Methodology of effective glucose-specific signal extraction in complicated sample

Wenliang Chen, Bin Deng, Rong Liu, Xiaoyu Gu, Kexin Xu*
State key laboratory of precision measuring technology and instrument
College of Precision Instrument & Opto-Electronics Engineering
Tianjin University, Tianjin 300072, China

ABSTRACT

In the area of noninvasive human blood glucose concentration detecting, it has always been a critical task to extract the glucose-specific signal from the highly overlapped and disturbed near-infrared spectrum. In this paper, the methodology of effective glucose-specific signal extraction in complicated non-scattering sample is studied. By analyzing the impact of water displacement upon dissolution of glucose, the relationship between glucose concentration and absorption coefficient of the sample is deduced. Then, the reference wavelength where the absorption coefficient is insensitive to the changes of glucose concentration is put forward theoretically. Accordingly, the validating experiments in aqueous glucose solutions are executed. Both the theoretical and laboratorial results show that the reference wavelength of glucose appears at 1525nm. Based on the reference wavelength, an effective method for extracting the glucose-specific signal in complicated non-scattering samples is proposed and the corresponding validating experiments are constructed with different glucose and albumin concentration. Two different methods, traditional and the novel reference wavelength method are used to extract glucose signal and the corresponding root mean square error of prediction are 19.86mg/dl and 9.87mg/dl respectively. The experiment results indicate that the reference wavelength method can effectively eliminate the influence of various noises on the glucose-specific signal extraction, and thus can remarkably improve the measuring precision in noninvasive near-infrared glucose detecting.

Keywords: glucose-specific signal, reference wavelength, near-infrared

1. INTRODUCTION

Over the past few decades, non-invasive blood glucose monitoring has become an increasingly interesting and challenging field in the area of biomedical engineering, particularly after optical approaches have stepped into this endeavor [1-3]. However, there is still not yet a clinically feasible instrumentation available for continuous monitoring of human blood glucose concentration *in-vivo*. This is largely due to the sensitivity of instrument response to the glucose specific signals that is unfortunately swamped within a complex physiological system. Those undesired noises from the instrumental drift, environmental condition variations, absorption from other physiological components and differences between individual people all make it tougher to parse the glucose specific signal, thus, lead to the relatively poor predicting precision. Therefore, an effective solution to this will be not only meaningful but also inspiring for shedding a new light into the field of noninvasive blood glucose detecting.

* E-mail: kexin@tju.edu.cn, Phone: 86-22-27403944, Fax: 86-22-27406379

Optical Diagnostics and Sensing VII, edited by Gerard L. Coté, Alexander V. Priezzhev,
Proc. of SPIE Vol. 6445, 64450L, (2007) · 1605-7422/07/$18 · doi: 10.1117/12.699496

Generally, in the quantitative NIR (Near-Infrared) measurement, background deducting method is widely used to eliminate the undesired noises. In the traditional data processing procedure, the sample and background spectra are simply deducted. This is practically deficient, since it assumes the sample and background spectra share the same noises which require that the two spectra must be obtained under the same optical path simultaneously. It is obvious impossible. Moreover, it only takes into consider the concentration variation of glucose and assumes the content of all the other components are unchangeable. In the real complicated sample, however, the random differences within the concentrations of other components are inevitable. Therefore, the efficiency of the traditional background deducting method is rather limited. Gu [4] studied the relationship between the predicting precision of glucose concentration and the complexity of sample. When the traditional background deducting method was adopted, the low RMSEP (Root Mean Square Error of Prediction) of 9mg/dl could be obtained in the simple glucose and water solution, while in the more complicated sample, the RMSEP increased along with the complexity.

In this paper, the methodology of effective glucose-specific signal extraction in complicated non-scattering sample is studied. By analyzing the impact of water displacement upon dissolution of glucose, the relationship between glucose concentration and absorption coefficient of the sample is deduced. Then, the reference wavelength where the absorption coefficient is insensitive to the changes of glucose concentration is put forward theoretically. Based on the analysis of the mechanism of noise influence and reference wavelength, an effective method for extracting the glucose-specific signal in complicated non-scattering samples is proposed. Moreover, the corresponding validating experiments are constructed in glucose solutions with different glucose and albumin concentration. Two different methods, traditional and the novel reference wavelength method are used to extract glucose signal and the efficiency are compared respectively to evaluate the feasibility and practicability of this novel method.

2. THEORIES

2.1 Theoretical analysis of reference wavelength

Absorption coefficient originates from the Beer-Lambert relationship and corresponds to the wavelength dependent absorption strength for a certain sample. In the absence of molecular interactions between solutes, the total absorption coefficient is simply the summation of the individual values for the composite solutes [5].

$$\mu_a = \sum_{i=1}^{n} \left(\varepsilon_i \cdot c_i \right) \tag{1}$$

where, μ_a is the total absorption coefficient of a sample, ε_i and c_i are molar absorptivities and concentrations for each of the *n* components in the sample, respectively.

In this study, we focus only on the influence of glucose concentration variation on the whole absorptive coefficient of the solution and regard the other components as undesired signals. So, we assume that the molar concentration of all the solutes except glucose and water keep constant, and only consider the displacement of water by the individual glucose molecules.

According to the displacement effect, the molar concentration of water decreases with the increase of glucose concentration. This directly affects the total absorption coefficient of the solution in two ways as expressed in the following equation:

$$\Delta\mu_a = \varepsilon_g \cdot \Delta c_g + \varepsilon_w \cdot \Delta c_w \tag{2}$$

where, $\Delta\mu_a$ is the variation of total absorption coefficient, ε_g and ε_w are the molar absorptivities of glucose and water respectively, Δc_g and Δc_w are the molar concentration variation of glucose and water respectively.

According to the study of Weast [6], at $20^\circ C$, the molar concentration of water is

$$c_w^0 = 55.4 \text{ mol/L}$$

When the glucose concentration increases by 1mM, the molar concentration of water will decrease by 0.0111% [6]. Thus the relationship between the molar concentration of glucose and water can be defined as follows:

$$\Delta c_w = c_w^0 \cdot (-0.0111\%) \cdot \Delta c_g = -6.1494 \Delta c_g \tag{3}$$

Then, the final equation for $\Delta\mu_a$ is:

$$\Delta\mu_a = (\varepsilon_g - 6.1494 \cdot \varepsilon_w) \cdot \Delta c_g \tag{4}$$

We define λ_r as the reference wavelength under which the absorption coefficient is insensitive to the variation of glucose concentration. Then at the wavelength λ_r, the following condition must be fulfilled:

$$\Delta\mu_a(\lambda_r) = 0$$

So,

$$\frac{\varepsilon_g(\lambda_r)}{\varepsilon_w(\lambda_r)} = 6.1494 \tag{5}$$

According to our previous study [7], at the wavelength of 1525nm, the ratio of the molar absorptivities of glucose and water fits equation (5) best, which is:

$$\left.\frac{\varepsilon_g}{\varepsilon_w}\right|_{\lambda=1525nm} = \frac{5.416\times10^{-3}}{0.866\times10^{-3}} = 6.254 \tag{6}$$

Therefore, in this study, we consider 1525nm as the reference wavelength of glucose.

2.2 Traditional background deducting method

Typically, the impact of solvent can be compensated by using a background spectrum of blank solvent in quantitative near-infrared spectroscopy technique. It is a common approach to deducting a background spectrum to improve the specificity in glucose signal. According to the Lambert-Beer Law, if we do not take into account the scattering effect, the relationship between incident and transmitted light is:

$$I = I_0 \exp(-\mu_a \cdot l) \tag{7}$$

where, I_0 is the intensity of incident light, I is the intensity of transmitted light obtained from the detector, l is optical path length.

It is assumed that the difference between the background and sample lies only in the variation of glucose concentration. All the other components' concentrations in different samples are constant and in consistent with those in the background. Then, the light intensity of background and sample with ΔG glucose concentration variation can be expressed as:

$$I_b = I_{0-b} \exp(-\mu_a \cdot l) \tag{8}$$

$$I_s = I_{0-s} \exp(-\mu_a \cdot l - \Delta\mu_{a-\Delta G} \cdot l) \tag{9}$$

where, I_b and I_s are transmitted light intensities of background and sample, respectively. I_{0-b} and I_{0-s} are incident light intensities of background and sample, respectively. $\Delta\mu_{a-\Delta G}$ is the absorption coefficient variation of the sample caused by ΔG variation of glucose concentration.

When the measuring condition is precisely and strictly controlled, we can make the approximation that $I_{0-b} = I_{0-s}$. Then, the corresponding glucose-specific signal A_{glu} is:

$$A_{glu} = \Delta\mu_{a-\Delta G} \cdot l = -\ln \frac{I_s}{I_b} \tag{10}$$

Simultaneously collecting the sample and background spectra is the prerequisite to guarantee the identical incident light intensity of background and sample in this approach. In the experimental aspects, however, it is hard to realize. Moreover, we assume that only glucose concentration changes during the measurement, but actually, the concentration of other components may vary randomly. Therefore, more efficient method is highly needed to extract the glucose-specific signal.

2.3 Reference wavelength based glucose-specific signal extracting method

In the pragmatic application, however, the perfect situation assumed in section 2.2 can never be fulfilled. When there is a ΔG glucose concentration variation, we must take into consideration η, which is the random proportional incident light intensity deviation of sample from background. Moreover, the concentration variations of other absorptive molecules will also influence the absorption coefficient. Thus, combined with these disturbances, the sample's transmitted light intensity can be written as:

$$I_s = I_{0-b} \cdot \eta \cdot \exp\left(-\mu_a \cdot l - \Delta\mu_{a-\Delta G} \cdot l - \Delta\mu_{a-\Delta other} \cdot l\right) \tag{11}$$

where, $\Delta\mu_{a-\Delta other}$ is the absorption coefficient variation caused by components other than glucose.

Thus, the real glucose-specific signal at the wavelength of λ can be expressed as:

$$A_{glu}(\lambda) = -\ln\frac{I_s(\lambda)}{I_b(\lambda)} + \ln\eta - A_{\Delta other}(\lambda) \tag{12}$$

where, $A_{\Delta other}(\lambda) = \Delta\mu_{a-\Delta other}(\lambda) \cdot l$

At the reference wavelength of λ_r, the transmitted light intensity is insensitive to the variation of glucose concentration. Thus, the small fluctuation of the transmitted light intensity at λ_r can be regarded totally as the consequence of background noises. Technically, the dissolution of any solute will displace molecules of co-solutes from the optical path. But, at milli-molar concentrations, the effects of co-solute displacement are negligible and, to an approximation, can be ignored [8]. Thus, $A_{glu}(\lambda)$ under the wavelength of λ_r still keeps zero:

$$A_{glu}(\lambda_r) = -\ln\frac{I_s(\lambda_r)}{I_b(\lambda_r)} + \ln\eta - A_{\Delta other}(\lambda_r) = 0 \tag{13}$$

Rearrangement of equation (13) gives an expression for the absorbance changes caused by glucose concentration variation, as shown in equation (14):

$$\ln\eta - A_{\Delta other}(\lambda_r) = \ln\frac{I_s(\lambda_r)}{I_b(\lambda_r)} \tag{14}$$

Generally, the random variations of other unconcerned components' concentration are limited in a very small range. So the wavelength dependency of $A_{\Delta other}$ can be ignored and we can make an approximation that $A_{\Delta other}(\lambda) \approx A_{\Delta other}(\lambda_r)$. Therefore, according to equation (12) and equation (14), we can obtain the glucose-specific signal from the complicated non-scattering sample using the following background deducting method:

$$A_{glu}(\lambda) = -\ln\left(\frac{I_s(\lambda)}{I_b(\lambda)} \cdot \frac{I_b(\lambda_r)}{I_s(\lambda_r)}\right) \tag{15}$$

According to equation (15), in this reference wavelength data processing method, we only need to know the actual transmitted light intensity under the measuring and reference wavelength. The glucose concentration information will not be affected by the drift of light source and random noise. Meanwhile, the disturbance brought by the diversification of other absorptive particles can be remarkably reduced in complex samples. This novel background deducting method will largely enhance the practicability and precision of glucose concentration detecting by NIR spectroscopy.

3. EXPERIMENTS

3.1 Instrumentation

In this system, a tungsten-halogen lamp（Model PG64623, OSRAM, German） is adopted as the light source. And a non-collinear TeO_2 acousto-optic tunable filter (Model TEAF10-1.0-1.8-S, with Model VFI-80-50-DDS-B1-C2-E RF driver, Brimrose Company, U.S.A) is adopted to perform the wavelengths scan. Thus, monochromatic light ranging from 1100nm to 1750nm wavelengths can be obtained by time-sharing. When the monochromatic light diffracted by AOTF enters the sample in the quartz sample cell, the transmitted intensity is converted into analog signal by the InGaAs PIN photodiode (Model G5851-21, Hamamatsu Photonics K.K, Japan). And then, the signal is input into the computer after being digitized by 16-bit data acquisition card (Model PCI-MIO-16XE-50, National Instrument Inc., U.S.A). The computer controls the closed-loop process. The detailed description and layout of this system has been discussed in our previous study [9].

3.2 Validating experiments of the reference wavelength

To validate the existence of the reference wavelength, we add different amount of glucose into the distilled water to obtain glucose concentration graduation from 1000mg/dl to 2000mg/dl. Before sample scanning, the spectrum of distilled water is recorded as background. According to equation (10), the absorbance values can be obtained. The absorptive characteristics of those two different glucose concentrations are analyzed to validate the existence of the reference wavelength.

3.3 Validating experiments of the reference wavelength method

Different concentrations of albumin are added into the glucose solution to increase the complexity of the sample. In this way, the effect of the reference wavelength based glucose-specific signal extracting method can be evaluated. The distribution of the glucose and albumin concentration in all samples is shown as fellows:

Table 1. Sample concentration design in the validating experiments

	Calibration Set	Prediction Set
Concentration interval of glucose,	15	Random
Albumin(mg/dl)	5	0
Concentration range of glucose,	10-445	30,80,140,180,230,270,320,380,410
Albumin(mg/dl)	10-155	80
Number of samples	30	9

In order to diminish the impact of instrument drift, random sampling is adopted. Besides, before collecting every samples spectrum, the spectrum of distilled water is measured in advance as the background. We also try to make as shorter intervals between the sample and background spectra collecting as we can to avoid obvious instrument drift. Two different ways of data processing are carried out to obtain the specific information relating to the glucose concentration variation:

a. Traditional background deducting method: For every sample, we use the spectrum of distilled water which is recorded exactly right before this sample as the background. Glucose concentration information can be calculated according to equation (10).

b. Reference wavelength method: The first measured distilled water spectrum is used as the background for all samples, and the glucose concentration information of different samples is calculated using equation (15).

The PLS (Partial Least Squares) method [10] is applied to build the calibration models for the calibration set. The absorbance values obtained respectively from these two methods were used as spectra variables. Then the prediction set is used to estimate the prediction precisions of the two calibration models.

4. RESULTS AND DISCUSSION

4.1 Validating experiments of the reference wavelength

According to the experiments and data processing methods described in section 3.2, the relative absorbance values derived from different glucose concentrations are shown in Fig. 1.

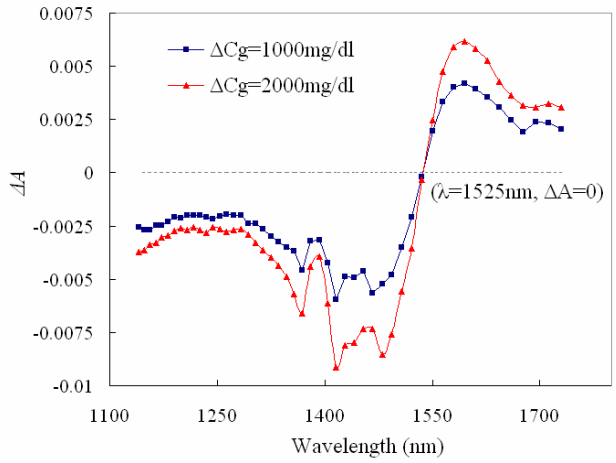

Fig. 1. The absorbance value curve with different glucose concentrations

Obviously, the curves intersect at the wavelength of 1525nm with relative absorbance values of zero. It approves that there is no overall influence of glucose concentration variation on the total absorption coefficient of the solution. Therefore, at the wavelength of 1525nm, the total absorption coefficient is insensitive to the variation of glucose concentration. This experiment result matches the theoretical deduction and establishes a solid foundation for the further application of this reference wavelength method in more complicated non-scattering sample.

4.2 Application of reference wavelength in complicated sample

We use the correlation coefficient, RMSEC (Root Mean Square Error of Calibration) and RMSEP (Root Mean Square Error of Prediction) to evaluate the efficiency of the two calibration models built on the traditional background deducting method and reference wavelength method respectively as described in section 3.3. Results are shown in the following Fig. 2 and Table 2, where “a” and “b” represent the two different methods in accordance with the description in section 3.3.

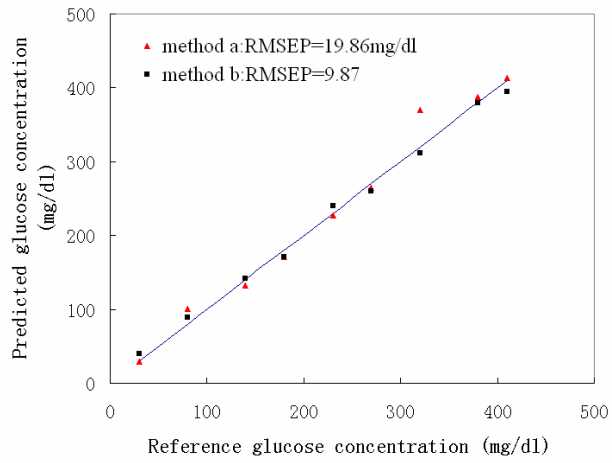

Fig. 2. The relation curves between the reference and predicted glucose concentration obtained from the traditional background deducting method (▲) and the novel reference wavelength method (■), respectively.

Table 2. Results of the two different calibration models

Methods for extracting A_{glu}	Calibration Models		Prediction results
	Correlation Coefficient	RMSEC (mg/dl)	RMSEP (mg/dl)
Method a	0.9960	10.07	19.86
Method b	0.9981	7.95	9.87

a). In the traditional background deducting method a, the classic "Sandwich" method is adopted. A background spectrum is collected immediately after the sample spectrum. Since we have rigidly confined the time intervals between all the sample and background spectra, the influences of the system shift and environmental noises are to some extent under control. However, the random disturbance is inevitable because the sample and background are not measured simultaneously after all. What's more, in the complicated variations of the background, this method can not eliminate the influences of other absorptive molecules, albumin as in this sample. Finally, larger predicting error is observed, as shown in Table 2. The RMSEC and RMSEP are 10.07mg/dl and 19.86 mg/dl, respectively.

b). Compared to method a, the additional reference wavelength based glucose-specific signal extracting method is adopted in the data pre-processing procedure. As what has been discussed before, it can effectively eliminate the disturbance brought by the time asynchronies and the concentration variation of other components. Its efficiency has been approved by the high predicting precision, with 7.95 mg/dl for RMSEC and 9.87 mg/dl for RMSEP respectively, which gives a better performance than the previous method.

5. CONCLUSIONS

In this study, the reference wavelength where the absorption coefficient is insensitive to the glucose concentration variation is put forward theoretically. Both the theoretical and laboratorial results show that the reference wavelength of glucose appears at 1525nm. At this wavelength, the small fluctuation of the transmitted light intensity can be regarded totally as the consequence of background noises. Thus, it can be used to eliminate the noises in the sample spectra. The

validating experiment in the complicated sample, which is composed of different glucose, albumin and water, approved that compared to the traditional background deducting method, the reference wavelength based glucose-specific signal extracting method can largely improve the predicting precision by 50%, with RMSEP of 19.86mg/dl and 9.87mg/dl respectively. The experimental results indicate that the reference wavelength method can effectively eliminate the influence of various noises on the glucose-specific signal extraction, and thus can remarkably improve the measuring precision in noninvasive near-infrared glucose detecting.

However, the application of this novel method is confined into the non-scattering samples at present. The further investigation of its wider use in complicated scattering sample is to be held in order to gain our final aim to apply this method in the in-vivo human blood glucose concentration detecting.

ACKNOWLEDGMENTS

Financial support for this work was provided by China Natural Science Foundation under Grant No. 30600126 and China Postdoctoral Science Foundation under Grant No. 20060390661.

REFERENCES

1. F. Sun, D. Kong, T. Mei, et al., "Near infrared spectral methods for non-invasively measuring blood glucose," *Proc. SPIE,* **5281:** 595-601, 2004.
2. O. S. Khalil, "Non-invasive glucose measurement technologies: an update from 1999 to the dawn of the new millennium," *Diabetes Technol. Ther.,* **6:**660-97, 2004.
3. J. T. Olesberg, L. Liu, M. A. Arnold, et al., "In Vivo Near-Infrared Spectroscopy of Rat Skin Tissue with Varying Blood Glucose Levels," *Anal. Chem.,* **78 :** 215-223, 2006
4. X. Gu, R. Liu Rong, W. Chen, et al., "The discussion about the complexity of samples and the prediction accuracy of glucose concentration in non-invasive blood glucose sensing," *Proc.SPIE,* **5696:** 222-231, 2005.
5. S. Fang. Sun, *Physical Chemistry of Macromolecules: Basic Principles and Issues,* 2nd Edition, John Wiley & Sons, Inc. Hoken, New Jersey, USA, 2004.
6. R. C. Weast and M. J. Astle, *CRC Handbook of Chemistry and Physics*, 63rd Edition, CRC press, Inc. Boca Raton, Florida, USA, 1982.
7. W. Chen, *Study on Theories and Experiments of Signal Acquisition for Non-Invasive Blood Glucose Sensing with Near-Infrared Spectroscopy*, Ph.D. dissertation, Tianjin University, Tianjin, China, 2005. (in Chinese)
8. A. K. Amerov, J. Chen and M. A. Arnold, "Molar Absorptivities of Glucose and Other Biological Molecules in Aqueous Solutions over the First Overtone and Combination Regions of the Near-Infrared Spectrum", *Appl. Spec.* **58:** 1195-1204, 2004.
9. K. Xu, Q. Qiu, J. Jiang, et al., "Non-invasive glucose sensing with near-infrared spectroscopy enhanced by optical measurement conditions reproduction technique," *Opt. Lasers Eng.,* **43:**1096-1106, 2005.
10. C. H. Spiegelman, M. J. Mcshane, M. J. Goetz, et al., "Theoretical justification of wavelength selection in PLS calibration: Development of a new algorithm," *Anal. Chem.,* **70:** 35-44, 1998.