Prediction of glucose in whole blood by near-infrared spectroscopy: Influence of wavelength region, preprocessing, and hemoglobin concentration

Yoen-Joo Kim
University of Texas at Austin
Center for Nano- and Molecular Science and Technology
Austin, Texas 78712

Gilwon Yoon
Seoul National University of Technology
Department of Electronics and Information
172 Kongneung-dong, Nowon-Gu
Seoul, Korea 139-743

Abstract. Measurement accuracy for predicting glucose in whole blood was studied based on near-infrared spectroscopy. Optimal wavelength regions, preprocessing, and the influence of hemoglobin were examined using partial least-squares regression. Spectra between 1100 and 2400 nm were measured from 98 whole blood samples. In order to study the influence of hemoglobin, which is the most dominant component in blood, 98 samples were arranged such that glucose and hemoglobin concentrations were distributed in their physiological ranges. Samples were grouped into three depending on hemoglobin level. The results showed that glucose prediction was influenced by hemoglobin concentrations in the calibration model. It was necessary for samples used in the calibration model to represent the entire range of hemoglobin level. The cross-validation errors were the smallest when the wavelength regions of 1390 to 1888 nm and 2044 to 2393 nm were used. However, prediction accuracy was not very dependent on preprocessing methods in this optimal region. The standard error of glucose prediction was 25.5 mg/dL and the coefficient of variation in prediction was 11.2%. [DOI: 10.1117/1.2342076]

Keywords: glucose; near-infrared; blood; partial least-squares; hemoglobin; noninvasive.

Paper 05250SSR received Aug. 28, 2005; revised manuscript received Feb. 19, 2006; accepted for publication Apr. 11, 2006; published online Aug. 31, 2006.

1 Introduction

Since the introduction of infrared as a dream beam,[1] infrared spectroscopy has been applied for measuring blood glucose noninvasively. There have been even some premature announcements of a noninvasive glucose monitor in the market, but still hope for and doubts of this technology prevail without a commercial product available at this time. There have been many investigations, for example, from an early scientific investigation by Robinson et al.[2] and several papers have reviewed this technology.[3–6]

Initial investigations for noninvasive glucose monitoring used a wavelength region of 700 to 1300 nm that contains higher orders of glucose overtone regions.[2,7,8] However, this wavelength region shows very little glucose absorption, for example, less than 0.1%, compared to the fundamental absorption region of 9 to 9.6 μm. Naturally, other glucose absorption regions were explored. They are the combination spectral region between 2.0 and 2.5 μm and the first overtone band of 1.52 to 1.85 μm. Based on the measurements at these bands, studies have been made with aqueous solutions mixed with some blood substances,[9–12] with blood[13,14] and *in vivo* experiments.[15,16]

The fundamental glucose absorption band lies in the mid-infrared (MIR) region. Due to interference with other substances, 9.0 and 9.6 μm are expected to be the most promising wavelengths to predict glucose absorption in the MIR region when interferences by other blood substances are also taken into account.[17] There have been various investigations on measuring glucose using the MIR region.[18–21] Unfortunately, the MIR region may not be used for *in vivo* monitoring because light penetration is limited to only several scores of micrometers depending on specific wavelengths.

The near-infrared (NIR) 1.5 to 2.5 μm band appears to be a suitable region for noninvasive glucose monitoring because it has higher glucose sensitivity compared with the second or third overtones and deeper penetration compared with the fundamental region. Basically, difficulties lie principally in weak glucose absorption, strong light scattering, and the interferences by other blood substances as well as other tissues. Initial enthusiasm from successful experiments with cuvette samples was often replaced by frustration when researchers performed *in vivo* experiments. A powerful tool, multivariate analysis, such as partial least-squares regression, may convert spectra to glucose fitting erroneously according to temporal or environmental correlation.[22,23]

Is it possible to achieve noninvasive glucose monitoring particularly using NIR absorption spectroscopy? What would

Address all correspondence to Gilwon Yoon, Seoul National University of Technology, Department of Electronics and Information, 172 Kongneung-dong, Nowon-Gu, Seoul, Korea 139-743; Tel: 82-2-970-6419; Fax: 82-2-979-7903; E-mail: gyoon@snut.ac.kr

1083-3668/2006/11(4)/041128/7/$22.00 © 2006 SPIE

be an order of achievable maximum accuracies? As one of the steps toward noninvasive glucose monitoring, whole blood samples were investigated in this study. Interestingly enough, there has been little investigation on glucose prediction using whole blood. Amerov et al. used bovine blood from a single blood matrix.[14] They had varied glucose concentrations. However, other blood components were the same. In our case, we used human whole blood. Our samples had different concentrations of glucose and hemoglobin as well as other blood substances. Because our research aim was to know how accurately blood glucose can be monitored, we used the different blood samples instead of a single matrix. Other human whole blood research that we are aware of was by Haaland et al.[13] They prepared blood samples from only four persons. Twenty samples from each person were made. Predictions in terms of the standard error of prediction (SEP) using the samples made from the same person ranged from 30.5 to 37.9 mg/dL. However, predictions based on the calibration model using a different individual were poor and they did not even reveal the numbers. They stated that different blood compositions were sufficiently different among the four subjects. In our study, the number of different blood samples was increased so that different blood chemistry was taken into account. We examined which optimal wavelength regions should be used to predict glucose concentrations in the NIR region. We also studied the effect of data preprocessing and the influence of hemoglobin that is the most dominant component in blood.

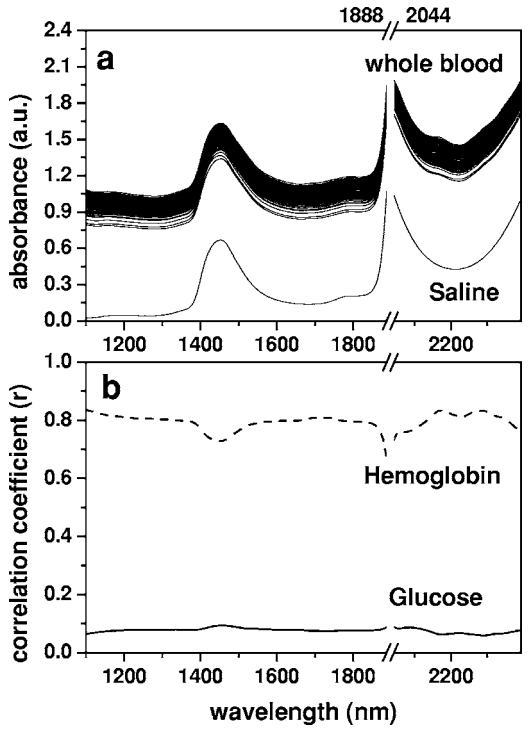

Fig. 1 (a) Whole blood spectra of 98 samples and saline spectrum; (b) whole blood spectra are correlated with hemoglobin and glucose concentrations at each wavelength and computed correlations coefficients are shown.

2 Experiments

A NIRSystems™ 6500 spectrometer equipped with silicon and lead-sulfide (PbS) detectors was used to measure the spectra of 98 blood samples. Each blood sample was made by pooling 3 to 5 EDTA whole blood samples where both blood types (ABO and Rh) and hemoglobin concentrations were being checked. Pooling blood was required to ensure that there was enough blood volume when preparing each sample that was used not only for spectrum measurement but also for reference value measurement. Glucose stock solution of 20 g/dL in saline was added to blood samples to control glucose concentrations. Highly concentrated glucose stock was added into a different blood sample to make a blood sample with a particular glucose concentration. No dilution of blood was made. First, we had information on hemoglobin concentration for every extracted blood sample. We mixed (or pooled) 3 to 5 extracted blood samples of similar hemoglobin concentrations to make one blood sample. By doing this, we had enough blood volume for each sample. Also during this process, we could arrange the samples so that their hemoglobin concentrations were distributed in the entire physiological range. After that, we added glucose to assign different glucose values such that glucose and hemoglobin concentrations are not correlated to each other.

Spectra were measured by a Foss™ NIR 6500 system between 1100 nm and 2500 nm with a step of 2 nm. One scan time was set to 1 s and 32 scan data were averaged to produce a spectrum. The system signal-to-noise ratio of measured spectrum was 10^{-5} absorbance that was computed from two consecutively acquired spectra. Each spectrum was obtained between 1100 and 2500 nm. Whole blood was contained in a 0.5 mm detachable cell. The spectrum of the blood sample was measured. The spectrum without the cell was also measured and used for reference. The absorbance spectrum was obtained from these two single beam intensities. Immediately after each measurement, a portion of blood was centrifuged and the plasma was frozen to measure glucose concentration. A Beckman™ chemistry analyzer based on the glucose hexokinase method was used to measure plasma glucose. Using another portion of the same blood, hemoglobin concentration was measured by the HiCN method using a SYSMEX™ instrument.

Measured glucose and hemoglobin for 98 samples ranged from 45 to 432 mg/dL and 7.5 to 16.6 g/dL, respectively. Figure 1(a) shows measured transmission spectra. Whole blood shows higher absorption than saline [see Fig. 1(a)]. In our figures, values around 1940 nm, a very strong water absorption peak, are not shown in order to increase the dynamic range of the *y* axis. Correlation coefficients between hemoglobin or glucose values with respect to whole blood absorbance for all the samples were calculated at each wavelength. Correlation coefficients of hemoglobin and glucose with respect to absorbance at each wavelength are shown in Fig. 1(b). Correlation coefficients for hemoglobin are around 0.8 and those for glucose are smaller than 0.1. This indicates that measured absorbance spectra varied mainly depending on hemoglobin level.

2.1 *Wavelength Selection*

The region between 1100 and 2400 nm includes the first overtone and combination bands. It is necessary to choose a specific wavelength region that minimizes prediction errors.

Table 1 Predictions of glucose concentrations at various spectral regions. Spectra of all 98 samples were used. The best SECV and VC_{CVal} were obtained when 1390 to 1888 and 2044 to 2392 nm were used.

Spectral region (nm)	N[a]	#f[b]	SECV[c] rCVal[d]	SEC[e] rCal[f]	VC_{CVal}[g] (%)
1100–2498	700	14	51.4	21.5	24.1
			0.9008	0.9860	
1100–1888	570	9	27.4	23.5	12.9
2044–2392			0.9728	0.9822	
1390–1888	425	8	26.1	21.6	12.3
2044–2392			0.9755	0.9847	
1516–1816	297	7	34.1	30.2	16.0
2062–2352			0.9575	0.9696	
1100–1888	395	10	42.2	31.7	19.8
			0.9345	0.9677	
1390–1888	250	9	39.6	32.2	18.6
			0.9425	0.9663	
2044–2392	175	6	40.8	33.1	19.2
			0.9389	0.9630	

[a]*N*: number of variables used for PLSR analysis.
[b]#*f*: optimal number of factors.
[c]SECV (mg/dL): standard error of cross validation.
[d]rCVal: correlation coefficient of cross validation.
[e]SEC (mg/dL): standard error of calibration.
[f]rCal: correlation coefficient of calibration.
[g]VC_{CVal} (%): coefficient of variation in cross validation, SECV/mean$\times$100.

We performed partial least-squares regression (PLSR) analysis by using PIROUETTE™ 2.6 software (Infometrix Inc). All 98 samples were examined. Before calibration, spectra were mean-centered. We used all the sample data. Loading vectors were analyzed to examine the influence of wavelength. Our previous work,[24] has more detailed descriptions on loading vector analysis and wavelength band selections. A similar approach was adapted in this investigation. In choosing the wavelength ranges, a region between 1.5 and 1.8 μm (first overtone band) and a region between approximately 2 to 2.4 μm (combination band) were considered. Also, the entire range of 1.1 to 2.4 μm was included as one of the regions. A region around 1940 nm has a higher water absorption peak and hemoglobin absorption increases toward 1100 nm. Therefore, the elimination of 1940 and 1100 nm peaks produced further wavelength regions (Table 1). Table 1 summarizes glucose prediction at various spectral regions. For each spectral region, we computed the standard error of cross validation (SECV), correlation coefficient of cross validation (rCVal), standard error of calibration (SEC), correlation coefficient of calibration (rCal), and coefficient of variation in cross validation (VC_{CVal}). An optimum number of factors used in the regression were determined by the leave-one-out cross validation and *F* test with a significance level of 5% among the factors. The best result was obtained using the regions of 1390 to 1888 and 2044 to 2392 nm where SECV is the least.

Over 1100 to 2392 nm, except for a water absorption peak around 1940 nm, we plotted the first through third loading vectors and regression vectors of a glucose calibration model that used all 98 samples (Fig. 2). Loading vectors are shown together with glucose spectrum [Fig. 2(b)] and hemoglobin spectrum [Fig. 2(c)] whose values were measured from saline solutions. Therefore, a spectrum of glucose or hemoglobin in Fig. 2 was calculated by subtracting saline spectra. Hemoglobin was prepared by the blood cell lysis method described in the Ref. 25.

2.2 *Data Preprocessing and Enhancement*

Various data preprocessing methods have been utilized to improve calibration and prediction modeling. In this study, multiplicative scatter correction (MSC)[26] and standard normal variate (SNV)[27] were tested in order to minimize the scattering effect of blood cells. In addition, the second derivative method that has been widely used to remove baseline variations was applied. Fifteen or 25 points smoothing was made before differentiation to reduce noises. After preprocessing, mean centering was applied for data enhancement. Figure 3 shows final spectra processed by MSC, SNV, and the second

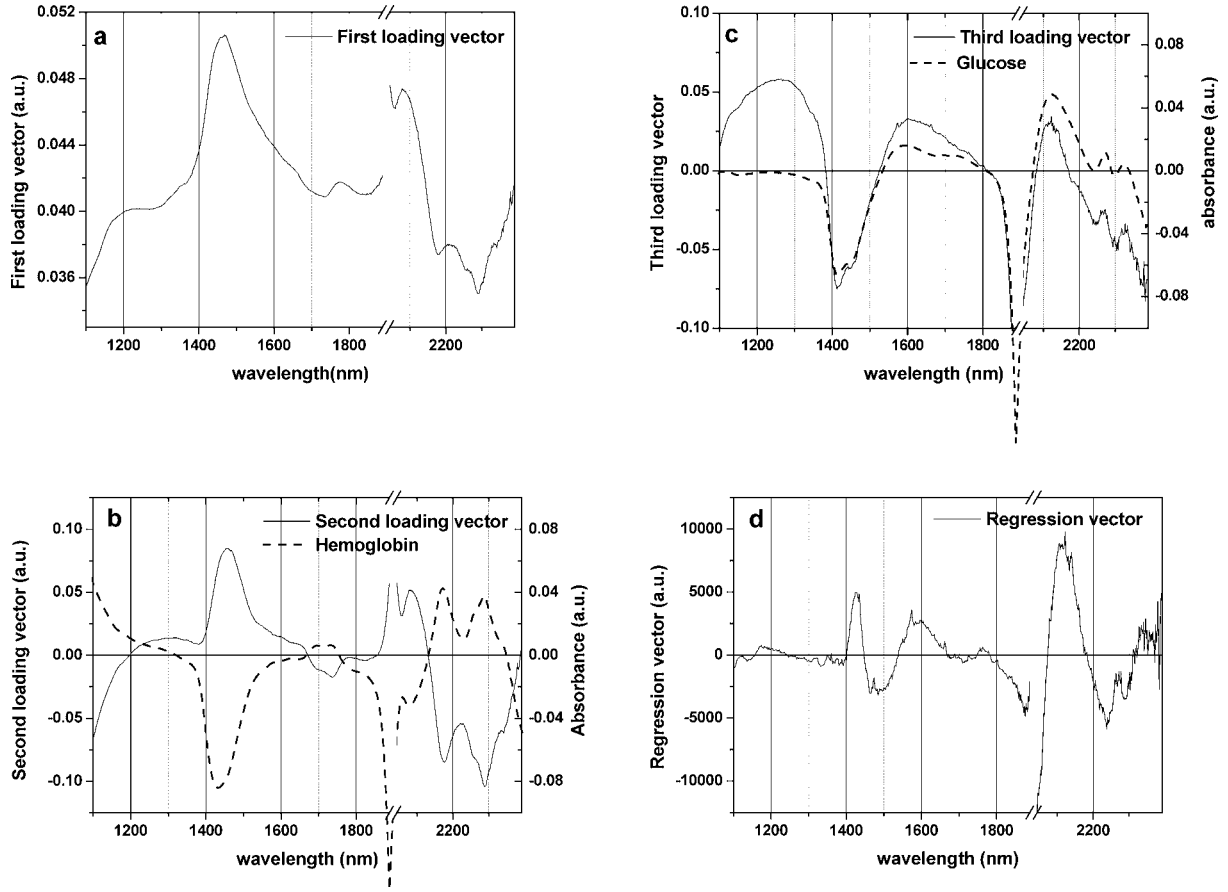

Fig. 2 Calibration modeling based on the PLSR analysis was done for all 98 blood samples: (a) first loading vector, (b) second loading vector, (c) third loading vector, and (d) regression vector.

derivatives. In order to study the effect of preprocessing, glucose concentrations were predicted. In this case, we used the wavelength bands of 1390 to 1888 and 2044 to 2392 nm that produced the best results in Sec. 2.1. The results were summarized in terms of SECV and VC_{cval} as shown in Table 2.

2.3 *Influence of Hemoglobin Level*

Hemoglobin is the most dominant component in blood, and its concentration level is more than 100 times of glucose. Its absorbance becomes increasingly strong toward short NIR and visible wavelengths. Even though hemoglobin absorption peaks do not interfere with the peaks of other components, its influence is by no means negligible due to its high concentration.[4,17] Therefore, it is expected that calibration and prediction modeling can be substantially influenced by hemoglobin level.

To study hemoglobin influence, 98 samples were divided into several groups. First, the entire samples were divided into two groups that are the calibration set (63 samples) and the prediction set (35 samples). Both groups were arranged so that hemoglobin and glucose concentrations were evenly distributed. Next, all the samples were grouped into three groups depending on hemoglobin level (Hb_{high}: 16.6 to 13 g/dL; Hb_{mid}: 12.8 to 10.9 g/dL; Hb_{low}: 10.7 to 7.5 g/dL). Each of three groups had glucose concentrations evenly distributed in the entire range. Table 3 summarizes the groups, ranges of

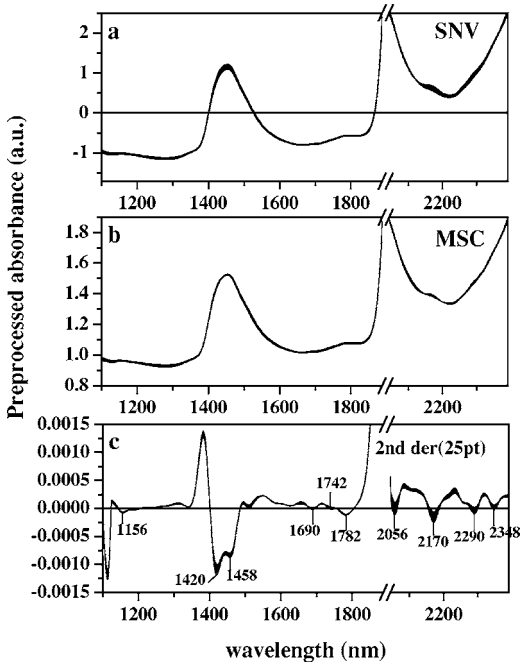

Fig. 3 Whole blood spectra preprocessed by (a) SNV, (b) MSC, (c) second derivatives.

Table 2 The effects of spectral data preprocessing in terms of SEC: all 98 samples were calibration modeled using different preprocessed spectra at the wavelengths of 1390 to 1888 and 2044 to 2392 nm.

Preprocessing method	#f	SECV rCVal	SEC rCal	VC_{CVal} (%)
Mean centering	8	26.1	21.6	12.3
		0.9755	0.9847	
MSC	6	26.7	23.8	12.5
mean centering		0.9746	0.9810	
SNV	7	26.9	22.0	12.6
mean centering		0.9738	0.9810	
Second derivative (15)	9	41.6	26.8	19.5
mean centering		0.9367	0.9768	
Second derivative (25)	9	55.1	27.5	25.9
mean centering		0.8850	0.9754	

hemoglobin and glucose, and their standard deviations. It is important that hemoglobin and glucose concentrations in each group are not correlated. All five groups were checked for the correlation between hemoglobin and glucose concentrations, and we verified that the correlations were negligible as can be seen in terms of the correlation coefficient, r (Table 3).

Calibration modeling was performed using the four calibration groups (Hb_{cal}, Hb_{high}, Hb_{mid}, and Hb_{low}). Wavelength bands of 1390 to 1888 nm and 2044 to 2392 nm with mean centering were applied in PLSR analysis. The results were shown in Table 4. Table 5 displays SEP and correlation coefficient of prediction (r_{Pre}). Because glucose values are different among the prediction sets, prediction accuracy was analyzed in terms of the coefficient of variation in prediction (VC_{Pre}). VC_{Pre} is defined as (SEP/mean value of glucose)

Table 3 Concentration distributions of hemoglobin and glucose in different sample groups.

Group	M[a]	Component[b]	Min	Max	Mean	Standard deviation	r[c]
Hb_{total}	98	Hemoglobin	7.5	16.6	12.1	2.2	−0.0504
		Glucose	45	432	213	119	
Hb_{cal}	63	Hemoglobin	7.9	16.6	12.2	2.2	−0.1310
		Glucose	45	432	205	121	
Hb_{pre}	35	Hemoglobin	7.5	16.3	12.0	2.2	0.1066
		Glucose	54	428	228	116	
Hb_{high}	36	Hemoglobin	13	16.6	14.4	1.1	−0.0514
		Glucose	50	424	207	115	
Hb_{mid}	31	Hemoglobin	10.9	12.8	11.9	0.7	0.0027
		Glucose	46	432	213	121	
Hb_{low}	31	Hemoglobin	7.5	10.7	9.7	0.8	0.0228
		Glucose	45	428	221	125	

[a]M is the number of samples.
[b]The units are g/dL for hemoglobin and mg/dL for glucose.
[c]r: correlation coefficient between hemoglobin and glucose.

Table 4 The result of calibration models for glucose determination from the four calibrations sets of different hemoglobin levels. PLSR was performed using the bands of 1390 to 1888 and 2044 to 2392 nm with mean centering.

Group	Mean value of glucose (mg/dL)	*M*	*#f*	SECV rCVal	SEC rCal	VC_{CVal} (%)
Hb_{cal}	205	63	8	27.5	20.9	13.4
				0.9829	0.9870	
Hb_{high}	207	36	7	36.4	25.1	17.6
				0.9477	0.9807	
Hb_{mid}	213	31	7	38.3	25.6	18.0
				0.9470	0.9826	
Hb_{low}	221	31	8	29.0	16.4	13.1
				0.9724	0.9937	

$\times$ 100 expressed as a percentage. The results are summarized in Table 5.

3 Results and Discussion

Before further analysis of data preprocessing and hemoglobin influence, an optimal wavelength region that provided the least calibration errors was obtained. SECV varied widely between 26.1 and 51.4 mg/dL, while various wavelength regions in 1100 to 2498 nm were tested (Table 1). The best results were achieved when the regions of 1390 to 1888 nm and 2044 to 2392 nm were used. The regions contain both first overtone and combination bands. The optimal region included a water absorption peak at 1440 nm in the first overtone band, but excluded a water absorption peak of 1940 nm and wavelengths shorter than 1390 nm.[28] As can be observed in Fig. 1(a), the region between 1100 and 1390 nm shows different slopes between hemoglobin and glucose. Absorption of saline decreases as the wavelength becomes shorter. This is a typical feature of the water absorption spectrum. On the other hand, blood absorption increases toward 1100 nm. This reflects hemoglobin absorption. When 1100 to 1390 nm was included, SECV increased from 21.6 mg/dL to 27.4 mg/dL.

Figure 2 shows loading vectors between 1100 and 2392 nm. The first loading vector appears to represent a spectral profile of blood to some degree [Fig. 2(a)]. The second loading vector is similar to hemoglobin spectrum, but it is a mirror image [Fig. 2(b)]. A spectral pattern of glucose looks similar to that of the third loading vector although there is a mismatch at wavelengths shorter than 1390 nm [Fig. 2(c)]. It is interesting to note that the exclusion of 1100 to 1390 nm produced better glucose prediction. Figure 2(d) illustrates regression vectors. The high absolute value of the regression vector indicates high contribution to glucose calibration at that wavelength. No contribution of 1100-1390 nm is again observed in Fig. 2(d). However, strong influences by two water absorption peaks (1440 and 1940 nm) are depicted in Fig. 2.

Applying scattering correction methods of MSC and SNV did not improve the prediction accuracy as can be seen in Table 2. Figure 3 shows preprocessed spectra by MSC, SNV, and second derivatives. In the case of the second derivative method that has been widely used for baseline correction, the results were the worst and produced higher numbers of factors. For the second derivatives, negative peaks appeared at 1420, 1458, 1690, 1742, 1782, 2056, 2170, 2290, and 2348 nm. Peaks at 1420 and 1458 nm belong to the water absorption band and the rest are close to the peaks in the second derivative spectra of hemoglobin (1690, 1740, 2056, 2170, 2290, and 2350 nm) given by Kuenstner and Norris.[25] This indicates that whole blood spectra are dominated by hemoglobin spectra. Hemoglobin features appear to be more enhanced than glucose features during differentiation.

The influence of hemoglobin concentrations in the samples was summarized in Table 4. Calibration modeling using Hb_{cal} had SECV of 27.5 mg/dL and VC_{CVal} of 13.4%. When calibration models from the sets of high, medium, and low hemoglobin levels (Hb_{high}, Hb_{mid}, and Hb_{low}, respectively) were performed, SECV ranged between 29 and 38 mg/dL and VC_{CVal} varied from 13.1 to 18.0%. Glucose concentrations were predicted and the results were summarized in Table 5. Based on the calibration model using 63 samples (Hb_{cal}), glucose values of the other 35 samples (Hb_{pre}) were predicted. SEP was 25.5 mg/dL where the mean value of glucose was 228 mg/dL and VC_{pre} was 11.2%. Cross predictions among the different groups of hemoglobin levels were made. SEPs varied a great deal depending on the groups. We observed SEPs of 23.1 to 74.2 mg/dL and VC_{pre}s of 10.8 to 22% (Table 5). The more difference in hemoglobin level between the sets, the higher prediction error appeared to be. For example, VC_{pre} was 35.8% when Hb_{high} was predicted based on the calibration model of Hb_{low}. When the calibration model based on Hb_{high} predicted glucose concentrations of Hb_{low}, VC_{pre} was 22%. The highest values were 35.8 and 22%. It is observed that hemoglobin distribution in the calibration or prediction model influenced the accuracy substantially. It is expected that the calibration model should use a sample set

Table 5 Prediction of glucose concentrations based on the four calibration models.

Calibration set	Prediction set	Mean value of glucose	SEP[a] (r_{Pre})[b]	VC_{Pre}[c] (%)
Hb_{cal}	Hb_{pre}	228	25.5 (0.9764)	11.2
Hb_{high}	Hb_{mid}	213	23.1 (0.9817)	10.8
	Hb_{low}	221	48.7 (0.9279)	22.0
Hb_{mid}	Hb_{high}	207	39.3 (0.9465)	19.0
	Hb_{low}	221	46.9 (0.9328)	21.2
Hb_{low}	Hb_{high}	207	74.2 (0.8672)	35.8
	Hb_{mid}	213	33.8 (0.9603)	15.9

[a]SEP (mg/dL): standard error of prediction for glucose.
[b]r_{Pre}: correlation coefficient of prediction.
[c]VC_{Pre} (%): coefficient of variation in prediction of glucose, SEP/mean $\times$ 100.

consisting of all physiological ranges for hemoglobin levels.

4 Summary

First overtone band or combination band alone was not a sufficient wavelength region in predicting glucose in whole blood. The region including both bands, but excluding a water absorption peak of 1940 nm, gave better prediction. A simple mean centering as a data preprocessing method produced good results in the optimal wavelength region. However, we may have to limit our statement to our particular case of PLSR analysis and whole blood samples because the generalization about preprocessing may be dependent on a particular multivariate method and samples. When whole blood was dealt with, hemoglobin concentrations in the calibration model should represent an entire range of hemoglobin. We have not found previous investigations where the actual hemoglobin concentrations were analyzed either for blood analysis or for *in vivo* experiment. We obtained a SEP of 25.5 mg/dL where blood glucose ranged between 45 and 432 mg/dL. The coefficient of variation in prediction was 11.2%. For noninvasive glucose monitoring, person-to-person blood chemistry as well as tissue variations make situations more difficult. When individual calibration (i.e., personal use) is adapted, the problem of person-to-person variation can be avoided. The personal calibration is recommended as a first step toward a noninvasive glucose monitor.

Acknowledgments

This work was supported in part by the Ministry of Science and Technology of Korea through the Cavendish-KAIST Cooperative Research Program. We thank Haemin Cho for spectrum measurements.

References

1. I. Amato, "Race quickens for non-stick blood monitoring technology," *Science* **258**, 892–893 (1992).
2. M. R. Robinsons, R. P. Eaton, D. M. Haaland, G. W. Koepp, E. V. Thomas, B. R. Stallard, and P. L. Robinson, "Noninvasive glucose monitoring in diabetic patients: A preliminary evaluation," *Clin. Chem.* **38**(9), 1618–1622 (1992).
3. H. M. Heise, "Non-invasive monitoring of metabolites using near infrared spectroscopy: state of the art," *Horm. Metab. Res.* **28**, 527–534 (1996).
4. O. S. Khalil, "Spectroscopic and clinical aspects of noninvasive glucose measurements," *Clin. Chem.* **45**(2), 165–177 (1999).
5. R. McNichols and G. L. Cote, "Optical glucose sensing in biological fluids: an overview," *J. Biomed. Opt.* **5**(1), 5–16 (2000).
6. A. Sieg, R. H. Guy, and M. B. Delgado-Charro, "Noninvasive and minimally invasive methods for transdermal glucose monitoring," *Diabetes Tech. Ther.* **7**(1), 174–197 (2005).
7. R. D. Rosenthal, "Method for providing general calibration for near infrared instruments for measurement of blood glucose," U.S. Patent No. 5,204,532 (1993).
8. U. A. Muller, B. Mertes, C. Fischbacher, K. U. Jageman, and K. Danzer, "Non-invasive blood glucose monitoring by means of near infrared spectroscopy: methods for improving the reliability of the calibration models," *Int. J. Artif. Organs* **20**(5), 285–290 (1997).
9. M. R. Riley, M. A. Arnold, and D. W. Murhammer, "Matrix-enhanced calibration procedure for multivariate calibration models with near-infrared spectra," *Appl. Spectrosc.* **52**(10), 1339–1347 (1998).
10. G. Yoon, A. K. Amerov, K. J. Jeon, and Y.-J. Kim, "Determination of glucose concentration in scattering medium using selected wavelengths at overtone absorption band," *Appl. Opt.* **41**(7), 1469–1475 (2002).
11. L. Zhang, G. W. Small, and M. A. Arnold, "Multivariate calibration standardization across instruments for the determination of glucose by Fourier transform near-infrared spectroscopy," *Anal. Chem.* **75**, 5905–5915 (2003).
12. 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**, 5405–5413 (2004).
13. D. M. Haaland, M. R. Robinson, G. W. Koepp, E. V. Thomas, and R. P. Eaton, "Reagentless near-infrared determination of glucose in whole blood using multivariate calibration," *Appl. Spectrosc.* **46**(10), 1575–1578 (1992).
14. A. K. Amerov, J. Chen, G. W. Small, and M. A. Arnold, "Scattering and absorption effects in the determination of glucose in whole blood by near-infrared spectroscopy," *Anal. Chem.* **77**, 4587–4594 (2005).
15. J. J. Burmeister, M. A. Arnold, and G. W. Small, "Noninvasive blood glucose measurements by near-infrared transmission spectroscopy across human tongues," *Diabetes Tech. Ther.* **2**(1), 5–16 (2000).
16. K. Maruo, M. Tsurugi, J. Chin, T. Ota, H. Arimoto, Y. Yamada, M. Tamura, M. Ishii, and Y. Ozaki, "Noninvasive blood glucose assay using a newly developed near-infrared system," *IEEE J. Sel. Top. Quantum Electron.*," **9**(2), 322–330 (2003).
17. H. Zeller, P. Novak, and R. Landgraf, "Blood glucose measurement by infrared spectroscopy," *Int. J. Artif. Organs* **12**(2), 129–135 (1989).
18. H. M. Heise, R. Marbach, G. Janatsch, and J. D. Kruse-Jarres, "Multivariate determination of glucose in whole blood by attenuated total reflection infrared spectroscopy," *Anal. Chem.* **61**, 2009–2015 (1989).
19. R. A. Shaw, S. Kotowich, M. Leroux, and H. H. Mantsch, "Multivariate serum analysis using mid-infrared spectroscopy," *Ann. Clin. Biochem.* **35**, 624–632 (1998).
20. S. Hahn, G. Yoon, G. Kim, and S.-H. Park, "Reagentless determination of human serum components using infrared absorption spectroscopy," *J. Opt. Soc. Korea* **7**(4), 240–244 (2003).
21. Y.-J. Kim, S. Hahn, and G. Yoon, "Determination of glucose in whole blood samples by mid-infrared spectroscopy," *Appl. Opt.* **42**(4), 745–749 (2003).
22. M. A. Arnold, J. J. Burmeister, and G. W. Small, "Phantom glucose calibration models from stimulated noninvasive human near-infrared spectra," *Anal. Chem.* **70**, 1773–1781 (1998).
23. R. Liu, W. Chen, X. Gu, R. K. Wang, and K. Xu, "Chance correlation in non-invasive glucose measurement using near-infrared spectroscopy, *J. Phys. D* **38**, 2675–2681 (2005).
24. Y.-J. Kim and G. Yoon, "Multicomponent assay for human serum using mid-infrared transmission spectroscopy based on component-optimized spectral region selected by first loading vector analysis," *Appl. Spectrosc.* **56**(5), 625–632 (2002).
25. J. T. Kuenstner and K. H. Norris, "Spectrophotometry of human hemoglobin in the near infrared region from 1000 to 2500 nm," *J. Near Infrared Spectrosc.* **2**, 59–65 (1994).
26. H. Martens and T. Næs, *Multivariate Calibration*, John Wiley & Sons Ltd., New York (1989).
27. M. S. Dhanoa, S. J. Lister, R. Sanderson, and R. J. Barnes, "The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformations of NIR spectra," *J. Near Infrared Spectrosc.* **2**, 43–47 (1994).
28. G. M. Hale and M. R. Querry, "Optical constants of water in the 200-nm to 200-μm wavelength region," *Appl. Opt.* **12**(3), 555–563 (1973).