Journal of Imaging Science and Technology® 62(5): 000000-1–000000-7, 2018.

Dual-Band Infrared Video-Based Measurement Using Pulse Wave Maps to Analyze Heart Rate Variability

Ryota Mitsuhashi▲
Graduate School of Science and Engineering, Chiba University, Chiba, Japan
E-mail: r.mitsuhashi@chiba-u.jp

Keiichiro Kagawa and Shoji Kawahito
Research Institute of Electronics, Shizuoka University, Hamamatsu, Japan

Chawan Koopipat
Department of Imaging and Printing Technology, Chulalongkorn University, Thailand

Norimichi Tsumura▲
Graduate School of Science and Engineering, Chiba University, Chiba, Japan

Abstract. *Remote photoplethysmography (rPPG) enables us to capture the vital signs such as pulse rate, respiratory rate, oxygen saturation, and even heart rate variability (HRV) without any contact devises. Although the papers of rPPG mainly focus on the use of standard RGB camera, it cannot used for the cases in night or under the dim light conditions. Therefore, in this paper, we propose a novel noncontact method for monitoring HRV without visible lighting. The proposed method uses dual-band infrared videos to ensure robustness to fluctuations in illumination. The hemoglobin component is extracted via a simple projection from the dual-band pixel values in logarithmic space. We demonstrate the accurate extraction of pulse wave signals using pulse wave maps. As the results, we indicated the effectiveness of HRV monitoring in the situation under the dim light condition.*
[DOI: 10.2352/J.ImagingSci.Technol.2018.62.5.000000]

1. INTRODUCTION

Remote photoplethysmography (rPPG) is a useful technique for monitoring vital signs such as pulse rate, respiratory rate, oxygen saturation, and even heart rate variability (HRV) using a standard RGB camera. The pulse rate reflects humans' physiological health, and plays an important role in fitness and health care monitoring. For further analysis of human physiological information, measurements of HRV (also known as the index of cardiac autonomic activity [1]) enable noncontact monitoring of the autonomic nervous system, which controls involuntary body functions. The low-frequency signals (LF) are widely known as one of the most reliable indicators of sympathetic activity [2], whereas the high-frequency (HF) signals are affected by breathing and are related to parasympathetic activity [3]. HRV monitoring provides many benefits in situations such as monitoring fatigue, concentration at work, and drowsiness when driving, and can also help to prevent sudden infant death syndrome (SIDS), heart attacks, or paroxysmal diseases in patients located either at home or in hospital.

2. RELATED WORKS

McDuff et al. [4] developed a method for detecting HRV performance using a five-band sensor with red, green, blue, cyan, and orange bands. They also developed a statistical model called blind source separation using independent component analysis (ICA), and demonstrated the effectiveness of remote HRV measurements. Poh et al. [5] found that HRV could be measured using a low-cost commerce webcam. Their method employs ICA to detect the HRV from variations in the spatially averaged pixel values of the region of interest (ROI) from standard RGB video recordings made under ambient light conditions. Although their method can easily detect the vital signs, thus allowing more practical uses than the approach of McDuff et al. [4], they assume that the videos contain a single periodic component. In other words, their method is not applicable if the subjects are moving with a specific frequency (e.g., when training in a gym). Alghoul et al. [6] proposed a method for estimating HRV using a single-band magnified video signal in the green channel based on an application of Eulerian video magnification [7]. They first proposed an HRV estimation technique using a single-band video signal, but only verified the effectiveness of this method in a laboratory setup. However, in real applications, the illumination often exhibits time-varying fluctuations. Thus, their estimations are inevitably affected by fluctuations in illumination. Kurita et al. [8] and Fukunishi et al. [9] proposed a remote HRV monitoring method for extracting hemoglobin information based on a skin optics model. Their approach considers the modified Lambert–Beer law

▲IS&T Members.

Received Sept. 4, 2018; accepted for publication Sept. 8, 2018; published online xx xx, xxxx. Associate Editor: Mathieu Hebert.
1062-3701/2018/62(5)/000000/7/$25.00

in standard RGB video recordings under ambient light conditions. The key point of their method is the removal of fluctuations in illumination (shading component) by projecting into a constant vector in a logarithmic color vector space.

As we described some applications of vital signs in the introduction part, it can be used for various monitoring systems. Especially, based on the above papers, we focus on the continuous monitoring of infants' vital signs to prevent SIDS and the monitoring of long-distance drivers for signs of fatigue, particularly at night. In both cases, physiological information such as vital signs must be monitored without visible lighting.

Inspired by several fundamental studies on pulse rate monitoring without visible lighting using single-band middle-wavelength video cameras [10], single-band near-infrared video cameras [11], and thermal video cameras [12], Mitsuhashi et al. [13, 14] proposed a method for monitoring the pulse rate using dual-band infrared video signals. They decomposed these video signals into hemoglobin, which is the component induced from blood flow changes, and shading, which is the component induced from intensity variations in illumination, based on an application presented in [8, 9]. As a result, the pulse rate can be accurately estimated even when the illumination is fluctuating. However, they could not capture the HRV because the pulse wave signal could not be extracted with sufficient accuracy. Regardless, we consider their method to have significant potential of further analysis for HRV monitoring if the extraction of the pulse wave signal induced from the heartbeat can be improved.

Kumar et al. [15] found that human faces can be distinguished as subregions that strongly reflect the pulsatile blood flow changes and other subregions that do not. They tracked faces using a Kanade–Lucas–Tomasi (KLT) tracker and divided the face region into several ROIs. In the frequency domain, they then computed the pulsatile components of the pulse wave signals arising from cardiac heartbeats in each subregion and generated a goodness metric (pulse wave map) that visualizes pulse wave component inside the face. This led to a robust pulse rate monitoring system that uses the single-band green channel in a low-cost standard RGB camera and enabled the precise extraction of pulse wave signals based on the spatial relative amplitude of the pulsatile components.

Mitsuhashi et al. [13, 14] attempted to extract the hemoglobin component from dual infrared bands via a simple projection in the logarithmic color vector space. Their ROI is manually selected to be part of the forehead, as many researchers typically use this region as the part of the ROIs. However, according to the work of [15], the strength of the pulse wave signal depends on the subregion inside the face. They did not consider this matter and the ROI will affect the accuracy of pulse rate estimation. Moreover, they could not capture the HRV parameters such as LF, HF, and LF/ HF ratio due to the quality of extracted pulse wave.

Therefore, in this paper, we aimed to improve the work presented by [13, 14] for extracting HRV parameters under the dim light condition. In order to achieve the extraction of HRV parameters, we consider the spatial information of extracted hemoglobin component. We replace the manual ROI selection with a process based on a weighted ratio calculated from the pulsatile components induced from heartbeat. By choosing ROIs based on this weighted ratio, we can detect the HRV parameters even when the illumination is fluctuating under dim light condition. The remainder of this paper is organized as follows. In Section 2, we briefly explain the conventional method for extracting the hemoglobin component from facial images [13, 14] as they presented and describe our proposed method for selecting the ROI based on the weighted ratio. In Section 3, we describe experiments to evaluate this approach, and present HRV estimation results in comparison with the ground truth measured by an electrocardiograph. The experimental results are discussed in Section 4, and we conclude by summarizing this study in Section 5.

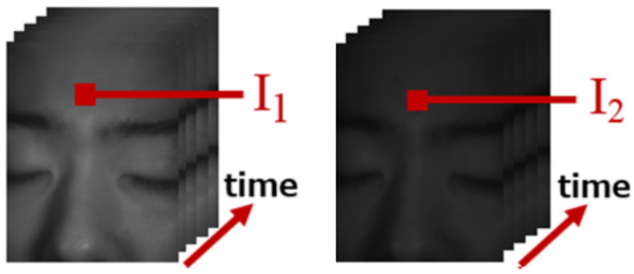
3. PROPOSED METHOD FOR ESTIMATING HRV USING DUAL-BAND INFRARED VIDEO SIGNALS

In this section, we briefly introduce the conventional method [13, 14] for extracting hemoglobin factor based on an application of [8, 9]. After the hemoglobin information has been extracted, the pulse wave signal is generated by integrating the pulse wave map and multiple inputs of rPPG signals, a process inspired by [15].

3.1 *Extraction of Hemoglobin Information Using Dual-Band Infrared Video Recordings*

In this section, we describe the procedure for obtaining pulse wave signals from two-band infrared video signals. This method uses a combination of bandpass filters with central wavelengths of 780 nm and 900 nm. This is because, in previous studies [13, 14], simulations based on spectral characteristics indicate that these wavelengths are effective filters for extracting the pulse rate. Therefore, these optical bandpass filters were attached to the front of each monochrome camera. Details about the experimental setup and experimental results are described in Section 3.

Figure 1 shows the facial video recordings captured from the two-band camera at the central wavelengths of 780 nm and 900 nm. The corresponding two-band pixel values of each wavelength were converted into points in the color vector space, where each pixel value is converted into a logarithmic value, as shown in Figure 2. The horizontal and vertical axes in Fig. 2 indicate the logarithmic pixel values in the two-band infrared videos. The studies of Kurita et al. [8] and Fukunishi et al. [9] used a two-layered skin model based on the modified Lambert–Beer law. Similarly, we applied the modified Lambert–Beer law in a single-layered model. The process of converting the two-band infrared video signals into their hemoglobin and shading components in each frame is as follows. Consider the two-dimensional plane constructed from the pixel values of two-band images



(a) 780 nm (b) 900 nm

Figure 1. Original dual-band infrared video recordings at different central wavelengths before the extraction of hemoglobin information.

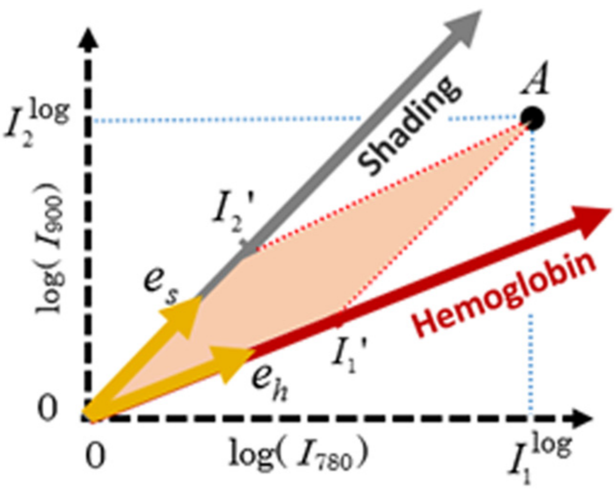


Figure 2. Outline of the proposed method for separating the hemoglobin and shading components in logarithmic color vector space

corresponding to the same location (e.g., forehead), as shown by the red point in Fig. 1(a), (b). An arbitrary vector **A** is expressed as the linear combination of basis vectors, as shown in the left-hand term of Eq. (1). Vector **A** can also be expressed as a linear combination of new basis vectors in the color space via a logarithmic transformation, as shown in Fig. 2. Therefore, vector **A** can be expressed by two patterns using the following equations.

$$\mathbf{A} = I_1\,\mathbf{e}_x + I_2\,\mathbf{e}_y = I_1'\,\mathbf{e}_h + I_2'\,\mathbf{e}_s, \tag{1}$$

where I_1 and I_2 are the logarithmic pixel values before the video signals are converted into their hemoglobin and shading components. I_1' and I_2' are the corresponding components after the video signals have been converted into their hemoglobin and shading components. $\mathbf{e}_x$ and $\mathbf{e}_y$ are the basis vectors in the color space; $\mathbf{e}_h$ and $\mathbf{e}_s$ are the new basis vectors for the hemoglobin and shading components. Equation (1) can be represented as

$$\begin{pmatrix}\mathbf{e}_x & \mathbf{e}_y\end{pmatrix}, \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix}\mathbf{e}_h & \mathbf{e}_s\end{pmatrix} \begin{pmatrix} I_1' \\ I_2' \end{pmatrix}. \tag{2}$$

Applying the two-dimensional inverse matrix of $(\mathbf{e}_h, \mathbf{e}_s)$ to Eq. (2) and noting that $(\mathbf{e}_x, \mathbf{e}_y)$ is the identity matrix, we can write as the following equation.

$$\begin{pmatrix} I_1' \\ I_2' \end{pmatrix} = \begin{pmatrix}\mathbf{e}_h & \mathbf{e}_s\end{pmatrix}^{-1} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}. \tag{3}$$

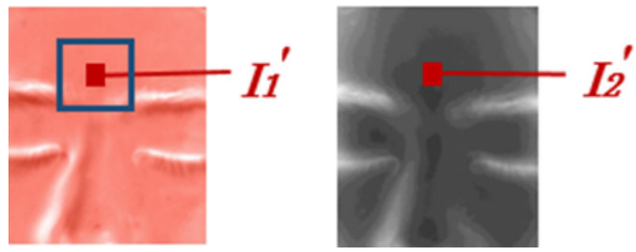
Suppose that the basis vector $\mathbf{e}_h$ is represented by the transpose expression ${}^t(\mathbf{h}_x, \mathbf{h}_y)$. The basis vector $\mathbf{e}_s$ is then given by the transpose matrix ${}^t(1/\sqrt{2} \quad 1/\sqrt{2})$, because the shading components are the same in any band. Thus, we have

$$\begin{pmatrix} I_1' \\ I_2' \end{pmatrix} = \begin{pmatrix}\mathbf{h}_x & 1/\sqrt{2} \\ \mathbf{h}_y & 1/\sqrt{2}\end{pmatrix}^{-1} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}. \tag{4}$$

In determining the elements of the hemoglobin vector, denoted by $\mathbf{h}_x$ and $\mathbf{h}_y$, Eq. (4) can be represented as

$$\begin{pmatrix} I_1' \\ I_2' \end{pmatrix} = \begin{pmatrix}\cos(\boldsymbol{\theta}) & 1 \\ \sin(\boldsymbol{\theta}) & 1\end{pmatrix}^{-1} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} \tag{5}$$

where $\boldsymbol{\theta}$ indicates the angle of the hemoglobin vector. We calculated the heart rate for each angle of the vector (91 steps from 0 to 90°) and then calculated the absolute error rate at every angle. Finally, we determined the effective hemoglobin vector as that which minimized the absolute error rate with respect to the ground truth heart rate. Decomposed videos of the separated hemoglobin and shading components are shown in Figure 3.



(a) Hemoglobin component (b) Shading component

Figure 3. Skin components after the separation of hemoglobin and shading by using our proposed algorithm.

3.2 *Pulse Wave Acquisition by Integrating the Pulse Wave Map and Multiple Inputs of rPPG Signals*

We evaluated the proposed method for extracting the hemoglobin component using the dual-band infrared video signals presented in [13, 14]. Using the resulting hemoglobin images, we focused on the automatic extraction of the pulse wave signal. With regard to the ROI, a face tracker is typically used to detect the face and set the ROI to the forehead or the cheek. However, the extracted pulse wave signal must be very clean to accurately estimate HRV. According to the maximum ratio combining (MRC) algorithm presented in [15], a clean rPPG signal can be generated by integrating the pulse wave map and multiple inputs of rPPG signals obtained from the subregions of the face. Let us introduce an example of how to generate clean pulse wave signals. Consider a facial region trimmed to a pixel resolution of

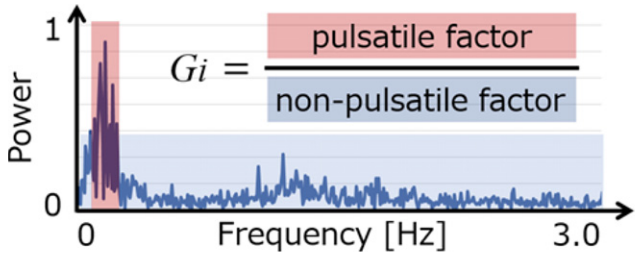
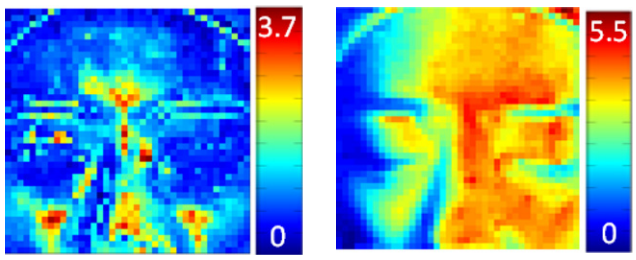


Figure 4. Calculation of the pulsatile component in the frequency domain.



(a) Single-band IR video (b) Hemoglobin component

Figure 5. Pulse wave map from single-band infrared video and extracted hemoglobin components using the proposed method.

400 × 400 squares, and divide the face into 1600 blocks of 10 × 10 pixels. We first compute the pixel value variations $y_i(t)$ by spatially averaging the pixel values within each subregion, where i denotes the corresponding subregion. All $y_i(t)$ are temporally filtered using a [0.75 Hz, 3.0 Hz] bandpass filter to remove noise from outside the band of interest.

Consider the sequence $y_1(t), y_2(t), \ldots, y_n(t)$ to be the different channels that receive signals and noise of different strengths. Finally, a clean pulse wave signal $p(t)$ is obtained by combining all of these different channels using a weighted average as follows:

$$p(t) = \sum_{i=1}^{n} G_i \, y_i(t) \tag{6}$$

where n denotes the number of subregions. G_i represents the weighted ratio, which is defined as the signal to noise ratio in the frequency domain. This is given by

$$G_i = \frac{\int_{\max Freq-0.3}^{\max Freq+0.3} Y(f)df}{\int_{\max Freq-0.3}^{2*\max Freq+0.3} Y(f)df - \int_{\max Freq-0.3}^{\max Freq+0.3} Y(f)df} \tag{7}$$

where $Y(f)$ denotes the power spectral density (PSD) of $y(i)$, as shown in Figure 4, and max $Freq$ represents the peak power spectrum density of $yi(t)$ in the frequency domain. The numerator represents the pulsatile factor induced from heartbeats and the denominator represents the non-pulsatile factor induced from the noise. We computed G_i for every subregion and then obtained a pulse distribution map, which reflects the proportion of the pulsatile component in each subregion (see Figure 5(a), (b)). Fig. 5(a) shows the pulse map obtained from single-band infrared video recordings under stable light conditions. This map indicates that the cheek and forehead subregions contain stronger pulsatile components than the other facial subregions, as expected. Fig. 5(b) shows the pulse map obtained from hemoglobin images extracted from dual-band infrared video recordings. As most subregions exhibit high values, this confirms the effectiveness of our separation method. After we have obtained the pulse map-based signal, a procedure based on the smoothness prior and an adaptive bandpass filter is performed. The window of the adaptive bandpass filter is set to [max $Freq - 0.3$ 2 ∗ max $Freq + 0.3$], because the detection of HRV is assumed to require a precise passband considering the harmonics. Finally, a clean pulse wave signal is obtained, as shown in Figure 6.

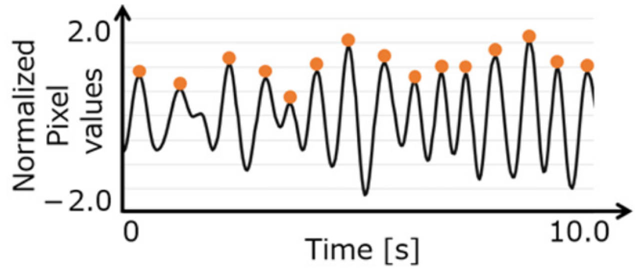


Figure 6. Pulse wave signal given by the proposed method considering the integration of the pulse wave map and multiple rPPG signals.

To compare the estimated HRV values from the extracted pulse wave signals with the ground truth obtained from an electrocardiograph, the pulse wave signal is interpolated with a cubic spline function at a sampling frequency of 50 Hz, as is the ground truth. The R–R wave (RR) intervals are obtained by detecting the peaks of the interpolated signal at every 20 frames. They are calculated as the intervals between neighboring peaks. The pulse rate is obtained by averaging these RR intervals according to HRV analysis is performed by PSD estimation using the Lomb periodogram. The LF and HF powers are calculated by integrating the PSD curve over the regions 0.04–0.15 Hz and

$$PR = \frac{60}{\overline{\text{RR intervals}}} \tag{8}$$

0.15–0.4 Hz, respectively. We calculate the ratio of LF and HF to verify the accuracy of our estimations.

4. EXPERIMENTAL SETUP AND RESULTS

In this section, we describe the experimental setup and procedure for verifying the effectiveness of our proposed method. The experimental results show that our proposed method provides good HRV estimates without visible lighting and when the illumination is fluctuating.

4.1 *Experimental Setup*

The experimental specification is shown in Table I. The experiments were performed indoors in a dark room with two artificial sunlight lamps acting as sources of illumination,

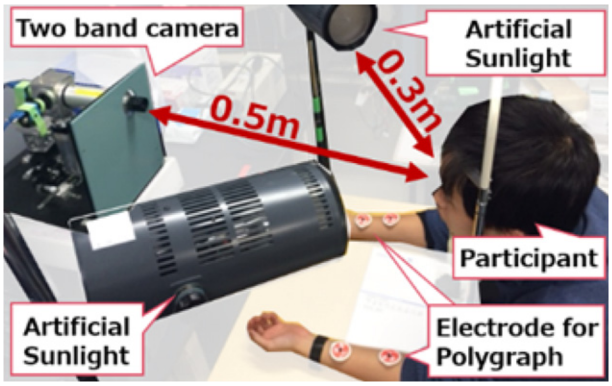


Figure 7. Two-band infrared video recording system

Table I. Experimental specification.

Camera	RGB CCD DMK 23UV024 (Imaging Source, Inc.)
Optical filter	FUJI FILTER OPTICAL IR78, IR90 (FUJIFILM, Inc.)
Light source	Artificial Light, XC-100 (SERIC, Inc.)
Electrocardiograph	RMT1000, Nihon Kohden Inc.

as shown in Figure 7. In this study, the flickering of the artificial sunlight played the role of fluctuations in the light source. Participants were seated in front of a table and fixed in position using a thin rest positioned in front of the two-band camera at distances of approximately 0.5 m from the camera and 0.3 m from each artificial sunlight source. The two-band camera system includes a beam splitter and monochrome cameras, with bandpass filters attached to the front of each camera.

As described in previous studies [13, 14], filters with central wavelengths of 780 nm and 900 nm were selected as the most effective combination for capturing variations in oxy-hemoglobin. Therefore, we limited the incident light to the monochrome camera with a central wavelength of 780 nm and a full width at half maximum of ±10 nm. We limited the incident light to the other monochrome camera with a central wavelength of 900 nm and a full width at half maximum of ±10 nm. We recorded videos using the following patterns. The first video was recorded without fluctuations in illumination using stable artificial sunlight. The stable artificial sunlight condition can be obtained by waiting for more than 30 min after turning on the light. The second video was recorded under conditions with fluctuations in illumination. These illumination fluctuations occur if the artificial sunlight source is turned up in less a few minutes. We assumed that the former environment had no illumination fluctuations and the latter suffered from illumination fluctuations. All videos were recorded using an 8-bit monochrome camera at 30 fps with pixel resolution of 640 × 480 and were saved in BMP format on a PC. We also recorded an electrocardiogram for each participant

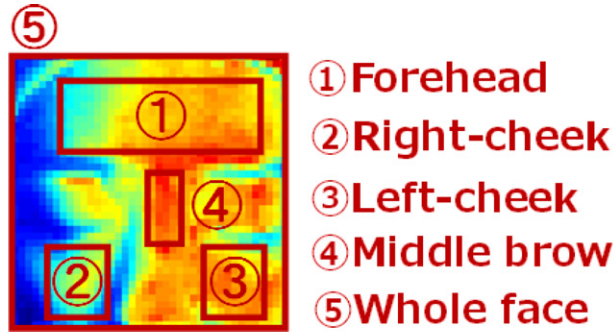


Figure 8. Several ROIs for validating the estimation accuracy of our proposed method.

Table II. PR comparison with the ground truth under conditions with fluctuations in illumination. AER: absolute error rate; MAE: mean absolute error; RMSE: root mean squared error.

	AER [%]	MAE [bpm]	RMSE [bpm]
One band 780 [nm]	2.74	2.51	2.56
One band 900 [nm]	2.68	2.57	2.58
Proposed Method	0.35	0.33	0.43

using a polygraph system at a sampling rate of 1 kHz with a cut-off frequency of 15 Hz. The ground truth heart rate was calculated by averaging the RR intervals obtained from the electrocardiogram and used to verify the accuracy of the proposed method.

4.2 *Experimental Results*

Since we obtained hemoglobin images as shown in Fig. 3(a), which means the spatial distribution of the hemoglobin components, we evaluated the estimated pulse rate at each parts of the face as following metrics: the absolute error rate (AER), the mean absolute error (MAE), and the root mean squared error (RMSE) which are given by following equations. And we evaluated the spatial parts of the face, which are from forehead, right cheek, left cheek, middle brow, and whole face. The average values are listed as following Tables II–V by averaging the estimated values at each parts of face as shown in Figure 8.

$$AER = \frac{|GT - EV|}{GT} \times 100 \tag{9}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|GT - EV| \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(GT - EV)^2}. \tag{11}$$

Here, GT represents the ground truth obtained from the electrocardiogram and EV represents the estimated

Table III. LF comparison with the ground truth under conditions with fluctuations in illumination. AER: absolute error rate; MAE: mean absolute error; RMSE: root mean squared error.

	AER [%]	MAE [n.u.]	RMSE [n.u.]
One band 780 [nm]	27.18	6.25	6.40
One band 900 [nm]	29.74	13.74	14.17
Proposed Method	10.78	2.48	2.79

Table IV. HF comparison with the ground truth under conditions with fluctuations in illumination. AER: absolute error rate; MAE: mean absolute error; RMSE: root mean squared error.

	AER [%]	MAE [n.u.]	RMSE [n.u.]
One band 780 [nm]	23.27	6.36	6.96
One band 900 [nm]	28.60	10.46	10.80
Proposed Method	8.14	6.26	6.84

Table V. LF/HF comparison with the ground truth under conditions with fluctuations in illumination. AER: absolute error rate; MAE: mean absolute error; RMSE: root mean squared error.

	AER [%]	MAE [n.u.]	RMSE [n.u.]
One band 780 [nm]	27.44	0.052	0.052
One band 900 [nm]	37.53	0.268	0.125
Proposed Method	11.70	0.035	0.050

value determined using the proposed method. The AER is normalized with respect to the ground truth. This gives an indication of how close the estimated value is to the ground truth. We calculated pulse rate (PR), LF, HF, and LF/HF for each single-band and dual-band infrared video. Table I compares the estimated PR with the ground truth.

According to Table I, the AER of the pulse rate computed from dual-band infrared video signals is lower than that from single-band infrared videos as you can see this Table. Tables II and III present comparisons of LF and HF with the ground truth. These results confirm that our proposed method performs well under fluctuations in illumination. This is because the corresponding pixel values of the dual-band videos are separated into the hemoglobin and shading components by fixing the fluctuation of illumination (shading) to (1, 1) in the color vector space. The results in Table IV indicate that our proposed method offers superior performance with comparison of conventional single-band estimation.

5. DISCUSSION

The experimental results demonstrate that the proposed method provides more accurate estimates of PR, LF, HF, and LF/HF in an environment without visible lighting than the conventional method based on noncontact pulse wave monitoring using a single-band infrared video camera. In particular, our method remains effective under fluctuations in illumination, as it uses the effective combination of infrared filters for monitoring HRV. As described in Section 1, the single-band infrared video camera method calculates the spatially averaged pixel values within the ROI selected by manual [11]. However, our proposed method considers the brief skin optics and weighted integration of spatial distribution of hemoglobin components. According to the results in Table II, we obtained superior AER, MAE, and RMSE values in comparison with the conventional method. Moreover, we generated the pulse wave signal by integrating a weighted ratio and multiple inputs of rPPGs. Based on this, the generated pulse wave signal depends on the pulse map. In other words, if the pulse maps are well generated, the pulse wave signal becomes cleaner and the PR, LF, HF, and LF/HF estimations will be more accurate. According to Fig. 5(a), the pulse maps are well generated because the subregions of the forehead and cheeks exhibit relatively higher amplitudes than the other subregions, as shown in Fig. 8(a). According to Fig. 5(b), such amplitudes were observed in the same subregions as Fig. 5(a). We consider the single-band video recordings of 900 nm to have quite low signal to noise ratio (SNR), as indicated in Fig. 1(b) and Fig. 8(b). The principle of the pulse map suggests that the method will be weaker with low SNRs. To address this issue, we attempted to generate a high SNR pulse map by magnifying the video presented in [7], but we did not obtain a sufficient pulse map because the SNR in the original video recording is too low. The noise was also magnified and appeared as artifacts. We will attempt to rebuild the experimental setup in future work.

Second, it can be considered that the participants moved slightly during the continuous 2 min recording. The accuracy could be improved by implementing facial tracking or mask processing of the face region during the video recording step. We considered the possibility that tiny movements could appear as artifacts and affect the estimation accuracy. The results in Tables III–V confirm that the single-band infrared video cannot be used to obtain the HRV with sufficient accuracy under fluctuations in the illumination, because the AER values of the pulse rates measured around 30% by using only single-band infrared video. We considered that the bandpass filter did not sufficiently remove the fluctuations in illumination because of aperiodic noise. Hence, the noise frequency was included within the frequency range of the bandpass filter and the pulse rate was affected by the inclusion of this noise. The proposed method obtained AER values for PR, LF, and HF of around 10%. These results demonstrate the robustness of the proposed method against the illumination fluctuations. This is a result of using dual-band infrared video signals, because the video signals are decomposed into the hemoglobin and shading components by the application of the basis translation matrix in the color vector space. As we mentioned earlier, better accuracy could be obtained by implementing facial tracking or mask processing of the face region.

6. CONCLUSIONS AND FUTURE WORKS

We have proposed a noncontact HRV monitoring method that is robust to illumination fluctuations. Separated hemoglobin and shading components were obtained by determining a new basis vector in the logarithmic color vector space. As shown in Table IV, our proposed method greatly improves the HRV estimation accuracy, especially LF/HF/ ratio, with an AER of 11.70 % compared with 27.44% and 37.53% using the conventional method based on single-band infrared video signals. To verify the effectiveness of our method, a large number of subjects will be recruited for future studies. It will be necessary to improve the accuracy of HRV estimation using the proposed method by implementing facial tracking and generating high-performance pulse wave maps. We obtained the optimal hemoglobin vector in the case of a small pixel distribution. Therefore, we need to decompose the hemoglobin component using a large distribution of the pixel values. By implementing these improvements, we expect to obtain further robust HRV monitoring that can compensate the motion of subjects.

ACKNOWLEDGMENT

This work was supported in part by the MEXT/JST COI STREAM program.

REFERENCES

[1] M. Malik, J. Bigger, A. Camm, R. Kleiger, A. Malliani, A. Moss, and P. Schwartz, "Heart rate variability: standards of measurement, physiological interpletation, and clinical use," Eur. Heart J. **17**, 354–381 (1996).

[2] M. Pagani, R. Furlan, P. Pizzinelli, W. Crivellaro, S. Cerutti, and A. Malliani, "Spectral analysis of R-R and arterial pressure variabilities to assess sympatho-vagal interaction during mental stress in humans," J. Hypertens **7** (1989) S14–5.

[3] S. Akselrod, D. Gordon, F. A. Ubel, D. C. Shannon, A. Berger, and R. J. Cohen, "Power spectrum analysis of heart rate fluctuation: a quantative probe of beat-to-beat cardiovascular control," Science **213**, 220–222 (1981).

[4] D. McDuff, S Gontarek, and R. W. Picard, *Improvements in Remote Cardio-Pulmonary Measurement Using a Five Band Digital Camera* (IEEE Transactions on Biomedical Engineering, 2014), pp. 2593–2601.

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

[6] K. Alghoul, S. Alharthi, H. A. Osman, and A. E. Saddik, "Heart Rate Variability Extraction From Videos Signals: ICA vs. EVM Comparison," IEEE Access **5**, 4711–4719 (2017).

[7] H. Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, "Eulerian video magnification for revealing subtle changes in the world," ACM Trans. Graph. **31**, 1–8 (2012).

[8] K. Kurita, T. Yonezawa, M. Kuroshima, and N. Tsumura, "Non-contact video based estimation for heart rate variability spectrogram using ambient light by extracting hemoglobin information," Color and Imaging Conference **2015**, 207–211 (2015).

[9] M. Fukunishi, K. Kurita, S. Kawahito, and N. Tsumura, "Non-contact video-based estimation of heart rate variability spectrogram from hemoglobin composition," J. Artif. Life Robot. (2017).

[10] M. Garbey, N. Sun, A. Merla, and I. Pavlidis, "Contact-free measurement of cardiac pulse based on the analysis of thermal imagery," IEEE Trans. Biomed. Eng. **54**, 1418–1426 (2007).

[11] W. Zeng, Q. Zhang, Y. Zhou, G. Xu, and G. Liang, "Infrared video based non-invasive heart rate measurement," IEEE Conf. Robotics Biomimetics (December 6–9, 2015) 1041–104. Q.1

[12] K. Hamedani, A. Veeraraghanvan, and A. Sabharwal, "Spatio-temporal filtering of thermal video sequences for heart rate," Expert Syst. Appl. **54**, 88–94 (2016).

[13] R. Mitsuhashi, G. Okada, K. Kurita, S. Kawahito, K. Kagawa, K. Chawan, and N. Tsumura, "Non-contact video-based method for monitoring pulse wave on face without visible lighting," *Color and Imaging Conference* (2017), Vol. 2017, pp. 207–211. Q.2

[14] R. Mitsuhashi, G. Okada, K. Kurita, S. Kawahito, K. Kagawa, K. Chawan, and N Tsumura, "Non-contact pulse wave detection by two-band infrared video-based measurement on face without visible lighting," J. Artif. Life Robot. (2018). Q.3

[15] M. Kumar, A. Veeraraghavan, and A. Sabharwal, "Distance PPG: Robust non-contact vital signs monitoring using a camera," Biomed. Opt. Express **6** (2015) 1565–88.

Queries for IS&T paper 0488

Journal:	**JIST**
Author:	**Mitsuhashi et al.**
Short title:	**Dual-band infrared video-based measurement using pulse wave maps to analyze heart rate variability**

Page 7

Query 1:
AU: Please check the page range of ref. [11].

Page 7

Query 2:
Au: Please provide place and publisher for ref. [13].

Page 7

Query 3:
Au: Please provide volume and page range for ref. [9] and [14].