEN-Brest

Corral L F, Paez G and Strojnik M 2011 Optimal wavelength selection for non-contact reflection photoplethysmography *22nd Congress of the Int. Commission for Optics (Puebla, Aug. 2011) (Proc. SPIE vol 8011)* pp 8011–91, 15–9