

近赤外瞬時差分分光法による非観血的血糖計測のための 多変量校正モデルの検討

山越 康弘^{*,**}・小川 充洋^{**}・山越 健弘^{***}・田村 俊世^{*}・山越 憲一^{***}

Multivariate Calibration Models for Non-invasive Prediction of Blood Glucose Level Using an Instantaneous Differential Near-infrared Spectrophotometry

Yasuhiro YAMAKOSHI^{*,**}, Mitsuhiro OGAWA^{**}, Takehiro YAMAKOSHI^{***}, Toshiyo TAMURA^{*},
Ken-ichi YAMAKOSHI^{***}

Abstract An optical method recently proposed for non-invasive *in vivo* blood glucose concentration (BGL) measurement, named "Pulse Glucometry", was combined and compared with four multivariate analyses for constructing calibration models: Principal Component Regression (PCR), Partial Least Squares Regression (PLS), Artificial Neural Network (ANN), Support Vector Machines Regression (SVMsR). A very fast spectrophotometer for "Pulse Glucometry" provides the total transmitted radiation spectrum (I_λ) and the cardiac-related pulsatile component (ΔI_λ) superimposed on I_λ in human fingertips over a wavelength range from 900 to 1700 nm with resolution of 8 nm in 100 Hz sampling. From a family of I_λ s measured, which include information relating to blood constituent such as BGL values, differential optical densities (ΔOD_λ s, where $\Delta OD_\lambda = \text{Log}(1 + \Delta I_\lambda / I_\lambda)$) were obtained and normalized by the ΔOD_λ values at 1100 nm. Finally, the 2nd derivatives of the normalized ΔOD_λ s ($\Delta^2 OD_\lambda$ s) along wavelengths were calculated as regressors. Subsequently, calibration models from paired data sets of regressors (the values of $\Delta^2 OD_\lambda$ s) and regressand (the corresponding known BGL values) were constructed with PCR, PLS, ANN and SVMsR. The results show that each calibration model provides a relatively good regression with a modified 5-fold cross validation for total 95 paired data, in which the BGLs ranged from 100.7–246.3 mg/dl. The results were evaluated by the Clarke error grid analysis and all data points obtained from all calibration models fell within the clinically acceptable regions (region A or B). Among them, ANN and SVMsR calibration provided the best plot distributions (in ANN; Region A: 77 plots (81.1%), B: 18 plots (18.9%). in SVMsR; Region A: 78 (82.1%), B: 17 (17.9%)). Total calculation time of SVMsR is about 100 times shorter than ANN. These results suggest that a calibration model using SVMsR is highly promising for "Pulse Glucometry."

Keywords: pulse glucometry, non-invasive blood glucose measurement, multivariate analysis, support vector machines, nonlinear problem.

1. は じ め に

生活習慣病の代表である糖尿病は、最近の高齢化や生活習慣の変化等から増加の一途を辿っている。糖尿病が進行した場合、食事・運動療法に加えインスリン等の薬剤治療が必要であり、患者は一日頻回にわたり血糖値を計測し、

血糖コントロールを余儀なくされる[1]。現在、糖尿病患者の日常の血糖値計測は微量の採血を伴う SMBG (self-monitoring of blood glucose: 血糖自己測定) 用装置を用いて行うのが一般的であるが、患者にとって痛みと煩わしさ、安全性等の問題があり、非観血で血糖計測できる方法が強く望まれている。採血しない方法としては、Glucose Watch Biographer が唯一アメリカ食品医薬品局 (FDA: Food and Drug Administration) より認可されている。これはイオントフォoresis [2, 3] を用いて組織液よりグルコースを含む成分をグルコース感応電極まで誘導して計測を行うものであり、完全に非侵襲的ではなく、繰り返し使用が皮膚の炎症の原因となることも知られている[4]。

非観血的な血糖計測法としては、光音響分光法によるもの[5]、ラマン分光法によるもの[6, 7]なども提案されてい

2007 年 8 月 20 日受付, 2007 年 11 月 20 日改訂
Received: August 20, 2007; revised November 20, 2007

* 千葉大学大学院工学研究科

Graduate School of Engineering, Chiba University

** (株) ユーシス

yu.sys Corporation

*** 金沢大学大学院自然科学研究科

Graduate School of Natural Science and Technology, Kanazawa University

るが、近赤外分光を利用した試みも多い[8-10]。しかし、単純な透過光あるいは反射光の分光分析を基本とした方法では、皮膚や骨組織等の血液以外の生体成分に由来する妨害物質の影響が支配的であると考えられ、未だ実用に供しえる精度で血糖値を計測することは困難であるというのが現状である[11]。

筆者らは、簡便で高精度な非観血血糖値計測を実現することを最終目的に、生体透過近赤外光の瞬時計測により生体組織内の血液容積変動に伴う透過光変化から血液成分に由来する信号成分を抽出し、その時間差分から吸光度スペクトルを算出して、これに多変量解析法の一つである PLS 法を適用して血糖値を予測する方法を提案した[12]。筆者らはこれをパルスグルコメトリ (Pulse Glucometry) と命名し、これまで方法論の基礎的な検討を行ってきた[13]。パルスグルコメトリ法による血糖値計測を実用にするためには、装置の小型化、糖尿病患者も含めた大規模なデータ収集や多変量校正モデルの改良などの問題を克服しなければならないが、その中でも血糖値予測に関わる校正モデルの構築は重要な課題として位置づけられる。本論文では、予測精度の向上を目的とした校正モデル構築法の改良のために複数の多変量解析手法を本法に導入し、それらの比較と検討を行ったので報告する。

2. 方 法

2.1 原理

パルスグルコメトリの原理の詳細は既に報告[12, 13]されているので、ここでは概要のみを示す。

心拍動による血液容積変動に伴う光電脈波(透過光変化)を観測できることは一般に知られているが、この現象について図 1 左上に示す単純な光学モデルを仮定する。このモデルに示すように、生体組織を「動脈」、「静脈」、「血管以外の組織」の 3 つの要素から成るものとした場合、波長 λ における透過光強度の脈動成分は動脈容積の変化のみに由来するものと考えられる。このとき、波長 λ における透過光強度の時系列を $I_{\lambda 1}(t)$ とし、入射光強度を $I_{0\lambda 1}$ とすると、時刻 t における生体の光学密度 (Optical Density) $OD_{\lambda 1}(t)$ は時間の関数であり、以下の式で与えられる。すなわち、

$$OD_{\lambda 1}(t) = \log \left\{ \frac{I_{0\lambda 1}}{I_{\lambda 1}(t)} \right\} \quad (1)$$

このとき、任意の時刻 t_1 と t_2 での光学密度の差分 $\Delta OD_{\lambda 1}(t_1, t_2)$ を考えると、以下のようになる。

$$\begin{aligned} \Delta OD_{\lambda 1}(t_1, t_2) &= \log \left\{ \frac{I_{0\lambda 1}}{I_{\lambda 1}(t_2)} \right\} - \log \left\{ \frac{I_{0\lambda 1}}{I_{\lambda 1}(t_1)} \right\} \\ &= \log \left\{ \frac{I_{\lambda 1}(t_1)}{I_{\lambda 1}(t_2)} \right\} \end{aligned} \quad (2)$$

また、このとき時刻 t_1 と t_2 での透過光強度の差分を考え、これを $\Delta I_{\lambda 1}(t_1, t_2)$ とすれば、

$$\Delta I_{\lambda 1}(t_1, t_2) = I_{\lambda 1}(t_1) - I_{\lambda 1}(t_2) \quad (3)$$

となり、 $\Delta OD_{\lambda 1}(t_1, t_2)$ は以下のように書ける。すなわち、

$$\Delta OD_{\lambda 1}(t_1, t_2) = \log \left\{ \frac{I_{\lambda 1}(t_1)}{I_{\lambda 1}(t_2)} \right\} = \log \left\{ 1 + \frac{I_{\lambda 1}(t_1, t_2)}{I_{\lambda 1}(t_2)} \right\} \quad (4)$$

これらの式より明らかなように、 $\Delta OD_{\lambda 1}(t_1, t_2)$ は入射光強度 $I_{0\lambda 1}$ を計測せずに、透過光強度の時間差分から求めることが出来る。この光学密度の差分 $\Delta OD_{\lambda 1}(t_1, t_2)$ は、図 1 に示したモデルでの動脈による変動に由来するものと考えられる。この波長 λ における $\Delta OD_{\lambda 1}(t_1, t_2)$ は、ラマン散乱や蛍光などの効果を見捨ててよい範囲では t_1 , t_2 および λ の関数として考えられ、 $\Delta OD(\lambda, t_1, t_2)$ と書けることは明らかである。ここで示した、ある 1 波長における光学密度の差分である $\Delta OD(\lambda, t_1, t_2)$ を複数の波長範囲で同時に得ることに、よって、光学密度の差分スペクトラムを得ることができ、この光学密度の差分スペクトラムを以下では $\Delta OD_{\lambda}(t_1, t_2)$ と表記する。

差分スペクトラムの導出においては、複数の t_1 と t_2 の選び方が考えられるが、既報では t_1 を心臓の拡張期に t_2 を時刻 t_1 の拡張期に対応する収縮期にとる方法が示されており、本報告においてもそれに従う。以下では、拡張期の時刻 t_1 を代表時間として $\Delta OD_{\lambda}(t_1, t_2)$ の代わりに $\Delta OD_{\lambda}(t_1)$ と表記し、さらに一般的に $\Delta OD_{\lambda}(t)$ と表現する。

以上のようにして得られた $\Delta OD_{\lambda}(t)$ を動脈変動由来のパラメータと考え、動脈中の血中成分の一つであるグルコース濃度と $\Delta OD_{\lambda}(t)$ との関係が多変量校正モデルによって記述し、血糖値未知の被験者より得られる $\Delta OD_{\lambda}(t)$ から血糖値を予測する方法がパルスグルコメトリの計測原理である(図 1)。

2.2 装置の概要

差分吸光スペクトラム測定のために、新たに開発した瞬時分光測光装置を使用した。本装置の詳細は既報[13]に譲るが、この装置は、図 2 に示すように、光源 (ハロゲンランプ: 最大定格 150 W)、光源用安定化電源、送・受光用光ファイバ、ポリクロメータ (M25-TP; 分光計器(株))、液体窒素冷却式 InGaAs フォトダイオードリニアアレー受光装置 (OMA-V: 512-1.7(LN), Princeton Instruments Co.)、および受光装置制御用 PC から構成されている。本システムにより、波長 900 ~ 1700 nm の近赤外光を約 8 nm の波長分解能、強度分解能 16 bit、最小露光時間 2 μ s、最大スペクトル速度 1800 spectra/s で計測可能である。

2.3 実験方法

今回は、健常人 9 名 (男 7 名, 女 2 名, 22 ~ 59 歳) を対象として、光計測と血糖計測を行った。できるだけ幅広い範囲の血糖値に対応したデータを取得するために、被験者に OGTT (Oral glucose tolerance test, 経口ブドウ糖負荷試験) を課して血糖値を変化させた。なお、被験者には実験前日午後 9 時より、実験当日開始時刻午前 9 時まで絶飲食を依頼した。被験者は安静座位にて OGTT 用グル

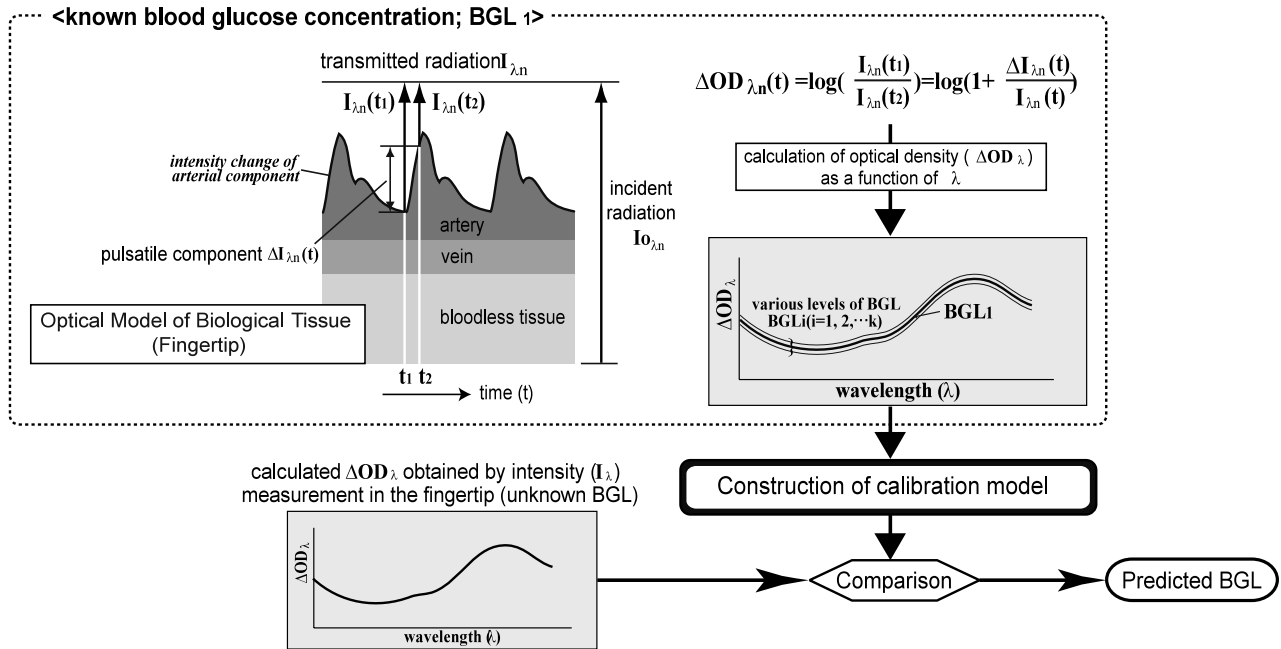


図 1 近赤外瞬時分光による非観血的血糖計測法（パルスグルコメトリ）の原理。図中の左上部に生体組織の光学モデルを示す。説明は本文参照。

Fig. 1 Basic principle of non-invasive blood glucose (BGL) measurement using instantaneous differential NIR spectrophotometry, named pulse glucometry. Optical model of the biological tissue (fingertip) is shown in the upper left part. See text for further explanation.

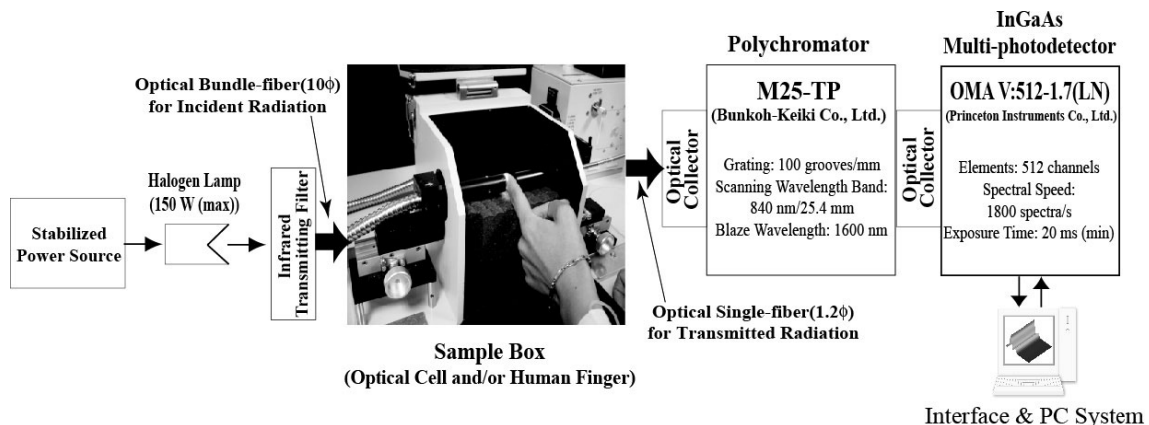


図 2 瞬時近赤外分光測光システムの概要図

Fig. 2 Schematic block diagram of the high-speed NIR spectrophotometric system.

コース溶液（75 g/225 ml，トレーラン-G75，清水製薬（株））を飲下後，実験中安静座位を保った。光計測と血糖計測は，OGTT 用グルコース溶液飲下後，10 分間隔で 120 分経過時点まで行った。血糖計測のために，医師指導下で被験者左前腕部橈側皮静脈より毎回約 3 ml を採血した。静脈血サンプルは生化学分析装置（富士ドライケム 7000 および SD スライド GLU-W III，（株）富士フィルムメディカル）を用いて分析され，血糖値を得た。

光計測は，被験者左第 2 指尖部を対象として行い，1 回の計測につき 20 ～ 30 秒間行い，約 20 ～ 30 心拍に相当する光強度データを計測した。この際，入射光強度を適宜調整し，波長 900 ～ 1700 nm の範囲において，透過光強度変化

（脈浪成分）が良好な S/N（20 dB 程度）を持つよう注意した。

なお，本実験に際して，その目的やプロトコルなどについて金沢大学医学倫理委員会の承認を受けた後に，各被験者の同意を得て実験を行った。

2・4 解析手法

得られた光強度データは（光強度（I），波長（λ），時間（t））の 3 次元より構成される。このデータを I-t 平面で観測すれば，各波長での光電脈波に相当するデータとなるが，さらにノイズ成分を除去するために，以下の前処理を行った。データは波長の次元では，900 ～ 1700 nm の範囲において約 8 nm の分解能で計測された約 100 個の変量を持つが，各

サンプル時間においてケモメトリックス分野でスペクトルの前処理のために一般的に使われている Savitzky-Golay フィルタを用いて I- λ 平面内のフィルタ処理を行い、ノイズ成分を除去した[14]。さらに、このようにして得られたデータの I-t 平面においても同様に Savitzky-Golay フィルタによる処理を行ってノイズ成分を除去した。なお、以上ふたつのフィルタ処理のためのパラメータは経験的に決定し、I- λ 平面内では多項式の次数を 2、フレームサイズを 3 とし、I-t 平面では多項式の次数を 2、フレームサイズを 15 とした。この処理は、各波長での光電脈波にローパスフィルタを適用することに相当するが、最終的に得られた透過光強度変化成分 ΔI の S/N は全波長に対して 25 dB 以上のデータを採用した。

以上のようにして得られた各波長での光強度データについて、心臓の拡張期と収縮期に相当するトップピークとボトムピークを求め、「原理」で示した方法によって光学密度の差分スペクトル $\Delta OD_{\lambda}(t)$ を求めた。既報[13]のように、この光学密度の差分スペクトル $\Delta OD_{\lambda}(t)$ を、1100 nm に相当する値を 1.0、最小値を 0 となるように正規化（スケール）した。

本論文では、以上のようにして得られた正規化光学密度差分スペクトル $\Delta OD_{\lambda}(t)_{scale}$ から、ベースライン変動を除去し、重畳する複数のピークを分離することを期待してケモメトリックス分野で一般的に用いられている手法に従い[14]、波長方向に 2 階差分し、2 階差分正規化 $\Delta OD_{\lambda}(t)$ スペクトルである $\frac{\delta^2(\Delta OD_{\lambda}(t)_{scale})}{\delta \lambda^2}$ を取得した（ただし、 $\frac{\delta}{\delta \lambda}$ は波長に対する差分オペレータを示す）。したがって、 $\frac{\delta^2(\Delta OD_{\lambda}(t)_{scale})}{\delta \lambda^2}$ は波長次元では 100 要素を持つことになる。以上のようにして得られた $\frac{\delta^2(\Delta OD_{\lambda}(t)_{scale})}{\delta \lambda^2}$ と光計測時の実測血糖値データと合わせてデータセットとしてモデル構築した。

血糖値を目的変量とし、2 階差分正規化 $\Delta OD_{\lambda}(t)$ スペクトル $\frac{\delta^2(\Delta OD_{\lambda}(t)_{scale})}{\delta \lambda^2}$ を説明変量とする多変量校正モデル構築のためには、 $\frac{\delta^2(\Delta OD_{\lambda}(t)_{scale})}{\delta \lambda^2}$ が 100 変量から成るデータであるため、何らかの多変量回帰計算が必要となる。多変量回帰においては重回帰分析が最も単純かつ明瞭であるが、本論文で用いるようなデータセット数がデータ変量数より小さいデータセットに適用することはできない。既報では多変量校正モデル構築のために PLS 法を用いているが[13]、本論文においては、主成分回帰 (PCR: Principal Component Regression), PLS 回帰 (Partial Least Squares Regression), 単一中間層の階層型ニューラルネットワーク (ANN: Artificial Neural Network), サポートベクター

マシン回帰 (SVMsR: Support Vector Machines Regression) をそれぞれ用いて比較することとした。なお、ANN の学習のためには可変計量勾配法（準ニュートン法）を採用し、過剰あてはめ (over-fitting) を避けるために weight decay (荷重衰退) を行った。SVMsR のカーネル関数には ANOVA radial basis kernel を採用した[15-17]。以上の計算には、ソフトウェア Matlab (version 7.0.x) と R (version 2.3.1) を用いた。サポートベクターマシン回帰には R 用モジュール kernlab (version 08-2) を[17]、主成分回帰と PLS 回帰のためには R 用モジュール pls (version 1.2-1) をそれぞれ用いた。なお、使用した PC は DELL 社の Dimension 9100 (CPU Pentium D 830, 2GB memory, Windows XP) である。

データセットを学習データと検証データに分割し、学習データによって多変量校正モデルを構築した後、検証データを入力して検証することとした。学習データあるいは検証データの偏りを避けるために、n 分割交差検定法 (n-fold cross validation) を導入した[18]。具体的には、最大血糖値と最少血糖値のデータを除いたデータセットをランダムに 5 分割し、A 群から E 群までの 5 個のデータセット群を作る。この分割された中の 1 データセット群を検証データとし、残りの 4 データセット群に最大血糖値および最少血糖値のデータセットを加えたものを学習データとして学習を行わせ、検証データを A 群としたものから E 群としたもののまでの 5 回の検証を行った。このような交差検定の方法は、5 分割交差検定法 (5-fold cross validation) の変形と考えることができる。なお、学習データに必ず最大血糖値および最少血糖値のデータセットを含めるようにしたのは、外挿を避けるためである。

検証データを多変量校正モデルに入力して得られた予測血糖値の評価のために、血糖計測精度の実用的な評価のために広く用いられている Clarke の Error Grid Analysis (EGA と略す) を用いた[19]。

3. 結 果

図 3 は、以上のような手順で求められた $\Delta OD_{\lambda}(t)$, $\Delta OD_{\lambda}(t)_{scale}$, $\frac{\delta^2(\Delta OD_{\lambda}(t)_{scale})}{\delta \lambda^2}$ の一例を示す。図 3 に示すように、 $\Delta OD_{\lambda}(t)$ の形状は、本実験での計測領域においては血漿（水分）の吸光特性が支配的である 1450 nm 付近にピークを持つ全血サンプルから得られるスペクトラム[20]の形状とは一致しない。この形状については、本論文では議論しないが、既報[13]での考察を参照されたい。得られた $\frac{\delta^2(\Delta OD_{\lambda}(t)_{scale})}{\delta \lambda^2}$ のうち、明らかにノイズが支配的なものなどをデータセットから排除し、97 個のデータセットを得て解析に供した。今回採用したデータセットの実測血糖値範囲は 89.0 ~ 251.0 [mg/dl]であった。前述の「方法」にお

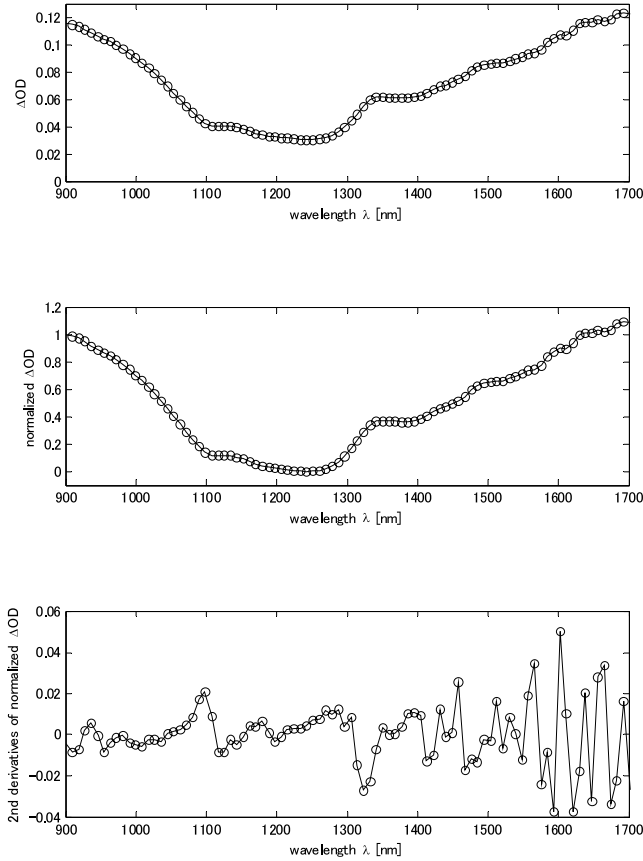


図3 瞬時分光測光システムによって得られた (a) 光学密度の差分 $\Delta OD_{\lambda}(t)$, (b) 正規化された光学密度 $\Delta OD_{\lambda}(t)$ scale および (c) 正規化された光学密度の波長方向 2 階

$$\text{差分} \frac{\delta^2(\Delta OD_{\lambda}(t)_{\text{scale}})}{\delta \lambda^2}$$

Fig. 3 Optical data obtained by the high-speed NIR spectrophotometric system, showing (a) differential optical density ($\Delta OD_{\lambda}(t)$), (b) normalized differential optical density ($\Delta OD_{\lambda}(t)_{\text{scale}}$), and (c) the 2nd derivatives of normalized optical density along wavelengths

$$\frac{\delta^2(\Delta OD_{\lambda}(t)_{\text{scale}})}{\delta \lambda^2}$$

いて示した交差検定を行うために、データセットをランダムに 5 分割し、A 群から E 群までの 5 個のデータセット群を得た。検証用データからは最大血糖値と最少血糖値を含むデータが除かれたため、検証用データセットには血糖値 100.7 ~ 246.3 [mg/dl] の範囲が含まれた。各データセット群には 19 個のデータセットが含まれた。

次に、各多変量解析手法においてパラメータを調整し、多変量校正モデルを得た。各校正モデルにおいて最終的に採用されたパラメータを表 1 に示す。ANN において学習のための計算時間は 2 時間程度であったが、その他の方法においては数分程度以内であった。検証のために必要な時間は、各手法とも 1 分以下であった。

得られた多変量校正モデルに検証データを入力し、血糖計測値と予測値の 2 乗平均誤差を示したのが表 2 である。

表 1 各校正モデルにおいて採用されたパラメータ

Table 1 Introduced parameters in each multivariate calibration models

Method	Parameter	Value
PCR	Applied principal components	1st to 20th
PLS	Number of latent variable	15
ANN	Number of units in the internal layer	50
	Weight decay	0.01
	ANOVA RBF kernel parameter sigma	0.01
SVMsR	ANOVA RBF kernel parameter degree	1
	ϵ in Vapnik's insensitive loss function	0.123

PCR: Principal Component Regression

PLS: Partial Least Squares Regression

ANN: Artificial Neural Network

SVMsR: Support Vector Machines Regression

表 2 血糖予測値と実測値の二乗平均誤差による各回帰モデルの評価

Table 2 Mean squared errors of predicted and measured values of blood glucose obtained by multivariate calibration models.

Method	MSE
PCR	28.64
PLS	31.90
ANN	30.12
SVMsR	28.89

PCR: Principal Component Regression

PLS: Partial Least Squares Regression

ANN: Artificial Neural Network

SVMsR: Support Vector Machines Regression

表 3 Clarke error grid analysis の結果

Table 3 Summarized results from Clarke error grid analysis.

	Region		
	A	B	C-E
Method			
PCR	70	25	0
PLS	74	21	0
ANN	77	18	0
SVMsR	78	17	0

PCR: Principal Component Regression

PLS: Partial Least Squares Regression

ANN: Artificial Neural Network

SVMsR: Support Vector Machines Regression

また、EGA の結果、全モデルの全検証データセット群による検証において、全てのプロットは臨床的良好精度範囲 (clinically accurate) とされる領域 A あるいは臨床的許容範囲 (benign error) とされる領域 B に存在した。表 3 に EGA の結果を示す。また、各校正モデル構築法による交差検定の結果を図 4 に示す。

4. 考 察

表 2 に示すように、平均二乗誤差の評価では、最も単純な方式である PCR が最も良好な結果を与えた。しかし、表

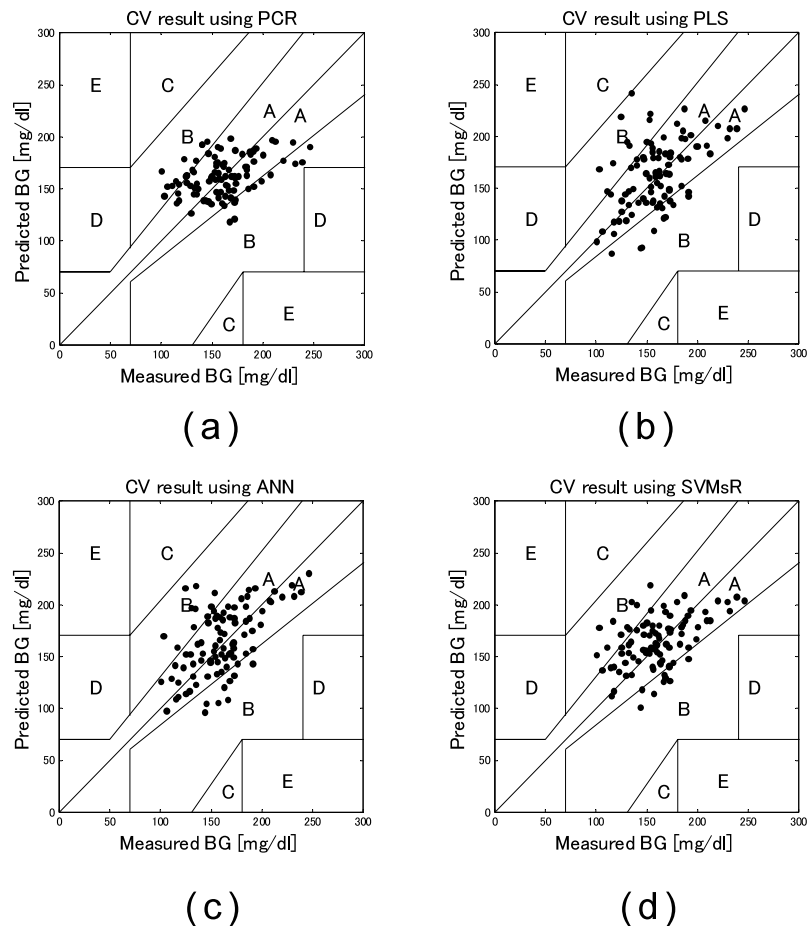


図 4 各手法による交差検定結果の error grid analysis; 図はそれぞれ (a) PCR, (b) PLS, (c) ANN, (d) SVMsR による解析結果.

Fig. 4 Results of cross validations plotted on Clarke error grid analytical diagram. (a) PCR, (b) PLS, (c) ANN and (d) SVMsR.

3に示した EGA の成績では, PCR は最も悪かった. さらに, EGA において, PLS は PCR に比較して良好であったが, ANN と SVMsR に対しては劣っていた. ANN と SVMsR はほぼ同様の成績を示したが, SVMsR の方が比較的良好な成績が得られたと考えられた. 基本的な統計で良好とされる結果と実用的に良好な結果とが両立しないことがあり得ることは, EGA の性質を以下のように解釈すれば明白である. すなわち, EGA は回帰の評価を A から E までのカテゴリに分割して, 離散的かつ非等方的な評価関数を与えるものであると解釈できる. 本研究においても, 平均二乗誤差の評価では SVMsR とほぼ同等の成績を示した PCR が, EGA 上では SVMsR に劣るという結果になった. 本研究では健常人のみを被験者としたため, 検証データ中に血糖値が 250[mg/dl] 以上や 100[mg/dl] 以下のデータはない. しかし, 糖尿病患者を対象とした場合, 血糖値 250[mg/dl] 以上やインシュリン投与後の血糖値 70[mg/dl] 以下の場合では, EGA の領域 D (potentially dangerous) が問題となる. とくに, 血糖値 70[mg/dl] 以下の領域においては, 同一の誤差であっても EGA では予測血糖値が実測値より低い場合は領域 A であるが, 高い場合に

領域 D となる場合があり, このような場合, 平均二乗誤差での評価では同一であっても, 実用における結果が大きく異なることとなる (図 5 を参照). このように, 血糖値の計測法の評価では EGA の非対称性や離散性を考慮することが重要であるが, 今回の結果も, 平均二乗誤差のみでの評価の持つ危険性を示唆したものと考えられた.

表 1 に示された「Vapnik の insensitive loss function における ε (ε in Vapnik's insensitive loss function)」などのパラメータは, すべて汎化性能 (または汎化性能の反義概念としての感度) に関連するものであるが, 各方法でのパラメータの数は 1 ないし 3 である. また, SVMsR のパラメータ数は 3 であるが, ANOVA radial basis カーネルのパラメータ sigma は整数のみをとるため, パラメータ調整に必要な時間は短い. 実際, いずれのモデルにおいてもパラメータの調整は問題とならず, 「高い性能を示したモデルの構築のためにはパラメータ調整が困難である」と言う実用上のトレードオフは発生しなかった. このことより, 各手法の評価は純粋に回帰性能によって行うことができるものと考えられた.

カーネル法の一種であるサポートベクターマシンはパ

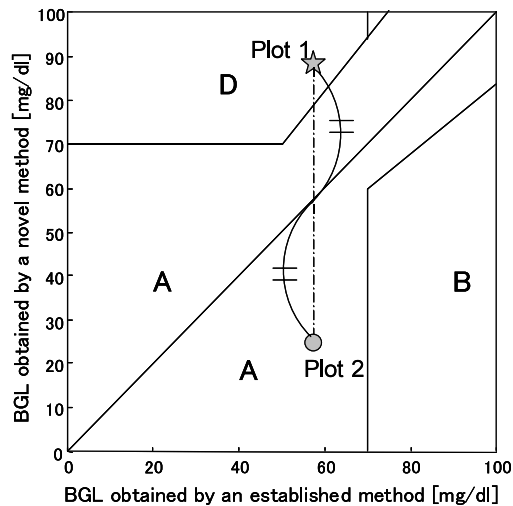


図5 同一の二乗誤差のデータプロットが、error grid analysisでは別の評価を与える例。Plot 1 (★印)はD領域にあるが、Plot 2 (●印)はA領域にあることに注意。

Fig. 5 An example of two data plots with the same squared error. Note that the data can provide different results on Clarke error grid analysis: The plot 1 (★) locates on the region D, while the plot 2 (●) on the region A.

ターンを高次元空間へ写像し、その空間において線形問題を解くものであり、近年、機械学習の分野で期待されている手法である[21]が、今回の結果ではPLSと比較して良好な結果を得た。Thissenら[22]は、ケモメトリクス分野にSVMsRを導入し、PLSとの比較を論じているが、その結果を見てもSVMsRがPLSに対して優位という結果が示されている。いずれの結果でもPLSでは全く不可能であった問題がSVMsRで可能になったという状況ではないが、これは実際の自然界に存在する問題は線形モデルのみで近似できる範囲が少なからずあるためと考えられ、合理的であると判断できる(そもそも、PLSではまったく結果を得られない問題とは、線形性がデータ空間中のいかなる範囲においても全く存在しない問題に対応する、と言うことに注意すべきである)。しかし、サポートベクターマシンの原理は明快であり、特に生体を対象とするような非線形性も仮定すべき分野においては魅力的である。今後は、糖尿病患者を含めてデータセット数を増やし、さらに検討を行う予定であるが、糖尿病患者においては末梢循環の状態に個人差が大きいと考えられ、この差が非線形成分として測定データに重畳する可能性が高いと筆者らは予想している。PLSは基本的に非線形性には対応できないため、非線形性が予測されるデータへの対応のために、SMRsRの持つ融通性の高い解析手法に期待している。

PLSを非線形問題に拡張したものとしてQPLS (Quadratic Partial Least Squares)があるが[23]、QPLSにおいては、どのような非線形関係が存在するかを仮定しなければならないため、瞬時分光測光システムで得られるような高次元のデータセットに対しては、必ずしも実用的ではな

いと筆者らは考えている。一方、SVMsRとANNにおいては、対象にどのような非線形関係があるかを予め仮定する必要はないと言う利点を持つ。今回のSVMsRとANNを比較した場合、SVMsRはANNと比較して顕著に優れた成績を得ることはできなかった。しかし、学習時の計算時間はSVMsRの方が100倍程度高速であり、モデル構築のための時間の観点ではSVMsRは明らかにANNに優越するものと考えられた。先述したように、SVMsRにおいても調整のためのパラメータはさほど多くはなく、学習時間がANNより顕著に短いため、パルスグルコメトリへの応用に限らず、幅広い分野において実用的な分析方法として期待できると考えられる。SVMsR用の計算パッケージも多く公開されており、今後は「高速なANN類似の方法」としてSVMsRを用いることも可能となると思われる。

今回の目的変数である血糖値は、GOD-POD法によって計測されたものであるが、使用した血液分析装置の性能によれば、その正確度はグルコース濃度100 mg/dl以上の全血サンプルにおいて $\pm 15\%$ 以内とされており、実測値にも誤差を含むと考えられる。そのため、仮にパルスグルコメトリの計測精度がその誤差よりも小さかった場合、校正モデルの誤差には血糖値計測誤差の影響が支配的となり得るため、正しい校正モデルの構築には限界が生じることになる。このような議論は極論であるが、目的変数に含まれる誤差に関する議論が今後必要になるものと思われる。例えば、誤差を含む目的変数に対する単回帰の方法としてPassing-Bablok法(Passing-Bablok regression)が知られている[24]。この方法は主として臨床生化学検査において2つの測定法で得られた測定結果の対応を解析することを目的としている。多変量回帰においても、このような手法に相当する方法を用いることができれば理想的であろう。筆者らも今後はこの手法も視野に入れながら、パルスグルコメトリに最適な解析手法を確定していく予定である。

臨床における現在の血糖値計測の標準は化学的手法によるものであるが、生体物質の化学的な計測は多くの場合、生体物質中の還元性物質などの影響を免れない。また、還元性物質の影響を受けない化学的グルコース計測法にはo-toluidine-borate法[25]があるが、グルコースに対する特異性が低いとも言われている。例えば、ヒトの生存に不可欠なビタミンC(アスコルビン酸)は還元性物質であるため、血糖値計測などにおいて誤差要因となることが考えられるが、一方で糖尿病患者においてビタミンCの摂取を勧める知見があることも知られている[26]。このような患者に対しても正しく血糖値管理を行うためには、化学反応に由来する原理に依らない方法で血糖値計測を行う体系化も有用と考えられる。本研究は非侵襲血糖計測に関するものであるが、全血サンプル中の血糖値の光学的計測も多く進んでおり[27]、将来的には化学計測に一切依存しない正確な*in vitro*および*in vivo*非侵襲の血糖計測法の確立も可

能と考えられる。in vivo での非侵襲血液成分計測の重要性は十分に認識されていると思われるが、in vitro での、全血サンプルの光による成分計測の重要性も、以上のことを考えると非侵襲血液成分計測に勝るとも劣らないと認識すべきかも知れない。

パルスグルコメトリを実用に供するためには、今後、装置の小型化、糖尿病患者も含めた大規模なデータ収集や多変量校正モデルの改良などの課題を克服しなければならない。本論文は課題の中でも校正モデルの構築に関わるものであるが、今後は健常者および糖尿病患者も対象とした幅広い計測例を蓄積することが必要である。また、現行の液体窒素冷却 InGaAs フォトダイオードアレーを用いる計測システムはパルスグルコメトリの理想的な計測性能を実証するためのものであり、本法を用いた実用化システムの実現のために、より安価かつ簡便な計測装置の実現に関する検討を今後行う予定である。

5. 結 論

新たな非観血的血糖計測法であるパルスグルコメトリによる血糖値予測のための多変量校正モデル構築に関して、主成分回帰 (PCR)、PLS 回帰、単一中間層の階層型ニューラルネットワーク (ANN)、サポートベクターマシン回帰 (SVMsR) のそれぞれの導入を試みた。その結果、それぞれについて良好な血糖値予測結果を与えるパラメータを得た。以上より、これまでに用いられていた PLS 回帰と同程度の性能を ANN および SVMsR で実現できることが示唆され、特に SVMsR の実用的有用性が示された。

謝辞 本研究の実施にあたってご協力を頂いた金沢大学自然科学研究科・田中志信教授、野川雅道助教、大学院生・佐藤真君に深謝したい。また、本研究の一部は SMC 株式会社の研究援助によるものであり、ここに謝意を表したい。

文 献

- Lawton J, Peel E, Douglas M, Parry O: 'Urine testing is a waste of time': newly diagnosed Type 2 diabetes patients' perceptions of self-monitoring. *Diabetic Med.* **21** (9): 1045-1048, 2004.
- Tamada J, Bohannon N, Potts R: Measurement of glucose in diabetic subjects on non/minimally invasive transdermal extraction. *Nat Med.* **1**: 1198-1201, 1995.
- Rao G, Glikfield P, Guy R: Reverse iontophoresis: Noninvasive glucose monitoring *in vivo* in humans. *Pharm Res.* **12**: 1869-1873, 1995.
- Tierney MJ, Tamada JA, Potts RO, Eastman RC, Pitzer K, Ackerman NR, Fermi SJ: The GlucoWatch biographer: a frequent automatic and noninvasive glucose monitor. *Ann Med.* **32** (9): 632-641, 2000.
- MacKenzie HA, Ashton HS, Spiers S, Shen Y, Freeborn SS, Hannigan J, Lindberg J, Rae P: Advances in Photoacoustic Noninvasive Glucose Testing. *Clin Chem.* **45**: 1587-1595, 1999.
- Berger AJ, Koo TW, Itzkan I, Horowitz G, Feld MS: Multi-component Blood Analysis by Near-Infrared Raman Spectroscopy. *Appl Opt.* **38**: 2916-2926, 1999.
- Berger AJ, Itzkan I, Feld MS: Feasibility of measuring blood glucose concentration by near-infrared Raman spectroscopy. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **53A** (2): 287-92, 1997.
- Robinson MR, Eaton RP, Haaland DM, Koepp GW, Thomas EV, Stallard BR, Robinson PL: Noninvasive glucose monitoring in diabetic patients: a preliminary evaluation. *Clin Chem.* **38**: 1618-1622, 1992.
- Sämann A, Fischbacher C, Jagemann KU, Danzer K, Schüller J, Papenkordt L, Müller UA: Non-invasive blood glucose monitoring by means of near infrared spectroscopy: investigation of long-term accuracy and stability. *Exp Clin Endocrinol Diabetes.* **108**: 406-413, 2000.
- Maruo K, Tsurugi M, Tamura M, Ozaki Y: In vivo noninvasive measurement of blood glucose by near-infrared diffuse-reflectance spectroscopy. *Appl Spectrosc.* **57**: 1236-1244, 2003.
- Khalil OS: Non-invasive glucose measurement technologies: An update from 1990 to the dawn of the new millennium. *Diabetes Technol Ther.* **6**: 660-697, 2004.
- Yamakoshi K: International Patent No. PCT/JP03/03587, 2003.
- Yamakoshi K, Yamakoshi Y: Pulse glucometry: a new approach for noninvasive blood glucose measurement using instantaneous differential near-infrared spectrophotometry. *J Biomed Opt.* **11** (5): 054028, 2006.
- 尾崎幸洋, 宇田明史, 赤井俊雄: 化学者のための多変量解析. 講談社, 東京, 2002, pp. 81-86
- Martens H, Næs T: Multivariate calibration. John Wiley & Sons, New York, 1989, pp.97-165.
- 相吉英太郎: ニューラルネットワークとその学習アルゴリズムの基礎. 電気学会 GA・ニューロを用いた学習法とその応用調査専門委員会編, 学習とそのアルゴリズム. 森北出版, 東京, 2002.
- Karatzoglou A, Smola A, Hornik K, Zeileis A: kernlab-An S4 Package for Kernel Methods in R, *J Stat Software.* **11** (9): 1-20, 2004.
- Ron K: A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence.* **2** (12): American Association for Artificial Intelligence, Montreal, 1995, pp. 1137-1143.
- Clarke WL, Cox D, Gonder-Frederick LA, Carter W: Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care.* **10** (5): 622-628, 1987.
- Nogawa M, Tanaka S, Yamakoshi K: Development of an optical arterial hematocrit measurement method: pulse hematometry. *Procs of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference (CD-ROM), IEEE EMBS, Shanghai, 2005.*
- 津田宏治: サポートベクターマシンとは何か. 電子情報通信学会誌. **83** (6): 460-466, 2000.
- Thissen U, Pepers M, Ustun B, Melssen WJ, Buydens LMC: Comparing support vector machines to PLS for spectral regression applications. *Chemometrics and Intelligent Laboratory Systems.* **73**: 169-79, 2004.
- Blanco M, Coello J, Iturriaga H, MasPOCH S, Pages J: NIR calibration in non-linear systems: different PLS approaches and artificial neural networks. *Chemometrics*

- and Intelligent Laboratory Systems. **50**(1): 75-82, 2000.
24. Passing H, Bablok W: A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I. *J Clin Chem Clin Biochem.* **21**: 709-720, 1983.
 25. Winckers PL, Jacobs P: A simple automated determination of glucose in body fluids using an aqueous o-toluidine-acetic acid reagent, *Clin Chim Acta.* **34**(3): 401-408, 1971.
 26. Will JC, Byers T: Does diabetes mellitus increase the requirement for vitamin C. *Nutr Rev.* **54**(7): 193-202, 1996.
 27. Turza S, Kurihara M, Kawano S: Near infrared analysis of whole blood and plasma in blood-collecting tubes. *J Near Infrared Spectroscopy.* **14**(3): 147-153, 2006.

山越 康弘 (ヤマコシ ヤスヒロ)

2000年北海道東海大学工学部卒業。2004年TYT 技研株式会社に研究員として入社, 2007年株式会社ユーシス代表取締役。同年, 千葉大学大学院工学研究科博士後期課程入学, 現在に至る。主に, 無侵襲生体計測, 健康・福祉工学などの研究に従事。

日本生体医工学会, IEEE などに所属。



小川 充洋 (オガワ ミツヒロ)

1994年早稲田大学大学院修士課程修了, 1998年東京医科歯科大学大学院医学系研究科博士課程修了。博士(医学)。1998年(財)長寿科学振興財団リサーチ・レジデント, 1999年国立療養所中部病院長寿医療研究センター流動研究員, 同年, 東京医科歯科大学生体材料工学研究所助手, 2005年Microvitae Technologies社(仏)主席技師, 2006年TYT(株)研究員, 2007年(株)ユーシス取締役, 現在に至る。無侵襲生体計測, バイオセンサなどの研究開発に従事。

日本生体医工学会, IEEE などに所属。



山越 健弘 (ヤマコシ タケヒロ)

1998年室蘭工業大学建設システム工学科卒, 2000年同大学院修士課程修了。2006年福井大学大学院博士課程修了。工学博士。2006年金沢大学大学院自然科学研究科 システム創成科学専攻 信頼性システム工学講座(寄附講座)特任准教授に就任, 現在に至る。主に, 生体計測と解析, 生体情報の無侵襲・無拘束計測法, 健康・福祉工学などの研究開発に従事。

日本生体医工学会, 日本自動車技術会, IEEE Engineering in Medicine & Biology Society に所属。



田村 俊世 (タムラ トシヨ)

1980年東京医科歯科大学大学院研究科修了(医学博士), 2004年4月千葉大学工学部メディカルシステム工学科教授。2007年4月千葉大学大学院工学研究科人工システム科学専攻メディカルシステムコースに改組, 現在に至る。主な研究テーマは生体計測, 生体信号処理, 福祉工学, 高齢者支援工学など。

日本生医工学会(理事), 日本生活支援工学会(理事), ライフサポート学会(理事) 計測自動制御学会, 米国電気学会, の会員



山越 憲一 (ヤマコシ ケンイチ)

1970年早稲田大学理工学部機械工学科卒, 1972年同大学院修士課程修了。工学博士, 医学博士。1972年東京女子医大助手, 1974年東京医科歯科大助手, 1980年同大講師, 同年北海道大学助教授, 1987年オックスフォード大学客員教授, 1994年金沢大学教授, 2002年中国大連大学客員教授, 2003年早稲田大学客員教授, 現在に至る。生体計測と制御, バイオメカニクス, 健康・福祉工学などの研究開発に従事。

日本生体医工学会, 計測自動制御学会, 日本機械学会, 日本人工臓器学会, 国際医用生体工学会議, IEEE などに所属。

