Non-contact Measurement of Respiration Rate With Camera-based Photoplethysmography During Rest and Mental Stress

Hannes Ernst, Hagen Malberg, Martin Schmidt

Institute of Biomedical Engineering, TU Dresden, Germany

Abstract

Cameras offer new possibilities in assessing human states like stress. This work addresses the measurement of respiration rate with several approaches. Data from a mental stress study ($>\ 40\,\mathrm{h}$ *recording time,* 54 *participants) were evaluated and color channel combinations were examined using a hemispherical surface grid search with mean absolute error* MAE *to optimize the approaches. The grid search converged towards the green channel in the baseline modulation approach (*$MAE = 2.53\,\mathrm{rpm}$*). However, best results were achieved with the frequency modulation approach (*$MAE = 2.13\,\mathrm{rpm}$*) with the color channel combination optimal for heart rate measurement. Respiration rate increased highly significant during stress (*$p < 0.001$*, Mann-Whitney U-test). Deliberate selection of the color channel combination is crucial for respiration rate measurement with cameras in order to assess human states or pathophysiological processes.*

1. Introduction

Respiration rate is an important parameter to comprehend the human state, e.g. in heart rate variability analysis. Respiration modulates photoplethysmographic signals in three ways: baseline modulation due to the variation of pressure in the thorax, amplitude modulation due to the variation of ventricular filling, and frequency modulation due to the respiratory sinus arrhythmia [1].

Camera-based photoplethysmography (cbPPG), a non-contact monitoring technique, allows the parallel assessment of heart rate, heart rate variability, and respiration rate [2]. RGB cameras provide three channels: red, green, and blue (R, G, and B). Many techniques for the combination of these signals have been proposed to improve signal quality (e.g. blind source separation, CHROM, and POS) [2]. However, algorithmic differences such as windowing, norming, or parameter tuning limit the explanatory power of direct comparisons. Therefore, a data-driven grid search was proposed to analyze the combination of R, G, and B directly in a systematic manner [3]. So far, however, such analysis has been performed only for heart rate measurement [3]. This work addresses non-contact respiration rate measurement with cbPPG during rest and mental stress and, for the first time, systematically investigates the influence of the combination of color channels provided by RGB videos in this regard.

2. Methods and Materials

Data from 54 healthy participants (22 female, $25.8 \pm 4.9\,\mathrm{yr}$, $22.3 \pm 2.7\,\mathrm{kg/m^2}$) of a stress study (Mannheim Multicomponent Stress Test [4]) provided synchronized chest expansion signals ($25.6\,\mathrm{Hz}$, Equivital, ADInstruments Ltd., Dunedin, New Zealand), earlobe photoplethysmograms ($1000\,\mathrm{Hz}$, MLT1020EC, ADInstruments Ltd.), and RGB video recordings ($100.49\,\mathrm{fps}$, $320\,\mathrm{px} \times 640\,\mathrm{px}$, $12\,\mathrm{bit}$, UI-3060CP-C-HQ Rev.2, IDS Imaging Development Systems GmbH, Obersulm, Germany) with a total duration of more than $40\,\mathrm{h}$. Details about study protocol and device synchronization are given in [5, 6].

R, G, and B were extracted from the video recordings as the mean within the level set region of interest (see [7]). They were combined to a camera-based photoplethysmographic signal sig_{cbPPG} by the linear combination:

$$sig_{cbPPG}[n] = c_r \cdot R[n] + c_g \cdot G[n] + c_b \cdot B[n] \quad (1)$$

with combination coefficients c_r, c_g, and c_b. The combination coefficients were varied systematically to investigate the effect of the channel combination on respiration rate measurement. They covered the surface of the unit hemisphere ($c_g \leq 0$), which was discretized with 5185 equally distributed grid points by means of Icosahedron-based S^2 Decomposition [8] as illustrated by Figure 1.

The signals were split into $10\,\mathrm{s}$ segments with $3\,\mathrm{s}$ shift, which led to $N_{seg} = 45796$ segment pairs. Signal segments were band-pass filtered in the range 0.167–$0.5\,\mathrm{Hz}$ to gain respiratory signals sig_{resp} that matched the physiological range of breathing for healthy adults (10–$30\,\mathrm{rpm}$). Photoplethysmograms from earlobe and camera were resampled to $25.6\,\mathrm{Hz}$ to match the sampling rate of the chest expansion signal. Respiration rate RR in rpm of each

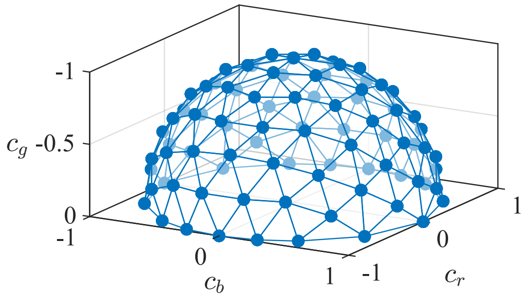

Figure 1: Structure of a hemispherical surface grid with 89 grid points. Each grid point represents a color channel combination, i.e. a set of $[c_r, c_g, c_b]$. The negative c_g axis was chosen to incorporate the signal inversion typical for photoplethysmography.

segment was gained from the frequency of the highest peak f_{peak} in the amplitude spectrum $|X[k]|$:

$$X[k] = \mathcal{F}(sig_{resp}[n]), \tag{2}$$

$$\begin{aligned} RR &= 60\,\mathrm{s/min} \cdot f_{peak} \\ \text{with} \quad f_{peak} &= \Delta f \cdot k_{max} \\ \text{and} \quad k_{max} &= \arg\max_k |X[k]|. \end{aligned} \tag{3}$$

To investigate the influence of measurement duration, this procedure was repeated for segment lengths of 30 s (9 s shift, $N_{seg} = 13675$), and 60 s (15 s shift, $N_{seg} = 8338$). In all cases, sufficient length of the fast Fourier transformation ensured a frequency resolution Δf lower than $0.0167\,\mathrm{Hz}$ (1 rpm).

Respiration rates from earlobe and camera were evaluated with reference to respiration rates from chest expansion RR_{ref} with the mean absolute error MAE:

$$MAE = \frac{1}{N_{seg}} \sum_{i=1}^{N_{seg}} |RR_{ref}(i) - RR(i)| \tag{4}$$

and the median $\widetilde{SNR}$ of the signal-to-noise ratio SNR:

$$\begin{aligned} SNR &= 10 \cdot \lg \frac{\sum_{\Delta f \cdot k = 10\,\mathrm{rpm}}^{30\,\mathrm{rpm}} \Pi(k) \cdot |X(k)|^2}{\sum_{\Delta f \cdot k = 10\,\mathrm{rpm}}^{30\,\mathrm{rpm}} (1 - \Pi(k)) \cdot |X(k)|^2} \\ \text{with} \quad \Pi(k) &= \begin{cases} 1 & \Delta f \cdot |k_{max,ref} - k| \leq 3\,\mathrm{rpm}, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{5}$$

In addition to the baseline modulation approach, camera-based respiration rates were extracted using the frequency modulation approach. Pulse onsets were detected for each segment with the MATLAB function `findpeaks`[1]. The pulse onsets provided a tachogram which was filtered[2] and equidistantly resampled (spline interpolation, 1000 Hz). Respiration rates were derived from the tachograms in the same manner as described by equations (2) and (3) and evaluated with MAE and $\widetilde{SNR}$.

Finally, to investigate the effect of stress on respiration rate, respiration rates under rest and during the stress test were compared using the Mann-Whitney U-test.

3. Results

Figure 2 shows the results of the hemispherical surface grid searches. The grids exhibited monotonic progression towards the optimum $\min(MAE)$, i.e. no isolated clouds of high or low MAE were found. Table 1 summarizes performance and location of the optimal color channel combinations $O3C_{resp}$. Normalization caused only marginal performance differences. MAE lowered for longer segments (10 s → 30 s: −20.2 %, 30 s → 60 s: −1.2 %) down to 2.53 rpm.

Table 1 additionally provides results of the earlobe measurements. Despite lower $\widetilde{SNR}$, camera-based measurements yielded a relative reduction of MAE by −26.2 % for 30 s segment length (10 s: −26.2 %, 60 s: −18.6 %) compared to earlobe measurements.

Table 2 contains the results of the frequency modulation approach. Tachogram filtering reduced the MAE (10 s: −2.3 %, 30 s: −6.2 %, 60 s: −7.4 %). In accordance with the baseline modulation approach, MAE lowered for longer segments (10 s → 30 s: −41.5 %, 30 s → 60 s: −6.6 %) down to 2.13 rpm.

Figures 3a and 3b show respiration rates by state (rest vs. stress) for reference and camera-based measurements, respectively. Highly significant differences ($p < 0.001$, Mann-Whitney U-test, Bonferroni correction) were found between respiration rates during rest and stress for reference as well as all settings of camera-based measurement. Respiration rate increased during mental stress (median during rest: 16.5 rpm, during stress: 19.5 rpm), which matched the expected physiological response.

4. Discussion

$O3C_{resp}$ is located close to the green channel and differs from the optimal color channel combination for heart rate measurement (see Figure 2). Thus, our results indicate that channel combinations such as POS or CHROM (both model-based approaches developed for heart rate measurement) may not be the optimal choice for respiration rate measurement with the baseline modulation approach. An interesting aspect of this finding is that it implies a limitation of the dichromatic reflection model on the background

[1] Minimum peak distance: 0.3 s. Minimum peak prominence: 0.5. sig_{cbPPG} was derived with the grid search approach optimized for maximum accuracy of heart rate measurement (see [3]) and was normed to the median of the moving standard deviation (window length: 2 s).

[2] A pulse was removed if its duration differed more than 20 % to the duration of the preceding one. Segments with more than 10 % of the pulses excluded were removed.

Table 1: Summary of results for respiration rate measurement with the earlobe photoplethysmogram (PPG) and the camera-based photoplethysmogram (cbPPG) from the grid search approaches (baseline modulation). If multiple color channel combinations yielded $\min(MAE)$ and identical $\widetilde{SNR}$, the combination closest to the centroid was selected as $O3C_{resp}$. T: Segment length.

	PPG		**Original cbPPG**					**Normalized cbPPG**				
T in s	MAE in rpm	$\widetilde{SNR}$ in dB	$\min(MAE)$ in rpm	$\widetilde{SNR}$ in dB	$O3C_{resp}$ c_r	c_g	c_b	$\min(MAE)$ in rpm	$\widetilde{SNR}$ in dB	$O3C_{resp}$ c_r	c_g	c_b
10	4.35	−3.98	3.21	−5.99	0.07	−1.00	0.00	3.21	−5.99	0.07	−1.00	0.00
30	3.47	−2.90	2.56	−3.49	0.08	−0.98	−0.20	2.56	−3.50	0.10	−0.99	−0.10
60	3.11	−2.03	2.53	−2.54	0.19	−0.97	−0.13	2.54	−2.75	0.23	−0.97	−0.10

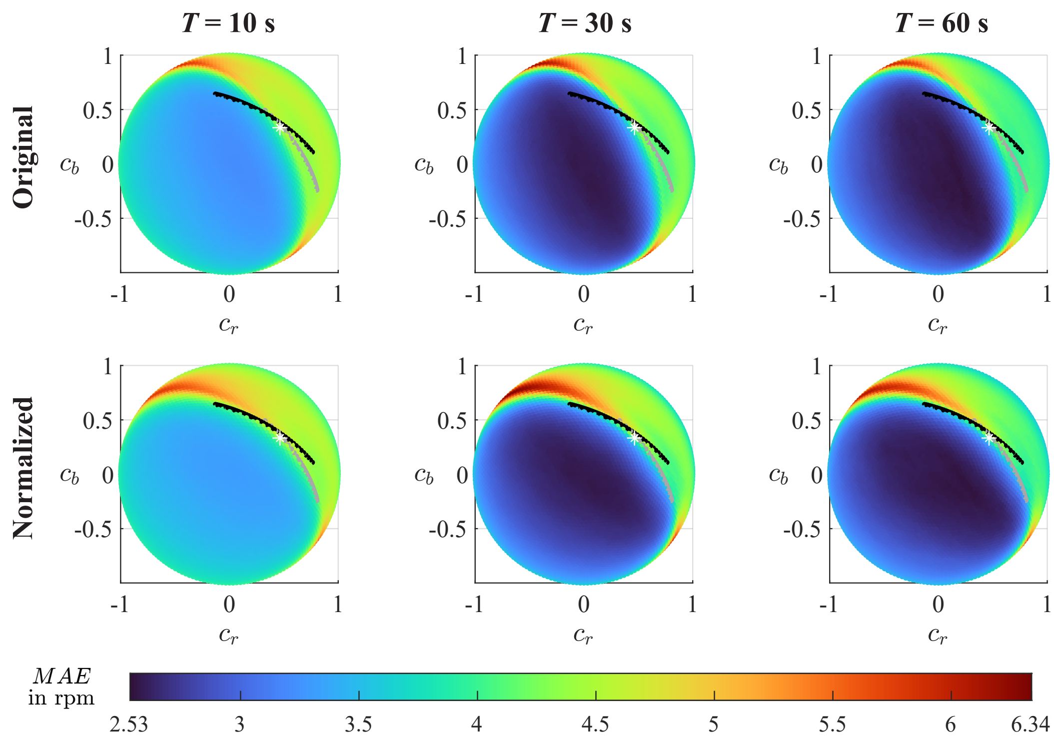

Figure 2: Mean absolute error MAE for all color channel combinations of the hemispherical surface grid searches (top view with c_g axis in the center). The grids show monotonic progression towards the optimum ($\min(MAE)$, see Table 1). For reference, typical color channel combinations to derive photoplethysmographic signals from videos are given in gray (POS) and black (CHROM). The white asterisk marks the optimal color channel combination for maximum accuracy of heart rate measurement used for the tachogram generation. T: Segment length.

of the blood volume effect, although the close proximity of $O3C_{resp}$ to the green channel points toward a causal effect of hemoglobin's dominating absorption in this spectral range. Nevertheless, the frequency modulation approach (utilizing the cardiac cbPPG signal) outperforms the baseline modulation approach (see Table 2). Here, color channel combinations close to POS and CHROM perform well.

Variations between $O3C_{resp}$ of different settings remain small (see Table 1), especially between the original and normalized signals. This might result from the orientation close to the green channel: As G dominates in (1), scaling becomes less important. However, differences can be seen in the orientation of the band of high MAE in Figure 2.

Considering the trade-off between temporal resolution and error, 30 s mark a sufficient measurement duration.

Figure 3d shows a representative course of camera-based respiration rate measurement which was effective to a large extent. However, even though most measurements were accurate, RR occasionally deviated, mostly to the lower bound. This was found to be the major source of errors as Figure 3b and Figure 3c illustrate. Dominant low-frequent interferences can be attributed to vasomotor

Table 2: Summary of the camera-based respiration rate measurement results from the pulse onset tachograms (frequency modulation). T: Segment length.

	Original		**Filtered**	
T in s	MAE in rpm	$\widetilde{SNR}$ in dB	MAE in rpm	$\widetilde{SNR}$ in dB
10	3.99	−0.11	3.90	−0.13
30	2.43	1.11	2.28	1.06
60	2.30	1.19	2.13	1.26

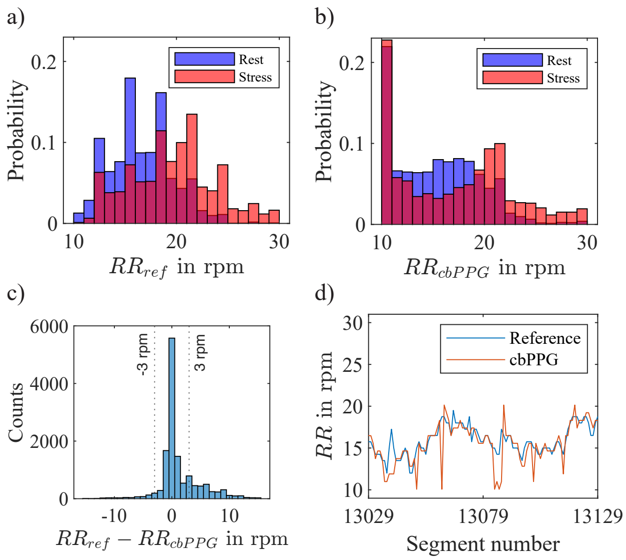

Figure 3: Excerpts from reference and cbPPG data (frequency modulation approach, 30 s segment length, tachograms filtered). a) RR by state from reference. b) RR by state from cbPPG. c) Errors between RR from reference and cbPPG. d) Exemplary measurement course.

activity. The probability of the lower bound drop was similar during rest and stress.

In the literature, similar performance for camera-based respiration rate measurement using the baseline modulation approach is reported: $MAE = 2.2$ rpm ([9], setting: face) and $MAE = 4.67$ rpm ([1], setting: spontaneous breathing with CHROM). However, based on our results, we recommend using the frequency modulation approach.

5. Conclusion

Our work demonstrated the importance of deliberate selection of the color channel combination in camera-based photoplethysmography and presents a systematic investigation for the quality of respiration rate measurement.

Further work should address the handling of low-frequent interferences as the main error source.

The non-contact measurement of vital signs such as respiration rate will help to comprehend human states or pathophysiological processes in settings that require unobtrusive measurements.

Acknowledgments

This work was supported by grants of the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG 319919706/GRK2323) and the TU Dresden Graduate Academy (financed from federal and state funds).

References

[1] van Gastel M, Stuijk S, de Haan G. Robust respiration detection from remote photoplethysmography. Biomedical Optics Express 2016;7(12):4941.

[2] Zaunseder S, Trumpp A, Wedekind D, Malberg H. Cardiovascular assessment by imaging photoplethysmography - a review. Biomedical Engineering Biomedizinische Technik jun 2018;63(5):617–634.

[3] Ernst H, Scherpf M, Malberg H, Schmidt M. Optimal color channel combination across skin tones for remote heart rate measurement in camera-based photoplethysmography. Biomedical Signal Processing and Control jul 2021; 68:102644.

[4] Reinhardt T, Schmahl C, Wüst S, Bohus M. Salivary cortisol, heart rate, electrodermal activity and subjective stress responses to the Mannheim Multicomponent Stress Test (MMST). Psychiatry Research 2012;198(1):106–111.

[5] Ernst H, Scherpf M, Malberg H, Schmidt M. Pulse Arrival Time - A Sensitive Vital Parameter for the Detection of Mental Stress. In Current Directions in Biomedical Engineering, volume 7. oct 2021; 419–422.

[6] Ernst H, Pannasch S, Helmert JR, Malberg H, Schmidt M. Cardiovascular Effects of Mental Stress in Healthy Volunteers. In 48th Conference Computing in Cardiology (CinC). Brno: IEEE, sep 2021; 1–4.

[7] Trumpp A, Rasche S, Wedekind D, Schmidt M, Waldow T, Gaetjen F, Plötze K, Malberg H, Matschke K, Zaunseder S. Skin Detection and Tracking for Camera-Based Photoplethysmography Using a Bayesian Classifier and Level Set Segmentation. In Bildverarbeitung für die Medizin 2017. Berlin, Heidelberg: Springer Vieweg, 2017; 43–48.

[8] Semechko A. Suite of functions to perform uniform sampling of a sphere (version 1.6.0.0), 2019.

[9] Villarroel M, Jorge J, Pugh C, Tarassenko L. Non-Contact Vital Sign Monitoring in the Clinic. In 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). IEEE, 2017; 278–285.

Address for correspondence:

Dipl.-Ing. Hannes Ernst
Institute of Biomedical Engineering, TU Dresden
Fetscherstr. 29, 01307 Dresden, Germany
hannes.ernst@tu-dresden.de