

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2007-199787
(P2007-199787A)

(43) 公開日 平成19年8月9日(2007.8.9)

(51) Int. Cl.	F I	テーマコード (参考)
G06N 3/00 (2006.01)	G06N 3/00 560A	4C117
G06N 3/08 (2006.01)	G06N 3/08 Q	
A61B 5/00 (2006.01)	A61B 5/00 G	

審査請求 未請求 請求項の数 16 O L (全 21 頁)

(21) 出願番号 特願2006-14396 (P2006-14396)
(22) 出願日 平成18年1月23日 (2006.1.23)

(71) 出願人 304020177
国立大学法人山口大学
山口県山口市吉田1677-1
(71) 出願人 504174180
国立大学法人高知大学
高知県高知市曙町二丁目5番1号
(72) 発明者 市原 清志
山口県宇部市南小串1-1-1 山口大学
医学部
(72) 発明者 片岡 浩巳
高知県南国市岡豊町小蓮 高知大学 医学部
Fターム(参考) 4C117 XA05 XA07 XB09 XJ34 XJ38 XJ48

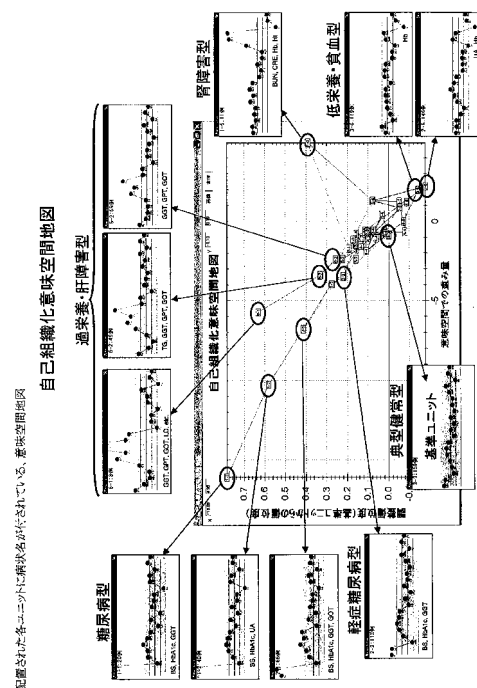
(54) 【発明の名称】 意味空間を表示した自己組織化マップ

(57) 【要約】 (修正有)

【課題】自己組織化マップ法の特徴地図における欠点であった、その空間配置の非一貫性の問題を克服し、再現性のあるクラスタリングを提供する。

【解決手段】自己組織化マップ法の特徴地図における各ユニットを、参照ベクトルの基線からの偏位量をY軸に、参照ベクトルの意味空間での重み量をX軸にとり、2次的に配置することにより、各ユニットは、再現性のある配置で分離して表示することができる。意味空間での重み量の算出の1例として、病態診断用の複数の検査項目の場合、このうち半分の項目を(1)糖尿病関連、過栄養、肝障害関連、の意味空間に配分し、残り半分の項目を(2)炎症関連、低栄養・腎機能関連、貧血関連の意味空間に配分し、この2つの意味空間でのベクトルの「重み量」を算出することにより、複数の意味が同一方の意味空間に配分されても、例えば、糖尿病関連のユニットと、肝障害関連のユニットとは分離された配置で表示される。

【選択図】 図5



【特許請求の範囲】

【請求項 1】

複数個の属性よりなる参照ベクトルを持つ複数のユニットの中から、前記参照ベクトルと同じ個数の属性よりなる入力ベクトルに最も類似した参照ベクトルを持つ勝者ユニットを探索し、学習回数に応じて順次狭められる前記勝者ユニットの近傍領域に含まれる各ユニットの参照ベクトルを、前記学習回数に応じて順次小さくされる学習係数と前記入力ベクトルとに基づいて更新し、前記学習回数が予め設定された回数に達するまで、前記勝者ユニットの探索及び前記参照ベクトルの更新を繰り返すことにより、特徴地図を作成する手順を含む自己組織化マップ法において、前記特徴地図の各ユニットの参照ベクトルについて、基線からの偏位量を算出し、参照ベクトルの属性について、1又は2以上の意味空間を設定して、前記参照ベクトルの属性の一部あるいは全部を、その設定した意味空間に配分し、前記特徴地図の各ユニットの参照ベクトルについて、その意味空間での重み量を算出し、参照ベクトルの基線からの偏位量を一方の軸に、及び、参照ベクトルの意味空間での重み量を他方の軸に、前記特徴地図の各ユニットを配置して、意味空間地図を作成する手順を含むことを特徴とする自己組織化マップ法。

10

【請求項 2】

特徴地図を作成する手順において、前記入力ベクトルの各属性の数値が変数変換により正規分布化されている、請求項 1 の自己組織化マップ法。

【請求項 3】

特徴地図を作成する手順において、前記入力ベクトルの各属性の数値が基準化されている、請求項 1 ~ 請求項 2 のいずれか 1 項の自己組織化マップ法。

20

【請求項 4】

意味空間地図を作成する手順において、基線からの偏位量が基線から上下のいずれの偏位であるかが加味された、調整偏位量である請求項 1 ~ 請求項 3 のいずれか 1 項の自己組織化マップ法。

【請求項 5】

意味空間地図を作成する手順において、設定する意味空間が 2 つである請求項 1 ~ 請求項 4 のいずれか 1 項の自己組織化マップ法。

【請求項 6】

意味空間地図を作成する手順において、意味空間地図に前記各ユニット間の類似性を結合線により表示する請求項 1 ~ 請求項 5 のいずれか 1 項の自己組織化マップ法。

30

【請求項 7】

特徴地図を作成する手順において、前記属性が病態診断における検査項目であり、作成する意味空間地図が病態診断用である請求項 1 ~ 請求項 6 のいずれか 1 項の自己組織化マップ法。

【請求項 8】

前記意味空間地図の前記ユニットに、対応する病状名を表示する、作成する意味空間地図が病態診断用である請求項 7 に記載の自己組織化マップ法。

【請求項 9】

前記意味空間地図に、時系列遷移により病態の悪化か治癒の方向かを視覚的に表示する請求項 7 ~ 請求項 8 のいずれか 1 項に記載の自己組織化マップ法。

40

【請求項 10】

前記意味空間地図に更に、生活習慣特性を付記する請求項 7 ~ 請求項 9 のいずれか 1 項に記載の自己組織化マップ法。

【請求項 11】

自己組織化マップ法の特徴地図の各ユニットについて、参照ベクトルの基線からの偏位量を一方の軸に、及び参照ベクトルの意味空間での重み量を他方の軸に、各ユニットを配置してなる意味空間地図であって、請求項 1 ~ 請求項 10 のいずれか 1 項に記載の自己組織化マップ法により作成された意味空間地図。

【請求項 12】

50

複数個の属性よりなるベクトルを入力する手段、前記入力ベクトルに最も類似した参照ベクトルを持つ勝者ユニットを探索する手段、学習回数に応じて順次狭められる前記勝者ユニットの近傍領域に含まれる各ユニットの参照ベクトルを、前記学習回数に応じて順次小さくされる学習係数と前記入力ベクトルとに基づいて更新する手段、前記学習回数が予め設定された回数に達するまで、前記勝者ユニットの探索及び前記参照ベクトルの更新を繰り返す手段、自己組織化マップの特徴地図を作成する手段、前記特徴地図の各ユニットの参照ベクトルについて、基線からの偏位量を算出する手段、前記属性の一部あるいは全部を意味空間に配分し、その意味空間に配分された属性から、前記参照ベクトルの意味空間での重み量を算出する手段、参照ベクトルの基線からの偏位量を一方の軸に、及び参照ベクトルの意味空間での重み量を他方の軸に、前記特徴地図の各ユニットを配置して、2次元的に表示する手段を具備することを特徴とする自己組織化マップ表示装置。 10

【請求項 13】

各入力ベクトルの属性の数値が正規分布化されるように変数変換する手段を有する、請求項 12 の自己組織化マップ表示装置。

【請求項 14】

各入力ベクトルの属性の数値を基準化する手段を有する、請求項 12 ~ 請求項 13 のいずれか 1 項の自己組織化マップ表示装置。

【請求項 15】

コンピュータを自己組織化マップ表示装置として機能させるためのプログラムであって、請求項 1 ~ 請求項 10 のいずれか 1 項の自己組織化マップ法の特徴地図を作成する手順、及び意味空間地図を作成する手順を実行するプログラム。 20

【請求項 16】

請求項 15 のプログラムを記録したコンピュータ読み取り可能な記録媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、自己組織化マップ法、自己組織化マップ、自己組織化マップ表示装置、その表示装置を作動させるプログラム、そのプログラムを記録した記録媒体に関する。自己組織化マップの、病態診断用への応用又はその他への応用に関する。

【背景技術】

【0002】

自己組織化マップ法

自己組織化マップ法は、非常に強力なクラスタリング能力を持った教師無し学習アルゴリズムである。多数のデータを解析する手段として優れている。

【0003】

自己組織化マップ法の特徴地図の作り方

参照ベクトルを持つ複数のユニットの中から、複数の項目に関するデータを利用した入力ベクトルに最も類似した参照ベクトルを持つ勝者ユニットを探索するステップ、学習回数に応じて順次狭められる前記勝者ユニットの近傍領域に含まれる各ユニットの参照ベクトルを前記学習回数に応じて順次小さくされる学習係数と前記入力ベクトルとに基づいて更新するステップ、前記学習回数が予め設定された回数に達するまで前記勝者ユニットの探索及び前記参照ベクトルの更新を繰り返すステップ、そして前記各ユニットを2次元的に表示するステップにより自己組織化マップの特徴地図が作られる。 40

【0004】

各ユニットの配置

このような特徴地図には、各ユニットは互いに類似したものが近接的に配置されるように表示されている。

【0005】

自己組織化マップ法の利用分野

自己組織化マップは、化学分析分野、定量化学分析の分野、ガラス分類問題の分野、人 50

事適正配置問題の分野、巡回セールスマン問題の分野、プリント基板電子部品配置問題などに利用される（非特許文献1）。

【0006】

そして、複数の項目に関するデータについてその各項目を基準値により正規化して健康診断用に応用することも提案されている（特許文献1）。

【0007】

従来自己組織化マップ法の問題点

しかしながら、従来自己組織化マップ法による特徴地図には、各ユニットの配置が同じデータセットを行っても処理を行うごとに異なる結果が得られてしまう問題があった（図6左図）。これは、学習の初期値がゼロあるいは乱数から出発して、多数の試行によって自己組織化現象を発生させる原理に基づくためである。特許文献1の方法でも、特徴地図を作り直すごとに位置的な情報が異なり、観測者に混乱を与える可能性があった。また、自己組織化マップ法による分類結果を、さらにクラスター分析で樹形図の形にクラス展開表示をしても、クラス間の基準が存在しないため、樹形図の解釈が困難であった（図8下図）。

10

【0008】

正規分布化

多数項目のデータ配列は、生命科学分野で取り扱われる計測値の場合、多様な分布型を取ることが多い。例えば、医学分野で取り扱われる検査値の分布型は、正規分布のほかに対数正規分布と、その中間にある平方根正規分布、3乗根正規分布などさまざまであるが、一般にべき乗値 $k(p)$ と変換原点 $a(p)$ を指定して、べき乗変換を行うと、その値は近似的に正規分布となる（Box-Cox方式変法：非特許文献2）。なお、同方式では、べき乗値 $k(p) = 0$ のときは対数変換を用いる。

20

【特許文献1】特開2003-263502号公報

【非特許文献1】徳高平蔵・岸田悟・藤村喜久郎著「自己組織化マップの応用（多次元情報の2次元可視化）」1999年2月20日海文堂出版株式会社発行

【非特許文献2】Ichi hara, K. and Kawai, T: Determination of reference intervals for 13 plasma proteins based on IFCC international reference preparation (CRM470) and NCCLS proposed guideline (C28-P, 1992): trial to select reference individuals by results of screening tests and application of maximal likelihood method. J Clin Lab Anal. 10 (2): 110-7, 1996.

30

【発明の開示】

【発明が解決しようとする課題】

【0009】

本発明は、自己組織化マップ法の特徴地図における欠点であった、その空間配置の非一貫性の問題を克服し、再現性のあるクラスタリングを提供することを目的とする。また、本発明は利用者に対し自己組織化マップ法の分類結果を自動的に系統樹の形で意味空間に展開して提供することを目的としている。さらに本発明は、医療分野の病態解析ツールを提供することを目的としている。

40

【0010】

本発明は、医学関連では下記のシステムなどに適用可能である。

- (1) 診断支援システムの病態解析および表示（病態の分類と進行度の系統樹表示）
- (2) 検診システムの健康度表示（健常から疾患への移行のどの段階にあるかを系統樹上で提示）
- (3) 検診システムにおいて、生活習慣の偏りが関与する病態の系統的表示
- (4) 時系列パターンの変化に関する病態ステージの意味的表示システム（病態パターン

50

の遷移の把握や、病原菌の耐性化の程度とその遷移課程過程の提示)

本発明は、医学分野以外にも幅広く適用可能なアルゴリズムである。

【課題を解決するための手段】

【0011】

本明細書では、自己組織化マップ法の特徴地図の各ユニットについて、参照ベクトルの基線からの偏位量を一方の軸に、及び参照ベクトルの意味空間での重み量を他方の軸に、各ユニットを配置してなる地図を、「意味空間地図」という。

【0012】

本発明の自己組織化マップ法は、特徴地図を作成する手順、及び参照ベクトルの基線からの偏位量を一方の軸に、及び参照ベクトルの意味空間での重み量を他方の軸に、前記特徴地図の各ユニットを配置して、意味空間地図を作成する手順を含むことを特徴とする。

10

【0013】

前記自己組織化マップ法の特徴地図は、従来技術により、入力ベクトルに最も類似した参照ベクトルを持つ勝者ユニットを探索し、参照ベクトルを更新し、前記勝者ユニットの探索及び前記参照ベクトルの更新を繰り返し、各ユニットを2次元的に配置することにより作成できる。

【0014】

本発明の自己組織化マップ法は、特徴地図を作成する手順において、前記入力ベクトルの各属性の数値を変数変換により正規分布化することができる。

【0015】

本発明の自己組織化マップ法は、特徴地図を作成する手順において、前記入力ベクトルの各属性の数値を基準化することができる。

20

【0016】

本発明の自己組織化マップ法は、意味空間地図を作成する手順において、基線からの偏位量が、基線から上下のいずれの偏位であるかが加味された調整偏位量であることができる。

【0017】

本発明の自己組織化マップ法は、意味空間地図を作成する手順において、設定する意味空間を2つに取ることができる。

【0018】

本発明の自己組織化マップ法は、意味空間地図を作成する手順において、前記各ユニット間の類似性を結合線により表示することができる。

30

【0019】

本発明の自己組織化マップ法は、特徴地図を作成する手順において、前記属性を病態検査項目とすることにより、病態診断用の意味空間地図を作成することができる。本明細書においては、「病態診断」という用語は、健常者に対して行われる健康診断も含む。本発明の自己組織化マップ法は、この病態診断用の意味空間地図に、さらに病態名を表示することができる。本発明の自己組織化マップ法は、この病態診断用の意味空間地図に、時系列遷移により病態の悪化が治癒の方向かを視覚的に表示することができる。本発明の自己組織化マップ法は、この病態診断用の意味空間地図に、生活習慣特性を付記することができる。これにより生活習慣と病態との関連性を直接的に示すことができる(図10)。

40

【0020】

本発明の意味空間地図は、参照ベクトルの基線からの偏位量を一方の軸に、及び参照ベクトルの意味空間での重み量を他方の軸に、前記特徴地図の各ユニットを配置することを特徴とする。

【0021】

本発明の、自己組織化マップ表示装置は、複数個の属性よりなるベクトルを入力する手段、前記入力ベクトルに最も類似した参照ベクトルを持つ勝者ユニットを探索する手段、学習回数に応じて順次狭められる前記勝者ユニットの近傍領域に含まれる各ユニットの参照ベクトルを、前記学習回数に応じて順次小さくされる学習係数と前記入力ベクトルとに

50

基づいて、更新する手段、前記学習回数が予め設定された回数に達するまで、前記勝者ユニットの探索及び前記参照ベクトルの更新を繰り返す手段、自己組織化マップの特徴地図を作成する手段、前記特徴地図の各ユニットの参照ベクトルについて基線からの偏位量を算出する手段、前記属性の一部あるいは全部を意味空間に配分し、その意味空間に配分された属性から前記参照ベクトルの意味空間での重み量を算出する手段、参照ベクトルの基線からの偏位量を一方の軸に、及び参照ベクトルの意味空間での重み量を他方の軸に、前記特徴地図の各ユニットを配置して、2次元的に表示する手段を具備することを特徴とする。

【0022】

本発明の自己組織化マップ表示装置は、各入力ベクトルが各属性の数値に対して正規分布化されるように変数変換する手段を有することができる。

10

【0023】

本発明の自己組織化マップ表示装置は、各入力ベクトルの属性の数値を基準化する手段を有することができる。

【0024】

本発明のプログラムは、コンピュータを自己組織化マップ表示装置として機能させるためのプログラムであって、本発明の自己組織化マップ法の特徴地図を作成する手順、及び意味空間地図を作成する手順を実行するプログラムである。

【0025】

本発明のコンピュータ読み取り可能な記録媒体は、本発明の自己組織化マップ法の特徴地図を作成する手順、及び意味空間地図を作成する手順を実行するプログラムを記録したものである。

20

【発明の効果】**【0026】**

本発明による自己組織化マップ法による意味空間地図では、各ユニットが参照ベクトルの基線からの偏位量が一方の軸（Y軸）に、及び参照ベクトルの意味空間での重み量を他方の軸（X軸）に配置されているので、データセットを何回行っても、各ユニットが同じ位置に再現され（図6右図）、従来の自己組織化マップ法による特徴地図の問題点（図6左図）を解消する。

【0027】

本発明による自己組織化マップ法による意味空間地図（図8上図）では、基準ユニット（参照ベクトルの基線からの偏位量が最小のユニット）を基準として再配置され、他のユニットが基準ユニット（標準）からどの程度乖離しているかの判断が容易となる。これに対し、従来のクラスター分析を使ってデンドログラムの形式で再配置（図8下図）した場合には、枝の配置に意味を持たせることができない。このため、基準ユニットも他のユニットと混在しており、分類結果の解釈が困難である。

30

【0028】

本発明では、与えられた多変量情報の全体像を分類し、その系列を可視化するため、利用者が直感的に分類された情報の意味を理解でき、かつ自己組織化マップ法による分類操作を繰り返してもその特徴地図の意味空間上への写像（系統樹）は変更されることが無い

40

【0029】

本発明の自己組織化マップ法は、前記属性の数値を正規分布化する態様においては、基線からの偏位量の数値が、又は意味空間での重み量の数値が、極端値の影響を受なくなる。

【0030】

本発明の自己組織化マップ法は、前記属性の数値が基準化されている態様においては、基準化するときの中央値を、「参照ベクトルの基線」とすることができる。

【0031】

50

本発明の自己組織化マップ法は、参照ベクトルの「基線からの偏位量」として、「基線から上下のいずれの偏位であるかが加味された調整偏位量」を使用する態様において、各ユニットの配置位置が更に差別化ができる。

【0032】

本発明の自己組織化マップ法は、設定する意味空間が2つである態様においては、前記各ユニットの配置がX軸のプラス側とマイナス側の両方に展開できるので、2次元表示の意味空間地図にとって好ましい。ベクトルの複数の属性が3以上の意味に対応する場合、設定する意味空間が2つであると、X軸の展開は、原理的には3以上の意味に対応できないはずであるが、意外なことにこの態様の意味空間地図のX軸の展開は、各ユニットが3つ以上の意味において分離されて表示される。この場合、意味空間地図は3つ以上の意味に分離されているが、その分離の程度はユニット間の類似性を表すものではない。

10

【0033】

本発明の自己組織化マップ法は、意味空間地図に前記各ユニット間の類似性を結合線により表示する態様においては、前記各ユニット間の類似性を意味空間地図から把握することができる。

【0034】

本発明の意味空間地図は、病態診断用に使用する態様では、各ユニットに対応する病態が同じ位置に再現されるので病態診断用に好適である。この場合、同じ病態の軽重のユニットが近接して配置されるので、病態診断用に応用する場合、病態の重症度と相互関連を明らかにできるので好都合である。これまでにない細やかな病型分類（健常または基準病態からの乖離の判断）が可能となる。さらに入力ベクトルの属性として、個人の生活習慣特性を臨床検査の測定値と平行して配置することにより、検査の異常と生活習慣との関連性を直接的に受診者に訴え、生活習慣病改善に向けた努力を促すことができるので、予防医学的見地から極めて好都合である（図10）。本発明の意味空間地図は、十分なデータ数に基づいて作成した場合、データを随時追加して再実行してもその分類結果の配置は基本的には変化しない。このため、同じ患者の検査結果をその上に随時表示することで、病態の変化を明確に把握できる（図9）。生活習慣に関するデータを付記すると、病態と生活習慣に関する相互関連を明らかにできるので好都合である。

20

【0035】

本発明の自己組織化マップ表示装置は、本発明の意味空間地図を自動的に作成することができる。

30

【0036】

本発明のプログラムは、コンピュータを、本発明の自己組織化マップ表示装置として機能させる。本発明のプログラムは、コンピュータ読み取り可能な記録媒体に記録することができる。

【発明を実施するための最良の形態】

【0037】

原データ配列 $x(p, i)$

本明細書では、P個の属性を持つN個のデータ配列を (p, i) と表記する。本明細書では原データ配列を $x(p, i)$ で表す。医学分野で取り扱われる検査項目には、血糖(BS)、HbA1c、HDLコレステロール(HDL)、総コレステロール(TC)、トリグリセライド(TG)、GT(GGT)、ALT(GPT)、AST(GOT)、乳酸脱水素酵素(LD)、アルカリフォスファターゼ(ALP)、血小板数(PLT)、白血球数(WBC)、総蛋白(TP)、アルブミン(ALB)、尿酸(UA)、尿素窒素(BUN)、クレアチニン(CRE)、ヘモグロビン(Hb)、ヘマトクリット(Ht)、赤血球数(RBC)などがある。これら例示の20項目の検査データの配列は、20個の属性を持つ原データ配列である。

40

【0038】

正規分布化値： $X(p, i)$

P個の属性を持つ原データ配列 $x(p, i)$ の各属性値は、生命科学分野で取り扱われ

50

る計測値の場合、多様な分布型を取ることが多い。このため、対数正規分布など、正規分布からの偏りが強い属性が存在すると、如何にその属性値を基準化しても、分布の裾の値が分類に強く影響する。この影響を除くには、属性毎にその分布型を調べ、それを正規分布に変換する。

【0039】

例えば、医学分野で取り扱われる検査項目の測定値の分布型は、正規分布のほかに対数正規分布と、その中間にある平方根正規分布、3乗根正規分布などさまざまであるが、一般にべき乗値 $k(p)$ と変換原点 $a(p)$ を指定して、 P 個の属性を持つ原データ配列 $x(p, i)$ の各属性値に、べき乗変換を行うと、その値 $X(p, i)$ の分布は、近似的に正規分布となる (Box - Cox 方式変法: 非特許文献 2)。なお、同方式では、べき乗値 $k(p) = 0$ のときは対数変換を用いる。

10

【0040】

【数1】

$k(p) \neq 0$ のとき、一般べき乗変換

$$X(p, i) = \frac{(x(p, i) - a(p))^{k(p)} - 1}{k(p)}$$

$k(p) = 0$ のとき、対数変換

20

$$X(p, i) = \log(x(p, i) - a)$$

図4は、血糖 (BS)、総コレステロール (TC)、トリグリセライド (TG)、ALT (GPT)、AST (GOT)、乳酸脱水素酵素 (LD)、アルカリフォスファターゼ (ALP)、血小板数 (PLT)、白血球数 (WBC) について、実際の分布図 (左図) とべき乗変換後の分布図 (右図) を示した図である。度数分布図に重ねて描いた曲線は、各検査値の平均値と標準偏差からもとめた理論正規分布曲線を示す。図4は、べき乗変換後の分布 (右図) が正規分布であることを示している。

30

【0041】

データの基準化値 $u(p, i)$

図3は、基準化データ作成部の基準化の概念を示した説明図である。上部から第 p 番目の属性の実際の分布図、正規分布化した分布図、そして基準化した基準範囲の求め方の原理図である。

【0042】

前項で正規分布化されたデータ配列 $X(p, i)$ の値を次のように変換して基準化値 $u(p, i)$ とする。

ここに、 $p = 1, 2, \dots, P,$
 $i = 1, 2, \dots, N,$

40

P は属性総数、 N はデータ総数である。

【0043】

基準範囲が存在しない数値情報の場合、データの基準化値 $u(p, i)$ は、正規分布化値 $X(p, i)$ とその平均値 $M(p)$ とその標準偏差 $SD(p)$ から次のように算出される。

【0044】

【数 2】

$$u(p, i) = \frac{X(p, i) - M(p)}{SD(p)}$$

医学分野の検査値の正常値範囲のように、その数値情報に基準範囲がある場合は、その下限値 $ll(p)$ 、上限値 $ul(p)$ のべき乗変換値、 $LL(p)$ 、 $UL(p)$ を用いて、次式で示す $M(p)$ 、 $SD(p)$ を、それぞれ、“平均値”相当値、“標準偏差”相当値として用いることができる。

【0045】

【数 3】

$$M(p) = (UL(p) + LL(p)) / 2$$

【0046】

【数 4】

$$SD(p) = (UL(p) - LL(p)) / 4$$

この数式 3 の $M(p)$ と、数式 4 の $SD(p)$ を使って、数式 2 により、基準化値 $u(p, i)$ を算出する。

【0047】

入力ベクトル $u(p)$

入力ベクトルは、 P 個の属性をもつベクトルである。入力ベクトルとして、データの基準化値 $u(p, i)$ の i 番面のデータ $u(p)$ を使用できる。前述したように、 p は属性の次元 ($p = 1, 2, \dots, P$) あらわす。

【0048】

入力ベクトルは、臨床検査値や波形のようなデータの場合は 1 次元の配列として取り扱うが、画像のようなデータの場合は 2 次元データ配列として利用できる。また、時間軸を加えた時系列検査値や立体画像のようなデータの場合は 3 次元データ配列として取り扱う。

【0049】

参照ベクトル $r(p, w)$

参照ベクトルは、ユニットごとに 1 個存在する、各ユニットを代表するベクトルである。参照ベクトルは、入力ベクトルと同じ数の要素、 P 個の要素を持つベクトルである。参照ベクトルを $r(p, w)$ と記載する。 p は属性の次元 ($p = 1, 2, \dots, P$)、 w は参照ユニット番号 ($w = 1, 2, \dots, W$) を表す。参照ベクトルは、2 次元あるいは多次元、または、6 角配列としてユニット間の近傍関係を持つ構造を有している。

【0050】

ベクトルの基線

本明細書において、ベクトルの基線とは、入力ベクトルが数式 2 により基準化されている場合、ベクトルの各属性の数値が 0 であることを意味している。入力ベクトルが数式 2 により、基準化されていない場合、ベクトルの各属性の数値の平均値 $M(p)$ または、数式 3 の $M(p)$ が、ベクトルの基線に相当する。

【0051】

10

20

30

40

50

参照ベクトルの基線からの偏位量

本明細書において、「参照ベクトルの基線からの偏位量」とは、参照ベクトルの各属性の数値が基線からどの程度偏位しているかを示す量である。基準値化されている入力ベクトルが使用された場合、参照ベクトルの各属性の数値は基線からの偏位量をあらわしている。「参照ベクトルの基線からの偏位量」としては、各要素の基線からの偏差平方和の平方根（数式 5）を採用することができるが、これに限定されない。

【 0 0 5 2 】

【 数 5 】

$$\text{基線からの偏位度 } D(w) = \sqrt{\sum_{p=1}^P r(p, w)^2}$$

10

基準ユニット

本明細書において、参照ベクトルの偏位度 $D(w)$ が最も小さいユニットを「基準ユニット」という。「参照ベクトルの基線からの偏位量」として、基準ユニットの参照ベクトル $r(p, w_0)$ の偏位度を 0 として位置づけて、偏位度 $D(w)$ を計算し直して使用しても良い。

20

【 0 0 5 3 】

調整偏位量

本明細書において、「参照ベクトルの基線からの偏位量」としては、基線からの「偏位量」が、基線から上下のいずれの偏位であるかが加味された、「調整偏位量」であることができる。基線から偏位の符号付き総和が、プラスかマイナスかで、「偏位量」にプラスかマイナス符号を付けて、「参照ベクトルの基線からの調整偏位量」とすることができる。この「調整偏位量」として、数式 6 の調整偏位度 $D'(w)$ を使用することができる。

【 0 0 5 4 】

このような「調整偏位量」を使用することにより、本発明の意味空間地図において、各ユニットは更に明確に分離されて表示されるようになる。

30

【 0 0 5 5 】

【 数 6 】

$$\sum_{p=1}^P r(p, w) > 0 \text{ のとき、調整偏位度 } D'(w) = \sqrt{\sum_{p=1}^P r(p, w)^2}$$

$$\sum_{p=1}^P r(p, w) < 0 \text{ のとき、調整偏位度 } D'(w) = -\sqrt{\sum_{p=1}^P r(p, w)^2}$$

40

本明細書において、「参照ベクトルの基線からの調整偏位量」としては、調整偏位度 $D''(w)$ （数式 7）を使用することもできる。

【 0 0 5 6 】

「参照ベクトルの基準ユニットからの調整偏位度 $D''(w)$ 」とは、参照ベクトルの基線からの偏位度 $D(w)$ を、各要素の基線からの偏差平方和の平方根として計算し、それ

50

を $D(w)$ が最も小さい基準ユニット (要素番号 w_0) の基線からの偏位度 $D(w_0)$ から差し引いた値を使用するものである。ここで、基線から偏位の符号付き総和がプラスかマイナスかで、この値にプラスかマイナスの符号を付けて次のように表す。

【0057】

【数7】

$$\sum_{p=1}^P r(p, w) \geq 0 \text{ のとき}$$

$$\text{調整偏位度 } D''(w) = \sqrt{\sum_{p=1}^P r(p, w)^2} - \sqrt{\sum_{p=1}^P r(p, w_0)^2}$$

$$\sum_{p=1}^P r(p, w) < 0 \text{ のとき}$$

$$\text{調整偏位度 } D''(w) = - \left(\sqrt{\sum_{p=1}^P r(p, w)^2} - \sqrt{\sum_{p=1}^P r(p, w_0)^2} \right)$$

10

20

参照ベクトルの偏位量の表す意味

「参照ベクトルの偏位量」は、本発明の意味空間地図において、例えば本発明の健康診断用の意味空間地図において、健常状態者からの乖離度、またはある疾患の重症度を示すといえる。これにより、視覚的に健常からの乖離の度合いや疾患の相互関係 (相互移行過程) を知ることができる。

【0058】

意味空間

P個の属性を持つデータ配列において、P個の項目のうちのいくつかを1つ又は2つ以上の特定の意味で分類できる。例えば、医学分野で取り扱われる検査項目には、血糖 (BS)、HbA1c、HDL、総コレステロール (TC)、トリグリセライド (TG)、GGT、GPT、GOT、乳酸脱水素酵素 (LD)、アルカリフォスファターゼ (ALP)、血小板 (PLT)、白血球数 (WBC)、TP、ALB、UA、BUN、CRE、Hb、Ht、RBC などがある。その検査項目が、糖尿病関連、過栄養・肝障害関連、炎症関連、低栄養・腎機能関連、貧血関連などに分類できる。血糖 (BS)、HbA1c は、糖尿病関連である。HDL、総コレステロール (TC)、トリグリセライド (TG) は、過栄養である。GGT、GPT、GOT、乳酸脱水素酵素 (LD)、アルカリフォスファターゼ (ALP) は、過栄養または肝障害関連である。血小板 (PLT)、白血球数 (WBC)、TP は、炎症関連である。ALB、UA、BUN、CRE は、低栄養・腎機能関連である。Hb、Ht、RBC は、貧血関連である。本件明細書では、データの項目を、このような意味に分類する場合、意味空間を設定するという。そして、本件明細書では、データの項目をこのような意味に分類することを、項目を意味空間に配分するという。

30

40

【0059】

本発明では、参照ベクトルの属性について意味空間を設定し、その意味空間に配分する。

【0060】

意味空間での重み量

本発明では、参照ベクトルについて、「意味空間での重み量」を算出する。本発明の自己組織化マップとして表示する場合、X軸に盛る量である。本明細書において、「意味空間での重み量」とは、参照ベクトルについて、ある意味空間における重みを評価する量で

50

ある。その意味空間に配分されたすべての属性についての数値を評価する量である。その意味空間に配分された属性の数値の自乗和の平方根などを「意味空間での重み量」とすることができるが、これに限定されない。また更にその意味空間に独自の係数を掛けても良い。

【0061】

事前にべき乗変換で正規分布化している場合、「意味空間での重み量」は正規分布化していない場合と異なり、極端値の影響を受けにくくなる。

【0062】

意味空間の数

病態診断では、参照ベクトルについて、(1)糖尿病関連、(2)過栄養、(3)肝障害関連、(4)炎症関連、(5)低栄養・腎機能関連、(6)貧血関連の、6個の意味空間について、「意味空間での重み量」を算出することもできる。すなわち、その意味空間に配分された属性の数値の自乗和の平方根に独自の係数 $a_1 \sim a_6$ を掛けることもできる。

【0063】

しかしながら、2つの意味空間を設定することは2次元地図のX軸に表示するのに好ましい。病態診断では、20個の属性の参照ベクトルについて、このうち半分の10個の属性について(1)糖尿病関連、過栄養、肝障害関連、の意味空間に配分し、残り半分の(2)炎症関連、低栄養・腎機能関連、貧血関連の意味空間に配分し、この2つの意味空間での「重み量」を算出することができる。

【0064】

意味空間を2つに設定した場合の「重み量」

このように意味空間を、参照ベクトルの属性の配列の左右により、2つに設定した場合、参照ベクトルの「重み量」としては数式8の $L(w)$ で表すことができる。その計算の基本要素は、符号付き偏差平方和 (SSS : signed sum of square) である。

【0065】

【数8】

$$\text{符号付き偏差平方和 } SSS = \sum_{p=1}^P \left(p - \frac{P}{2} \right) \cdot r(p, w)^2$$

$$SSS \geq 0 \text{ の場合、「重み量」指数 } L(w) = \sqrt{SSS}$$

$$SSS < 0 \text{ の場合、「重み量」指数 } L(w) = -\sqrt{-SSS}$$

「重み量」指数

本発明の実施例の意味空間地図においては、「重み量」指数が、大きくマイナスの場合は糖尿病関連、大きくプラスの場合は、栄養障害・貧血関連を表し、その中間の場合は、その部位だけに異常があるか複数の系統の異常が存在することを表す。

【0066】

ユニット間の類似性の表示

前記各ユニット間の類似性の求め方は、全ユニット間の距離(ユークリッド距離など)を求め、それが最小となる上位のユニットを選んで、その距離(類似度)に応じて実線の太さ、あるいは色で各ユニット間の類似性を表示する。

【実施例1】

【0067】

10

20

30

40

50

自己組織化マップ作成および表示の処理装置

図 1 は、本発明の自己組織化マップの作成および表示の装置のハードウェア構成を示す。

【0068】

本発明の自己組織化マップの作成および表示の装置は、コンピュータで構成される。コンピュータは、ハードディスク 11 と、ROM・RAM 12 と、CPU 13 とで構成された処理装置と、操作部 14 と、表示部 15 と、入出力インタフェース部 16 とを備えている。尚、操作部 14 には、マウスなどのポインティングデバイスなどを取り付けるようにしてもよい。また、自己組織化マップの作成にデータフロー形（非ノイマン型）あるいはハードウェアによる自己組織化処理装置を用いても良い。

10

【0069】

図 2 は、自己組織化意味空間表示を行うための処理の流れを示した図である。

【0070】

本発明の装置は、正規分布化データ作成部、基準化データ作成部、自己組織化マップ作成部、意味空間地図作成部で構成される。本発明の自己組織化マップを作成するプログラムは、図 2 で示される処理の流れを実行する。本発明の自己組織化マップを作成するプログラムは記録媒体に保存されていても良い。

【0071】

正規分布化データ作成部は、実際の属性データの分布型を解析し、正規分布化に必要なパラメータ（べき乗値 $k(p)$ 、変換原点 $a(p)$ ）を選択し正規分布化データを作成する。また、自己組織化マップ特徴地図作成処理部は、変換後のデータを基準化後、自己組織化マップ法のアルゴリズムに従って自己組織化マップ特徴地図を作成する。意味空間地図作成部は、本発明のクラスタリングのアルゴリズムでその各ユニットを再配置して意味空間地図を作成する。さらに、意味空間地図作成部は、意味空間地図上に各ユニット間の類似性を表示する。そして、意味空間地図上に各ユニットに対応する病名などのクラスタリングされたグループの意味を表示する。

20

【0072】

ハードディスク 11 は、アプリケーションやデータを記憶する記憶装置であって、ハードディスク 11 には、本発明の自己組織化マップを作成するためのプログラムが記憶されている。

30

【0073】

ROM には、CPU などが最初に実行するコードや、OS をディスクからロードするためのコードなどが記憶されている。RAM には、CPU が作業を行うワークエリアなどがある。

【0074】

本発明の装置は、正規分布化データ作成部と基準化データ作成部、自己組織化マップ特徴地図作成部と意味空間地図作成部として機能する。

【0075】

正規分布化データ作成部

正規分布化データ作成部は、分布型判定部と分布型パラメータ修正部、そして、正規分布化データ変換部で構成している。分布型判定部は、入力データの属性の分布型を解析して、適切な変換パラメータを決定する。変換法には、べき乗変換などを適用できる。

40

【0076】

分布型パラメータ修正部は、解析した変換パラメータの確認と微調整が可能である。

【0077】

正規分布化データ変換部では、変換パラメータにしたがってデータを変換し、正規分布化を行う。

【0078】

正規分布化データ作成部は、 P 個の属性を持つ原データ配列 $x(p, i)$ から正規分布化値 $X(p, i)$ を作成する。

50

【0079】

基準化データ作成部

データベクトル間で類似度を調べる場合、データを構成する属性間で単位が異なると絶対値の大きな属性が類似度に大きな影響を持つ。このため、データベクトルを類似度に基づいて分類をする場合、その基準化が必要となる。

正規分布化されたデータ配列 $X(p, i)$ の値を変換して基準化値 $u(p, i)$ とする。この i 番目の基準化値を i 番目の入力ベクトル $u(p)$ ($p = 1, 2, \dots, P$) とする。

【0080】

自己組織化マップ特徴地図作成部

10

自己組織化マップ特徴地図作成部は、自己組織化マップ特徴地図の作成処理を行うブロックであって、初期値設定部と入力ベクトル提示部と勝者ユニット探索部と参照ベクトル更新部と学習回数判定部として機能する。

【0081】

初期値設定部は、事前設定としてネットワークの大きさ（ユニットの配列の要素数） M 、1つの入力ベクトルに対して行われる参照ベクトルの更新の回数（総学習回数） T 、ユニットの位相近傍形状（例えば、直角格子型及び六角格子型の何れか）、近傍領域の初期値 $N_c(0)$ 、及び学習率係数の初期値（0）を設定する。また、初期値設定部は複数の参照ベクトルを初期化するとともに、複数ある入力ベクトルの提示順を決定する。参照ベクトルの初期値は、すべてのニューロンを0とした場合と $-1 \sim 1$ の乱数を満たす場合

20

【0082】

入力ベクトル提示部は、初期値設定部によって決定された提示順に従って、基準化値 $u(p, i)$ から i 番目のものを選択して、 i 番目の入力ベクトル $u(p)$ を提示する。入力ベクトルが少数であった場合は、多重回数提示することによって事例数の増大を図る。

【0083】

勝者ユニット探索部は、入力ベクトル $u(p)$ に最も類似したパターンを参照ベクトル $r(p, w)$ の中から探す。類似距離の指標は、ユークリッド距離を用いる方法が用いられるが、ハミングの距離やマハラノビスの汎距離などの距離関数も適用できる。

【0084】

30

近傍領域演算部の近傍距離は、初期値を全体の学習回数 T の $1/2$ から $2/3$ の範囲とし、現在行なわれている学数回数から算出する。学数回数 t 回目の近傍を N_c 、初期値 $N_c(0)$ とすると、下記式9の式で近傍距離が決定できる。

【0085】

【数9】

$$N_c(t) = N_c(0) \times (1 - t/T)$$

学習率係数演算部は、初期値設定部によって設定された学習率係数の初期値（0）と初期値設定部によって設定された総学習回数 T と現在行われている学習の学習回数 t とを利用して下記数式10を演算し、現在行われている学習での学習率係数（ t ）を算出する。この学習率係数（ t ）は学習が進むにつれて小さくなっていく。

40

【0086】

【数10】

$$\alpha(t) = \alpha(0) \times (1 - t/T)$$

参照ベクトル演算部は、近傍領域演算部によって得られた近傍領域 $N_c(t)$ 内に含ま

50

れるユニットの参照ベクトルの夫々に対して、学習率係数演算部によって得られた学習率係数 $\alpha(t)$ を利用して下記数式 11 を演算して、ユニット w における学習回数 t に対する参照ベクトル $r(p, w)$ を $R(t, p)$ として、それを更新し、更新後の参照ベクトル $R(t+1, p)$ を得る。また、参照ベクトル演算部 44c は、近傍領域 $N_c(t)$ 内に含まれないユニットの参照ベクトルの夫々に対して下記式数式 12 を演算して参照ベクトル $r(t+1, p)$ を得る。

【0087】

【数11】

$$R(t+1, p) = R(t, p) + \alpha(t) \times (u(p) - R(t, p))$$

10

【0088】

【数12】

$$R(t+1, p) = R(t, p)$$

学習回数判定部は、現在行われている学習の学習回数 t が初期値設定部によって設定された総学習回数 T に達したか否かを判定する。

20

【0089】

意味空間地図作成部

意味空間地図作成部は、参照ベクトルの基線からの調整偏位量算出部、参照ベクトルの意味空間での偏位量算出部、意味空間地図作成部からなる。

【0090】

参照ベクトルの基線からの調整偏位量算出部

参照ベクトルの基線からの調整偏位量算出部は、全ユニット ($w = 1, 2, \dots, W$) について、数式 5 により、基線からの偏位度 $D(w)$ を算出する。そして、数式 7 により、調整偏位度 $D''(w)$ を算出する。

30

【0091】

参照ベクトルの意味空間での「重み量」指数 $L(w)$ 算出部

参照ベクトルの意味空間での偏位量算出部は、数式 8 により、2つの意味空間での「重み量」指数 $L(w)$ を算出する。

【0092】

意味空間地図作成部

意味空間地図作成部は、各ユニットの参照ベクトルの調整偏位量 $D''(w)$ を Y 軸に、各ユニットの参照ベクトルの「重み量」指数 $L(w)$ を X 軸にとり、各ユニットを 2次元的に展開して表示する。

【0093】

意味空間地図作成部は、各ユニットの類似性を表示する。前記各ユニットについて、他のユニットの全類似距離を求め、それが最も短いものを 3つ選び、類似の程度を例えば次のように表現する。その距離が極めて短い場合太い線で、それよりも長いが一応レベル以下の場合細い線で、さらにそれよりも長い場合は結ばない。また線の色は、たとえば類似距離が最短なものに対しては赤色、2番目のものに対しては橙色、3番目のものに対しては灰色とする。

40

【実施例 2】

【0094】

4323 例の病態診断用データ

本例の P 個の属性を持つ原データ配列 $x(p, i)$ は、病態診断用のデータで、性別、

50

年齢と20項目検診検査結果よりなる4323例のデータである。本例の20項目の検診検査項目は、血糖(BS)、HbA1c、HDL、総コレステロール(TC)、トリグリセライド(TG)、GGT、GPT、GOT、乳酸脱水素酵素(LD)、アルカリフォスファターゼ(ALP)、血小板(PLT)、白血球数(WBC)、TP、ALB、UA、BUN、CRE、Hb、Ht、RBCである。これらの項目のうち、血糖(BS)、HbA1cは、糖尿病関連である。HDL、総コレステロール(TC)、トリグリセライド(TG)は、過栄養である。GGT、GPT、GOT、乳酸脱水素酵素(LD)、アルカリフォスファターゼ(ALP)は、過栄養または肝障害関連である。血小板(PLT)、白血球数(WBC)、TPは、炎症関連である。ALB、UA、BUN、CREは、低栄養・腎機能関連である。Hb、Ht、RBCは、貧血関連である。この20の検診検査項目は、左側から糖尿病関連、過栄養・肝障害関連、炎症関連、低栄養・腎機能関連、貧血関連の順に配列されている。表1のデータは、その内の10例を示したものである。

10

【0095】

【表1】

病態診断用データ

性	年	BS	HbA1c	HDL	TC	TG	GGT	GPT	GOT	LD	ALP	PLT	WBC	TP	ALB	UA	BUN	CRE	Hb	Ht	RBC
0	49	83	4.5	68	157	45	65	18	27	161	150	11.8	3100	6.9	4.1	4.4	16	1.1	13	38.1	406
0	53	103	5.5	41	147	117	53	23	31	225	213	22.3	4400	6.7	4.1	5.8	16	0.8	15.2	44.2	455
0	85	92	5.1	45	138	104	71	16	26	204	179	25.6	10300	7	3.7	4.3	9	0.8	14.6	43.2	446
0	61	96	5.7	38	204	118	44	9	21	170	250	17.7	7800	7.2	4.3	8	16	1.5	15.5	45	501
0	61	109	5.3	37	181	69	27	15	14	134	253	30.3	8100	7.2	4.4	4.2	17	0.9	15.1	46.4	494
0	64	99	6.1	37	165	69	32	14	17	169	291	42.7	10400	7	3.9	4.2	10	0.8	14.1	41.4	461
0	58	100	4.6	88	226	66	342	64	93	185	238	18.5	5600	6.5	3.9	4.2	8	1	14.5	43.9	416
0	53	87	5.3	60	142	39	43	17	22	210	174	23.8	7500	7	4.1	4.5	14	0.8	13.8	43.7	446
0	64	122	4.3	68	221	88	85	27	21	181	266	18.7	8300	6.7	4.1	6.8	13	1.1	16.9	47.1	507
0	48	96	5.2	29	172	223	24	15	14	179	219	19.6	6500	6.6	4.2	4.8	13	1.1	15.5	43.8	478

20

表2は属性情報を示す。表2には、項目名、利用有無のフラグ(0=使わない、1:使う)、べき乗変換値と変換原点、基準範囲(男性用下限上限値)、基準範囲(女性用下限上限値)の順で示されている。

【0096】

【表2】

属性情報

属性名	SOM利用	べき乗	変換原点	男下限	男上限	女下限	女上限
性別	0	0	0	0	0	0	0
年	0	0	0	20	60	20	60
BS	1	0.5	0	74	110	74	110
HbA1c	1	1	0	4.6	5.8	4.6	5.8
HDL	1	1	0	35	85	40	100
TC	1	1	0	130	250	130	250
TG	1	0	20	40	200	40	200
GGT	1	0	6	12	90	12	50
GPT	1	0	7	8	45	8	32
GOT	1	0.3	8	12	30	12	35
LD	1	0.3	90	130	240	130	240
ALP	1	0.3	100	120	370	120	370
PLT	1	1	0	15	35	15	35
WBC	1	0.3	2000	3500	9500	3500	9500
TP	1	1	0	6.5	8	6.5	8
ALB	1	1	0	3.8	5	3.8	5
UA	1	1	0	4.5	8	3.5	6.8
BUN	1	1	0	7	21	7	21
CRE	1	1	0	0.5	1.25	0.4	0.9
Hb	1	1	0	14	18	11	15
Ht	1	1	0	38	49	34	45
RBC	1	1	0	380	580	340	500

40

50

入力ベクトル $u(p)$ の作成

この例の原データ $x(p, i)$ は、 $p: 1 \sim 20$ 、 $i: 1 \sim 4323$ である。本発明の実施例 1 の装置により、原データ $x(p, i)$ を正規分布化値 $X(p, i)$ と基準化値 $u(p, i)$ する。この i 番目の基準化値を、入力ベクトル $u(p)$ とする。

【0097】

自己組織化マップ法による特徴地図の作成

本発明の実施例 1 の装置の勝者ユニットの探索部により、20 個の属性よりなる参照ベクトル参照ベクトル $r(p, w)$ を持つ複数のユニットの中から、前記入力ベクトル $u(p)$ に最も類似した参照ベクトル $r(p, w)$ を持つ勝者ユニットを探索し、本発明の実施例 1 の装置の参照ベクトル更新部により、学習回数に応じて順次狭められる前記勝者ユニットの近傍領域に含まれる各ユニットの参照ベクトル $r(p, w)$ を前記学習回数に応じて順次小さくされる学習係数と前記入力ベクトル $u(p)$ とに基づいて更新し、前記学習回数が 4323 回に達するまで前記勝者ユニットの探索及び前記参照ベクトル $r(p, w)$ の更新を繰り返すことにより、 6×6 のユニットが配列した、自己組織化マップ法による特徴地図を作成した。

10

【0098】

自己組織化マップ法による特徴地図

このようにして作成した特徴地図には、 6×6 のユニットが配列し、各ユニットは互いに類似したものが近接的に配置されるように表示されている。しかしながら、このように作成された特徴地図は作り直すごとに位置的な情報が異なってしまう。

20

【0099】

図 6 の左側には、このようにして作成された 3 例の特徴地図を例示する。この特徴地図においては、ユニットが試行する毎に異なる配列になる。

【0100】

意味空間地図の作成

実施例 1 の装置で、意味空間地図作図部において、各ユニットの参照ベクトルの調整偏位量 $D''(w)$ を Y 軸に、各ユニットの参照ベクトルの「重み量」指数 $L(w)$ を X 軸にとり、特徴地図の 6×6 の各ユニットを 2 次元的に再配置する、意味空間地図を作成した。

【0101】

作成した意味空間地図

図 5 は、このようにして作成した意味空間地図である。この意味空間地図には、元の特徴地図のユニットの配列位置 $m - n$ が示されている。図 5 には、このうちのいくつかのユニットについて、健常群型 (3 - 3)、糖尿病型 (2 - 2、2 - 1、1 - 2、1 - 1)、肝障害型 (5 - 2、6 - 2、6 - 1)、腎障害型 (1 - 6)、低栄養・貧血型 (3 - 6、2 - 6) であることが示されている。これらのユニットについて、その参照ベクトルが図示されている。

30

【0102】

各ユニット間の類似性の表示

図 5 の意味空間地図には、各ユニット間の類似性が表示されている。前記各ユニットについて、他のユニットの全類似距離を求め、それが最も短いものを 3 つ選び、そしてその距離が極めて短い場合太い線で、それよりも長いが一レベル以下の場合細い線で、さらにそれよりも長い場合は結ばない。また線に着色する場合、類似距離が最短なものに対しては赤色、2 番目のものに対しては橙色、3 番目のものに対しては灰色とする。

40

【0103】

意味空間地図の再現性

図 6 は、本発明の意味空間地図の再現性を示すものである。左側の特徴地図においては、ユニットの配置の再現性が悪いのに対して、右側の意味空間地図では再現性が良いことが図示されている。

【0104】

本発明の意味空間地図では、マップ上の各ユニットはマップ上での配列の再現性がよい

50

ので、マップ上に自動的に病態名を表示することができる。

【0105】

べき乗変換による、意味空間地図出力の最適化

図7は、べき乗変換による意味空間地図の最適化を示した図である。図7上図は、べき乗変換を加えた本実施例のものであり、図7下図はべき乗変換を加えなかったものである。それぞれの属性固有のべき乗変換を加えた方が明確に意味を反映した地図を得ることができる。本実施例では、べき乗変換を施さないと異常値を多く持つ群のパターンが強調され過ぎて、図7のごとく早期発見が求められるより軽微な異常を持つパターンを、健常典型パターンから分離しにくくなる。

【0106】

従来のクラスター分析との対比

図8は、本実施例の意味空間地図と従来のクラスター分析による再配置結果を比較した図である。図8上図は、本実施例の意味空間地図である。図8下図は、従来のクラスター分析による再配置結果である。本実施例(図8上図)では、基準ユニットである5-4を中心として再配置され、他のユニットが標準(基準ユニット)からどの程度乖離しているかの判断が容易となる。これに対し、従来のクラスター分析を使ってデンドログラムの形式で再配置(図8下図)した場合には、枝の配置に意味を持たせることができない。このため、基準ユニット(5-4)も含めた健常パターンを示すユニット(灰色のクラス番号)と他の病態パターンを示すユニットが混在しており、かつ病態群の中でも重症の程度を区別できないなど、分類結果の解釈が困難である。

【0107】

意味空間地図による経時遷移の観察

図9は、ある患者が属するパターンの経時変化を示した図である。本発明の意味空間地図は、十分なデータ数に基づいて作成した場合、データを随時追加して再実行してもその分類結果の配置は基本的には変化しない。このため、同じ患者の検査結果をその上に随時表示することで、病態の変化を明確に把握できる。

【0108】

図10は、検診検査結果に生活習慣調査から得た生活習慣特性値を組み合わせて、自己組織化意味空間地図を作成した結果である。同地図を利用すれば、検査値の異常パターンと生活習慣特性の関連性を直感的に把握でき、検診後の指導に好都合である。また、病態パターン間の繋がりも容易となり、生活習慣病に対してきめ細やかな予防対策を立てるのに役立つ。

【図面の簡単な説明】

【0109】

【図1】本発明の自己組織化マップの表示装置のハードウェア構成

【図2】本発明の意味空間地図の表示を行うための処理の流れを示した図

【図3】基準化の概念の説明図 上部から、第p番目の属性の実際の分布図、正規分布化した分布図、そして基準化した基準範囲の求め方の原理図。

【図4】べき乗変換前後の計測値分布の比較

【図5】本発明の意味空間地図

【図6】意味空間地図の再現性を示す図。左側は特徴地図、右側は意味空間地図

【図7】べき乗変換による、意味空間地図の最適化を示した図。図7上図は、べき乗変換を加えた本実施例のものであり、図7下図は、べき乗変換を加えなかったもの。

【図8】自己組織化マップ分類結果に対する再配置法の比較図8の上図は、本実施例の意味空間地図。図8の下図は、従来のクラスター分析のデンドログラムによる再配置結果。

【図9】ある患者が属するパターンの経時変化を示した意味空間地図。

【図10】検診検査結果と生活習慣特性を組み合わせて作成した意味空間地図

10

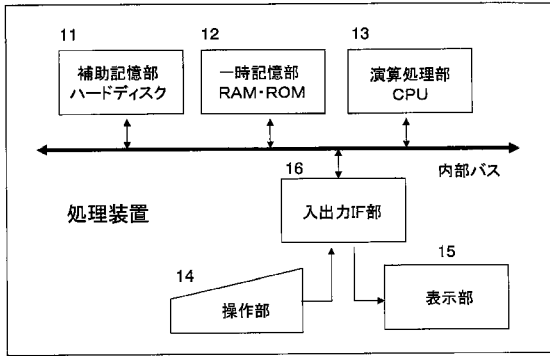
20

30

40

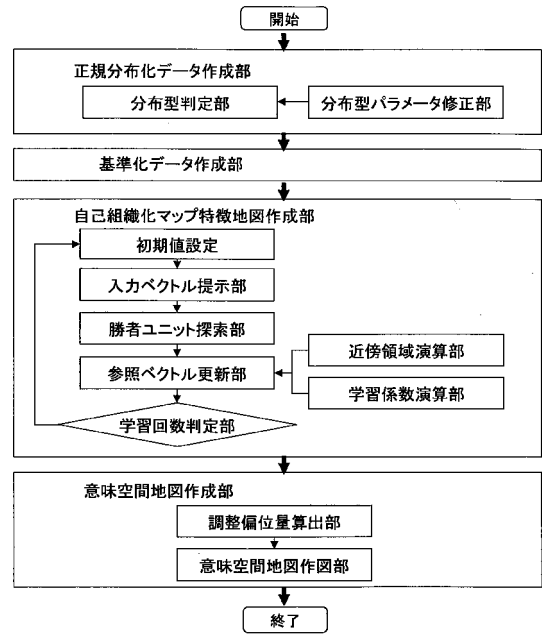
【 図 1 】

自己組織化意味空間表示および処理装置



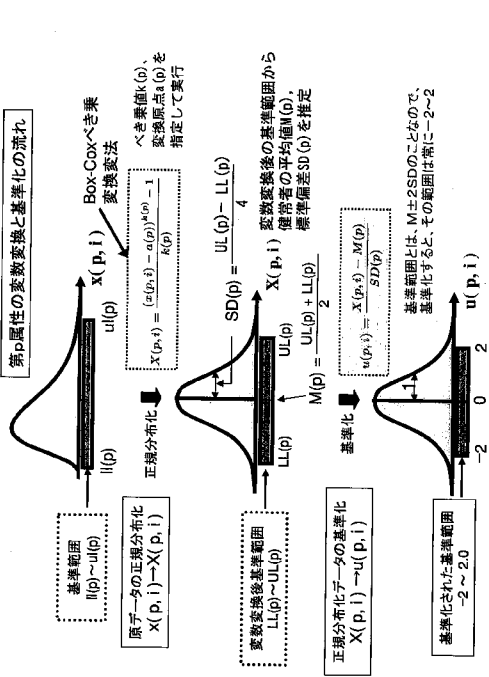
【 図 2 】

意味空間地図表示を行なうための処理の流れを示した図



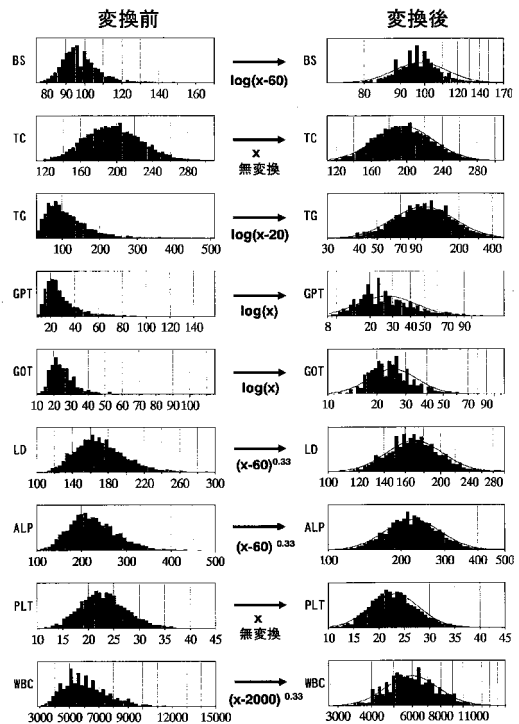
【 図 3 】

基準化データ作成部の基準化の概念を示した説明図。上部から、第p番目の属性の属性の分布図、正規分布化した分布図そして、基準化した基準範囲の求め方の原理図。

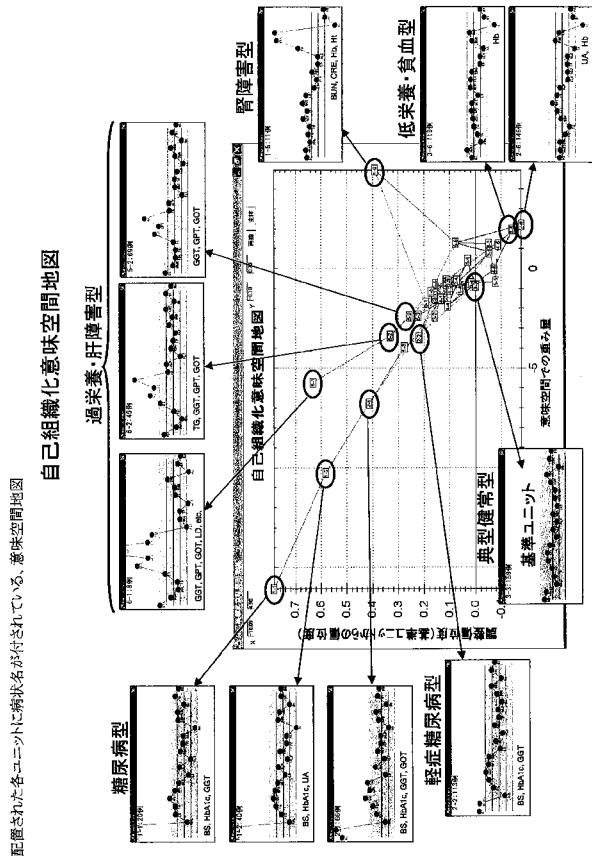


【 図 4 】

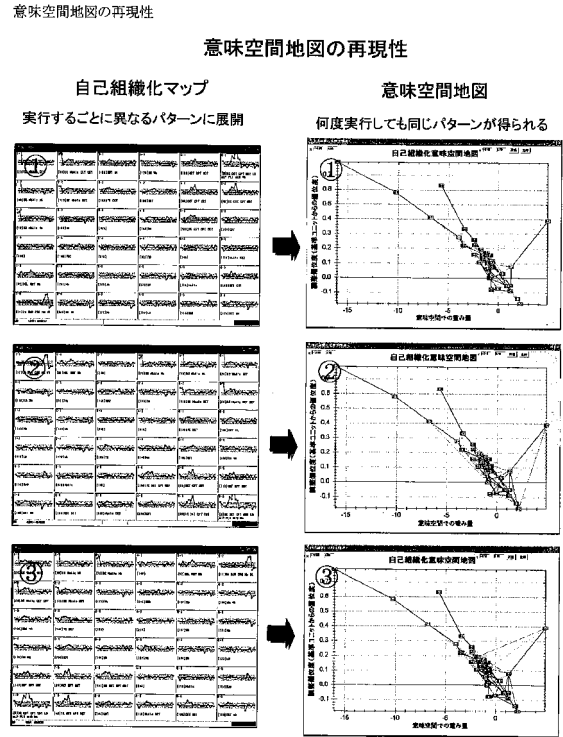
べき乗変換前後の計測値分布の比較



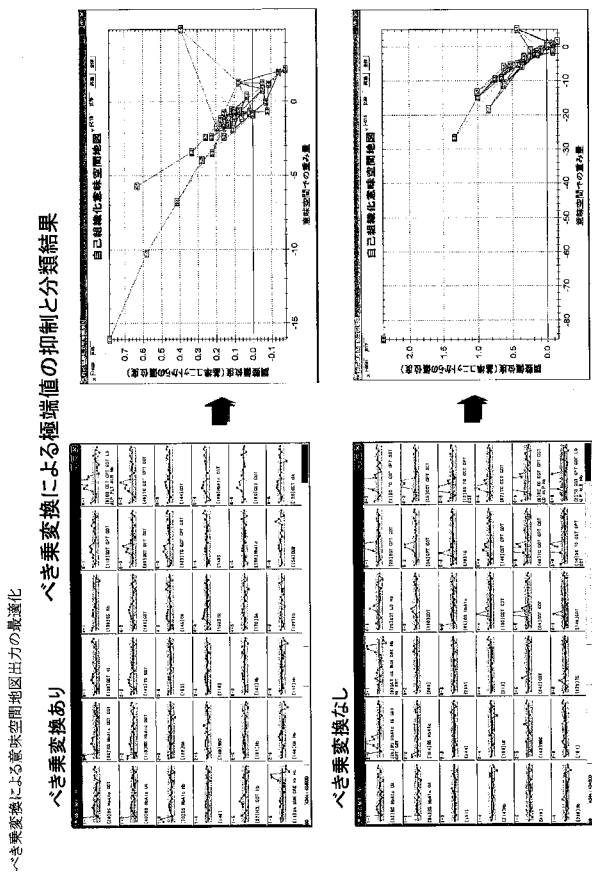
【 図 5 】



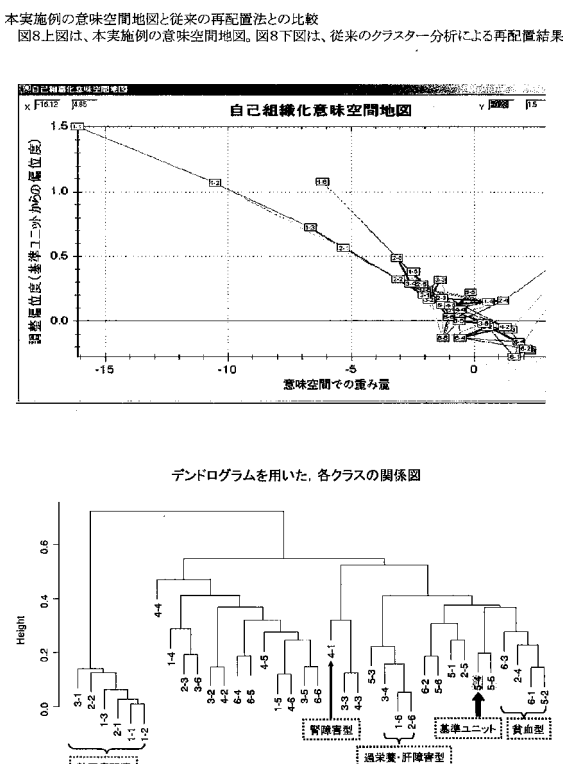
【 図 6 】



【 図 7 】

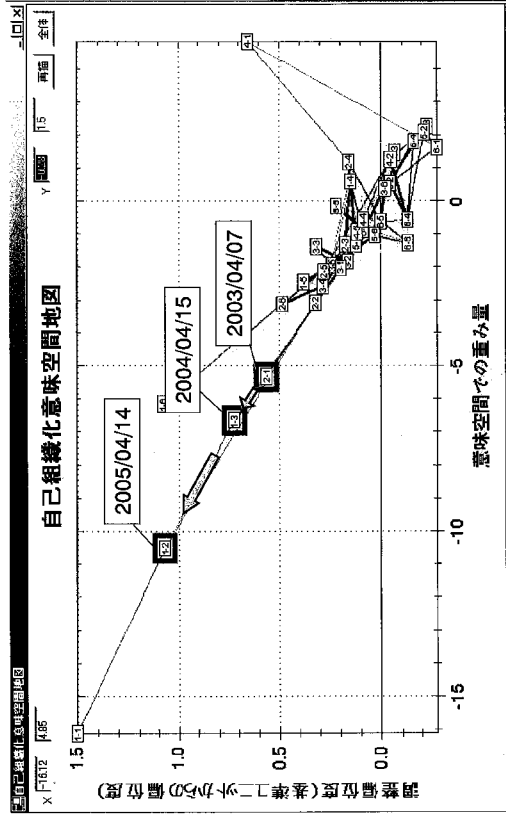


【 図 8 】



【 図 9 】

意味空間地図による種時遷移の表示



【 図 10 】

検診検査結果と生活習慣特性を組み合わせて作成した意味空間地図

