

# Non-contact Video Based Estimation of Heart Rate Variability Spectrogram from Hemoglobin Composition

MUNENORI FUKUNISHI\*<sup>1</sup>, KOUKI KURITA\*<sup>1</sup>, SHOJI YAMAMOTO\*<sup>2</sup> AND  
NORIMICHI TSUMURA\*<sup>1</sup>

<sup>1</sup>Graduate School of Advanced Integration Science, Chiba University, 1-33 Yayoi-cho, Inage-Ku, Chiba 263-8522, JAPAN

<sup>2</sup>Tokyo Metropolitan College of Industrial Technology, 8-17-1 Minami-Senju, Arakawa-ku, Tokyo 116-0003, JAPAN

**Abstract:** Non-contact HR measurement is getting active research area. Recently remote photoplethysmography (rPPG) measurement based on simple skin optics model has been proposed and it showed the effectiveness. In this paper, we propose an accurate remote observation of the heart rate (HR) and heart rate variability (HRV) based on hemoglobin component estimation which is based on detail skin optics model. We perform experiments to measure subjects at rest and under cognitive stress with the proposed method putting a polarized filter in front of camera to evaluate the principal of the framework. From the results of the experiments, the proposed method shows a high correlation with the electrocardiograph (ECG) which is assumed as the ground truth. We also evaluate the proposed method without putting any polarized filter and confirm the usefulness for the remote observation of HRV which require accurate detection of HR.

**Keywords:** heart rate, heart rate variability, non-contact measurement

## 1 INTRODUCTION

Non-contact video based measurement of physiological status has great potential for healthcare applications, medical diagnosis, and affective computing. With recent advances in mobile technology, non-contact heart rate (HR) detection has become an active area of research. In particular, various techniques have been proposed for the measurement of the blood volume pulse (BVP). Verkruysse *et al.* [1] demonstrated the measurement of BVP under ambient light using the G channel of movies captured by a consumer camera. Poh *et al.* [2] also developed a remote BVP measurement technique using a low-cost webcam, based on blind source separation. Haan *et al.* [7] proposed remote photoplethysmography (rPPG) measurement which is based on simple skin optics model and showed the effectiveness recently.

Heart rate variability spectrograms (HRVS) are useful for non-invasive monitoring of the autonomic nervous system, which controls involuntary body functions, such as breathing, blood pressure, and heartbeat. The low-frequency (LF) power in HRVS (0.05-0.15 Hz) is widely known as one of the most reliable indicators of sympathetic activity since the power increases under cognitive stress [3]. The high frequency (HF) power in HRVS (0.15-0.40 Hz) is affected by breathing and is related to parasympathetic activity. McDuff *et al.* [4] developed a remote HRVS measurement technique using a special sensor with five color channels (16 bits/channel): red, green, blue, cyan, and

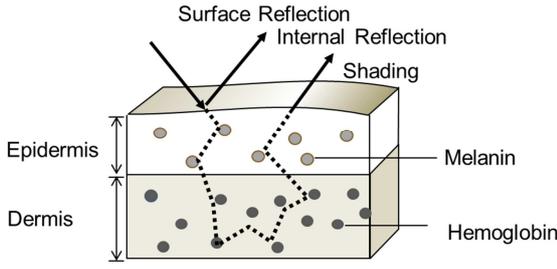
orange (RGBCO) [5, 6]. They showed the effectiveness of the remote HRVS measurement.

Estimation methods for skin components have been proposed based on skin optics. Tsumura *et al.* [8] proposed a method for estimating hemoglobin, melanin, and shading components from a skin image captured by a standard RGB camera. The method based on the detail skin optics model. However, the skin component separation technique has not been applied to non-contact HR and HRV detection until now.

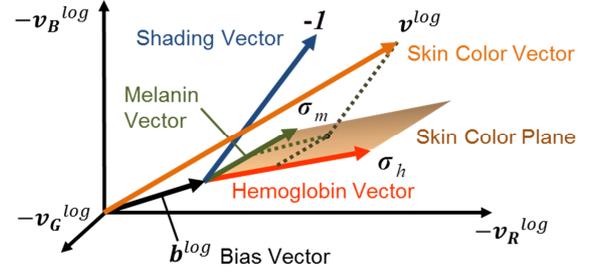
In this paper, therefore, we propose a method to obtain accurate remote observations of BVP and HRVS using a standard RGB camera based on the extraction method for the hemoglobin component. We evaluated the proposed method with polarized filter in front of camera and without polarized filter respectively. The rest of this paper is organized as follows. In Section 2, we outline the method for extracting hemoglobin information from skin images. In Section 3, we adapt the method from Section 2 to measure BVP and HRVS. In Section 4, we describe the experimental setup and show the results of BVP detection with the proposed method, and compare them with ECG data as the ground truth and with the conventional method using a five-band camera. We also show experimental results from stress monitoring. In Section 5, we present our conclusions.

## 2 Extraction of Hemoglobin Information from a Skin Color Image

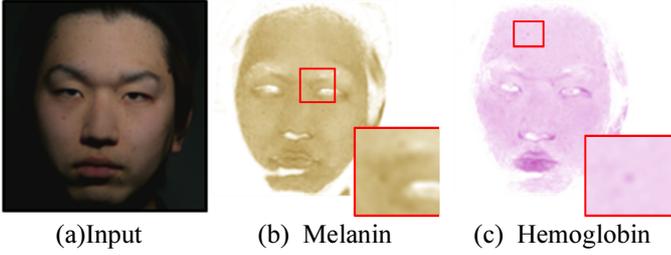
Figure 1 shows the skin model for the extraction of hemoglobin component. We use two-layered skin model



**Fig. 1** Skin Model for the Estimation of Hemoglobin, Melanin, Shading



**Fig. 2** Skin Color Vector and Melanin, Hemoglobin and Shading Vectors



**Fig. 3** Estimation Results of Melanin, Hemoglobin and Shading Component

composed of epidermis and dermis. We simplify the model assuming that epidermis only has chromophores of melanin and dermis only has chromophores of hemoglobin. The reflected light from the surface of skin consist of surface reflection and internal reflection. The modified Lambert-Beer law [9] is used to model the behavior of internal reflection. The spectral radiance  $L(x, y, \lambda)$  at the position  $(x, y)$  on the surface is described by

$$L(x, y, \lambda) = e^{-\rho_m(x,y)\sigma_m(\lambda)l_e(\lambda) - \rho_h(x,y)\sigma_h(\lambda)l_d(\lambda)} E(x, y, \lambda), \quad (1)$$

where  $\lambda$  is the wavelength,  $E(x, y, \lambda)$  denotes the spectral irradiance of incident light at point  $(x, y)$ , and  $\rho_m(x, y)$ ,  $\rho_h(x, y)$ , and  $\sigma_m(\lambda)$ ,  $\sigma_h(\lambda)$  are the densities of the chromophores and the spectral cross-sections of the melanin and hemoglobin,  $l_s(\lambda)$ ,  $l_d(\lambda)$  denote the mean path lengths of photons in the epidermis and dermis layers, respectively. Polarization filters are put in front of the illumination and camera in the position as crossed nicols so that we can ignore surface reflection. Camera signal  $v_i(x, y)$ ,  $i = R, G, B$  can be modeled as

$$\begin{aligned} v_i(x, y) &= k \int L(x, y, \lambda) s_i(\lambda) d\lambda \\ &= k \int e^{-\rho_m(x,y)\sigma_m(\lambda)l_e(\lambda) - \rho_h(x,y)\sigma_h(\lambda)l_d(\lambda)} E(x, y, \lambda) s_i(\lambda) d\lambda \\ &\quad (i = R, G, B) \end{aligned}$$

where  $s_i(\lambda)$  denotes the spectral sensitivity of a camera, and  $k$  denotes the coefficient of camera gain. Since the spectral reflectance curve of skin is smooth and roughly correlated with camera sensitivity, we can approximately assume  $s_i(\lambda) \approx \delta(\lambda - \lambda_i)$ . We assume the spectral irradiance of incident light  $E(\lambda)$  can be written as the following equation.

$$E(x, y, \lambda) = p(x, y) \bar{E}(\lambda).$$

Here, the factor  $p(x, y)$  is related to shading information and  $\bar{E}(\lambda)$  indicates that the basic color of illumination is the same at any point on the surface of the object. The camera signal can be rewritten as

$$\begin{aligned} v_i(x, y) &= k e^{-\rho_m(x,y)\sigma_m(\lambda_i)l_e(\lambda_i) - \rho_h(x,y)\sigma_h(\lambda_i)l_d(\lambda_i)} p(x, y) \bar{E}(\lambda_i) \end{aligned}$$

By taking the logarithm of both sides of Equation (4), we can derive the following equation.

$$\mathbf{v}^{log}(x, y) = -\rho_m(x, y) \boldsymbol{\sigma}_m - \rho_h(x, y) \boldsymbol{\sigma}_h + p^{log}(x, y) \mathbf{1} + \mathbf{b}^{log}$$

where

$$\mathbf{v}^{log}(x, y) = [\log(v_R(x, y)) \quad \log(v_G(x, y)) \quad \log(v_B(x, y))]^T$$

$$\boldsymbol{\sigma}_m = [\sigma_m(\lambda_R)l_e(\lambda_R) \quad \sigma_m(\lambda_G)l_e(\lambda_G) \quad \sigma_m(\lambda_B)l_e(\lambda_B)]^T$$

$$\boldsymbol{\sigma}_h = [\sigma_h(\lambda_R)l_d(\lambda_R) \quad \sigma_h(\lambda_G)l_d(\lambda_G) \quad \sigma_h(\lambda_B)l_d(\lambda_B)]^T$$

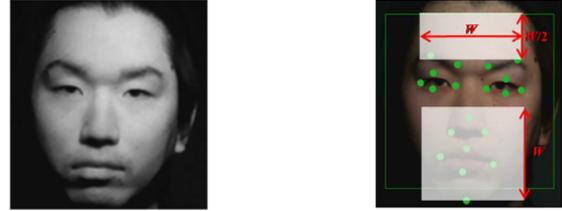
$$\mathbf{1} = [1 \quad 1 \quad 1]^T$$

$$p^{log}(x, y) = \log(p(x, y)) + \log(k)$$

$$\mathbf{b}^{log}(x, y) = [\log(E_R(\lambda_R)) \quad \log(E_G(\lambda_G)) \quad \log(E_B(\lambda_B))]^T$$

Therefore, the logarithm of the captured RGB signals  $\mathbf{v}^{log}$  can be represented by the weighted linear combination of the three vectors  $\boldsymbol{\sigma}_m$ ,  $\boldsymbol{\sigma}_h$  and  $\mathbf{1}$  with the bias vector  $\mathbf{b}^{log}$  as shown in Figure 2. We predefine a skin color plane using training data set. The logarithm of the captured RGB signals  $\mathbf{v}^{log}$  is projected onto the skin color plane along with the shading vector  $\mathbf{1}$ . From the position on the skin plane, we obtain the hemoglobin vector  $\boldsymbol{\sigma}_h$ .

Figure 3 provides the estimation results for the melanin, hemoglobin, and shading components from the input image. We can see the mole and pigmented spot in the melanin component and pimples in the hemoglobin component. The shading image provides a reasonable representation of the facial structure.



**Fig. 4.** Facial Feature Points and ROI

### 3 Extraction of Hemoglobin Information from a Skin Color Image

In this section, we describe the procedure for obtaining the BVP and HRVS based on the extracted hemoglobin information.

Figure 4 (a) shows an example of an input image for the estimation of the hemoglobin component. First, we estimate the hemoglobin component in each frame from the facial capture video using the method described in Section 2. In order to compensate for movement, we detect feature points by LEAR [10] facial landmark detector and determine the region of interest (ROI) using the feature points. We calculate the mean value of the hemoglobin component in the two ROIs shown in Figure 4 (b) in each frame.

Figure 5 shows the signal processing procedure for detecting the BVP and HRVS. In Step 1, we calculate the hemoglobin component in each frame, in the manner described above, and obtain the temporal variation of the hemoglobin component for the face. The waveform included both the BVP component, which corresponds to HR, and the fluctuation during the observation period which is caused by changes in the direction the subject is

facing. In Step 2, a detrending technique based on the smoothness prior approach [11] is applied to the waveform in order to eliminate the LF fluctuations. In Step 3, we apply detrending [14] and a band-pass filter is applied to extract the heart beat components which have a frequency between 0.75 Hz (45 beats-per-minute) and 3 Hz (180 beats-per-minute). Accordingly, we can obtain the BVP waveform. In Step 4, we detect the local peaks of the BVP waveform by comparing the signal value of the waveform with the neighborhoods. The red circles on the BVP waveform indicate peak heart beats. In Step 5, we calculate the time intervals of each peak by subtracting the peak time from the previous timing and form the waveform which shows the temporal alteration of the peak timings. Due to this process, the time intervals of each peak are sampled sparsely and unevenly. In order to apply a frequency analysis with sparsely and unevenly sampled data, we utilize Lomb-Scargle periodograms [12, 13] for the calculation of the power spectrum in Step 6. In the analysis, we used a 60-second window to calculate the power spectrum with steps sizes of 1 second, across the 120-second sampling period.

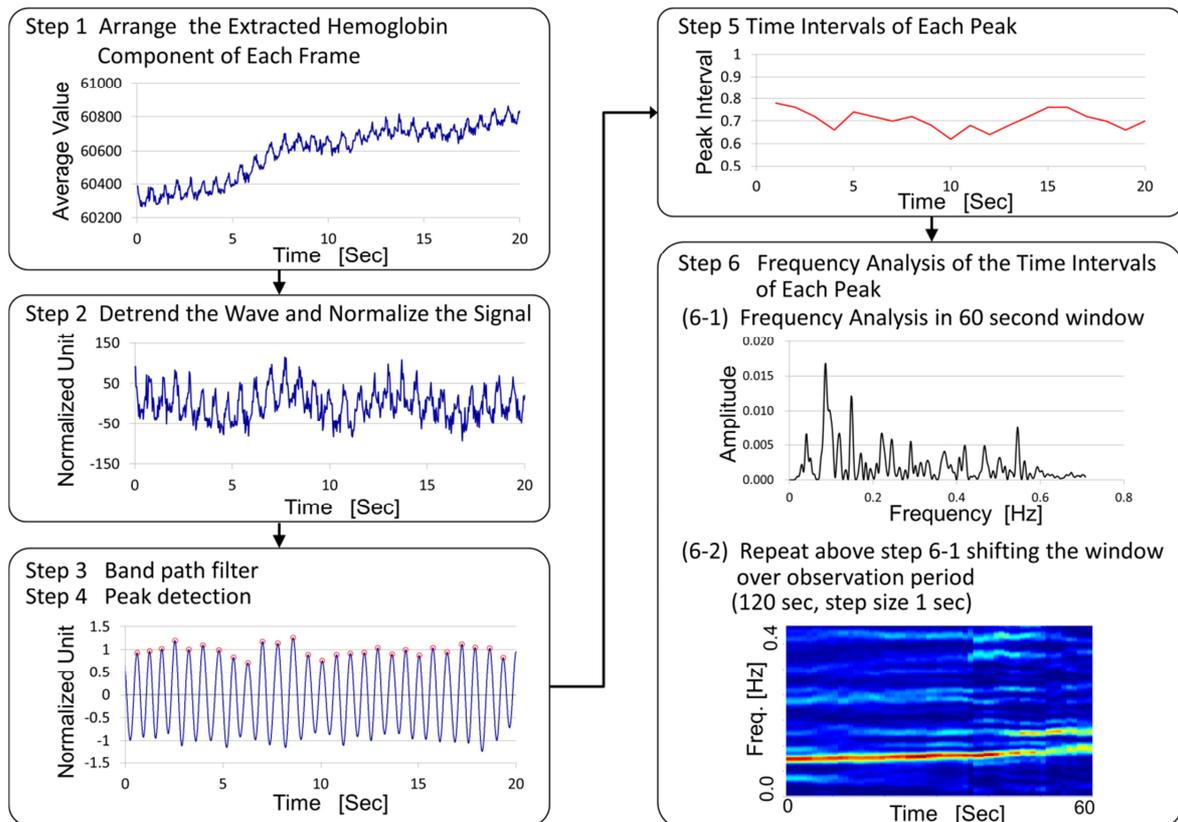
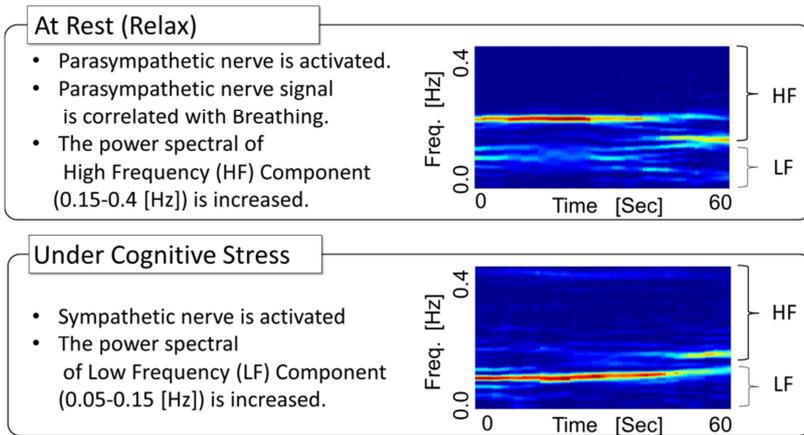
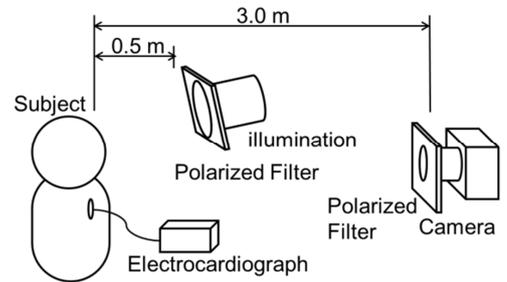


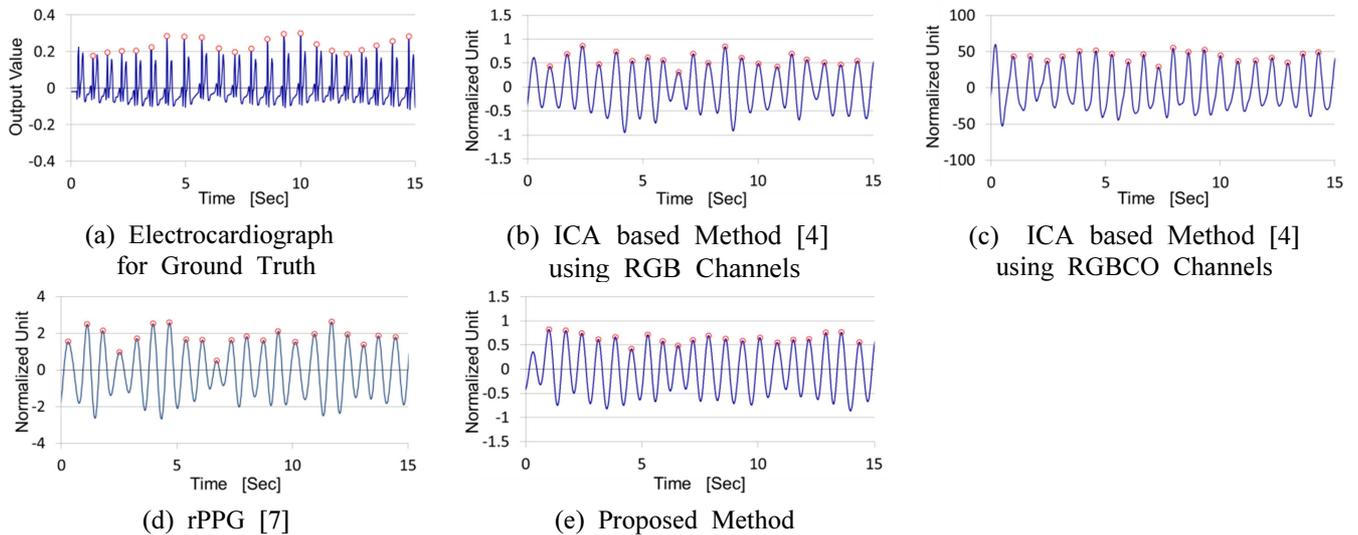
Fig. 5 Signal Processing to Detect Blood Volume Pulse and Heart Rate Variability Spectrogram



**Fig 6.** Relationship between the sympathetic nervous system and Heart Rate Variability Spectrogram



**Fig 7.** Experimental Setup



**Fig. 8** Results of Blood Volume Pulse (BVP) Detection and Electrocardiograph Data for Ground Truth

Figure 6 shows examples of the HRV spectrograms obtained by the above method and the relationship between the HRV spectrogram and the sympathetic nervous system. At rest, the parasympathetic nerve is activated. The parasympathetic nerve is influenced by the activity of respiratory sinus arrhythmia (RSA). Since usual breathing rate is between 10 and 25 inhalations per minute, the HF (0.15-0.4 Hz) power spectrum is increased at rest. On the other hand, under cognitive stress, the sympathetic nerve is activated. The sympathetic nerve is related to fluctuations in the blood pressure, although it is not influenced by the breathing rate, since the sympathetic nerve can transmit only very low frequency signals. Hence, the LF (0.04-0.15 Hz) powers were modulated under cognitive stress. By observing the HF and LF components, we can estimate the subject's stress levels.

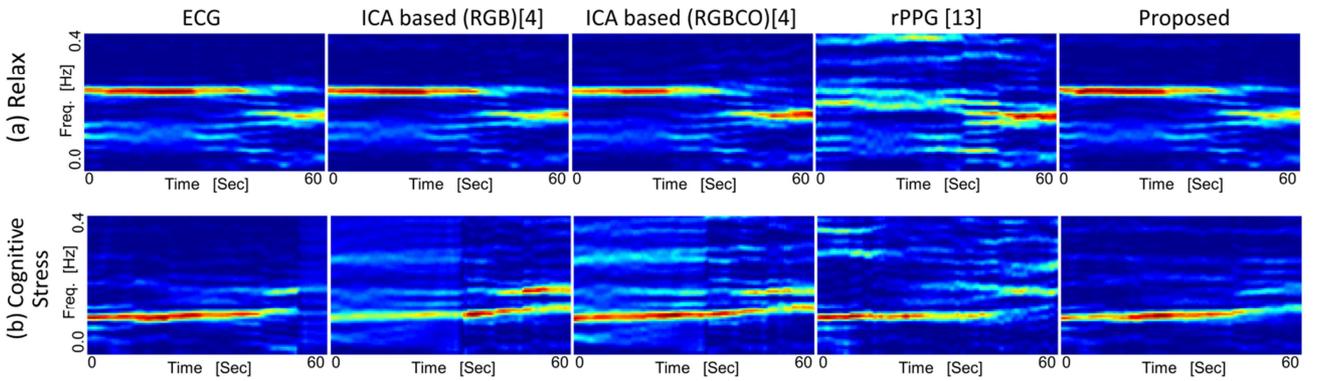
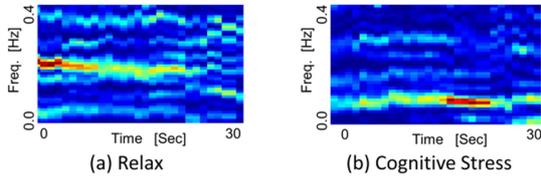
## 4 Experiments

### 4.1 Experimental Setup

Figure 7 shows the experimental setup used to obtain the BVP and HRVS with our method. The video data of a subject's face were taken from a distance of 3 meters with a digital single-lens reflex (DSLR) camera with a sensor of the five color (i.e., RGBCO) channels (12 bits/channel) [6]. We use the data of RGB channels to evaluate the proposed method, the Independent Component Analysis (ICA) based method [4] and chrominance based rPPG [7]. We also use the data of RGBCO channels for Independent Component Analysis (ICA) based method [4]. The frame rate of the camera was 30 frames per second (fps). Each frame was  $640 \times 480$  pixels. A standard Zuiko 50 mm lens was used in our experiment. Each frame was saved on a laptop PC (Dell Inc. Latitude E6530, 2.4 GHz, 3 MB cache). Artificial

**Table 1** Accuracy of the Measured Heart Rates at Rest and under Cognitive Stress

		(a) Relax				(b) Cognitive Stress			
		Subject #1	Subject #2	Subject #3	Subject #4	Subject #1	Subject #2	Subject #3	Subject #4
Hart Rate [bpm] *beat per minute	Electrocardiogram	83.63	64.40	60.78	72.75	84.61	71.82	60.97	72.19
	ICA based [4] w/ RGB	83.28	63.81	61.16	73.02	84.26	71.23	61.39	72.02
	ICA based [4] w/ RGBCO	83.43	63.85	61.04	72.99	84.38	72.10	61.31	72.10
	rPPG [7]	83.58	63.85	61.17	77.92	84.01	71.30	61.39	72.08
	Proposed method	83.50	63.87	60.08	72.81	84.74	72.06	61.30	72.15
Accuracy [%]	ICA based [4] w/ RGB	99.58	99.08	99.37	99.63	99.59	99.18	99.31	99.76
	ICA based [4] w/ RGBCO	99.76	99.15	99.57	99.67	99.73	99.61	99.44	99.88
	rPPG [7]	99.94	99.15	99.36	92.89	99.29	99.28	99.31	99.85
	Proposed method	99.84	99.18	98.85	99.92	99.85	99.67	99.46	99.94

**Fig. 9** Heart Rate Variability Spectrograms**Fig. 10** Heart Rate Variability Spectrograms without polarized filter

solar light (SERIC Ltd. SOLAX XC-100) was used to illuminate the face at a distance of 0.5 m from the subject. In the first experiment, we put polarized filters (crossed nicols) in front of the source of illumination and camera to simplify the estimation of the hemoglobin component by removing surface reflection. In the second experiment, we evaluate without using polarized filters assuming actual usage. The measurements taken with the electrocardiograph (NIHON KOHDEN RMT-1000) were used for the ground truth. In the experiments, we obtained videos from 4 participants. The subjects were three Japanese males and one female aged from 21 to 48 years old. The experiments for each subject were conducted under the two conditions, at rest (not under cognitive stress) and under cognitive stress (The subjects were required to keep subtracting 7

from 4000) respectively. In the first experiment with polarized filter, the duration is 120 second and the window size of spectral analysis is 60 seconds. In the second experiment without polarized filter, the duration is 60 second and the window size of spectral analysis is 30 seconds.

## 4.2 Experimental Results

### A. Experiment with Polarized Filter.

Figure 8 shows electrocardiogram (ECG) as ground truth and the results of the BVP detection obtained from proposed method and conventional methods [4][7]. The red circles indicate local peaks of the BVP. The ECG data has two peaks in each pulse interval. We extracted the first peak as the local peak for the calculation of peak interval. We can confirm that the proposed method and all other conventional methods successfully detected the BVP signal.

Tables 1 show the HR detected at rest and under cognitive stress respectively. The HR is obtained by the following equation.

$$heart\ rate = 60 / \overline{peak\ interval}$$

Here,  $\overline{peak\ interval}$  is the average of the time intervals of each peak. The figures in Tables 1 also show the accuracy of HR comparing with ECG data. The results

show that the proposed method had around over 99% accuracy for the HR both at rest and under cognitive stress, which is better than the ICA based method [4]. The result of rPPG at rest shows slightly better accuracy than proposed method in term of HR.

Figure 9 shows the results of the HRVS of Subject 1. Each spectrogram is described in heat map format. Red indicates high powers, and blue indicates low power. Each method shows high power of LF (0.04-0.15 Hz) under cognitive stress; whereas each method shows high power of HF (0.15-0.4 Hz) at rest. These features agree with a prior study on sympathetic activity [3]. However, in detail, the results of conventional methods show some difference from that of ECG. LF peaks of rPPG have lower value than ECG. (The peak of rPPG is yellow, whereas ECG's peak is red) The results of ICA based methods also has lower peak than ECG in the results of conative stress. On the other hand, the proposed method well agrees with that of ECG.

#### B. Experiment without Polarized Filter

Figure 10 shows the results of the HRVS of Subject 1 without putting polarized filter. We can confirm that high power of LF (0.04-0.15 Hz) under cognitive stress and HF (0.15-0.4 Hz) at rest successfully. The results was low resolution compared with 9, since he duration and window size of spectral analysis were smaller than the condition of Figure 9 as described in section 4.1.

### 5 Conclusion

In this paper, we proposed a novel framework for the estimation of HR and HRVS based on the visual estimation of hemoglobin components. In a study with four subjects, the results show that the proposed method could accurately estimate the HR and HRV. Especially the proposed method can get more accurate HRV than rPPG. The rPPG use first order approximation of Taylor expansion to describe internal reflection. On the other hand, the proposed method uses the modified Lambert-Beer law, more detail model. It could be a factor for the improvement. We also confirmed the effectiveness of our method for remote observation of the autonomic nervous system at rest and under cognitive stress without putting polarized filter.

At last we mention the limitation and the future work. Our experiment was conducted in an environment with no change in illumination and no large motions of the subjects. Evaluation under various conditions of illumination and movement is necessary for better bench mark with conventional method. An evaluation under ambient light will be done in a future work. We evaluated only Asian subjects in the experiment. We have to confirm the

effectiveness of the skin component extraction for Caucasian and Negroid subjects as well.

### REFERENCES

- [1] W. Verkruysse, L. O. Svaasand and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Opt. Exp.*, vol. 16, no. 26, pp. 21434-21445, (2008).
- [2] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Optics Express*, vol. 18, no. 10, pp. 10 762–10 774, (2010).
- [3] Pagani M, Furlan R, Pizzinelli P, Crivellaro W, Cerutti S, Malliani A. Spectral analysis of R-R and arterial pressure variabilities to assess sympatho-vagal interaction during mental stress in humans. *J Hypertens*; 7 (Suppl): S14–5, (1989).
- [4] D. McDuff, S. Gontarek, R. W. Picard, "Improvements in Remote Cardio-Pulmonary Measurement Using a Five Band Digital Camera" *IEEE Transactions on Biomedical Engineering*, pp. 2593-2601, (2014).
- [5] Y. Monno, M. Tanaka, and M. Okutomi, "Multispectral demosaicking using guided filter," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, pp. 82 990O–82 990O.(2012)
- [6] Y. Monno, S.Kikuchi, M. Tanaka, and M. Okutomi, "A Practical One-Shot Multispectral Imaging System Using a Single Image Sensor", *IEEE Trans. Image Processing*, vol.24, no.10, pp.3048-3059, Oct. (2015)
- [7] G. De Haan, V. Jeanne. "Robust pulse rate from chrominance-based rPPG," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, pp. 2878–2886 (2013).
- [8] N. Tsumura, N. Ojima, K. Sato, M. Shiraishi, H. Shimizu, H. Nabeshima, S. Akazaki, K. Hori, Y. Miyake, "Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin," *ACM Trans. on Graphics*, Vol.22, No.3, 770-779, (2003).
- [9] M.Hiroka, Firbank M., Essenpreis M., Cope M., Arrige S. R., Zee P. V. D., Delpy D. T., "A Monte Carlo investigation of optical pathlength in inhomogeneous tissue and its application to near-infrared spectroscopy." *Phys. Med. Biol.* 38, 1859-1876. (1993)
- [10] B. Martinez, M. F. Valstar, X. Binefa and M. Pantic "Local evidence aggregation for regression-based facial point detection", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 5, pp. 1149-1163, May, (2013).
- [11] M. P. Tarvainen, P. O. Ranta-aho and P. A. Karjalainen "An advanced detrending method with application to HRV analysis", *IEEE Trans. Biomed. Eng.*, vol. 49, no. 2, pp. 172-175, Feb., (2002)
- [12] Scargle, J. D. Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. *ApJ* 1:263 pp. 835-853 (1982)
- [13] Press, William H., and George B. Rybicki. "Fast Algorithm for Spectral Analysis of Unevenly Sampled Data", *Astrophysical Journal*. Vol. 338, pp. 277–280.(1989)
- [14] M. P. Tarvainen, P. O. Ranta-aho and P. A. Karjalainen " An advanced detrending method with application to HRV analysis", *IEEE Trans. Biomed. Eng.*, vol. 49, no. 2, pp. 172-175, Feb.(2002)
- [15] S. Akselrod, D. Gordon, F. A. Ubel, D. C. Shannon, A. Berger and R. J. Cohen "Power spectrum analysis of heart rate fluctuation: A quantitative probe of beat-to-beat cardiovascular control" *Science*, vol. 213, no. 4504, pp. 220-222, (1981)