An Improvement for Video-based Heart Rate Variability Measurement

Peixi Li[1], Yannick Benezeth[1], Keisuke Nakamura[2], Randy Gomez[2], Chao Li[3], Fan Yang[1]
[1] ImViA EA7535, Univ. Bourgogne Franche-Comt´e, Dijon, France
[2] Honda Research Institute Japan Co., Ltd., 8-1 Honcho, Wako-shi, Saitama, Japan
[3] State Key Laboratory of Acoustics, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China
e-mail: *Peixi_Li@etu.u-bourgogne.fr

Abstract—**Remote photoplethysmography (RPPG) is a technique to measure the physiological signs (such as Heart Rate) remotely. Compared with the conventional photoplethysmography (PPG), the RPPG provides long term monitoring without contact equipment and makes the application possible and comfortable in daily life. Heart Rate Variability (HRV) is a medical index which is calculated as the interval of the heartbeats. It has been proved that the HRV can be used as a biomarker for the autonomic nervous system (ANS) so that the features of HRV are possible to be utilized to detect the human stress and other emotion states [1] [2]. To make the RPPG method works more effectively in human emotion states detection, the measurement of HRV should be improved. In this paper, we propose to use the slope sum function (SSF) to improve the interbeat interval (IBI) detection. The performance of this new method was evaluated with the HRV features which can be used for emotion detection. The results showed that this new method has improved the accuracy of the HRV measurement in both frequency domain and time domain.**

Keywords-remote photoplethysmography (RPPG); heart rate variability (hrv); slope sum function (SSF)

I. Introduction

Photoplethysmography (PPG) is a widely used technique in medical applications to measure the physiological signs such as heart rate (HR). It has two basic components. One is a specified light source and the other one is the light sensor. By measuring the variation of the light transmission or reflection via human body parts such as the fingers, the variation of blood concentration can be measured so that the blood volume pulse (BVP) can be detected. The PPG is a low-cost device, but it cannot work when long term monitoring and non-contact conditions are required. So the video-based remote PPG (RPPG) method has been recently emerged [3]. Instead of using the specified light source and photoelectric sensor, the natural light and the low-cost web camera are used in the RPPG method. With similar principle, the blood concentration variation in the face are measured with RPPG, thus the BVP can be calculated.

With the BVP signal and some pre-processing methods, the HR and Heart Rate Variability (HRV) can be calculated based on peaks of the BVP. HRV has been investigated by the medical researchers as a biomaker for the autonomic nervous system (ANS) [1] [2]. Several features of the HRV can be used to detect the human emotion states. For instance, the low frequency (LF) feature of HRV reflects both sympathetic and parasympathetic activity of the ANS and the high frequency (HF) feature reflects the parasympathetic branch of the ANS. In some RPPG research, the power spectral density ratio LF/HF was used to detect the cognitive stress [4]. And some other HRV features obtained by RPPG in time domain such as Standard Deviation (STD) and Root Mean Square of the Successive Differences (RMSSD) are possible to be used to determine different human emotion states as well [5] [6].

The HRV is sensitive and not always precisely measured with this remote method due to the noise caused by the sensors and possibly the human tissues. Therefore, improving the precision of the HRV detection is a significant task.

Slope Sum Function (SSF) has been used in the contact equipment for HR measurement such as Electrocardiography (ECG) [7] [8] [9]. But it has not been investigated in the RPPG framework. The HRV measured by RPPG was also investigated. But the SSF and the precision of relevant HRV features for emotion detection have not been studied in this framework [10]. In this paper, we proposed to use the SSF to improve the remote measurement of HRV. The performance of this new method was compared with the conventional MATLAB peak detection method using several HRV features.

Section 2 explains the method. Section 3 describes the experiments. Section 4 is the conclusion.

II. Method

A. Slope Sum Function (SSF)

Most of the BVP signals are periodic or quasi-periodic signals where rising phases and falling phases appear alternately and sequentially. To reduce the noise of the BVP signal which may affect the detection of systolic peaks, it is possible to consider enhancing the rising trend of the signal and reducing the downward trend of it, so that the signal may be more clear for the conventional peak detection method. And this is the purpose of using the SSF, which is expressed as:

978-1-7281-3660-8/19/$31.00 ©2019 IEEE

$$S[i] = \sum_{i-w}^{i} \Delta U_k \ \text{ and } \ i = w, w+1, w+2, ..., N \qquad (1)$$

And ΔU_k is expressed as:

$$\Delta U_k = \begin{cases} \Delta Y_k & \text{if} \quad \Delta Y_k > 0 \\ 0 & \text{if} \quad \Delta Y_k \leq 0 \end{cases} \qquad (2)$$

Formulas 1 and 2 show the calculation of the new signal S transformed from original signal by SSF, where i represents the time stamp of the signal and w is the window size. In formula 2, $\Delta Y_k = Y_k - Y_{k-1}$, where Y is the original signal. To effectively transform the signal with maximum upward trend, the window size w should be approximately the same with the length of the rising phase of the original signal. In our case, the window size was carefully selected as the time of 3 samples for the frame rate of 30 fps.

Fig.1 shows an example of the BVP signal processed by SSF. The blue curve is a well filtered BVP signal obtained by RPPG. The red curve is the new signal transformed by SSF. It can be seen that after using the SSF, the upslope part is enhanced and it becomes cleaner and much more straightforward for peak detection with fewer potential false peaks.

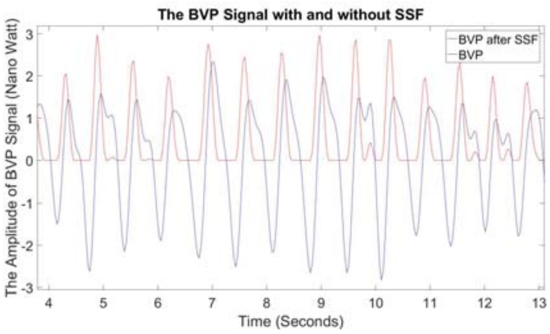

Figure 1. An example of the original BVP signal (blue) and the BVP signal after SSF (red)

B. System FrameWork

The system framework of video-based HRV measurement is presented in Fig. 2. To get the temporal RGB traces, firstly the Viola-Jones face detector and the Kanade-Lucas-Tomasi tracker provided by MATLAB toolbox were used for each video frame. Then, the skin classification method proposed by Conaire et al. [11] was used to select the skin pixels and discard the non-skin pixels. We used this method because it performed very well in the RPPG based HR measurement and worked very fast according to our previous research [12]. The results of these two approaches was the segmentation of region of interest (ROI) for the video frames. Then, the pixels inside ROI are spatially averaging to obtain the RGB time series. These temporal traces are then pre-processed by zero-mean and unit variance normalization and band-pass filtered with Butterworth filter. The frequency band was 0.7 to 4 HZ which represented the heart rate of 42 beats per minute (bpm) to 240 bpm. The rPPG signal is then extracted using the chrominance-based method [13]. This method applies simple linear combinations of RGB channels and obtains very interesting performance with low computational complexity. Let $y^c(t)$ be the RGB time series obtained after pre-processing, with where $c \in \{R, G, B\}$ is the color channel, this method projects RGB values onto two orthogonal chrominance vectors X and Y:

$$\begin{aligned} X(t) &= 3y^R(t) - 2y^G(t), \\ Y(t) &= 1.5y^R(t) + y^G(t) - 1.5y^B(t). \end{aligned} \qquad (3)$$

The BVP signal P is finally calculated with $P(t) = X(t) - \alpha Y(t)$ where $\alpha = \sigma(X)/\sigma(Y)$. Because X and Y are two orthogonal chrominance signals, PPG-induced variations will likely be different in X and Y, while motion affects both chrominance signals identically.

After this, we used SSF to convert the BVP signal to a new signal. Then the MATLAB peak detection was applied on this new signal and the interbeat intervals (IBI) were calculated based on the time stamps of the peaks. Then the HRV was calculated by interpolating IBI with the video frame rate.

III. Experiments

The performance of the new method was compared with the conventional BVP peak detection method in the framework of remote HRV measurement. In this section, first of all the video dataset and experimental set up is described in details. Secondly, the HRV features used as the evaluation metrics are introduced and explained. The results are finally presented and discussed in details.

A. Experiment Set Up

For the experiments, we used our in-house video datasets *UBFC-rPPG* [14] which have 46 videos. The volunteers were asked to sit at the distance of one meter from the web camera and were asked to play a time sensitive video game. Ambient light was used in the experiment to create diffuse reflections. The experimental set up is shown in Fig. 3.

A C++ program was used to record the videos and synchronize the videos with the signals of the contact PPG sensor which was used as the ground truth. The web camera is a Logitech C920. The resolution of the video frame is 640x480 and the frame rate is 30/second. The format is 8-bit uncompressed RGB. The contact PPG sensor is a CMS50E transmissive pulse oximeter. The experimental dataset with the ground truth can be downloaded from our project webpage [1]. Some sample images are shown in Fig. 4.

[1] https://sites.google.com/view/ybenezeth/ubfcrppg

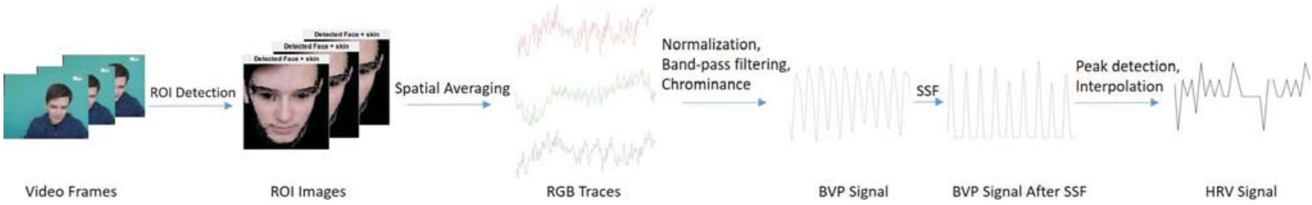

Figure 2. System Framework

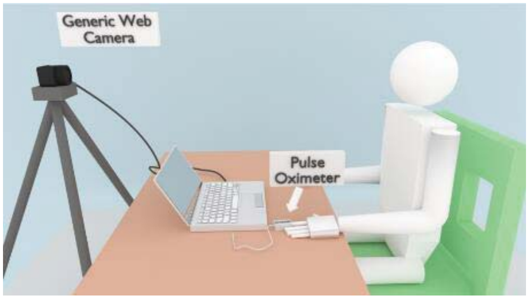

Figure 3. Experimental Setup

Figure 4. Experimental sample images from the *UBFC-RPPG* database.

B. Evaluation Metrics

As explained before, frequency features such as HF and LF features can be used to detect the human stress [4]. Temporal features such as STD, RMSSD and IBI can potentially determine the emotion states as well [4] [5] [6]. So we used the errors of the HF, LF, STD, RMSSD and IBI to assess the accuracy of HRV measurement. The ground truth is the HRV measured by contact PPG. These are the metrics:

- The relative HF error ($\%HF_{er}$): the HF is calculated as the area of power spectral density between 0.15-0.4 HZ and normalized by the sum of (HF + LF). The relative error was defined as:

$$\%HF_{er} = |\frac{HF_{RPPG}}{HF_{PPG}} - 1| \times 100\% \tag{4}$$

- The relative LF error ($\%LF_{er}$): the LF is calculated as the area of spectral power density between 0.04-0.15 HZ and normalized by the sum of (HF + LF). The relative error was defined as:

$$\%LF_{er} = |\frac{LF_{RPPG}}{LF_{PPG}} - 1| \times 100\% \tag{5}$$

- The absolute error and relative error of IBI between RPPG and the ground truth, denoted as $AbIBI_{er}$ and $\%IBI_{er}$.
- The absolute error of STD and RMSSD, denoted as $AbSTD_{er}$ and $AbRMSSD_{er}$. The RMSSD was defined as:

$$RMSSD = \sqrt{\frac{1}{N-1}(\sum_{i=1}^{N-1}(IBI_{(i+1)} - IBI_i)^2)} \tag{6}$$

Similarly, we calculated the STD and RMSSD of the HRV measured by RPPG and the ground truth and then obtained the absolute error.

The results were the average values of the metrics of the entire datasets.

C. Results and Discussion

Fig. 5 shows a comparative example of the peak detection for the BVP signal with and without SSF. It can be seen that there was a slight phase shift between the RPPG and PPG (ground truth) signals which was possibly caused by the difference of the time that blood pulse reached the finger and face along with the sensor latency. But this does not affect the HRV calculation because it is the interval values (IBI) that build up the HRV signal. According to the figure, the IBI stayed almost the same with the ground truth when SSF was applied. And in the figure without SSF, the IBI between the fourth and fifth peaks is underestimated due to the problematic shape of the signal. Table I shows the error of the IBI measurement. Both the absolute error and the relative error of the SSF method were lower than the previous method without SSF. It is worth mentioning that the relative error of HR measurement in this study was about 2% in both methods which was much lower than the IBI error. It

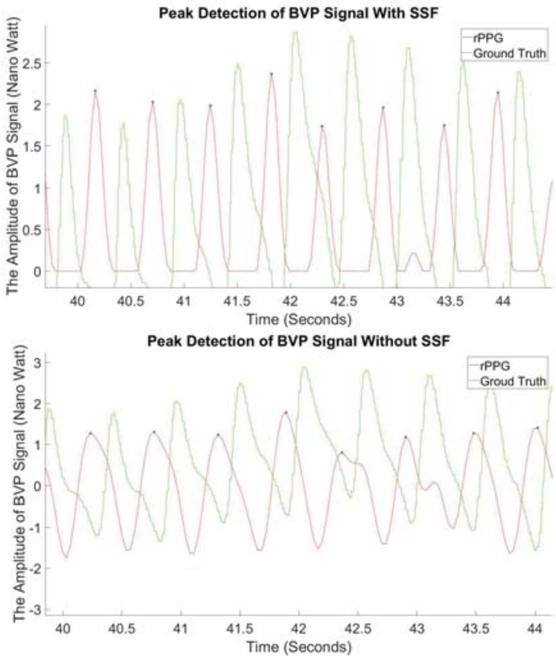

Figure 5. A comparative example of peak detection with and without SSF. Red curve is RPPG and green curve is the ground truth.

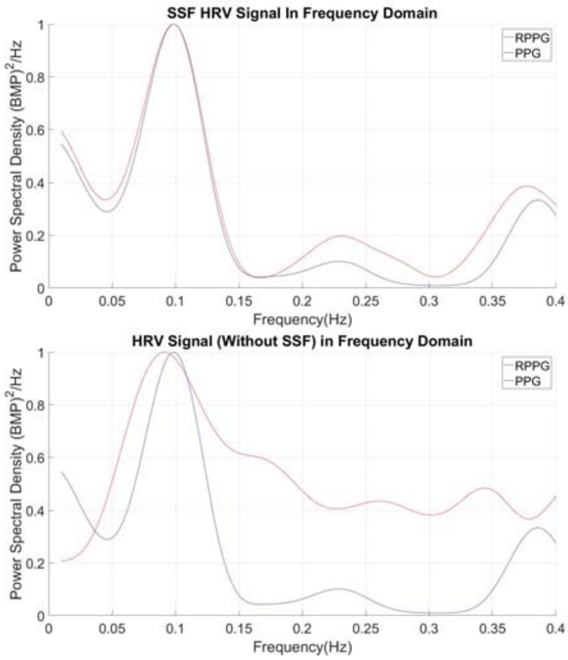

Figure 6. RPPG HRV signal (red curve) compared with the ground truth (blue curve) in frequency domain

means the HRV measurement is indeed a more challenging task for RPPG research.

Table I THE AVERAGE ERRORS OF IBI

Methods	$AbIBI_{er}(s)$	$\%IBI_{er}$
SSF	0.051	7.47 %
No SSF	0.070	10.6 %

Fig. 6 shows an example of HRV signal in frequency domain. It can be seen that the HRV calculated with SSF matched the curve of the ground truth better than the previous method in frequency domain. And the area of LF part (0.04-0.15 HZ) of the ground truth is bigger than the HF part (0.15-0.4 HZ). This is not surprising because the volunteers in the experiment were playing a time sensitive game which might increase their stress expressed by the LF part of HRV [4]. The relative errors of the frequency features are shown in the Table II. The relative error of LF was much lower in the case of SSF. And it was slightly higher in HF, but this is not significant because the LF reflects the emotion conditions in this specified video datasets. On the other hand, the relative error of these frequency features are still bigger than ideal expectation, and this may be the reason that the remote emotion detection by RPPG has not reached the accuracy of more than 90% [4].

Table III shows the errors of the time domain features, STD and RMSSD. Similarly, the SSF performed much better than the original signal.

Table II THE AVERAGE RELATIVE ERRORS OF HRV FEATURES IN FREQUENCY DOMAIN

Methods	$\%LF_{er}$	$\%HF_{er}$
SSF	23.8%	38.6 %
No SSF	32.8 %	34.2 %

Table III THE AVERAGE ERRORS OF HRV FEATURES IN TIME DOMAIN

Methods	$AbSTD_{er}(s)$	$AbRMSSD_{er}(s)$
SSF	0.025	0.047
No SSF	0.037	0.093

IV. CONCLUSION AND FUTURE WORK

The HRV measured by RPPG has great potential for remote emotion detection and the accuracy of HRV estimation should reach an ideal level. In this paper we achieved two main tasks. Firstly we used SSF in our RPPG framework to improve HRV measurement by enhancing the upslope trend and reducing the downward phase of the BVP signal. Secondly we investigated the errors of HRV features which are possibly used in emotion detection. The experiment was done with a low-cost camera and a contact PPG as the ground truth. 46 videos were collected to test our framework. The results showed that our new method performed better in almost all of the features than the previous method. However, as the results showed, the relative errors of some HRV features were still quite big. This method and relevant features have not been tested in the real emotion detection scenario yet. So in our future work, we will firstly work further to improve the RPPG framework to improve the HRV measurement. Secondly we will test our methods and relevant features in emotion detection experiments.

REFERENCES

[1] Evans, S., Seidman, L. C., Tsao, J. C., Lung, K. C., Zeltzer, L. K., Naliboff, B. D. (2013). Heart rate variability as a biomarker for autonomic nervous system response differences between children with chronic pain and healthy control children. Journal of pain research, 6, 449.

[2] Quintana, D. S., Guastella, A. J., Outhred, T., Hickie, I. B., Kemp, A. H. (2012). Heart rate variability is associated with emotion recognition: direct evidence for a relationship between the autonomic nervous system and social cognition. International Journal of Psychophysiology, 86(2), 168-172.

[3] Poh, M. Z., McDuff, D. J., Picard, R. W. (2011). Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE transactions on biomedical engineering, 58(1), 7-11.

[4] McDuff, D., Gontarek, S., Picard, R. (2014, August). Remote measurement of cognitive stress via heart rate variability. In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 2957-2960). IEEE.

[5] Rakshit, R., Reddy, V. R., Deshpande, P. (2016, November). Emotion detection and recognition using HRV features derived from photoplethysmogram signals. In Proceedings of the 2nd workshop on Emotion Representations and Modelling for Companion Systems (p. 2). ACM.

[6] Delaney, J. P. A., Brodie, D. A. (2000). Effects of short-term psychological stress on the time and frequency domains of heart-rate variability. Perceptual and motor skills, 91(2), 515-524.

[7] Elgendi, M., Norton, I., Brearley, M., Abbott, D., Schuurmans, D. (2013). Systolic peak detection in acceleration photoplethysmograms measured from emergency responders in tropical conditions. PLoS One, 8(10), e76585.

[8] Rankawat, S. A., Rankawat, M., Dubey, R. (2015, May). ECG artifacts detection in Noncardiovascular signals using Slope Sum Function and Teager Kaiser energy. In 2015 International Conference on BioSignal Analysis, Processing and Systems (ICBAPS) (pp. 6-10). IEEE.

[9] Jang, D. G., Park, S., Hahn, M., Park, S. H. (2014). A real-time pulse peak detection algorithm for the photoplethysmogram. Int. J. Electron. Electr. Engin, 2, 45-49.

[10] Huang, R. Y., Dung, L. R. (2016). Measurement of heart rate variability using off-the-shelf smart phones. Biomedical engineering online, 15(1), 11.

[11] Conaire, C. O., O'Connor, N. E., Smeaton, A. F. (2007, June). Detector adaptation by maximising agreement between independent data sources. In 2007 IEEE Conference on Computer Vision and Pattern Recognition (pp. 1-6). IEEE.

[12] Li, P., Benezeth, Y., Nakamura, K., Gomez, R., Li, C., Yang, F. (2018, October). Comparison of region of interest segmentation methods for video-based heart rate measurements. In 2018 IEEE 18th International Conference on Bioinformatics and Bioengineering (BIBE) (pp. 143-146). IEEE.

[13] De Haan, G., Jeanne, V. (2013). Robust pulse rate from chrominance-based rPPG. IEEE Transactions on Biomedical Engineering, 60(10), 2878-2886.

[14] Bobbia, S., Macwan, R., Benezeth, Y., Mansouri, A., Dubois, J. (2017). Unsupervised skin tissue segmentation for remote photoplethysmography. Pattern Recognition Letters.