ROI Selection for Remote Photoplethysmography

Georg Lempe[1], Sebastian Zaunseder[1], Tom Wirthgen[2], Stephan Zipser[2], Hagen Malberg[1]

[1]Institut für Biomedizinische Technik, TU Dresden
[2]Fraunhofer IVI, Dresden
georg.lempe@tu-dresden.de

Abstract. Camera-based remote photoplethysmography (rPPG) is a technique that can be used to measure vital signs contactlessly. In order to optimize the extraction of photoplethysmographic signals from video sequences, we investigate the spatial dependence of the photoplethysmographic signal. For an evaluation of the suitability of various regions of interest for rPPG measurements, we conducted a study on 20 healthy subjects. We analysed the videos using a refined pulse amplitude mapping approach. Our results show that the signal-to-noise ratio of rPPG signals can be improved by limiting the region of interest to certain regions of the face.

1 Introduction

Photoplethysmography (PPG) is a technique used to non-invasively determine blood volume changes at a measurement site by measuring the absorption of light. In a clinical setting PPG is used to measure oxygen saturation, but cardiac and respiratory activity can be derived from PPG signals as well. Remote photoplethysmography (rPPG) makes use of optical imaging sensors to measure subtle skin-color changes resulting from blood volume changes in near-surface vessels [1]. The contactless measurement principle makes rPPG an interesting technology in applications outside the clinical environment where traditional biomedical monitoring is not feasible today. The decreasing prices and the abundant availability of digital cameras and processing power over the last years have contributed to a growing number of rPPG related publications.

However, the image and signal-processing methods used today are not yet robust enough to allow for a reliable monitoring of cardiac activity in everyday situations. Common artifacts due to movement or changes in illumination can surpass the signal by orders of magnitude and pose the main problem that hinders rPPG. Several signal processing approaches to improve the quality of rPPG signals were therefore proposed in the past [2, 3].

Another aspect of the signal extraction, the selection of the region of interest (ROI), has received little attention. In recent publications most commonly the entire face or static areas within are being used as ROI. Poh et al. determine

the ROI using face detection. Hereby the rectangle enclosing the detected face is shrunk and then used as ROI [2]. Sun et al., Lewandowska et al. and Takano et al. use rectangular ROIs centered on the forehead and the cheeks [4, 3, 5]. While in the latter cases the selection was done manually, Lewandowska et al. also propose a method to geometrically determine a ROI on the forehead in relation to the position of the eyes. Verkruysse as well as Hülsbusch and Blažek have shown that not all parts of the skin's surface undergo the same amount of color change [6, 7]. Even though these phenomena have been described qualitatively, no quantitative analyses have been conducted yet.

2 Materials and methods

We hypothesize that it is possible to improve the rPPG signal's quality by determining a well-suited ROI. Building upon the aforementioned studies by Hülsbusch, Blažek and Verkruysse we investigate whether it is possible to define an optimized ROI using Lewandowska's approach of defining landmarks in the face. To achieve this we apply Verkruysse's method of pulse amplitude mapping [6], extending it by using precise reference signals.

2.1 Measurement setup

We conducted a study on 20 healthy subjects (age: 26.5 ± 4 years, sex: 8 female, 12 male). The subjects were asked to lie down on a tilt table and to hold still during the measurement. The camera (IDS uEye 5240 CP-C) was mounted 1 m above the table's surface, in front of the face. The videos were recorded synchronously with data of conventional, contacting medical sensors. We recorded RGB videos of 3 minutes length with a color depth of 10 Bit per channel, a resolution of 300×200 px and a frame rate of 100 Hz. The subject was illuminated frontally by four fluorescent tubes (NARVA LT 58 W/025 universalwhite). The distance between the lamps and the subject's face was 1.5 m. The lamps were equipped with electronic ballast circuits that drive the tubes with a frequency of 40 kHz avoiding the flickering with the line frequency. As ground truth reference

Fig. 1. Pulse amplitude maps showing the distribution of the rPPG signal's quality. Blue stands for a low value of the signal quality, red for a high value.

we recorded a PPG from an ear clip on the subjects right earlobe as well as an electrocardiogram (ECG) with clamp electrodes with a sampling frequency of 1 kHz. An ADInstruments PowerLab 16/35 was used to record the output of these sensors.

2.2 Mapping the signal quality

For the following analyses only the green color channel was considered, since this channel contains the strongest rPPG signal. The videos were scaled down by a factor of 4 by averaging spatially over neighborhoods of 16 px. By evaluating the signal quality of each pixel of the scaled video, maps can be created that show the spatial distribution of the signal quality (Fig. 1). In this context we defined the signal quality based on the spectral features of the rPPG signal x_r

$$q = \frac{A_{hr}}{\sum_{i=0}^{N/2} \mathcal{F}\{x_r(i)\}} = \frac{\mathcal{F}\{x_r(i_{hr})\}}{\sum_{i=0}^{N/2} \mathcal{F}\{x_r(i)\}} \tag{1}$$

where A_{hr} is the signal's amplitude at the heart rate f_{hr} and $\mathcal{F}(x)$ is the discrete Fourier transform (Fig. 2). The heart rate is determined using the reference

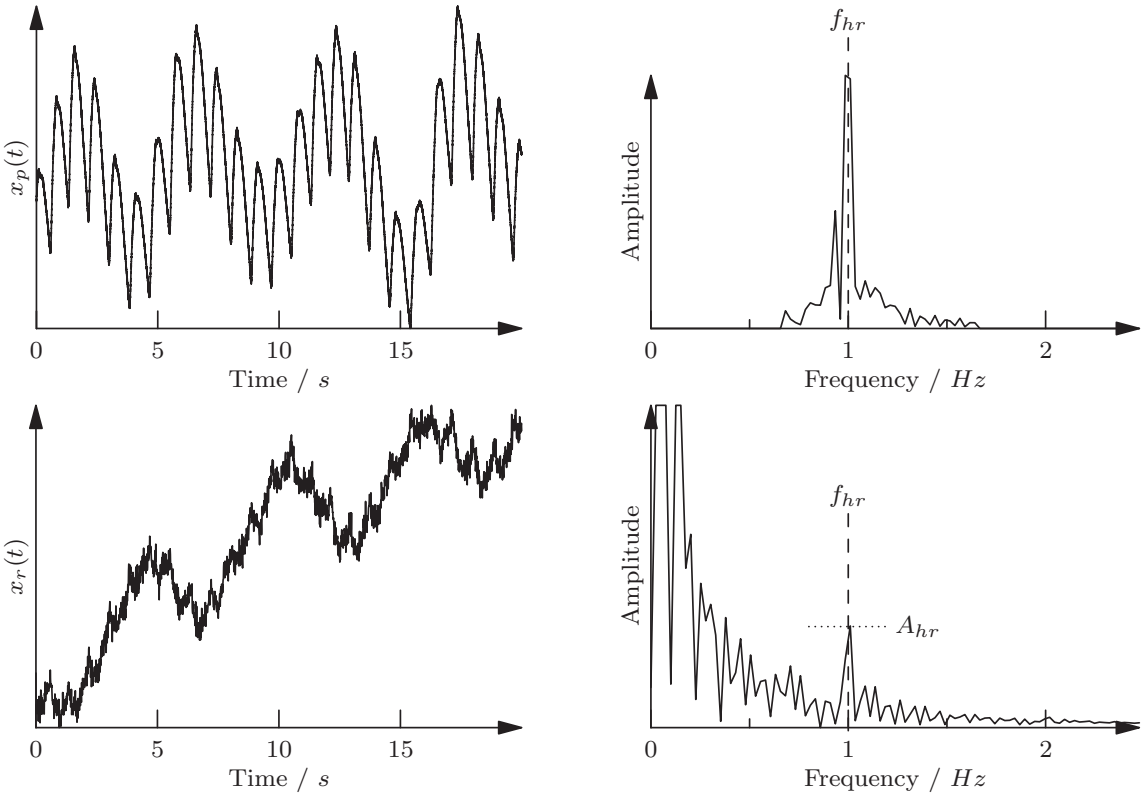

Fig. 2. Time signals and spectra of the reference PPG (top) and the rPPG (bottom) signals. The rPPG was derived using a 16 px neighborhood as ROI. The high frequent noise visible in its time signal is due to the comparatively small ROI, the flickering of the light source and the short exposure times / high sampling rates.

PPG. For each subject a total of 100 partially overlapping windows with a length of 20 seconds each were analysed in terms of Eq. 1.

2.3 Comparison of ROIs

In order to determine favorable ROIs we compared average signal qualities. As a reference, we manually labeled an ROI containing the subject's face. We also manually labeled landmarks whose coordinates were used to construct several square ROIs within the face. Fig. 3 shows the position of the landmarks and the ROIs. We determined each ROI's signal quality by averaging the values of the signal quality maps within the ROI. By applying this method to all windows of each measurement and averaging over all the results we determined one single signal quality parameter μ_q for each ROI.

3 Results

The methods described earlier lead to the generation of maps of the signal qualities as depicted in Fig. 1. These maps show a non uniform distribution of the signal quality. The averaged signal qualities μ_q of the single ROIs, as well as their standard deviations are shown in Fig. 3. It can be seen that similar ROIs show a similar signal quality. Comparing the highest and lowest signal quality the results show an improvement of 50 percent.

4 Discussion

In this article we have shown that the choice of the ROI has a significant influence on the quality of the rPPG signal. The resulting signal qualities suggest the

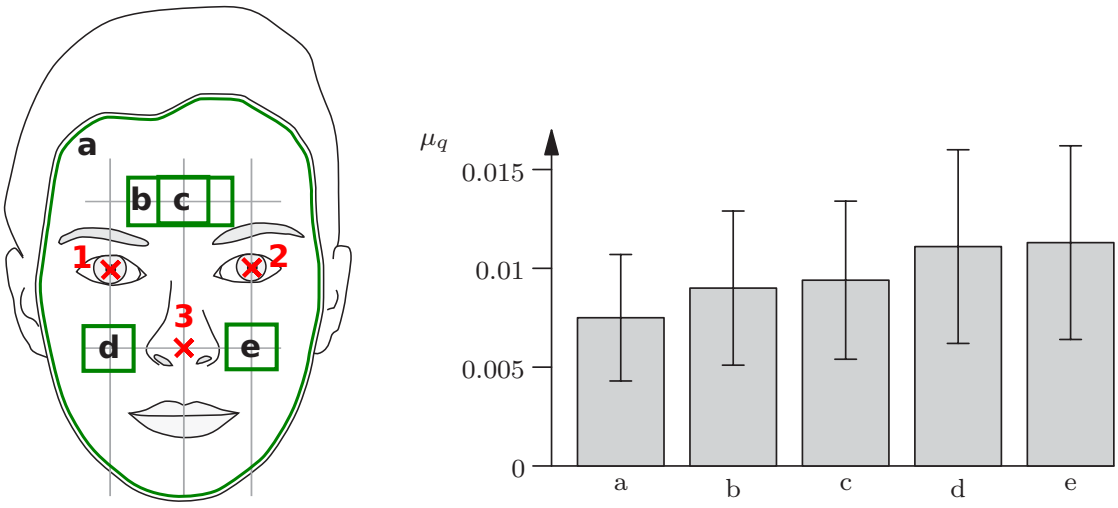

Fig. 3. Left: Position of the ROIs (green) and landmarks (red) within the face. Right: Mean signal quality μ_q and standard deviation for all measurements and windows. The cheeks, ROIs d and e, show the highest mean signal quality.

selection of the cheeks as ROIs for the extraction of rPPG signals. This is favorable since practical considerations support the use of the cheeks as ROIs as well. They are rarely covered by clothing or facial hair. We conclude that the selection of the entire face as an ROI, as it is used in state of the art publications, is not the optimal base for the extraction of rPPG signals. Using only parts of the face, such as the cheeks, for the signal extraction however introduces additional challenges. In future work segmentation algorithms capable of robustly tracking the desired ROIs have to be designed and tested for their applicability to rPPG. It has to be noted that these conclusions are valid for uniform illumination. The application of rPPG in everyday situations may require dynamic ROI selection in order to adapt the signal extraction to unfavorable lighting.

References

1. Blazek V, Wu T, Hoelscher D. Near-infrared CCD imaging: possibilities for non-invasive and contactless 2D mapping of dermal venous hemodynamics. Proc SPIE. 2000;3923:2.
2. Poh MZ, McDuff DJ, Picard RW. Advancements in noncontact, multiparameter physiological measurements using a webcam. Biomed Eng. 2011;58(1):7–11.
3. Lewandowska M, Nowak J. Measuring pulse rate with a webcam. J Med Imaging Health Inform. 2012;2(1):87–92.
4. Sun Y, Hu S, Azorin-Peris V, et al. Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise. J Biomed Opt. 2011;16:077010.
5. Takano C, Ohta Y. Heart rate measurement based on a time-lapse image. Med Eng Phys. 2007;29(8):853–7.
6. Verkruysse W, Svaasand LO, Nelson JS. Remote plethysmographic imaging using ambient light. Opt Expr. 2008;16(26):21434–45.
7. Huelsbusch M, Blazek V. Contactless mapping of rhythmical phenomena in tissue perfusion using PPGI. Proc SPIE. 2002;4683:110.