

Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

Near-infrared noninvasive blood glucose prediction without using multivariate analyses: introduction of imaginary spectra due to scattering change in the skin

Katsuhiko Maruo
Yukio Yamada

Near-infrared noninvasive blood glucose prediction without using multivariate analyses: introduction of imaginary spectra due to scattering change in the skin

Katsuhiko Maruo^{a,*} and Yukio Yamada^b

^aPanasonic Healthcare Co., Ltd., Corporate Research & Development Center, 2131-1, Minamigata, Toon City, Ehime 791-0395, Japan

^bThe University of Electro-Communications, Brain Science Inspired Life Support Research Center, 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan

Abstract. A noninvasive measurement method is proposed and examined to continuously predict blood glucose contents using near-infrared diffuse reflection difference spectra measured at the skin tissue without using multivariate analyses. Using the modified Beer's law, the difference spectra are assumed to be synthesized from four major components in the human skin (water, protein, glucose, and fat) and a scattering equivalent component called baseline. As a result, one of the origins of the errors in blood glucose prediction using near-infrared is found to be the similarity of the shapes of the absorption spectrum between glucose and baseline. After separating the glucose contributions from the difference spectra at the characteristic wavelengths of baseline and fat, an imaginary component combining baseline and fat is introduced by considering that both the change in the fat contribution and the generation of baseline originate from the change in scattering in the skin. The imaginary component enables us to reduce the errors in blood glucose prediction. In contrast to the methods using multivariate analyses, the calculation process of the blood glucose contents from the measured reflection spectra is clear in this method, thus, it is easy to estimate the origins of the changes and contributions of the components in the measured difference spectra. The proposed method may become a useful tool for realization of noninvasive blood glucose prediction using near-infrared spectroscopy. © 2015 Society of Photo-Optical Instrumentation Engineers (SPIE)

[DOI: [10.1117/1.JBO.20.4.047003](https://doi.org/10.1117/1.JBO.20.4.047003)]

Keywords: noninvasive measurement; blood glucose; near-infrared spectroscopy; diffuse reflectance; modified Beer's law; difference absorbance spectra.

Paper 140838PR received Dec. 16, 2014; accepted for publication Mar. 9, 2015; published online Apr. 10, 2015.

1 Introduction

Although noninvasive blood glucose prediction using near-infrared (NIR) spectroscopy has been studied for many years at various companies and research laboratories in the world, there is no medical device approved by the Food and Drug Administration of the United States at present.¹⁻³ One of the technical problems is that the glucose signal contained in the NIR reflection spectra from living tissue is significantly small compared to the signals of other components in living tissue, such as water, protein, fat, etc.^{4,5} Therefore, it is very difficult to prevent many kinds of disturbances (caused by the changes in the contents of the components other than glucose) from being present in the measured spectra and also very difficult to secure a prediction accuracy sufficient for clinical applications. It is common that NIR spectroscopy for noninvasive blood glucose prediction applies chemometrics, which uses multivariate analyses to obtain calibration models. Generally speaking, in chemometrics, a researcher should collect spectra as many as times possible with pre-experiments, and a set of the spectra must contain various disturbances covering wide ranges of their changes. Then the set of the spectra is used for a multivariate analysis to obtain a robust and accurate calibration model.⁶ However, collecting spectra as many times as possible for that purpose

adversely impairs the building of a good calibration model for accurate prediction. In particular, for blood glucose prediction, due to the poor signal-to-noise ratios of the measured spectra against the variation of glucose content, which is essentially very small in tissues, the measureable number of disturbances and the measurable ranges of their changes are limited. Then, to reduce the number of disturbances, many researchers have tried to measure spectra under constrained or controlled conditions, such as building calibration models for individuals, improving the reproducibility of the positioning of sensors, controlling the skin temperature, etc.^{7,8}

Multivariate analyses enable us to automatically and easily build calibration models that predict an objective variable by removing the influences of the disturbances. However, when the calibration models built by multivariate analyses are used, the processes to predict the blood glucose contents from measured NIR spectra are so-called black boxes. Therefore, many of the former studies using multivariate analyses have reported only the prediction accuracy of the blood glucose content, and have not evaluated the results from a spectroscopic point of view with consideration of the glucose signals or disturbances contained in the measured NIR spectra. As a result, unfortunately, trial and error studies with concepts or methods similar to those already found to be unsuccessful by former studies have been

*Address all correspondence to: Katsuhiko Maruo, E-mail: maruo.katsuhiko@jp.panasonic.com

repeatedly conducted. This fact might be one of the reasons for the delayed realization of noninvasive blood glucose prediction by NIR spectroscopy.

Arnold et al.⁹ applied a net analyte signal (NAS) method to noninvasive NIR glucose prediction for evaluation of glucose signals and disturbances contained in the measured spectra. They obtained NIR transmission spectra by *in vivo* experiments using rats and by *in vitro* experiments using aqueous glucose solutions in the wavelength range from 2040 to 2380 nm and predicted the glucose contents from the spectra. They showed that the calculated NAS vectors agreed well with the regression vectors obtained from the partial least squares (PLS) analysis which correctly predicted glucose content, and they concluded that analysis of the regression vectors could provide valuable insights into the chemical basis of selectivity for multivariate calibration models.

Alexeeva and Arnold¹⁰ made a heterogeneity distribution map of the ingredients contained in rat skin tissue, i.e., water, collagen, fat, keratin, etc., by a microspectroscopy technique measuring NIR transmission spectra in the wavelength range from 2040 to 2380 nm. They concluded that such a heterogeneity resulted in a significant disturbance for the blood glucose prediction and that appropriate corrections were needed to reduce the influence of tissue heterogeneity for accurate noninvasive glucose prediction in the future.

Maruo et al.^{11,12} proposed a unique method for building a calibration model from NIR diffuse reflectance spectra obtained numerically by a Monte Carlo simulation using the absorption and scattering coefficients of living tissues. Their method made it possible to arbitrarily numerically incorporate the disturbances into the spectra by properly adjusting the optical properties according to the changes in the contents of tissue components and other disturbing factors. They generated a total of 64 NIR diffuse reflectance spectra of human skin in the wavelength range from 1400 to 1850 nm by changing the glucose, water, fat, and protein contents as well as the scattering coefficient and temperature in the skin. A calibration model for noninvasive blood glucose prediction was built by a multivariate analysis from the 64 spectra, and the performance of this method was examined with clinical measurements. It is suggested that the major factors disturbing the spectra were water, fat, protein, scattering, and temperature changes in the skin tissue.

Toward the realization of noninvasive glucose prediction, we believe that the origins and magnitudes of the disturbances contained in the measured spectra should be understood and that the disturbances should be removed by calibration models obtained by multivariate analysis, or by controlling the measurement conditions, or by other methods.

The purpose of this study is to develop a method that can clearly show the process of predicting the blood glucose contents from measured spectra for understanding the origins and magnitudes of the disturbances and which can successfully trace and predict the changes in the blood glucose content. The blood glucose prediction method developed in this study is similar to the methods based on the conventional classical least squares (CLS) method,¹³ and is summarized in the following.

NIR absorbance spectra are obtained periodically from the reflectance spectra measured at the fixed measurement position of human skin irradiated by NIR light. Then, the absorbance difference spectra are calculated by subtracting the basis absorbance spectrum, which is the first absorbance spectrum measured at the beginning of measurement, from the measured absorbance

spectra obtained every 5 min. Using the modified Beer's law,¹⁴ we assume that the absorbance difference spectra can be synthesized by a linear combination of NIR absorption spectra of four components in the skin tissue (water, protein, glucose, and fat) and scattering equivalent absorption (so-called baseline¹⁵) component, and we can obtain the temporal changes in the contents of the components by solving the linear equations for five wavelengths characteristic to the five components. The temporal changes in the glucose content in the skin have been found to reflect the temporal changes in the glucose content in the blood with a time delay of about several minutes.^{16,17} Therefore, the temporal changes in the glucose content obtained above are expected to trace the temporal changes in the blood glucose content, but the results show failures. The reasons for the failures can be partly attributed to the changes in the scattering property in the skin after attaching the probe on the skin and to the difficulty of separating the glucose and baseline components. In order to incorporate the gradual changes in the scattering property during measurement in real time and to separate the glucose and baseline components, we develop a new approach to introduce a temporally varying imaginary component, which combines and replaces the baseline and fat components. The obtained changes in the glucose contents successfully trace and predict the changes in the blood glucose contents.

In Sec. 1, the background and purpose of this study are described, and the method proposed in this study is summarized. In Sec. 2, the prediction method, which is a simple application of the standard CLS method based on the modified Beer's law, is formulated by considering the five components (water, protein, glucose, fat, and baseline), and the experimental setup and data acquisition are explained. Then the results using the five components are described and the reasons for the failures of the prediction are explained. In Sec. 2.2, the imaginary component is introduced, and the results using the imaginary component are described and show the success in prediction. Sections 3 and 4 give discussion and conclusions, respectively.

2 Methods Based on the Modified Beer's Law

2.1 Formulation of the Blood Glucose Prediction

Based on the modified Beer's law, we first define three dimensionless quantities, i.e., the absorbance difference spectrum (expressed as the absorbance difference or the difference spectra for simplicity where it is understood without confusion),

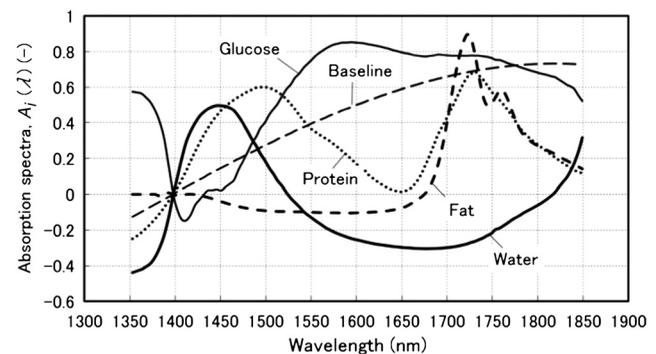


Fig. 1 Absorption spectra, $A_i(\lambda)$, of the four components in living tissues and baseline for synthesizing the absorbance difference spectra.

$\Delta OD(\lambda, t)$, the absorption spectrum of component i , $A_i(\lambda)$, and the content index of component i , $N_i(t)$, where λ and t are the wavelength and time after the start of the measurement, respectively. Appendix A gives the detailed derivations of the three quantities and the relationship among them, Eq. (15), which is rewritten below as Eq. (1), meaning that $\Delta OD(\lambda, t)$ is synthesized (or constructed) by the sum of the products of $A_i(\lambda)$ and the difference in $N_i(t)$, $\Delta N_i(t)$, over the components of concern

$$\Delta OD(\lambda, t) = \sum_i A_i(\lambda) \Delta N_i(t). \quad (1)$$

In this study, a total of five components (water, glucose, protein, fat, and baseline) are chosen to synthesize the absorbance difference spectra. The methods at determining $A_i(\lambda)$ of the five components are given in Appendix B, and Fig. 1 shows $A_i(\lambda)$ of the five components. The baseline component, which is the scattering equivalent absorption spectrum, is one of the key

components in this study. In our previous study, we created diffuse reflectance spectra from a Monte Carlo simulation of light propagation in the skin tissue and built a calibration model from these spectra using a PLS analysis.¹¹ Six components that influenced the spectra were chosen in the simulation, and five of them are also chosen in this study. One component that is not used in this study is temperature. During the experiments conducted in this study, the skin temperature was controlled and kept at $35 \pm 0.1^\circ\text{C}$ by fixing the optical probe to the skin surface by an adhesive tape as explained in Sec. 2.2, while in the previous study, the probe was attached to the skin surface by pushing at each spectral measurement.¹² Thus, temperature is excluded from the disturbance components to reduce the number of disturbances for a better demonstration of the method in this study.

Equation (1) is applied at the five characteristic wavelengths of the five components to obtain a set of simultaneous equations, Eq. (2):

$$\begin{bmatrix} \Delta OD(1450, t) \\ \Delta OD(1510, t) \\ \Delta OD(1600, t) \\ \Delta OD(1650, t) \\ \Delta OD(1727, t) \end{bmatrix} = \begin{bmatrix} A_W(1450)A_P(1450)A_G(1450)A_S(1450)A_F(1450) \\ A_W(1510)A_P(1510)A_G(1510)A_S(1510)A_F(1510) \\ A_W(1600)A_P(1600)A_G(1600)A_S(1600)A_F(1600) \\ A_W(1650)A_P(1650)A_G(1650)A_S(1650)A_F(1650) \\ A_W(1727)A_P(1727)A_G(1727)A_S(1727)A_F(1727) \end{bmatrix} \begin{bmatrix} \Delta N_W(t) \\ \Delta N_P(t) \\ \Delta N_G(t) \\ \Delta N_S(t) \\ \Delta N_F(t) \end{bmatrix}, \quad (2)$$

where the subscripts of $A_i(\lambda)$ and $\Delta N_i(t)$, i.e., $i = W, P, G, S$, and F , indicate water, protein, glucose, baseline, and fat, respectively. The characteristic wavelengths of the four tissue components are selected at their absorption peaks of 1450 nm for water, 1510 nm for protein, 1600 nm for glucose, and 1727 nm for fat, respectively. For the baseline, that has no absorption peak, we

have selected 1650 nm for the reason that the influences of water, protein, and fat on the measured spectral changes would be relatively small at $\lambda = 1650$ nm, where absorption by these components is weak.

By solving Eq. (2), the unknowns, $\Delta N_W(t)$, $\Delta N_P(t)$, $\Delta N_G(t)$, $\Delta N_S(t)$, and $\Delta N_F(t)$, are obtained as Eq. (3):

$$\begin{bmatrix} \Delta N_W(t) \\ \Delta N_P(t) \\ \Delta N_G(t) \\ \Delta N_S(t) \\ \Delta N_F(t) \end{bmatrix} = \begin{bmatrix} A_W(1450)A_P(1450)A_G(1450)A_S(1450)A_F(1450) \\ A_W(1510)A_P(1510)A_G(1510)A_S(1510)A_F(1510) \\ A_W(1600)A_P(1600)A_G(1600)A_S(1600)A_F(1600) \\ A_W(1650)A_P(1650)A_G(1650)A_S(1650)A_F(1650) \\ A_W(1727)A_P(1727)A_G(1727)A_S(1727)A_F(1727) \end{bmatrix}^{-1} \begin{bmatrix} \Delta OD(1450, t) \\ \Delta OD(1510, t) \\ \Delta OD(1600, t) \\ \Delta OD(1650, t) \\ \Delta OD(1727, t) \end{bmatrix}. \quad (3)$$

There are several reasons why we do not use the measured spectra fully over the whole wavelength range but use only the characteristic wavelengths: (1) the relation between the difference spectra and the quantities associated with the components is intuitively understood, (2) numerical calculation of Eq. (3) is easy, (3) the differences between the results using the full spectra and those using the characteristic wavelengths are found to be very small from our experiences, and (4) using only several wavelengths is more practical than using the full spectra toward the realization of noninvasive blood glucose monitoring devices.

The absolute values of the contents of the components cannot be obtained because the modified Beer's law includes unknown light path lengths in the skin, but the temporal changes in the contents of the components relative to those at the beginning can be obtained. The predicted blood glucose content, $G_p(t)$, is obtained by Eq. (4) from the change in the glucose content

index, $\Delta N_G(t)$, using a conversion factor, r , and the measured (true) initial blood glucose content, $G_m(t=0)$

$$G(t) = r\Delta N_G(t) + G_m(t=0) [\text{mg/dL}]. \quad (4)$$

2.2 Experiments

We have developed a new diffuse reflectance spectroscopic instrument to measure the NIR reflectance spectra from the human skin. Figure 2 shows a schematic diagram of the instrument, which consisted of a 150-W halogen lamp light source (TYPE 6550, Philips, Netherlands), an optical fiber bundle that included both source and detector fibers (cladding diameter 200 μm , core diameter 175 μm , NA 0.2, Fujikura, Japan), and a compact spectrometer containing a grating and a 256-channel linear-array photodetector (C9914GB, Hamamatsu Photonics,

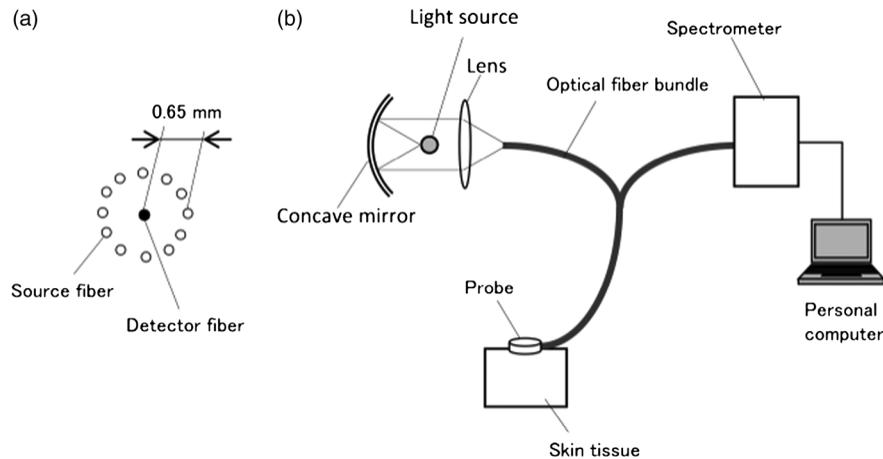


Fig. 2 (a) Arrangement of the source and detector optical fibers at the probe end and (b) schematic diagram of the experimental setup.

Japan). The optical signals from the spectrometer were transferred to a computer (Let's Note, Panasonic, Japan) to be processed for prediction of the blood glucose contents. The source and detector optical fibers were combined to make a single optical fiber bundle with a probe that had a cylindrical shape of

20 mm diameter and 8 mm length. Figure 2(a) shows the arrangement of the optical fibers at the end plane of the probe, where 12 source fibers surrounded one detector fiber in a circle with a radius of $650 \mu\text{m}$. The temperature of the interface between the probe and the skin surface was proportional-integral-differential controlled to be kept at $35 \pm 0.1^\circ\text{C}$ with a ribbon heater and a thermocouple attached to the probe. Light from the halogen lamp was collected by a concave mirror and transmitted through the source fibers to irradiate the skin. Then diffusely reflected light that reached the detector fiber was transmitted to the spectrometer. The reflected light intensity from a standard reflectance target, $I_s(\lambda)$ defined in Eq. (10), was measured by irradiating a standard 10% reflectance target (Labsphere, USA) at the beginning of the experiment. The spectra from 1350 to 1850 nm with a wavelength step of ~ 4 nm were used for data processing. The accumulation time of each linear-array photodetector was 140 ms, and 288 spectra were averaged per measurement to obtain one spectrum. The quality of the spectra was also assessed by measuring spectra every 5 min over the period of 120 min and by conducting a root-mean-square (rms) noise analysis of 100% lines on the collected data. Results showed that the average rms of the 100% line was $88.0 \mu\text{AU}$, which we predicted was sufficient for the purpose of predicting the blood glucose contents.¹⁸

True blood glucose contents and NIR diffuse reflectance spectra were measured continuously during the experiment in which blood glucose contents were changed artificially by oral glucose intake. True blood glucose contents were measured by using a portable blood glucose meter (DIA meter, Arkray, Japan) from a blood drop obtained by puncturing a fingertip every 15 min. The optical probe for spectral measurement was attached on the inside skin surface of the left forearm of a subject by a medical double-sided adhesive tape, and the NIR diffuse reflectance spectra were measured every 5 min during the experiment. A 200 ml of liquid-type nutrition food (Calorie Mate, Otsuka Pharmaceutical Co., Ltd., Japan) containing ~ 30 g carbohydrates, 10 g protein, and 4.5 g lipid was used for oral glucose intake. Each measurement was taken ~ 1.5 h before glucose intake, and ~ 2.0 h after glucose intake. The subject was a healthy man in his fifties and was kept seated calmly on a chair with a natural posture. His left forearm was horizontally placed on a soft pad without fixation, and he was encouraged to move as little as possible. No remarkable motion was observed during the measurement.

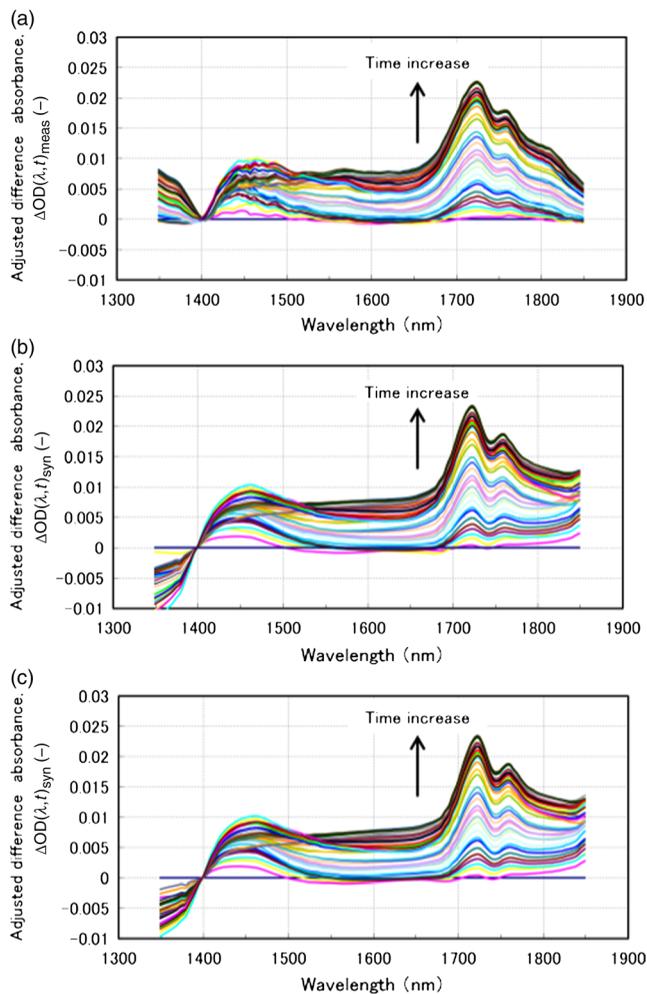


Fig. 3 Absorbance difference spectra obtained by (a) measurement, (b) synthesis using the five components, and (c) synthesis using the four components including the imaginary component.

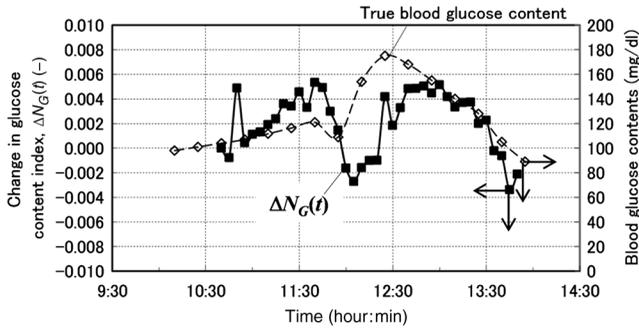


Fig. 4 Time series of the glucose content indices (left ordinate) and the true blood glucose contents (right ordinate) when using the five components.

Because the measured difference spectra were unstable immediately after the optical probe was attached to the forearm, spectral measurements started 30 min after attaching the probe when the measured absorbance difference spectra became stable. The first measured absorbance spectrum was used as the basis spectrum, $OD_1^*(\lambda, t_0)$ defined in Appendix A. Measurements of the true blood glucose contents every 15 min started immediately after attaching the probe, so the third true blood glucose content and the first absorbance spectrum were measured at almost at the same time.

The experiment was approved by the ethical committee of the company where the first author worked.

2.3 Results Using Five Components

Measured and synthesized difference spectra, $\Delta OD(\lambda, t)$, are shown in Fig. 3. Figure 3(a) plots 40 spectra of $\Delta OD(\lambda, t)$ measured every 5 min, and it is seen that the measured $\Delta OD(\lambda, t)$ increases with the increase in time to depart from the λ axis, $\Delta OD = 0$. Figure 3(b) plots 40 spectra of $\Delta OD(\lambda, t)$ synthesized from the calculated $\Delta N_i(t)$ of the five components using the measured $\Delta OD(\lambda, t)$ in Fig. 3(a) and Eq. (3). It is seen that the shapes of the synthesized $\Delta OD(\lambda, t)$ in Fig. 3(b) agree well with those of the measured ones in Fig. 3(a). The change in the glucose content index, $\Delta N_G(t)$, and the true blood glucose content, $G_m(t)$, are plotted in Fig. 4. From the good agreement between the measured and synthesized $\Delta OD(\lambda, t)$, it is expected that the profile of $\Delta N_G(t)$ would trace that of $G_m(t)$, but it does not, as is shown in Fig. 4.

Not only in the experiment stated above, but also in other oral glucose intake experiments, similar disagreements between the profiles of $\Delta N_G(t)$ and $G_m(t)$ were observed although good agreements between the measured and synthesized $\Delta OD(\lambda, t)$ were obtained. Resultantly, we conclude that it is difficult to predict the blood glucose contents correctly by a simple calculation of Eq. (3), which considers the five components of water, protein, glucose, fat, and baseline. From these results, we hypothesize the following:

- (1) As shown in Fig. 1, the absorption spectra of both glucose and baseline have no clear absorption peak with an approximately monotonic increase as the wavelength increases. Therefore, it is difficult to separate these two components from the shapes of their absorption spectra.
- (2) The absorption spectra of both glucose and baseline reveal strong absorption at their characteristic

wavelengths, 1600 and 1650 nm, respectively. Therefore, they influence each other, and a good prediction of $\Delta N_G(t)$ becomes difficult.

Now we propose a method in the next section to solve the problems described above so as to separate the glucose and baseline contributions in the measured $\Delta OD(\lambda, t)$, and to simultaneously incorporate the gradual change in scattering in the skin by introducing an imaginary component.

3 Introducing an Imaginary Component and Results

3.1 Separation of Glucose Contributions from the Baseline and Fat Contributions in the Measured $\Delta OD(\lambda, t)$

Before introducing an imaginary component, the glucose contribution in the measured $\Delta OD(\lambda, t)$ at the characteristic wavelength of baseline must be separated to solve the problem stated above. The separation can be conducted by utilizing the differences in the temporal response between the glucose and baseline components.

While the probe is attached on the skin by a medical double-sided adhesive tape with a very small contact pressure, the contribution of baseline increases monotonically and slowly during the measurement period of ~ 5 h. On the other hand, the blood glucose contents of healthy people increase during the period of ~ 1 h after oral glucose intake and then decrease in the following period of ~ 1 h to return to the initial values with some variations among individuals. This difference in the temporal response suggests the possibility of a separation of the contributions of glucose and baseline by temporal averaging of $\Delta OD(\lambda, t)$ at $\lambda = 1650$ nm, where the absorption spectra of both glucose and baseline have large values and influence each other. The proposed averaging process is as follows.

The time series of $\Delta OD(1650, t)$ are given by measurements every 5 min from $t = t_0$ to $t = t_n (= 5j \text{ min}; j = 0, 1, 2, \dots, n)$, and the difference between $\Delta OD(1650, t_j)$ and $\Delta OD(1650, t_0)$ is denoted as $\Delta s(t_j) = \Delta OD(1650, t_j) - \Delta OD(1650, t_0)$. Then, at each measurement time, $t = t_n$, the average, $\overline{\Delta s(t_n)}$, is given by Eq. (5):

$$\begin{aligned} \overline{\Delta s(t_n)} &= [\Delta s(t_0) + \Delta s(t_1) + \dots + \Delta s(t_n)] / (n + 1) \\ &= \sum_{j=0}^n \Delta s(t_j) / (n + 1). \end{aligned} \quad (5)$$

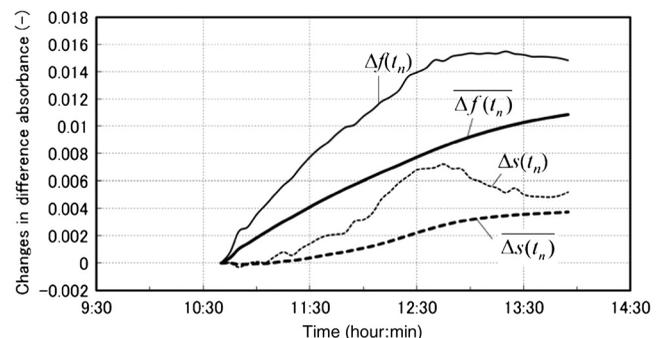


Fig. 5 Time series of $\Delta s(t_n)$, $\overline{\Delta s(t_n)}$, $\Delta f(t_n)$, and $\overline{\Delta f(t_n)}$.

Figure 5 shows an example of the time series of $\Delta s(t_n)$ and $\overline{\Delta s(t_n)}$. In this experiment, the measurement started at 10:40 and oral glucose intake was done at $\sim 12:00$. In the time series of $\Delta s(t_n)$, there exists a rise beginning at $\sim 12:00$ and a drop beginning at $\sim 13:00$, which are added to the monotonic increase with time, suggesting the change in the blood glucose content by the oral glucose intake. On the other hand, the time series of $\overline{\Delta s(t_n)}$ does not show the rise and drop after the oral glucose intake, but increases monotonically, suggesting that $\overline{\Delta s(t_n)}$ mainly reflects the temporal change in the baseline component. Thus, the contributions of the glucose and baseline components are considered to be separated to some extent.

At the characteristic wavelength of fat, $\lambda = 1727$ nm, glucose also has strong absorption, and it is necessary to reduce the contribution of glucose in the measured $\Delta OD(1727, t)$. This is made by taking the difference in $\Delta OD(\lambda, t)$ between $\lambda = 1727$ nm and $\lambda = 1650$ nm, with denoting $\Delta f(t) = \Delta OD(1727, t) - \Delta OD(1650, t)$ because glucose has a strong absorption at $\lambda = 1650$ nm similar to that at $\lambda = 1727$ nm. Again, $\Delta f(t)$ is averaged in the same manner as Eq. (5) to further reduce the contribution of glucose by Eq. (6):

$$\begin{aligned} \overline{\Delta f(t_n)} &= [\Delta f(t_0) + \Delta f(t_1) + \cdots + \Delta f(t_n)] / (n + 1) \\ &= \sum_{j=0}^n \Delta f(t_j) / (n + 1). \end{aligned} \quad (6)$$

The time series of $\Delta f(t)$ and $\overline{\Delta f(t_n)}$ are also plotted in Fig. 5, and it is seen that the contribution of glucose is reduced in $\overline{\Delta f(t_n)}$, which reflects the change in the fat component to some extent, although it is not clearly as seen as $\Delta s(t_n)$.

3.2 Introducing an Imaginary Component with Creation of Its Spectrum

During the measurement period of ~ 5 h, it is reasonable to assume that the fat content in the skin tissue does not change. Therefore, we suppose that the large growth of the peak at $\lambda = 1727$ nm appearing in $\Delta OD(\lambda, t)$ shown in Fig. 3(a), which is assigned as a fat absorption peak, is not due to the increase in the fat content in the skin, but due to the increase in the light intensity propagating in the subcutaneous layer, which contains fat with a larger content than in the skin. The increase in the light intensity in the subcutaneous layer is understood to be caused by the decrease in the scattering coefficient of the stratum corneum and epidermis. Attaching the probe to the skin surface would

reduce evaporation of moisture from the skin surface and would increase the water content in the stratum corneum and epidermis. The increase in the water content causes the decrease in the scattering coefficient because the difference in the refractive index between the cells and intercellular liquid decreases. The decrease in the scattering coefficient in the stratum corneum and epidermis makes more light propagate in the subcutaneous layer. From another point of view, because baseline is an expression of wavelength dependency of scattering by the skin with the equivalent absorption, the decrease in the scattering coefficient in the stratum corneum and epidermis causes the decrease in the reflected light intensity, which means the increase in the absorption of baseline. As a result, contributions of both the fat and baseline components, $\Delta N_F(t)$ and $\Delta N_S(t)$, to $\Delta OD(\lambda, t)$ increase with the decrease in scattering in the stratum corneum and epidermis.

Now we introduce an imaginary component with a content index of $N_I(t)$ and an absorption spectrum of $A_I(\lambda, t)$, respectively. Here, $A_I(\lambda, t)$ varies with time, and at time t_n , it is created by adding the absorption spectra of baseline and fat, $A_S(\lambda)$ and $A_F(\lambda)$, to the weights of $\overline{\Delta s(t_n)}$ and $\overline{\Delta f(t_n)}$, respectively, as in Eq. (7):

$$A_I(\lambda, t_n) = \overline{\Delta s(t_n)} A_S(\lambda) + \overline{\Delta f(t_n)} A_F(\lambda). \quad (7)$$

Equation (7) is more or less heuristically obtained, but there are some rationales. Combining the absorption spectra of baseline and fat comes from the fact that the changes in both the baseline and fat components originate from the changes in scattering in the skin. By combining them, it is possible to consider only one component, which is dependent on scattering in the skin. By weighing them with the averaged quantities of $\overline{\Delta s(t_n)}$ and $\overline{\Delta f(t_n)}$, it is possible to incorporate the intensity (or significance) of the baseline and fat components in the imaginary component at each measurement time, respectively. Note that the absorption spectra of baseline and fat, $A_S(\lambda)$ and $A_F(\lambda)$, are independent of time, while the absorption spectrum of the imaginary component, $A_I(\lambda, t_n)$, is dependent on time so that the imaginary component can reflect the gradual change in scattering in the skin after attaching the probe on the skin.

Using the imaginary component that combines and replaces the baseline and fat components, Eq. (1) for the characteristic wavelengths of the four components, i.e., water, glucose, protein, and imaginary, are obtained as Eq. (8) and solved for the content indices of the four components as Eq. (9):

$$\begin{bmatrix} \Delta OD(1450, t) \\ \Delta OD(1510, t) \\ \Delta OD(1600, t) \\ \Delta OD(1727, t) \end{bmatrix} = \begin{bmatrix} A_W(1450)A_P(1450)A_G(1450)A_I(1450, t) \\ A_W(1510)A_P(1510)A_G(1510)A_I(1510, t) \\ A_W(1600)A_P(1600)A_G(1600)A_I(1600, t) \\ A_W(1727)A_P(1727)A_G(1727)A_I(1727, t) \end{bmatrix} \begin{bmatrix} \Delta N_W(t) \\ \Delta N_P(t) \\ \Delta N_G(t) \\ \Delta N_I(t) \end{bmatrix}, \quad (8)$$

$$\begin{bmatrix} \Delta N_W(t) \\ \Delta N_P(t) \\ \Delta N_G(t) \\ \Delta N_I(t) \end{bmatrix} = \begin{bmatrix} A_W(1450)A_P(1450)A_G(1450)A_I(1450, t) \\ A_W(1510)A_P(1510)A_G(1510)A_I(1510, t) \\ A_W(1600)A_P(1600)A_G(1600)A_I(1600, t) \\ A_W(1727)A_P(1727)A_G(1727)A_I(1727, t) \end{bmatrix}^{-1} \begin{bmatrix} \Delta OD(1450, t) \\ \Delta OD(1510, t) \\ \Delta OD(1600, t) \\ \Delta OD(1727, t) \end{bmatrix}. \quad (9)$$

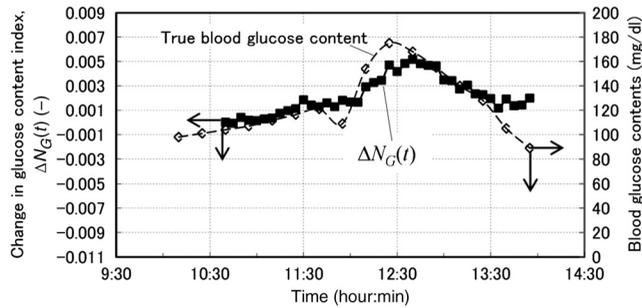


Fig. 6 Time series of the glucose content indices (left ordinate) and the true blood glucose contents (right ordinate) when using the four components including the imaginary component.

Here, the subscript of time, n , is omitted for simplicity, and the characteristic wavelength of the imaginary component is selected as 1727 nm, which is the same as that of fat. The characteristic wavelength of baseline, 1650 nm, is closer to that of glucose, 1600 nm, and better results are expected with the wavelength of 1727 nm farther from the characteristic wavelength of glucose.

3.3 Results Using the Imaginary Component

The changes in the content indices of the four components, $\Delta N_i(t)$, are obtained using Eq. (9) from the measured data used in Sec. 3, and the absorbance difference spectra, $\Delta OD(\lambda, t)$, are synthesized by substituting the obtained $\Delta N_i(t)$ into Eq. (8), as plotted in Fig. 3(c). The time series of the change in the glucose content index, $\Delta N_G(t)$, is compared with that of the true blood glucose content, $G_m(t)$, in Fig. 6. Again, the synthesized spectra using the imaginary component in Fig. 3(c) agree well with the measured ones in Fig. 3(a), even if the number of components is reduced from five to four. The change in the glucose content index, $\Delta N_G(t)$, traces the change in the true blood glucose content, $G_m(t)$, in Fig. 6 showing the rise and fall during the oral glucose intake, while in Fig. 4, $\Delta N_G(t)$ did not trace $G_m(t)$ well. Using Eq. (4), $\Delta N_G(t)$ in Fig. 6 is converted to the predicted blood glucose contents, $G_p(t)$, with the conversion factor, $r = 1.54 \times 10^4$ mg/dL, and $G_p(t)$ is compared with $G_m(t)$, the true blood glucose content, in Fig. 7 after shifting $G_p(t)$ by 5 min to earlier times due to the time delay. The value of r was determined from the results of several experiments conducted in advance. The correlation coefficient of the predicted and true blood glucose contents is 0.93 and the standard error of

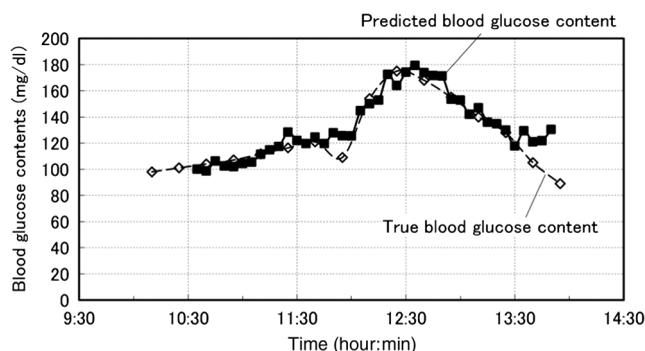


Fig. 7 Time series of the true blood glucose contents and predicted ones using the four components.

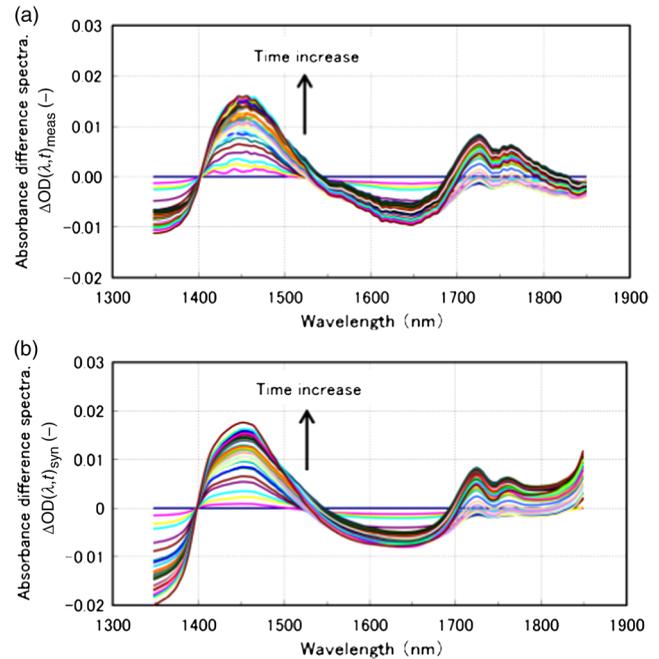


Fig. 8 (a) Measured absorbance difference spectra, and (b) synthesized absorbance difference spectra using the four components including the imaginary component for the second experiment.

prediction (SEP) is 9.9 mg/dL. The synthesized $\Delta OD(\lambda, t)$ using the imaginary component shown in Fig. 3(c) is only slightly different from those without using the imaginary component shown in Fig. 3(b). But, only a slight difference in the synthesized spectra leads to a significant improvement in the prediction of the blood glucose contents from Fig. 4 to Figs. 6 and 7, thus demonstrating the good performance of the approach using the imaginary component.

Another type of the temporal development of $\Delta OD(\lambda, t)$ different from that showing the growth at $\lambda = 1727$ nm in Fig. 3 is also examined. That type shows a large growth at $\lambda = 1450$ nm, which is the characteristic wavelength of water, and 40 measured $\Delta OD(\lambda, t)$ are plotted in Fig. 8(a). While $\Delta OD(1450, t)$ shows a large increase, $\Delta OD(1600, t)$ at the characteristic wavelength of glucose shows a decrease. Figure 8(b) shows the synthesized $\Delta OD(\lambda, t)$ calculated from the solutions of Eq. (9) using the four components, and the shapes of the synthesized

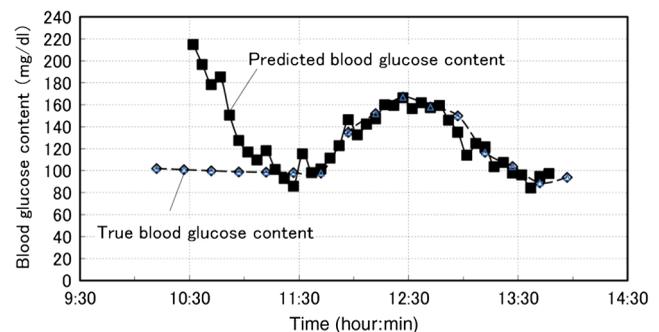


Fig. 9 Time series of the true blood glucose contents and predicted ones using the four components for the second experiment. The predicted blood glucose content is matched with the true one at the time of the beginning of the oral glucose intake.

$\Delta OD(\lambda, t)$ are very similar to those of the measured $\Delta OD(\lambda, t)$ in Fig. 8(a). The calculated $\Delta N_G(t)$ are converted to $G_p(t)$ with $r = 1.54 \times 10^4$ mg/dL, and Fig. 9 compares $G_p(t)$ with $G_m(t)$ after shifting $G_p(t)$ for compensation of the 5-min time delay. In this case, the true initial blood glucose content, $G_m(t = 0)$, was 100 mg/dL, which should have been used in Eq. (4). However, the predicted glucose content, $G_p(t)$, decreased rapidly from the beginning of the spectral measurement to the time of the oral glucose intake, while the true blood glucose content, $G_m(t)$, remained constant during that period. As a result, $G_p(t)$ significantly deviated from $G_m(t)$. In this case, $G_m(t = 0)$ ($= 100$ mg/dL) in Eq. (4) is replaced by $G_m(t = 70)$ ($= 100 + 115$ mg/dL) so that $G_p(t)$ is equal to $G_m(t)$ at 11:40, the time of the oral glucose intake. As shown in Fig. 9, $G_p(t)$ agrees well with $G_m(t)$ after the oral glucose intake. The disagreement in the early period has to be investigated and solved in the future. The correlation coefficient and SEP for the whole period are 0.56 and 27.7 mg/dL.

4 Discussion

From this study, it is found that the difficulty in predicting the blood glucose content from the reflectance measurements mainly originates from the change in baseline (the change in scattering in the skin) and from the similarity of the absorption spectra between glucose and baseline. Their influences are reduced by introducing the imaginary component. In the process of introducing the imaginary component, the difference in the time response between the glucose content and scattering (baseline) is used to separate their contributions to the measured $\Delta OD(1650, t)$. By creating the absorption spectrum of the imaginary component, the performance of predicting the blood glucose contents is significantly improved. In other words, utilizing temporal information in addition to intensity information in the measured spectra greatly improves the prediction performance.

Because the imaginary component is created using the absorbance difference spectra measured at the fixed position on the skin sequentially from the start of measurement to the time of concern, it reflects the gradual changes in scattering in the skin irrespective of the differences in individuals, measurement positions, seasons, etc. Therefore, the method using the imaginary component is appropriate for real-time prediction of blood glucose contents. However, the four components including the imaginary component are not necessarily universal for all cases as demonstrated by the failure shown in Fig. 9. Therefore, other key components must be found by further developments of the approach proposed in this study toward the clinical use of NIR noninvasive blood glucose monitoring.

Instead of introducing the imaginary component, which combines and replaces the baseline and fat components, modification of the absorption spectra of baseline and fat by multiplying the temporal averages, such as $\Delta s(t_n)A_S(\lambda)$ and $\Delta f(t_n)A_F(\lambda)$, respectively, may be another choice for separating the glucose contributions and for incorporating the scattering changes. In this case, the number of the components is kept at five, and the results of the synthesized $\Delta OD(\lambda, t)$ and $G_m(t)$ may provide more information for improving the prediction. More appropriate modifications than the method proposed in this study may be possible and developed in future studies.

Conventional methods using multivariate analyses, such as principal component regression and PLS, also utilize temporal information. However, in the case where there exist some components that occasionally correlate with the blood glucose

content, there is the possibility of a so-called chance temporal correlation,¹⁹ because the calibration models are built after the whole measurement data are obtained. On the other hand, the method proposed here calculates the glucose content index from the absorbance difference spectra at every measurement time (every 5 min) and predicts the blood glucose content independent of the measurements after the measurement at that time of concern. So the possibility of chance temporal correlation is very small in the proposed method.

Because the temporal averaging used in this study presumes the time response of the blood glucose content for healthy people, it will be necessary to study further for application to diabetes patients, who might have time responses different from healthy people.

Better methods other than the temporal averaging used in this study will also be possible because its purpose is to remove the influences of the glucose contributions from the measured $\Delta OD(1650, t)$ and $\Delta OD(1727, t)$ at the characteristic wavelengths of baseline and fat, respectively. It is well known that the influence of baseline can be reduced by pretreatment of the measured spectra such as differentiation and multiplicative scatter correction. Former studies of the quantification of blood glucose content in living bodies²⁰ and in red blood cell products²¹ reported that the performances of the quantification were improved by the pretreatment. Introducing the imaginary component to replace baseline may be similar to the pretreatment in the former studies.

In the two experiments described in this study, we used the conversion factor, $r = 1.54 \times 10^4$ mg/dL, for predicting $G_m(t)$ from $\Delta N_G(t)$. The glucose content index is given as $N_G(t) = c_G(t)/c_{G0}$, and the conversion factor, r , can be determined from the measured glucose content, $c_G(t)$, and the standard glucose content, c_{G0} . However, because it is difficult to estimate the value of the standard glucose content, c_{G0} , the conversion factor, r , was empirically determined by comparing the change in the true blood glucose content with the change in the glucose content index. Although it is difficult to determine the value of c_{G0} , the value is constant, and it is reasonable to use the same conversion factor consistently. Good results of predicting the blood glucose content after the oral glucose intake in the two experiments using the same conversion factor illustrate the description above.

In CLS methods that are similar to our method, it is generally understood that estimation of the number of components for synthesizing spectra is very important, and that correct numbers of the components provide accurate predictions, while incorrect numbers of the components with unexpected disturbances lead to inaccurate predictions. Therefore, the large prediction errors from the beginning of the experiments to the oral glucose intake in Fig. 9 are thought to be caused by unexpected disturbances contained in the measured spectra. By attaching the probe to the skin surface, the increase in the water content and smoothing in the cell shape in the stratum corneum and epidermis drastically and simultaneously take place, especially at the beginning of the probe tip attachment. The large growth of the water absorption peak around $\lambda = 1450$ nm in $\Delta OD(\lambda, t)$ shown in Fig. 8 indicates that some phenomena related to water are attributed to the large prediction errors shown in Fig. 9.

In this study, it is assumed that the effective optical path length, l , does not depend on the wavelength, λ , but l actually increases with the increase in the wavelength because l depends on the scattering coefficient similar to baseline. According to numerical simulation of light propagation in the skin, l increases

from 1.2 to 2.0 mm as the wavelength increases from 1350 to 1800 nm when the distance between the source and detector fibers is 0.65 mm.²² It will be possible to incorporate this wavelength dependency of l in a future study.

In contrast to the methods using multivariate analyses, the proposed method based on the modified Beer's law does not build a calibration model, which needs many premeasurement data in advance, but needs only the absorption spectra of the components in living tissues and baseline. This feature of the proposed method will be very useful for the realization of practical devices for accurate blood glucose prediction although it is necessary to discuss precision, error and detection limit, and so on in comparison with chemometrics methods before the proposed method is really available for clinical use.

5 Conclusions

The glucose prediction method based on the modified Beer's law to synthesize the measured absorbance difference spectra from the absorption spectra of the components in living tissues and baseline is investigated. Synthesizing the measured spectra provides clear understanding of the process of blood glucose prediction from the measured spectra. It is found that the main factors of large prediction errors are the gradual changes in scattering in the skin and the similarity in the shapes of the absorption spectra between glucose and baseline. The prediction errors caused by these two factors are found to be improved significantly by introducing the imaginary component combining and replacing the baseline and fat components after separating the glucose contributions from the measured absorbance differences at the characteristic wavelengths of the baseline and fat components. The results using the imaginary component demonstrate the possibility of simple and real-time blood glucose prediction using NIR spectroscopy based on the modified Beer's law, although there still exist unknown disturbances that affect the prediction accuracy.

Because the imaginary components created from the spectra measured at the fixed position on the skin sequentially from the start of measurement to the time of concern are free from the differences in individuals, measurement positions, seasons, etc., the method using the imaginary component is appropriate for real-time prediction of blood glucose contents. Thus, the proposed method will become a useful tool for realizing non-invasive blood glucose prediction using NIR spectroscopy with further development for incorporating unexpected disturbances in addition to the disturbances considered in this study.

Appendix A Relation Between the Difference Spectra and the Contents of the Components Based on the Modified Beer's Law

The absorbance difference spectra at time t and wavelength λ , i.e., $\Delta OD^*(\lambda, t)$, is defined by the difference between the measured absorbances at state 2 and state 1, $OD_2^*(\lambda, t)$ and $OD_1^*(\lambda, t_0)$, respectively:

$$\begin{aligned} \Delta OD^*(\lambda, t) &= OD_2^*(\lambda, t) - OD_1^*(\lambda, t_0) \\ &= -\ln \frac{I_2(\lambda, t)}{I_0(\lambda)} + \ln \frac{I_1(\lambda, t_0)}{I_0(\lambda)} \\ &= -\ln \frac{I_2(\lambda, t)}{I_1(\lambda, t_0)} = -\ln \frac{I_2(\lambda, t)/I_s(\lambda)}{I_1(\lambda, t_0)/I_s(\lambda)}, \end{aligned} \quad (10)$$

where $I_0(\lambda)$ is the source light intensity, $I_1(\lambda, t_0)$ and $I_2(\lambda, t)$ are the reflected light intensities at a reference time (state 1, $t = t_0$) and an arbitrary time (state 2, $t = t$), respectively, and $I_s(\lambda)$ is the reflected light intensity from a standard reflectance target.

Theoretically, the absorbances, $OD_1^*(\lambda, t_1)$ and $OD_2^*(\lambda, t)$, are given by the modified Beer's law, which is applied to light scattering media as Eqs. (11) and (12), respectively:

$$OD_1^*(\lambda, t_0) = \sum_i \varepsilon_i^*(\lambda) c_{i1}(t_0) l_1(\lambda, t_0) + B(\lambda, t_0), \quad (11)$$

$$OD_2^*(\lambda, t) = \sum_i \varepsilon_i^*(\lambda) c_{i2}(t) l_2(\lambda, t) + B(\lambda, t), \quad (12)$$

where subscript i indicates the tissue component, such as glucose, water, fat, and protein, subscripts 1 and 2 indicate the states 1 and 2, respectively, $\varepsilon^*(\lambda)$ is the molar extinction coefficient ($\text{mm}^{-1} \text{M}^{-1}$), c is the content of each component (M), l is the effective optical path length (mm), and B is the attenuation of light intensity by scattering. As the effective optical path length l and the attenuation by scattering B essentially depend not only on the wavelength but also on the scattering and absorption coefficients in the skin tissue, they depend on time. From Eqs. (10), (11), and (12), the absorbance difference, $\Delta OD^*(\lambda, t)$, is given as Eq. (13):

$$\begin{aligned} \Delta OD^*(\lambda, t) &= \sum_i \varepsilon_i^*(\lambda) [c_{i2}(t) l_2(\lambda, t) - c_{i1}(t_0) l_1(\lambda, t_0)] \\ &\quad + B(\lambda, t) - B(\lambda, t_0) \\ &= \sum_i \varepsilon_i^*(\lambda) l \Delta c_i(t) + \varepsilon_B^*(\lambda) l \Delta c_B(t). \end{aligned} \quad (13)$$

Here, $\Delta c_i(t) = c_{i2}(t) - c_{i1}(t_0)$, and the effective optical path lengths are assumed not to change by the change from state 1 to state 2 and not to depend on the wavelength, i.e., $l_1(\lambda, t_0) = l_2(\lambda, t) = l$, and the change in the attenuation by scattering is assumed to be expressed in the same manner as the other components in living tissue, as $B(\lambda, t) - B(\lambda, t_0) = \varepsilon_B^*(\lambda) l \Delta c_B(t)$ when introducing the equivalent molar extinction coefficient, $\varepsilon_B^*(\lambda)$, and an equivalent content, $c_B(t)$. Now, the content of each component, $c_i(t)$, is expressed as a product of each standard content, c_{i0} , and a coefficient, $N_i(t)$, i.e., $c_i(t) = c_{i0} N_i(t)$, and the standard absorption spectrum of the component i , including attenuation by scattering, is defined as $A_i^*(\lambda) = \varepsilon_i^*(\lambda) c_{i0} l$. Then, Eq. (13) is simplified to Eq. (14):

$$\Delta OD^*(\lambda, t) = \sum_i A_i^*(\lambda) \Delta N_i(t), \quad (14)$$

where $\Delta N_i(t) = N_{i2}(t) - N_{i1}(t_0)$, and N_i is called the content index of the component i .

An additional procedure is performed on the absorbance difference spectrum, $\Delta OD^*(\lambda, t)$, and the standard absorption spectrum, $A_i^*(\lambda)$. They are adjusted to be equal to zero at $\lambda = 1400$ nm and denoted as $\Delta OD(\lambda, t)$ and $A_i(\lambda)$, and expressed by Eqs. (15) and (16), respectively:

$$\begin{aligned}\Delta OD(\lambda, t) &= \Delta OD^*(\lambda, t) - \Delta OD^*(1400, t) \\ &= \sum_i A_i^*(\lambda) \Delta N_i(t) - \sum_i A_i^*(1400) \Delta N_i(t) \\ &= \sum_i A_i(\lambda) \Delta N_i(t),\end{aligned}\quad (15)$$

$$A_i(\lambda) = A_i^*(\lambda) - A_i^*(1400). \quad (16)$$

The idea of adjustment at $\lambda = 1400$ nm came from our preliminary experiment using liquid phantoms. The liquid phantoms made of Intralipid solution (Fresenius Kabi) had the reduced scattering coefficients of either 1.35 or 1.65 mm^{-1} and contained glucose with five glucose contents from 100 to 800 mg/dL simulating the range of the human blood glucose content. The absorbance spectra of the total of 10 phantoms in the wavelength range from 1300 to 1850 nm were measured and analyzed. As a result, when the absorbance spectra were adjusted to be equal to zero at $\lambda = 1400$ nm, their values at $\lambda = 1600$ nm (a peak wavelength of glucose absorption) for all the phantoms were plotted on one regression line as a function of the glucose content. This means that the values at $\lambda = 1600$ nm of the adjusted (at $\lambda = 1400$ nm) standard absorbance are independent of the reduced scattering coefficient of the background medium.²³ Therefore, adjustment at $\lambda = 1400$ nm was found to reduce the effect of the disturbance by scattering the change of predicting blood glucose contents. So in this study, we also use adjustment at $\lambda = 1400$ nm.

In the main text of this study, the adjusted absorbance difference, $\Delta OD(\lambda, t)$, is simply called the absorbance difference spectrum, and the adjusted standard absorption spectrum, $A_i(\lambda)$, is also simply called the absorption spectrum.

Appendix B Determination of the Absorption Spectra

The absorption spectra, $A_i(\lambda)$, of the five components of water, protein, glucose, fat, and baseline are calculated from the standard absorption spectra, $A_i^*(\lambda)$, by adjusting at $\lambda = 1400$ nm. $A_i^*(\lambda)$ of the five components are determined as described in the following.

$A_i^*(\lambda)$ of water and glucose were obtained from the transmittance measurements of pure water and an aqueous glucose solution (with the glucose content of 10 g/dL) in cells with a thickness of 1.0 mm. The measured absorbance ($-\ln[I_i(\lambda)/I_0(\lambda)]$) of the pure water sample was equal to $A_i^*(\lambda)$, and the absorption coefficient, $\mu_a(\lambda)$, of water was obtained by dividing $A_i^*(\lambda)$ by the cell thickness. Note that the absorption coefficient is given as $\mu_a(\lambda) = \varepsilon^*(\lambda)c$, and $A^*(\lambda) = \mu_a(\lambda) \cdot l = \varepsilon^*(\lambda) \cdot cl$ in general. $\mu_a(\lambda)$ of the aqueous glucose solution was also obtained by dividing the measured absorbance by the cell thickness, and $\mu_a(\lambda)$ of pure glucose was obtained by subtracting $\mu_a(\lambda)$ of pure water from $\mu_a(\lambda)$ of the aqueous glucose solution. Then, the extinction coefficient, $\varepsilon_i^*(\lambda)$, of glucose was obtained by dividing $\mu_a(\lambda)$ of glucose by the concentration of the glucose, i.e., 10 g/dL = 0.56 M.

The samples of protein and fat were collagen and beef tallow, respectively, which are scattering media. Then the absorption coefficients of protein and fat, $\mu_{ai}(\lambda)$ (subscript i represents protein or fat), were obtained by an inverse Monte Carlo calculation from the measured spectra of the transmittance and reflectance using a spectrometer (UV-3100, Shimadzu, Japan) equipped with

an integrating sphere.²⁴ The cell thicknesses of the samples, l_i , were 0.5 mm for collagen and 1.0 mm for beef tallow, but the concentration of the samples, c_{i0} , were unknown. Therefore, the absorption coefficient of collagen and beef tallow was given as $\mu_{ai}(\lambda) = \varepsilon_i^*(\lambda)c_{i0}\alpha_i$ with arbitrary values of $c_{i0}\alpha_i$. Then the standard absorption spectra were given as $A_i^*(\lambda) = \varepsilon_i^*(\lambda)c_{i0}l_i = \mu_{ai}(\lambda)l_i/\alpha_i$.

$A_i^*(\lambda)$ of baseline was estimated from the scattering coefficient of epidermis. As indicated in the papers of Troy and Thennadil²⁵ and Simpson et al.,²⁶ the scattering coefficient of epidermis decreases monotonically with the increase in the wavelength. When the scattering coefficient decreases, the diffuse reflectance also decreases, resulting in the increase in the absorbance. Therefore, $A_i^*(\lambda)$ of baseline was given to increase monotonically with the wavelength.

In general, the absorption spectrum of component i , $A_i(\lambda)$, is expressed by Eq. (17) with the extinction coefficient normalized by an arbitrary factor α_i as $\varepsilon_i(\lambda) = \varepsilon_i^*(\lambda)/\alpha_i$:

$$A_i(\lambda) = A_i^*(\lambda) - A_i^*(1400) = \alpha_i \varepsilon_i(\lambda) c_{i0} l - \alpha_i \varepsilon_i(1400) c_{i0} l. \quad (17)$$

References

- G. C. Cote, "Noninvasive and minimally-invasive optical monitoring technologies," *J. Nutr.* **131**(5), 1596S–1604S (2001).
- O. S. Khalil, "Noninvasive glucose measurement technologies: an update from 1999 to the dawn of the new millennium," *Diabetes Technol. Ther.* **6**(5), 660–697 (2004).
- S. K. Vashist, "Non-invasive glucose monitoring technologies in diabetes management: a review," *Anal. Chim. Acta* **750**, 16–27 (2012).
- H. M. Heise, "Near-infrared spectroscopy for *in vivo* glucose sensing," Chapter 3 in *Biosensors in the Body Continuous In Vivo Monitoring*, D. M. Fraser, Ed., pp. 79–116, John Wiley and Sons, New York (1997).
- J. Chen, M. A. Arnold, and G. W. Small, "Comparison of combination and first overtone spectral regions for near infrared calibration models for glucose and other biomolecules in aqueous solutions," *Anal. Chem.* **76**(18), 5405–5413 (2004).
- H. Martens and T. Naes, *Multivariate Calibration*, John Wiley, New York (1991).
- S. F. Malin et al., "Noninvasive prediction of glucose by near-infrared diffuse reflectance spectroscopy," *Clin. Chem.* **45**(9), 1651–1658 (1999).
- K. Maruo et al., "In vivo noninvasive measurement of blood glucose by near-infrared diffuse-reflectance spectroscopy," *Appl. Spectrosc.* **57**(10), 1236–1244 (2003).
- M. A. Arnold, L. Liu, and J. T. Olesberg, "Selectivity assessment of noninvasive glucose measurements based on analysis of multivariate calibration vectors," *J. Diabetes Sci. Technol.* **1**(4), 454–462 (2007).
- N. V. Alexeeva and M. A. Arnold, "Near-infrared microspectroscopic analysis of rat skin tissue heterogeneity in relation to noninvasive glucose sensing," *J. Diabetes Sci. Technol.* **3**(2), 219–232 (2009).
- K. Maruo et al., "New methodology to obtain a calibration model for noninvasive near-infrared blood glucose monitoring," *Appl. Spectrosc.* **60**(4), 441–449 (2006).
- K. Maruo et al., "Noninvasive near-infrared blood glucose monitoring using a calibration model built by a numerical simulation method: trial application to patients in an intensive care unit," *Appl. Spectrosc.* **60**(12), 1423–1431 (2006).
- H. M. Heise and R. Winzen, "Chemometrics in near-infrared spectroscopy," Chapter 7 in *Near-Infrared Spectroscopy: Principles, Instruments, Applications*, H. W. Siesler et al., Eds., pp. 125–162, Wiley-VCH, Weinheim (2002).
- D. T. Delpy et al., "Estimation of optical pathlength through tissue from direct time of flight measurement," *Phys. Med. Biol.* **33**(12), 1433–1442 (1988).
- P. Ritthiruangdej et al., "Non-destructive and rapid analysis of chemical compositions in Thai steamed pork sausages by near-infrared spectroscopy," *Food Chem.* **129**, 684–692 (2011).

16. S. N. Thennadil et al., "Comparison of glucose content in interstitial fluid, and capillary and venous blood during rapid change in blood glucose levels," *Diabetes Technol. Ther.* **3**(3), 357–365 (2001).
17. J. P. Bantle and W. Thomas, "Glucose measurement in patients with diabetes mellitus with dermal interstitial fluid," *J. Lab. Clin. Med.* **130**, 436–441 (1997).
18. H. Arimoto, M. Tarumi, and Y. Yamada, "Estimation of absorption change in scattering media by glucose concentration for spectroscopic measurements," in *Proc. of Asian Symp. on Biomedical Optics and Photomedicine*, pp. 214–215 (2002).
19. M. A. Arnold, J. J. Burmeister, and G. W. Small, "Phantom glucose calibration models from simulated noninvasive human near-infrared spectra," *Anal. Chem.* **70**(9), 1773–1781 (1998).
20. Y. P. Du et al., "Improvement of the partial least squares model performance for oral glucose intake experiments by inside mean centering and inside multiplicative signal correction," *Anal. Sci.* **21**(8), 979–984 (2005).
21. Y. Suzuki et al., "Preliminary evaluation of optical glucose sensing in red cell contents using near-infrared diffuse-reflectance spectroscopy," *J. Biomed. Opt.* **17**(1), 017004 (2012).
22. K. Iino et al., "Monte Carlo simulation of NIR reflectance spectroscopy in the wavelength range from 1000 nm to 1900 nm," *Opt. Rev.* **10**(6), 600–606 (2003).
23. K. Maruo and Y. Suzuki, "The method for blood glucose measurement and its apparatus," Japan Patent Kokai 2014-18478 (2014). (in Japanese).
24. T. Ota et al., "Measurement of optical properties of blood components using an inverse Monte Carlo technique," in *Proc. of Asian Symp. on Biomedical Optics and Photomedicine*, pp. 230–231 (2002).
25. T. L. Troy and S. N. Thennadil, "Optical properties of human skin in the near infrared wavelength range of 1000 to 2200 nm," *J. Biomed. Opt.* **6**(2), 167–176 (2001).
26. C. R. Simpson et al., "Near-infrared optical properties of *ex vivo* human skin and subcutaneous tissues measured using the Monte Carlo inversion technique," *Phys. Med. Biol.* **43**(9), 2465–2478 (1998).

Katsuhiko Maruo is a chief engineer at Panasonic Healthcare Co., Ltd. He received his BS and MS degrees in chemical engineering from Kobe University in 1983 and 1985, respectively, and his PhD from the University of Electro-Communications in 2007. His current research interests include near-infrared spectroscopy and fluorescent detection of living materials.

Yukio Yamada is a professor of the University of Electro-Communications. He received his BS, MS, and PhD degrees from Tokyo Institute of Technology in 1970, 1973, and 1983, respectively. He joined a national research laboratory (Mechanical Engineering Laboratory, Japan) in 1974, and moved to the University of Electro-Communications in 2001. His research interests include light propagation of near-infrared light in turbid media and its applications to biomedical photonics.