Effect of Time Window Size for Converting Frequency Domain in Real-Time Remote Photoplethysmography Extraction

Yu Jin Shin[1], Woo Jung Han[1], Kun Ha Suh[2], and Eui Chul Lee[1](✉)

[1] Department of Human-Centered Artificial Intelligence, Sangmyung University, Seoul, South Korea
eclee@smu.ac.kr
[2] R&D Team, Zena Inc., Seoul 04782, Korea

Abstract. Remote-photoplethysmography (rPPG) is an attractive technology that can measure vital signs at a distance without contact. Previous remote-photoplethysmography studies focused mainly on eliminating the artifact such as motion but finding the optimal setup or hyperparameters are also an important factor influencing the performance. As one of them, window size is the length of the signal used to calculate the vital signs once in a spectral method and has not been analyzed in detail in previous works. In general, the use of a long window size increases the re-liability of the estimations, but it cannot reflect continuously changing physiological responses of human. Also, using too short window size increases uncertainty. In this paper, we compare and analyze the pulse rate estimation results according to window sizes from short to long using CHROM, which is one of the popular rPPG algorithms. Results on the PURE dataset showed that the longer the window size, the higher the SNR and the lower the RMSE. At a window size of about 4 s (120 frames), the SNR was switched from negative to positive and an acceptable error rate (RMSE $<$ 5) was observed.

Keywords: Remote PPG · Cardiac pulse · Window size · Pulse rate · Physiological signal

1 Introduction

Remote-photoplethysmography (rPPG) is a promising optical technology that can extract physiological signals to obtain vital information such as a pulse rate in contactless monitoring. The subtle color changes on the skin surface are synchronized with cardiac activities. The fundamental principle of rPPG is to measure blood flow changes by detecting light reflected from the skin tissues by a camera sensor. Unlike conventional contact-based PPG approaches that require additional devices, rPPG has the advantage of being able to remotely measure vital signs using a camera, which is ubiquitous today, under ambient lighting conditions. The rPPG is an emerging technology and can be applied to various fields such as sleep monitoring, home healthcare, elderly care, fitness, entertainment, and forensic science.

J.-H. Kim et al. (Eds.): IHCI 2021, LNCS 13184, pp. 145–149, 2022.
https://doi.org/10.1007/978-3-030-98404-5_14

Studies on the easiness of rPPG extraction for each channel of digital images, noise-resistant component analysis methods, optical/mathematical models of light sources and skin, and performance degradation due to image compression have been conducted [1–5].

On the other hand, most rPPG studies have only estimated a single pulse rate from a long-length input video. However, such measurements cannot reflect a person's active physiological responses which change continuously. Moreover, in practical use, subjects will not want to wait that long for measurements. Therefore, it is necessary to consider short-term measurements for rPPG-based vital sign estimation. The window size refers to the length of the signal for estimating a single estimation and has not been covered in detail in previous studies. In this paper, we aim to contribute to confirming the impact of window size in rPPG-based pulse rate measurement and finding the optimal window size. For rPPG extraction, we used the CHROM method as the default rPPG algorithm from facial skin.

2 Methods

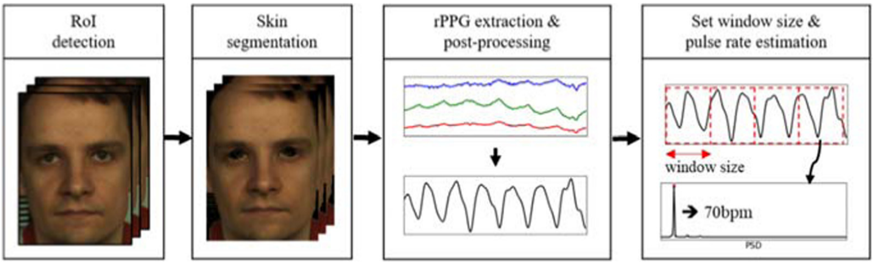

Fig. 1. Overview of the proposed rPPG-based pulse-rate estimation approach

Figure 1 shows the overview of the proposed rPPG-based pulse-rate estimation approach. First, the frontal face is detected by the face detector from OpenCV DNN module [6] at the first frame. Then Kernelized correlation filter (KCF) tracker [7] is used to track the detected face in subsequent frames. This step enables stable facial localization while minimizing the background pixels. Next, since the background has no pulsatile information, a statistical filtering method based on the YCbCr color space [8] is applied to filter out background pixels from the facial rectangle. The range of the skin was determined in a heuristic method. Then the CHROM method [9] is used for rPPG extraction. This algorithm is robust to the subject's motion and is therefore frequently adopted for rPPG extraction. Since the signal still contains noise components, to further improve the signal quality, we apply two post-processing steps. As the first post-processing step, we re-move the breath-like trend from the signal and apply detrending to obtain zero-centered normalized signals. In addition, a Butterworth bandpass filtering [10] with cut-off frequencies of 42 bpm and 240 bpm is applied, which removes components irrelevant to cardiac activities. Finally, to calculate pulse rate from the signal, a spectral method is used. It is considered for rPPG waveforms that are not relatively clean. At this time, to mitigate spectral leakage, we apply the Hann window before the Fourier transform.

The window size determines the interval of the signal to participate in the calculation of a single pulse-rate. In general, the use of a long window size increases the re-liability of the estimation in rPPG-based vital signs estimation, but it cannot reflect continuously changing physiological responses, and if it is too short, an incorrect estimation is obtained. In our experiment, a total of 15 window size candidates (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, and 60 s) were considered. In the PSD-based pulse rate estimation, a low frequency resolution is obtained if short-time signal is provided. When calculating the pulse rate from a signal with the length of 10-s, it means that the frequency resolution is 0.1 Hz in which the interval of each bin indicates 6 bpm. Therefore, to increase the frequency resolution (i.e., calculated with 1 bpm precision), some tricks could be introduced. The zero-padding technique is one of the ways to increase the frequency resolution and was used in our experiment. More specifically, we applied windowing the signal before applying zero-padding. Consecutive pulse rates were obtained by applying sliding window at 1 frame intervals from the signal then the statistical performance was evaluated.

3 Results

The algorithms were validated on PURE dataset [11]. This dataset comprises 10 participants performing different, controlled head motions in front of a camera. It was recorded in 6 different setups such as steady, talking, slow translation, fast translation, small rotation, small rotation, and medium rotation resulting in a total number of 60 sequences of 1 min each. The videos were captured with an eco274CVGE camera by SVS-Vistek GmbH at a frame rate of 30 Hz with a cropped resolution of 640 $\times$ 480 pixels and a 4.8 mm lens. Reference data have been captured in parallel using a finger clip pulse oximeter (pulox CMS50E) that delivers pulse rate wave and SpO2 readings with a sampling rate of 60 Hz. The test subjects were placed in front of the camera with an average distance of 1.1 m. Lighting condition was daylight through a large window frontal to the face with clouds changing illumination conditions slightly over time.

Table 1. Result of SNR and RMSE of each window length

Window size (sec)	1	2	3	4	5	6	7	8	9	10	20	30	40	50	60
SNR	−8.92	−5.57	−2.88	−1.00	0.60	4.72	2.93	3.80	4.50	5.01	6.05	6.17	6.26	6.24	5.91
RMSE	22.45	13.19	6.94	4.49	3.16	2.69	2.36	2.23	2.14	1.94	1.60	1.52	1.50	1.45	1.13

We measured SNR values by setting them differently to a total of 15 window sizes. Figure 2 and Table 1 show the SNR and RMSE comparison as a function of window length, respectively. As the window size increases, higher SNR and lower RMSE were obtained. At the window size of about 4 s (120 frames), the SNR is switched from negative to positive and a low error rate (RMSE < 5) was observed. Additionally, it was confirmed that using a window of sufficient length has the effect of reducing the estimation error in moving subject scenarios.

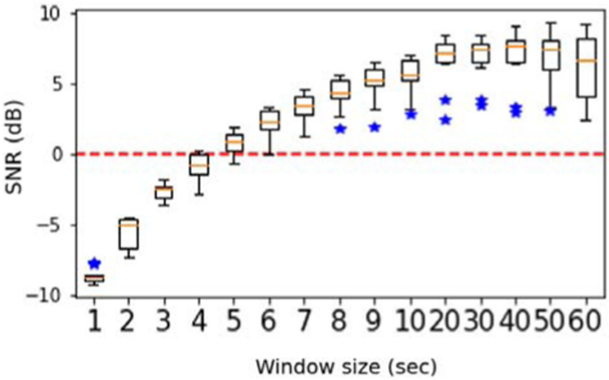

Fig. 2. The SNR comparison as a function of window length on subjects

Figure 2 shows the SNR of the length of the window is a boxplot comparison, the orange line is the mean value of the SNR, the blue star is the outlier, and the dotted line is the reference line when the SNR is zero. A comparison of window sizes by setup according to SNR shows that overall, the larger the window size, the higher the SNR. The larger the window size, the more cardiac information is included. The window size and SNR are proportional because we extract signals based on the most dominant information. As expected, the small window size shows low accuracy due to insufficient information to analyze.

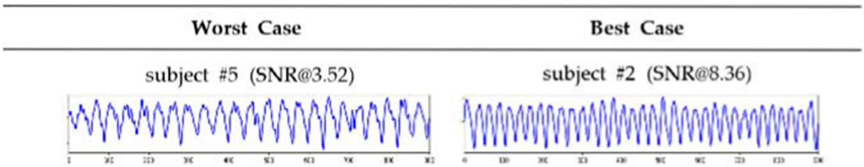

Fig. 3. Best and Worst cases of CHROM waveform when window size is 30

Figure 3 shows the CHROM waveforms for the worst and best cases when window size is 30. The video used is a signal waveform with a total of 900 frames and is 30 s long at 30 fps. Best and worst are extracted based on SNR values, and the values of SNR are written above each waveform. As can be seen from the figure, the best cases show that waveform results are more regular and stable.

4 Conclusion

In this paper, we analyzed how window size affects pulse rate estimation. To extract the rPPG signal and find the most optimal window size for estimating heart rate, we analyzed the accuracy of the estimated pulse-rate according to window size. The pulse-rate was calculated by applying a spectral-based estimation method that multi-plied the maximum power peak by 60 after the FFT was applied. We extracted the pulse signal with varying window-size for the PURE dataset. Each pulse-rate according to different window sizes was compared with SNR and RMSE. Experiments show that the smaller the window size, the less information, the lower the performance, and the longer the window size, the better overall performance. However, it has been observed that long window sizes are not always optimal, and sometimes performance gradually degrades when exceeding a certain size. Consequently, the window size affects the rPPG signal extraction performance, so it is necessary to find the optimal window size according to the application domains.

Acknowledgement. This work was supported by the NRF (National Research Foundation) of Korea funded by the Korea government (Ministry of Science and ICT) (NRF-2019R1A2C4070681).

References

1. de Hann, G., Jeanne, V.: Robust pulse rate from chrominance-based rppg. IEEE Trans. Biomed. Eng. **60**, 2878–2886 (2013)
2. Verkruysse, W., Svaasand, L.O., Nelson, J.S.: Remote plethysmographic imaging using ambient light. Opt. Express **16**(26), 21434–21445 (2008)
3. Lewandowska, M., Rumiński, J., Kocejko, T.: Measuring pulse rate with a webcam – a non-contact method for evaluating cardiac activity. In: IEEE Federated Conference on Computer Science and Information Systems, pp. 405–410 (2011)
4. Ming, P., Daniel, J., McDuff, R., Picard, W.: Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Opt. Express **18**(10), 10762–10774 (2010)
5. Wang, W., Brinker, A.C., Stuijk, S., de Hann, G.: Algorithmic principles of remote PPG. IEEE Trans. Biomed. Eng. **64**, 1479–1491 (2017)
6. OpenCV deep learning module samples 'OpenCV dnn face detector'. https://github.com/opencv/opencv/tree/master/samples/dnn. Accessed 01 Oct 2021
7. Henriques, J.F., et al.: High-speed tracking with kernelized correlation filters. IEEE Trans. Pattern Anal. Mach. Intell. **37**, 583–596 (2014)
8. Phung, S.L., Bouzerdoum, A., Chai, D.: A novel skin color model in YCbCr color space and its application to human face detection. In: Proceedings. International Conference on Image Processing, vol. 1. IEEE (2002)
9. De Haan, G., Jeanne, V.: Robust pulse rate from chrominance-based rPPG. IEEE Trans. Biomed. Eng. **60**(10), 2878–2886 (2013)
10. Chatterjee, A., Roy, U.K.: PPG based heart rate algorithm improvement with Butterworth IIR filter and Savitzky-Golay FIR filter. In: 2018 2nd International Conference on Electronics, Materials Engineering and Nano-Technology (IEMENTech). IEEE (2018)
11. Stricker, R., Müller, S., Gross, H.M.: Non-contact video-based pulse rate measurement on a mobile service robot. In: IEEE International Symposium on Robot and Human Interactive Communication, pp. 1056–1062 (2014)