

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4363792号
(P4363792)

(45) 発行日 平成21年11月11日(2009.11.11)

(24) 登録日 平成21年8月28日(2009.8.28)

(51) Int.Cl.		F I			
G06F	17/30	(2006.01)	G06F	17/30	370Z
G06T	1/00	(2006.01)	G06F	17/30	350C
			G06F	17/30	380F
			G06T	1/00	200E

請求項の数 9 (全 16 頁)

(21) 出願番号	特願2001-85759 (P2001-85759)	(73) 特許権者	000005223 富士通株式会社
(22) 出願日	平成13年3月23日 (2001.3.23)		神奈川県川崎市中原区上小田中4丁目1番1号
(65) 公開番号	特開2002-288219 (P2002-288219A)	(74) 代理人	100095555 弁理士 池内 寛幸
(43) 公開日	平成14年10月4日 (2002.10.4)	(72) 発明者	上原 祐介 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
審査請求日	平成18年5月18日 (2006.5.18)	(72) 発明者	増本 大器 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
		(72) 発明者	指田 直毅 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

最終頁に続く

(54) 【発明の名称】 情報検索システムおよび方法

(57) 【特許請求の範囲】

【請求項1】

ユーザが複数の検索キーを指定する検索キー指定部と、
各検索キーに対応づけられた特徴量ベクトルを抽出する特徴量抽出部と、
前記複数の検索キーに対応づけられた複数の前記特徴量ベクトルについて、対応し合う要素間の平均距離を比較し、小さい方から規定個の特徴量ベクトルの要素を検索に用いる検索特徴量として選択する検索特徴量選択部と、
前記検索特徴量を検索実行キーとして、情報格納部を検索する検索部と、
各検索キーの特徴量ベクトルの要素のうち、前記検索特徴量として抽出されなかった特徴量ベクトルの要素を分類配置特徴量として選択する分類配置特徴量選択部と、
前記分類配置特徴量を分類配置の特徴として用いて前記検索部の検索により得られたデータを分類配置する分類配置部と、
前記分類配置部による分類配置にしたがって前記データを表示する表示部を備えたことを特徴とする情報検索システム。

【請求項2】

前記分類配置特徴量選択部は、前記分類配置特徴量として、前記検索結果に対する分散が高い特徴量ベクトルの要素を規定個選択する請求項1に記載の情報検索システム。

【請求項3】

前記検索キーの指定において、利用者が、前記指定した各検索キーを平面上に配置し、前記検索特徴量選択部は、検索キー間の特徴量ベクトル要素の比が前記平面上に配置さ

れた検索キー間の相対的距離比と近い特徴量ベクトル要素を規定数選び、前記検索特徴量として選択する請求項 1 に記載の情報検索システム。

【請求項 4】

前記分類配置部は、前記検索キーとして指定された情報の分類配置位置を指定する分類配置位置指定部を備え、

前記検索部により得られたデータの分類配置において、前記分類配置部は、前記検索キーとして指定された情報を、前記分類配置位置指定部で指定された分類配置位置に固定する請求項 1 に記載の情報検索システム。

【請求項 5】

前記検索特徴量選択部と前記分類配置特徴量選択部が、ユーザが選択した検索キーの履歴情報を保持し、参照すべき情報検索処理の範囲の指定を受け、前記検索特徴量選択部は、検索特徴量の選択にあたり、参照すべき情報検索処理の範囲において前記履歴情報が示すユーザが選択した検索キーも加えて検索特徴量を選択し、前記分類配置特徴量選択部は、参照すべき情報検索処理の範囲において前記履歴情報が示すユーザが選択した検索キーも加えて分類配置の特徴として利用する分類配置特徴量を選択する請求項 1 に記載の情報検索システム。

10

【請求項 6】

前記表示部が、分類配置された検索結果を仮想的な三次元空間における任意の表示視点から見た状態として表示する機能と、

利用者による前記表示視点の変更操作を受け付け、仮想的な三次元空間のウォークスルーを可能とする機能を備え、

前記検索キー指定部は、前記表示視点の位置との距離が規定値以内となった情報を利用者が選択した検索キーとみなす請求項 1 に記載の情報検索システム。

20

【請求項 7】

前記検索キー指定部において、利用者が指定する検索キーがサムネイル画像であり、

前記特徴量抽出部は、検索キーとして指定されたサムネイル画像に対応する原画像に対応づけられた画像特徴量を特徴量として抽出する請求項 1 から 6 のいずれかに記載の情報検索システム。

【請求項 8】

コンピュータが実行する情報検索方法であって、

前記コンピュータの検索キー指定部が、ユーザから複数の検索キーの指定を受け、

前記コンピュータの特徴量抽出部が、各検索キーに対応づけられた特徴量ベクトルを抽出し、

30

前記コンピュータの検索特徴量選択部が、前記複数の検索キーに対応づけられた複数の前記特徴量ベクトルについて、対応し合う要素間の平均距離を比較し、小さい方から規定個の特徴量ベクトルの要素を検索に用いる検索特徴量として選択し、

前記コンピュータの検索部が、前記検索特徴量を検索実行キーとして、情報格納部を検索し、

前記コンピュータの分類配置特徴量選択部が、各検索キーの特徴量ベクトルの要素のうち、前記検索特徴量として抽出されなかった特徴量ベクトルの要素を分類配置特徴量として選択し、

40

前記コンピュータの分類配置部が、前記分類配置特徴量を分類配置の特徴として用いて前記検索部の検索により得られたデータを分類配置し、

前記コンピュータの表示部が、前記分類配置部による分類配置にしたがって前記データを表示することを特徴とする情報検索方法。

【請求項 9】

コンピュータに情報検索処理を実行させる情報検索処理プログラムであって、

ユーザから複数の検索キーの指定を受け付ける検索キー指定処理と、

各検索キーに対応づけられた特徴量ベクトルを抽出する特徴量抽出処理と、

前記複数の検索キーに対応づけられた複数の前記特徴量ベクトルについて、対応し合う

50

要素間の平均距離を比較し、小さい方から規定個の特徴量ベクトルの要素を検索に用いる検索特徴量として選択する検索特徴量選択処理と、

前記検索特徴量を検索実行キーとして、情報格納部を検索する検索処理と、

各検索キーの特徴量ベクトルの要素のうち、前記検索特徴量として抽出されなかった特徴量ベクトルの要素を分類配置特徴量として選択する分類配置特徴量選択処理と、

前記分類配置特徴量を分類配置の特徴として用いて前記検索処理により得られたデータを分類配置する分類配置処理と、

前記分類配置処理による分類配置にしたがって前記データを表示する表示処理とをコンピュータに実行させる情報検索処理プログラム。

【発明の詳細な説明】

10

【0001】

【発明の属する技術分野】

本発明は、テキストデータあるいは画像データなどの大量の情報からユーザが所望の情報に近いと思う複数の情報をキーとして指定し、指定された複数の検索キー集合の特徴に基づいて検索し、検索結果を類似した情報同士ほど近くに集まるように表示画面上に配置して表示することで、ユーザが所望の情報を素早く探すことが可能な情報検索システムに関する。

【0002】

【従来の技術】

大量の情報から所望の情報を検索する従来の情報検索システムでは、検索要求と検索対象の情報とを照合し、適合度が高い順に並べたリストとして検索結果を表示する。この場合、適切な検索要求を与えなければ精度の高い検索結果は得られない。検索対象が画像情報である場合などは、一般的に、利用者が適切な検索要求を想起して指定するのは困難であり、曖昧な検索要求を入力することが多い。そのため所望の情報を見つけるためには大量の検索結果を取得して、その中身を目視して判断する必要があり、多大な労力を要していた。

20

【0003】

検索キー選択の技術として、検索結果の情報を各々の特徴を表す特徴量に基づいて仮想的な3次元空間に視覚的に配置表示することで、検索の効率化をはかる技術が特開2000-250943号公報に開示されている。この技術では、表示画面に表示されている情報の中から利用者が選択した一つの検索キーが検索要求として用いられる。この場合、検索キーや検索対象の情報の内容を特徴量ベクトル値として表し、特徴量ベクトル値間の差で情報間の類似度を計算して適合度を求める。その際、特徴量ベクトルの要素は均一に扱われている。

30

【0004】

また、検索結果の分類配置の技術として、2つあるいは3つの特徴量を時間的に変更しながら、それら特徴量が張る軸で構成される2次元あるいは3次元の空間上で情報集合を一覧表示することで、閲覧の効率化をはかる技術が特開2000-29885号公報に開示されている。

【0005】

【発明が解決しようとする課題】

40

利用者が所望の情報に近いとして検索キーを選択するとき、特徴量ベクトルの要素の一部に注目していることが多い。従って、真にユーザが求める情報を高い精度で検索して得るためには、ユーザが注目している特徴量ベクトルの要素を的確に推定し、注目している特徴量ベクトルの要素のみを用いて検索する必要がある。特開2000-250943号公報記載の技術では特徴量ベクトルの要素は均一であるため、適切な検索キーの選択ができない。

【0006】

また、異なる種類の特徴量を用いる場合は、同様にその中でユーザが注目する特徴量のみを用いて検索する必要がある。これに対し、異なる種類の特徴量を用いる場合に、ユーザが選択した複数の検索キーの特徴量の中で、同種の特徴量同士の差を求め、差が小さい特徴量を検索に用いる特徴量とする技術が、特開平11-238078号公報に開示されている。こ

50

ここで、検索結果の一覧性を高め、閲覧効率を向上させるためには、類似した情報同士ほど近くに集まるように配置する必要があり、さらに、情報は検索結果表示画面上において適切に分散していることが好ましい。配置の分散が高いほど、共通の特徴を持った情報の塊が現れやすくなり、閲覧の効率化が図られるためである。しかしながら、上記の特開平11-238078号記載の技術では、検索に用いる特徴量を選別する方法は開示されているものの、検索結果をユーザに表示するための分類配置に用いる特徴量の選別方法には言及していない。分類配置に用いる特徴量として検索に用いた特徴量を選択すると、検索結果は当該特徴量に対して類似しているため適切に検索結果が分散することができない。この問題は、一つの特徴量のみを用いる場合の、特徴量のベクトル表現における要素の選択でも同様である。

10

【0007】

また、上記の特開2000-29885号記載の技術では、分類配置表示において、空間軸に使用する特徴量を時間的に変更することで、特徴量選択の問題に代替的に対処している。しかしながら、特徴量の時間的な変更はユーザの意図とは無関係であり、意図に合った特徴量を選ばれるまで待たされるという問題があり、必ずしも利用者の意図に合った特徴量による検索を行うことはできない。

【0008】

本発明は、上記課題を解決すべく、検索にあたっては、ユーザが選択した複数の検索キーに基づいて検索に用いる情報を適切に選択し、当該選択された情報を用いて検索を実行し、さらに、検索結果の分類配置にあたっては、選択した検索キーに基づいて分類配置に用いる情報を適切に選択して検索結果を分類配置することにより、ユーザの検索意図に合った検索結果表示画面を生成し、効率的な検索を可能とすることを目的とする。

20

【0009】**【課題を解決するための手段】**

上記課題を解決するために、本発明の情報検索システムは、ユーザが複数の検索キーを指定する検索キー指定部と、各検索キーに対応づけられた特徴量ベクトルを抽出する特徴量抽出部と、前記特徴量ベクトルの対応し合う要素間の平均距離を比較し、小さい方から規定個の特徴量ベクトルの要素を検索に用いる検索特徴量として選択する検索特徴量選択部と、前記検索特徴量を検索実行キーとして、情報格納部を検索する検索部と、各検索キーの特徴量ベクトルの要素のうち、前記検索特徴量として抽出されなかった特徴量ベクトルの要素を分類配置特徴量として選択する分類配置特徴量選択部と、前記分類配置特徴量を分類配置の特徴として用いて前記検索結果を分類配置する分類配置部と、前記分類配置結果を表示する表示部を備えたことを特徴とする。

30

【0010】

上記構成により、検索において、特徴量の中でユーザが注目する特徴量ベクトル要素のみを使うことでユーザの意図に合った精度の高い検索が可能であり、かつ、検索結果の分類配置において検索で用いた特徴量ベクトル要素を除いた特徴量ベクトル要素を用いるので適切に検索結果を分散させることができ、一覧性の高い配置で検索結果の表示が可能となる。

【0011】

なお、分類配置特徴量選択部が、分類配置特徴量として、前記検索結果に対する分散が高い特徴量ベクトルの要素を規定個選択することで一覧性の高い配置で検索結果の表示が可能となる。

40

【0012】

次に、検索キーの指定において、利用者が、前記指定した各検索キーを平面上に配置し、前記検索特徴量選択部は、検索キー間の特徴量ベクトル要素の比が前記平面上に配置された検索キー間の相対的な距離比と近い特徴量ベクトル要素を規定数選び、前記検索特徴量として選択することが好ましい。

【0013】

上記構成により、複数の検索キー間の類似関係を検索や分類配置に反映することが可能と

50

なり、検索キーとして指定する特徴量の重み付けを行うことができ、ユーザが特徴量間の類似関係を意図している場合に高い精度で検索することができる。

【0014】

なお、分類配置において、検索キーとして指定した情報の分類配置位置を指定することができれば、ユーザが検索キーとして指定した個々の情報と他の検索結果である情報との関係に興味があるときに、分類配置表示画面における各情報の相対的關係が分かりやすくなる。

【0015】

また、分類配置表示画面において、いわゆるウォークスルー技術を適用できれば、検索結果の閲覧が行い易くなり、また、前記検索キー指定部が、前記表示視点の位置との距離が規定値以内となった配置情報を利用者が選択した検索キーとみなすことにより、検索キー指定が速やかに行うことができる。

10

【0016】

また、検索にあたり、過去の検索処理のうち、参照すべき検索処理の範囲を指定し、その履歴情報を用いつつ検索処理を実行できれば、一度の検索では十分な精度が出ないときでも、いわゆる試行錯誤的に、検索対象を絞り込んで行くことができ、検索精度を上げることができる。

【0017】

なお、上記の構成において、前記検索キー指定部において、利用者が指定する検索キーがサムネイル画像であり、前記特徴量抽出部は、検索キーとして指定されたサムネイル画像に対応する原画像に対応づけられた画像特徴量を特徴量として抽出することが好ましい。

20

【0018】

検索処理においては、サムネイル画像を用いた方が、閲覧性に優れていることが多いからである。

【0019】

なお、上記の情報検索システムを実現する処理ステップを備えた処理プログラムコードを提供することができ、当該処理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から処理プログラムを読み込むことにより、コンピュータを用いて構築することができる。

【0020】

30

【発明の実施の形態】

本発明の実施形態1にかかる情報検索システムを述べる。

【0021】

実施形態1にかかる情報検索システムは、複数の検索キーの特徴量ベクトル要素のうち、対応し合う特徴量ベクトル要素同士の平均距離が小さいものを規定個検索特徴量として選択し、検索特徴量とされた特徴量ベクトル値を検索実行キーとして、情報格納部を検索し、分類配置においては、検索特徴量として抽出されなかった特徴量ベクトル要素を分類配置特徴量として選択して分類配置を行うものである。この例では、テキストデータあるいは画像データからサムネイル画像を生成し、検索結果としてサムネイル画像を分類配置し、利用者の閲覧に供する例とする。

40

【0022】

図1は実施形態1にかかる情報検索システムの構成例を示す図である。

【0023】

情報格納部10は、検索対象となる情報を格納している部分である。この例では検索対象情報がテキストデータあるいは画像データであり、それら情報が格納されている。

【0024】

サムネイル生成部13は、情報格納部10に格納されたテキストデータあるいは画像データからサムネイルを生成する部分である。

【0025】

データを表示するために用いるサムネイルとしては、画像データの場合、画像データその

50

もの、あるいは、その縮小画像がある。また、テキストデータの場合、図書ならば書名、その他文書ならばタイトル、あるいは、ファイル名などがある。これらを含め、データの中身を判断できる情報を視覚的に表すものであればよい。

【 0 0 2 6 】

表示部 1 4 は、検索結果として分類配置された情報を表示する部分である。この例では、検索されたテキストデータあるいは画像データに対応するサムネイルを表示する部分である。また、検索キー指定部 1 5 における検索キーの指定の際にサムネイルを用いる場合、表示部 1 4 に表示されたサムネイルを利用することも可能である。

【 0 0 2 7 】

検索キー指定部 1 5 は、ユーザが複数の検索キーを指定する部分である。この例では、検索キーがサムネイル画像であり、検索キーの指定方法として、表示部 1 4 により表示されたサムネイルから所望の情報に近い複数の情報を選択することとし、検索キー指定部 1 5 は、表示部 1 4 上に表示されたサムネイルを指定する手段を利用者に提供する。

【 0 0 2 8 】

特徴量抽出部 1 6 は、検索キー指定部 1 5 により指定された各検索キーに対応づけられた特徴量を抽出する部分である。この例では、利用者が指定した検索キーとしたサムネイルに対応するテキストデータあるいは画像情報の内容を特徴付ける特徴量ベクトルの値を特徴量として抽出する。

【 0 0 2 9 】

特徴量としては、画像データの場合、画像中の画素について R G B 色空間や H S V 色空間を分割したときの各ブロックに含まれる画素の割合である色ヒストグラムや画像を縦横格子状に分割したときの各ブロックの平均色を並べた色空間分布などがある。また、テキストデータの場合、含まれる単語の頻度分布などがある。これらを含め、データの特徴を表すベクトル値であればよい。

【 0 0 3 0 】

検索特徴量選択部 1 7 は、検索キー指定部 1 5 を介して指定された複数の検索キーの特徴量ベクトル要素のうち、特徴量ベクトルの対応し合う要素間の平均距離を比較し、小さい方から規定個の特徴量ベクトルの要素を検索に用いる検索特徴量として選択する部分である。検索特徴量選択部 1 7 は、複数の検索キー画像の特徴量ベクトル要素同士を比較して平均距離を算出し、その平均距離が小さいものから規定数の特徴量ベクトル要素を、検索特徴量として選択する。

【 0 0 3 1 】

検索部 1 8 は、検索特徴量選択部 1 7 により選択された検索特徴量を検索実行キーとして用いて、情報格納部 1 0 を検索する部分である。この例では、検索特徴量となった特徴量ベクトル要素のみに着目し、情報格納部 1 0 に格納されたテキストデータあるいは画像データの当該特徴量ベクトル値とを比較し、両者間の差が小さい方から順に規定数個のテキストデータあるいは画像データを検索結果として情報格納部 1 0 から取得する。

【 0 0 3 2 】

分類配置特徴量選択部 1 9 は、検索部 1 8 により取得された検索結果の特徴量について、検索特徴量選択部 1 7 により選択された検索特徴量以外の特徴量ベクトル要素に着目し、検索結果の情報集合内での分散が高い方から規定数個の特徴量ベクトル要素を分類配置特徴量として選択する。この例では、検索特徴量として選択されなかった特徴量ベクトル要素のみに着目し、検索結果のテキストデータあるいは画像データ集合内での分散が高い方から規定数個の特徴量ベクトル要素を分類配置特徴量として選択する。

【 0 0 3 3 】

分類配置部 2 0 は、検索部 1 8 により検索された情報を、分類配置特徴量のみに着目して情報間の差が小さいほど近くに集まるよう平面上に配置し、結果をサムネイル生成部 1 3 により生成したサムネイルによって表示部 1 4 を介して利用者に表示する。

【 0 0 3 4 】

次に、上記情報検索システムの処理の流れを図 2 ~ 図 7 のフローチャートを参照しつつ説

10

20

30

40

50

明する。

【 0 0 3 5 】

処理に先立ち、あらかじめ、情報格納部 1 0 に格納されているすべてのテキストデータあるいは画像データから、ベクトル値として数量化された特徴量を抽出して保持しておく。また、処理に先立ち、サムネイル生成手段 1 3 により情報格納部 1 0 に格納されたテキストデータあるいは画像データからサムネイルを生成しておく。

【 0 0 3 6 】

まず、利用者が複数の検索キーを選択する（ステップ S 2 0 1）。このステップ S 2 0 1 の詳細ステップを図 3 のフローチャートとして示している。

【 0 0 3 7 】

この例では、表示部 1 4 によってサムネイルの一覧が表示されている画面上で、検索キーとして選択したいデータを画面上の位置で指定する。表示部 1 4 の画面上での位置の指定方法としては、例えば、操作入力装置としてマウスを用い、マウスをクリックしたときのカーソルの現在位置とする方法がある。検索キー情報として選択するデータを表すサムネイルを表示装置の表示面上の位置で操作入力装置により指定する（ステップ S 2 0 1 - 1）。

【 0 0 3 8 】

検索キー指定部 1 5 は、指定された位置から対応するデータを検索キーとして特定する（ステップ S 2 0 1 - 2）。2 つ以上の検索キーが選択された上で検索キー指定部 1 5 に対して利用者から検索キー選択終了が指示されると検索キー選択処理を終了する（ステップ S 2 0 1 - 3）。

【 0 0 3 9 】

表示部 1 4 上に示される表示画面の例を模式的に示した図が図 8 である。図 8 のように複数のサムネイルが表示され、利用者がポインティングデバイスなどで検索キーとするサムネイルを指定する。この例では、8 0 1、8 0 2、8 0 3 の 3 つのサムネイルが検索キーとして指定されたとする。

【 0 0 4 0 】

次に、検索特徴量を選択する（ステップ S 2 0 2）。このステップ S 2 0 2 の詳細ステップを図 4 のフローチャートとして示している。

【 0 0 4 1 】

まず、検索特徴量選択部 1 7 は、特徴量抽出部 1 6 により抽出された検索キーとして選択された各テキストデータあるいは画像データの特徴量ベクトル値を受け取る（ステップ S 2 0 2 - 1）。なお、処理に先立って、あらかじめ、情報格納部 1 0 に格納されているすべてのテキストデータあるいは画像データから特徴量ベクトル値を保持しているのので当該保持されている特徴量ベクトル値から該当する特徴量ベクトル値を取り出す。

【 0 0 4 2 】

次に、検索特徴量選択部 1 7 は、取り出した特徴量ベクトル値集合に対し、対応する特徴量ベクトル要素ごとに差の平均値を計算する（ステップ S 2 0 2 - 2）。検索特徴量選択部 1 7 は、計算した差の平均値が小さい方から規定個の特徴量ベクトル要素のみを残すように選択する（ステップ S 2 0 2 - 3）。

【 0 0 4 3 】

例えば、いま、選択する特徴量ベクトルの要素数を N とすると、検索キーとして選択されたデータ集合の特徴量ベクトル値について、同じ位置の特徴量ベクトル値の差の平均値を各位置ごとに計算し、小さい方から m 番目までの特徴量ベクトル要素を求める。

【 0 0 4 4 】

例えば、検索キーとされた情報 1 の特徴量ベクトル値が $(D_1(1), D_2(1), D_3(1), D_4(1), D_5(1))$ であり、検索キーとされた情報 2 の特徴量ベクトル値が $(D_1(2), D_2(2), D_3(2), D_4(2), D_5(2))$ であり、検索キーとされた情報 3 の特徴量ベクトル値が $(D_1(3), D_2(3), D_3(3), D_4(3), D_5(3))$ であるとする。情報 1 ~ 3 までの特徴量ベクトルの要素同士を比較して、その平均距離を求める。例えば、各要素同士

10

20

30

40

50

の差の絶対値の平均として(数1)より求める。

【0045】

【数1】

$$(|D_i(1) - D_i(2)| + |D_i(2) - D_i(3)| + |D_i(3) - D_i(1)|) / 3$$

ここで、 $i = 1, \dots, 5$

今、検索特徴量のベクトル要素として選択される規定個が2つとすると、(数1)より求めた平均距離が小さい方から2つの特徴量ベクトル要素が選ばれる。例えば、 D_1, D_5 とする。つまり、情報1から($D_1(1), D_5(1)$)が選択され、情報2から($D_1(2), D_5(2)$)が選択され、情報3から($D_1(3), D_5(3)$)が検索特徴量として選択される。

【0046】

次に、検索部18により選択された検索特徴量を用いて検索を実行する(ステップS203)。このステップS203の詳細ステップを図5のフローチャートとして示している。

【0047】

検索部18は、情報格納部10に格納されているテキストデータあるいは画像データの特徴量ベクトル値を逐次取り出す(ステップS203-1)。取り出した特徴量ベクトル値のうち検索特徴量として選択された特徴量ベクトル要素のもののみを残すよう選択する(ステップS203-2)。検索キーの選択された特徴量ベクトル値と情報格納部10から取り出した各データの選択された特徴量ベクトル値とを照合し、両者の差をユークリッド距離として計算し、各検索キーごとに得られた差を求める(ステップS203-3)。情報格納部10に格納されているすべてのデータについて求めた各検索キーごとの差の平均が小さい順に並べ、先頭から規定個のデータを検索結果とする(ステップS203-4)。

【0048】

図8に示した3つの検索キーから選択された検索特徴量を基に検索した結果を模式的に示した例が図9である。画像要素901がすべての検索結果に含まれている。

【0049】

次に、分類配置特徴量選択部19により分類配置特徴量を選択する(ステップS204)。このステップS204の詳細ステップを図6のフローチャートとして示している。

【0050】

分類配置特徴量選択部19は、検索結果のデータに対応する特徴量ベクトル値から、検索時に選択した特徴量ベクトル要素以外の要素の値のみを残すよう選択する(ステップS204-1)。ここでは、 $D_1(n)$ と $D_5(n)$ が検索特徴量として選択されていたので、 $D_2(n)$ 、 $D_3(n)$ 、 $D_4(n)$ が残る。

【0051】

次に、選択された特徴量ベクトル値の集合において対応する特徴量ベクトル値ごとに分散値を計算する(ステップS204-2)。各特徴量ベクトル要素について分散値の高い方から規定個の特徴量ベクトル要素の値のみを残すよう変更し、分類配置特徴量とする(ステップS204-3)。例えば、規定個を2つとし、 $D_2(n)$ 、 $D_3(n)$ が分散特徴量として選択される。

【0052】

次に、分類配置部20は、分類配置特徴量を用いて検索結果を分類配置する(ステップS205)。このステップS205の詳細ステップを図7のフローチャートとして示している。分類配置部20は、検索結果のデータ集合について対応する特徴量ベクトル値間の差が小さいほど近くに集まるような平面上の配置を計算する(ステップS205-1)。分類配置部20は、分類配置結果を元にサムネイル生成手段13によって生成された各データに対応するサムネイルの表示平面上での配置位置を決める(ステップS205-2)。例えば、特徴量ベクトル値間の差が小さいほど近くに集まるような平面上の配置の計算方法としては、自己組織化マップ、多次元尺度構成法などを用いる。

【0053】

分類配置部20は、分類配置データを表示部14に対して出力する(ステップS205-

10

20

30

40

50

3)。

【0054】

次に、表示部14は、分類配置結果に基づきサムネイルを用いて検索結果を表示する(ステップS206)。

【0055】

図10は、図9に示したように得られた検索結果を分類配置特徴量を分類配置の特徴として自己組織化マップにより分類配置した模式図である。画像要素901を含み、さらに他の画像要素の有無において似通った画像同士が近くに分類配置されている様子が分かる。

【0056】

以上、ステップS201からステップS206の処理により、情報分類処理を実行する。

10

【0057】

ここで、ステップS201からステップ206を繰り返すことが可能であり、このとき、検索特徴量選択部17および分類配置特徴量選択部19は、ユーザが選択した検索キーの履歴を保持すれば、特徴量の選択処理の時に過去の特徴量選択の情報を利用できる。つまり、過去の情報検索処理のうち、参照すべき情報検索処理の範囲の指定を受け、検索特徴量選択部17は、検索特徴量の選択にあたり、参照の指定を受けた検索処理において利用された情報も加えて検索特徴量を選択し、分類配置特徴量選択部19は、参照指定を受けた検索処理において分類配置の特徴として利用された分類配置特徴量も加えて分類配置の特徴として利用する分類配置特徴量を選択する。

【0058】

20

また、表示部14が、分類配置された検索結果を仮想的な三次元空間における任意の表示視点から見た状態として表示する機能と、利用者による表示視点の変更操作を受け付け、仮想的な三次元空間のウォークスルーを可能とする機能を備えることができる。さらに、このとき、検索キー指定部15は、サムネイルに対する視点の位置を表示空間における規定距離内に移動することで、検索キーの選択操作とすることも可能である。また、検索キー指定部15は、視点の位置がサムネイルの表示位置を画面の奥行き方向に突き抜けることで検索キーの選択とみなすことにより、新たにステップS201からステップ206を繰り返すことが可能である。

【0059】

以上、本実施形態1の情報検索システムによれば、検索において特徴量の要素の中でユーザが注目する要素のみを使うことでユーザの意図に合った精度の高い検索が可能であり、かつ、検索結果の分類配置において検索で用いた要素を除いた特徴量の中で分散の高い要素のみを用いることで高速かつ一覧性の高い配置で表示が可能となり、大量な情報を効率的に検索することができる。

30

【0060】

(実施形態2)

実施形態2の情報検索装置は、検索キーの指定において、利用者が、前記指定した各検索キーを平面上に配置し、検索特徴量選択部は、検索キー間の特徴量の値の比が前記平面上に配置された検索キー間の配置平面上での相対的な距離比に近い特徴量を規定個選び、前記検索特徴量として選択するものである。

40

【0061】

この例では、利用者は検索キー指定部15により複数の検索キーを選択後、検索キーに対応するサムネイルを表示部14の表示平面上の任意の位置に配置することで、検索キー間の類似関係を指定できることとする。図11は、3つの検索キーA、B、Cを配置した例である。この例では、利用者は検索キーAとBの間で近く、AとC、あるいは、AとBの間では近くないような類似関係を指定している。

【0062】

検索特徴量選択部17は、検索特徴量となる特徴量ベクトル要素の選択または検索特徴量となる種類の特徴量の選択において、まず、対応する特徴量間の差の平均値が規定値以内の要素を求め、次に、その中で、対応する要素間の差の比を求める。この差の比が、利用

50

者が配置したサムネイル間の画面上の差の比と近いほうから規定個の要素を選択する。例えば、いま、図 1 1 のように、利用者が画面上に配置した検索キー A B 間の距離が L_{AB} であり、B C 間の距離が L_{BC} であり、C A 間の距離が L_{CA} であったとする。また、検索キー A B 間の特徴量ベクトル値の i 番目の要素の差が li_{AB} であり、B C 間の要素の差が li_{BC} であり、C A 間の要素の差が li_{CA} であったとする。このとき、 i 番目の要素に対する評価値 E_i を (数 2) より計算する。この評価値が小さい方から規定個の要素を選択する。

【 0 0 6 3 】

【 数 2 】

$$E_i = \left| \frac{L_{BC}}{L_{AB}} - \frac{li_{BC}}{li_{AB}} \right| + \left| \frac{L_{CA}}{L_{AB}} - \frac{li_{CA}}{li_{AB}} \right|$$

10

【 0 0 6 4 】

検索特徴量の選択処理以降の、検索部 1 8 での検索処理、分類配置部 2 0 による分類配置処理、表示部 1 4 での表示処理などは、実施形態 1 と同様で良い。

【 0 0 6 5 】

以上のように、本実施形態 2 の情報検索システムによれば、複数の検索キー間の類似関係を検索特徴量の選択に反映することが可能となり、検索キーの配置を通じて検索キーの特徴量に対する重み付けを行うことができ、検索キー間の類似関係を利用者の意図に従って検索処理に反映させ、高い精度で検索することができる。

20

【 0 0 6 6 】

(実施形態 3)

実施形態 3 の情報検索システムは、分類配置部が、検索キーとして指定された情報の分類配置位置を指定する分類配置位置指定部を備えたものであり、検索結果の分類配置において、分類配置位置指定部の指定に従い、検索キーとなった情報の分類位置を指定位置とすることができるシステムである。

【 0 0 6 7 】

このように、分類配置において、検索キーとして指定した情報の分類配置位置を指定することができれば、ユーザが検索キーとして指定した個々の情報と他の検索結果である情報との関係に興味があるときに、分類配置表示画面における各情報の相対的關係が分かりやすくなる。

30

【 0 0 6 8 】

図 1 2 は、実施形態 3 の情報検索システムの構成例を示す図である。

【 0 0 6 9 】

分類配置部 2 0 a は、分類配置位置指定部 2 1 を備えている。利用者は、分類配置位置指定部 2 1 を介して、検索キーとして選択したデータのうち分類配置位置を固定したい検索キーを指定し、さらにその表示装置 1 4 上での表示位置を指定する。この検索キーの指定と表示位置の指定のタイミングは、例えば、検索キー指定部 1 5 を介した検索キーの指定時とする。

40

【 0 0 7 0 】

検索部 1 8 による検索処理後、分類配置部 2 0 は、検索結果として得られた情報を分類配置特徴量により分類配置するが、分類配置位置の固定指定された情報は、その分類配置位置に固定し、他の情報を自己組織化マップなどによる分類配置を行う。

【 0 0 7 1 】

例えば、実施形態 2 において説明した、複数の検索キーを平面上に配置することにより検索キーの特徴量に対する重み付けを行う仕組みと組み合わせた例を示す。利用者が図 1 1 に示したように、3 つの検索キー A , B , C を配置し、さらに、分類配置位置指定部 2 1

50

により、これら3つの検索キーA、B、Cの分類配置位置を図11に示した相対位置において固定する指定を行う。図13は当該指定を受けた場合の分類配置の結果を示す図である。3つの検索キーA、B、Cのサムネイルが図11に示した相対位置において固定されており、他のサムネイルが自己組織化マップにより分類配置されている。

【0072】

(実施形態4)

本発明の利用者インタフェースシステムは、上記に説明した構成を実現する処理ステップを記述したプログラムをコンピュータ読み取り可能な記録媒体に記録して提供することにより、各種コンピュータを用いて構築することができる。本発明の利用者インタフェースシステムを実現する処理ステップを備えたプログラムを記録した記録媒体は、図14に図示した記録媒体の例に示すように、CD-ROM1002やフレキシブルディスク1003等の可搬型記録媒体1001だけでなく、ネットワーク上にある記録装置内の記録媒体1000や、コンピュータのハードディスクやRAM等の記録媒体1005のいずれであっても良く、プログラム実行時には、プログラムはコンピュータ1004上にローディングされ、主メモリ上で実行される。

【0073】

【発明の効果】

本発明の情報検索システムによれば、検索処理において、特徴量の中でユーザが注目する特徴量ベクトル要素のみを使うことでユーザの意図に合った精度の高い検索が可能であり、かつ、検索結果の分類配置において検索で用いた特徴量ベクトル要素を除いた特徴量ベクトル要素を用いるので適切に検索結果を分散させることができ、一覧性の高い配置で検索結果の表示が可能となる。

【0074】

本発明の情報検索システムによれば、検索処理において、特徴量の中でユーザが注目する種類の特徴量のみを使うことでユーザの意図に合った精度の高い検索が可能であり、かつ、検索結果の分類配置において検索で用いた特徴量を除いた特徴量のみを用いるので検索結果が適切に分散することとなり、一覧性の高い配置で検索結果の表示が可能となり、大量な情報を効率的に検索することができる。

【0075】

また、分類配置特徴量のうち、検索結果に対する分散が高い特徴量ベクトル要素または特徴量の種類を選択して分類配置実行に用いるので、一覧性の高い配置で検索結果の表示が可能となる。

【図面の簡単な説明】

【図1】 本発明の実施形態1にかかる情報検索システムの構成例を示す図

【図2】 本発明の情報検索システムの処理の流れを示すフローチャート

【図3】 図2のステップS201の詳細ステップを示したフローチャート

【図4】 図2のステップS202の詳細ステップを示したフローチャート

【図5】 図2のステップS203の詳細ステップを示したフローチャート

【図6】 図2のステップS204の詳細ステップを示したフローチャート

【図7】 図2のステップS205の詳細ステップを示したフローチャート

【図8】 表示部14上に示される表示画面の例を模式的に示した図

【図9】 図8に示した3つの検索キーから選択された検索特徴量を基に検索した結果を模式的に示した図

【図10】 図9に示したように得られた検索結果を、分類配置特徴量を分類配置の特徴として自己組織化マップにより分類配置した模式図

【図11】 3つの検索キーA、B、Cを配置した例を示す図

【図12】 本発明の実施形態3にかかる情報検索システムの構成例を示す図

【図13】 指定を受けた場合の分類配置の結果を示す図

【図14】 本発明の実施形態4にかかる利用者インタフェースシステムを実現する処理プログラムを記録した記録媒体を示した図

10

20

30

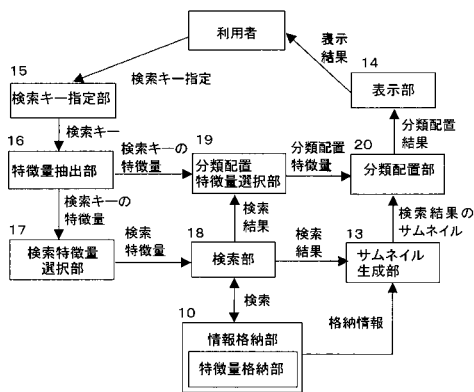
40

50

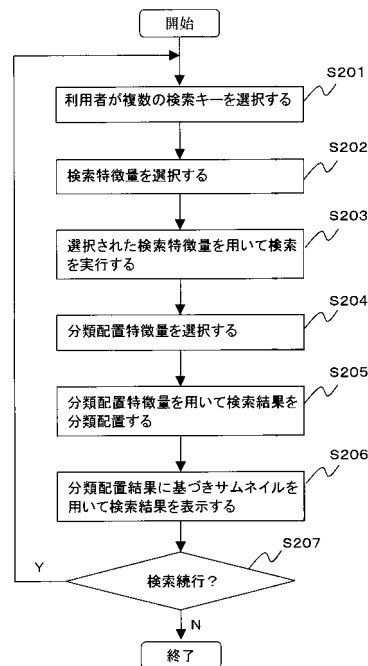
【符号の説明】

- 10 情報格納部
- 13 サムネイル生成部
- 14 表示部
- 15 検索キー指定部
- 16 特徴量抽出部
- 17 検索特徴量選択部
- 18 検索部
- 19 分類配置特徴量選択部
- 20, 20a 分類配置部
- 21 分類配置位置指定部
- 1000 ネットワーク上にある記録装置内の記録媒体
- 1001 可搬型記録媒体
- 1002 CD-ROM
- 1003 フレキシブルディスク
- 1004 コンピュータ
- 1005 コンピュータのハードディスクやRAM等の記録媒体

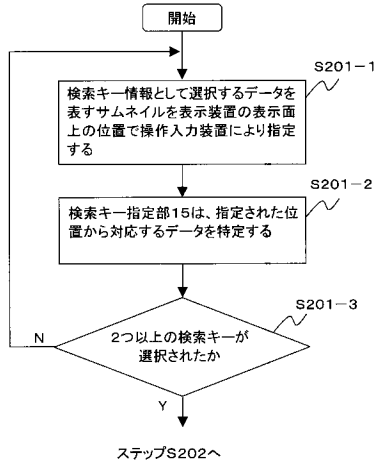
【図1】



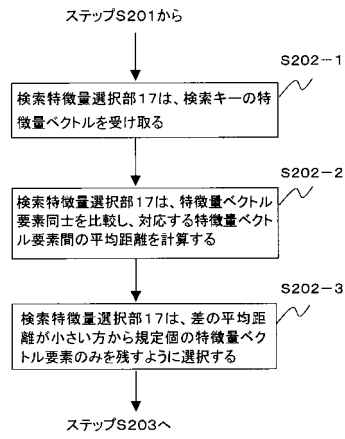
【図2】



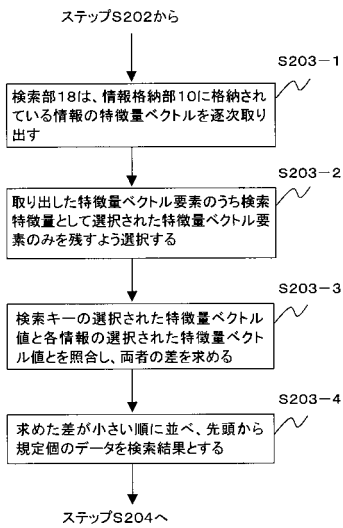
【図3】



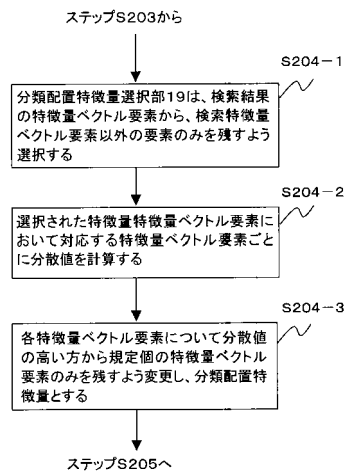
【図4】



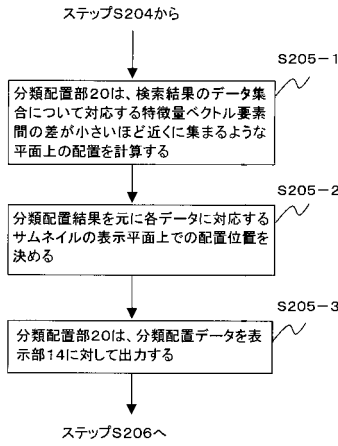
【図5】



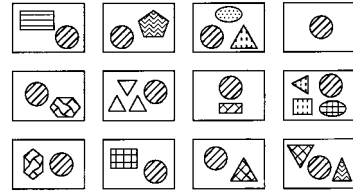
【図6】



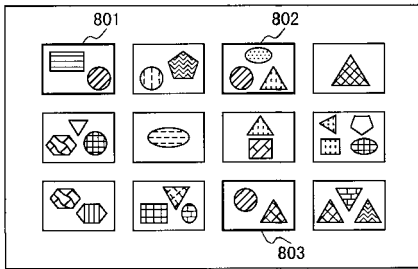
【図7】



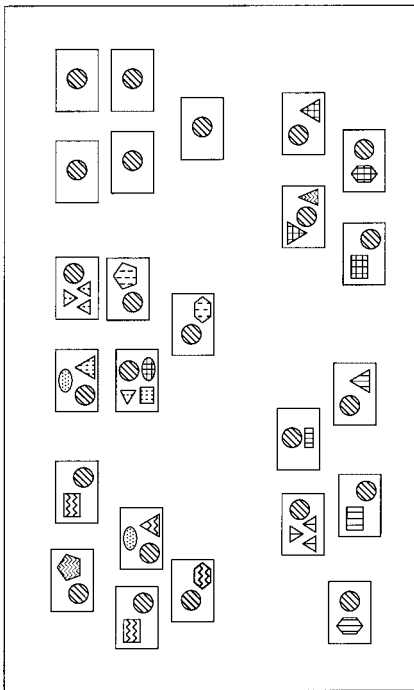
【図9】



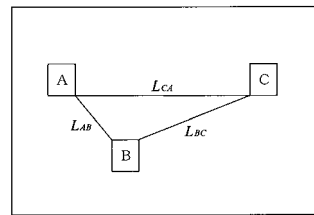
【図8】



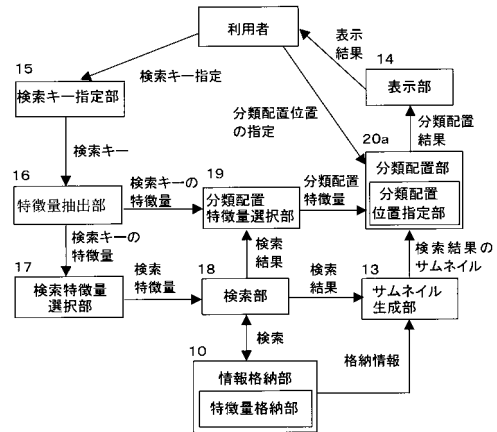
【図10】



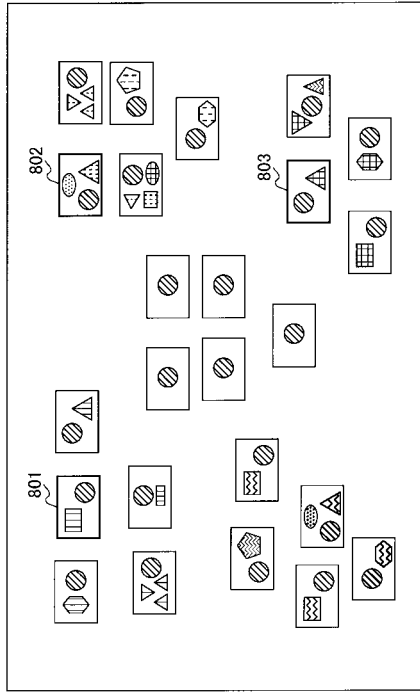
【図11】



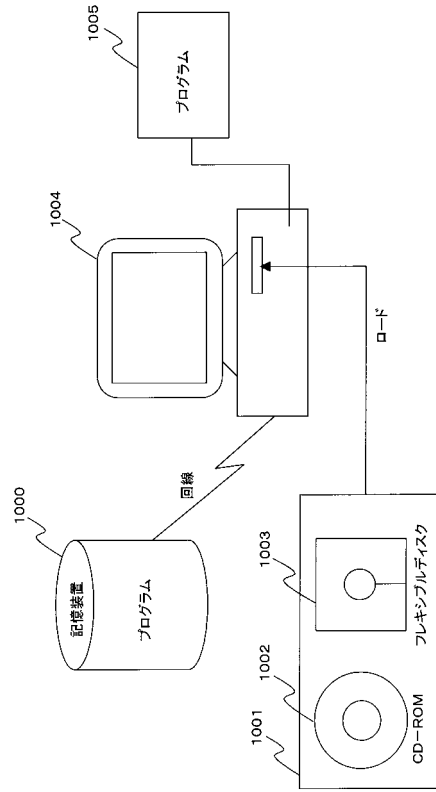
【図12】



【 図 1 3 】



【 図 1 4 】



フロントページの続き

- (72)発明者 椎谷 秀一
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
- (72)発明者 遠藤 進
神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

審査官 波内 みさ

- (56)参考文献 特開平11-238078(JP,A)
特開平09-204440(JP,A)
特開2000-242648(JP,A)
特開平10-240765(JP,A)
特開平11-175535(JP,A)
特開平09-259130(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 17/30

G06T 1/00