Measuring Heart Rate and Heart Rate Variability with Smartphone Camera

Donghao Qiao[1], Farhana Zulkernine[1], Raihan Masroor[2], Roshaan Rasool[3], Nauman Jaffar[3]

[1]*School of Computing*, *Queen's University*, Kingston, Ontario, Canada

[2]*Your Doctors Online*, Toronto, Ontario, Canada

[3]*SenSights.AI*, *MarkiTech*, Toronto, Ontario, Canada

{d.qiao, farhana.zulkernine}@queensu.ca; raihan@yourdoctors.ca; {roshaan.rasool, nauman.jaffar}@markiteck.ca

***Abstract*—Vital signs are important parameters that can reflect people's physiological status and help physicians provide medical advice. Remote Photoplethysmography (rPPG) is a fast, low-cost and convenient method to remotely collect biometric data, and requires only a facial video recorded using a smartphone or other camera. Remote medical service provisioning proved to be a dire need during the COVID-19 pandemic. To leverage the cloud-based medical advice provisioning platform of Your Doctors Online, we propose a rPPG methodology to measure people's Heart Rate (HR) and Heart Rate Variability (HRV) based on a facial video recorded by the users using a smartphone. We validate our model on the TokyoTech remote PPG dataset.**

***Keywords*—rPPG, vital signs, heart rate, heart rate variability**

I. Introduction

Measurement of cardiac parameters like Heart Rate (HR) and Heart Rate Variability (HRV) are important indicators of a person's physiological and emotional well-being. HR is the number of times the heart beats in a minute. The time interval between two heartbeats is called the Inter-Beat Interval (IBI) and the variation of consecutive IBIs is called HRV [1]. Oximeter uses the principle of Photoplethysmography (PPG) [2] wherein the skin is illuminated with light, and based on the proportion of the blood volume flowing through the tissues, some light is absorbed by the body while the rest is reflected. By monitoring the amount of the reflected light, the Blood Volume Pulse (BVP) signal is computed to measure vital signs. The remote rPPG method for measuring vital signs is based on this principle of PPG but allows a contactless camera-based method for vital signs measurement. When light is reflected from the skin surface there is a slight color change on the skin of the face, which can be tracked with face videos captured with cameras. However, the videos recorded using mobile devices have motion and light noise. In this project, we try to diminish these influences and build an application to measure users' HR and HRV in real time using the front camera of a smartphone.

The overall workflow is demonstrated in Fig. 1 which includes 5 steps. First, a user's face is detected in the video frame by frame and part of the facial area is selected (Fig. 1. (a)) as the Region of Interest (RoI). Second, raw BVP signals, the average values of pixels in each of the three channels in every frame, are extracted from the RoI. Plots of the raw BVP signals of the three channels are shown in Fig. 1. (b). Third, multiple filters are applied to the green channel raw BVP signal (Fig. 1. (c)) to extract clean and desired BVP signal (Fig. 1. (d)). Fourth, peak detection is applied on the BVP signal to calculate the HRV in time domain. Fifth, the Power Spectral Density (PSD) of the signal is computed to estimate HR in frequency domain (Fig. 1. (e)). We validated our method on the publicly available TokyoTech remote PPG dataset [3] containing videos recorded under realistic conditions and achieved a mean absolute error of 1.49 ± 2.20 bpm. The detailed methodology and experimental results are presented in Section II and III respectively.

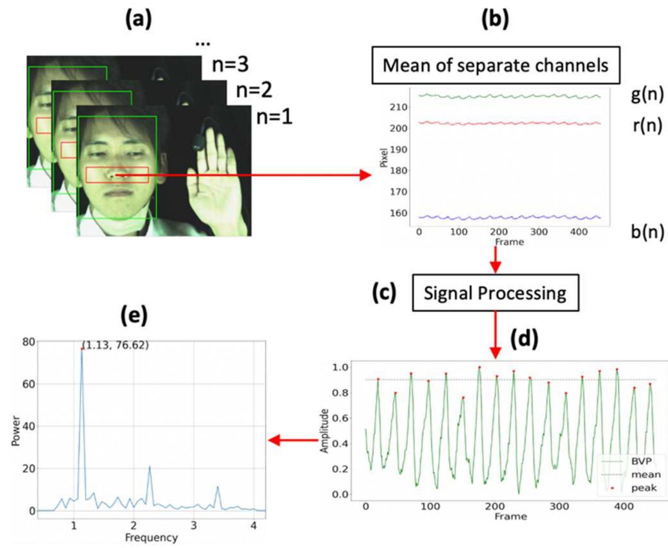

Fig. 1. Overall workflow. (a) Detect face and extract RoI. (b) Calculate the mean of separate channels in every frame and extract the raw BVP signal (g(n)). (c) Process the raw BVP signal with multiple filters to get the final BVP. (d) Detect the peaks of the BVP signal for HRV claculation. (e) Calculate the PSD of BVP in frequency domain for HR calculation.

II. Methodology

A. RoI Detection

The OpenCV Deep Neural Network (DNN) face detector [4] is applied to detect user's face, which is based on the Single Shot Detector (SSD). Face detection can localize the bounding box (green box in Fig. 1. (a)) of the entire face in the image and is used to detect the face frame by frame from videos. A smart phone or video camera is used to either pre-record and upload or stream the video to a cloud server. After detecting the face, part of the face area is selected as the RoI in every video frame. In order to reduce motion noise caused by blinking eyes or mouth movements, we only leverage the face area around the nose.

As shown in Fig. 1. (b), each RoI has three RGB channels: red, green and blue. The raw BVP signal of each frame is

This project is funded by Mitacs and Your Doctors Online.

[2] https://yoursdoctors.online [3] https://sensights.ai/ https://markitech.ca

2375-0324/21/$31.00 ©2021 IEEE
DOI 10.1109/MDM48529.2021.74

represented by the mean pixel values of the green channel since it contains the strongest plethysmography signal [5]. Hence, the RoI generates a raw BVP signal $g(n)$ where $n \in [1, N]$ represents the number of frame.

B. Signal Processing BVP Extraction

1) Denoise filter: After obtaining the raw BVP signals, a denoise filter is applied to remove the large jumps and steps caused by the motion noise such as head movements and hand shaking. We calculate the absolute differences between the consecutive signals and remove the signals for which the difference is above a predefined threshold to reduce the noise.

2) Normalization: The signal is normalized by subtracting it from the mean and dividing the result by the standard deviation.

3) Independent Component Analysis (ICA): ICA [6] is always used to separate the independent signals from a mixed signal set. Assume that a linear mixture of the observed signal $x = As$. The purpose of ICA is to find a matrix W which is the inverse of A and obtain the independent component by $s = Wx$. ICA randomly returns positive or negative signal. So, the extracted signal may be reversed sometimes. We select the signal that has a higher correlation with the input signal.

4) Detrending filter: A detrending filter [7] is applied to the signal obtained from ICA for plethysmography signal processing. The method is based on formulation of a smoothness prior and can filter low frequency trend. A smoothing parameter λ is applied to adjust the frequency response.

5) Moving Average Filter: Finally, a moving average filter (Eq. (1)) with $L = 5$ is utilized to remove random noise.

$$g(n) = \frac{1}{L}\sum_{k=0}^{L-1} g(n-k) \qquad (1)$$

C. Vital Signs Calculation

1) Heart Rate (HR): To calculate the HR, a bandpass filter is applied to select frequency-of-interest [2, 5]. The cutoff frequencies are 0.7 and 4 Hz, which correspond to 42 and 240 bpm. Welch's method [8] is applied to compute the PSD in frequency domain. In Fig. 1. (e), the peak coordinate (1.13, 76.62) indicates the HR frequency f_{HR} as 1.13 and the maximum power as 76.62. The final predicted HR of the video is $60 \times f_{HR}$ bpm.

2) Heart Rate Variability (HRV): IBI is the time period between the consecutive heartbeats. We detect the peaks (red dots in Fig. 1. (d)) of the extracted BVP signal. $IBI = t_n - t_{n-1}$ where t_n is the time of the n^{th} detected peak. The Root Mean Square of Successive RR Interval Differences (RMSSD) [1] as shown in Eq. (2), is used to calculate HRV.

$$RMSSD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(IBI_i - IBI_{i+1})^2} \qquad (2)$$

where N is the number of frames.

III. Experiments

The model is implemented using Python3 and is evaluated on TokyoTech Remote PPG Dataset [3].

A. TokyoTech Remote PPG Dataset

This dataset consists of 9 subjects (8 male and 1 female) and each subject has three 1-minute videos corresponding to three sessions: relax, exercise and relax. The participants perform hand grip before the exercise session. The resolution is 640×480 and the frame rate is 30 Frame-Per-Second (fps). A finger clip contact PPG sensor (Procomp Infinity T7500M, Thought Technology Ltd., Canada) is used to gather the BVP signals which represent the gold standard measurements in the dataset.

B. Results

Mean Absolute Error (MAE) and Standard Deviation (std) are applied to evaluate our model as shown in Table. I.

TABLE I. EXPERIMENT RESULTS

	HR (bpm)	HRV (ms)
	MAE* ± *std	***MAE* ± *Std***
Relax	1.68 ± 2.51	19.58 ± 22.65
Exercise	1.57 ± 2.32	27.42 ± 32.54
Relax	1.22 ± 1.62	25.97 ± 29.63
Overall	1.49 ± 2.20	24.33 ± 28.66

IV. Conclusions and Future Works

In this project, we developed a rPPG application to measure users' HR and HRV using a smartphone's front camera. and validated the same using a public rPPG dataset. We achieved 1.49 bpm MAE on HR calculation. However, HRV is more sensitive to noise such as illumination and motion noise, and the future work will focus on improving noise filtering to obtain a more accurate HRV and measuring other biometric data such as blood pressure and oxygen saturation.

References

[1] Shaffer, F. and Ginsberg, J.P., 2017. An overview of heart rate variability metrics and norms. Frontiers in public health, 5, p.258.

[2] Sun, Y. and Thakor, N., 2015. Photoplethysmography revisited: from contact to noncontact, from point to imaging. IEEE transactions on biomedical engineering, 63(3), pp.463-477.

[3] Maki, Y., Monno, Y., Yoshizaki, K., Tanaka, M. and Okutomi, M., 2019, July. Inter-Beat Interval Estimation from Facial Video Based on Reliability of BVP Signals. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 6525-6528). IEEE.

[4] OpenCV, opencv face detector, GitHub repository. Available online: https://github.com/opencv/opencv/tree/4.0.0-beta/samples/dnn/face_detector. (accessed on February 10, 2021)

[5] Rouast, P.V., Adam, M.T., Chiong, R., Cornforth, D. and Lux, E., 2018. Remote heart rate measurement using low-cost RGB face video: a technical literature review. Frontiers of Computer Science, 12(5), pp.858-872.

[6] Hyvärinen, A. and Oja, E., 2000. Independent component analysis: algorithms and applications. Neural networks, 13(4-5), pp.411-430.

[7] Tarvainen, M.P., Ranta-Aho, P.O. and Karjalainen, P.A., 2002. An advanced detrending method with application to HRV analysis. IEEE Transactions on Biomedical Engineering, 49(2), pp.172-175.

[8] Welch, P., 1967. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Transactions on audio and electroacoustics, 15(2), pp.70-73.