CameraHRV: Robust measurement of Heart Rate Variability using a Camera

Amruta Pai, Ashok Veeraraghavan, and Ashutosh Sabharwal

Scalable Health Labs, Department of Electrical and Computer Engineering,
Rice University, 6100 Main Street, Houston, Texas, USA 77005

ABSTRACT

The inter-beat-interval (time period of the cardiac cycle) changes slightly for every heartbeat; this variation is measured as Heart Rate Variability (HRV). HRV is presumed to occur due to interactions between the parasympathetic and sympathetic nervous system. Therefore, it is sometimes used as an indicator of the stress level of an individual. HRV also reveals some clinical information about cardiac health. Currently, HRV is accurately measured using contact devices such as a pulse oximeter. However, recent research in the field of non-contact imaging Photoplethysmography (iPPG) has made vital sign measurements using just the video recording of any exposed skin (such as a person's face) possible. The current signal processing methods for extracting HRV using peak detection perform well for contact-based systems but have poor performance for the iPPG signals. The main reason for this poor performance is the fact that current methods are sensitive to large noise sources which are often present in iPPG data. Further, current methods are not robust to motion artifacts that are common in iPPG systems. We developed a new algorithm, CameraHRV, for robustly extracting HRV even in low SNR such as is common with iPPG recordings. CameraHRV combined spatial combination and frequency demodulation to obtain HRV from the instantaneous frequency of the iPPG signal. CameraHRV outperforms other current methods of HRV estimation. Ground truth data was obtained from FDA-approved pulse oximeter for validation purposes. CameraHRV on iPPG data showed an error of 6 milliseconds for low motion and varying skin tone scenarios. The improvement in error was 14%. In case of high motion scenarios like reading, watching and talking, the error was 10 milliseconds.

Keywords: Imaging photoplethysmography, frequency demodulation, heart rate variability

1. INTRODUCTION

Heart rate variability (HRV) captures the variation in the time interval between consecutive heartbeats.[1] Currently, HRV is accurately measured using the electrocardiographic device (ECG). Pulse rate variability (PRV) obtained from the Photoplethysmographic (PPG) signal uses pulse cycles as a substitute to the RR intervals obtained from the QRS complex of the ECG signal. PRV is equivalent to HRV obtained from the ECG signal when the subject is at rest.[2] HRV parameters are potential indicators for stress levels and cardiac health. Thus, HRV has gained significance over the past few decades.

Contact-based systems like the ECG and pulse oximeter have limited practicability. Such systems are predisposed to infection-related concerns due to their contact nature. Contact devices are inconvenient to use during HRV data collection procedures lasting up to 24 hours. Non-contact based HRV measurements obtained from imaging photoplethysmography can potentially provide a solution to such limitations.

Imaging Photoplethysmography (iPPG)[3–5] uses the phenomenon that as blood volume changes in the arteries and capillaries beneath the skin surface, the intensity of light reflected by the blood also changes. This change in the reflected light can be measured using a simple consumer grade camera even under an ambient light setting. Thus, by using a simple video recording of a person's face (or exposed skin surface) various vital signs like heart rate can be extracted accurately.[6] Few successful attempts[7–9] have been made on extracting HRV from the iPPG signal. Custom peak detection algorithms used in these attempts can lead to incorrect inter-beat interval

Send correspondence to Amruta Pai at ap52@rice.edu. This work was partially supported by NSF ERC Grant, EEC-1648451 (PATHS-UP).

estimates. Incorrect inter-beat interval estimates arise due to false peaks and artifacts present in the iPPG signal of low signal quality. HRV parameters are directly derived from the values of the inter-beat interval estimates. Incorrect beat-to-beat estimates introduce a large error in the HRV parameters.

Research over past few decades has explained how the HRV arises due to modulation of the heart rate frequency by the sympathetic and parasympathetic nervous system.[10] In this paper, we propose CameraHRV, which leverages the concept that HRV information is related to the instantaneous frequency variation of the periodic pulse signal.[11–13] This relation allowed us to extract the phase of the iPPG signal by frequency demodulation. Also, using the prior information known about HRV, we obtained a continuous beat to beat HRV measurement. We compared the performance of CameraHRV to current HRV estimating methods like peak detection,[8] and continuous wavelet transform filtering.[14] We showed that for iPPG signals that are of low signal quality, obtaining HRV information through instantaneous frequency is more reliable than current time domain approaches like peak detection.

We quantified the error in all the methods for different iPPG signal quality resulting due to varying skin tones, and for different high motion scenarios like reading, watching videos, and talking. CameraHRV improves the accuracy of time domain HRV parameters by 10 milliseconds for dark skin tone subjects, and by 5 milliseconds for motion scenarios. CameraHRV shows a percent error of 20%. The improvement is large over continuous wavelet filtering method which has an error of 34%. The peak detection approach breaks with an error of 98%. Under scenarios of low motion, CameraHRV presents an error of 6 milliseconds across different skin tones. But under high motion scenarios, CameraHRV also fails to provide accurate HRV measurements along with the other methods. Hence, there is scope for further improvement of the algorithm for high motion scenarios.

CameraHRV is a step towards increasing the potential of non-contact based HRV measurements. Potential scenarios include the use of the front facing cameras on the laptops or smartphones during online activities without the requirement of a wrist-worn device.

2. CAMERAHRV

The proposed algorithm CameraHRV had three main steps. First, was the detection of feature points of the face in the frames of the video, and dividing the detected face area into smaller regions of interest (ROIs). Second, was the extraction of the iPPG signal from the ROIs. Third, was the extraction of instantaneous frequency, and reconstruction of the first harmonic of the iPPG signal.

2.1 Face detection and ROI selection

We used the OpenFace[15,16] toolkit for facial landmark detection. It tracks 68 feature points on the face. We obtained the 68 feature points for every frame of a particular subject's video recording. For every frame, a boundary was created with the feature points as shown in the Figure 1. The 8 feature points marking the eyebrows were shifted upwards to include the forehead region as well. We masked all the regions outside the boundary created with the feature points. We also masked the eye areas, and the lip regions. The unmasked area of the frame in the videos was divided into smaller grids each of 20×20 pixels. Each grid was tracked using the KLT tracker.[17,18] The tracked grids confined in the unmasked region of the video frames were selected as the regions of interest (ROIs).

2.2 Extraction of imaging photoplethysmography signal from videos

After identifying the ROIs, we robustly extracted the iPPG signals from these regions. It is established in earlier research[6] that the absorption spectra of the prominent blood chromophores are maximum in the passband range of the green filters used in color cameras. Hence, we used the pixel intensities of the single green channel video recordings to obtain the iPPG signal. We first averaged the pixel intensities within each ROI to obtain the reflected light intensity signal as

$$R(r,t) = c(r) + x(r,t) + w(r,t), \tag{1}$$

where $R(r,t)$ was the reflected light intensity signal. $c(r)$ was the constant light reflection component by skin surface, and tissue, which depends only on ROI r. $x(r,t)$ represents the pulsatile component, which was our signal of interest. $w(r,t)$ was the additive noise component dominated by motion artifacts, and changes in illumination.

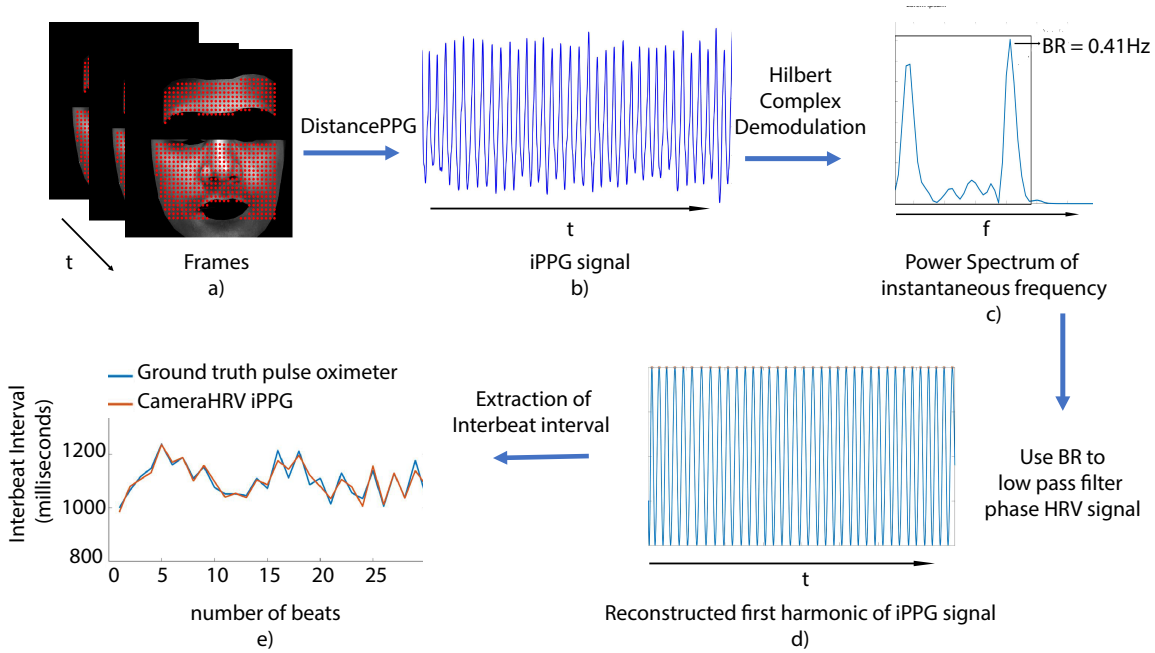

Figure 1. a) Tracking ROIs bounded by facial landmark points using KLT. b) iPPG signal waveform obtained. c) Power Spectrum of instantaneous frequency after Hilbert transform, and complex demodulation. d) Reconstructed first harmonic of iPPG signals after low pass filtering of phase. e) Inter beat intervals estimated from video compared with inter beat intervals from ground truth pulse oximeter.

.

The $w(r,t)$ component was stronger for subjects with darker skin tone due to higher melanin content present in the skin. High melanin content hinders absorptions of light by the chromophores present in the blood.

To obtain an iPPG signal of good signal quality we adopted the Maximal Ratio Combining algorithm,[8] summarized as follows. First, the reflected light intensity signal was temporally filtered using the bandpass filter [0.5,4] Hz which corresponds to the heart rate range of 30 bpm to 240 bpm. The filtering also helped to remove the noise present outside the band of interest. The filtered signal $y(r,t)$ was dependent on the ROI, and time. $x(r,t)$ was the pulsatile component in the passband of the filter, and $n(r,t)$ was the noise present with the signal in the passband, resulting in,

$$y(r,t) = x(r,t) + n(r,t). \tag{2}$$

We defined an amplitude range for every ROI as $A(r) = \max(y(r,t)) - \min(y(r,t))$ in accordance with the earlier work.[8] We rejected ROIs where $A(r) > 8$. We found a coarse estimate of pulse rate PR by spatially averaging $y(r,t)$ from all non-rejected ROIs, and detecting the peak in the power spectral density. The PR was used to define a goodness metric $G(r)$ for every ROI as follows:

$$G(r) = \frac{\int_{PR-b}^{PR+b} Y(r,f)}{\int_{0.5}^{4} Y(r,f) - \int_{PR-b}^{PR+b} Y(r,f)}, \tag{3}$$

where $Y(r,f)$ was the power spectral density of the filtered signal $y(r,t)$. The HRV information associated with the fundamental harmonic was present in the sidebands of the PR. HRV information is generally present between [0.04,0.45] Hz in the spectrogram of the instantaneous frequency. Hence, we estimated the power of the signal of interest around the PR with b as 0.45 Hz.

We calculated the $A(r)$, and $G(r)$ for every non-rejected region, and computed its mean as A_{mean}, and G_{mean} respectively. The spatially combined iPPG signal $p(t)$ is then computed using all the non-rejected ROIs

as follows,

$$p(t) = \sum y(r,t) I(A(r) < A_{mean} \text{ and } G(r) > G_{mean}), \tag{4}$$

where $I(.)$ was the indicator function. Note that we did not compute the product of $G(r)y(r,t)$ in the summation as done in earlier work.[8] The pulse signal is modulated by the HRV information. It is present in the sidebands around the first harmonic in the spectrogram of the iPPG signal. We were only interested in the first harmonic of the signal so computing the product did not hold relevance in our application. The goodness metric did not reflect the signal to noise ratio of the first harmonic alone. Hence, we just spatially averaged regions as shown in equation 4.

2.3 Extraction of instantaneous frequency and reconstruction iPPG signal first harmonic

Once the iPPG signal from the videos are obtained, the next step was to perform frequency demodulation on the iPPG signals to extract the HRV information. The iPPG signal $p(t)$ was a quasi-periodic signal with a slight variation in the heart rate frequency. This slight variation about the constant frequency was represented as the variation in phase of the signal as follows,

$$\frac{1}{2\pi}\frac{\mathrm{d}\phi(t)_{hrv}}{\mathrm{d}t} = f_{hrv}. \tag{5}$$

Thus, we modeled the first harmonic of the PPG signal $p_{h1}(t)$ as

$$p_{h1}(t) = a(t)\cos(2\pi f_{hr} t + \phi(t)_{hrv}), \tag{6}$$

f_{hr} was computed by detecting the peak in the power spectral density of $p(t)$. For obtaining the first harmonic $p_{h1}(t)$, $p(t)$ is bandpass filtered around the constant frequency, $[f_{hr} - 0.6, f_{hr} + 0.6]$ Hz.

Using the principle of complex demodulation, and Hilbert transform[11–13] we obtained $\phi(t)_{hrv}$. First, we computed the Hilbert transform of $p_{h1}(t)$ as $H(p_{h1}(t))$, which converts the real signal $p_{h1}(t)$ into an analytic signal $H(p_{h1}(t))$ as follows,

$$H(p_{h1}(t)) = a(t)e^{i(2\pi f_{hr} t + \phi(t)_{hrv})}. \tag{7}$$

We then demodulated the analytic signal as

$$H(p_{h1}(t))e^{(-i2\pi f_{hr} t)} = a(t)e^{i(\phi(t)_{hrv})}. \tag{8}$$

As the demodulated phase HRV signal was now in the zero-frequency band, we temporally filtered it with a low pass filter with a cut-off frequency of 0.55 Hz. The HRV information is generally present between [0, 0.45] Hz but we used a larger band for filtering so that the roll off of the narrow low pass filter does not alter the HRV information. The beat to beat variations is often in the range of a few ten milliseconds. Thus, the signal strength of HRV is very low. Noise present in the iPPG signal easily corrupts the instantaneous phase, and frequency variation of the signal. Thus, peak detection on the iPPG signal provides erroneous estimates.

Filtering between [0, 0.55] Hz gives us a coarse estimate of the phase hrv signal $\phi(t)_{hrv}$ which was noisy. To obtain a better estimate we filtered it using a low pass filter with a cut-off frequency dependent on the breathing rate. The high-frequency spectra of HRV [0.15 0.45]Hz is associated with respiratory arrhythmia, and tracks the breathing rate across a range of frequencies. The respiratory arrhythmia is the main modulator in the high-frequency range of the HRV information. Hence, we used the breathing rate frequency BR to calculate the cut-off frequency of the low pass filter. The cut-off frequency was calculated in the following manner.

We obtained f_{hrv} from the property of equation 5. The power spectral density of the f_{hrv} was obtained as F_{hrv}. Although BR is the primary modulator, slight power in frequencies beyond the BR also contribute to the value of the HRV parameters. But the power in the frequencies of range [BR, 0.55] Hz was less compared to the power associated with BR peak. These frequency peaks were easily corrupted by noise. But the power of the noise peaks was not higher than the power of the BR peak. It was also not as low as other peaks due to HRV information. Using this observation we used a power threshold $pow_{th} = 0.2 \cdot F_{hrv}(BR)$. The final cut-off frequency was defined as follows.

$$f_{cutoff} = \max(f) \text{ such that } f > BR \text{ and } F_{hrv}(f) < pow_{th}. \quad (9)$$

where f was the range of frequencies [BR, 0.55] Hz. A value of 0.2 was chosen because it worked well for our dataset consisting of 30 data points.

As the bandwidth of the filter was very narrow typically of 0.4 Hz, a gradual roll-off leads to an inaccurate estimation of HRV metrics. Thus, we used the finite impulse response filter (FIR filter) of order 350 for a sampling rate of 30 Hz. After detecting the BR peak, if the cut-off frequency was found to be less than 0.4 Hz then we set it to a default of 0.4 Hz. Otherwise, it would lead to considerable HRV information being filtered out. We do lose some HRV information but we prevent too much error in our parameter estimates.

The filtered demodulated phase HRV signal is lp. As lp is an analytic signal the corrected phase HRV signal is obtained as follows:

$$\hat{\phi}(t)_{hrv} = \tan^{-1} \frac{imag(lp)}{real(lp)}. \quad (10)$$

We filtered $\hat{\phi}(t)_{hrv}$ using the same cut-off frequency discussed above to remove any distortions arising due to the nonlinear inverse tangent transformation. Then, we reconstructed the first harmonic component of the iPPG signal as follows:

$$p_{camera}(t) = \cos(2\pi f_{hr} + \hat{\phi}(t)_{hrv}). \quad (11)$$

$p_{camera}(t)$ was then used to obtain the inter beat interval, and HRV time domain parameters.

3. RESULTS AND DISCUSSIONS

3.1 Methods for comparison

To test the performance of the proposed algorithm CameraHRV, we compared it with two other HRV estimating methods.[8,14] For all the three methods the distancePPG algorithm[8] was used to obtain the iPPG signal. The HRV estimation methods adopted for comparison purposes are the following. In the first method,[8] a simple peak detection was used on the iPPG signals. We named the first method as DistancePPG-Peak. Peak detection methods are susceptible to false peaks that completely skew HRV measurements. So, we compared with a second method[14] that was a continuous wavelet filtering method. In the continuous wavelet filtering method, a wavelet energy curve was used to filter continuous wavelet transform coefficients of the iPPG signal. After filtering, an inverse continuous wavelet transform was used to obtain a less noisy iPPG signal. We named this second method as DistancePPG-CWT. We used the PPG signal obtained from an FDA approved pulse oximeter as ground truth.

3.2 HRV Time domain parmeters

Comparison of the three algorithms was performed for two-time domain HRV parameters. The parameters were $SDNN$ which is standard deviation of interbeat intervals, and $SDSD$ which is standard deviation of successive difference in interbeat intervals.[19]

The interbeat interval (IBI) of the n^{th} heartbeat is defined as $IBI(n) = Time_{peak}(n) - Time_{peak}(n-1)$. Successive differences of interbeat intervals is defined as $SD(n) = IBI(n) - IBI(n-1)$. $SDNN_{hrv}$ is standard deviation of IBI. $SDNN_{hrv}$ is related to the total power of HRV, and the long-term HRV variation. $SDSD_{hrv}$ is standard deviation of SD. $SDSD_{hrv}$ is related to the HF power of HRV, and the short-term HRV variation.

To calculate the interbeat interval, $findpeaks$ function in Matlab was used to detect the peaks on the PPG signal estimates. Spurious peaks located too close to each other were rejected. We imposed a condition that consecutive peaks should be at least $0.5/f_{hr}$ seconds apart.

3.3 Dataset

To understand the effect of different signal quality on the accuracies of the time domain HRV parameters, we used two different datasets. The two different datasets were the skin tone dataset, and the motion dataset from the DistancePPG paper.[8] The skin tone dataset consists of 18 single channel(green) videos taken from 9 subjects of varying skin tones ranging from pale white to dark, 5 male, and 4 female. Each subject provided two 40 seconds videos. The subjects were asked to remain stationary but the involuntary movement was not restricted. Involuntary motion seen in the videos consisted of slight facial movements. The motion dataset consists of a total of 12 single channel(green) videos of 40 seconds duration each taken from 4 subjects of a fair skin tone. Each subject volunteered 3 videos where they were either only reading, watching or talking. The motion dataset had an additional 6 motion videos from 2 subjects of a darker skin tone which we used separately. These dataset collections were approved by the Rice University Institutional Review Board (Protocol number: 14-45E, Approval Date: 3/04/2014). The motion seen in the motion dataset consists of a large amount of facial movement, and head motion. Table 1 shows the range of HRV present in both the datasets across different videos of all the subjects.

Table 1. Range of values of the $SDNN_{hrv}$,$SDSD_{hrv}$ in the different datasets. The parameters are calculated using the gold standard pulse oximeter

Ground truth HRV Time Domain Parameters over 40 second time window Mean ± standard deviation in milliseconds		
Dataset	Skin Tone (18 videos, 9 subjects)	Motion (12 videos, 5 subjects)
$SDNN_{hrv}$	(45 ± 23) ms	(51 ± 16) ms
$SDSD_{hrv}$	(52 ± 30) ms	(47 ± 16) ms

3.4 CameraHRV Performance

We divided the performance analysis of CameraHRV into two parts. First, we analyzed error in the two-time domain parameters for different skin tones ranging from pale white to dark brown skin tone as shown in Figure 2, and 3. Second, we analyzed the error for high motion affected scenarios such as reading, watching videos, and talking.

3.4.1 Skin Tone

As the quality of the iPPG signal varies considerably with skin tone, it is essential to test how the algorithm performs for different skin tones. Figure 2, and Figure 3 shows the absolute error in milliseconds for $SDNN_{hrv}$, and $SDSD_{hrv}$ parameters respectively, calculated from iPPG signals acquired from skin tone dataset.

The DistancePPG-Peak detection method gives erroneous values for darker skin tones as simple peak detection method is vulnerable to noisy peaks, and motion artifacts. Although the peak detection algorithm rejected some spurious peaks, DistancePPG-Peak performs poorly, especially for dark skin tone subjects. It presented a mean absolute error of 19 ms, and root mean square error of 33 ms for $SDNN_{hrv}$.

The DistancePPG-CWT method presented a mean absolute error of 12 ms, and root mean square error of 17 ms for $SDNN_{hrv}$. The continuous wavelet filtering method attenuated the power in the sidebands around the main heart rate frequency. As the HRV information was present in the sidebands, in some cases this attenuation led to the suppression of the high frequency component of HRV. Thus, the standard deviation of HRV is found to be lower than the actual value.

The mean absolute error, and root mean square error for $SDNN_{hrv}$ calculated by CameraHRV method is 5 ms, and 7 ms respectively which is less compared to other methods.

As the standard deviation of successive differences of HRV $SDSD_{hrv}$ is a more sensitive parameter, all the three algorithms presented a larger error. CameraHRV provided a mean absolute error of 8 ms, and root mean square error 10 ms compared to DistancePPG-CWT which had an error of 24 ms, 35 ms, and DistancePPG-Peak 39 ms, 64 ms respectively.

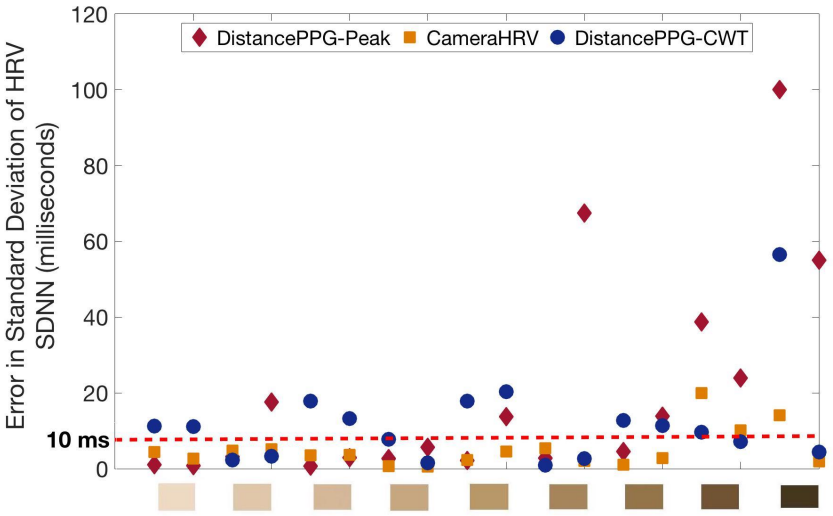

Figure 2. Absolute Error in standard deviation of inter-beat intervals ($SDNN_{hrv}$) from iPPG (two videos of each of the 9 subject) of different skin tones. Error is obtained by comparing with $SDNN_{hrv}$ from pulse oximeter PPG. CameraHRV reduces the error in Standard deviation of HRV below 10 ms for most skin tones.

.

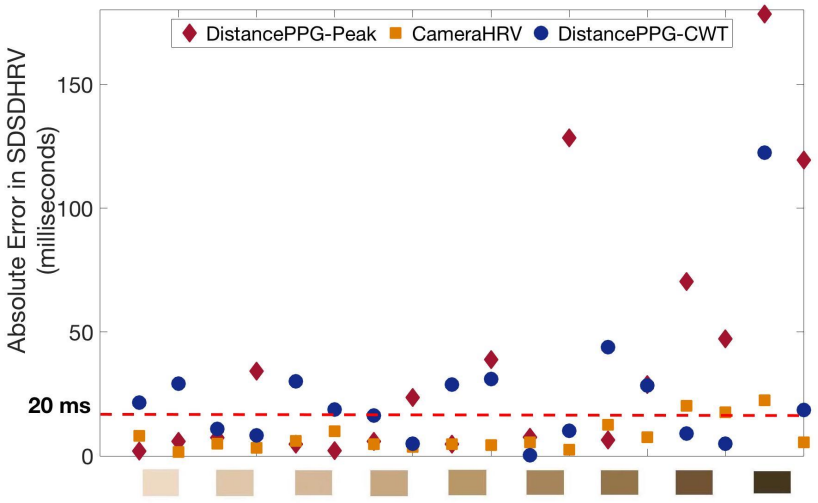

Figure 3. Absolute Error in standard deviation of successive differences in inter-beat intervals ($SDSD_{hrv}$) from iPPG (two videos of each of the 9 subject) of different skin tones. Error is obtained by comparing with $SDSD_{hrv}$ from pulse oximeter PPG. CameraHRV reduces error below 20 ms for most skin tones.

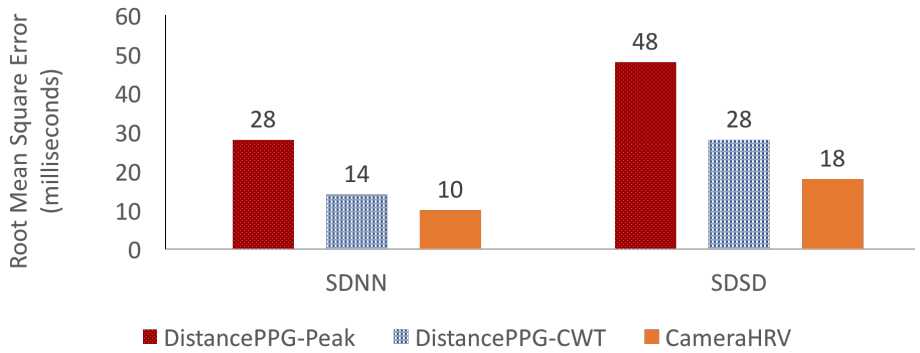

Figure 4. Root mean square error in standard deviation of inter-beat intervals ($SDNN_{hrv}$), and standard deviation of successive differences in inter-beat intervals ($SDSD_{hrv}$) calculated between iPPG signals for the motion dataset (12 videos comprising of subjects talking, watching or reading from 4 fair tone subjects), and ground truth pulse oximeter. CameraHRV has a slightly better performance than the other methods.

3.4.2 Motion

The motion was another main contributor to noise in the iPPG signals. Figure 4 shows the root mean square error in $SDNN_{hrv}$, and $SDSD_{hrv}$ parameters.

The mean absolute error, and root mean square error for $SDNN_{hrv}$ in case of CameraHRV was 8 ms, 10 ms respectively, DistancePPG-CWT was 9 ms, 14 ms respectively, and DistancePPG-Peak was 19 ms, and 28 ms respectively. The mean absolute error, and root mean square error for $SDSD_{hrv}$ in case of CameraHRV was 11 ms, 18 ms respectively, DistancePPG-CWT was 17 ms, 28 ms respectively, and DistancePPG-Peak was 34 ms, 48 ms respectively.

The CameraHRV method gave less error than other two methods. The subjects were all of fair skin tone. Thus, the error is largely due to motion artifacts while reading, watching, or talking. We combined all the motion scenarios together as different subjects had a different amount of head motion for different activities.

We also tested the algorithm for motion videos with subjects of a dark skin tone. All the three algorithms performed poorly for such videos. Although CameraHRV performed better in presence of motion compared to other current methods, there is scope for improvement in the algorithm. The reason CameraHRV failed was that the noise present in the sidebands overpowered the BR peak which led to incorrect filtering, and larger error in HRV parameter values.

4. CONCLUSION AND FUTURE WORK

We present CameraHRV an algorithm for non-contact measurement of heart rate variability. We validated our algorithm against the ground truth pulse oximeter for different skin tones, and motion scenarios. We also compared the performance of CameraHRV to other algorithms which are currently used. We showed CameraHRV performed significantly better than the standard algorithms for different skin tones with low motion scenario, and slightly better in case of high motion scenarios. It has a mean absolute error of 5 ms, and 8 ms for $SDNN_{hrv}$, and $SDSD_{hrv}$ respectively when tested for different skin tones. In case of motion, it presents a mean absolute error of 8 ms, 11 ms respectively. There is a need to improve the performance of CameraHRV in case of motion scenarios. Future work may include developing an iPPG signal estimator aimed at attaining a higher accuracy in HRV parameters. A better technique for reducing noise in the phase of the iPPG signal may also be investigated.

ACKNOWLEDGMENTS

This work was partially supported by NSF ERC Grant EEC-1648451 (for PATHS-UP ERC). We would also like to thank Mr. Mayank Kumar for his valuable suggestions, and discussions.

REFERENCES

[1] Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M., and Suri, J. S., "Heart rate variability: a review," *Med Biol Eng Comput* **44(12)**, 1031–1051 (2006).

[2] Schafer, A. and Vagedes, J., "How accurate is pulse rate variability as an estimate of heart rate variability? a review on studies comparing photoplethysmographic technology with an electrocardiogram," *International Journal of Cardiology* **166(1)**, 15–29 (2013).

[3] Allen, J., "Photoplethysmography and its application in clinical physiological measurement.," *Physiological Measurement* **28(3)**, R1–R39 (2007).

[4] Verkruysse, W., Svaasand, L. O., and Nelson, J. S., "Remote plethysmographic imaging using ambient light.," *Optics Express* **16(26)**, 21434–21445 (2008).

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

[6] Sun, Y. and Thakor, N. V., "Photoplethysmography revisited: From contact to noncontact, from point to imaging.," *IEEE Trans. Biomed. Engineering* **63(3)**, 463–477 (2016).

[7] Sun, Y., Hu, S., Azorin-Peris, V., Kalawsky, R., and Greenwald, S., "Noncontact imaging photoplethysmography to effectively access pulse rate variability.," *Journal Biomedical Optics* **18(16)**, 0161205 (2013).

[8] Kumar, M., Veeraraghavan, A., and Sabharwal, A., "DistancePPG: Robust non-contact vital signs monitoring using a camera," *Biomedical Optics Express* **6(5)**, 1565 (2015).

[9] McDuff, D., Gontarek, S., and Picard, R. W., "Remote measurement of cognitive stress via heart rate variability.," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* **2014**, 2957–60 (2014).

[10] Saul, J., "Beat-to-beat variations of heart rate reflect modulation of cardiac autonomic outflow," *Physiology* **5(1)**, 32–37 (1990).

[11] Hayano, J., Barros Filho, A. K., Kamiya, A., Ohte, N., and Yasuma, F., "Assessment of pulse rate variability by the method of pulse frequency demodulation," *Biomedical engineering online* **4**, 62 (2005).

[12] Li, H., Kwong, S., Yang, L., Huang, D., and Xiao, D., "Hilbert-huang transform for analysis of heart rate variability in cardiac health," *IEEE/ACM Trans. Comput. Biol. Bioinformatics* **8(6)**, 1557–1567 (2011).

[13] Chou, Y., Zhang, A., Wang, P., and Gu, J., "Pulse rate variability estimation method based on sliding window iterative dft and hilbert transform," *Journal of Medical and Biological Engineering* **34(4)**, 1557–1567 (2013).

[14] Bousefsaf, F., Maaoui, C., and Pruski, A., "Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate," *Biomedical Signal Processing and Control* **8(6)**, 568 – 574 (2013).

[15] Baltrusaitis, T., Robinson, P., and Morency, L.-P., "Openface: An open source facial behavior analysis toolkit.," *WACV* , 1–10 (2016).

[16] Baltrusaitis, T., Robinson, P., and Morency, L.-P., "Constrained local neural fields for robust facial landmark detection in the wild," *Proceedings of the 2013 IEEE International Conference on Computer Vision Workshops* , 354–361 (2013).

[17] Lucas, B. D. and Kanade, T., "An iterative image registration technique with an application to stereo vision," *Proceedings of the 7th International Joint Conference on Artificial Intelligence* **2**, 674–679 (1981).

[18] Tomasi, C. and Kanade, T., "Detection and tracking of point features," *International Journal of Computer Vision* (1991).

[19] "Heart rate variability standards of measurement, physiological interpretation, and clinical use task force of the european society of cardiology the north american society of pacing electrophysiology," *Circulation* **93(5)**, 1043–1065 (1996).