

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-528083

(P2004-528083A)

(43) 公表日 平成16年9月16日(2004.9.16)

(51) Int. Cl. <sup>7</sup>	F I	テーマコード (参考)
<b>A 6 1 B 5/145</b>	A 6 1 B 5/14 3 1 0	2 G 0 5 9
<b>G 0 1 N 21/35</b>	G 0 1 N 21/35 Z	4 C 0 3 8

審査請求 有 予備審査請求 有 (全 64 頁)

(21) 出願番号	特願2002-578807 (P2002-578807)	(71) 出願人	503016164 センシス メディカル インク アメリカ合衆国 85226 アリゾナ チャンドラー ウェスト チャンドラー ブルヴァード 7470
(86) (22) 出願日	平成14年3月22日 (2002.3.22)	(74) 代理人	100092048 弁理士 沢田 雅男
(85) 翻訳文提出日	平成15年10月1日 (2003.10.1)	(72) 発明者	ブランク トーマス ビー 米国 85225 アリゾナ州 チャンド ラー タルサ ストリート 2922 イ ー
(86) 国際出願番号	PCT/US2002/009114	(72) 発明者	モンフレ ステファン エル 米国 85296 アリゾナ州 ギルバー ト イーストパロ ブランコ ウェイ 1 289
(87) 国際公開番号	W02002/080767		
(87) 国際公開日	平成14年10月17日 (2002.10.17)		
(31) 優先権主張番号	09/825,687		
(32) 優先日	平成13年4月3日 (2001.4.3)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 非侵襲的血液検体予測のための局所化された校正モデルを構築する多段階方法

## (57) 【要約】

【課題】組織の吸収スペクトルを分類する多段階方法であって、スペクトルの干渉物質の共変動、試料の不均一性、状態の変動、および構造的変動に起因する試料のスペクトルにおける変動を低減するために使用されるローカルグループに、校正と試料のスペクトルを局所化する方法を提供すること。

【解決手段】非侵襲的血液検体予測における分類および校正の多段階方法は、共変動するスペクトルの干渉物質を制限することによって予測誤差を最小にする。組織試料は、被検者の人口統計情報と、生体内の近赤外線スペクトル測定などの機器による皮膚測定とに基づいてカテゴリ化される。多段階インテリジェントパターン分類手順は、スペクトルデータを、組織の特性の一貫性の程度が高いクラスターに編成する。各段階において、被検者の人工統計データと、スペクトル測定情報と、組織の分類を確立するのに適したその他のデバイス測定値を使用して、分類の精度が連続的に高められる。校正のためのこの多段階分類方式は、多変量の統計手法と、スペクトルの特徴を使用する多段階分類とを利用する。多段階分類において使用される変数は、皮膚表面の水和状態、皮膚表面の温度、組織の水和状態、近赤外線の脂肪帯による真皮の光学的相対厚さの評価でよい。すべての組織パラメータは、重要な波長部分に沿って近赤外線スペクトル信号を使用して評価される。

## 【特許請求の範囲】

## 【請求項1】

測定された組織スペクトルから対象の血液検体の濃度を推定するための多段階較正モデルを構築する方法であって、当該方法が、  
 較正セットを用意するステップであって、当該較正セットが、被検者の母集団の代表的試料採取からの典型的なスペクトル測定値のデータセットを有する、ステップと、  
 当該スペクトル測定値を、当該被検者に属する先験的な情報に基づいて、事前に定義されるクラスに最初に分類するステップと、  
 当該測定されたスペクトルを、組織の測定場所における少なくとも1つの機器測定値に基づいて、事前に定義されるクラスにさらに分類するステップと、  
 当該測定されたスペクトルから、さらなる分類のための少なくとも1つの特徴を抽出するステップと、  
 を有する方法であって、  
 決定規則によってクラスの割り当てが行われる、  
 方法。

10

## 【請求項2】

当該最初の分類ステップが、  
 第一段階において、当該測定されたスペクトルを、被検者の年齢に基づいて、事前に定義されるクラスに分類ステップと、  
 第二段階において、当該測定されたスペクトルを、被検者の性別に基づいて、事前に定義されるクラスにさらに分類するステップと、  
 を有する、請求項1の方法。

20

## 【請求項3】

当該次の分類ステップが、  
 第三段階において、当該測定されたスペクトルを、組織の測定場所における角質層の水和状態の推定に基づいて、事前に定義されるクラスに分類するステップと、  
 第四段階において、当該測定されたスペクトルを、当該組織の測定場所における皮膚の温度に基づいて、事前に定義されるクラスにさらに分類するステップと、  
 を有する、請求項1の方法。

## 【請求項4】

当該角質層の水和状態の推定が、当該組織の測定場所における周囲湿度の測定値に基づく、請求項3の方法。

30

## 【請求項5】

当該特徴抽出ステップが、組織の測定場所の構造的な特性と生理的な状態を簡潔に表すために、解釈のための試料の測定の品質または特定の側面を向上させる数学的変形を有し、その結果の一連の特徴が、被検者を分類して血液検体予測に最も有用な較正モデルを決定するために使用される、請求項1の方法。

## 【請求項6】

当該特徴が、前処理される測定値から、

40

$$z \in \mathcal{R}^M$$

(

$$z = f(\lambda, x)$$

は、測定値空間から特徴空間への写像)  
 により求められるベクトル

$$f: \mathcal{R}^N \rightarrow \mathcal{R}^M$$

により表され、

50

**$f(\bullet)$** 

を分解することにより、特定の特徴を求めるための特定の変形

$$f_i(\bullet): \mathcal{R}^N \rightarrow \mathcal{R}^M_i$$

が得られ、次元

**$M_i$**

10

が、 $i$ 番目の特徴がスカラーまたはベクトルでありかつすべての特徴の集積が前記ベクトル $\underline{z}$ であるか否かを示し、かつ、特徴がベクトルまたはパターンとして表されるときに、その特徴が、内在する身体的現象の指標となる特定の構造を示す、請求項5の方法。

【請求項7】

個々の特徴が、カテゴリに分割され、当該カテゴリが、身体系に関係する特定の解釈を必ずしも持たない抽象的特徴と、試料の先験的な理解から導かれ、かつ身体的現象に直接的に関係付けることができる単純特徴と、

を有する、請求項6の方法。

20

【請求項8】

近赤外線スペクトル吸収の測定値から計算できる特徴が、

脂肪組織の厚さ

ヘマトクリット・レベル

組織の水和状態

タンパク質吸収の大きさ

当該組織の散乱特性

皮膚の厚さ

温度に関係する効果

年齢に関係する効果

スペクトルの特質

経路長の推定値

組織内の血液の体積分率

環境の影響に関するスペクトルの特質

のいずれかを含む、請求項7の方法。

30

【請求項9】

既知のスペクトル吸収パターンに関係する特徴を求めるためにスペクトルの分解を使用するステップ、

をさらに有する、請求項1の方法。

【請求項10】

人口統計的変数に関係する、測定された吸収スペクトルにおける変動を表すことのできるモデルを構築するために、因子ベースの方法を使用するステップ、

をさらに有する、請求項1の方法において、

当該モデルへの測定された吸収スペクトルの投影が、当該人口統計的変数に関係するスペクトル変動を表す特徴を構成する、方法。

40

【請求項11】

当該特徴抽出ステップが、事前定義された多数のクラスの1つに測定値を割り当てる、請求項1の方法。

【請求項12】

事前定義されたクラスとの類似性について特徴を評価するステップと、

50

クラスメンバーシップを割り当てるステップと、  
をさらに有する、請求項1の方法。

【請求項13】

測定値とクラス割り当てを使用して、特徴からクラス割り当てへの写像を求めるステップ  
、  
をさらに有する、請求項1の方法。

【請求項14】

クラスを当該特徴から監視式に定義するステップであって、特徴の各セットが、2つ以上の  
領域に分割され、かつ、クラスが、特徴の分割の組合せによって定義される、ステップ  
と、

10

組み合わせることのできる当該定義されたクラスのグループを求めるために、前記スペク  
トルデータに対してクラスター分析を実行するステップであって、クラス定義の最終的な  
数が、大幅に低減される、ステップと、

誤分類の数を最小にする、前記特徴空間からクラス推定への最適な写像または変形を求め  
ることによって、クラス定義の後に監視式パターン認識により分類器を設計するステップ  
と、

測定された特徴のセットを推定された分類に変換するモデルを、クラス定義に基づいて作  
成するステップであって、当該クラス定義が、当該測定値を得るために使用される測定シ  
ステムの仕様を満たすように最適化される、ステップと、

をさらに有する、請求項13の方法。

20

【請求項15】

当該最適化されたクラスが、グループ内の測定値の間の類似性がグループ間の類似性より  
大きい測定値のグループを有する、請求項14の方法。

【請求項16】

当該スペクトル測定値の重みを、

$$W = (X^T X)^{-1} X^T y$$

( $X$ はスペクトル測定値の行列を表し、かつ $y$ は各測定用の当該対象の検体濃度の基準値を  
表す)

30

に従って計算するステップ、

をさらに有する、請求項15の方法。

【請求項17】

当該グループの1つの中のスペクトル測定値の重みのベクトルが、当該グループの回帰ベ  
クトルを有する請求項16の方法において、

当該回帰ベクトルが、当該グループの較正モデルを有する、方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、近赤外線組織吸収スペクトルを使用する非侵襲的血液検体予測に関する。よ  
り具体的には、本発明は、スペクトルの干渉物質(interferent)に起因する予測誤差を最  
小にするために、グループ内における一貫性の程度が高い試料のスペクトルのグループに  
対する局在化された較正モデルを構築する方法に関する。

40

【背景技術】

【0002】

非侵襲的血液検体測定の目的は、皮膚を貫通せずに対象の血液検体の濃度を求めること  
である。近赤外(NIR)分光法は、被検者に伝達される低エネルギーの近赤外線の吸収の測定  
値に基づく、有望な非侵襲的技術である。光は、皮膚の小さい領域にフォーカスされ、皮  
下組織を介して伝播する。反射または伝達により皮膚から逸脱して分光計によって検出さ  
れる光は、近赤外線が貫通して試料採取される組織の内容物に関する情報を提供する。各

50

波長における光の吸収は、組織の構造的な特性と化学組成によって決まる。組織層は、それぞれ種類の異なる固有の化学的性質と粒子分布を持ち、光の吸収と入射放射線の散乱を発生させる。水、タンパク質、脂肪、血液検体などの化学的成分は、固有の吸収特性に従ってそれぞれの濃度に比例して光を吸収する。試料組織のスペクトルには、対象の検体のみならず、検体の測定と干渉する多数の他の物質に関する情報も含まれる。この結果、検体信号を分析するためには、多数の干渉物質が重なったスペクトル標識から検体のスペクトル信号を抽出するための数学的モデルを構築する必要がある。膨大な干渉物の正確な補正を行うモデルを定義するには、相当多数の組織試料について100以上の周波数におけるスペクトル測定値が必要になる。

#### 【0003】

血液検体の正確な非侵襲的推定は、試料、すなわち患者の皮膚と生きている組織の動的な性質によっても制限される。化学的、構造的、および生理的な変動が起きると、測定される組織試料の光学的特性に劇的な変化が生じる。R. Anderson、J. Parrishの「人の皮膚の光学 (The optics of human skin)」 (Journal of Investigative Dermatology, Vol. 77(1)、p.13~19、1981年)、W. Cheong、S. Prahl、A. Welchの「生物学的組織の光学的特性の検討 (A review of the optical properties of biological tissues)」 (IEEE Journal of Quantum Electronics, Vol.26(12)、p.2166~2185、1990年12月)、D. Benaron、D. Hoの「時間分解分光測定法を使用する組織内の撮像 (NIRI) および定量化 (NIRS) : 統計的および動的に可変な光路長の影響 (Imaging (NIRI) and quantitation (NIRS) in tissue using time-resolved spectrophotometry: the impact of statically and dynamically variable optical path lengths)」 (SPIE, vol.1888、p10~21、1993年)、J. Conway、K. Norris、C. Bodwellの「身体組成の推定の新しい手法: 赤外相互作用 (A new approach for the estimation of body composition: infrared interactance)」 (The American Journal of Clinical Nutrition, vol.40、p.1123~1140、1984年12月)、S. Homma、T. Fukunaga、A. Kagayaの「人の筋肉の測定における近赤外分光信号における脂肪組織の厚さの影響 (Influence of adipose tissue thickness in near infrared spectroscopic signals in the measurement of human muscle)」 (Journal of Biomedical Optics, Vol.1(4)、p.418~424、1996年10月)、A. Profioの「組織内の光の伝送 (Light transport in tissue)」 (Applied Optics, vol.28(12)、p.2216~2222、1989年6月)、M. Van Gemert、S. Jacques、H. Sterenborg、W. Staの「皮膚の光学 (Skin optics)」 (IEEE Transactions on Biomedical Engineering, vol.36(12)、p.1146~1154、1989年12月)、B. Wilson、S. Jacquesの「組織の光学的反射および透過: 原理と応用 (Optical reflectance and transmittance of tissues: principles and applications)」 (IEEE Journal of Quantum Electronics, vol.26(12)、p.2186~2199) を参照。

#### 【0004】

スペクトル変動の全体的な原因として、次の一般的なカテゴリがあげられる。

#### 【0005】

1. スペクトル的に干渉する種類 (species) の共変動。血液検体の近赤外スペクトル吸収プロファイル (absorption profiles) は、重なって、短い時間間隔で同時に変化する傾向にある。この重なりに起因してスペクトル干渉が生じ、干渉する種類の数よりも多くの独立的に変化する波長における吸収の測定が必要となる。

#### 【0006】

2. 試料の不均一性。組織の測定場所には、さまざまな組成と散乱特性を有する複数の層と区画 (compartment) がある。波長に対するスペクトル吸収の測定は、これらの組織成分の光学的特性と組成の複雑な組合せに関係する。従って、血液検体の濃度が変化しているときのスペクトル応答は、単純な線形モデルから逸脱する可能性が高い。

#### 【0007】

3. 状態の変動。被検者の生理的な状態の変動は、相対的に短い時間間隔で組織層と区画の光学的特性に影響する。このような変動は、例えば、水和レベル、組織中の血液の体積分率の変化、ホルモンの刺激、皮膚温度の変動、血液のヘモグロビンレベルに関係するこ

10

20

30

40

50

とがある。光学プローブとの接触に反応するわずかな変動が予測されることさえある。

【0008】

4. 構造的な変動。個人の組織の特質は、遺伝、環境的な影響、老化 (aging process)、性別、体組成などの要因の結果として差がある。これらの差異は、大部分が解剖学的なものであり、多様な組織形状を形成するゆっくりと変化する構造的な特性として記述することができる。この結果として、被検者の組織が、真皮厚さ、タンパク質レベル、体脂肪率などの特定の特性に直接関係付けることのできる、スペクトル吸収の明確かつ組織体系的な特徴またはパターンを持つことがある。吸収の特徴は一人の患者においては再現されるが、患者の母集団における構造的な変動は、単一の数学的較正モデルの使用によって正しく表すことができないことがある。従って、患者間の差異は、近赤外線スペクトル吸収による血液検体の非侵襲的測定にとって重大な障害である。

【0009】

非分散的 (non-dispersive) システムにおいては、上記(1)に類似する変動は、多重線形回帰と因子ベースのアルゴリズムなど多変量手法を使用して容易にモデル化される。拡散反射における組織の散乱特性をモデル化するために過去に多大な努力がなされているが、上記(2)に概説されている問題はほとんど探求されていない。上記(3)と(4)にリストされているタイプの変動は、有効な解決策が報告されていない重大な非線形スペクトル応答を引き起こす。例えば、非侵襲的なブドウ糖測定のいくつかの報告されている方法は、短い時間間隔に対して個人に固有な較正モデルを構築する。K. Hazenの「近赤外分光法を使用した生物学的基質におけるブドウ糖の定量 (Glucose determination in biological matrices using near-infrared spectroscopy)」(アイオワ大学博士論文、1995年8月)と、J. Burmeisterの「人の非侵襲的血液ブドウ糖測定値のための生体外モデル (In vitro model for human noninvasive blood glucose measurements)」(アイオワ大学博士論文、1997年12月)と、M. Robinson、R. Eaton、D. Haaland、G. Koepp、E. Thomas、B. Stallard、P. Robinsonの「糖尿病患者におけるブドウ糖の非侵襲的モニタリング：予備評価 (Noninvasive glucose monitoring in diabetic patients: a preliminary evaluation)」(Clin. Chem. Vol.38(9)、p.1618~1622、1992年)を参照。この手法により、患者間の差異がモデル化されてしまうことが回避され、従って、多数の個人に一般化されることがない。しかしながら、この較正モデルは、タイプ(4)の変動に関連して再較正が必要となりうる長い時間間隔でのテストが行われていない。さらに、報告されている方法は、タイプ(3)の一連の変動に対して有効であることは、開示されていない。

【0010】

【特許文献1】

米国特許出願番号09/359,191 (S. Malin、T. Ruchtiの「非侵襲的血液検体予測用のインテリジェントシステム (An Intelligent System For Noninvasive Blood Analyte Prediction)」(1999年7月22日))

【非特許文献1】

R. Anderson、J. Parrishの「人の皮膚の光学 (The optics of human skin)」(Journal of Investigative Dermatology、Vol.77(1)、p.13~19、1981年)

【非特許文献2】

W. Cheong、S. Prael、A. Welchの「生物学的組織の光学的特性の検討 (A review of the optical properties of biological tissues)」(IEEE Journal of Quantum Electronics、Vol.26(12)、p.2166~2185、1990年12月)

【非特許文献3】

D. Benaron、D. Hoの「時間分解分光測定法を使用する組織内の撮像 (NIRI) および定量化 (NIRS) : 統計的および動的に可変な光路長の影響 (Imaging (NIRI) and quantitation (NIRS) in tissue using time-resolved spectrophotometry: the impact of statically and dynamically variable optical path lengths)」(SPIE、vol.1888、p10~21、1993年)

【非特許文献4】

10

20

30

40

50

- J. Conway、K. Norris、C. Bodwellの「身体組成の推定の新しい手法：赤外相互作用 (A new approach for the estimation of body composition: infrared interactance)」 (The American Journal of Clinical Nutrition、vol.40、p.1123~1140、1984年12月)  
【非特許文献5】
- S. Homma、T. Fukunaga、A. Kagayaの「人の筋肉の測定における近赤外分光信号における脂肪組織の厚さの影響 (Influence of adipose tissue thickness in near infrared spectroscopic signals in the measurement of human muscle)」 (Journal of Biomedical Optics、Vol.1(4)、p.418~424、1996年10月)  
【非特許文献6】
- A. Profioの「組織内の光の伝送 (Light transport in tissue)」 (Applied Optics、vol. 1.28(12)、p.2216~2222、1989年6月) 10  
【非特許文献7】
- M. Van Gemert、S. Jacques、H. Sterenborg、W. Staの「皮膚の光学 (Skin optics)」 (IEEE Transactions on Biomedical Engineering、vol.36(12)、p.1146~1154、1989年12月)  
【非特許文献8】
- B. Wilson、S. Jacquesの「組織の光学的反射および透過：原理と応用 (Optical reflectance and transmittance of tissues: principles and applications)」 (IEEE Journal of Quantum Electronics、vol.26(12)、p.2186~2199) 20  
【非特許文献9】
- K. Hazenの「近赤外分光法を使用しての生物学的基質におけるブドウ糖の定量 (Glucose determination in biological matrices using near-infrared spectroscopy)」 (アイオワ大学博士論文、1995年8月)  
【非特許文献10】
- J. Burmeisterの「人の非侵襲的血液ブドウ糖測定値のための生体外モデル (In vitro model for human noninvasive blood glucose measurements)」 (アイオワ大学博士論文、1997年12月)  
【非特許文献11】
- M. Robinson、R. Eaton、D. Haaland、G. Koeppe、E. Thomas、B. Stallard、P. Robinsonの「糖尿病患者におけるブドウ糖の非侵襲的モニタリング：予備評価 (Noninvasive glucose monitoring in diabetic patients: a preliminary evaluation)」 (Clin. Chem、Vol.38(9)、p.1618~1622、1992年) 30  
【非特許文献12】
- K. Fukunagaの「統計的パターン認識の紹介 (Intro to Statistical Pattern Recognition)」 (Academic Press、カリフォルニア州サンディエゴ、1990年) J. Hertz、A. Krogh、R. Palmerの「ニューラル計算の理論の紹介 (Introduction To The Theory Of Neural Computation)」 (Addison-Wesley Publishing Co.、カリフォルニア州レッドウッドシティ、1991年)  
【非特許文献13】
- R. Duda、P. Hartの「パターン分類と情景分析 (Pattern Classification and Scene Analysis)」 (John Wiley and Sons、ニューヨーク、1973年) 40  
【非特許文献14】
- J. Conway、K. Norris、C. Bodwellの「身体組成の推定の新しい手法：赤外相互作用 (A new approach for the estimation of body composition: infrared interactance)」 (The American Journal of Clinical Nutrition、vol.40、p.1123~1140、1984年12月)  
【非特許文献15】
- K. Martinの「近赤外線分光法による皮膚内の水分の直接的な測定 (Direct measurement of moisture in skin by NIR spectroscopy)」 (J. Soc. Cosmet. Chem.、Vol.44、p.249~261、1993年9/10月)  
【非特許文献16】 50

W. Cheong, S. Prael, A. Welchの「生物学的組織の光学的特性の検討 (A review of the optical properties of biological tissues)」(IEEE Journal of Quantum Electronics, vol.26(12)、p.2166~2185、1990年12月)

【非特許文献17】

W. Andrew, R. Behnke, T. Satoの「加齢による人の真皮の細胞集団における変化 (Changes with advancing age in the cell population of human dermis)」(Gerontologia, vol.10、p.1~19、1964/65年)

【非特許文献18】

W. Montagna, K. Carlisleの「老化する人の皮膚における構造的な変化 (Structural changes in aging human skin)」(The Journal of Investigative Dermatology, vol.73、p.47~53、1979年)

10

【非特許文献19】

J. Brocklehurstの「老人病医薬と老人病学 (Textbook of Geriatric Medicine and Gerontology)」(p.593~623、Churchill Livingstone, Edinburgh and London、1973年)

【非特許文献20】

T. Ruchtiの「内部報告とプレゼンテーション (Internal Reports and Presentations)」(Instrumentation Metrics, Inc)

【非特許文献21】

S. Matcher, M. Cope, D. Delpyの「近赤外分光法における組織の発色団の濃度変化を定量化するための水の吸収スペクトルの使用 (Use of water absorption spectrum to quantify tissue chromophore concentration changes in near-infrared spectroscopy)」(Phys. Med. Biol., vol.38、p.177~196、1993年)

20

【非特許文献22】

Y. Paoの「適応性パターン認識とニューラルネットワーク (Adaptive Pattern Recognition and Neural Networks)」(Addison-Wesley Publishing Company, Inc., マサチューセッツ州リーディング、1989年)

【非特許文献23】

J. Bezdek, S. Pal編「パターン認識のためのファジィモデル (Fuzzy Models for Pattern Recognition)」(IEEE Press、ニュージャージー州ピスカタウェイ、1992年)

【非特許文献24】

30

S. Wold, M. Sjostromの「SIMCA:類似性と相似の観点における化学データの分析方法 (SIMCA: A method for analyzing chemical data in terms of similarity and analogy)」(Chemometrics: Theory and Application, B. Kowalski編、ACS Symposium Series, Vol.52、1977年)

【非特許文献25】

S. Haykinの「ニューラルネットワーク:包括的な基本 (Neural Networks: A Comprehensive Foundation)」(Prentice-Hall、ニュージャージー州アッパーサドルリバー、1994年)

【非特許文献26】

C. Chen編「ファジィ理論とニューラルネットワークハンドブック (Fuzzy Logic and Neural Network Handbook)」(IEEE Press、ニュージャージー州ピスカタウェイ、1996年)

40

【非特許文献27】

L. Zadehの「ファジィ集合 (Fuzzy Sets)」(Inform. Control, Vol.8、p.338~353、1965年)

【非特許文献28】

H. Martens, T. Naesの「多変量較正 (Multivariate Calibration)」(John Wiley and Sons、ニューヨーク、1989年)

【非特許文献29】

D. Bommannan, R. Potts, R. Guyの「赤外分光法による生体内の角質層のバリア機能の検査 (Examination of the Stratum Corneum Barrier Function In Vivo by Infrared Spec

50

troscopy)」(J. Invest. Dermatol., vol.95, p.403~408, 1990年)

【非特許文献30】

J. Middleton, B. Allenの「角質層に対する温度および湿度の影響と皮膚のひび割れとの関係 (Influence of temperature and humidity on stratum corneum and its relation to skin chapping)」(J. Soc. Cosmet. Chem., vol.24, p.239~43, 1973年)

【非特許文献31】

N. Draper, H. Smithの「応用回帰分析 (Applied Regression Analysis)」(第二版, John Wiley and Sons, ニューヨーク, 1981年)

【発明の開示】

【課題を解決するための手段】

10

【0011】

本発明は、組織の吸収スペクトルを分類する多段階方法であって、スペクトルの干渉物質の共変動、試料の不均一性、状態の変動、および構造的変動に起因する試料のスペクトルにおける変動を低減するために使用されるローカルグループに、較正と試料のスペクトルを局所化する (localize) 方法を提供する。測定スペクトルは、測定時における患者に最も正確な推定を生成するために設計されている局所化された較正モデルに関連付けられる。分類は、現在の患者の状態と構造に関連して抽出される、組織の吸収スペクトルの特徴により行われる。

【0012】

本発明は、患者の代表的母集団からの組織の吸収スペクトルから、またはグループに分けられた個々の患者の生理的な状態から、局所化された較正モデルを構築する方法も提供する。グループまたはクラスは、クラス内の組織の特質の変動がクラス間の変動よりも小さくなるように、構造的な類似性と状態の類似性とに基づいて定義される。

20

【0013】

【発明を実施するための最良の形態】

【0014】

{多段階分類}

スペクトルと他の電子的情報および人口統計的情報を使用しての組織試料の分類は、各種のアルゴリズムを使用することができる。組織の状態をグループ内における類似性の高いグループに分けるための各種の分類器 (classifier) が存在し、例えば、統計的分布情報を利用するベイジアン分類器、先験的な情報をほとんど想定しないノンパラメトリック・ニューラルネットワーク分類器である。K. Fukunagaの「統計的パターン認識の紹介 (Intro to Statistical Pattern Recognition)」(Academic Press, カリフォルニア州サンディエゴ, 1990年)、J. Hertz, A. Krogh, R. Palmerの「ニューラル計算の理論の紹介 (Introduction To The Theory Of Neural Computation)」(Addison-Wesley Publishing Co., カリフォルニア州レッドウッドシティー, 1991年)を参照。本文書において選択されている多段階分類手法は、より多くのデータが利用可能になるに従って分類データベースを発展、拡張することができる。多段階分類器は、階層的な分類ツリーに似ているが、分類ツリーと違い、明確な関数またはファジー関数によって決定規則を定義することができ、決定規則を定義するために使用される分類アルゴリズムは、ツリー構造全体の中で変わることができる。

30

40

【0015】

次に図1を参照する。この図には、多段階分類方式の例が表されている。第一段階11は、事前定義されている年齢グループ、すなわち18~27歳(15)、28~40歳(14)、40~54歳(13)、55~80歳(12)に従って、試料のスペクトルを割り当てる。示されているように、試料は28~40の年齢グループに割り当てられている。第二段階16は、性別に従って試料をクラス17、18に割り当て、この場合には女性である。第三段階19は、角質層の水和状態、すなわち角質水分量31~60(20)、<30(21)、>61(22)に従ってグループ分けし、この場合には>61である。第四段階23は、皮膚温度、すなわち88~90度(24)、86~88度(25)、84~86度、<84度に従ってグループ分けし、この場合には84~86度である。このようにして、多

50

段階構造における各段階の中でクラスメンバーシップ（帰属）の決定が行われる。最後に、最終段階28において、真皮の光学的相対厚さに従って、事前定義されているグループ29、30、31のうちの1つへの最後のクラス割り当てが行われる。

【0016】

図を簡潔にするため、図1には、選択されているクラスからの分岐のみが完全に示されているが、完全な多段階分類構造には、さらに多数の中間および最終的な分類カテゴリが存在することになる。例えば、図の第四段階23には組織の測定スペクトルの96の分類が存在し、最終段階には288の分類が存在することになる。多段階分類構造の上記の説明は、一例を目的とするのみである。当業者には、実際の分類構造はこれより多いかまたは少ない段階を有することができ、かつ各段階においてこの例において使用されているものと異なる決定規則を利用できることが理解されるであろう。

10

【0017】

{特徴の抽出}

前述されているように、分類構造内の各段階において、試料の先験的な知識に基づいて、または組織の測定場所において行われる機器測定値に基づいて、分類が行われる。図1の例においては、最初の2つの段階は、試料に関する先験的な情報、すなわち被検者の年齢と性別を利用する。それ以降の段階は、組織の測定場所における機器測定値から得られる情報を利用する。さらなる分類は、組織の吸収スペクトル自体から抽出される特徴に基づいて行われる。

【0018】

20

特徴の抽出は、解釈を目的として試料の測定の品質または特定の側面を向上させる数学的変形である。R. Duda、P. Hartの「パターン分類と情景分析 (Pattern Classification and Scene Analysis)」(John Wiley and Sons、ニューヨーク、1973年)を参照。図2は、非侵襲的血液検体予測用のインテリジェント測定システムのブロック線図を示し、このシステムについては、本出願の親出願であるS. Malin、T. Ruchtiの「非侵襲的血液検体予測用のインテリジェントシステム (An Intelligent System For Noninvasive Blood Analyte Prediction)」(米国特許出願番号09/359,191:1999年7月22日)に完全に説明されている。図2における特徴の抽出41の目的は、組織の測定場所の構造的な特性と生理的な状態を簡潔に表すことである。一連の特徴は、患者を分類して血液検体予測に最も有用な較正モデルを決定するために使用される。

30

【0019】

特徴は、ベクトル

$$z \in \mathcal{R}^M$$

に表され、このベクトルは、前処理された測定値から、次式により求められる。

$$z = f(\lambda, x) \quad (1)$$

この式において、

40

$$f: \mathcal{R}^N \rightarrow \mathcal{R}^M$$

は、測定値空間から特徴空間への写像である。

$$f(\cdot)$$

を分解することにより、特定の特徴を求めるための特定の変形

$$f_i(\cdot): \mathcal{R}^N \rightarrow \mathcal{R}^{M_i}$$

が得られる。次元

$M_i$

は、 $i$ 番目の特徴がスカラーまたはベクトルでありかつすべての特徴の集積がベクトル $\underline{z}$ であるか否かを示す。特徴がベクトルまたはパターンとして表されるときには、その特徴は内在する身体的現象の指標となる特定の構造を示す。

10

【0020】

個々の特徴は、次の2つのカテゴリに分けられる。

1. 抽象的
2. 単純

【0021】

抽象的特徴は、身体系に関係する特定の解釈を必ずしも持たない。具体的には、主成分分析のスコアは、有用な特徴であるが、その身体的な解釈は必ずしも既知でない。主成分分析の有用性は、組織の吸収スペクトルの性質に関係する。組織のスペクトル吸収における最も有意の変動は、血液検体に起因するのではなく、測定場所の状態、構造、および組成に関係する。この変動は、主要な主成分によってモデル化される。従って、主要な主成分は、組織の測定場所の構造的な特性と生理的な状態に関係する変動を表す傾向にある。

20

【0022】

単純特徴は、試料の先験的な理解から導かれ、身体的現象に直接的に関係付けることができる。近赤外線スペクトル吸収の測定値から計算できる有用な特徴として、次の特徴があげられる（ただしこれらに制限されない）。

【0023】

1. 脂肪組織の厚さ。J. Conway、K. Norris、C. Bodwellの「身体組成の推定の新しい手法：赤外相互作用（A new approach for the estimation of body composition: infrared interactance）」（The American Journal of Clinical Nutrition、vol.40、p.1123 ~ 1140、1984年12月）と、S. Homma、T. Fukunaga、A. Kagayaの「人の筋肉の測定における近赤外分光信号における脂肪組織の厚さの影響（Influence of adipose tissue thickness in near infrared spectroscopic signals in the measurement of human muscle）」（Journal of Biomedical Optics、vol.1(4)、p.418 ~ 424、1996年10月）を参照。

30

【0024】

2. 組織の水和状態。K. Martinの「近赤外線分光法による皮膚内の水分の直接的な測定（Direct measurement of moisture in skin by NIR spectroscopy）」（J. Soc. Cosmet. Chem.、Vol.44、p.249 ~ 261、1993年9 / 10月）を参照。

【0025】

3. タンパク質吸収の大きさ。J. Conwayらの前出の文献を参照。

【0026】

4. 組織の散乱特性。A. Profioの「組織内の光の伝送（Light transport in tissue）」（Applied Optics、vol.28(12)、p.2216 ~ 2222、1989年6月）、W. Cheong、S. Prael、A. Welchの「生物学的組織の光学的特性の検討（A review of the optical properties of biological tissues）」（IEEE Journal of Quantum Electronics、vol.26(12)、p.2166 ~ 2185、1990年12月）、R. Anderson、J. Parrishの「人の皮膚の光学（The optics of human skin）」（Journal of Investigative Dermatology、vol.77(1)、p.13 ~ 19、1981年）を参照。

40

【0027】

5. 皮膚の厚さ。R. Andersonらの前出の文献と、Van Gemmertらの前出の文献を参照。

【0028】

50

6. 温度に関係する効果。Funkungaの前出の文献を参照。

【0029】

7. 年齢に関係する効果。W. Andrew、R. Behnke、T. Satoの「加齢による人の真皮の細胞集団における変化 (Changes with advancing age in the cell population of human dermis)」(Gerontologia, vol.10、p.1~19、1964/65年)、W. Montagna、K. Carlisleの「老化する人の皮膚における構造的な変化 (Structural changes in aging human skin)」(The Journal of Investigative Dermatology, vol.73、p.47~53、1979年)、J. B rocklehurstの「老人病医薬と老人病学 (Textbook of Geriatric Medicine and Gerontology)」(p.593~623、Churchill Livingstone, Edinburgh and London、1973年)を参照。

10

【0030】

8. 性別に関係するスペクトルの特質。T. Ruchtiの「内部報告とプレゼンテーション (Internal Reports and Presentations)」(Instrumentation Metrics, Inc)を参照。

【0031】

9. 経路長の推定値。R. Andersonらの前出の文献と、S. Matcher、M. Cope、D. Delpyの「近赤外分光法における組織の発色団の濃度変化を定量化するための水の吸収スペクトルの使用 (Use of water absorption spectrum to quantify tissue chromophore concentration changes in near-infrared spectroscopy)」(Phys. Med. Biol., vol.38、p.177~196、1993年)を参照。

【0032】

10. 組織内の血液の体積分率。Wilsonらの前出の文献を参照。

20

【0033】

11. 環境の影響に関係するスペクトルの特質

スペクトルの分解は、既知のスペクトル吸収パターンに関する特徴を求めるために使用される。例えば、タンパク質と脂肪には、組織のスペクトル吸収へのそれぞれの貢献 (contribution)を求めるために使用できる既知の吸収標識を持つ。測定された貢献は、特徴として使用され、単一値による基本的な変数を表す。

【0034】

年齢などの人口統計的情報に関する特徴は、単一の吸収プロファイルによって表すことのできない、多数の異なる効果の組合せである。さらに、人口統計的変数と組織のスペクトル吸収の関係は、確定的ではない。例えば、真皮の厚さとその他の多数の組織の特性は、統計的に年齢に関係するが、遺伝的および環境的な影響の結果としても相当に変化する。従って、人口統計的変数に関する測定された吸収における変動を表すことのできるモデルを構築するために、因子ベースの方法が使用される。測定された吸収スペクトルをそのモデルに投影すると、その投影は、人口統計的変数に関するスペクトル変動を表す特徴を構成する。抽象的特徴と単純特徴の集積は、M次元の特徴空間を構成する。一連の特徴全体にわたる情報は冗長であるため、分類器の信頼性を向上させるために最適な特徴選択および/またはデータ圧縮が適用される。

30

【0035】

{分類}

特徴抽出の目的は、分類に関与する測定値の顕著な特質を定義することである。特徴抽出は、多段階分類ツリー構造の分岐点において実行される。分類ステップの目的は、特定の非侵襲的測定に最適な較正モデルを割り当てることである。このステップにおいては、患者は、較正モデルが構築およびテスト済みである多数の事前定義されたクラスのうちの1つに割り当てられる。適用される較正モデルは、類似する組織の吸収スペクトルに対して構築されたものであるため、血液検体予測は、一般的な較正モデルから得られるものより正確である。

40

【0036】

図3に示されているように、パターン分類は、一般的には次の2つのステップを含む。

【0037】

50

1. 分類モデル53が、抽出された特徴について、事前定義されたクラスとの類似性を評価する、写像ステップ。

【0038】

2. 決定エンジン54がクラスメンバーシップを割り当てる、割り当てステップ。

【0039】

この枠組みの中で、2つの一般的な分類方法が提案されている。1つは、相互に排他的なクラスを使用し、従って、各測定を1つのクラスに割り当てる。第二の方式は、同時に複数のクラスのメンバーであることが可能となるファジー分類システムを利用する。いずれの方法も、以下に説明されるように、事前に定義されるクラスに基づく。

【0040】

{クラス定義}

分類システムを構築するには、母集団の代表的試料採取 (representative sampling) から典型的なスペクトル測定値のデータセットが必要である。クラス定義とは、調査データセット内の測定値をクラスに割り当てることである。クラス定義の後、測定値とクラス割り当てが使用されて、特徴からクラス割り当てへの写像が求められる。

【0041】

クラス定義は、監視式手法または非監視式手法のいずれかによって実行される。Y. Paoの「適応性パターン認識とニューラルネットワーク (Adaptive Pattern Recognition and Neural Networks)」(Addison-Wesley Publishing Company, Inc., マサチューセッツ州リーディング、1989年)を参照。監視式手法の場合には、クラスは、データ内の既知の差異によって定義される。このような先験的情報の使用は、監視式パターン認識における最初のステップであり、これにより、クラス割り当てが既知であるときに分類モデルが構築される。例えば、観察されるスペクトル変動の大部分は、体脂肪、組織の水和状態、皮膚の厚さなどのいくつかの身体的特性に関係する3つの抽象的因子によってモデル化できる。これらの3つの特徴に基づいて患者をカテゴリ化すると、各特徴に値「高」と「低」が割り当てられる場合には、8つの異なるクラスが生成される。この手法の欠点は、スペクトルの類似性に注意が払われていないことと、特徴の数が増えるとクラスの数が増加する傾向にあることである。

【0042】

非監視式方法は、特徴空間におけるデータのクラスターまたは自然なグループ分けを探索して構築するのに、スペクトル測定値のみに依存する。このような分析においては、クラスター内の同質性とクラスター間の分離が最適化される。身体的意味を伴う特徴から成るクラスターは、特徴空間における変動の原因となる既知の基本的な現象に基づいて解釈できる。しかしながら、クラスター分析においては、先験的情報が利用されず、矛盾する結果が得られることがある。

【0043】

これら2つの手法の組合せにおいては、自然に発生するスペクトルクラスを得るために、先見的な知識と、特徴空間の探索とが利用される。この手法においては、クラスは、最初に監視式に特徴から定義される。特徴の各セットが2つ以上の領域に分割され、この分割された特徴群の組合せによってクラスが定義される。データに対してクラスター分析が実行され、2つの手法の結果が比較される。クラスターが体系的に使用され、組み合わせることのできるクラスのグループが求められる。この組み合わせの後の最終的なクラス定義の数は、データ内の自然な分割によって大幅に減少する。

【0044】

クラスの定義の後、監視式パターン認識によって分類器が設計される。測定された特徴のセットを推定された分類に変換するモデルが、クラス定義に基づいて作成される。分類器の最終的な目的は、信頼性が高くかつ正確な較正モデルを生成することであるため、反復的な手法、すなわち測定システムの仕様を満たすようにクラス定義が最適化される手法に従う必要がある。

【0045】

10

20

30

40

50

## { 統計的分類 }

統計的分類方法は、変動を統計的に記述することのできる相互に排他的なクラスに適用される。J. Bezdek、S. Pal編「パターン認識のためのファジィモデル (Fuzzy Models for Pattern Recognition)」(IEEE Press、ニュージャージー州ピスカタウェイ、1992年)を参照。クラス定義が一連の典型的試料に割り当てられると、誤分類の数が最小になる、特徴空間からクラス推定への最適な写像または変形を求めることによって、分類器が設計される。この写像の形式は、「最適な」の定義が変わると方法ごとに変わる。既存の方法としては、線形判別分析、SIMCA、k最近隣法(k nearest-neighbor)、人工ニューラルネットワークの各種形式があげられる。Funkunagaの前出の文献、Hertzらの前出の文献、Martinの前出の文献、Dudaらの前出の文献、Paoの前出の文献、S. Wold、M. Sjostromの「SIMCA:類似性と相似の観点における化学データの分析方法 (SIMCA: A method for analyzing chemical data in terms of similarity and analogy)」(Chemometrics: Theory and Application、B. Kowalski編、ACS Symposium Series、Vol.52、1977年)、S. Haykinの「ニューラルネットワーク:包括的な基本 (Neural Networks: A Comprehensive Foundation)」(Prentice-Hall、ニュージャージー州アッパーサドルリバー、1994年)を参照。結果は、

$$c = f(z) \quad (2)$$

に従って特徴をクラスcに写像する関数またはアルゴリズムである。この式において、cは間隔[1,P]の中の整数、Pはクラスの数である。クラスは、較正の節に説明されているように、較正モデルを選択するかまたは適合させるために使用される。

## 【 0 0 4 6 】

## { ファジィ分類 }

統計ベースのクラス定義により、血液検体の推定に適用可能な一連のクラスが生成されるが、スペクトル変動につながる組織試料の光学的特性は、連続的な値として変化する。従って、特に組織の厚さ、水和レベル、体脂肪率の自然な変動の結果として、クラスの重なりが生じる。明確なクラス境界は存在せず、多数の測定値がクラスの間位置することも多く、いくつかのクラスのいずれかに分類される統計的な確率は等しい。従って、厳密なクラス境界と相互に排他的なメンバーシップ関数は、対象の母集団の性質に適さない。

## 【 0 0 4 7 】

より適応性の高いクラスの割り当て方法は、ファジィ集合理論に基づく。J. Bezdekらの前出の文献と、C. Chen編「ファジィ理論とニューラルネットワークハンドブック (Fuzzy Logic and Neural Network Handbook)」(IEEE Press、ニュージャージー州ピスカタウェイ、1996年)、L. Zadehの「ファジィ集合 (Fuzzy Sets)」(Inform. Control、Vol.8、p.338~353、1965年)を参照。一般的に、ファジィ集合におけるメンバーシップは、グレードの連続体と、各クラスについて特徴空間を間隔[0,1]に写像する一連のメンバーシップ関数とによって定義される。割り当てられたメンバーシップグレードは、そのクラスに帰属する度合いを表し、「1」が最高の度合いに対応する。従って、1つの試料が同時に複数のクラスのメンバとなることがある。

## 【 0 0 4 8 】

特徴空間からクラスメンバーシップのベクトルへの写像は、式

$$c_k = f_k(z) \quad (3)$$

によって与えられ、この式において、 $k=1, 2, \dots, P$ ,

$$f_k(\cdot)$$

はk番目のクラスのメンバーシップ関数であり、すべてのkについて

$$c_k \in [0, 1]$$

10

20

30

40

50

であり、ベクトル

$$c \in \mathcal{R}^P$$

はクラスメンバーシップの集合である。メンバーシップベクトルは、事前定義されたクラスそれぞれのメンバーである度合いを与え、校正アルゴリズムに渡される。

【0049】

メンバーシップ関数の設計においては、前述されている方法に似たファジークラス定義が利用される。ファジークラスター分析を適用することができ、構造と最適化手法の異なるいくつかの方法を使用してファジー分類器を構築できる。いずれの方法も、試料の母集団におけるクラスメンバーシップの推定誤差を最小化する方向に作用する。

10

【0050】

{多段階校正}

血液検体予測は、図2に示されているように、前処理された測定値に校正モデルを適用することによって行われる。提案されている予測システムは、分類ステップに基づいて適合化できる、または選択される1つまたは一連の校正モデルを使用する。

【0051】

局所化された校正モデルの構築

正確な血液検体予測のためには、測定時に遭遇する、共変動する干渉物質、試料の不均一性、状態および構造的な変動を補正することのできる校正モデルが必要である。化学的な吸収物質の複雑な混合物（システムの構成要素の間でスペクトルの相当な重なりを示す）は、多変量の統計的モデルを使用することによってのみ解釈できる。しかしながら、予測誤差は、校正データにおいて検体濃度と共変動する干渉物質の変動が増すに従って増大する。従って、血液検体予測は、校正セットデータにおいて検体の濃度との相関関係の薄い、干渉の変動が小さい測定値に対して実行されるのが最良である。干渉の変動すべてをランダムにできない場合があるため、一般には、スペクトルの干渉物質の変動の範囲を制限することが望ましい。

20

【0052】

多段階の分類および校正システムの背後にある原理は、必要な検体信号の存在下において重なっている干渉信号を補正するために必要である、アルゴリズムの一般化されたクラス特性に基づく。H. Martens、T. Naesの「多変量校正 (Multivariate Calibration)」(John Wiley and Sons、ニューヨーク、1989年)を参照。この用途において使用されるモデルは、( $\underline{y}$ として表される)単一の従属変数を推定するために( $\underline{x}$ として表される)複数の独立変数の測定を必要とする。例えば、 $\underline{y}$ は組織のブドウ糖濃度であり、 $\underline{x}$ は、 $n$ 個の波長のそれぞれにおける非侵襲的スペクトル信号の強度から成るベクトル $[\underline{x}_1 \underline{x}_2 \dots \underline{x}_n]$ を表す。

30

【0053】

単一変数であるブドウ糖の推定の計算において使用されるモデルの一般化された形式では、方程式(4)に示されているように非侵襲的スペクトルの加重総和が使用される。重み $w$ は、回帰ベクトルと称される。

40

$$y = \sum w_i x_i \quad (4)$$

この重みは、校正モデルを定義し、スペクトル行列 $X$ 内の非侵襲的スペクトルの特定の校正セットと、各スペクトルの関連付けられた基準値 $\underline{y}$ とから、次のように計算する必要がある。

$$w = (X^T X)^{-1} X^T y W \quad (5)$$

方程式(5)を使用する多変量システムにおいて予測されるモデリング誤差は、線形加法ミクスチャモデル(linear additive mixture model)を使用して推定できる。線形加法ミクスチャは、ミクスチャ内の個々の構成要素の純粋なスペクトルの合計がミクスチャのスペ

50

クトルと等しいという定義を特徴とする。線形ミクスチャモデルは、線形加法システムに基づく多変量モデルと非侵襲的血液分析の一般的な制限（例えば線形加法の挙動からいくらか逸脱することを予測することができる制限）を評価するうえで有用である。

【 0 0 5 4 】

図4は、典型的な非侵襲的吸収スペクトルを示す。一連のスペクトル測定値は、各行が個々の試料のスペクトルに対応し、かつ各列が単一波長における信号の大きさを表す行列  $X$  として表すことができる。この測定値行列は、機器の基準線の変動の行列  $B_0$ 、純粋成分のスペクトルの行列  $K$ 、純粋成分の濃度  $Y$ 、各スペクトルの測定に存在するランダム測定ノイズ  $E$  の、線形加法ミクスチャモデルとして表すことができる。

$$X = B_0 + YK^T + E \quad (6)$$

10

【 0 0 5 5 】

線形加法モデルは、拡張ミクスチャモデルとして干渉物質と検体とにさらに分解することができる。

$$X = B_0 + YK^T + TP^T + E \quad (7)$$

方程式7において、 $T$ は全試料における干渉物質の濃度または大きさを表す行列であり、 $P$ は存在する干渉物質または干渉効果の純粋なスペクトルを表す。この定式化においては、スペクトルの変形すべてを干渉物質とみなすことができる。例えば、光学的試料採取の組織体における可変の試料の拡散および偏差の効果を、干渉源としてこの定式化に含める必要がある。検体  $Y$  に対する一般化最小二乗モデルの直接的な較正  $y$  は、

20

$$y_{OIS} = (K^T \sum^{-1} K)^{-1} K^T \sum^{-1} (x - k_0) \quad (8)$$

であり、この式において、 $\sum$  は干渉物質またはスペクトル効果の共分散行列として定義され、 $\sum^{-1}$  は測定ノイズとして定義され、 $x$  はスペクトル測定値であり、 $k_0$  はスペクトル測定値に存在する機器基準線成分である。

$$\sum = P^T (H^T)^{-1} P + \text{diag}(\sigma^2) \quad (9)$$

このような一般化最小二乗予測式から導かれる平均二乗誤差 (MSE) は、Martensらの前出の文献に記載されている。

$$MSE(y_{OIS}) = \text{trace}(K^T \sum^{-1} K)^{-1} \quad (10)$$

30

方程式10は、干渉物質の存在時における最小二乗予測式の一般化された制限を記述する。 $K$ が血糖の濃度を表すなら、方程式7の基本的な解釈は次のようになる。すなわち、ブドウ糖推定における平均二乗誤差は、較正データにおけるブドウ糖濃度に共変動する干渉物質の変動が増すに従って増大する。従って、ブドウ糖の正確な推定は、較正セットデータにおいてブドウ糖濃度と相関関係の薄い、干渉の変動が小さい測定値に対して実行するのが最良である。干渉の変動すべてをブドウ糖に対してランダムにできない場合があるため、一般には、スペクトルの干渉物質の変動の範囲を制限することが望ましい。

【 0 0 5 6 】

この多段階分類は、試料の測定値をグループ内における一貫性の程度が高いグループに配置することによって、スペクトルの干渉物質の変動を制限する方法を提供する。グループは、試料の先験的知識と、組織の測定場所における機器測定値と、抽出された特徴とに基づいて定義される。連続する段階それぞれにおいて、グループ内のスペクトル間の変動が逐次制限されるように、試料はさらに分類される。クラスの定義において利用される組織パラメータとしては、角質層の水和状態、組織の温度、真皮の厚さがあげられる。

40

【 0 0 5 7 】

組織の水和状態

角質層 (SC)、すなわち角層は、腕の裏側の約 10 ~ 15  $\mu\text{m}$  の厚さを占める。この角質層は、主としてケラチンの死んだ細胞、水、および脂質から成る。D. Bommannan、R. Potts、R. Guyの「赤外分光法による生体内の角質層のバリア機能の検査 (Examination of the S

50

stratum Corneum Barrier Function In Vivo by Infrared Spectroscopy)」(J. Invest. Dermatol., vol.95, p.403~408, 1990年)を参照。角質層の水和状態は、室温と相対湿度の関数として時間とともに変化することが知られている。J. Middleton, B. Allenの「角質層に対する温度および湿度の影響と皮膚のひび割れとの関係(Influence of temperature and humidity on stratum corneum and its relation to skin chapping)」(J. Soc. Cosmet. Chem., vol.24, p.239~43, 1973年)を参照。角質層は分光計の入射ビームが貫通する最初の組織であるため、組織試料の他のどの部分よりもより光子が多い。従って、組織試料のこの最初の層の中の水など、近赤外線を強く吸収する物質の変動は、角質層の下に貫通する光子の波長および深さ-強度プロファイルを変化させるように作用することがある。

10

#### 【0058】

角質層の水和状態の変化の影響は、簡単な実験によって観察できる。実験の第一部においては、角質層の水和状態は、周囲条件に応じて自由に変動できるようにされる。実験の第二部においては、角質層の水和状態の変動は、皮膚の表面における相対湿度を測定前に高レベルに制御することによって制限される。水和状態が制御されない実験と制御される実験を使用しての非侵襲的測定値は、それぞれ、図5と図6にプロットされている。1900 nmにおける水帯(water band) 61の変化は、変化する表面の水和状態を評価するのに使用できる。1900 nmにおける水帯61の変動の範囲は、図5の場合よりも図6の場合の方がかなり狭い。表面の水和状態はスペクトル測定における重要な変数を表すため、これは組織試料における類似性をカテゴリ化するときを使用できる有用な要素である。

20

#### 【0059】

##### 組織の温度

測定される組織の温度は、貫通の最深レベルにおける体の中心部の温度から、皮膚の表面の温度まで変化し、皮膚の表面の温度は、一般的には周囲温度、組織の測定場所の位置、測定場所における衣服の量に関係する。生きている人間の組織の約65%を占める水のスペクトルは、1100~2500 nmの波長範囲において試料採取される深さすべてにおいて最も主要なスペクトル成分である。これらの2つの事実と、水帯が1450 nmにおいて温度に起因して移動するという公知の現象のために、ブドウ糖など多数の血液検体に関する情報の解釈が相当に複雑になる。試料採取された生きた組織内に一連の温度状態が存在し、組織内の状態の範囲と分布が皮膚表面の温度に依存することは明らかである。さらに、皮膚の屈折率は、温度と共に変化することが知られている。従って、皮膚の温度は、較正モデルの生成と予測のためのグループを識別するために、多段階分類において使用するための重要なカテゴリ変数とみなされる。

30

#### 【0060】

##### 真皮の光学的厚さ

血液の成分を較正するためには、組織の光学的試料採取を繰り返す必要がある。血液は人の組織の一部であり、血液検体は組織の一部に属すのみであるため、組織の光学的試料採取における変化によって、血液検体のレベルを変化させないための検体信号の大きさが変化することがある。試料採取におけるこの種類の効果に起因して、検体の特定のレベルの信号強度が変化することにより、較正において混乱が生じることがある。

40

#### 【0061】

光学的試料採取深さのカテゴリ化は、異なる層のスペクトルの標識帯(marker bands)を分析することによって行われる。例えば、皮膚の下の最初の組織層(the first tissue layer)は、主として脂肪から成る皮下脂肪組織である。脂肪吸収帯の強度は、皮下組織レベルまで貫通した相対的な光子束を評価するために使用できる。脂肪帯が明確であるほど、太い光子束が脂肪組織に達して検出器に戻ったことを意味する。図7には、明確な脂肪帯71と通常の脂肪帯72を有するスペクトルが示されている。光学的厚さの最も重要な用途は、光学プローブによって試料採取された内部組織における水和状態の程度を評価することである。光学的厚さは、性別と体型の強い関数でもあり、従ってこの特性の測定値は、一人の個人の中の内部の水和状態を評価するのに有用となる。

50

## 【0062】

以下の節では、2種類の分類器、すなわち相互に排他的な分類器とファジーな分類器の較正システムについて説明する。

## 【0063】

{ 相互に排他的なクラス }

一般的な場合においては、決定された分類は、患者の分類とスペクトル測定に基づいて血液検体予測を生成する非線形モデルに渡される。図8に示されているこのプロセスでは、吸収スペクトルに明らかに現れている組織の構造的な特性と生理的な状態に従って、現在の被検者の推定方式が修正される。

## 【0064】

この一般的なアーキテクチャは、非線形PLS (partial least squares) または人工ニューラルネットワークなどの非線形較正モデル103を必要とし、なぜなら写像がきわめて非線形なためである。前処理された測定値  $\underline{x}$  に対する血液検体予測は、分類が  $c$  によって指定されるとき、次式によって与えられる。

$$\hat{y} = g(c, \underline{x}) \quad (1)$$

この式において、

$$g^{(*)}$$

は、 $\underline{x}$  と  $c$  を血液検体の濃度の推定

■

に写像する非線形較正モデルである。

## 【0065】

望ましい具体化においては、各クラスに対して異なる較正が具体化される。推定されるクラスは、現在の測定値を使用しての血液検体予測に最も適した、 $p$  個の較正モデルのうちの1つを選択するために使用される。 $k$  が測定値に対するクラス推定であるとすると、血液検体予測は、

$$\hat{y} = g_k(\underline{x}) \quad (12)$$

であり、この式において

$$g_k^{(*)}$$

は、 $k$  番目のクラスに関連付けられている較正モデルである。

## 【0066】

較正は、血液検体の基準値と事前に割り当てられる分類定義とを使用して、一連の典型的な吸収スペクトルから構築される。このセット(「較正セット」と称される)は、患者の母集団において遭遇する一連の生理的な状態を完全に表すために十分な試料を有する必要がある。 $p$  個の異なる較正モデルは、 $p$  個のクラスのそれぞれに割り当てられる測定値から個々に構築される。モデルは、主成分回帰、PLS回帰、および人工ニューラルネットワークなどの既知の方法を使用して具体化される。Hertzらの前出の文献と、Paoの前出の文献と、Haykinの前出の文献と、Martensらの前出の文献と、N. Draper、H. Smithの「応用回帰分析 (Applied Regression Analysis)」(第二版、John Wiley and Sons、ニューヨーク、1981年)を参照。各クラスに関連付けられているさまざまなモデルは、独立したテストセットまたは交差検証(cross validation)に基づいて評価され、モデルの「最良の」セットが多段階分類に組み込まれる。これにより、患者の各クラスは、そのクラスに固有な較正モデルを持つ。

## 【0067】

{ ファジークラスメンバーシップ }

ファジー分類が使用されるときには、較正は、単一の推定されたクラスではなく、メンバーシップのベクトルに渡される。ベクトル  $c$  は、血液検体予測またはいくつかの血液検体

10

20

30

40

50

予測の最適な組み合わせに適する較正モデルの適合化を求めるために利用される。図9に示されている一般的な場合には、メンバーシップベクトルと前処理された吸収スペクトルの両方が、血液検体予測のための単一の較正111によって使用される。この計算は、次式によって与えられる。

$$y = g(c, x) \quad (13)$$

この式において

$g(x)$

は、非線形回帰、非線形PLS、または人工ニューラルネットワークによって求められる非線形写像である。この写像は、前述されている較正セットから構築され、一般的に複雑である。

10

【0068】

図10に示されている望ましい具体化は、各クラスに対して個別の較正121を有する。しかしながら、各較正は、各測定値に割り当てられているメンバーシップベクトルを利用することによって、較正セット内のすべての測定値を使用して生成される。さらに、メンバーシップベクトルは、非ファジー化(defuzzification) 122によって、すべてのクラスからp個の血液検体予測の最適な組合せを求めるために使用される。従って、較正の構築時に、較正セットの1つの測定値が、複数の較正モデルに影響を与えることがある。同様に、予測時に、血液検体推定を生成するために複数の較正モデルが使用される。

20

【0069】

p個の較正モデルのそれぞれは、較正データのセット全体を使用して構築される。しかしながら、k番目の較正モデルが計算されるとき、較正の測定値は、k番目のクラスにおけるそれぞれのメンバーシップによって重み付けされる。この結果として、特定のクラスの較正モデルに対する試料の影響は、そのクラスにおけるメンバーシップの関数である。

【0070】

線形の場合には、回帰係数を計算するために重み付き最小二乗(weighted least squares)が適用され、因数ベースの方法の場合には、共分散行列が適用される。Dudaらの前出の文献を参照。吸収スペクトルの行列

$$X_k \in \mathcal{R}^{r \times w}$$

30

、基準血液検体濃度

$$Y \in \mathcal{R}^r$$

(rは測定スペクトルの数、wは波長の数(number wavelengths))であり、各吸収スペクトルのクラスkにおけるメンバーシップが、

$$C_k \in \mathcal{R}^r$$

40

の要素であるとする。この場合、主成分は、次式によって与えられる。

$$F = X_k M \quad (14)$$

この式において、MはPの最初のn個の固有ベクトルの行列である。重み付けされた共分散行列Pは、次式によって求められる。

$$P = X_k V X_k^T \quad (15)$$

この式において、Vは対角線上の要素 $c_k$ を持つ正方行列である。回帰行列Bは、次式によっ

50

て求められる。

$$B = (F^T V F)^{-1} F^T V Y \quad (16)$$

人工ニューラルネットワークなどの反復式方法が適用されるときには、学習するアルゴリズムに試料が提示される頻度を求めるためにメンバーシップが使用される。これに代えて、大きさが $\gamma$ に従って決められる共分散行列を使用して、拡張カルマンフィルタが適用される。

【0071】

非ファジー化の目的は、正確な血液検体予測を生成する測定値のメンバーシップベクトルに基づいて、 $p$ 個の異なる血液検体予測の最適な組み合わせを見つけることである。従って、非ファジー化は、血液検体予測のベクトルとクラスメンバーシップのベクトルから単一の検体予測への写像である。非ファジー化器(defuzzifier)は、次式のような変形として示すことができる。

$$\hat{y} = d(c, \gamma_1, \gamma_2, \gamma_3, \dots, \gamma_p) \quad (17)$$

この式において、

$d(\cdot)$

は非ファジー化関数、 $c$ はクラスメンバーシップベクトル、 $\gamma_k$ は $k$ 番目の較正モデルの血液検体予測である。重心または加重平均などの非ファジー化の既存の方法は、小さな較正セットの場合に適用される。しかしながら、試料の数が十分な場合には、

$d(\cdot)$

は条件付き非線形モデルによって生成される。

【0072】

{ 機器の説明 }

多段階の分類および較正は、拡散反射測定によって被検者の前腕の近赤外線吸収スペクトルを求める走査分光計内に実施される。この機器は、石英ハロゲンランプと、単色光分光器と、InGAs検出器とを使用する。試料から検出された強度は、アナログ電子装置によって電圧に変換され、16ビットA/Dコンバータによってデジタル化される。スペクトルは処理用のインテリジェント測定システム(IMS)に渡され、その結果として、ブドウ糖予測か、または無効な走査を示すメッセージのいずれかが得られる。

【0073】

本発明は、本明細書においては望ましい実施例を参照しながら説明されているが、当業者には、他の実施例を、本発明の精神および範囲から逸脱することなく、本明細書に記載されている実施例の代わりに使用できることが容易に理解されるであろう。従って、本発明は、添付される請求項によって制限されるのみである。

【図面の簡単な説明】

【0074】

【図1】本発明による多段階分類ツリーの構造を表す。

【図2】本発明による、血液検体の非侵襲的測定用インテリジェントシステムのアーキテクチャのブロック線図である。

【図3】本発明による、パターン分類システムのブロック線図である。

【図4】拡散反射近赤外線分光計を使用して集められた非侵襲的吸収スペクトルである。

【図5】組織の水和状態を制御する方策がない場合の、何度か繰り返された非侵襲的測定

10

20

30

40

50

のスペクトルを示す。

【図 6】本発明による、周囲湿度を使用して水和状態を制御する場合の、何度か繰り返された非侵襲的測定のスペクトルを示す。

【図 7】1710 nmに明確な脂肪帯を持つ非侵襲的吸収スペクトルを示す。

【図 8】本発明による、相互に排他的なクラスの場合の一般的な較正システムのブロック線図である。

【図 9】本発明による、ファジーなクラス割り当ての場合の一般的な較正システムのブロック線図である。

【図 10】本発明による、ファジー集合割り当ての場合の平行較正モデルの例を示すブロック線図である。

10

【符号の説明】

【0075】

11 第一段階

12 年齢55～80

13 年齢40～54

14 年齢28～40

15 年齢18～27

16 第二段階

17、18 クラス

19 第三段階

20 角質水分量31～60

21 角質水分量 < 30

22 角質水分量 > 61

23 第四段階

24 皮膚温度88～90

25 皮膚温度86～88

26 皮膚温度84～86

27 皮膚温度 < 84

28 最終段階

29、30、31 グループ

41 特徴の抽出

53 分類モデル

54 決定エンジン

61 水帯

71 明確な脂肪帯

72 通常の脂肪帯

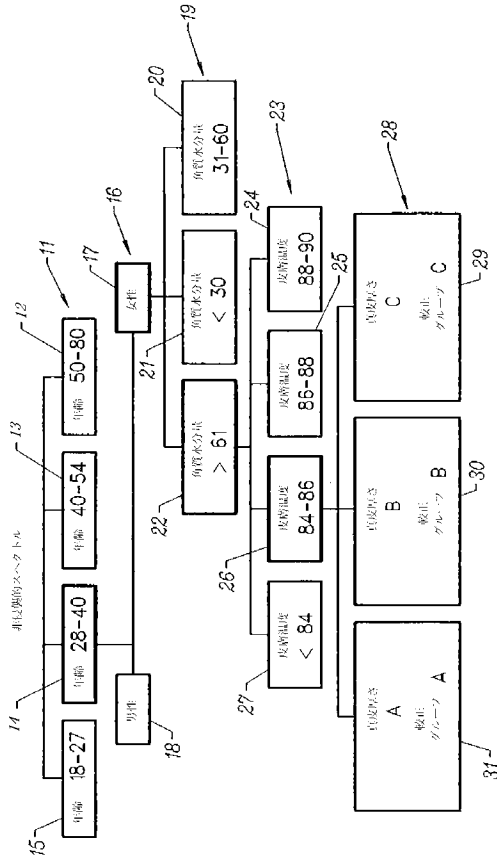
111、121 較正

122 非ファジー化

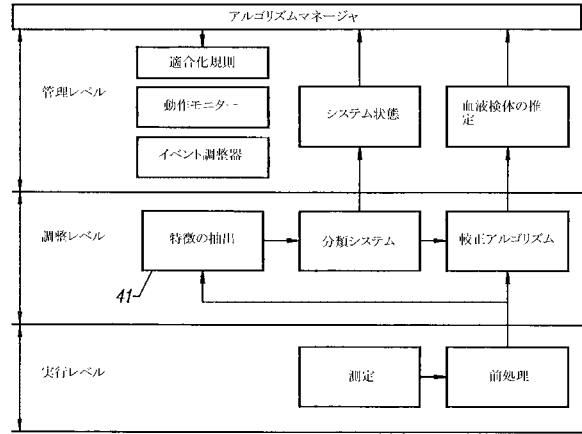
20

30

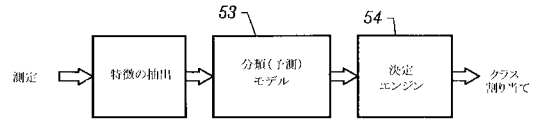
【図1】



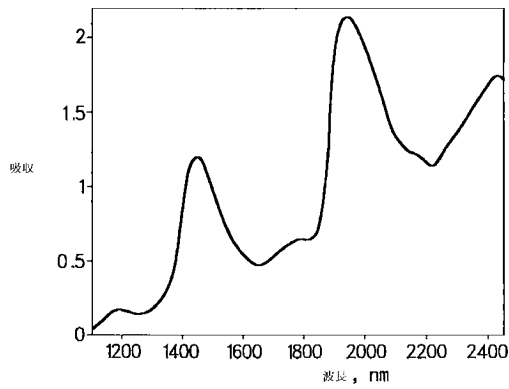
【図2】



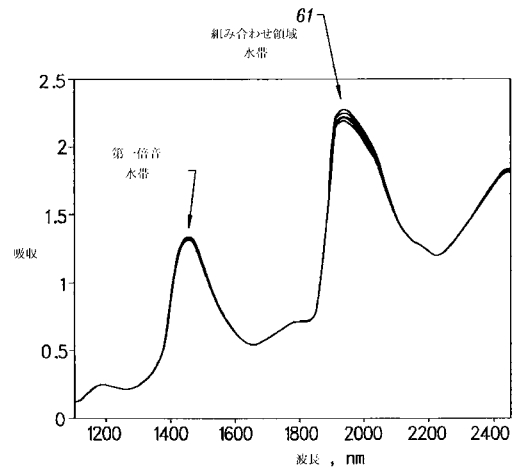
【図3】



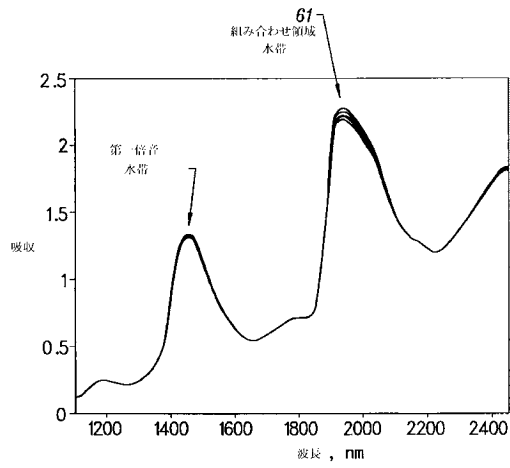
【図4】



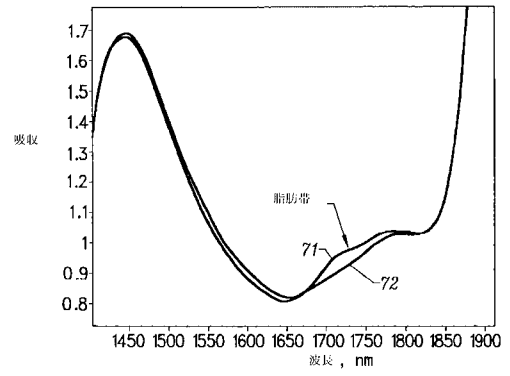
【図5】



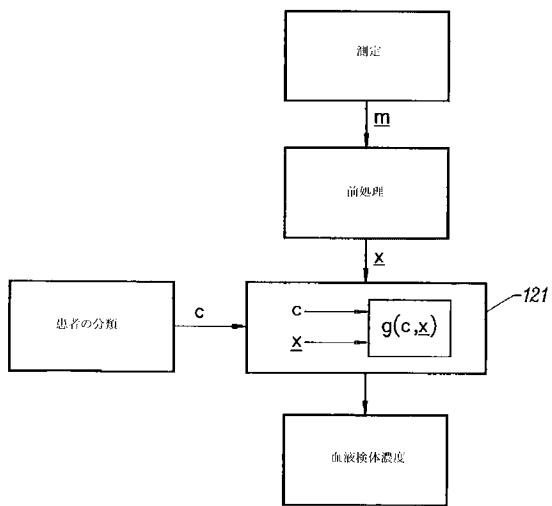
【図6】



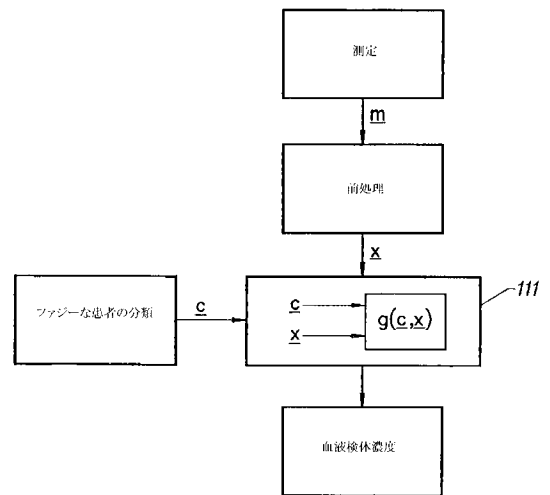
【図7】



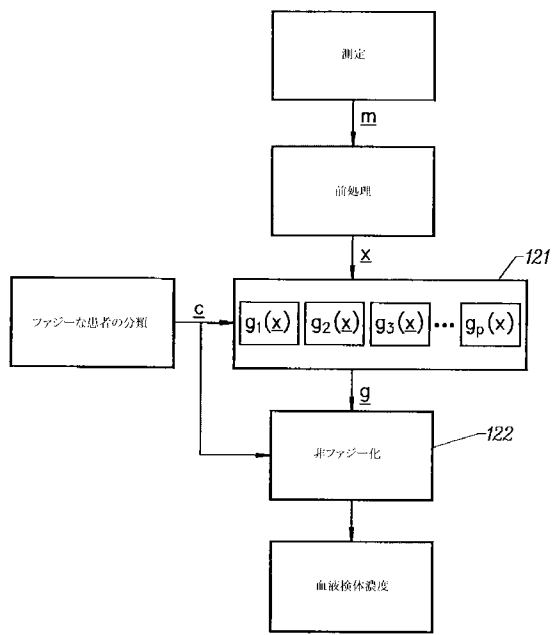
【図8】



【図9】



【図10】



## 【国際公開パンフレット】

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
17 October 2002 (17.10.2002)

PCT

(10) International Publication Number  
WO 02/080767 A1

- (51) International Patent Classification: A61B 5/00 (81) Designated States (*national*): AF, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CI, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GI, GM, GR, GU, HD, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.
- (21) International Application Number: PCT/US02/09114
- (22) International Filing Date: 22 March 2002 (22.03.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 09825.687 3 April 2001 (03.04.2001) US
- (71) Applicant: INSTRUMENTATION METRICS, INC. [US/US]; 7470 West Chandler Blvd., Chandler, AZ 85226 (US).
- (72) Inventors: BLANK, Thomas, B.; 2922 E. Tulsa Street, Chandler, AZ 85225 (US); MONFRE, Stephen, L.; 1289 East Palo Blanco Way, Gilbert, AZ 85296 (US); RUCHTI, Timothy, L.; 1501 West Sea Haze Drive, Gilbert, AZ 85233 (US); THENNADIL, Suresh.; 7250 S. Kyrene Road, Apt. 207, Tempe, AZ 85283 (US).
- (74) Agents: GLENN, Michael, A. et al.; Glenn Patent Group, 3475 Edison Way, Suite L, Menlo Park, CA 94025 (US).
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**  
with international search report  
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



WO 02/080767 A1

(54) Title: A MULTI-TIER METHOD OF DEVELOPING LOCALIZED CALIBRATION MODELS FOR NON-INVASIVE BLOOD ANALYTE PREDICTION

(57) Abstract: A method of multi-tier classification and calibration in noninvasive blood analyte prediction minimizes prediction error by limiting co-varying spectral interferences. Tissue samples are categorized based on subject demographic and instrumental skin measurements, including *in vivo* near-IR spectral measurements. A multi-tier intelligent pattern classification sequence organizes spectral data into clusters having a high degree of internal consistency in tissue properties. In each tier, categories are successively refined using subject demographics, spectral measurement information and other device measurements suitable for developing tissue classifications. The multi-tier classification approach to calibration utilizes multivariate statistical arguments and multi-tiered classification using spectral features. Variables used in the multi-tiered classification can be skin surface hydration, skin surface temperature, tissue volume hydration, and an assessment of relative optical thickness of the dermis by the near-IR fat band. All tissue parameters are evaluated using the NIR spectrum signal along key wavelength segments.



WO 02/080767

PCT/US02/09114

from the heavily overlapped spectral signatures of interfering substances. Defining a model that produces accurate compensation for numerous interferents may require spectral measurements at one hundred or more frequencies for a sizeable number of tissue samples.

5

- Accurate noninvasive estimation of blood analytes is also limited by the dynamic nature of the sample, the skin and living tissue of the patient. Chemical, structural and physiological variations occur produce dramatic changes in the optical properties of the measured tissue sample. See R. Anderson, J. Parrish, *The optics of human skin*, Journal of Investigative Dermatology, vol. 77(1), pp. 13-19 (1981); and W. Cheong, S. Prahl, A. Welch, *A review of the optical properties of biological tissues*, IEEE Journal of Quantum Electronics, vol. 26(12), pp. 2166-2185 (Dec. 1990); and D. Benaron, D. Ho, *Imaging (NIR) and quantitation (NIRS) in tissue using time-resolved spectrophotometry: the impact of statically and dynamically variable optical path lengths*, SPIE, vol. 1888, pp.10-21 (1993); and J. Conway, K. Norris, C. Bodwell, *A new approach for the estimation of body composition: infrared interactance*, The American Journal of Clinical Nutrition, vol. 40, pp. 1123-1140 (Dec. 1984); and S. Homma, T. Fukunaga, A. Kagaya, *Influence of adipose tissue thickness in near infrared spectroscopic signals in the measurement of human muscle*, Journal of Biomedical Optics, vol. 1(4), pp. 418-424 (Oct. 1996); and A. Profio, *Light transport in tissue*, Applied Optics, vol. 28(12), pp. 2216-2222 (June 1989); and M. Van Gemert, S. Jacques, H. Sterenborg, W. Sta, *Skin optics*, IEEE Transactions on Biomedical Engineering, vol. 36(12), pp. 1146-1154 (Dec. 1989); and B. Wilson, S. Jacques, *Optical reflectance and transmittance of tissues: principles and applications*, IEEE Journal of Quantum Electronics, vol. 26(12), pp. 2186-2199.

Overall sources of spectral variations include the following general categories:

1. Co-variation of spectrally interfering species. The near infrared spectral absorption profiles of blood analytes tend to overlap and vary simultaneously over brief time periods. This overlap leads to spectral

30

WO 02/080767

PCT/US02/09114

interference and necessitates the measurement of absorbance at more independently varying wavelengths than the number of interfering species.

2. Sample heterogeneity. The tissue measurement site has multiple layers and compartments of varied composition and scattering. The spectral absorbance versus wavelength measurement is related to a complex combination of the optical properties and composition of these tissue components. Therefore, the spectral response with changing blood analyte concentration is likely to deviate from a simple linear model.  
5  
10
3. State Variations. Variations in the subject's physiological state effect the optical properties of tissue layers and compartments over a relatively short period of time. Such variations, for example, may be related to hydration levels, changes in the volume fraction of blood in the tissue, hormonal stimulation, skin temperature fluctuations and blood hemoglobin levels. Subtle variations may even be expected in response to contact with an optical probe.  
15
4. Structural Variations. The tissue characteristics of individuals differ as a result of factors that include hereditary, environmental influences, the aging process, sex and body composition. These differences are largely anatomical and can be described as slowly varying structural properties producing diverse tissue geometry. Consequently, the tissue of a given subject may have distinct systematic spectral absorbance features or patterns that can be related directly to specific characteristics such as dermal thickness, protein levels and percent body fat. While the absorbance features may be repeatable within a patient, the structural variations in a population of patients may not be amenable to the use of a single mathematical calibration model. Therefore, differences between patients are a significant obstacle to the noninvasive measurement of blood analytes through NIR spectral absorbance.  
20  
25  
30

WO 02/080767

PCT/US02/09114

In a non-dispersive system, variations similar to (1) above are easily modeled through multivariate techniques such as multiple linear regression and factor-based algorithms. Significant effort has been expended to model the scattering properties of tissue in diffuse reflectance, although the problem outlined in (2) above has been largely unexplored. Variation of the type listed in (3) and (4) above causes significant nonlinear spectral response for which an effective solution has not been reported. For example, several reported methods of noninvasive glucose measurement develop calibration models that are specific to an individual over a short period of time. See K. Hazen, *Glucose determination in biological matrices using near-infrared spectroscopy*, Doctoral Dissertation, University of Iowa (Aug. 1995); and J. Burmeister, *In vitro model for human noninvasive blood glucose measurements*, Doctoral Dissertation, University of Iowa (Dec. 1997); and M. Robinson, R. Eaton, D. Haaland, G. Koepp, E. Thomas, B. Stallard and P. Robinson, *Noninvasive glucose monitoring in diabetic patients: a preliminary evaluation*, *Clin. Chem.*, vol. 38 (9), pp. 1618-1622 (1992). This approach avoids modeling the differences between patients and therefore cannot be generalized to more individuals. However, the calibration models have not been tested over long time periods during which variation of type (4) may require recalibration. Furthermore, the reported methods have not been shown to be effective over a range of type (3) variations.

#### **SUMMARY OF THE INVENTION**

The invention provides a Multi-Tier method for classifying tissue absorbance spectra that localizes calibration and sample spectra into local groups that are used to reduce variation in sample spectra due to co-variation of spectral interferences, sample heterogeneity, state variation and structural variation. Measurement spectra are associated with localized calibration models that are

WO 02/080767

PCT/US02/09114

designed to produce the most accurate estimates for the patient at the time of measurement. Classification occurs through extracted features of the tissue absorbance spectrum related to the current patient state and structure.

- 5 The invention also provides a method of developing localized calibration models from tissue absorbance spectra from a representative population of patients or physiological states of individual patients that have been segregated into groups. The groups or classes are defined on the basis of structural and state similarity such that the variation in tissue characteristics within a class is smaller than the  
10 variation between classes.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

- Figure 1 provides a representation of a Multi-Tiered Classification Tree structure,  
15 according to the invention;

Figure 2 is a block diagram of the architecture of an intelligent system for the noninvasive measurement of blood analytes, according to the invention;

- 20 Figure 3 is a block diagram of a pattern classification system, according to the invention;

Figure 4 is a noninvasive absorbance spectrum collected using a diffuse reflectance NIR spectrometer;

- 25 Figure 5 shows the spectra of repeated noninvasive measurements with no attempt to control tissue hydration;

- 30 Figure 6 shows the spectra of repeated noninvasive measurements using ambient humidity to control hydration, according to the invention;

Figure 7 shows a noninvasive absorbance spectrum having a pronounced fat band at 1710nm;

WO 02/080767

PCT/US02/09114

Figure 8 is a block schematic diagram of a general calibration system for mutually exclusive classes, according to the invention;

- 5 Figure 8 is a block schematic diagram of a general calibration system for fuzzy class assignments, according to the invention; and

Figure 10 is a block schematic diagram showing an example of parallel calibration models for fuzzy set assignments, according to the invention.

10

#### **DETAILED DESCRIPTION**

##### **MULTI-TIERED CLASSIFICATION**

- The classification of tissue samples using spectra and other electronic and demographic information can be approached using a wide variety of algorithms. A wide range of classifiers exists for separating tissue states into groups having high internal similarity: for example, Bayesian classifiers utilizing statistical distribution information; or nonparametric neural network classifiers that assume little *a priori* information. See K. Fukunaga, Intro to Statistical Pattern Recognition, Academic Pres, San Diego, CA (1990); and J. Hertz, A. Krogh, R. Palmer, Introduction To The Theory Of Neural Computation, Addison-Wesley Publishing Co., Redwood City CA (1991). The multi-tiered classification approach selected here provides the opportunity to grow and expand the classification database as more data become available. The multi-tiered classifier is similar to a hierarchic classification tree, but unlike a classification tree, the decision rules can be defined by crisp or fuzzy functions and the classification algorithm used to define the decision rule can vary throughout the tree structure.
- 15  
20  
25

- Referring now to Figure 1, an example of a Multi-Tiered Classification scheme is represented. A first tier 11 assigns sample spectra according to pre-defined age groups: 18 – 27 (15), 28 – 40 (14), 40 – 54 (13) and 55 - 80 years old (12). As
- 30

WO 02/080767

PCT/US02/09114

indicated, a sample has been assigned to the 28 - 40 age group. A second tier 16 assigns samples to classes 18, 17 according to sex, in this case female. A third tier 19, groups according to stratum corneum hydration: 31 - 60 (20); <30 (21) and >61 corneometer units (22); in this case, >61. A fourth tier 23, groups according to skin temperature: 88 - 90 (24); 86 - 88 (25); 84 - 86 and <84 degrees; in this case 84 - 86 degrees. In this way, a determination of class membership is made within each tier in the multi-tiered structure. Finally, in a last tier 28, a final class assignment is made into one of three pre-defined groups 29, 30 and 31 according to relative optical thickness of the dermis.

10 For economy's sake, only the branching adjacent the selected classes is completely shown in Figure 1, though there would be many more intermediate and final classification categories in a full multi-tiered classification structure. For example, at the fourth tier 23 of Figure, there would be ninety-six possible classifications for a tissue measurement spectrum; at the final tier, there would be two hundred eighty-eight possible classifications. The foregoing description of a Multi-Tier Classification structure is meant to be exemplary only. One skilled in the art will appreciate that an actual classification structure could have more or fewer tiers, and different decision rules could be utilized at each tier than have been utilized in the example.

#### FEATURE EXTRACTION

As previously indicated, at each tier in the classification structure, classification is made based on a *priori* knowledge of the sample, or on the basis of instrumental measurements made at the tissue measurement site. In the example of Figure 1, the first two tiers utilize a *priori* information about the sample: subject age and sex. Successive tiers utilize information gained from instrumental measurements at the tissue measurement site. Further classification occurs on the basis of extracted features from the tissue absorbance spectra themselves.

Feature extraction is any mathematical transformation that enhances a quality or aspect of the sample measurement for interpretation. See R. Duda, P. Hart,

WO 02/080767

PCT/US02/09114

Pattern Classification and Scene Analysis, John Wiley and Sons, New York (1973).  
 Figure 2 shows a block diagram of an intelligent measurement system for noninvasive blood analyte prediction, fully described in the parent application to the current application: S. Malin and T. Ruchti, *An Intelligent System For Noninvasive*  
 5 *Blood Analyte Prediction*, U.S. Patent Application Ser. No. 09/359,191; July 22, 1999, The purpose of feature extraction 41 in Figure 2 is to concisely represent the structural properties and physiological state of the tissue measurement site. The set of features is used to classify the patient and determine the calibration model(s) most useful for blood analyte prediction.

10

The features are represented in a vector,  $z \in \mathfrak{R}^M$  that is determined from the preprocessed measurement through

$$z = f(\lambda, x) \quad (1)$$

where  $f: \mathfrak{R}^N \rightarrow \mathfrak{R}^M$  is a mapping from the measurement space to the feature space.

15

Decomposing  $f(\cdot)$  will yield specific transformations,  $f_i(\cdot): \mathfrak{R}^N \rightarrow \mathfrak{R}^{M_i}$  for determining a specific feature. The dimension,  $M_i$ , indicates whether the  $i^{\text{th}}$  feature is a scalar or a vector and the aggregation of all features is the vector  $z$ . When a feature is represented as a vector or a pattern, it exhibits a certain structure indicative of an underlying physical phenomenon.

20

The individual features are divided into two categories:

1. abstract and
2. simple.

25

Abstract features do not necessarily have a specific interpretation related to the physical system. Specifically, the scores of a principal component analysis are useful features although their physical interpretation is not always known. The utility of the principal component analysis is related to the nature of the tissue absorbance spectrum. The most significant variation in the tissue spectral absorbance is not caused by a blood analyte but is related to the state, structure  
 30 and composition of the measurement site. This variation is modeled by the primary principal components. Therefore, the leading principal components tend

WO 02/080767

PCT/US02/09114

to represent variation related to the structural properties and physiological state of the tissue measurement site.

Simple features are derived from an *a priori* understanding of the sample and can be related directly to a physical phenomenon. Useful features that can be calculated from NIR spectral absorbance measurements include but are not limited to:

1. Thickness of adipose tissue. See J. Conway, K. Norris, C. Bodwell, *A new approach for the estimation of body composition: infrared interactance*, The American Journal of Clinical Nutrition, vol. 40, pp. 1123-1140 (Dec. 1984) and S. Homma, T. Fukunaga, A. Kagaya, *Influence of adipose tissue thickness in near infrared spectroscopic signals in the measurement of human muscle*, Journal of Biomedical Optics, vol. 1(4), pp. 418-424 (Oct. 1996).
2. Tissue hydration. See K. Martin, *Direct measurement of moisture in skin by NIR spectroscopy*, J. Soc. Cosmet. Chem., vol. 44, pp. 249-261 (Sept./Oct. 1993).
3. Magnitude of protein absorbance. See J. Conway, *et al.*, *supra*.
4. Scattering properties of the tissue. See A. Profio, *Light transport in tissue*, Applied Optics, vol. 28(12), pp. 2216-2222 (June 1989) and W. Cheong, S. Pahl, A. Welch, *A review of the optical properties of biological tissues*, IEEE Journal of Quantum Electronics, vol. 26(12), pp. 2166-2185 (Dec. 1990); and R. Anderson, J. Parrish. *The optics of human skin*, Journal of Investigative Dermatology, vol. 77(1), pp. 13-19 (1981).
5. Skin thickness. See Anderson, *et al.*, *supra*; and Van Gemmert, *et al.*, *supra*.
6. Temperature related effects. See Funkunga, *supra*.
7. Age related effects. See W. Andrew, R. Behnke, T. Sato, *Changes with advancing age in the cell population of human dermis*, Gerontology, vol. 10, pp. 1-19 (1964/65); and W. Montagna, K. Carlisle, *Structural changes in aging human skin*, The Journal of Investigative Dermatology, vol. 73, pp. 47-53 (1979); and

WO 02/080767

PCT/US02/09114

19 J. Brocklehurst, Textbook of Geriatric Medicine and Gerontology, pp.593-623, Churchill Livingstone, Edinburgh and London (1973).

- 5 8. Spectral characteristics related to sex. See T. Ruchti, Internal Reports and Presentations, Instrumentation Metrics, Inc.
9. Pathlength estimates. See R. Anderson, *et al.*, *supra* and S. Matcher, M. Cope, D. Delpy, *Use of water absorption spectrum to quantify tissue chromophore concentration changes in near-infrared spectroscopy*, Phys. Med. Biol., vol. 38, pp. 177-196 (1993).
- 10 10. Volume fraction of blood in tissue. See Wilson, *et al.*, *supra*.
11. Spectral characteristics related to environmental influences.
- 15

Spectral decomposition is employed to determine the features related to a known spectral absorbance pattern. Protein and fat, for example, have known absorbance signatures that can be used to determine their contribution to the tissue spectral absorbance. The measured contribution is used as a feature and represents the underlying variable through a single value.

20

Features related to demographic information, such as age, are combinations of many different effects that cannot be represented by a single absorbance profile. Furthermore, the relationship of demographic variables and the tissue spectral absorbance is not deterministic. For example, dermal thickness and many other tissue properties are statistically related to age but also vary substantially as a result of hereditary and environmental influences. Therefore, factor based methods are employed to build models capable of representing variation in the measured absorbance related to the demographic variable. The projection of a measured absorbance spectrum onto the model constitutes a feature that represents the spectral variation related to the demographic variable. The compilation of the abstract and simple features constitutes the  $M$ -dimensional feature space. Due to redundancy of information across the set of features, optimum feature selection and/or data compression is applied to enhance the robustness of the classifier.

25

30

35

WO 02/080767

PCT/US02/09114

**CLASSIFICATION**

5 The goal of feature extraction is to define the salient characteristics of measurements that are relevant for classification. Feature extraction is performed at branching junctions of the multi-tiered classification tree structure. The goal of the classification step is to assign the calibration model(s) most appropriate for a particular noninvasive measurement. In this step the patient is assigned to one of many predefined classes for which a calibration model has been developed and tested. Since the applied calibration model is developed for similar tissue absorbance spectra, the blood analyte predictions are more accurate than those obtained from a universal calibration model.

As depicted in Figure 3, pattern classification generally involves two steps:

- 15
1. a mapping step in which a classification model 53 measures the similarity of the extracted features to predefined classes; and
  2. an assignment step in which a decision engine 54 assigns class membership.

20 Within this framework, two general methods of classification are proposed. The first uses mutually exclusive classes and therefore assigns each measurement to one class. The second scheme utilizes a fuzzy classification system that allows class membership in more than one class simultaneously. Both methods rely on previously defined classes, as described below.

25

**Class Definition**

The development of the classification system requires a data set of exemplar spectral measurements from a representative sampling of the population. Class definition is the assignment of the measurements in the exploratory data set to classes. After class definition, the measurements and class assignments are used to determine the mapping from the features to class assignments.

30

WO 02/080767

PCT/US02/09114

Class definition is performed through either a supervised or an unsupervised approach. See Y. Pao, Adaptive Pattern Recognition and Neural Networks, Addison-Wesley Publishing Co., Reading MA (1989). In the supervised case, 5 classes are defined through known differences in the data. The use of a *priori* information in this manner is the first step in supervised pattern recognition, which develops classification models when the class assignment is known. For example, the majority of observed spectral variation can be modeled by three abstract factors, which are related to several physical properties including body fat, 10 tissue hydration and skin thickness. Categorizing patients on the basis of these three features produces eight different classes if each feature is assigned a "high" and "low" value. The drawback to this approach is that attention is not given to spectral similarity and the number of classes tends to increase exponentially with the number of features.

15 Unsupervised methods rely solely on the spectral measurements to explore and develop clusters or natural groupings of the data in feature space. Such an analysis optimizes the within cluster homogeneity and the between cluster separation. Clusters formed from features with physical meaning can be 20 interpreted based on the known underlying phenomenon causing variation in the feature space. However, cluster analysis does not utilize *a priori* information and can yield inconsistent results.

A combination of the two approaches utilizes *a priori* knowledge and exploration of 25 the feature space for naturally occurring spectral classes. In this approach, classes are first defined from the features in a supervised manner. Each set of features is divided into two or more regions and classes are defined by combinations of the feature divisions. A cluster analysis is performed on the data and the results of the two approaches are compared. Systematically, the clusters 30 are used to determine groups of classes that can be combined. After conglomeration, the number of final class definitions is significantly reduced according to natural divisions in the data.

WO 02/080767

PCT/US02/09114

Subsequent to class definition, a classifier is designed through supervised pattern recognition. A model is created, based on class definitions, that transforms a measured set of features to an estimated classification. Since the ultimate goal of the classifier is to produce robust and accurate calibration models, an iterative  
5 approach must be followed in which class definitions are optimized to satisfy the specifications of the measurement system.

#### Statistical Classification

The statistical classification methods are applied to mutually exclusive classes  
10 whose variation can be described statistically. See J. Bezdek, S. Pal, eds, Fuzzy Models for Pattern Recognition, IEEE Press, Piscataway NJ (1992). Once class definitions have been assigned to a set of exemplary samples, the classifier is designed by determining an optimal mapping or transformation from the feature space to a class estimate which minimizes the number of misclassifications. The  
15 form of the mapping varies by method as does the definition of "optimal". Existing methods include linear Discriminant analysis, SIMCA, k nearest-neighbor and various forms of artificial neural networks. See Fukunaga, *supra*; and Hertz, *et al.*, *supra*; and Martin, *supra*; and Duda, *et al.*, *supra*; and Pao, *supra*; and S. Wold, M. Sjöström, *SIMCA: A method for analyzing chemical data in terms of similarity and analogy*, Chemometrics: Theory and Application, ed. B.R. Kowalski, ACS  
20 Symposium Series, vol. 52 (1977); and S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice-Hall, Upper Saddle River NJ (1994). The result is a function or algorithm that maps the feature to a class,  $c$ , according to

$$c = f(z) \quad (2)$$

25 where  $c$  is an integer on the interval  $[1, P]$  and  $P$  is the number of classes. The class is used to select or adapt the calibration model as discussed in the Calibration Section.

#### Fuzzy Classification

30 While statistically based class definitions provide a set of classes applicable to blood analyte estimation, the optical properties of the tissue sample resulting in spectral variation change over a continuum of values. Therefore, the natural

WO 02/080767

PCT/US02/09114

variation of tissue thickness, hydration levels and body fat content, among others, results in class overlap. Distinct class boundaries do not exist and many measurements are likely to fall between classes and have a statistically equal chance of membership in any of several classes. Therefore, "hard" class boundaries and mutually exclusive membership functions appear contrary to the nature of the target population.

A more versatile method of class assignment is based on fuzzy set theory. See Bezdek, *et al.*, *supra*; and C. Chen, ed., *Fuzzy Logic and Neural Network Handbook*, IEEE Press, Piscataway NJ (1996); and L. Zadeh, *Fuzzy Sets, Inform. Control*, vol. 8, pp. 338-353 (1965). Generally, membership in fuzzy sets is defined by a continuum of grades and a set of membership functions that map the feature space into the interval [0,1] for each class. The assigned membership grade represents the degree of class membership with "1" corresponding to the highest degree. Therefore, a sample can simultaneously be a member of more than one class.

The mapping from feature space to a vector of class memberships is given by

$$c_k = f_k(z) \quad (2)$$

where  $k = 1, 2, \dots, P$ ,  $f_k(\cdot)$  is the membership function of the  $k^{\text{th}}$  class,  $c_k \in [0, 1]$  for all  $k$  and the vector  $c \in \mathcal{R}^P$  is the set of class memberships. The membership vector provides the degree of membership in each of the predefined classes and is passed to the calibration algorithm.

The design of membership functions utilizes fuzzy class definitions similar to the methods previously described. Fuzzy cluster analysis can be applied and several methods, differing according to structure and optimization approach can be used to develop the fuzzy classifier. All methods attempt to minimize the estimation error of the class membership over a population of samples.

30

#### **MULTI-TIERED CALIBRATION**

WO 02/080767

PCT/US02/09114

Blood analyte prediction occurs by the application of a calibration model to the preprocessed measurement as depicted in Figure 2. The proposed prediction system involves a calibration or a set of calibration models that are adaptable or selected on the basis of the classification step.

5

#### DEVELOPMENT OF LOCALIZED CALIBRATION MODELS

Accurate blood analyte prediction requires calibration models that are capable of compensating for the co-varying interferences, sample heterogeneity, state and structural variations encountered. Complex mixtures of chemically absorbing species that exhibit substantial spectral overlap between the system components are solvable only with the use of multivariate statistical models. However, prediction error increases with increasing variation in interferences that also co-vary with analyte concentration in calibration data. Therefore, blood analyte prediction is best performed on measurements exhibiting smaller interference variations that correlate poorly with analyte concentration in the calibration set data. Since it may not be possible to make all interference variations random, it is desirable to limit the range of spectral interference variation in general.

The principle behind the multi-tiered classification and calibration system is based on the properties of a generalized class of algorithms that are required to compensate for overlapped interfering signals in the presence of the desired analyte signal. See H. Martens, T. Naes, Multivariate Calibration, John Wiley and Sons, New York (1989). The models used in this application require the measurement of multiple independent variables, designated as  $x$ , to estimate a single dependent variable, designated as  $y$ . For example,  $y$  may be tissue glucose concentration, and  $x$  may represent a vector,  $[x_1 \ x_2 \ \dots \ x_n]$ , consisting of the noninvasive spectrum signal intensities at each of  $n$  wavelengths.

The generalized form of a model to be used in the calculation of a single glucose estimate uses a weighted summation of the noninvasive spectrum as in Equation 4. The weights,  $w$ , are referred to as the regression vector.

$$y = \sum w_i x_i \quad (4)$$

WO 02/080767

PCT/US02/09114

The weights define the calibration model and must be calculated from a given calibration set of noninvasive spectra in the spectral matrix  $X$ , and associated reference values  $y$  for each spectrum:

$$w = (X^T X)^{-1} X^T y \quad (5)$$

The modeling error that might be expected in a multivariate system using Equation 5 can be estimated using a linear additive mixture model. Linear additive mixtures are characterized by the definition that the sum of the pure spectra of the individual constituents in a mixture equals the spectra of the mixture. Linear mixture models are useful in assessing the general limitations of multivariate models that are based on linear additive systems and those, noninvasive blood analysis, for example, that can be expected to deviate somewhat from linear additive behavior.

Figure 4 shows an exemplary noninvasive absorbance spectrum. A set of spectral measurements may be represented as a matrix  $X$  where each row corresponds to an individual sample spectrum and each column represents the signal magnitude at a single wavelength. The measurement matrix can be represented as a linear additive mixture model with a matrix of instrument baseline variations  $B_0$ , a matrix of spectra of the pure components  $K$ , and the concentrations of the pure components,  $Y$ , and random measurement noise present in the measurement of each spectrum,  $E$ .

$$X = B_0 + YK' + E \quad (6)$$

The linear additive model can be broken up further into interferences and analytes as an extended mixture model.

$$X = B_0 + YK' + TP' + E \quad (7)$$

WO 02/080767

PCT/US02/09114

In equation 4,  $T$  is a matrix representing the concentration or magnitude of interferents in all samples, and  $P$  represents the pure spectra of the interfering substances or effects present. Any spectral distortion can be considered an interferent in this formulation. For example, the effects of variable sample scattering and deviations in optical sampling volume must be included as sources of interference in this formulation. The direct calibration for a generalized least squares model on analyte  $y$  is

$$y_{GLS} = (K^T \Sigma^{-1} K)^{-1} K^T \Sigma^{-1} (x - k_0); \quad (8)$$

10

where  $\Sigma$  is defined as the covariance matrix of the interfering substances or spectral effects,  $\hat{U}$  is defined as the measurement noise,  $x$  is the spectral measurement, and  $k_0$  is the instrument baseline component present in the spectral measurement.

$$\Sigma = P^T (H^T)^{-1} P + \text{diag}(\sigma^2) \quad (9)$$

20 The derived mean squared error (MSE) of such a generalized least squares predictor is found in Martens, *et al.*, *supra*.

$$MSE(y_{GLS}) = \text{trace}(K^T \Sigma^{-1} K)^{-1} \quad (10)$$

25 Equation 7 describes the generalized limitations of least squares predictors in the presence of interferents. If  $K$  represents the concentrations of blood glucose, a basic interpretation of Equation 7 is: the mean squared error in glucose estimates increases with increasing variation in interferences that also co-vary with glucose concentration in calibration data. Therefore, the accurate estimation of glucose is best performed on measurements exhibiting smaller interference variations that poorly correlate with glucose concentration in the calibration set data. Since it may

WO 02/080767

PCT/US02/09114

not be possible to make all interference variations random with glucose, it is desirable to limit the range of spectral interference variation in general.

5 The Multi-Tier Classification provides a method for limiting variation of spectral  
interferents by placing sample measurements into groups having a high degree of  
internal consistency. Groups are defined based on *a priori* knowledge of the  
sample, instrumental measurements at the tissue measurement site, and  
extracted features. With each successive tier, samples are further classified such  
that variation between spectra within a group is successively limited. Tissue  
10 parameters to be utilized in class definition may include: stratum corneum  
hydration, tissue temperature, and dermal thickness.

#### TISSUE HYDRATION

The stratum corneum (SC), or horny cell layer covers about 10-15  $\mu\text{m}$  thickness  
15 of the underside of the arm. The SC is composed mainly of keratinous dead  
cells, water and some lipids. See D. Bommannan, R. Potts, R. Guy, *Examination of  
the Stratum Corneum Barrier Function In Vivo by Infrared Spectroscopy*, J. Invest.  
Dermatol., vol. 95, pp 403-408 (1990). Hydration of the SC is known to vary over  
time as a function of room temperature and relative humidity. See J. Middleton, B.  
20 Allen, *Influence of temperature and humidity on stratum corneum and its relation to  
skin chapping*, J. Soc. Cosmet. Chem., vol. 24, pp. 239-43 (1973). Because it is the  
first tissue penetrated by the spectrometer incident beam, more photons sample  
the SC than any other part of the tissue sample. Therefore, the variation of a strong  
near IR absorber like water in the first layer of the tissue sample can act to change  
25 the wavelength and depth intensity profile of the photons penetrating beneath the  
SC layer.

The impact of changes in SC hydration can be observed by a simple experiment.  
In the first part of the experiment, the SC hydration is allowed to range freely with  
30 ambient conditions. In the second part of the experiment, variations in SC  
hydration are limited by controlling relative humidity to a high level at the skin  
surface prior to measurement. Noninvasive measurements using uncontrolled

WO 02/080767

PCT/US02/09114

and controlled hydration experiments on a single individual are plotted in Figures 5 and 6, respectively. Changes in the water band 61 at 1900nm can be used to assess changing surface hydration. It is apparent that the range of variation in the water band 61 at 1900nm is considerably narrower in Figure 6 than in Figure 5.

- 5 Since surface hydration represents a large variable in the spectral measurement, it is a valuable component for use in categorizing similarity in tissue samples.

#### TISSUE TEMPERATURE

- The temperature of the measured tissue volume varies from the core body temperature, at the deepest level of penetration, to the skin surface temperature, which is generally related to ambient temperature, location and the amount of clothing at the tissue measurement site. The spectrum of water, which comprises about 65% of living human tissue is the most dominant spectral component at all depths sampled in the 1100-2500 nm wavelength range. These two facts, along with the known temperature-induced shifting of the water band at 1450 nm, combine to substantially complicate the interpretation of information about many blood analytes, including glucose. It is apparent that a range of temperature states exist in the volume of sampled living tissue and that the range and distribution of states in the tissue depend on the skin surface temperature. Furthermore, the index of refraction of skin is known to change with temperature. Skin temperature may therefore be considered an important categorical variable for use in the Multi-Tier Classification to identify groups for the generation of calibration models and prediction.

#### 25 OPTICAL THICKNESS OF DERMIS

- Repeated optical sampling of the tissue is necessary to calibrate to blood constituents. Because blood represents but a part of human tissue, and blood analytes only reside in fractions of the tissue, changes in the optical sampling of tissue may change the magnitude of the analyte signal for unchanging levels of blood analytes. This kind of a sampling effect may confound efforts at calibration by changing the signal strength for specific levels of analyte.

WO 02/080767

PCT/US02/09114

Categorization of optical sampling depth is pursued by analyzing spectral marker bands of the different layers. For example, the first tissue layer under the skin is the subcutaneous adipose tissue, consisting mainly of fat. The strength of the fat absorbance band can be used to assess the relative photon flux that has penetrated to the subcutaneous tissue level. A more pronounced fat band means that a greater photon flux has reached the adipose tissue and returned to the detector. In Figure 7, spectra with pronounced 71 and normal 72 fat bands are presented. The most important use of the optical thickness is to assess the degree of hydration in the interior tissue sampled by the optical probe. Optical thickness may also be a strong function of gender and body type, therefore this property measurement would be useful for assessing interior hydration states within a single individual.

The following sections describe the calibration system for the two types of classifiers, mutually exclusive and fuzzy.

#### MUTUALLY EXCLUSIVE CLASSES

In the general case, the designated classification is passed to a nonlinear model that provides a blood analyte prediction based on the patient classification and spectral measurement. This process, illustrated in Figure 8, involves the modification of the estimation strategy for the current subject according to the structural tissue properties and physiological state manifested in the absorbance spectrum.

This general architecture necessitates a nonlinear calibration model such as nonlinear partial least squares or artificial neural networks since the mapping is highly nonlinear. The blood analyte prediction for the preprocessed measurement  $x$  with classification specified by  $c$  is given by

$$y = g(c, x) \quad (11)$$

where  $g(\cdot)$  is a nonlinear calibration model which maps  $x$  and  $c$  to an estimate of the blood analyte concentration,  $y$ .

WO 02/080767

PCT/US02/09114

In the preferred realization, a different calibration is realized for each class. The estimated class is used to select one of  $p$  calibration models most appropriate for blood analyte prediction using the current measurement. Given that  $k$  is the class estimate for the measurement, the blood analyte prediction is

$$\hat{y} = g_k(x), \quad (12)$$

where  $g_k(\cdot)$  is the calibration model associated with the  $k^{\text{th}}$  class.

The calibrations are developed from a set of exemplar absorbance spectra with reference blood analyte values and pre-assigned classification definitions. This set, denoted the "calibration set", must have sufficient samples to completely represent the range of physiological states to be encountered in the patient population. The  $p$  different calibration models are developed individually from the measurements assigned to each of the  $p$  classes. The models are realized using known methods including principal component regression, partial least squares regression and artificial neural networks. See Hertz, *et al., supra*; and Pao, *supra*; and Haykin, *supra*; and Martens, *et al., supra*; and N. Draper, H. Smith, Applied Regression Analysis, 2<sup>nd</sup> ed., John Wiley and Sons, New York (1981). The various models associated with each class are evaluated on the basis of an independent test set or cross validation and the "best" set of models are incorporated into the Multi-tier Classification. Each class of patients then has a calibration model specific to that class.

#### **FUZZY CLASS MEMBERSHIP**

When fuzzy classification is employed the calibration is passed a vector of memberships rather than a single estimated class. The vector,  $c$ , is utilized to determine an adaptation of the calibration model suitable for blood analyte prediction or an optimal combination of several blood analyte predictions. In the general case, illustrated in Figure 9, the membership vector and the preprocessed absorbance spectrum are both used by a single calibration 111 for blood analyte prediction. The calculation is given by

$$y = g(c, x) \quad (13)$$

WO 02/080767

PCT/US02/09114

where  $g(\cdot)$  is a nonlinear mapping determined through nonlinear regression, nonlinear partial least squares or artificial neural networks. The mapping is developed from the calibration set described previously and is generally complex.

- 5 The preferred realization, shown in Figure 10, has separate calibrations 121 for each class. However, each calibration is generated using all measurements in the calibration set by exploiting the membership vector assigned to each measurement. In addition, the membership vector is used to determine an optimal combination of the  $p$  blood analyte predictions from all classes through
- 10 defuzzification 122. Therefore, during calibration development, a given measurement of the calibration set has the opportunity to impact more than one calibration model. Similarly, during prediction more than one calibration model is used to generate the blood analyte estimate.
- 15 Each of the  $p$  calibration models is developed using the entire set of calibration data. However, when the  $k^{\text{th}}$  calibration model is calculated, the calibration measurements are weighted by their respective membership in the  $k^{\text{th}}$  class. As a result, the influence of a sample on the calibration model of a particular class is a function of its membership in the class.

20

In the linear case, weighted least squares is applied to calculate regression coefficients and, in the case of factor based methods, the covariance matrix. See Duda, *et al.*, *supra*. Given a matrix of absorbance spectra  $X_k \in \mathbb{R}^{r \times w}$  and reference blood analyte concentrations  $Y \in \mathbb{R}^r$  where  $r$  is the number of measurement spectra

25 and  $w$  is the number wavelengths, let the membership in class  $k$  of each absorbance spectrum be the elements of  $C_k \in \mathbb{R}^r$ . Then the principal components are given by

$$F^* = X_k M, \quad (14)$$

where  $M$  is the matrix of the first  $n$  eigenvectors of  $P$ . The weighted covariance

30 matrix  $P$  is determined through

$$P = X_k^T V X_k, \quad (15)$$

WO 02/080767

PCT/US02/09114

where  $V$  is a square matrix with the elements of  $C_k$  on the diagonal. The regression matrix,  $B$ , is determined through

$$B = (F^T V F)^{-1} F^T V Y. \quad (16)$$

- 5 When an iterative method is applied, such as artificial neural networks, the membership is used to determine the frequency the samples are presented to the learning algorithm. Alternatively, an extended Kalman filter is applied with a covariance matrix scaled according to  $V$ .
- 10 The purpose of defuzzification is to find an optimal combination of the  $p$  different blood analyte predictions, based on a measurement's membership vector that produces accurate blood analyte predictions. Therefore, defuzzification is a mapping from the vector of blood analyte predictions and the vector of class memberships to a single analyte prediction. The defuzzifier can be denoted as
- 15 transformation such that

$$\hat{y} = d(c, [y_1 \ y_2 \ y_3 \ \dots \ y_p]), \quad (17)$$

- where  $d(\cdot)$  is the defuzzification function,  $c$  is the class membership vector and  $y_k$  is the blood analyte prediction of the  $k^{\text{th}}$  calibration model. Existing methods of defuzzification, such as the centroid or weighted average, are applied for small
- 20 calibration sets. However, if the number of samples is sufficient,  $d(\cdot)$  is generated through a constrained nonlinear model.

#### INSTRUMENT DESCRIPTION

- The Multi-Tiered Classification and Calibration is implemented in a scanning
- 25 spectrometer which determines the NIR absorbance spectrum of the subject forearm through a diffuse reflectance measurement. The instrument employs a quartz halogen lamp, a monochromator and InGAs detectors. The detected intensity from the sample is converted to a voltage through analog electronics and digitized through a 16-bit A/D converter. The spectrum is passed to the Intelligent
- 30 Measuring System (IMS) for processing and results in either a glucose prediction or a message indicating an invalid scan.

WO 02/080767

PCT/US02/09114

Although the invention is described herein with reference to the preferred embodiment, one skilled in the art will readily appreciate that other applications may be substituted for those set forth herein without departing from the spirit and scope of the present invention. Accordingly, the invention should only be limited by the claims included below.

WO 02/080767

PCT/US02/09114

**CLAIMS**

What is claimed is:

5

1. A method of developing a multi-tiered calibration model for estimating concentration of a target blood analyte from measured tissue spectra, comprising the steps of:
  - providing a calibration set, wherein said calibration set comprises a data set of exemplar spectral measurements from a representative sampling of a subject population;
  - initially, classifying said spectral measurements into previously defined classes based on *a priori* information pertaining to said subject;
  - further classifying said measured spectrum into previously defined classes based on at least one instrumental measurement at a tissue measurement site;
  - extracting at least one feature from said measured spectrum for still further classification;
  - wherein a decision rule makes class assignments.
- 20 2. The method of Claim 1, wherein said initial classification step comprises the steps of:
  - in a first tier, classifying said measured spectrum into previously defined classes based on subject's age; and
  - in a second tier, further classifying said measured spectrum into previously defined classes based on subject's sex.
- 25 3. The method of Claim 1, wherein said subsequent classification step further comprises the steps of:
  - in a third tier, further classifying said measured spectrum into previously defined classes based on an estimation of stratum corneum hydration at said tissue measurement site; and
  - in a fourth tier, further classifying said measured spectrum into previously defined classes based on skin temperature at said tissue measurement site.
- 30 4. The method of Claim 3, wherein said stratum corneum hydration estimate is based on a measurement of ambient humidity at said tissue measurement site.
- 35

WO 02/080767

PCT/US02/09114

5. The method of Claim 1, wherein said feature extraction step comprises any mathematical transformation that enhances a quality or aspect of sample measurement for interpretation to represent concisely structural properties and physiological state of a tissue measurement site, wherein a resulting set of features is used to classify a subject and determine a calibration model that is most useful for blood analyte prediction.

6. The method of Claim 5, wherein said features are represented in a vector,  $z \in \mathfrak{R}^M$  that is determined from a preprocessed measurement through:

$$z = f(\lambda, x)$$

10 where  $f: \mathfrak{R}^N \rightarrow \mathfrak{R}^M$  is a mapping from a measurement space to a feature space, wherein decomposing  $f(\bullet)$  yields specific transformations,  $f_i(\bullet): \mathfrak{R}^N \rightarrow \mathfrak{R}^{M_i}$  for determining a specific feature, wherein the dimension  $M_i$  indicates whether an  $i^{\text{th}}$  feature is a scalar or a vector and an aggregation of all features is the vector  $z$ , and wherein a feature exhibits a certain structure indicative of an underlying physical phenomenon when said feature is represented as a vector or a pattern.

7. The method of Claim 6, wherein individual features are divided into categories, said categories comprising:

20 abstract features that do not necessarily have a specific interpretation related to a physical system; and  
simple features that are derived from an *a priori* understanding of a sample and that can be related directly to a physical phenomenon.

8. The method of Claim 7, wherein features that can be calculated from NIR spectral absorbance measurements include any of:

25 thickness of adipose tissue;  
hematocrit level;  
tissue hydration;  
30 magnitude of protein absorbance;  
scattering properties of said tissue;  
skin thickness;  
temperature related effects;  
age related effects;  
spectral characteristics;  
35 pathlength estimates;  
volume fraction of blood in tissue; and  
spectral characteristics related to environmental influences.

WO 02/080767

PCT/US02/09114

9. The method of Claim 1, further comprising the step of:  
employing spectral decomposition to determine features related to a known spectral  
absorbance pattern.
- 5 10. The method of Claim 1, further comprising the step of:  
employing factor-based methods to build a model capable of representing  
variation in a measured absorbance spectrum related to a demographic variable;  
wherein projection of a measured absorption onto said model constitutes a  
feature that represents spectral variation related to said demographic variable.
- 10
11. The method of Claim 1, wherein said feature extraction step assigns a  
measurement to one of many predefined classes.
- 15 12. The method of Claim 1, further comprising the steps of;  
measuring the similarity of a feature to predefined classes; and  
assigning class membership.
- 20 13. The method of Claim 1, further comprising the step of;  
using measurements and class assignments to determine a mapping from  
features to class assignments.
- 25 14. The method of Claim 13, further comprising the steps of:  
defining classes from said features in a supervised manner, wherein each set of  
features is divided into two or more regions, and wherein classes are defined by  
combination of feature divisions;  
performing a cluster analysis on the spectral data to determine groups of said  
defined classes that can be combined, wherein the final number of class definitions is  
significantly reduced;  
30 designing a classifier subsequent to class definition through supervised pattern  
recognition by determining an optimal mapping or transformation from the feature space  
to a class estimate that minimizes the number of misclassifications; and  
creating a model based on class definitions that transforms a measured set of features to  
an estimated classification, wherein said class definitions are optimized to satisfy  
35 specifications of a measurement system used to take said measurements.
15. The method of Claim 14, wherein said optimized classes comprise groups of  
measurements wherein similarity between measurements within a group is greater than  
similarity between groups.

WO 02/080767

PCT/US02/09114

16. The method of Claim 15, further comprising the step of:  
calculating weights,  $w$ , for said spectral measurements according to:
- $$W = (X^T X)^{-1} X^T y$$
- 5 where  $X$  represents a matrix of spectral measurements, and  $y$  represents a  
reference value of said target analyte concentration for each measurement.
17. The method of Claim 16, wherein a vector of weights of spectral measurements  
within one of said groups comprises a regression vector for said group;
- 10 wherein said regression vector comprises a calibration model for said  
group.

+

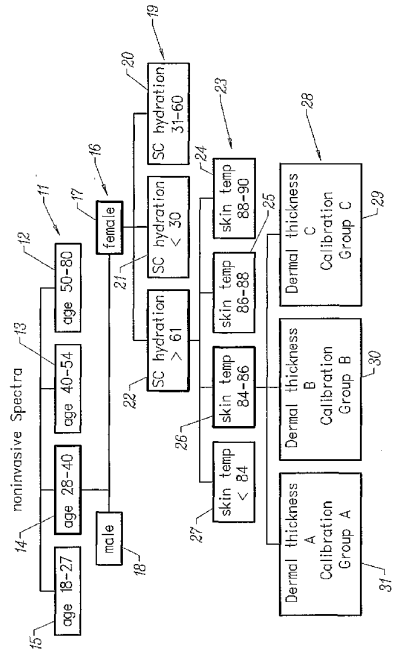


FIG. 1

+

+

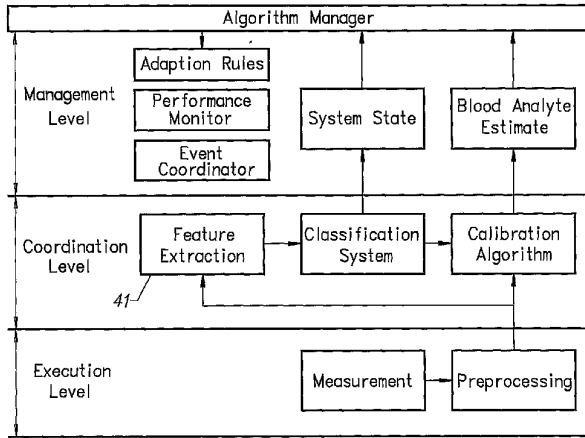


FIG. 2

+

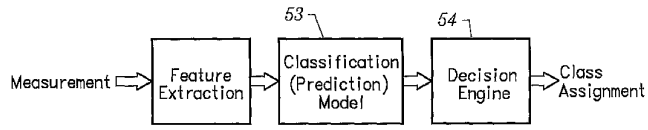


FIG. 3

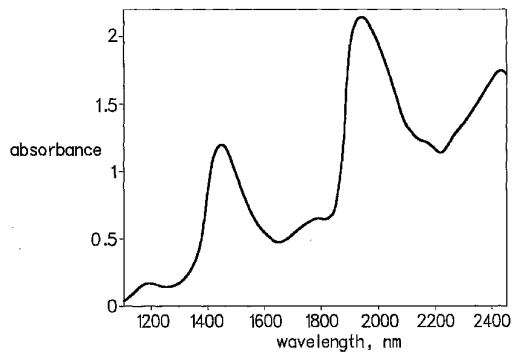


FIG. 4

WO 02/080767

PCT/US02/09114

+

4/9

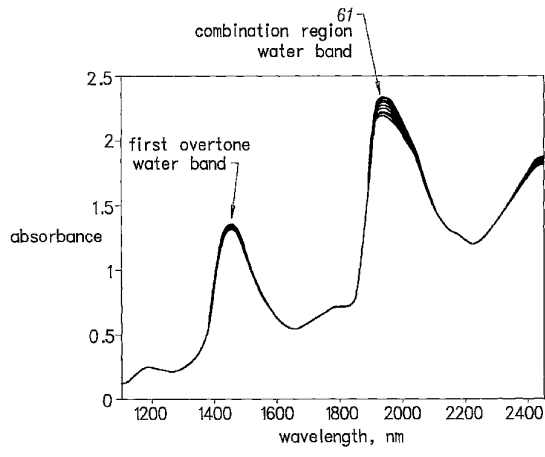


FIG. 5

WO 02/080767

PCT/US02/09114

+

5/9

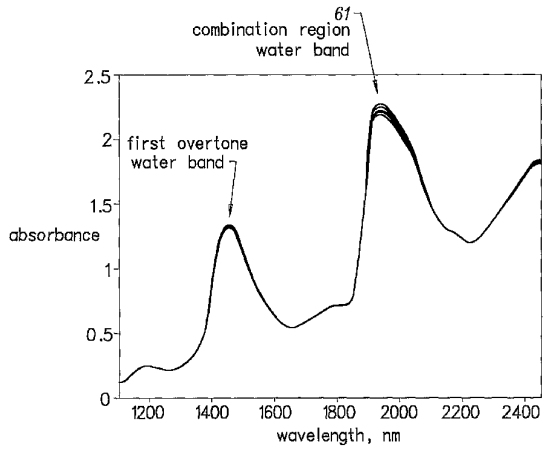


FIG. 6

WO 02/080767

PCT/US02/09114

+

6/9

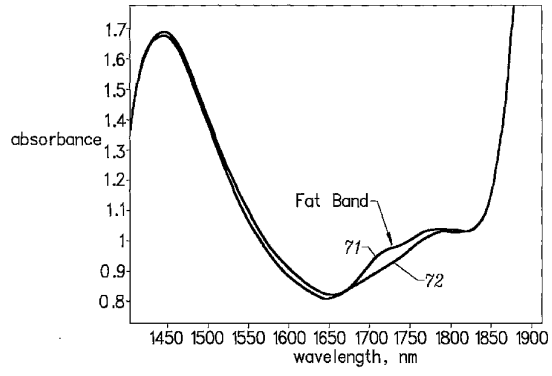


FIG. 7

+

7/9

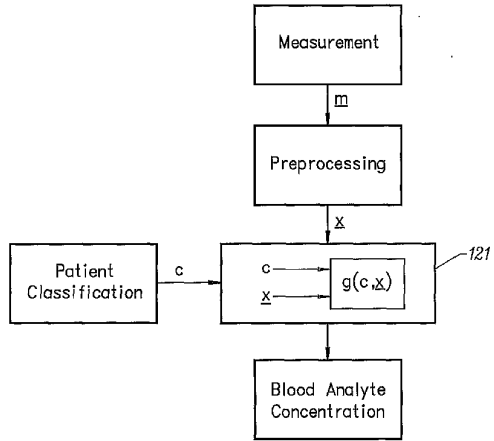


FIG. 8

+

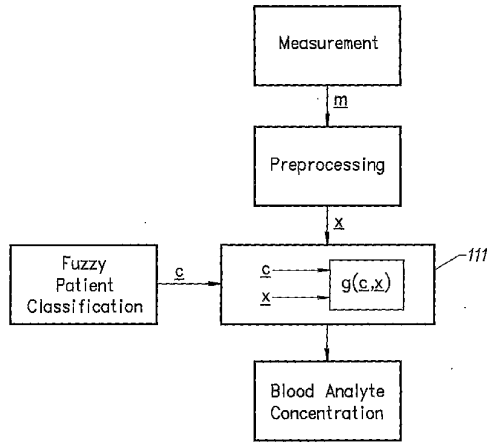


FIG. 9

+

9/9

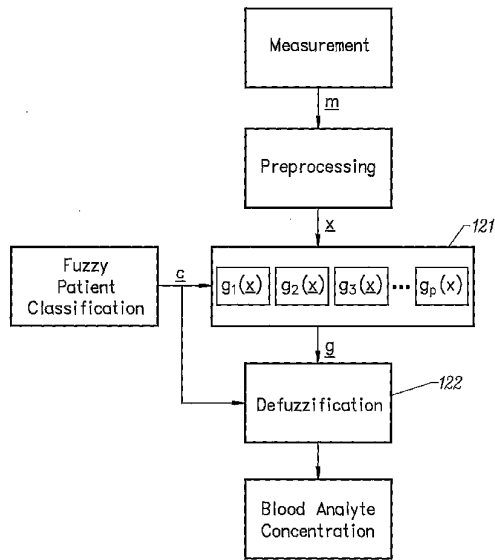


FIG. 10

【國際調查報告】

<b>INTERNATIONAL SEARCH REPORT</b>		International application No. PCT/US02/09114		
<b>A. CLASSIFICATION OF SUBJECT MATTER</b>				
IPC(7) : A61B 5/00 US CL : 609/322 According to International Patent Classification (IPC) or to both national classification and IPC				
<b>B. FIELDS SEARCHED</b>				
Minimum documentation searched (classification system followed by classification symbols) U.S. : 609/322, 310, 323, 330, 336, 473, 128/920, 923, 924, 925; 356/402, 250/340, 339.01, 341.1				
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched				
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)				
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>				
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.		
A	US 5,553,616 A (HAM et al) 10 September 1996 (10.09.1996), column 6, line 26 - column 9, line 47.	1-17		
A	US 5,576,544 A (ROSENTHAL) 19 November 1996 (19.11.1996), column 2, line 43 - column 3, line 13.	1-17		
A	US 5,725,480 A (OOSTA et al) 10 March 1998 (10.03.1998), column 1, line 59 - column 2, line 17.	1-17		
A	US 5,798,526 A (SHENK et al) 25 August 1998 (25.08.1998), column 2, lines 1 - 26.	1-17		
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.				
* Special categories of cited documents: <table border="0" style="width: 100%;"> <tr> <td style="width: 50%;">         "A" document defining the general state of the art which is not considered to be of particular relevance          "B" earlier application or patent published on or after the international filing date          "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)          "O" document referring to an oral disclosure, use, exhibition or other means          "P" document published prior to the international filing date but later than the priority date claimed       </td> <td style="width: 50%;">         "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention          "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone          "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art          "Z" document member of the same patent family       </td> </tr> </table>			"A" document defining the general state of the art which is not considered to be of particular relevance "B" earlier application or patent published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family
"A" document defining the general state of the art which is not considered to be of particular relevance "B" earlier application or patent published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family			
Date of the actual completion of the international search 31 May 2002 (31.05.2002)		Date of mailing of the international search report 06 AUG 2002		
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230		Authorized officer Eric F. Wingfield Telephone No. 703/308-0858		

## フロントページの続き

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN, TD,TG),AE,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,CA,CH,CN,CR,CU,CZ,DE,DK,DM,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID, IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,NO,NZ,PL,PT,RO,RU,SD,SE,SG,SI,S K,SL,TJ,TM,TR,TT,TZ,UA,UG,UZ,VN,YU,ZA,ZW

(72)発明者 ルチティ ティモシー エル

米国 85233 アリゾナ州 ギルバート ウェスト シー ヘイズ ドライブ 1501

(72)発明者 セナディル スレッシュ

米国 85283 アリゾナ州 テンペ アプト 207 エス キレーン ロード 7250

Fターム(参考) 2G059 AA01 BB13 EE01 EE12 FF06 FF08 HH01 MM01 MM02 MM14

4C038 KK00 KL05 KL07 KM03 KX02

专利名称(译)	构建无创血液样本预测的局部校正模型的多步法		
公开(公告)号	<a href="#">JP2004528083A</a>	公开(公告)日	2004-09-16
申请号	JP2002578807	申请日	2002-03-22
申请(专利权)人(译)	Forsythensis医疗墨		
[标]发明人	ブランクトーマスビー モンフレステファンエル ルチティティモシーエル セナデイルスレッシュ		
发明人	ブランクトーマスビー モンフレステファンエル ルチティティモシーエル セナデイルスレッシュ		
IPC分类号	G01N21/35 A61B5/00 A61B5/107 A61B5/145 A61B5/1455 A61B5/1495 G01N21/27 G01N21/47 G01N21/49 G06F19/00		
CPC分类号	A61B5/1075 A61B5/14532 A61B5/1455 A61B5/1495 A61B5/7264 A61B5/7267 A61B2560/0223 A61B2560/0233 G01N21/274 G01N21/359 G01N21/4785 G01N21/49 G16H50/20 Y10S128/92		
FI分类号	A61B5/14.310 G01N21/35.Z		
F-TERM分类号	2G059/AA01 2G059/BB13 2G059/EE01 2G059/EE12 2G059/FF06 2G059/FF08 2G059/HH01 2G059 /MM01 2G059/MM02 2G059/MM14 4C038/KK00 4C038/KL05 4C038/KL07 4C038/KM03 4C038/KX02		
优先权	09/825687 2001-04-03 US		
其他公开文献	JP4216077B2		
外部链接	<a href="#">Espacenet</a>		

摘要(译)

一种用于对组织吸收光谱进行分类的多步骤方法，该方法可减少由于干扰物质在光谱中的协方差，样品的不均匀性，状态的变异性及结构的变异性而导致的样品光谱的变异性。提供一种将校准和样品光谱定位到所使用的本地组的方法。非侵入性血液分析物预测中的分类和校准的多步骤方法通过限制共变频谱干扰物将预测误差降至最低。根据受试者的人口统计学信息和使用诸如体内近红外光谱测量之类的设备进行的皮肤测量，对组织样本进行分类。多阶段智能模式分类过程将光谱数据组织成簇，在组织特征方面具有高度一致性。在每个阶段，使用对象的人口统计数据，光谱测量信息和其他适用于建立组织分类的设备测量值，将不断提高分类的准确性。这种用于校准的多阶段分类方案利用了多元统计方法和利用频谱特征的多阶段分类。在多步骤分类中使用的变量可以是皮肤表面水合作用，皮肤表面温度，组织水合作用，具有近红外脂肪带的真皮的光学相对厚度的评估。使用沿感兴趣波长部分的近红外光谱信号评估所有组织参数。

$$B = (A^T V A)^{-1} A^T V Y$$