https://doi.org/10.2352/CIC.2022.30.1.12

A noise-robust pulse wave estimation from NIR video using Wiener estimation method

Yuta Hino, Koichi Ashida, and Norimichi Tsumura

Graduate School of Science and Engineering, Chiba University, CHIBA, JAPAN

Abstract

In this paper, we propose a noise-robust pulse wave estimation method from NIR video. Pulse wave estimation in the near-infrared region is expected to be applied to non-contact monitoring in dark areas. Conventional method cannot consider noise when performing estimation, so the accuracy of pulse wave estimation in noisy environment is not very high. This may adversely affect the accuracy of heart rate and other data obtained from pulse waves. Therefore, the objective of this study is to perform pulse wave estimation robust to noise. The Wiener estimation method was used in this study. The Wiener estimation method is a simple linear computation that can consider noise. The proposed method is expected to enable non-contact and accurate estimation of pulse wave from near-infrared video images. Experimental results show that the proposed method estimates the pulse wave more robustly to noise than the conventional method. Furthermore, the heart rate was estimated from the estimated pulse wave and the proposed method was able to obtain a value closer to the ground truth.

1. Introduction

Recently, there has been a growing demand for non-contact monitoring technology in dark areas. Examples of applications of non-contact monitoring technology include detection of sudden changes in state during sleep and detection of drowsiness while driving at night. In order to detect a sudden change in state during sleep, it is necessary to attach a contact-type measurement device. However, some people cannot attach a contact-type device for various reasons. If physiological information such as heart rate can be obtained by non-contact monitoring in the dark, it is possible to detect abnormalities quickly even during sleep for people who cannot attach a device.

In order to detect physiological information, it is necessary to estimate the pulse wave from video images. Kurita et al. proposed a method for pulse wave estimation by separating RGB video images into melanin, hemoglobin, and shadow components for analysis [1]. However, RGB cameras have problems in capturing video images in dark areas when light levels are insufficient and in pulse wave estimation. Garbey et al. proposed a method to measure pulse waves using a single band medium wavelength infrared camera [2]. Zeng et al. proposed a method to estimate the pulse rate using single-band near-infrared camera [3]. However, single-band cameras have a problem of being easily affected by noise such as body motion. Mitsuhashi et al. proposed a pulse wave estimation method that can be used in dark areas by using 2-band near-infrared video images [4]. However, in the near infrared environment, the absorbance change of hemoglobin is small and thus the effect of noise is large. This conventional method does not consider noise for pulse wave estimation and thus the problem of acquiring a distorted waveform when estimating the pulse wave. A distorted waveform may adversely affect the accuracy of physiological information obtained from pulse waves.

In this paper, therefore, we propose a method for estimating pulse waves from 2-band near-infrared video images while considering the effect of noise. The Wiener estimation method was used for pulse wave estimation to consider noise. The Wiener estimation method can consider noise added to the video image by a simple linear operation. In section 2, we discuss skin model in the near-infrared region used in this paper. In section 3, we discuss separation of hemoglobin and shading by using Wiener estimation method. In section 4, we discuss capturing face video image and calculation of noise added to image. In section 5, we discuss estimating pulse wave signal, processing pulse signals and methods for calculating heart rate from pulse waves. In section 6, we discuss the results of pulse waveform and heart rate obtained by conventional method and proposed method. Finally, in section 7, we present some discussions and concluded our research in this paper.

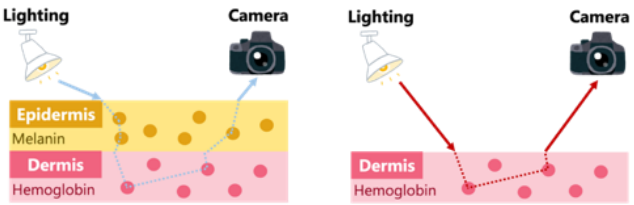
2. Skin model in the near-infrared region

Light is classified into three categories: ultraviolet light, visible light and infrared light. Infrared light has a long wavelength range and is classified into near-infrared rays, mid-infrared rays and far-infrared rays according to its wavelength range. In this study, pulse wave estimation is performed using near-infrared video images in the wavelength range of approximately 780[nm] to 1,000[nm].

Human skin is a multilayered structure that can be divided into three main parts: epidermis, dermis and subcutaneous tissue. The skin contains various pigments such as melanin, hemoglobin and bilirubin. Among these pigments, changes in melanin and hemoglobin have a significant effect on skin coloration. Melanin is found in the epidermis, while hemoglobin is found in the dermis, where capillaries are located. Kurita et al. [1] used a two-layer skin model composed of the epidermis and the dermis. They simplified the skin model by assuming the epidermis is the layer containing only melanin and the dermis is the layer containing only hemoglobin. This allows us to assume that the layer containing only melanin and the layer containing only hemoglobin are spatially independent. Therefore, under visible light illumination, human skin can be treated as a two-layer skin model consisting of a melanin layer and a hemoglobin layer, as shown in Figure 1(a). In a previous study by Kurita et al., melanin and hemoglobin components were estimated by independent component analysis after removing shading components from the observed signal in the density space.

Kurita et al. used an RGB camera; images obtained from an RGB camera contain only information obtained from visible light. On the other hand, near-infrared light has a longer wavelength than visible light. Therefore, it has a deeper penetration depth into the inside of a living body and is said to be able to measure blood vessels deep inside the body, called microvessels, which exist even deeper than capillaries [6]. Therefore, it can be assumed that when light in the near-infrared region enters the skin, reflection occurs only in the

deep skin, as shown in Figure 1(b). In other words, reflection in the epidermis is not considered, and reflection by hemoglobin near the arterioles is assumed to occur. In this study, we assumed a one-layer skin structure model as shown in Figure 1(b) and conducted our research based on this assumption.



(a) *Visible light* **(b)** *Infrared light*

Figure 1. *Skin model at different light wavelengths*

3. Hemoglobin and shade separation by using Wiener estimation method

The pixel value v_i when captured by the camera is expressed by the following equation.

$$v_i = \int \; t_i(\lambda)E(\lambda)S(\lambda)r(x,y,\lambda)d\lambda, \;\; i = 1,\dots,m, \qquad (1)$$

where $t_i(\lambda)$ is the spectral transmittance of the ith filter, $E(\lambda)$ is the spectral radiance of the illumination, $S(\lambda)$ is the spectral sensitivity of the sensor and $r(x,y,\lambda)$ is the spectral reflectance of the object at coordinates (x,y). Equation (1) can be represented graphically as shown in Figure 2.

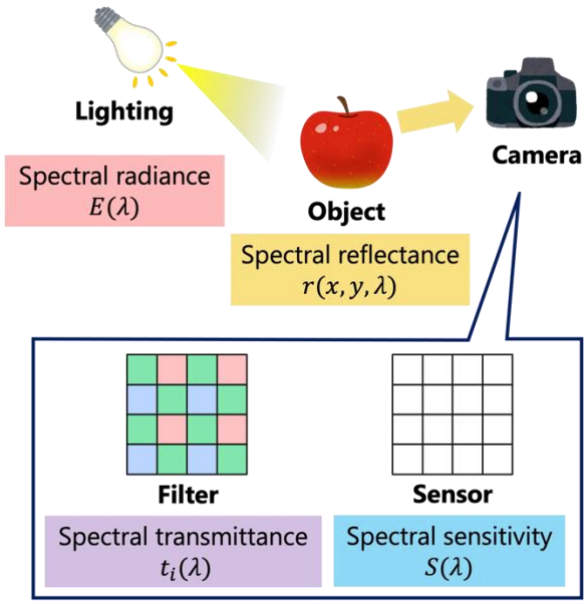


Figure 2. *Results of heart rate estimation for each method*

Expressing equation (1) in matrix form, equation (1) can be expressed as follows.

$$\boldsymbol{v} = F\boldsymbol{r}, \qquad (2)$$

where $\boldsymbol{v}$ is a vector of pixel values, F is a matrix of vectors of hemoglobin and shade components and $\boldsymbol{r}$ is a vector of object reflectance. This equation is also valid by replacing $\boldsymbol{v}$ with the spatial value of pixel values and $\boldsymbol{r}$ with a vector of hemoglobin and shadow components. In the following sections, we will use the above replacement.

Equation (2) is the equation for the case where no noise is added. However, noise is added to the image during actual image capture due to thermal noise and other causes. Therefore, when noise is added, Equation (2) can be expressed as follows.

$$\boldsymbol{v} = F\boldsymbol{r} + \boldsymbol{n}, \qquad (3)$$

where $\boldsymbol{n}$ denotes noise.

When estimating the pigment component $\boldsymbol{r}$ from the pixel value $\boldsymbol{v}$, the following equation is used for estimation.

$$\boldsymbol{r} = F^{-1}\boldsymbol{v}. \qquad (4)$$

It is noted that this estimation by using simple inverse matrix (Equation (4)) is used as conventional method in this paper. However, Equation (4) does not consider the noise added in Equation (3). Therefore, the estimated pigment components may differ significantly from the original ones.

To estimate the pigment component $\boldsymbol{r}$ from the pixel value $\boldsymbol{v}$ after considering noise, the following matrix G can be used.

$$\tilde{\boldsymbol{r}} = G\boldsymbol{v}, \qquad (5)$$

where $\tilde{\boldsymbol{r}}$ is the pigment component estimated by using Wiener estimation method. It is desirable to minimize the error between the correct value $\boldsymbol{r}$ and the estimated value $\tilde{\boldsymbol{r}}$. For this purpose, the mean squared error between $\boldsymbol{r}$ and $\tilde{\boldsymbol{r}}$ is first calculated. The mean square error e can be expressed as follows.

$$e = \langle(\boldsymbol{r} - \tilde{\boldsymbol{r}})^T(\boldsymbol{r} - \tilde{\boldsymbol{r}})\rangle, \qquad (6)$$

where $\langle\rangle$ denotes the ensemble mean for the pigment component vector. The estimation matrix that minimizes the mean squared error shown in Equation (6) is then expressed in Equation (7).

$$G = R_{rv}R_{vv}^{-1}, \qquad (7)$$

Where R_{rv} denotes the cross-correlation matrix of $\boldsymbol{r}$ and $\boldsymbol{v}$, and R_{vv} denotes the autocorrelation matrix of $\boldsymbol{v}$.

$$R_{rv} = \langle\boldsymbol{r}\boldsymbol{v}^T\rangle, \qquad (8)$$

$$R_{vv} = \langle\boldsymbol{v}\boldsymbol{v}^T\rangle. \qquad (9)$$

The autocorrelation matrix of $\boldsymbol{r}$ is expressed as in Equation (10).

$$R_{rr} = \langle\boldsymbol{r}\boldsymbol{r}^T\rangle. \qquad (10)$$

Using this, Equation (7) can also be expressed as in Equation (11).

$$G = R_{rr}F^T(FR_{rr}F^T)^{-1}. \qquad (11)$$

In this case, we assume that noise is added to the image, as shown in Equation (3). As a result, the estimation matrix is given by following equation.

$$G = R_{rr}F^T(FR_{rr}F^T + R_{nn})^{-1}, \qquad (12)$$

where R_{nn} denotes the autocorrelation matrix of the noise.

$$R_{nn} = \langle\boldsymbol{n}\boldsymbol{n}^T\rangle. \qquad (13)$$

As described above, the Wiener estimation method gives the estimation matrix that minimizes the mean-square error between the correct and estimated values by a simple linear operation when the signal and noise statics are known.

4. Capturing facial video and calculating autocorrelation matrixs

The experimental condition is shown in Figure 3. The subject, whose face was fixed with a chin rest in a dark room, was illuminated with an infrared LED light, and facial images were captured for 30[s] using a 4-band near-infrared camera at 66.5 [fps]. The artificial skin patch was captured at the same time. The artificial skin patch was used to obtain the autocorrelation matrix R_{nn} of the noise added to the captured images. Simultaneously with the imaging, a pulse wave was measured by attaching a photoelectric pulse wave meter to the tip of the index finger of the subject's left hand. The subjects were instructed to keep their faces and fingers as still as possible. Three male subjects in their 20s participated in this experiment.

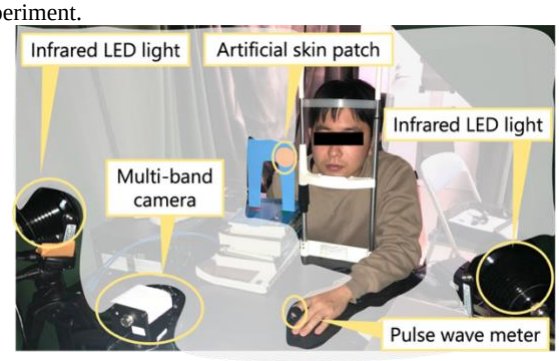
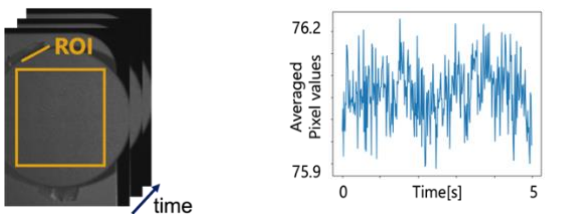


Figure 3. *Experimental Condition for capturing facial video*

As described in Section 3, the noise autocorrelation matrix R_{nn} is required when using Wiener estimation method. To obtain the noise autocorrelation matrix, it is necessary to obtain the noise added to the captured video image. In this study, the noise added to the video image was obtained by simultaneously capturing artificial skin patch with the face. Ideally, the pixel values of artificial skin patch do not change because the surface condition of the patches does not change over time. However, the pixel values of the artificial skin patch changes due to the influence of noise. This change in pixel value was used to determine the magnitude of the noise.

The region of interest was set for the image of the artificial skin patch as shown in Figure 4 and the time variation of the average pixel value was calculated within the region of interest. The standard deviation was also calculated on the time variation of the mean pixel value. The above process was performed on the two bands used for imaging and the larger standard deviation was used in subsequent procedures. Then, Gaussian noise was created by setting the mean value to 0 and the standard deviation to the obtained value, and the autocorrelation matrix of the noise R_{nn} was obtained using this noise.



(a) *Video Image of artificial skin patch* **(*b*)** average of pixel value

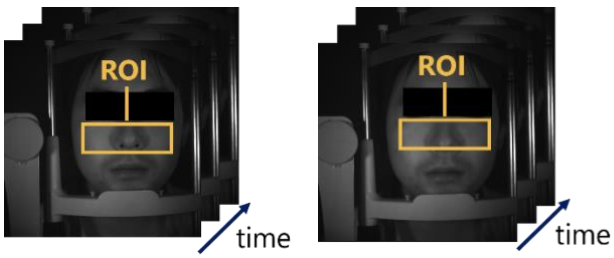
Figure 4. *Near-infrared image with a center wavelength of 800 nm*

The autocorrelation matrix of the skin pigment R_{rr} is also required when using Wiener estimation method. When calculating R_{rr}, the value that the hemoglobin and shade component vector $\boldsymbol{r}$ can take is necessary. However, since this value is unknown, we calculated the possible values of $\boldsymbol{r}$ using the pixel values obtained from a subject's face image and Equation (4). Then, R_{rr} was obtained using this value.

5. Acquisition of the original pulse wave signal and signal processing

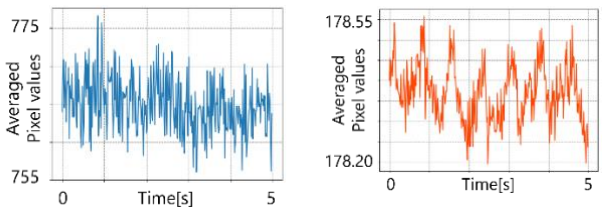
The temporal variation of the average pixel value in the region of interest was analyzed by selecting two bands from the four-based near-infrared images taken under the imaging environment described in Section 4 and setting the region of interest as shown in Figure 5. In this study, 2 bands images were used, one with a central wavelength of 800 nm and the other with a central wavelength of 930 nm. The horizontal axis was set to time variation, and the vertical axis was set to mean pixel value and the analysis was performed as the variation of mean pixel value from frame to frame. Figure 6 shows the original pulse wave signals obtained by applying the conventional and proposed methods to a subject's two-band near-infrared image.

Figure 6 shows that the original pulse wave signal obtained by the proposed method shows the peaks and valleys of the pulse wave, whereas the original pulse wave signal obtained by the conventional method is affected by noise and the waveform is distorted. This indicates that the proposed method can estimate the pulse wave robustly to noise.



(a) *Center wavelength of 800 nm* **(b)** *Center wavelength of 930 nm*

Figure 5. *Near-infrared video image of 2 bands selected from 4 bands*

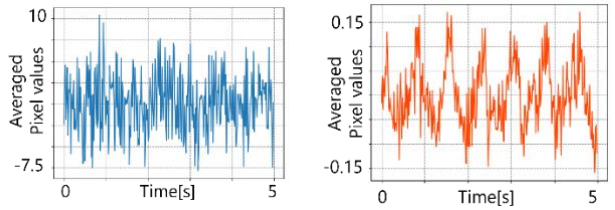


(a) *Conventional method* **(*b*)** *Proposed method*

Figure 6. *Original pulse wave signal*

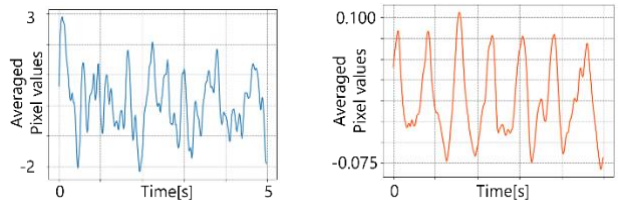
Detrending [7] was performed on the original pulse wave signal and the results are shown in Figure 7. A bandpass filter was then applied. The frequency range transmitted by the bandpass filter was

set to 0.75Hz ~ 4.0Hz [8]. Figure 8 shows the results of applying the bandpass filter.



(a) *Conventional method* **(b)** *Proposed method*

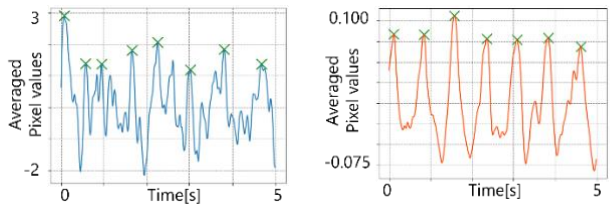
Figure 7. *Pulse wave after detrending*



(a) *Conventional method* **(b)** *Proposed method*

Figure 8. *Pulse wave after applying bandpass filter*

The upper peak points were detected by finding the local maximum values for each waveform in the bandpass-filtered pulse wave. The peak points are used to estimate the heart rate.



(a) *Conventional method* **(b)** *Proposed method*

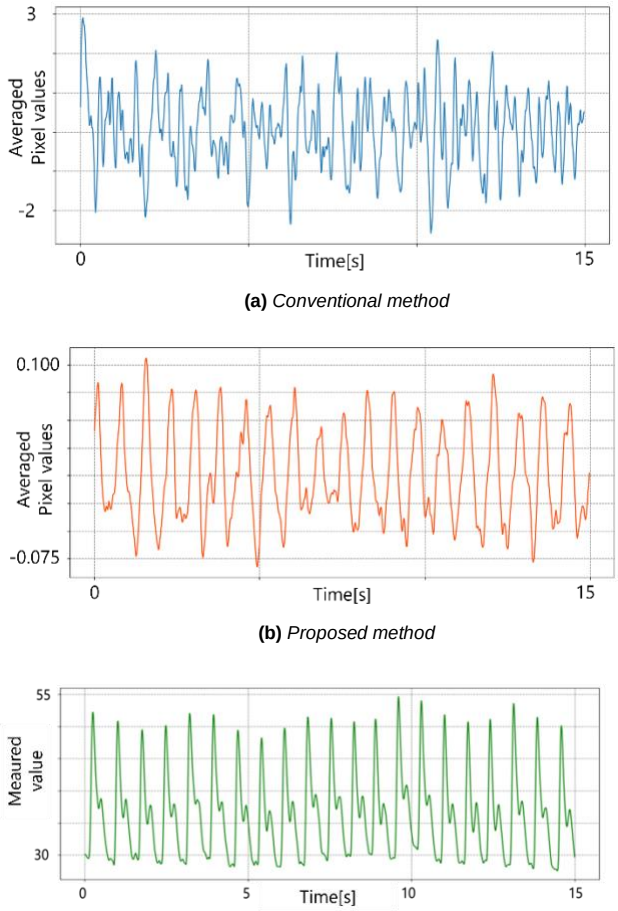
Figure 9. *Results of detecting peak points*

The pulse wave interval $RR_{interval}$ were calculated as the intervals between the neighboring peaks. Using Equation (14), the pulse rate was obtained by averaging $RR_{interval}$.

$$HR = \frac{60}{RR_{interval}} \tag{14}$$

6. Result

Pulse waves estimated by the two methods (conventional and proposed) and measured by an attached pulse wave meter were obtained from three subjects. Figure 10 shows the pulse waveform of one of the three subjects. This figure shows that the pulse waveform obtained by the proposed method is closer to the ground truth.



(a) *Conventional method*

(b) *Proposed method*

(c) *Ground truth*

Figure 10. *Estimated pulse waves and ground truth*

Heart rate was estimated from the pulse waves obtained by the two methods and the pulse wave obtained by the pulse wave meter in three subjects, and a comparison was performed. Figure 11 shows the results of heart rate estimation for the three subjects. This figure shows that the heart rate estimated from the pulse wave obtained by the proposed method is closer to the ground truth than that estimated from the pulse wave obtained by the conventional method.

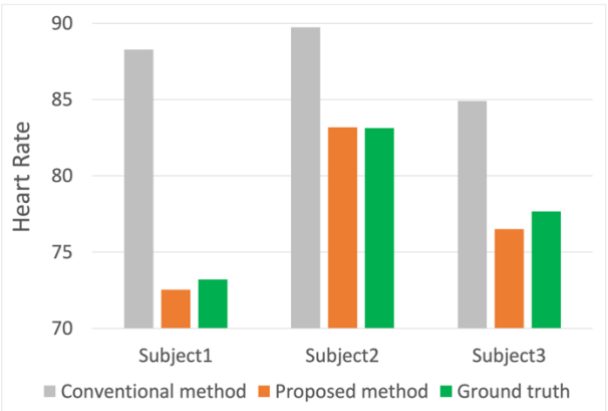


Figure 11. *Results of heart rate estimation for each method*

7. Conclusion and Future works

In this study, we proposed a noise-robust pulse wave estimation method from RGB video using Wiener estimation method. We compared the proposed method with a conventional method in a near infrared environment. For pulse wave estimation in the near-infrared environment, pulse waves can be obtained by separating hemoglobin and shading in a two-band face video image. While the conventional method uses the inverse matrix of the hemoglobin and shade component vectors, Wiener estimation method approximates the inverse matrix used in the conventional method by using the autocorrelation matrix of the pigment components and the autocorrelation matrix of the noise. Therefore, the estimation matrix obtained by Wiener estimation method can consider noise, whereas the conventional method does not have any information about noise in the matrix used for component separation.

The original pulse wave signals were obtained from 2-band near-infrared video images selected from face videos captured with a near-infrared 4-band camera by using both the conventional and the proposed methods. After applying a detrending and a bandpass filter to the original pulse wave signal, the pulse wave interval was obtained by peak detection to estimate the heart rate. As a result, the pulse waveform obtained by the proposed method was closer to the ground truth than that obtained by the conventional method. In addition, the heart rate estimated by the proposed method had a smaller error than that estimated by the conventional method. In conclusion, the proposed method can estimate the pulse wave robustly to noise.

Although the results in the near-infrared environment were better than those obtained with the conventional method, it is necessary to verify the use of this method in environments other than the near-infrared environment, such as RGB photography. It is also necessary to verify the method when the noise is larger than the noise added in this experiment.

References

[1] 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, Volume 2015, Number 1, October 2015, pp.207-211

[2] M. Garbey, N.Sun, A.Merla and I.Pavlidis, “Contact-Free Measurement of Cardiac Pulse Based on the Analysis of Thermal Imagery”, IEEE Transactions on Biomedical Engineering, July 2007, pp.1418-1426

[3] W. Zeng, Q.Zhang, Y.Zhou, G.Xu and G.Liang, “Infrared Video based Non-invasive Heart Rate Measurement”, IEEE Conference on Robotics and Biomimetics, December 2015, pp.1041-1046

[4] R. Mitsuhashi, G.Okada, K.Kurita, S.Kawahito, C. Koopipat and N. Tsumura, “Noncontact pulse wave detection by two-band infrared video-based measurement on face without visible lighting”, Artificial Life and Robotics, April 2018, pp.345-352

[5] N. Tsumura, H.Haneishi and Y.Miyake, “Estimation of Spectral Reflectances from Multi-Band Images by Multiple Regression Analysis”, Japanese Journal of Optics, Vol.27, no.7, pp.384-391, 1998

[6] R.rox Anderson, B.S. and John A.Prrish M.D., “The optics of human skin”, The journal of investigate dermatology, pp.13-19, 1981

[7] M. P. Tarvainen, P. O. Ranta-aho and P. A. Karjalainen, “An advanced detrending method with application to hrv analysis”, Biomedical Engineering, IEEE Transactions on, vol.49, no.2, pp.172-175, 2002.

[8] 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.10762-10774, 2010.