

Two-Band Infrared Video-Based Measurement for Non-Contact Pulse Wave Detection on Face without Visible Lighting

Ryota Mitsuhashi¹⁾, Genki Okada¹⁾, Koki Kurita¹⁾, Keiichiro Kagawa²⁾, Shoji Kawahito²⁾, Chawan Koopipat³⁾, Norimichi Tsumura¹⁾

1) Graduate School of Advanced Integration Science, Chiba University, CHIBA, JAPAN

2) Research Institute of Electronics, Shizuoka University, HAMAMATSU, JAPAN

3) Department of Imaging and Printing Technology, Chulalongkorn University, Thailand

Abstract

In this paper, we propose a novel non-contact pulse wave monitoring method which is robust to fluctuation of illumination by using two-band infrared videos. Since the proposed method uses the infrared light for illumination, we can detect the pulse wave on the face without visible lighting. The corresponding two-band pixel values in the videos can be separated into hemoglobin and shading components by applying separation matrix in logarithmic space for two pixel values. Since the shading component is separated, the extracted hemoglobin will be robust to the fluctuation of illumination. The pixel values of the region of interest (ROI) were spatially averaged all over the pixels for each frame. The averaged values are used to form the raw trace signal. Finally, the pulse wave and pulse rate were obtained from raw trace signal through several signal processing such as detrend, adaptive bandpass filter, and peak detection. We evaluated the absolute error rate for pulse rate between the estimated value and the ground truth obtained by the electrocardiogram. From the experiment, we found that the performance of our method was greatly improved in comparison with the conventional method by means of one-band infrared video.

1. Introduction

Recently, non-contact methods were proposed to measure physiological information using video camera. Pulse rate (PR) is one of the vitalsigns to reflect physiological information of human, which plays an important role in health care monitoring. Pulse rate monitoring can be used to monitor fatigue, concentration and drowsiness in driving, and to prevent sudden infant death syndrome or paroxysmal disease of the patients in the house and the hospital.

M. Z. Poh *et al.* [1] [2] developed a non-contact pulse wave monitoring method by applying independent component analysis to the variation of the spatially averaged pixel in the ROI by RGB video recordings in daylight. K. Kurita *et al.* [3] proposed a non-contact pulse wave monitoring method by extracting hemoglobin information from RGB video recordings on ambient light situation.

In the above applications, it can be required to measure the face in night while the participants are driving or sleeping without visible lighting. The conventional RGB video recordings cannot be used in this situation. M. Garbey *et al.* [4] proposed a method to measure the pulse on neck from thermal video by recording the variation of the pixel values. They conducted the experiment using one-band middle wavelength infrared video. W. Zeng *et al.* [5] suggested a method to estimate the pulse rate using one-band infrared video. Although methods by Garbey and Zeng [4, 5] can detect pulse rate in dim light condition, they cannot be used in the situation where illumination is fluctuated on real environment such as the situation in driving. Since the pixel values were affected by fluctuation of illumination, The results of measurements can also be affected.

In this paper, therefore, we propose the non-contact pulse wave monitoring method which is robust to fluctuation of illumination by using two-band infrared videos. Two-band infrared videos can be separated into hemoglobin and shading components by applying separation matrix in logarithmic space. Since the shading component is separated, the extracted hemoglobin information will be robust to the fluctuation of illumination. The pixel values of the ROI are spatially averaged all over the pixels for each frame. The averaged pixel values are used to form the raw trace signal. Finally, the pulse wave and pulse rate are obtained from raw trace signal through several signal processing. We will evaluate the absolute error rate for pulse rate between the estimated value and the ground truth obtained by the electrocardiogram as one of the contact method.

2. Proposed Method for Robust Extraction of Pulse Wave from Two-band Infrared Videos

In this section, we describe the procedure for acquiring the pulse wave from two-band infrared videos. For the explanation in this section, we suppose to use the combination of bandpass filters whose central wavelength are 780nm and 900nm. The filters were attached in front of each monochrome camera. It is noted this combination of filters is found in the selection process in Section 3.

Fig. 1 (a), (b) shows the captured facial video recordings from two-band camera whose central wavelength are 780nm and 900nm, respectively. The corresponding two-band pixel values of each wavelength were converted into a point in color vector space where each pixel value is converted into the logarithmic value as shown in Fig. 2. The horizontal and vertical axis in Fig. 2 indicate the logarithmic pixel values in two-band infrared videos, respectively.

Let me introduce a brief example of converting from two-band infrared videos into hemoglobin and shading components at each frame. The arbitrary point *A* in Fig. 2 was expressed by the coordinates of (I_1, I_2) in the Fig. 1 (a), (b), respectively.

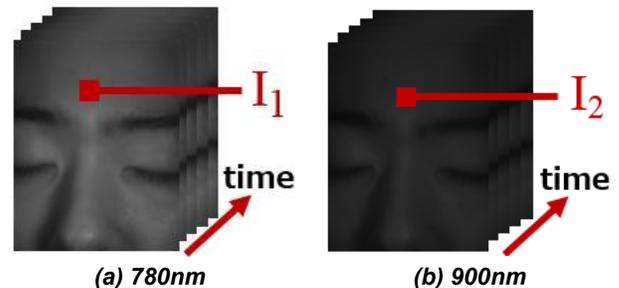


Fig. 1. Infrared video recordings (a) Central wavelength is 780nm, (b) Central wavelength is 900nm

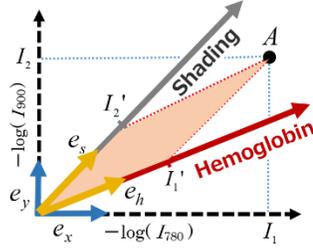


Fig. 2. The outline of our method how to separate the hemoglobin and shading in color vector space.

The point A was also expressed by the coordinates of (I_1', I_2') by acquiring the new basis vectors of hemoglobin and shading components in the color vector space. Therefore, the point A was expressed by two patterns as the following equation.

$$A = I_1 e_x + I_2 e_y = I_1' e_h + I_2' e_s \quad (1)$$

Where I_1 and I_2 are logarithmic pixel values before the videos were converted into hemoglobin and shading components. I_1' and I_2' are components after the videos were converted into hemoglobin and shading components. e_x and e_y are the basis vector in color vector space. e_h and e_s are the new basis vector for hemoglobin and shading components. Equation (1) is represented by the expression of the brief two dimensional metrics as the following equation.

$$\begin{pmatrix} e_x & e_y \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} e_h & e_s \end{pmatrix} \begin{pmatrix} I_1' \\ I_2' \end{pmatrix} \quad (2)$$

By applying the two dimensional inverse matrix of $\begin{pmatrix} e_h & e_s \end{pmatrix}$ to eq (2), the following eq (3) is obtained since the $\begin{pmatrix} e_x & e_y \end{pmatrix}$ is identity matrix.

$$\begin{pmatrix} I_1' \\ I_2' \end{pmatrix} = \begin{pmatrix} e_h & e_s \end{pmatrix}^{-1} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} \quad (3)$$

It is supposed that the basis vector e_h is represented by transpose expression, ${}^t(h_x \ h_y)$. The basis vector e_s can be given by transpose matrix, ${}^t(1 \ 1)$ since the shading components are same at any bands.

$$\begin{pmatrix} I_1' \\ I_2' \end{pmatrix} = \begin{pmatrix} h_x & 1 \\ h_y & 1 \end{pmatrix}^{-1} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} \quad (4)$$

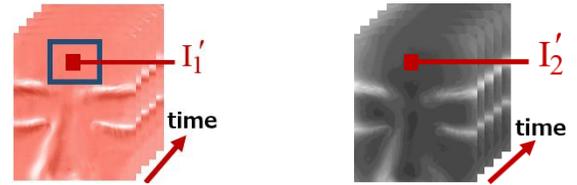
When we determine the each element of hemoglobin vector, h_x and h_y , the eq (4) was represented as the following equation.

$$\begin{pmatrix} I_1' \\ I_2' \end{pmatrix} = \begin{pmatrix} c \circ \xi \theta & 1 \\ s i \not\theta & 1 \end{pmatrix}^{-1} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} \quad (5)$$

Where θ indicates the angle of hemoglobin vector as the range from 0 to 360 degrees. The hemoglobin vector was set by simulating the case of all angles every one degree. We selected the most suitable vector of hemoglobin to estimate the pulse rate with the highest

accuracy. The decomposed videos of separated hemoglobin and shading components are shown in Fig. 3 (a), (b), respectively.

We focused on hemoglobin component and selected the *ROI* at the forehead of the participants with pixel resolution of 100×200 . The raw trace signal was formed by spatially averaging all over the pixels for each frame as shown in Fig. 4 (a). The raw trace signal shows parts of the results for only 5 seconds in continuous 2 minutes. As shown in Fig. 4 (b), the raw trace signal was detrended using a procedure based on a smoothness priors approach [5]. The detrended signal was bandpass-filtered as the following two steps. Firstly, the window of bandpass-filter was decided based on the power spectrum density (PSD) of the detrended signal. The maximum peak in the range of from 0.75Hz to 3.0Hz (45 to 180 bpm) was set as the central frequency of the window with the width of 0.4Hz in the PSD. Secondly detrended signal was bandpass-filtered by above window.

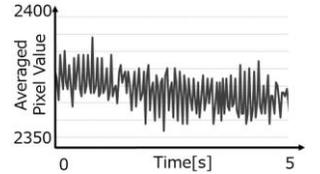


(a) Hemoglobin Component (a) Shading Component

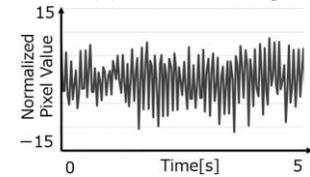
Fig. 3. Extracted components obtained by basis translation (a) Hemoglobin component, (b) Shading component.

Fig. 4 (c) shows the pulse wave which was bandpass filtered by adaptive window as mentioned on the above. Moreover, the peaks of the pulse wave were detected by calculating the maximum of the whole normalized pixel values every 15 frames. The RR intervals were calculated as intervals between the neighboring peaks. *PR* was calculated by averaging the RR intervals as the following equation.

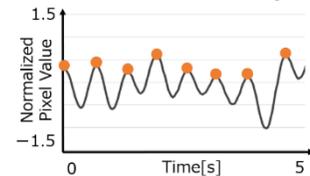
$$PR = \frac{60}{\text{RR intervals}} \quad (6)$$



(a) Raw trace signal



(b) Detrended signal



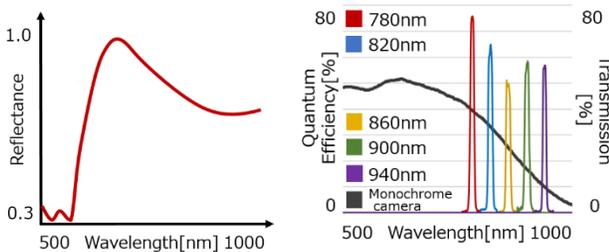
(c) Peak detected pulse wave

Fig. 4. Pulse wave processing (a) The raw trace that was spatially averaged pixel values in ROI in 5 seconds, (b) The detrended signal, (c) The bandpass filtered and the peak detected pulse wave

3. Selection of the effective two-band Infrared Filters

The pulse wave was obtained by extracting hemoglobin information by converting from two-band infrared video to decomposed video as we mentioned above section. In this regard, it is necessary to find the effective combination of the two-band infrared wavelength to capture the enough variation of hemoglobin component. We simulated the combinations of the infrared bandpass filters considering the relation among the reflectance of the oxy-hemoglobin, the sensitivity of the monochrome camera, and the spectral curves of the bandpass filters as shown in Fig. 5 (a), (b), respectively. We used the bandpass filters whose central wavelengths are 780nm, 820nm, 860nm, 900nm, and 940nm, with the full width at half maximum of $\pm 10\text{nm}$, respectively. (Edmund Optics Japan Inc.) According to the relation about the spectral curves of the two-band filters and the reflectance of the oxy-hemoglobin, a lot of information about the absorbance characteristics were obtained by selecting the filters at a far wavelength each other with the range of from 780nm to 940nm. Therefore, two-band infrared videos can be got easier the enough variation of the oxy-hemoglobin than one band infrared video. On the other hand, according to the relation between the sensitivity of the monochrome camera and the reflectance of the oxy-hemoglobin, the two-band are preferred to be separated each other to obtain the different trend at each band. However, the monochrome camera has a sensitivity that tends to decrease with closing to 1,000nm and is heavily affected to the random noise of the camera in higher wavelength. Based on these situations, the effective combination of the bandpass filters were determined by the tradeoff relationships between the reflectance of the oxy-hemoglobin and the sensitivity of the monochrome camera as you can see from Fig. 5 (a) and (b).

We simulated the experiment as following steps to capture the variation of oxy-hemoglobin. Firstly, we calculated the reflectance of the skin to close to the real skin. Secondly, the pixel values were calculated from integrating the reflectance of the skin, the spectral curve of oxy-hemoglobin, the sensitivity of the monochrome camera, and the spectral curves of the bandpass filters. Thirdly, the pixel values were separated into the hemoglobin and shading components. Finally, we evaluated the separated components using the evaluated function and obtained the effective combination of the bandpass filters to capture enough variation of the oxy-hemoglobin.



(a) Oxy-hemoglobin

(b) Spectral curves

Fig. 5. Spectral curves (a) The reflectance of the oxy-hemoglobin. (b) The sensitivity of the monochrome camera and the spectral curves of the bandpass filters

The hemoglobin distributes to all over the skin by blood circulation. In this paper, we simulated the situation that the value of the density of oxy-hemoglobin were distributed to each pixel values. We calculated the distribution of the oxy-hemoglobin as a

range of $0.200 \sim 0.213[\text{cm}^2/\text{mol}]$, with 100 steps, to close to the reflectance on the real face. In addition, the vessel is slightly enlarged by the heartbeat. Therefore, the density of oxy-hemoglobin is varied due to the heartbeat. We calculated the variation of oxy-hemoglobin by the heartbeat as a range of from 0.00000 to $0.00068[\text{cm}^2/\text{mol}]$, with two steps, to close to the reflectance on the real face. Based on these parameters, the virtual skin was calculated, and simulated the virtual model for selection of the effective filters.

The pixel values that were captured from the virtual skin were determined by integrating each component as the following equation.

$$I = \int sR(\lambda)C(\lambda)B(\lambda) d\lambda \quad (7)$$

Where $R(\lambda)$ is the reflectance of the virtual skin, $C(\lambda)$ is the sensitivity of the monochrome camera, $B(\lambda)$ is the spectral curves of each bandpass filter, s is the factor of the shading, and λ is wavelength. We calculated the intensity of the illumination as the constant in this paper. The pixel values of the two-band infrared videos were calculated by quantizing with the 8 bit in the simulation.

The component of the oxy-hemoglobin was estimated by determining the new basis vector in the color vector space composed from the two-band infrared videos. The corresponding two pixel values of each band were converted into the color vector space from RGB space by taking logarithm. The actual pixel values include the dark current noise. Dark current noise gives a heavy influence on the estimation of the pulse rate. In this paper, we simulated the dark current noise by assuming the normalized distribution function.

We evaluated the combinations of the filters by the following function to capture the variation of the oxy-hemoglobin effectively.

$$E = \frac{|I_{h+\Delta h} - I_h|}{(V_{h+\Delta h} + V_h) / 2} \quad (8)$$

where I_h and V_h are the pixel values and the variance of pixel values before oxy-hemoglobin were varied from specified density level, respectively. $I_{h+\Delta h}$ and $V_{h+\Delta h}$ are also the pixel values and the variance of pixel values after oxy-hemoglobin were varied from it. The evaluation value was closed to zero when the variation in oxy-hemoglobin was not captured due to the increase of dark current noise. Therefore, the set of the filters which give high evaluation value in spite of increase of noise were determined as the effective combination for easier detecting pulse wave and PR in this paper.

As a result of the simulation, the evaluation curves in each combination of the filters were obtained as shown in Fig. 6.

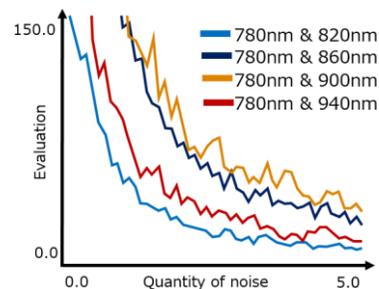


Fig. 6. Evaluation curves of each combination of the filters using evaluated function as we mentioned as the Eq (8)

The combinations of the filters whose wavelength are close to each other were heavily affected even if the noise was slight amount. The combinations whose wavelength are far from each other were tend to have robustness to the noise. However, the combination whose wavelength are 780nm and 940nm had the low evaluation because of the low sensitivity of the monochrome camera around 1,000nm as Fig. 5 (a). Therefore, the combination whose wavelength are 780nm and 900nm was selected as the effective combination to capture the variation of the oxy-hemoglobin in this paper.

4. Experimental Setup

The experiments were conducted at indoors in dark room with two artificial sunlight as the source of illumination as shown in Fig. 7. In this paper, the flicker of the artificial sunlight was played a role as the fluctuation of the light source environment. Participants were seated in front of a table and fixed by thin rest in front of the two-band camera at a distance of approximately 0.5 m from camera and 0.3 m from the each artificial sunlight, respectively. Two-band camera system is composed from the beam splitter and monochrome cameras where bandpass filters are attached in front of each camera.

As we mentioned above in section 3, the filters with the central wavelength of 780nm and 900nm were selected as the effective combination to capture the variation of the oxy-hemoglobin. Therefore, we limited the incident light of the monochrome camera whose central wavelength is 780nm with the full width at half maximum of ± 10 nm. On the other monochrome camera, we also limited the incident light of the range whose central wavelength is 900nm with the full width at half maximum of ± 10 nm. During the experiment, participants were asked to keep still as possible and breathe spontaneously. Moreover, participants were also asked to face the two-band camera while their video was recorded for two minutes. We recorded the videos as the following two patterns. Firstly, the video was recorded in the situation without fluctuation of illumination using artificial sunlight in stable condition. The stable condition of artificial sunlight can be obtained by waiting more than 30 minutes after the start of lighting. Secondly, the videos was also recorded in the situation with fluctuation of illumination using artificial sunlight. This fluctuation of illumination can be observed if the artificial sunlight is turned up in less a minute. We conducted the experiment by assuming the former as an environment without fluctuation of illumination and the latter as an environment with fluctuation of illumination, respectively. All videos were recorded in 8-bit monochrome camera at 30 fps, the pixel resolution of 640×480 , and saved in BMP format on the PC.

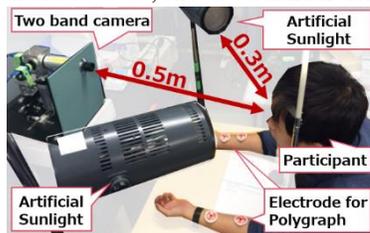


Fig. 7. Two-band infrared videos were obtained by recording the reflectance light from the skin using artificial sunlight.

We also recorded their electrocardiogram using the polygraph system at a sampling rate of 50 Hz and a cut-off frequency of 15Hz. [RMT1000: NIHON KOHDEN. Inc.] The ground truth of the heart rate was calculated by averaging RR-interval obtained from the electrocardiogram for the verification of the accuracy.

5. Experimental Results

We evaluated the estimated pulse rate by our method using absolute error rate (*AER*) as the following evaluation equation.

$$AER = \frac{|GT - EV|}{GT} \times 100 \quad (9)$$

Where *GT* is the ground truth obtained by the electrocardiogram. *EV* is the estimation values using our method. The *AER* between the estimation values and the ground truth normalized against the ground truth. This gives an indication of how close the estimation value against the ground truth in one band case and two-band case.

Table 1 shows the profile in comparison of the estimated pulse rate and the ground truth without varying the illumination. According to the Table 1, *AER* of the pulse rate using two-band infrared videos indicate the higher accuracy in comparison of the conventional method of Zeng *et al.* [5] using one band infrared video.

Table 2 shows the profile of the results of comparing the estimation of the pulse rate and the ground truth with varying the illumination. We confirmed the results that the estimation of pulse rate were heavily affected by the fluctuation of illumination. The results of our proposed method also indicate our method greatly improves the performance in comparison of the conventional method since the corresponding pixel values of videos are separated into hemoglobin and shading components by fixing to (1, 1) vector the fluctuation of illumination (shading) in the color vector space.

6. Discussion

As we mentioned results of our experiment, we estimated the *PR* with high accuracy even in an environment without visible lighting comparing with the conventional method of non-contact pulse wave monitoring using one band infrared video camera.

As we mentioned in section 1, Zeng *et al.* proposed the non-contact measurement method of the heart rate using one-band infrared video camera [5]. The raw trace signal was obtained by calculating the spatially averaged pixel values in the *ROI*. Through the frequency analysis, the *PR* was estimated with high accuracy. The performance of Zeng's method was indicated under 1% of *AER*.

According to the Table.1, we obtained the *AER* of the estimated *PR* with accuracy using only one band video, 8.17%, 6.98%, respectively. The accuracy can be improved by implementing the face tracking in the step of recording the videos. The *ROI* without tracking were captured to the different place in each frame. In spite of participant's faces were fixed to the thin rest and set the *ROI* of forehead, participants slightly moved in the continuous two minutes during the recordings. We considered that the raw trace signal was affected due to the slight movement and the *PR* was also affected by the slight movement. According to the Table 2, we confirmed that one band infrared video could not been obtained sufficiently to the *PR* with fluctuation of illumination since *AER* of the *HR* using only one band infrared video were indicated as the low performance whose accuracy of 18.66% and 29.56%, respectively. We considered that bandpass filter was not removed sufficiently to the fluctuation of illumination that has an aperiodic noise. Therefore, the frequency of the noise was included to the frequency of the bandpass filter and the *PR* was also affected by the included noise.

We obtained the *AER* of the *PR* with the accuracy of 0.45% without the fluctuation of illumination and 3.73% with the fluctuation of illumination. Our method indicates the robustness to

fluctuation of illumination using two-band infrared videos since the videos are separated to hemoglobin and shading components by applying the basis translation matrix in the color vector space. As we mentioned above, we can be obtained to the higher accuracy by implementing the face tracking. Moreover, we selected the hemoglobin vector as the optimal vector to estimate the pulse rate with the highest accuracy among whole vector from 0 degree to 360 degrees. However, we obtained the low accuracy in comparison of the results of the Zeng's method. As shown in Fig. 3 (d), the hemoglobin vector can be determined incorrectly since the shading of the nose on hemoglobin components were not separated perfectly into the shading components. However, we obtained the optimal hemoglobin vector in case of the situation without fluctuation of illumination. We need to try to decompose the videos with hemoglobin vector in case of without fluctuation of illumination.

Table 1. Comparison of ground truth and estimated pulse rate in each band without varying the illumination GT (Ground Truth), EV (Estimation value), AER (Absolute Error Rate)

	GT [bpm]	EV [bpm]	AER [%]
One band 780nm	89.12	96.46	8.17
One band 900nm	89.12	95.40	6.98
Two band 780nm, 900nm	89.12	88.77	0.45

Table 2. Comparison of ground truth and estimated pulse rate in each band with varying the illumination GT (Ground Truth), EV (Estimation value), AER (Absolute Error Rate)

	GT [bpm]	EV [bpm]	AER [%]
One band 780nm	98.94	117.40	18.66
One band 900nm	98.94	69.69	29.56
Two band 780nm, 900nm	98.94	102.63	3.73

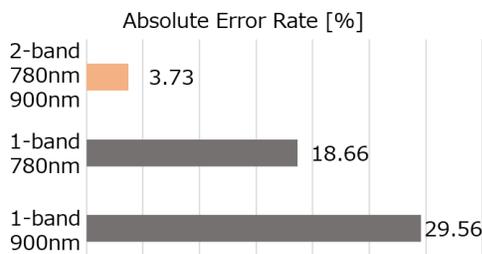


Fig. 8. Comparison of the AER in case of each band and two-band

7. Conclusions and Future Works

We proposed the non-contact pulse wave monitoring method which is robust to fluctuation of illumination. The separated components of hemoglobin and shading were obtained by determining the new basis vector in the color vector space. As shown in Fig. 8, our proposed method was greatly improved with the accuracy from 29.56% to 3.73% in comparison of the conventional method [5] by using one band infrared video.

It is necessary to improve the accuracy to estimate the pulse rate by implementing the face tracking. We obtained the optimal hemoglobin vector in case of the situation without fluctuation of illumination. We need to try to estimate the pulse rate using the optimal hemoglobin vector in case of the situation without fluctuation of illumination. Through these improvements, we attempt to measure the pulse rate variability which is robust to fluctuation of illumination using two-band infrared camera system in the situation without visible light sources as a future work.

Acknowledgements

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

References

- [1] M. Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Exp.*, vol. 18, no. 10, pp.10762–10774, May 2010.
- [2] M. Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7–11, Jan. 2011.
- [3] Kurita K, Yonezawa T, Kuroshima M and Tsumura N, "Non-Contact Video Based Estimation for Heart Rate Variability Spectrogram using Ambient Light by Extracting Hemoglobin Information," *Color and Imaging Conference*, Volume 2015, Number 1, October 2015, pp. 207-211
- [4] 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.*, vol. 54, no. 8, pp. 1418–1426, Aug. 2007
- [5] Wei Zeng, Qi Zhang, Yimin Zhou, Guoqing Xu, Guoyuan Liang. "Infrared Video based Non-Invasive Heart Rate Measurement": *IEEE Conference on Robotics and Biomimetics*, December 6-9, 2015, pp. 1041-104