High-resolution motion-compensated imaging photoplethysmography for remote heart rate monitoring

Audrey Chung, Xiao Yu Wang, Robert Amelard, Christian Scharfenberger, Joanne Leong, Jan Kulinski, Alexander Wong, David A. Clausi

Department of Systems Design Engineering, University of Waterloo, 200 University Ave W, Waterloo, Canada

ABSTRACT

We present a novel non-contact photoplethysmographic (PPG) imaging system based on high-resolution video recordings of ambient reflectance of human bodies that compensates for body motion and takes advantage of skin erythema fluctuations to improve measurement reliability for the purpose of remote heart rate monitoring. A single measurement location for recording the ambient reflectance is automatically identified on an individual, and the motion for the location is determined over time via measurement location tracking. Based on the determined motion information motion-compensated reflectance measurements at different wavelengths for the measurement location can be acquired, thus providing more reliable measurements for the same location on the human over time. The reflectance measurement is used to determine skin erythema fluctuations over time, resulting in the capture of a PPG signal with a high signal-to-noise ratio. To test the efficacy of the proposed system, a set of experiments involving human motion in a front-facing position were performed under natural ambient light. The experimental results demonstrated that skin erythema fluctuations can achieve noticeably improved average accuracy in heart rate measurement when compared to previously proposed non-contact PPG imaging systems.

Keywords: Imaging photoplethysmography, skin erythema, motion compensation, non-contact heart rate

1. INTRODUCTION

With the advent of new medical imaging techniques, non-invasive medical assessment and diagnosis has become increasingly feasible. A commonly used non-invasive medical assessment technique is photoplethysmography (PPG), the technique of optically acquiring a volumetric measure of an organ. While non-invasive, PPG systems typically require physical contact, e.g., an oximeter sensor placed directly on the subject's extremity. This limits PPG devices to patients where constant physical contact is feasible. More recently, there have been advancements in non-contact PPG or PPG imaging, referring to the acquisition of PPG measurements through video. A step towards distance medicine, PPG imaging enables people in remote locations to receive a basic type of medical assessment without the physical presence of a medical professional. Non-contact PPG is also more sanitary, efficient, and cost-effective than contact PPG systems.

Hulsbusch and Blazek[1] developed one of the first PPG imaging devices. Using a cooled near infrared (NIR) CCD camera and an LED array, the system was used to evaluate rhythmic blood volume changes in wounds where a contact PPG device would be infeasible; however, the camera was bulky and expensive. In 2005, Wieringa *et al.*[2] proposed the use of a monochrome CMOS camera to capture three independent videos using 300 LEDs at different wavelengths in the red and NIR electromagnetic spectra (600 nm, 810 nm, 940 nm). While the results were promising, the camera had poor SNR at longer wavelengths and motion artefacts due to independently acquire videos. Humphreys *et al.*[3] later expanded on this design by automatically illuminating and capturing video using software triggers. The PPG imaging device was able to capture PPG-like signals using an array of 36 LEDs at 760 nm and 880 nm, but require a large amount of power to overcome the low camera sensitivity in the NIR spectrum.

Further author information: (Send correspondence to A.C.)
A.C.: E-mail: agchung@uwaterloo.ca

State-of-art video PPG techniques focus on better extracting heart rate information from video recordings using ambient illumination.[4,5] Wu *et al.*[4] proposed an algorithm called Eulerian Video Magnification (EVM) to visualise human pulse. EVM uses spatial decomposition and temporal filtering to amplify micro-changes in video that are otherwise invisible to the human eye. However, EVM amplifies all changes in a video including any ambient noise or motion, making it difficult to distinguish a heart rate in non-ideal video. Poh *et al.*[5,6] proposed the use of Independent Component Analysis (ICA) to determine physiological parameters such as heart rate, respiratory rate, and heart rate variability from blood volume pulse (BVP). While relatively robust to non-ideal video via facial tracking, ICA is a statistical technique for transforming a set of observations into components that are minimally dependent. ICA tends to generate inconsistent results due to the randomness in the decoupling algorithm, and has no direct biological interpretation. Thus, a consistent and robust method that is biologically motivated is desirable.

A novel approach for remote heart rate monitoring is proposed via motion-compensated erythema fluctuation analysis. As skin erythema has an excellent linearity with hemoglobin concentration,[7] it can be used to measure a subject's blood flow and allows for a biologically inspired method for video photoplethysmography. The paper will be structured as follows. Methodology is described in Section 2. Experimental setup and results are shown in Section 3. Lastly, conclusions and future work are discussed in Section 4.

2. METHODS

We present a novel, biologically-inspired method for remote heart rate monitoring through skin erythema fluctuation analysis. Using high-resolution video recordings of human bodies in natural ambient light, the proposed method compensates for body motion and extracts skin erythema information to calculate a subject's heart rate. To analyse the potential use of skin erythema fluctuations in determining heart rate, the following general algorithm framework (shown below in Fig. 1) was developed.

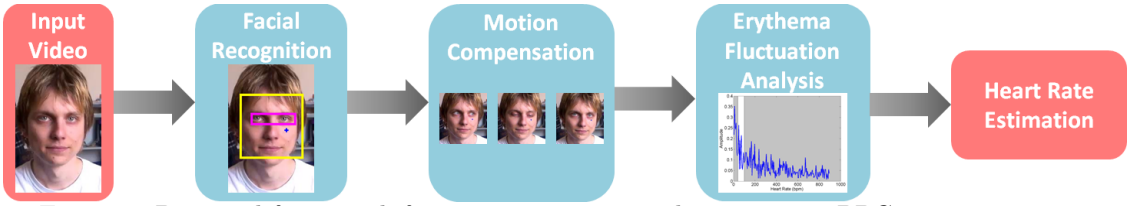

Figure 1: Proposed framework for motion compensated non-contact PPG imaging system.

2.1 Motion Compensation

Motion compensation via point tracking is used to make the acquisition of biometric signals from video more robust to temporal noise (e.g., natural human motion, and lighting fluctuations). To reduce computation time and background noise, a single sample $\underline{x}$ is selected and tracked:

$$\underline{x}_t = f(\underline{x}_{t-1}) \tag{1}$$

where $f(\underline{x}_{t-1})$ is the point tracking function (i.e., Kanade-Lucas-Tomasi (KLT) algorithm[8]). To eliminate large variations caused by point noise, the sample was expanded to an $n \times n$ pixel window centred at location $\underline{x}_t$. From Eq. 2, at a given time t,

$$r(t) = \mathrm{E}(r|\Omega(\underline{x}_t)) \qquad g(t) = \mathrm{E}(g|\Omega(\underline{x}_t)) \tag{2}$$

subject to

$$r(t_0) = \mathrm{E}(r|\Omega(\underline{x}_0)) \qquad g(t_0) = \mathrm{E}(g|\Omega(\underline{x}_0)) \tag{3}$$

where $\underline{x}_0$ is the initial sample location, $\Omega(\underline{x}_t)$ is the set of pixels in the $n \times n$ pixel window centred at $\underline{x}_t$, and $r(t)$ and $g(t)$ represent the expectation (denoted as E(.)) of the red and green values, respectively, given $\Omega(\underline{x}_t)$.

To ensure that a reasonable heart rate can be extracted from $\underline{x}$, the Viola-Jones algorithm[9] is used to register the subject's face and sub-features (i.e., eyes, mouth, and nose), and the initial sample location $\underline{x}_0$ is determined with respect to the sub-features. The initial sample point (denoted in Fig 2 using the blue "+") is selected to be on the subject's upper cheek based on facial skin thickness[10] and flatness of the area.

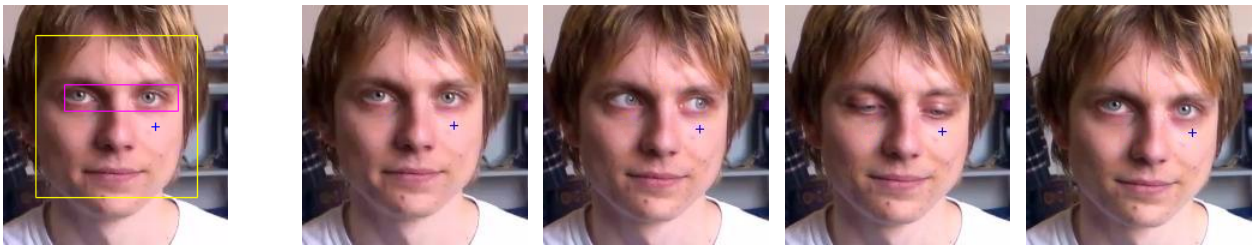

Figure 2: Facial recognition[9] and tracking[8] of selected sample point (denoted by blue "+"). The proposed method is robust to natural human motion via KLT point tracking.

2.2 Erythema Fluctuation Calculation

Given the motion compensated sample point $\underline{x}$, skin erythema fluctuation analysis of the sample is used to extract a heart rate for the subject. Skin erythema has an excellent linearity with hemoglobin concentration,[7] and can be used to measure a subject's blood flow. As such, skin erythema fluctuation analysis allows for a biologically inspired method for video photoplethysmography. A modification on Stamatas *et al.*,[11] Gong *et al.* proposed the calculation of skin erythema values for the quantification of hemoglobin and melanin from skin images using the following equation:

$$e(t) = \log_{10} \frac{1}{g(t)} - \log_{10} \frac{1}{r(t)} \tag{4}$$

where $r(t)$ and $g(t)$ are the red and green channels, respectively, as defined in Eq. 2.

As fluctuation in skin erythema corresponds to the flow of blood through a subject's face, an analysis of the frequency of erythema fluctuation can be used to estimate a subject's heart rate. Given the time series representation of skin erythema $e(t)$, the frequency representation of the erythema signal $E(u) = \mathcal{F}\{e(t)\}$ can be used to determine an estimation of the subject's heart rate in the video:

$$u_{\text{HR}} = \arg\max_u \ |E(u)| \qquad \text{subject to} \qquad \alpha \leq H(u) \leq \beta \tag{5}$$

where $H(u) = 60u$ is a function for converting frequency (Hz) to heart rate (bpm).

The average resting heart rate is between 60 bpm and 100 bpm,[12] with well-trained athletes averaging resting heart rate between 40 bpm and 60 bpm. As a result, the lower limit α is set to 40 bpm and the upper limit β is set to 100 bpm. The bpm corresponding to the highest amplitude within the range of plausible heart rates in the frequency domain is selected as the subject's estimated heart rate HR, i.e., $\text{HR} = H(u_{\text{HR}})$.

3. RESULTS

3.1 Experimental Setup

The proposed method determines the mean heart rate of a subject using eight videos 11 to 16 seconds in length. The test videos used in this study were recorded of five human subjects, S1 (author A.C.), S2 (author J.L.), S3 (author J.K.), S4 (author X.Y.W.), S5 (author A.W.), who have full knowledge of the study. All subjects were healthy at the time of the recordings. The test videos were taken of the subjects at rest, and all motion was assumed to be from the subject. All videos used for testing feature a single front-facing subject in natural ambient light, and were recorded in 1080p at 30fps using a static mobile phone (HTC One S).

Heart rate measurements were recorded for each video using a consumer level pulse oximeter (the Easy Pulse Sensor version 1.1). The Easy Pulse uses a HRM-2511E sensor, an infrared sensor in physical contact with the subject's finger to provide a heart rate reading of the test subject.[13] To evaluate the potential of erythema fluctuation analysis (EFA) for heart rate computation, it was compared to EVM[4] and ICA[5] using the set test videos. Both EVM and ICA were implemented as closely as possible to described methods. For the purposes of testing, a sample window size of $n = 11$ was used in the EFA algorithm, and the proposed EFA method was averaged over 10 iterations using various samples distributed across the subject's upper cheeks and central lateral forehead.

3.2 Experimental Results

The proposed EFA method, EVM,[4] and ICA[5] were evaluated using a set of eight videos, each with a single front-facing subject at rest. The estimated heart rate was computed for all test videos using each algorithm, and percentage errors were calculated relative to the heart rate measurements obtained via the Easy Pulse Sensor. The percentage error of each algorithm was analysed, and the mean and standard deviation of each is presented below in Table 1.

Table 1: Comparison of percentage errors for subjects at rest for proposed EFA method, EVM,[4] and ICA.[5] The proposed method has the lowest mean percentage error and standard deviation (std).

Algorithm	**Percentage Error (mean ± std)**
EFA	$\mathbf{12.1 \pm 10.6}$
EVM[4]	28.9 ± 16.6
ICA[5]	20.0 ± 17.3

Table 1 shows that EVM[4] has the highest mean percentage error. While the proposed EFA method and ICA[5] use facial recognition and tracking to compensate for natural human motion and limit the heart rate estimation to areas containing the subject's face, EVM amplifies micro-changes across the full video frame. Thus, EVM is less robust to background noise, lighting fluctuations, and subject motion, resulting in a relatively high percentage error.

Table 1 also shows that the proposed EFA method clearly has the lowest mean percentage error, as well as the lowest standard deviation. The relatively low percentage error is likely due to the use of a sample pixel window on the subject's face (as opposed to averaging red, green, and blue channels across the entire face[5]), further reducing the effect of temporal noise. In addition, the proposed EFA method yields consistent and reproducible results, while ICA[5] produces varying heart rate estimations due to the random nature of the statistical method.

To determine statistical significance, t-tests were performed comparing the proposed EFA method, EVM,[4] and ICA.[5] As EVM had the worst performance, it was selected as a baseline distribution. All t-tests were heteroscedastic and conducted assuming two-tailed distributions. The t-test between EVM and ICA resulted in a p-value of 31.2%, indicating no statistical significance in the difference in percentage error distribution. However, the t-test between EVM and EFA had a p-value of 3.3%, indicating that there is a significant difference in percentage error distributions. Thus, the proposed EFA method shows a significant improvement in estimating a subject's heart rate from video.

4. DISCUSSION

In this paper, we presented a novel motion-compensated erythema fluctuation-based method for remote heart rate monitoring. The proposed method was compared with other state-of-art video PPG algorithms using four single-subject test videos taken in ambient light. The experimental results indicate that the proposed method (EFA) performs better than state-of-art methods (t-test significance of 3.3%), and is a good candidate for robust non-contact heart rate estimation. Future works include more extensive testing of subjects in different lighting conditions, the incorporation of more robust and dynamic sampling methods, and examining the feasibility of EFA to capture other cardiovascular metrics such as heart rate variability and breathing rate.

ACKNOWLEDGMENTS

This research was undertaken, in part, thanks to funding from the Canada Research Chairs program. The study was also funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Ontario Ministry of Economic Development and Innovation.

REFERENCES

[1] Huelsbusch, M. and Blazek, V., "Contactless mapping of rhythmical phenomena in tissue perfusion using PPGI," in *Proc. SPIE*, **4683**, 110–117 (2002).

[2] Wieringa, F., Mastik, F., and van der Steen, A., "Contactless multiple wavelength photoplethysmographic imaging: A first step towards "SpO2 camera" technology," *Annals of Biomedical Engineering* **33**(8), 1034–1041 (2005).

[3] Humphreys, K., Ward, T., and Markham, C., "Noncontact simultaneous dual wavelength photoplethysmography: a further step toward noncontact pulse oximetry," *Review of scientific instruments* **78**(4), 044304 (2007).

[4] Wu, H.-Y., Rubinstein, M., Shih, E., Guttag, J. V., Durand, F., and Freeman, W. T., "Eulerian video magnification for revealing subtle changes in the world," *ACM Transactions on Graphics* **31**(4), 65 (2012).

[5] Poh, M.-Z., McDuff, D. J., and Picard, R. W., "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Transactions on Biomedical Engineering* **58**(1), 7–11 (2011).

[6] Poh, M.-Z., McDuff, D. J., and Picard, R. W., "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Optics Express* **18**(10), 10762–10774 (2010).

[7] Gong, H. and Desvignes, M., "Hemoglobin and melanin quantification on skin images," in *Image Analysis and Recognition*, Campilho, A. and Kamel, M., eds., *Lecture Notes in Computer Science* **7325**, 198–205, Springer (2012).

[8] Tomasi, C. and Kanade, T., [*Detection and tracking of point features*], School of Computer Science, Carnegie Mellon Univ. Pittsburgh (1991).

[9] Viola, P. and Jones, M., "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, **1**, I–511, IEEE (2001).

[10] Ha, R. Y., Nojima, K., Adams Jr, W. P., and Brown, S. A., "Analysis of facial skin thickness: defining the relative thickness index," *Plastic and Reconstructive Surgery* **115**(6), 1769–1773 (2005).

[11] Stamatas, G. N., Zmudzka, B. Z., Kollias, N., and Beer, J. Z., "Non-invasive measurements of skin pigmentation in situ," *Pigment Cell Research* **17**(6), 618–626 (2004).

[12] American Heart Association, "Target Heart Rates." http://www.heart.org/. (22 March 2013).

[13] R-B., "Easy pulse (version 1.1) sensor overview (part 1)." http://embedded-lab.com/blog/?p=7336. (20 April 2013).