Heart Beat Detection and Analysis from Videos

Sebastian Zaunseder, Andreas Heinke, Alexander Trumpp, Hagen Malberg
TU Dresden, Faculty of Electrical and Computer Engineering
Institute of Biomedical Engineering
Dresden, Germany
sebastian.zaunseder@tu-dresden.de

Abstract — **Camera-based photoplethysmography provides a mean to monitor cardiovascular parameters remotely. This contribution tackles heart beat detection from remote photoplethysmograms (rPPG) by using a Wavelet-based detection algorithm and time delay estimation. Using experimental recordings of 18 healthy volunteers under resting conditions we demonstrate that depending on the used region of interest detection accuracies of 95% to 99% can be obtained. Time delay estimation is shown to improve the temporal exactness of the detections. Future work will be directed at the detection in non-resting periods and the automated identification of usable signal segments.**

Keywords — remote photoplethysmography, camera, beat detection, pulse wave, wavelet transform, time delay estimation, biosignal processing

I. Introduction

Photoplethymography is a technique which is used to detect blood volume changes [1]. Clinically the photoplethysmogram (PPG) is measured by applying a clip to the finger or the earlobe. The analysis of the PPG allows to derive statements on the pulse rate, respiration and oxygen saturation.

Recent investigations have shown that by using cameras photoplethysmography is possible at a distance [2], [3]. In principle the *remote photoplethysmogram* (rPPG) contains equivalent information to the conventional PPG. The possibility to record a PPG without any contact renders the rPPG an interesting option for applications were conventional sensors are not applicable (e.g. in skin-injuries or for monitoring neonates) or would mean a non-tolerable obstacle (e.g. in fitness applications). Moreover, differing from the conventional PPG which performs a punctual measurement, the rPPG carries information related to areas which gives the opportunity to analyze spatio-temporal features related to the hemodynamics.

The main problem concerning the rPPG is its low signal-to-noise ratio (SNR). Besides small amplitudes regarding the measurement signal the rPPG suffers from its sensitivity to artifacts. In the past solutions have been presented to overcome existing limitations concerning the applicability of camera- based photoplethysmography. Amongst others, the use of Inde- pendent Component Analysis (ICA) [2], Principal Component Analysis (PCA) [4] and chrominance-based methods [5] have been described. Using such methods researchers could show that the extraction of the mean heart rate (HR) becomes feasible even in situations of moderate movements. Most often the estimation of mean HR bases on a spectral representation of the preprocessed and segmented rPPG. A detailed analysis of hemodynamics, however, will require more than the mean heart rate. For example the detection of single heart beats from the rPPG and their morphological analysis must be pursued to add important diagnostic information.

Our work is concerned with the detection of single heart beats in rPPGs. Regarding this aim basic research concerning the feasibility and exactness must be conducted. Within this contribution we propose a method to detect heart beats by using a Wavelet-based approach and improve the temporal accuracy of the detections by applying time delay estimation. The impact of defining proper regions of interest (ROI) to construct the rPPG is evaluated by incorporating different facial areas in our analysis.

II. Materials And Methods

A. Measurement Setup and Used Recordings

Our analysis bases on recordings which were performed in our laboratory using the measuring fixture presented in [6] (see Fig. 1). The measurement site allows for video recordings by different cameras together with various synchronously recorded reference signals like the electrocardiogram (ECG), PPG (typically recorded from the finger or the earlobe) and the respiration.

The current contribution uses data from a study incorporating 20 healthy volunteers (aged 26.5 ± 4 years, no known cardiovascular disorders). Due to temporal inaccuracies in the reference recordings two subjects could not be used for the current analysis so that 18 subjects entered the analysis. During the experiment the volunteers remained in supine position and were asked to perform an experimental protocol consisting of head movements followed by a resting period. Fig. 2 illustrates the protocol and its resulting signals. In the resting period the frontal view of subjects' faces was completely visible. Videos were recorded using an industrial RGB camera (UI-5240CP-CHQ, IDS) at a resolution of 300x200 pixels and 100 frames per second. As reference signal we recorded the ECG (Einthoven I lead) and a PPG from the earlobe. Heart beat detections in the reference signals were generated automatically and manually corrected in order to provide a gold standard for our analysis. As our main concern was the feasibility of heart beat detection and the achievable accuracy the data used in this study originated from the resting period. From each subject a signal excerpt of 60 s was used for the evaluation.

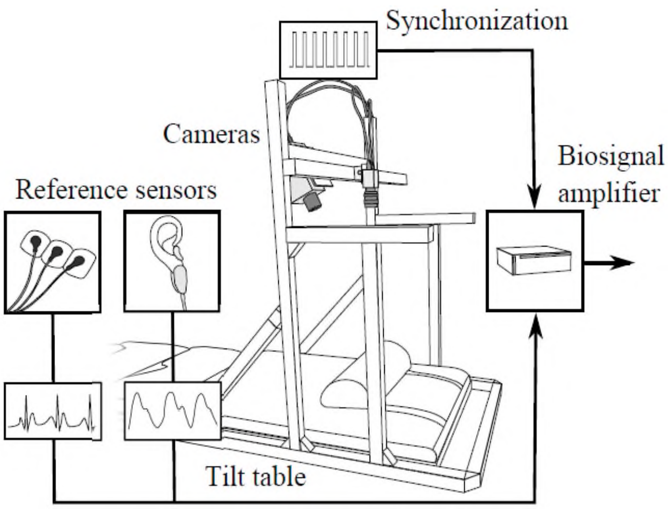

Fig. 1. Measuring fixture that is used to conduct basic research on the rPPG [7]

B. Processing

Signal and ROI selection: As the green channel was previously shown to carry the stongest information related to the pulsatile blood volume [3] our analysis only considers this channel. To allow a spatially resolved analysis we defined different ROIs semi-automatically. Fig. 3 contains the localization of ROIs. Region b-e were defined using fixed points in the face, namely the eyes and the nose [7]. Region a and f were manually defined. From each ROI an rPPG was derived. According to the definition of six ROIs six rPPG signals were created for each subject.

Heart beat detection: Detection of heart beats was done using the Wavelet Transform (WT). Details concerning the applied method can be found in [8]. Briefly, by using the quadratic spline wavelet and a shift invariant dyadic WT the rPPG was composed in scales up to scale 5. Each scale of the WT represents a smoothed and differentiated version of the input signal. By comparing scale 5 of the WT to an adaptive threshold relevant extrema were identified. Scale 2 was used to differentiate extrema which can be attributed to the pulse wave from the ones which can be assumed to have artificial origin. The application of the algorithm results in a set of beat detections located at time instants t_i.

Temporal alignment: Time delay estimation following Improved Woody's Method [9] was used to align the beat detections. The applied method maximizes the correlation between signal segments $x_i(k)$ located around the detections at t_i by iteratively shifting t_i together with its belonging signal segments. Fig. 4 contains the basic scheme which underlies TDE following Improved Woody's Method.

We applied TDE to all detections which occurred in the signals (note that from each subject 60 s were used). The length of x_i was chosen to be 84% of the median RR-interval which resulted from the detections in the 60 s. Differing from [9] we restricted the area in which a detection was allowed to be shifted to a window sized 120% of the median RR-interval. This restriction was introduced in order to avoid strong drifting of detections.

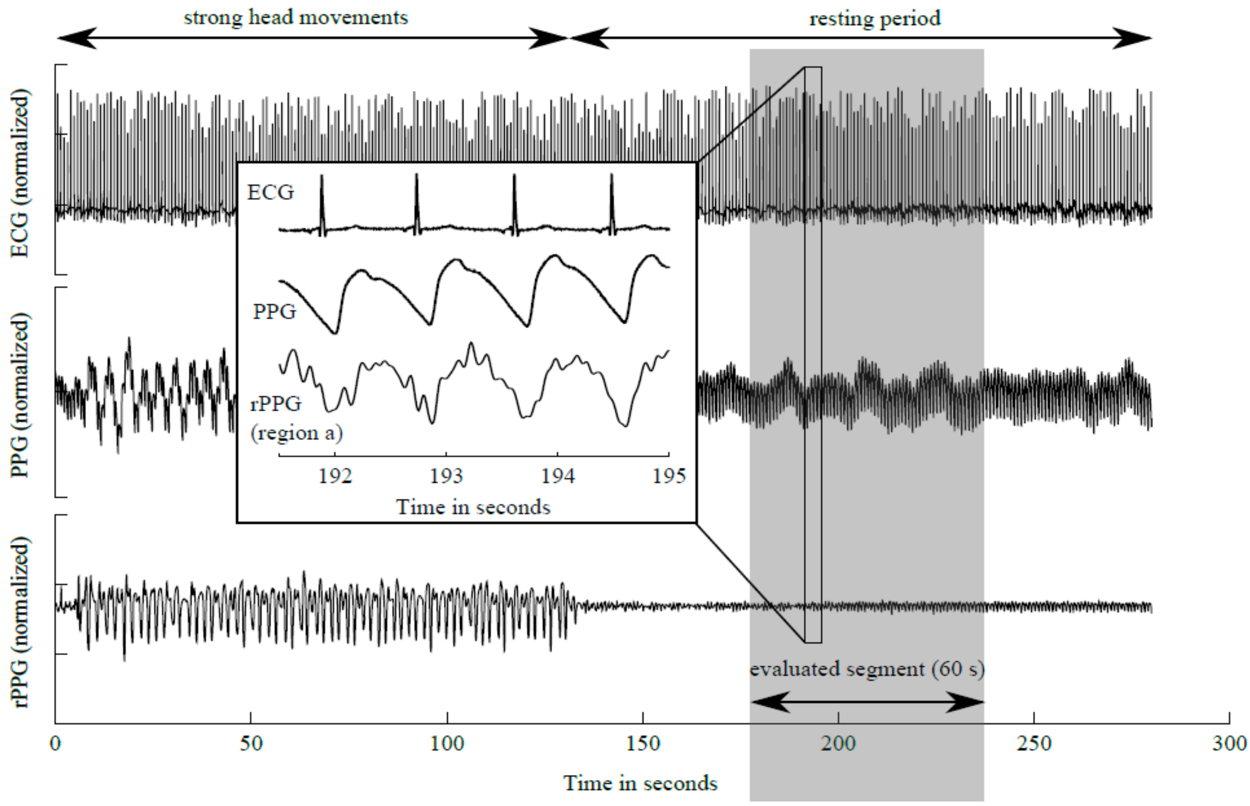

Fig. 2. Experimental protocol. Shown are the reference signals (ECG and earlobe PPG) together with the rPPG derived from region a (whole face, see Fig. 3)

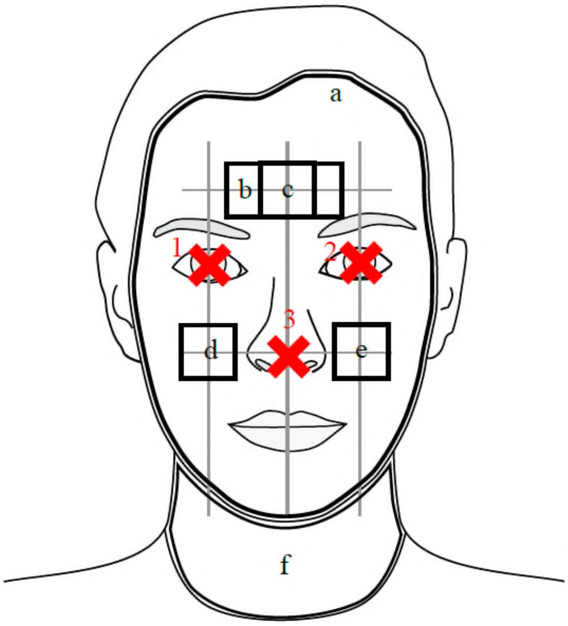

Fig. 3. ROI definition (a - whole face, b - forehead big, c - forehead small, d - right cheek, e - left cheek, f - neck).

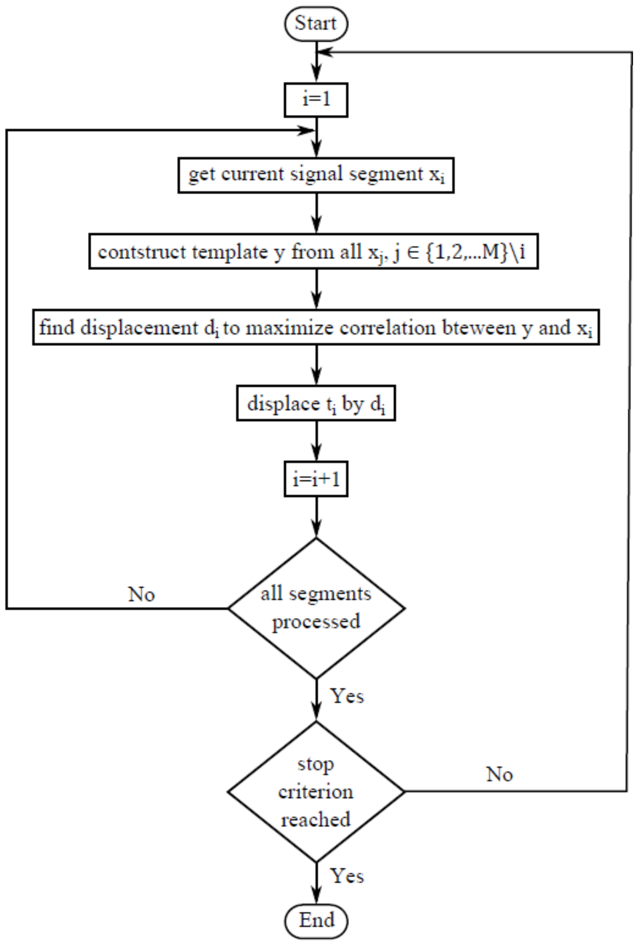

Fig. 4. Principal idea underlying Improved Woody's Method. The algorithm was used to temporally align heart beat detections. *M* denotes the number of beat detections occurring in the signal segments under consideration.

C. *Quality assessment*

The quality was assessed by means of beat detection accuracy and temporal exactness.

Table I contains the measures related to the beat detection accuracy. An algorithmic detection was considered as true positive (TP) detection if it occurred within 150 ms to an annotated beat in the reference PPG [10]. Detections which occur outside that range or dual detections were considered as false positive (FP). Missing detections were counted as false negative detections (FN).

In order to assess the temporal accuracy of the beat detections we compared the rPPG detections to the ones obtained from the earlobe PPG and the ECG. As quantitative measure we used the detection variability. The detection variability was computed by calculating the temporal difference of each true positive detection and its belonging reference detection. The standard deviation over such differences constitutes the detection variability. Ideally the detection variability should equal zero. Note that the detection variability should be considered together with the beat detection accuracy (Se, pP, Acc) as a low detection variability can be introduced by few detection.

TABLE I. APPLIED MEASURES TO ASSESS THE BEAT DETECTION ACCURACY (TP - TRUE POSITIVE DETECTIONS, FP - FALSE POSITIVE DETECTIONS, FN - FALSE NEGATIVE DETECTIONS).

Measure	**Calculation**
Sensitivity	$Se = \frac{TP}{TP + FN}$
Positive Prediction	$+P = \frac{TP}{TP + FP}$
Accuracy	$Acc = \frac{TP}{TP + FP + FN}$

TABLE II. MEAN BEAT DETECTION ACCURACY IN TERMS OF SE, +P AND ACC. GIVEN ARE THE MEAN VALUES AND STANDARD DEVIATION *STD* OVER ALL SUBJECTS AND FACIAL REGIONS A-F AS DEPICTED IN FIG. 3.

Region	**Se**	**+P**	**Acc**
	(all values given as mean ± *std*)		
a	0.993 ± 0.012	0.985 ± 0.032	0.979 ± 0.041
b	0.998 ± 0.005	0.992 ± 0.013	0.990 ± 0.015
c	0.995 ± 0.014	0.989 ± 0.019	0.985 ± 0.028
d	0.988 ± 0.022	0.980 ± 0.026	0.969 ± 0.042
e	0.991 ± 0.013	0.987 ± 0.022	0.979 ± 0.032
f	0.993 ± 0.012	0.956 ± 0.056	0.949 ± 0.058

III. RESULTS AND DISCUSSION

Table II contains the mean beat detection accuracy. Fig. 5 illustrates the results for Acc. Our analysis identifies the forehead to be the most appropriate region concerning the detection of single heart beats. For region b an accuracy of 99% could be obtained. The most pronounced drop concerning the detection accuracy occurs for region f, i.e. the neck, where an accuracy of 95% could be reached. The difference between both cheeks can be attributed to a single subject in which the detector performs considerably worse using right cheek than the left one.

Table III contains the mean detection variability before and after alignment of detections by TDE compared to the reference PPG and reference ECG. Fig. 6 illustrates the difference in detection variability before and after alignment for all subjects and regions (as reference the ECG was used).

In Fig. 6 positive values correspond to an improved temporal accuracy, i.e. the detection variability before alignment exceeds the one after alignment. As can be seen, TDE has a positive effect on the detection variability. As previously noted the detection variability should be considered together with the detection accuracy. However, as the accuracy was found to be not less than 80%, in most cases considerably above this value, the decreased detection variability is not erroneously generated.

In single cases a degradation of the results can be observed.

TABLE III. DETECTION VARIABILITY COMPARED TO THE REFERENCE PPG AND ECG. GIVEN ARE THE MEAN VALUES AND STANDARD DEVIATION STD OVER ALL SUBJECTS AND FACIAL REGIONS A-F AS DEPICTED IN FIG. 3.

Region	Compared to PPG Not-aligned	Compared to PPG Aligned	Compared to ECG Not-aligned	Compared to ECG Aligned
	(all values given as $mean \pm std$ in ms)			
a	13.45 ± 3.64	11.86 ± 3.71	14.09 ± 3.99	12.64 ± 4.01
b	11.08 ± 3.39	9.77 ± 2.51	11.61 ± 3.76	10.36 ± 3.00
c	10.14 ± 3.53	8.94 ± 3.16	10.84 ± 3.50	9.61 ± 3.08
d	12.84 ± 4.74	11.03 ± 3.40	13.45 ± 5.14	11.21 ± 3.50
e	11.79 ± 4.18	10.41 ± 3.33	12.31 ± 5.13	10.58 ± 3.37
f	15.10 ± 2.93	14.48 ± 3.73	15.49 ± 3.22	14.78 ± 3.61

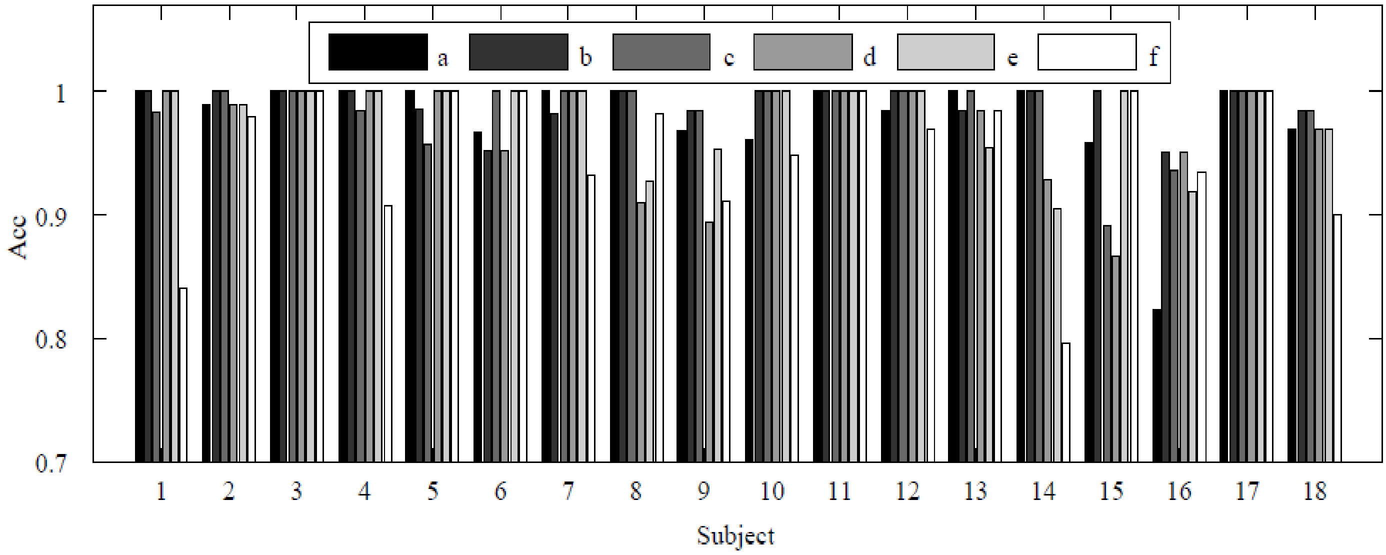

Fig. 5. Beat detection accuracy in terms of Acc for all subjects and facial regions a-f as defined in Fig. 3.

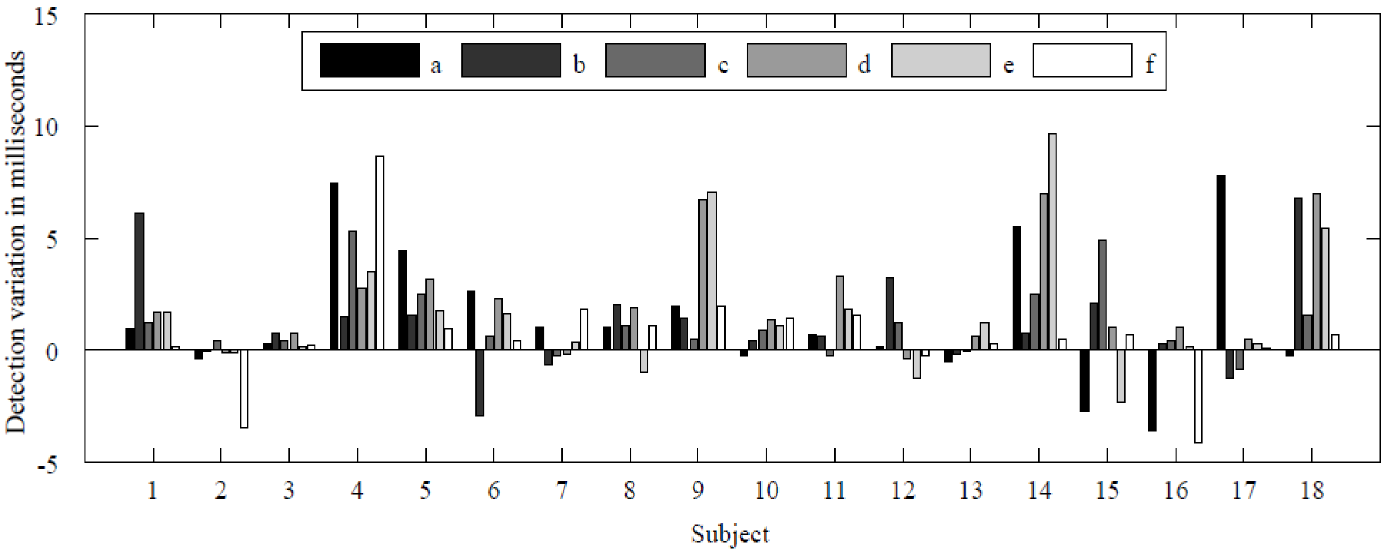

Fig. 6. Difference of the detection variation compared to the reference ECG before and after detection alignment by TDE for all subjects and facial regions a-f as defined in Fig. 3.

A deeper analysis revealed that in most cases where degradation occurs the rPPG is characterized by its particular low signal-to-noise ratio (SNR). The sensitivity of the proposed method to low SNRs also can be observed in region f. Here TDE fails to improve the poor temporal exactness by the same percentage than it does in other regions. Again, the forehead turns out to provide the most accurate results. The difference in detection variability using the ECG and the PPG was shown to be small. Regarding future applications which aim at a characterization of the autonomous nervous system this finding is of particular interest.

A general limitation of our study results from the employed data, i.e. included subjects and experimetal condition. Including subjects at rest simplyfies data processing significantly. Moreover, only young healthy subjects have been included. Both aspects must be taken into account when interpreting the results and considering an application under real world condition. The shown pilot study was intended to draw statements on the feasibilty of beat detection and establish the technical and procedural background for recordings in a clinical environment. Those recordings have recently started to face aspects related to the studied population, namely include a wider age distribution as well as cardiovascular diseases. So far, obtained results should be understood as proof of concept.

IV. Conclusion

We presented a method which can be used to detect and align single beats using the rPPG. Despite certain limitations our work shows that a highly accurate detection of single heart beats in the rPPG is feasible and TDE can contribute to improve the temporal alignment. The combination of the proposed processing strategy with suited preprocessing tech- niques, i.e. chrominance-based algorithms, should be evaluated to cope with situations of moderate movements in the future.

Acknowledgment

The authors would like to thank all volunteers participating in the experiments.

References

[1] J. Allen, “Photoplethysmography and its application in clinical physiological measurement.” Physiological measurement, vol. 28, no. 3, pp. R1–39, Mar. 2007.

[2] M.-Z. Poh, D. J. McDuff, and R. W. Picard, “Non-contact, automated cardiac pulse measurements using video imaging and blind source separation.” Optics express, vol. 18, no. 10, pp. 10 762–74, May 2010.

[3] W. Verkruysse, L.O. Svaasand, and J.S. Nelson, “Remote plethysmographic imaging using ambient light.” Optics express, vol. 16, no. 26, pp. 21 434–45, Dec. 2008.

[4] M. Lewandowska, J. Ruminski, T. Kocejko, and J. Nowak, “Measuring Pulse Rate with a Webcam - a Non-contact Method for Evaluating Cardiac Activity.” in FedCSIS, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., 2011, pp. 405–410.

[5] G. de Haan and V. Jeanne, “Robust pulse rate from chrominance-based rPPG.” IEEE transactions on bio-medical engineering, vol. 60, no. 10, pp. 2878–86, Oct. 2013.

[6] G. Lempe, S. Zaunseder, T. Wirthgen, S. Zipser, and H. Malberg, “ROI Selection for Remote Photoplethysmography,” in Bildverarbeitung fur die Medizin 2013. Springer, 2013, pp. 99–103.

[7] ——, “Kamerabasierte Erfassung kardiorespiratorischer Signale,” tm - Technisches Messen, vol. 80, no. 5, pp. 179–184, May 2013.

[8] A. Weder, M. Pietzsch, S. Zaunseder, M. Zimmerling, and S. Netz, “A mobile system for precise wireless pulse transit time (PTT) monitoring,” in Proceedings of the First ACM MobiHoc Workshop on Pervasive Wireless Healthcare - MobileHealth ’11, ser. MobileHealth ’11. New York, New York, USA: ACM Press, 2011, p. 1.

[9] A. Cabasson and O. Meste, “Time Delay Estimation: A New Insight Into the Woody’s Method,” IEEE Signal Processing Letters, vol. 15, pp. 573–576, 2008.

[10] “ANSI/AAMI EC57:1998 - Testing and reporting performance results of cardiac rhythm and ST-segment measurement algorithms.”