(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第6644767号
(P6644767)
(45)発行日　令和2年2月12日(2020.2.12)　(24)登録日　令和2年1月10日(2020.1.10)

(51)Int.Cl.　　F I
G06F 16/9035 (2019.01)　　G06F 16/9035
G06F 16/28 (2019.01)　　G06F 16/28

請求項の数 11　(全 12 頁)

(21)出願番号　特願2017-506965(P2017-506965)
(86)(22)出願日　平成27年8月27日(2015.8.27)
(65)公表番号　特表2017-526065(P2017-526065A)
(43)公表日　平成29年9月7日(2017.9.7)
(86)国際出願番号　PCT/EP2015/069574
(87)国際公開番号　WO2016/030436
(87)国際公開日　平成28年3月3日(2016.3.3)
審査請求日　平成30年8月23日(2018.8.23)
(31)優先権主張番号　PCT/CN2014/085560
(32)優先日　平成26年8月29日(2014.8.29)
(33)優先権主張国・地域又は機関
中国(CN)
(31)優先権主張番号　14194063.5
(32)優先日　平成26年11月20日(2014.11.20)
(33)優先権主張国・地域又は機関
欧州特許庁(EP)

(73)特許権者　590000248
コーニンクレッカ　フィリップス　エヌヴェ
KONINKLIJKE PHILIPS N.V.
オランダ国　５６５６　アーエー　アインドーフェン　ハイテック　キャンパス　５
High Tech Campus 5, NL-5656 AE Eindhoven
(74)代理人　110001690
特許業務法人M&Sパートナーズ
(74)代理人　100122769
弁理士　笛田　秀仙

最終頁に続く

(54)【発明の名称】相互相関に基づかれる階層的なデータ分析に関する方法及び装置

(57)【特許請求の範囲】
【請求項１】
相互相関に基づく、複数の属性を含むデータの階層的なデータ分析に関する装置であって、
データセットにおける各データの属性を名義値へと標準化する標準化部と、
前記属性の標準化された前記名義値に基づき、前記データセットにおける各データの前記属性の間の相関を算出する計算器と、
カテゴリ及び前記カテゴリの間の相関の第１のグラフを生成する第１の生成器であって、各カテゴリが、所定の規則に基づき分類された属性を含み、前記カテゴリの間の各相関は、個別のカテゴリの属性の間の平均的相関である、第１の生成器、又は推奨された属性の第１のグラフを生成する第１の生成器と、
前記第１のグラフからユーザにより選択された第１の属性、相関される属性及び前記第１の属性と前記相関される属性との間の相関の第２のグラフを生成する第２の生成器であって、前記第１の属性と各相関される属性との間の相関が、所定の相関閾値以上である、第２の生成器と、
前記第２のグラフからユーザにより選択される前記第１の属性及び少なくとも第２の属性の値に基づき、相関されるデータの統計分布の第３のグラフを生成する第３の生成器であって、前記相関されるデータが、前記第１の属性及び少なくとも前記第２の属性を含む、第３の生成器とを有し、
前記データが、医療データである、装置。

【請求項２】
前記名義値が、各データの前記属性に係る名義値とスケール値との間のマッピングを規定する所定の診断規則に基づき決定される、請求項１に記載の装置。
【請求項３】
前記第１のグラフの属性が、各属性のユーザによる選択頻度に基づき推奨される、請求項１又は２に記載の装置。
【請求項４】
前記第１の属性及び少なくとも前記第２の属性のユーザにより選択された値に基づき、関連付けられるデータのリストを生成する第４の生成器を更に有し、
前記関連付けられるデータが、前記第１の属性及び少なくとも前記第２の属性を含む、請求項１乃至３のいずれか一項に記載の装置。
【請求項５】
２つのカテゴリ又は属性の間の前記相関が、前記２つのカテゴリ又は属性を接続する相関インジケーターにより提示され、
前記相関インジケーターの視覚的な特性は、前記２つのカテゴリ又は属性の間の前記相関の値に基づく、請求項１乃至４のいずれか一項に記載の装置。
【請求項６】
相互相関に基づく、複数の属性を含むデータの階層的なデータ分析に関する、コンピュータ実現による方法であって、
データセットにおける各データの属性を名義値へと標準化するステップと、
前記属性の標準化された前記名義値に基づき、前記データセットにおける各データの前記属性の間の相関を算出するステップと、
カテゴリ及び前記カテゴリの間の相関の第１のグラフを生成するステップであって、各カテゴリが、所定の規則に基づき分類された属性を含み、前記カテゴリの間の各相関は、個別のカテゴリの属性の間の平均相関である、ステップ、又は推奨された属性の第１のグラフを生成するステップと、
前記第１のグラフからユーザにより選択された第１の属性、相関される属性及び前記第１の属性と前記相関される属性との間の相関の第２のグラフを生成するステップであって、前記第１の属性と各相関される属性との間の相関が、所定の相関閾値以上である、ステップと、
前記第２のグラフからユーザにより選択される前記第１の属性及び少なくとも第２の属性の値に基づき、相関されるデータの統計分布の第３のグラフを生成するステップであって、前記相関されるデータが、前記第１の属性及び少なくとも前記第２の属性を含む、ステップとを有し、
前記データが、医療データである、方法。
【請求項７】
前記名義値が、各データの前記属性に係る名義値とスケール値との間のマッピングを規定する所定の診断規則に基づき決定される、請求項６に記載の方法。
【請求項８】
前記第１のグラフの属性が、各属性のユーザによる選択頻度に基づき推奨される、請求項６又は７に記載の方法。
【請求項９】
前記第１の属性及び少なくとも前記第２の属性の値に基づき、関連付けられるデータのリストを生成するステップを更に有し、
前記関連付けられるデータが、前記第１の属性及び少なくとも前記第２の属性を含む、請求項６乃至８のいずれか一項に記載の方法。
【請求項１０】
２つのカテゴリ又は属性の間の前記相関が、前記２つのカテゴリ又は属性を接続する相関インジケーターにより提示され、
前記相関インジケーターの視覚的な特性は、前記２つのカテゴリ又は属性の間の前記相

関の値に基づく、請求項６乃至９のいずれか一項に記載の方法。
【請求項１１】
コンピュータに、請求項６に記載の方法のステップを実行させるためのコンピュータプログラムコード手段を含む、コンピュータプログラム。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は一般に、相関分析に基づき、関心データにアクセスすること、特にマスデータの相関分析に基づき、関心臨床データにアクセスすることに関する。
【背景技術】
【０００２】
今日では、病院における一般的な電子情報システムは、分析のためマスデータを集めることを可能にする。相関は、観察、予測、診断及び他の目的に有益な新しい知識を生成するため集められるデータの間の相互のインパクトを調査する重要な分析法である。しかしながら、あるデータ型（例えば数値、名詞等）のデータベースから抽出されるデータは、異なる種類の相関算出方法を用いて処理される必要がある。これは比較に適していない。更に、斯かる大量の情報、例えば、患者当たり２００以上のデータ属性を持つＣＶＩＳ（心血管情報システム）は、個別の特徴及び影響を調査することに関心のあるユーザにデータ及びデータ間の相関を提示するのに、うまく設計された構造を必要とする。
【発明の概要】
【発明が解決しようとする課題】
【０００３】
ＵＳ２０１３／０１３８５９２Ａ１号は、複数の属性を使用することにより関係グラフを生成し、仮説を表すために関係グラフからサブグラフを抽出する、マスデータ処理に関する方法を開示する。ここで、相関は、データ属性の依存分類に基づき生成される。また、異なる統計的検定により推定される相関を一様に表すため、ｐ値として表される相関値が使用される。これは関連付けられる属性の特定のデータ型に基づき決定される。しかしながら、ｐ値として表される相関値が、異なる仮説を処理するさまざまな統計的検定から生成されることができるにもかかわらず、いわゆる統一された相関値は、一貫した定量的値又は仮説を反映しておらず、及び従って比較に関して適切ではない。依存分類は、提供される相関を減らし、これによりユーザ利便性を改良する。しかし、それらは、データ型の潜在的依存の調査を抑制し、データに含まれる情報の部分を逃す。更に、階層的な分析が、データ処理に関して提供されるものではなく、すべてのデータ処理が、属性レベルに対して行われる。これは、分析を非効率的かつ不完全にする。
【０００４】
ＵＳ２０１２／２１５４５５Ａ１号は、ある方法を開示する。これは通信モジュールを用いて少なくとも１つの位置信号を受信し、位置信号から得られる地球空間的データをタイムスタンプと共にメモリに格納し、通信モジュールを用いてセンサから時間にわたり生医学的信号を受信するステップを含む。受信されたバイオ信号からの生医学的データは、タイムスタンプと共にメモリに格納される。位置信号を受信すること及び位置から地球空間的データを格納することが、異なる地理的な位置において繰り返される。
【０００５】
Patricio Soares Costaらによる「The use of multiple correspondence analysis to explore associations between categories of qualitative variables in healthy ageing」、Journal of aging research、vol. 2013、302163、2013、XP55190591は、コグニティブ老化を調査するために用いられる大きなデータセットにおいて基礎をなす構造を検出及び表すことにおいて、多重対応分析（ＭＣＡ）の適用性を示す研究を開示する。
【０００６】
従って、データの完全な調査を容易にし、明白で単純な態様においてユーザの関心情報を提示する効率的な方法及び装置を提供することが望ましい。

【課題を解決するための手段】
【０００７】
これらの懸念の１つ又は複数を好適に解決するため、本発明の１つの側面のある実施形態によれば、相互相関に基づかれる階層的なデータ分析に関する装置及び方法が提供される。
【０００８】
相互相関に基づき、複数の属性を含むデータのデータ分析に関する装置が提供され、この装置は、
データセットにおける各データの属性を名義値へと標準化する標準化部と、
上記属性の標準化された名義値に基づき、上記データセットにおける各データの上記属性の間の相関を算出する計算器と、
カテゴリ及び上記カテゴリの間の相関の第１のグラフを生成する第１の生成器であって、各カテゴリが、所定の規則に基づき分類された属性を含み、上記カテゴリの間の各相関は、個別のカテゴリの属性の間の平均的相関である、第１の生成器、又は推奨された属性の第１のグラフを生成する第１の生成器と、
上記第１のグラフからユーザにより選択された第１の属性、関連付けられる属性及び上記第１の属性と上記関連付けられる属性との間の相関の第２のグラフを生成する第２の生成器であって、上記第１の属性と各関連付けられる属性との間の相関が、所定の相関閾値以上である、第２の生成器と、
上記第２のグラフからユーザにより選択される上記第１の属性及び少なくとも第２の属性の値に基づき、上記関連付けられるデータの統計分布の第３のグラフを生成する第３の生成器であって、上記関連付けられるデータが、上記第１の属性及び少なくとも上記第２の属性を含む、第３の生成器とを有する。
【０００９】
統計分布が、座標平面において提示される。そこでは、第１の属性及び少なくとも第２の属性の各値の組み合わせ及び各値組み合わせに対応する統計が、軸の値及び少なくとも統計インジケーターの特徴的な視覚的な特性により表され、統計インジケーターが、第１の属性及び少なくとも第２の属性の値組み合わせ並びに値組み合わせに対応する統計を示す。
【００１０】
本発明において、属性値の標準化及び属性の間の相互相関に基づかれる階層的なデータ分析装置を導入することが提案される。属性のスケール値の名義値への標準化は、属性の相関の仮説に関する基礎を提供し、これは、更なる観察及び比較を科学的に正当化する。複数層の階層的な調査は、属性のレベルに関する分析だけでなく、関連付けられるデータの分析も可能にする。これはより詳細な観察を提供し、マスデータ分析を効率的で完全なものにする。
【００１１】
ある実施形態において、標準化は、ドメイン知識に基づかれる。
【００１２】
ドメイン知識に基づかれるスケール値の名義値への標準化は、データ分析を医学的により意味があるものにし、効率的にする。スケール値の代わりに、名義値は、例えば「正常」又は「異常」といった属性の状態の直接的で単純な規定を与える。これは分析をより知覚可能なものにする。
【００１３】
ある実施形態において、推奨は、選択頻度又は医療ガイドラインに基づかれる。
【００１４】
ある実施形態において、上記装置は、上記第１の属性及び少なくとも上記第２の属性のユーザにより選択された値に基づき、関連付けられるデータのリストを生成する第４の生成器を更に有し、上記関連付けられるデータが、上記第１の属性及び少なくとも上記第２の属性を含む。

【００１５】
この装置は、関連付けられるデータの内容を調査するため、１つの追加的な層を提供する。これは属性のカテゴリ／上位属性、属性、関連付けられるデータ及びデータ内容の完全な調査を完成させる。それは、利用可能なデータに含まれるすべての情報をユーザが完全に利用することを可能にする。
【００１６】
ある実施形態において、２つの属性の間の上記相関が、上記２つの属性を接続する相関インジケーターにより提示され、上記相関インジケーターの視覚的な特性は、上記相関値に基づかれる。
【００１７】
各相関インジケーターの視覚的な特性を用いた、属性の間の相関値の即時の視覚化は、属性の間の複雑な関係の便利な理解を容易にする。
【００１８】
本発明は、相互相関に基づき、複数の属性を含むデータのデータ分析に関する方法を有する。この方法は、
データセットにおける各データの属性を名義値へと標準化するステップと、
上記属性の標準化された名義値に基づき、上記データセットにおける各データの上記属性の間の相関を算出するステップと、
カテゴリ及び上記カテゴリの間の相関の第１のグラフを生成するステップであって、各カテゴリが、所定の規則に基づき分類された属性を含み、上記カテゴリの間の各相関は、個別のカテゴリの属性の間の平均的相関である、ステップ、又は推奨された属性の第１のグラフを生成するステップと、
上記第１のグラフからユーザにより選択された第１の属性、関連付けられる属性及び上記第１の属性と上記関連付けられる属性との間の相関の第２のグラフを生成するステップであって、上記第１の属性と各関連付けられる属性との間の相関が、所定の相関閾値以上である、ステップと、
上記第２のグラフからユーザにより選択される上記第１の属性及び少なくとも第２の属性の値に基づき、上記関連付けられるデータの統計分布の第３のグラフを生成するステップであって、上記関連付けられるデータが、上記第１の属性及び少なくとも上記第２の属性を含む、ステップとを有する。
【００１９】
本開示のさまざまな側面及び特徴が、以下により詳細に表される。本発明の他の目的及び利点は、詳細な説明及び添付の図面を参照してより明らかになり、及び容易に理解される。
【図面の簡単な説明】
【００２０】
【図１】本発明の実施形態の相互相関に基づかれる３層データ分析に関する装置を示す概略図である。
【図２】推奨された属性の第３のグラフを示す概略図である。
【図３ａ】属性のカテゴリ及びカテゴリの間の相関の第３のグラフを示す概略図である。
【図３ｂ】属性のカテゴリ及びカテゴリの間の相関の第３のグラフを示す概略図であり、選択されたカテゴリの属性が更に表示される図である。
【図４ａ】第１の属性、関連付けられる属性、及び第１の属性と第１の関連付けられる属性との間の相関の第１のグラフを示す概略図である。
【図４ｂ】第１のグラフの第２の属性の値に基づかれる関連付けられるデータの統計の第２のグラフを示す概略図であり、関連付けられるデータが、第１の属性及び第２の属性を有する、図である。
【図５ａ】第１の属性、関連付けられる属性及び第１の属性と第１の関連付けられる属性との間の相関の第１のグラフを示す概略図である。
【図５ｂ】第１のグラフの第２の属性及び第３の属性の値に基づかれる関連付けられるデ

ータの統計の第２のグラフを示す概略図であり、関連付けられるデータが、第１の属性、第２の属性及び第３の属性を有する、図である。
【図６】本発明の実施形態の相互相関に基づかれる３層データ分析に関する方法を示す概略図である。
【発明を実施するための形態】
【００２１】
本発明が、以下、実施形態及び図面を参照し、これらを組み合わせてより詳細に記載及び説明される。
【００２２】
図面における同じ参照符号は、類似する又は対応する特徴及び／又は機能を示す。
【００２３】
本発明は、特定の実施形態及び特定の図面を参照して説明されることになるが、本発明はそれらに限定されるものではなく、請求項によってのみ定まるものである。記載された図面は、概略的なものに過ぎず非限定的なものである。図面において幾つかの要素の大きさが誇張されている場合があり、説明目的のため実際のスケール通りに描かれていない場合がある。
【００２４】
図１は、相互のインパクトを調査するための本発明のある実施形態による相互相関に基づかれる、３層（カテゴリ／推奨－属性－データ）データ分析に関する装置を示す概略図である。本発明の分析に関する臨床データは、複数の属性を有する。各属性は、特定の患者の人口統計学的情報、生活様式情報、医療情報、ケアプロバイダ情報、歴史及びリスク要素情報、過去の訪問情報、手順情報等の１つのアイテムを含む。医療情報は、患者の基本的な健康情報、病変情報、デバイス情報及びフォローアップ情報を含む。各属性の値は、名義又はスケールタイプのいずれかとすることができる。名義タイプは、大きさに関して連続的でなく、測定可能でなく及び識別可能でない一種の値である。例えば性、故郷、仕事状態いったほとんどの人口統計学的情報及び薬物タイプ、病変タイプ、使用されたデバイスといったいくつかの医療歴史情報は、名義である。これは数値的に測定されることができない。対照的に、スケールタイプは、大きさに関して連続的で、測定可能で及び識別可能である一種の値である。例えば、例えば年齢といった人口統計学的情報及び薬物量、病変説明パラメータといった医療歴史情報は、スケールタイプ情報であり、これは、数値的に測定されることができる。上述した複数のデータは、本発明の分析対象としてのデータセットを構成する。標準化部１０１は、更なる分析に関して普遍的に比較できる基礎を提供するため、統一された標準下の名義値へとすべての属性の値を標準化する。統一された標準は、ドメイン知識に基づかれる。例えば、スケール値は、例えばアメリカ心臓病学会（ＡＣＣ）ガイドラインといった臨床ガイドラインに基づき、及び／又は局所標準を考慮する心臓専門医による入力に基づき、「正常」及び「異常」へと変換される。ガイドライン及び／又は専門家の入力を用いて、複数の属性を組み合わせることから、追加の属性が得られることができる。例えば、ＣＴＯが実行されたかどうか（はい／いいえ）、及びポスト手順、バイオマーカー、ＴＩＭＩが３であるかどうかから、名義ＣＴＯ結果（成功／失敗／ＣＴＯ未実行）が得られることができる。統一された標準化（名義値に変換されるスケール値）を用いて、属性の値は、すべての属性に関連付けられる１つの仮説下において生成される。これは、属性の相関分析に関する正当化された基礎を証明する。属性の変換された値に基づき、計算器１０２は、属性の間の相関を算出する。名義値に関する適切な統計方法が、この算出のために採用されることができ、それは例えばカイ二乗検定方法、フィッシャー正確な試験方法、二項試験方法、クラスカル－ウォリス試験方法などである。すべての属性に関する汎用仮説に基づき生成される相関は、科学的に意味があり、及び比較可能である。
【００２５】
第１の生成器１０３は、カテゴリ及びカテゴリの間の相関の第１のグラフを生成する。属性は、所定の規則又はデータレジストリ分類に基づきカテゴリに分類される。これは臨

床活動の定義、経済的要因に関連付けられる情報、ライフスタイル分類、フォローアップ情報、歴史及びリスク要素、生体構造情報、病変情報、デバイス情報、インシデント／合併症情報などに基づかれることができる。その後、カテゴリ及びそれらの間の相関は、カテゴリに関する依存関係の概要を与えるために提示される。カテゴリの間の相関は、各カテゴリへと分類される属性の相関値に基づかれる。１つの実現に関して、各カテゴリへと分類される属性の間の平均相関値が、カテゴリの間の相関を表すために利用されることができる。１つのカテゴリが選択された後、ユーザにより選択されたカテゴリの属性が表示される。属性のカテゴリは、データ分析のために処理される上部層として実現される。これは選択及び観察に関する選択肢を減らす。関心カテゴリの属性の更なる表示と共に、分析手順は、関心属性を捜し出す観点から、ユーザにとってより効率的になる。代替例として、データ分析に関する第１の層は、例えば相関又は他の基準に基づかれる臨床推奨、専門家の示唆又は計算された短リストからの、限られた推奨された属性のリストとして実現されることもできる。追加的に、データ分析に関する必要条件としてデータの構造を統一するため、データのプリプロセッサが採用されることができる。例えばＣＩＳ（臨床情報システム）、ＬＩＳ（研究室情報システム）、ＲＩＳ（放射線学情報システム）等のさまざまな電子情報システムが、病院での使用に利用可能である。これはさまざまなデータフォーマットを生じさせる。異なる情報システムにわたるデータ分析のため、すべてのデータに関して共通の基礎を提供するべく、統一された構造が望ましい。こうして、すべてのデータに関して特定の属性の相関分析が可能にされる。統一された構造は、利用可能な情報システムに関する可能な限りのすべての属性の一体化として設計されることができ、元の１つと比較して失った属性に関する新しい統一データを形成するため、値詰め（value stuffing）が実行される。例えば、新しい生成されたデータに関して欠けている属性に、ゼロが詰め込まれることができる。

【００２６】

第２の生成器１０４は、第１の属性、関連付けられる属性、及び第１の属性と第１の関連付けられる属性との間の相関の第２のグラフを生成する。第１の属性は、プリファレンスからユーザにより選択される属性である。関連付けられる属性は、第１の属性との相関が所定の相関閾値を越える属性である。例えば、名義値に適した統計方法の相関値が、統計的有意性によりｐ値として提示され、一般に受け入れられた閾値が０．０５にセットされる。それらの間の相関が、更なる調査のため提示される。提供されるのは、ユーザにより選択される属性及びその関連付けられる属性を明白で単純な態様において視覚化することである。

【００２７】

第３の生成器１０５は、ユーザにより選択される第２のグラフの第１の属性及び少なくとも第２の属性の値に基づき、関連付けられるデータの統計分布の第３のグラフを生成する。ここで、関連付けられるデータは、第１の属性及び少なくとも第２の属性を有する。第２の生成器１０４は、ユーザにより選択される属性に関連付けられるデータの詳細な調査を実現する。これは、関連付けられるデータのより多くの情報を統計的観点から提供する。第４の生成器（図１において図示省略）は、第１の属性、第２の属性及び／又は第３の属性に関してユーザにより選択された値に基づき、データリストを提示するよう配備されることができる。

【００２８】

図２、図３ａ及び図３ｂは、第３層データ分析のユーザインタフェースの実現である。図２は、推奨された属性の第１のグラフを示す概略図である。選択ウィンドウ３０１が、第３層分析の選択に関してセットされる。これは上位５つの結果測定又は分類とすることができる。トップ５結果測定に関しては、それらは、所定の規則に基づき、例えば、それらが選択される頻度又は医療ガイドラインに基づき、推奨される。その後、ディスプレイ領域３０２が、推奨される属性（属性０１～属性０５）に基づき、提示する。図３ａ及び図３ｂは、属性のカテゴリ及びカテゴリの間の相関の第１のグラフを示す概略図であり、それらは、ユーザにより選択されるカテゴリの属性を更に表示する。カテゴリが選択ウィ

ンドウ３０１を通して選択される場合、すべての属性は、ユーザがプリファレンスを選択するため、分類されたカテゴリ（カテゴリ０１～カテゴリ０５）において提示される。そして、カテゴリの間の相関が、両方のカテゴリを接続する相関インジケーターにおいて提示される。実施形態の相関インジケーターは、ラインの形である。ラインの厚さは、カテゴリの間の相関値を表す。特定の閾値を下回るあまりに弱い相関を持つカテゴリは、接続ラインを持たない。例えば、カテゴリ０２及びカテゴリ０５の間のラインは、カテゴリ０２及びカテゴリ０４の間のラインより薄い。これはカテゴリ０２が、カテゴリ０５よりカテゴリ０４とより強い相関を持つことを示す。相関値は、他の視覚的な特性又はインジケーターの他の形状により、提示されることもできる。視覚的な特性は、色、輝度、充填パターン又はその他とすることができる。形状は、バー、チェーン又はその他とすることができる。１つのカテゴリが、例えばカテゴリ０３が選択された後、カテゴリ０３に分類されるすべての属性（属性０３、属性０６、属性０７、属性０８、属性０９）のリスト３０２１が、ユーザによる更なる選択のためカテゴリ０３の下に表示される。この場合、ユーザは、属性０７を選択する。図２、図３ａ及び図３ｂは、この効率を拡張するデータ分析階層の上部層の実施形態である。

【００２９】

図４ａ及び図４ｂは、ユーザにより選択される第１の属性及び第２の属性を備える、第２及び第３層データ分析のユーザインタフェースの実現である。図４ａは、第１の属性、関連付けられる属性及び第１の属性と関連付けられる属性との間の相関の第２のグラフを示す概略図である。このインタフェースは、属性ディスプレイ領域４０１、属性選択ディスプレイウィンドウ４０２及びチャートボタン４０３を含む。属性ディスプレイ領域４０１は、生成された第１のグラフを表示するために用いられる。ユーザにより選択された第１の属性は属性０７である。これは中心に配置される。鎖点入りのライン４０１１～４０１５によりセグメント化される各領域は、１つのカテゴリの関連付けられる属性に割り当てられ、特定の基準に基づきソートされる。例えば、ある実施形態において、統計的有意性の昇順にされる。例えば、鎖点入りのライン４０１２及び鎖点入りのライン４０１３によりセグメント化される領域は、カテゴリ０３（属性０３、属性０６、属性０７、属性０８、属性０９）の関連付けられる属性に割り当てられる領域である。更に、分類された関連付けられる属性が、両側に散乱させられる。左側に配置される関連付けられる属性は、ユーザにより選択された属性０７とだけ相関する属性である。右側に配置される関連付けられる属性は、ユーザにより選択された属性０７を含む複数の属性と相関する属性である。その後、属性０２は、第２のグラフからユーザにより選択される第２の属性として選択される。任意の属性が図４ａにおいて選択される前に、属性上でのホバリングが、ライン（図示省略）に沿って詳細な情報（例えばｐ値及び相関強さといった統計的有意性）が表示されることをトリガーする。属性がユーザにより選択される属性として選択されるときはいつでも、それは、属性選択ディスプレイウィンドウ４０２に表示される。チャートボタン４０３は、関連付けられる属性の統計分布を示すことを可能にする。図４ｂは、第１のグラフから選択される第１の属性、第２のグラフから選択される第２の属性及び第１の属性を含む関連付けられるデータの値に基づき、関連付けられるデータの統計の第３のグラフを示す。ここで、関連付けられるデータは、第１の属性及び第２の属性を有する。インタフェースは、統計分布ディスプレイ領域５０１及び属性選択ディスプレイウィンドウ５０２を含む。チャートは、属性０７及び属性０２の異なる値に基づかれるバーチャートである。属性０７の値は、「正常」又は「異常」であり、属性０２の値は、「はい」又は「いいえ」である。これは４つの組み合わせを生じさせる。それぞれ、４つの組み合わせに関するバー形状の統計インジケーター５０１１～５０１４により提示される関連付けられるデータ分布が、座標平面に示される。ここで、ｙ軸は、対応する組み合わせに関して関連付けられるデータの番号を表し、ｘ軸は、第１の属性０７の値を表し、色が、第２の属性０２の値を表す。調査のためユーザ（図示所略）により選択される特定の組み合わせのデータのリストを示すため、更なる処理が行われることができる。この処理は、組み合わせを表すバーインジケーター上でのクリックにより又はユーザからの入力により実現さ

れることができる。

【００３０】

図５ａ及び図５ｂは、ユーザにより選択される第１の属性、第２の属性及び第３の属性を持つ第１及び第２層データ分析のユーザインタフェースの実現である。図６ａに関して、唯一の差は、ユーザにより第３の属性が選択される点にある。ここで、ユーザにより選択される第３の属性は、値が「はい」又は「いいえ」のいずれかである属性０９である。これは、８つの組み合わせを生じさせる。図５ｂに関して、関連付けられるデータ分布及び８つの組み合わせが、座標平面に示される。ここで、ｙ軸は、対応する組み合わせに関して関連付けられるデータの番号を表し、ｘ軸は、第１の属性の値を表し、色が、第２及び第３の属性の値を表す。

【００３１】

第１の属性に関連付けられる複数の属性が、統計分布分析に関して含まれることができ、例えば強度及び充填パターンといった統計特性の複数の視覚的な特性が、属性の値の複数の組み合わせを表すために利用されることができる。

【００３２】

図６は、本発明の実施形態において相互相関に基づかれる３層データ分析に関する方法を示す概略図である。本発明は、相互相関に基づかれるデータ分析の方法を有する。データは、複数の属性を有する。この方法は、

ステップ１０１：データセットにおける各データの属性を名義値へと標準化するステップと、

ステップ１０２：属性の標準化された名義値に基づき、データセットにおける各データの属性の間の相関を算出するステップと、

ステップ１０３：カテゴリ及びカテゴリの間の相関の第１のグラフを生成するステップであって、各カテゴリが、所定の規則に基づき分類された属性を含み、カテゴリの間の各相関は、個別のカテゴリの属性の間の平均的相関である、ステップ、又は推奨された属性の第１のグラフを生成するステップと、

ステップ１０４：第１のグラフからユーザにより選択された第１の属性、関連付けられる属性及び第１の属性と関連付けられる属性との間の相関の第２のグラフを生成するステップであって、第１の属性と各関連付けられる属性との間の相関が、所定の相関閾値以上である、ステップと、

ステップ１０５：第２のグラフからユーザにより選択される第１の属性及び少なくとも第２の属性の値に基づき、関連付けられるデータの統計分布の第３のグラフを生成するステップであって、関連付けられるデータが、第１の属性及び少なくとも第２の属性を含む、ステップとを有する。

【００３３】

図面、開示及び添付された請求項の研究から、開示された実施形態に対する他の変形が、請求項に記載の本発明を実施する当業者により理解され、実行されることができる。請求項において、単語「有する」は他の要素又はステップを除外するものではなく、不定冠詞「ａ」又は「ａｎ」は複数性を除外するものではない。シングルプロセッサ又は他のユニットが、請求項に記載される複数のアイテムの機能を満たすことができる。特定の手段が相互に異なる従属項に記載されるという単なる事実は、これらの手段の組み合わせが有利に使用されることができないことを意味するものではない。コンピュータプログラムは、他のハードウェアと共に又はその一部として供給される光学的記憶媒体又は固体媒体といった適切な媒体に格納／配布されることができるが、インターネット又は他の有線若しくは無線通信システムを介してといった他の形式で配布されることもできる。請求項における任意の参照符号は、発明の範囲を限定するものとして解釈されるべきではない。

【図１】

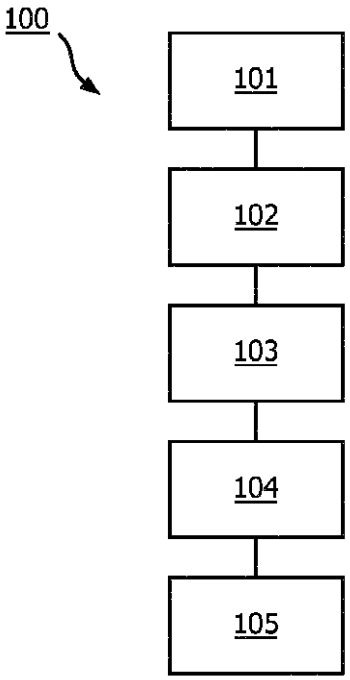
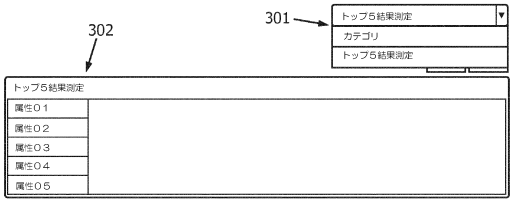
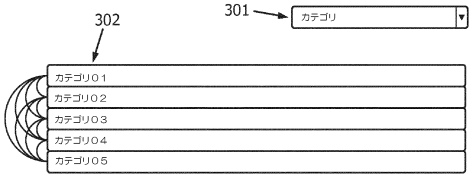


FIG. 1

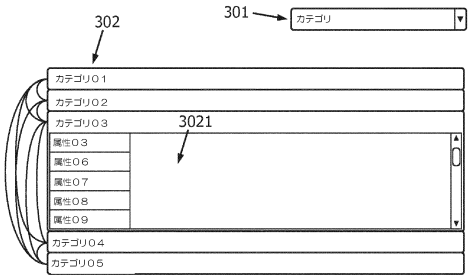
【図２】



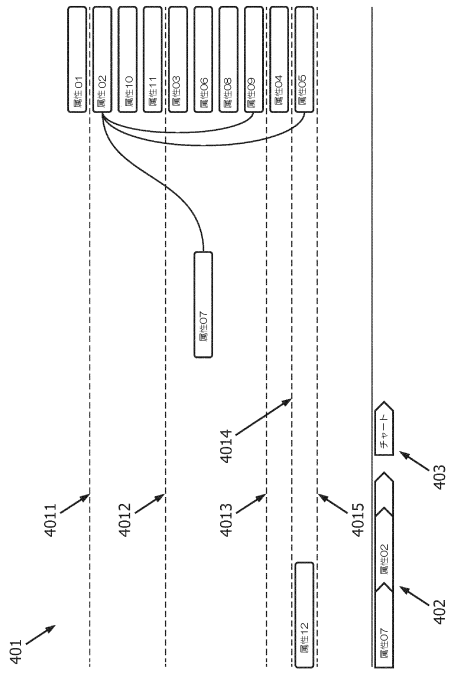
【図３ａ】



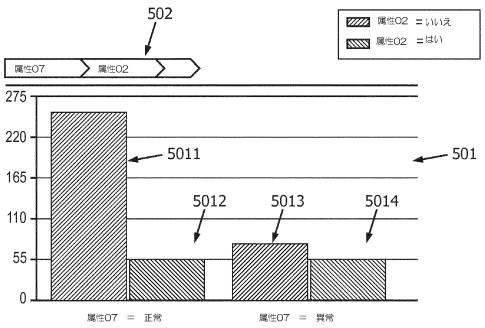
【図３ｂ】



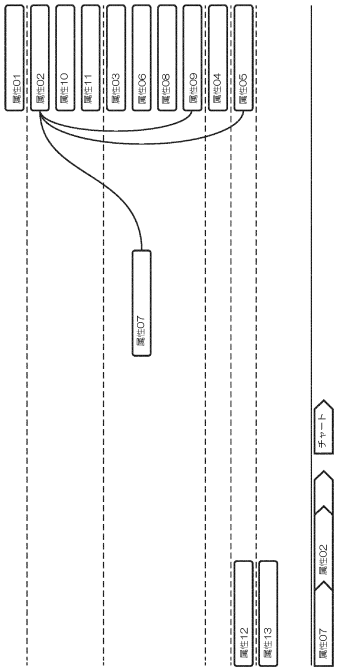
【図４ａ】



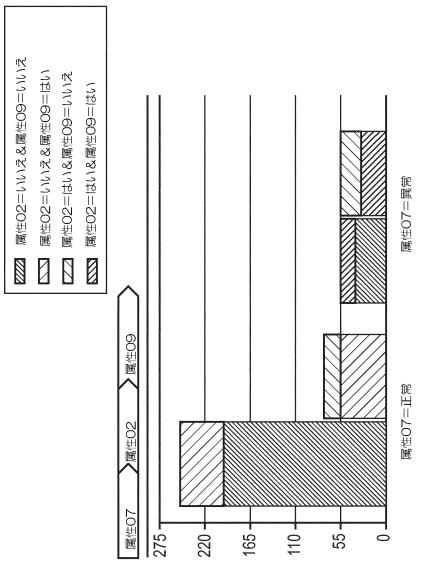
【図４ｂ】



【図５ａ】



【図５ｂ】



【図６】

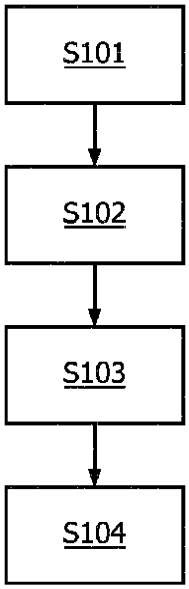


FIG. 6

フロントページの続き

(74)代理人 100163809
弁理士 五十嵐 貴裕
(72)発明者 チャウ チュー チャップ
オランダ国 ５６５６ アーエー アインドーフェン ハイ テック キャンパス ５
(72)発明者 リン チー ジョン
オランダ国 ５６５６ アーエー アインドーフェン ハイ テック キャンパス ５
(72)発明者 チャン タック ミン
オランダ国 ５６５６ アーエー アインドーフェン ハイ テック キャンパス ５
(72)発明者 ジア ユーガン
オランダ国 ５６５６ アーエー アインドーフェン ハイ テック キャンパス ５

審査官 鹿野 博嗣

(56)参考文献 米国特許出願公開第２０１３／０１３８５９２（ＵＳ，Ａ１）
国際公開第２０１０／０６４３２３（WO，Ａ１）
特開２００２－１９７０９８（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ １６／９０３５
Ｇ０６Ｆ １６／２８