This CVPR 2021 workshop paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

Oxygen Saturation Estimation Based on Optimal Band Selection from Multi-band Video

Ryo Takahashi
Graduate School of Advanced Integration Science, Chiba University
1-33 Yayoi, Inage, Chiba, Japan
takahashi@chiba-u.jp

Koichi Ashida
Graduate School of Advanced Integration Science, Chiba University
1-33 Yayoi, Inage, Chiba, Japan

Yasuo Kobayashi
ASTRODESIGN, Inc.
1-5-2, Minami-yukigaya, Ota-ku, Tokyo, Japan

Rumi Tokunaga
Graduate School of Global and Transdisciplinary Studies, Chiba University
1-33 Yayoi, Inage, Chiba, Japan

Shuhei Kodama
ASTRODESIGN, Inc.
1-5-2, Minami-yukigaya, Ota-ku, Tokyo, Japan

Norimichi Tsumura
Graduate School of Engineering, Chiba University
1-33 Yayoi, Inage, Chiba, Japan

Abstract

In this study, we propose a method to estimate oxygen saturation by selecting the best bands from video images captured by a multiband camera. Oxygen saturation is one of the most important bioindicators for measuring human health. For example, when a person contracts COVID-19, which is currently prevalent, oxygen uptake does not work properly and oxygen saturation drops without the person being aware of it, which may lead to severe symptoms. Monitoring oxygen saturation is very important so that the person receives treatment before such a situation occurs. The commonly used contact sensor is uncomfortable because of its pressure and it is difficult to wear on a daily basis, so non-contact estimation of oxygen saturation is desirable. To estimate oxygen saturation using a contact sensor, the difference in the absorption coefficients of oxidized hemoglobin and deoxidized hemoglobin is used. Using the same principle, it is possible to estimate oxygen saturation without contact using the signals from two channels obtained by an RGB camera. Currently, many smartphones are equipped with infrared cameras for face recognition, and increasingly more models are equipped with multi-camera systems consisting of RGB and infrared cameras. In such cases, it is difficult to take advantage of the multiple bands because the optimal combination of bands for oxygen saturation estimation varies depending on the imaging environment and the subject. In this study, to select the optimal combination of bands from multi-band video images, we used a Monte Carlo simulation of light scattering on the skin to simulate pulse waves during oxygen saturation changes while measuring the signals with a multi-band camera. We further propose a method to select the most accurate combination for estimating the oxygen saturation based on the features obtained from the pulse wave.

1. Introduction

One of the most important biological indicators of human health is oxygen saturation. Oxygen saturation is a marker that indicates the percentage of oxygen combined with hemoglobin contained in the red blood cells of the blood carried from the heart to the entire body. It is used as an indicator to diagnose respiratory failure because the oxygen saturation decreases when the lungs and heart lose their ability to take in oxygen. An example of the importance of oxygen saturation is the monitoring of patients with COVID-19, which is currently prevalent. When a person is infected with the SARS-CoV-2 virus, oxygen uptake does not work properly, but the person is unaware of this fact and oxygen saturation drops, which may lead to a serious condition. Monitoring oxygen saturation is very important to determine the symptoms so that the person may receive treatment before this situation occurs. The current mainstream method of measuring oxygen saturation is to use a contact sensor, which is uncomfortable due to its pressure, and it is difficult to live with this device on a daily basis. Given this situation, the non-contact estimation of oxygen saturation is desirable. To estimate oxygen saturation, a non-contact method has been proposed in which pulse waves are acquired from the R and B channels of an RGB camera and oxygen saturation is estimated using a method equivalent to that of a contact pulse oximeter. In this paper, we propose a method of estimating oxygen saturation using multi-band video images to obtain pulse waves from two channels in the same way as a contact pulse oximeter. Mobile devices such as iPhones and other smartphones are frequently equipped with infrared cameras for face recognition. In addition, sensors with a four-band Bayer array (RGB plus infrared) have been developed for use in surveillance cameras for both daytime and nighttime, and multiband cameras are expected to become more popular in the future. The use of such multiband cameras is expected to enable the estimation of oxygen saturation in various environments because the number of bands is larger than that of RGB cameras. However, when estimating oxygen saturation using a multiband camera, it is difficult to take full advantage of the multiple bands because the best combination of two bands for estimating oxygen saturation among the four bands varies depending on the imaging environment and the subject. Therefore, the purpose of this study is to construct a method to select the optimal combination of bands from the bands available in multi-band video images.

2. Non-contact method of oxygen saturation estimation

In the conventional method, oxygen saturation is estimated using a contact sensor. However, recent studies have proposed a method for estimating oxygen saturation using a camera.

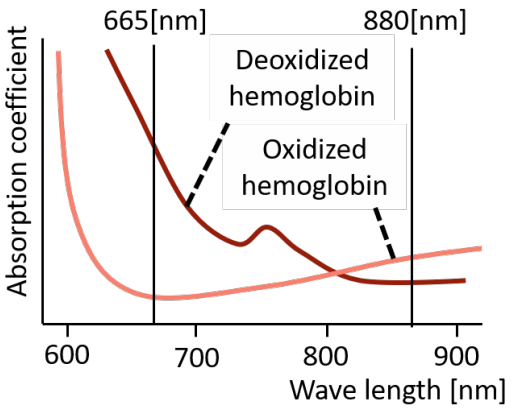

Figure 1. Absorption coefficients of Hb and HbO2

2.1. Principle of oxygen saturation estimation

The amplitude of a pulse wave, which indicates the change in arterial volume that occurs with the beating of the heart, changes with respect to oxygen saturation. This pulse wave variation is derived from the color of the blood. In blood, there are two types of hemoglobin: oxidized hemoglobin (HbO2), which is bound to oxygen, and deoxidized hemoglobin (Hb), which is not combined with oxygen. As shown in Figure 1, the absorption coefficients of these two types of hemoglobin differ according to the wavelength of light. Therefore, as shown in Figure 2, when the oxygen saturation decreases (HbO2 decreases and Hb increases), the amplitude of the pulse wave obtained by red light increases and the baseline decreases in the case of same blood volume. Conversely, the amplitude decreases and the baseline increases when infrared light is used. Therefore, the ratio of the AC and DC components of the pulse wave (ACDC ratio) is calculated, and the ratio of the ratio (RoR) calculated by dividing the ACDC ratio of red light by the ACDC ratio of infrared light is an index that correlates with oxygen saturation. The relationship between this RoR and oxygen saturation can be directly estimated by collecting data through repeated pulse wave measurement experiments while varying oxygen saturation and performing regression using the following equation.

$$\mathrm{SpO}_2 = \alpha \mathrm{S} + \beta \tag{1}$$

In Equation (1), S represents the RoR, and α and β represent the regression coefficients.

2.2. Camera-based method for measuring oxygen saturation

Verkruysse et al. reported that it is possible to obtain pulse waves by calculating the change over time of the aver-

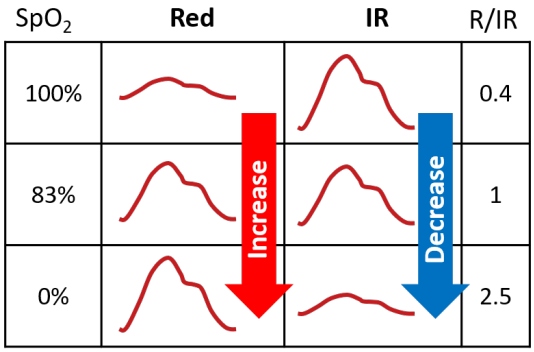

Figure 2. Pulse wave variation in red and infrared light

age pixel value of the G signal in a video taken with an RGB camera [1]. Currently, various methods have been proposed to extend this method, such as methods using principal component analysis and methods based on independent component analysis, which can robustly estimate pulse waves in response to illumination light and body motion [2, 3, 4, 5]. Of the RGB wavelengths that could be used for measurement, the most basic method by Verkruysse et al. uses the G signal. This is because green light has a larger absorption coefficient than red or blue light and is the most suitable for pulse wave measurement. In pulse wave acquisition using an RGB camera, this method is the most common because it requires the least amount of computation and can be implemented easily. In addition, Guazzi et al. reported that the non-contact estimation of oxygen saturation was possible using a two-band pulse wave acquired using an RGB camera with a method similar to that of the estimation using a contact sensor [6]. In this study, we apply this method with a multi-band camera to estimate oxygen saturation in a way that is robust to the lighting environment and other factors.

3. Optimal band selection for multi-band video

The proposed method selects the optimal bands using a regression model that is created from the optimal band selection features for pulse waves generated by computer simulation. When estimating the oxygen saturation from the pulse waves obtained from two bands, for a camera with n bands, a combination of two pulse waves is selected from the n pulse waves, and a method of calculating the ${}_nC_2$ pattern RoR is considered. Of these ${}_nC_2$ patterns of RoR, the optimal pattern for oxygen saturation estimation is expected to vary depending on the imaging environment and the subject. One reason for this is that the amount of noise in the pulse wave obtained from each band varies depending on the spectral distribution of the illumination and the characteristics of the subject's skin. Therefore, in this study, signal-to-noise ratio (SNR) is used as a feature to measure the amount of noise in the two pulse waves obtained from the two bands. If the spectral sensitivities of the two bands are close each other, the RoR change when the oxygen saturation changes tend to be small value. This is because the change in ACDC ratio of the two bands becomes similar when their spectral sensitivities are close each other. Therefore, as a feature that indicates the degree to which a pair of bands of the camera is suitable for oxygen saturation estimation, we use the difference in the RoR (normalized at 100% oxygen saturation) when the oxygen saturation decreases, which is calculated from the spectral sensitivity of the two bands. To select the best band using these features, a regression model is created using the correlation coefficient between the RoRs of the two bands and the oxygen saturation as the objective variable. This regression model is used to estimate the correlation coefficients of each of the ${}_nC_2$ pattern pulse wave combinations in the multi-band video image, and the optimal bands are the combination that yields the highest correlation.

3.1. Simulation of pulse waves on a computer

To select the optimal bands, it is necessary to investigate what characteristics of the pulse wave are optimal for oxygen saturation estimation. In this study, we used the Monte Carlo model of light transport in multi-layered tissues (MCML) to simulate the pulse waves during changes in oxygen saturation [7]. It is necessary to set the parameters of the skin to use MCML. The melanin concentration and hemoglobin concentration were assumed to vary by 3% and 0.4%–0.6%, respectively, and the oxygen saturation was assumed to vary by 80%–100%. For the spectral reflectance change of the skin obtained by MCML, the spectral sensitivity of the camera was set to 51 patterns from 400–900 nm in 10 nm steps, and the pulse waves were sampled. The sampling process is illustrated in Figure 3. In addition, the actual measured pulse wave is subject to noise. Noise from the camera's sensor and noise from changes in ambient light and other light sources are examples in this study. However, we focus on the noise caused by the camera's sensor, without considering the noise caused by changes in light in this paper, since we are using only Gaussian noise. For this reason, a white Gaussian noise was added to the pulse wave. An outline of this process is shown in Figure 4. The processing described in the following sections was performed on the pulse waves obtained through this simulation.

3.2. Simulation of pulse waves on a computer

In this study, the SNR is used as a measure of the amount of noise in a pulse wave. The SNR is the ratio of signal to noise; if the SNR is high, the influence of noise is small, and if it is low, the influence of noise is large. In general, the SNR is expressed using the common logarithm. If the power of the signal is PS and the power of the noise is PN,

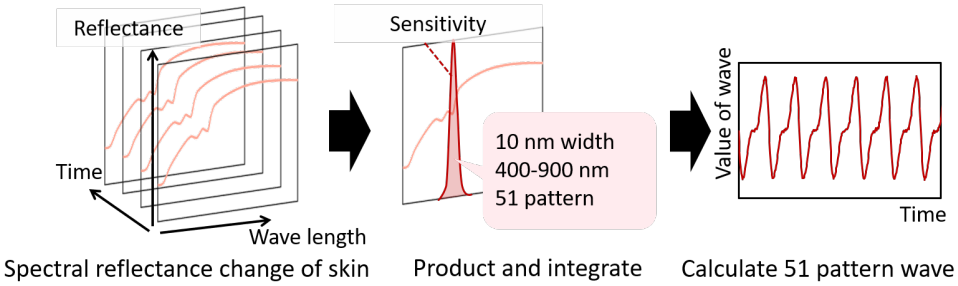

Figure 3. Overview of pulse wave simulation

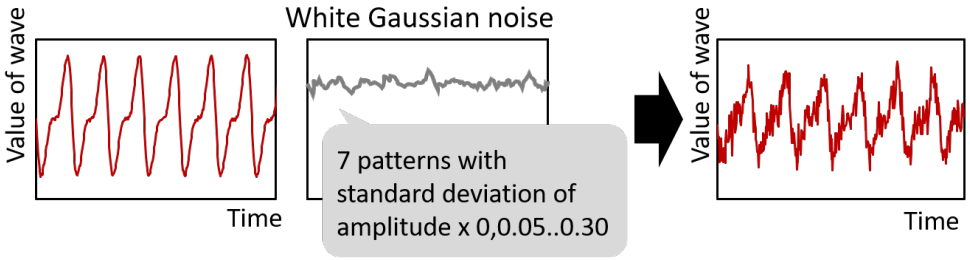

Figure 4. Adding noise to the simulated pulse wave

the SNR can be calculated as follows.

$$\mathrm{SNR} = 10\log\left(\frac{\mathrm{P_S}}{\mathrm{P_N}}\right) \tag{2}$$

In this study, the human heart rate is set to be 30–180 bpm (0.5–3.0 Hz), the peak of the frequency component in the interval of 0.5–3.0 Hz] is defined to be the heart rate, and the sum of the frequency components in the interval of that frequency ±0.1 Hz] is considered to be the signal component. To take harmonics into account, the second and third harmonics of the heart rate frequency are also considered as signal components, and the sum of the frequency components in the interval of the frequency of the second and third harmonics ±0.1 Hz is also considered to be a signal component. The noise component is the frequency component other than the signal component in the range of 0.5–10 Hz. Figure 5 shows how the signal and noise components were determined. The difference between the normalized RoR when the oxygen saturation is 100% and when the oxygen saturation is 80% is used as a feature to indicate how suitable a pair bands of the camera is for oxygen saturation estimation. The larger this difference, the larger the change in RoR when the oxygen saturation decreases, indicating that the combination is suitable for oxygen saturation estimation. A larger difference in the ACDC ratios of the two pulse waves when the oxygen saturation changes means that the difference in normalized RoR is larger, and these bands are hence more suitable for estimating oxygen saturation. Conversely, if the difference in the ACDC ratio of the two pulse waves when the oxygen saturation changes is small, the difference in the normalized RoR is small, and these bands are not suitable for oxygen saturation estimation. To calculate the difference in normalized RoR, MCML is used in the same way as it is in pulse wave simulation. The calculation of the difference in normalized RoR is presented in Figure 6, where the spectral reflectance of the skin, which is the result of six simulations, is calculated by combining three patterns of hemoglobin concentration (0.4%, 0.5%, and 0.6%) and two patterns of oxygen saturation (80% and 100%) using MCML. The AC component corresponding with the amplitude of the pulse wave is obtained by subtracting the signal value obtained for 0.6% hemoglobin concentration from the signal value obtained for 0.4% hemoglobin concentration. Although the hemoglobin concentration varies depends on person, it is generally about 0.2-1.0%, therefore, in this study, the baseline of the hemoglobin concentration is set to be 0.5%[8] as the typical value. In the a future issue, it is necessary to verify the change of baseline of hemoglobin concentration with other values because it varies depends on person. The ACDC ratio is calculated by dividing this AC component by the DC component. The ACDC ratio can be calculated for each pair of bands, and the RoR can be calculated by dividing the ACDC ratio of the second band by the ACDC ratio of the first band. This RoR is calculated at both 100% and 80% oxygen saturation, and the difference in normalized RoR, $\mathrm{d\hat{R}oR}$, which is the difference between the RoR at 100% and the RoR at 80%

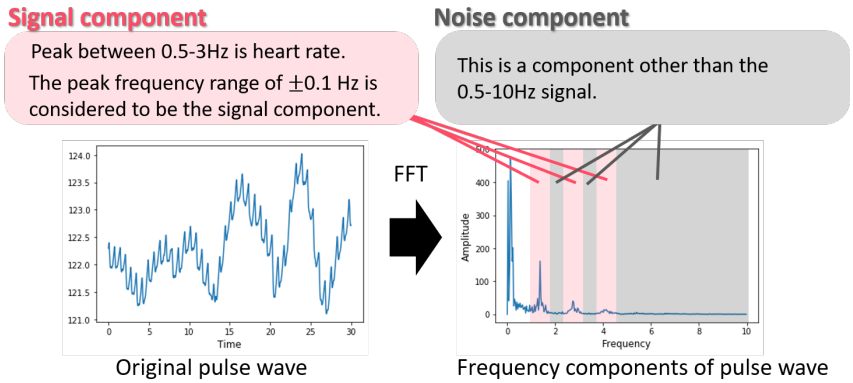

Figure 5. SNR calculation for the pulse wave

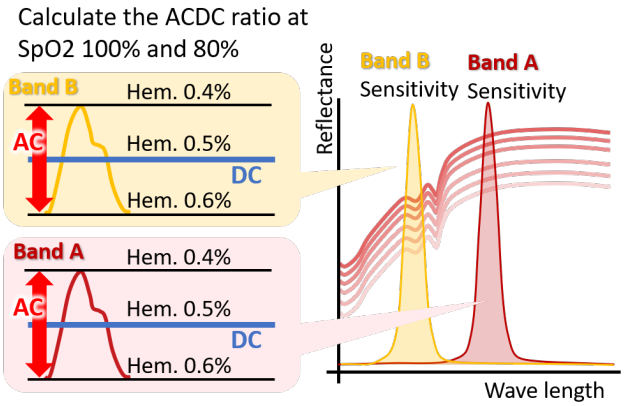

Figure 6. Overview of the difference in normalized RoR calculation

normalized by the RoR at 100%, is calculated as follows.

$$\mathrm{d\hat{R}oR} = \frac{\mathrm{RoR}_{100\%} - \mathrm{RoR}_{80\%}}{\mathrm{RoR}_{100\%}} \tag{3}$$

If the difference in normalized RoR is positive, the RoR at 100% oxygen saturation is greater than the RoR at 80% oxygen saturation. Therefore, when the oxygen saturation is 100%, the RoR is large, and when the oxygen saturation is 80%, the RoR is small, therefore it has a positive correlation with the oxygen saturation.

3.3. Calculation of the correlation coefficient between the RoR and oxygen saturation

The correlation coefficient between the RoR and oxygen saturation is the objective variable used to create the regression model. In this study, 51 pulse wave patterns were simulated, and the number of all combinations of two of the 51 patterns is ${}_{51}\mathrm{C}_2 = 1,275$ combinations. Because seven patterns of noise were added to these two pulse waves, a total of $1,275 \times 7 \times 7 = 62,475$ patterns of pulse waves were obtained. From these patterns, the RoR was calculated and its correlation coefficient with respect to oxygen saturation was calculated. There are various methods for calculating the AC and DC components of the pulse wave used in the RoR, but in this study, the pulse wave envelope and trend are used. The trend is used as the DC component, and the method proposed by Tarvainen et al. is used to calculate it [9]. Next, after the pulse wave has been detrended, a bandpass filter is applied to the pulse wave to remove the high-frequency noise component that was not removed by detrending. The bandpass filter passes signals with frequencies in the range of 0.5–10 Hz. The AC component is the periodic variation component of the pulse wave, and can be calculated by determining the pulse wave amplitude. To determine this amplitude, the pulse wave envelope is calculated for the upper and lower sides of the pulse wave. By subtracting the lower envelope from the upper envelope, the amplitude is determined, and this is the AC component. The AC component is then divided by the DC component to obtain the ACDC ratio, and the ACDC ratio of the first band is divided by the ACDC ratio of the second band to calculate the RoRs. The correlation coefficient between the calculated RoR and the oxygen saturation is used obtain the objective variables necessary for creating the model. Figure 7 shows an overview of the calculation of the RoR. In Figure 7, the frame on the x-axis represents the number of the image since this graph is the result of processing one image at a time.

3.4. Development of model for optimal band selection

A regression model was created using the three information variables (the SNR of the first band, the SNR of the second band, and the difference in normalized RoR) as explanatory variables and the correlation coefficient between the RoR and oxygen saturation as the objective variable.

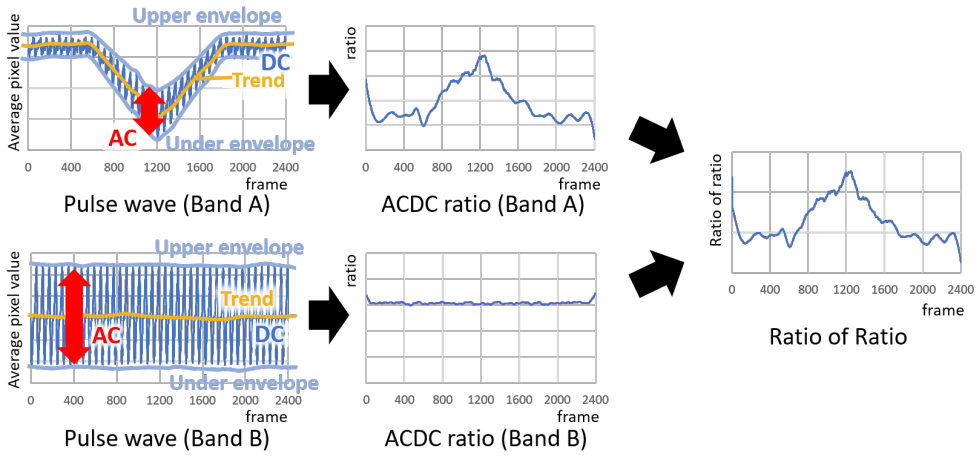

Figure 7. Overview of the calculation of the RoR

The data obtained from the simulation were divided into two parts, one for model creation and the other for testing, and the model was created and verified. Support vector regression (SVR) was used for the regression. To implement SVR, we used scikit-learn, a Python library for machine learning that provides the radial basis function (RBF) as the default function kernel for SVR. In this study, we chose this RBF to create the regression model. The tuning parameter C which determines how much misclassification is allowed, and parameter gamma, which specifies the size of the RBF kernel, can also be set to determine the degree of error tolerance. The optimal values of these parameters C and γ depend on the data set and need to be adjusted by cross-validation or similar procedures. In this study, cross-validation and grid search were used to optimize the parameters. Figure 8 shows the optimization process for the parameters by grid search and the verification procedure using the test data. Grid search is a method to determine the optimum parameters of the model. In this method, the coefficient of determination R^2 is obtained for all combinations while changing the parameters, and the parameter with the best result is adopted. The coefficient of determination was used as the evaluation index of the model for optimization in this process. The model is validated using test data that is not used for parameter optimization.

4. Model development and validation results

Table 1 shows the results of the verification with the test data for each parameter. A graphical representation of the results is shown in Figure 9. For the model with the highest coefficient of determination (0.8664), the mean absolute error is 0.1329, indicating that the correlation coefficient can be estimated with good accuracy. This result suggests that the optimal bands can be selected by choosing the pair with

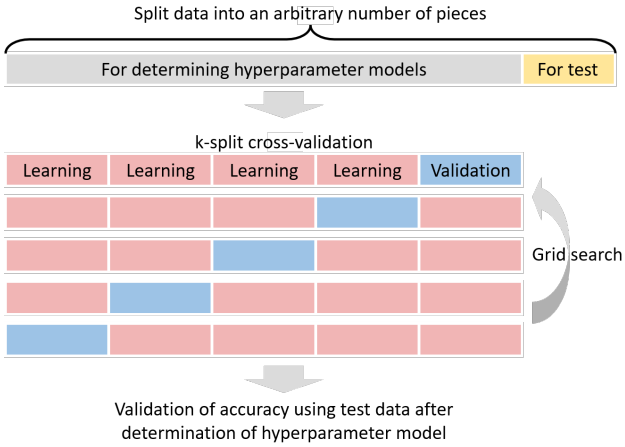

Figure 8. Optimization of the parameters by grid search, and the verification procedure

the highest correlation coefficient from all possible pairs. This approach should enable oxygen saturation to be accurately estimated.

5. Conclusion and future work

In this study, we proposed a method to select the optimal bands for oxygen saturation estimation from multiband video images. Pulse waves were generated by simulation, and the information necessary for selecting the optimal band was obtained from the pulse waves. From this information, a model was developed to estimate the correlation coefficient between the oxygen saturation and RoR. The model was validated using test data to check the prediction accuracy and confirm its effectiveness. In future work, we will need to verify the applicability of this method to real data by conducting actual measurement experiments. It is

Table 1. Verification results using the data for testing for each parameter

C	γ	R^2	C	γ	R^2
0.01	0.0001	0.0026	10	0.0001	0.7137
0.01	0.001	0.0582	10	0.001	0.7472
0.01	0.01	0.4168	10	0.01	0.8215
0.01	0.1	0.7790	10	0.1	0.8475
0.01	1	0.5784	10	1	0.8544
0.01	10	0.1348	10	10	0.7197
0.1	0.0001	0.0587	100	0.0001	0.7090
0.1	0.001	0.4436	100	0.001	0.8064
0.1	0.01	0.7537	100	0.01	0.8246
0.1	0.1	0.8305	100	0.1	0.8519
0.1	1	0.8366	100	1	0.8307
0.1	10	0.5095	100	10	0.6365
1	0.0001	0.4458	1000	0.0001	0.7448
1	0.001	0.7190	1000	0.001	0.8148
1	0.01	0.8096	1000	0.01	0.8276
1	0.1	0.8366	1000	0.1	0.8534
1	**1**	**0.8597**	1000	1	0.7680
1	10	0.7514	1000	10	0.3654

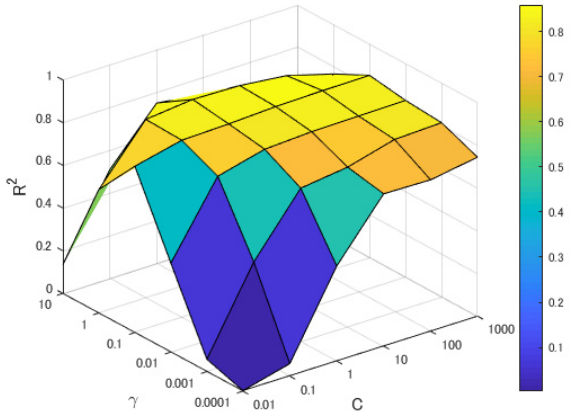

Figure 9. Verification results using the test data for each parameter

necessary to conduct simulations assuming an environment where ambient light and other factors change with more realistic noise statics.

References

[1] Verkruysse, W., Svaasand, L. O., and Nelson, J. S. Remote plethysmographic imaging using ambient light. Optics Express, 16(26):21434–21445, 2008. 3

[2] Yang, Y., Liu, C., Yu, H., Shao, D., Tsow, F., and Tao, N. Motion robust remote photoplethysmography in CIELab color space. Journal of Biomedical Optics, 21(11):117001, 2016. 3

[3] McDuff, D., Gontarek, S., and Picard, R. W. Improvements in remote cardiopulmonary measurement using a five band digital camera. IEEE Transactions on Biomedical Engineering, 61(10):2593–2601, 2014. 3

[4] Poh, M. Z., McDuff, D. J., and Picard, R. W. Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Transactions on Biomedical Engineering, 58(1):7–11, 2010. 3

[5] Shao, D., Yang, Y., Liu, C., Tsow, F., Yu, H., and Tao, N. Noncontact monitoring breathing pattern, exhalation flow rate and pulse transit time. IEEE Transactions on Biomedical Engineering, 61(11):2760–2767, 2014. 3

[6] Guazzi, A. R., Villarroel, M., Jorge, J., Daly, J., Frise, M. C., Robbins, P. A., and Tarassenko, L. Non-contact measurement of oxygen saturation with an RGB camera. Biomedical Optics Express, 6(9):3320–3338, 2015. 3

[7] Wang, L., and Jacques, S. L. Monte Carlo modeling of light transport in multi-layered tissues in standard C. The University of Texas, MD Anderson Cancer Center, Houston, 4–11, 1992. 3

[8] Nishidate, I., Maeda, T., Niizeki, K., & Aizu, Y. (2013). Estimation of melanin and hemoglobin using spectral reflectance images reconstructed from a digital RGB image by the Wiener estimation method. Sensors, 13(6), Nishidate7902-7915. 4

[9] Tarvainen, M. P., Ranta-Aho, P. O., and Karjalainen, P. A. An advanced detrending method with application to HRV analysis. IEEE Transactions on Biomedical Engineering, 49(2):172–175, 2002. 5