

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第4579638号  
(P4579638)

(45) 発行日 平成22年11月10日(2010.11.10)

(24) 登録日 平成22年9月3日(2010.9.3)

(51) Int.Cl.

F 1

GO6F 17/30	(2006.01)	GO6F 17/30	380E
G10L 15/00	(2006.01)	G06F 17/30	310Z
G10L 15/22	(2006.01)	G06F 17/30	350C
HO4N 5/225	(2006.01)	G10L 15/00	200P
HO4N 101/00	(2006.01)	G10L 15/22	460Z

請求項の数 7 (全 11 頁) 最終頁に続く

(21) 出願番号	特願2004-292606 (P2004-292606)
(22) 出願日	平成16年10月5日 (2004.10.5)
(65) 公開番号	特開2006-107108 (P2006-107108A)
(43) 公開日	平成18年4月20日 (2006.4.20)
審査請求日	平成19年10月3日 (2007.10.3)

(73) 特許権者	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(74) 代理人	100126240 弁理士 阿部 琢磨
(74) 代理人	100124442 弁理士 黒岩 創吾
(72) 発明者	久保山 英生 東京都大田区下丸子3丁目30番2号キヤノン株式会社内

審査官 長 由紀子

最終頁に続く

(54) 【発明の名称】データ検索装置及びデータ検索方法

## (57) 【特許請求の範囲】

## 【請求項 1】

各々のデータが所定の音声データと対応付けて蓄積されている複数のデータから、所望のデータを検索し、検索結果を表示手段に表示させるデータ検索装置のデータ検索方法であって、

前記複数のデータと、前記複数のデータの各々に対応付けられた所定の音声データとを蓄積する蓄積ステップと、

各々の前記音声データを音声認識することによって得られた第1の音素列を取得する取得ステップと、

ユーザによる操作に応答して、検索条件に相当する検索キーを入力する入力ステップと、

前記検索キーを形態素解析して単語列に分割し、さらに当該単語列に読みを付与し、第2の音素列を得る変換ステップと、

各々の前記音声データから得られた前記第1の音素列について、音素マッチングを行うことにより、前記第2の音素列との相関度をそれぞれ決定する決定ステップと、

前記相関度が所定の閾値以上である前記第1の音素列に対応する音声データが対応付けられた前記データを前記相関度でランキングした順位順に並べて表示させると共に、前記相関度が前記閾値未満である前記第1の音素列に対応する音声データが対応付けられた前記データを、前記データの名称の名前順、前記データが有する時間情報順、前記データのデータサイズ順、前記データの表示サイズ順のうちいずれかに従って並べて表示させるよ

10

20

う前記表示手段を制御する表示制御ステップを有するデータ検索方法。

【請求項 2】

各々のデータが所定の音声データと対応付けて蓄積されている複数のデータから、所望のデータを検索し、検索結果を表示手段に表示させるデータ検索装置のデータ検索方法であって、

前記複数のデータと、前記複数のデータの各々に対応付けられた所定の音声データとを蓄積する蓄積ステップと、

各々の前記音声データを音声認識することによって得られた第1の単語列を取得する取得ステップと、

ユーザによる操作に応答して、検索条件に相当する検索キーを入力する入力ステップと

10

前記検索キーを形態素解析して第2の単語列を得る変換ステップと、

各々の前記音声データから得られた前記第1の単語列について、単語マッチングを行うことにより、前記第2の単語列との相関度をそれぞれ決定する決定ステップと、

前記相関度が所定の閾値以上である前記第1の単語列に対応する音声データが対応付けられた前記データを前記相関度でランキングした順位順に並べて表示させると共に、前記相関度が前記閾値未満である前記第1の単語列に対応する音声データが対応付けられた前記データを、前記データの名称の名前順、前記データが有する時間情報順、前記データのデータサイズ順、前記データの表示サイズ順のうちいずれかに従って並べて表示させるよう前記表示手段を制御する表示制御ステップを有するデータ検索方法。

20

【請求項 3】

前記表示制御ステップは、前記相関度が所定の閾値以上である前記第1の音素列に対応する音声データが対応付けられた前記データを前記相関度でランキングした順位順に並べて第1のウインドウに表示させると共に、前記相関度が前記閾値未満である前記第1の音素列に対応する音声データが対応付けられた前記データを、前記データの名称の名前順、前記データが有する時間情報順、前記データのデータサイズ順、前記データの表示サイズ順のうちいずれかに従って並べて前記第1のウインドウと異なる第2のウインドウに表示させるよう前記表示手段を制御することを特徴とする請求項1に記載のデータ検索方法。

【請求項 4】

請求項1乃至は請求項3のいずれか1項に記載のデータ検索方法をコンピュータに実行させるプログラム。

30

【請求項 5】

請求項4に記載のプログラムを記憶したコンピュータ読み取り可能な記憶媒体。

【請求項 6】

各々のデータが所定の音声データと対応付けて蓄積されている複数のデータから、所望のデータを検索し、検索結果を表示手段に表示させるデータ検索装置であって、

前記複数のデータと、前記複数のデータの各々に対応付けられた所定の音声データとを蓄積する蓄積手段と、

各々の前記音声データを音声認識することによって得られた第1の音素列を取得する取得手段と、

40

ユーザによる操作に応答して、検索条件に相当する検索キーを入力する入力手段と、

前記検索キーを形態素解析して単語列に分割し、さらに当該単語列に読みを付与し、第2の音素列を得る変換手段と、

各々の前記音声データから得られた前記第1の音素列について、音素マッチングを行うことにより、前記第2の音素列との相関度をそれぞれ決定する決定手段と、

前記相関度が所定の閾値以上である前記第1の音素列に対応する音声データが対応付けられた前記データを前記相関度でランキングした順位順に並べて表示させると共に、前記相関度が前記閾値未満である前記第1の音素列に対応する音声データが対応付けられた前記データを、前記データの名称の名前順、前記データが有する時間情報順、前記データのデータサイズ順、前記データの表示サイズ順のうちいずれかに従って並べて表示させるよ

50

う前記表示手段を制御する表示制御手段を有するデータ検索装置。

【請求項 7】

各々のデータが所定の音声データと対応付けて蓄積されている複数のデータから、所望のデータを検索し、検索結果を表示手段に表示させるデータ検索装置であって、

前記複数のデータと、前記複数のデータの各々に対応付けられた所定の音声データとを蓄積する蓄積手段と、

各々の前記音声データを音声認識することによって得られた第1の単語列を取得する取得手段と、

ユーザによる操作に応答して、検索条件に相当する検索キーを入力する入力手段と、

前記検索キーを形態素解析して第2の単語列を得る変換手段と、

各々の前記音声データから得られた前記第1の単語列について、単語マッチングを行うことにより、前記第2の単語列との相関度をそれぞれ決定する決定手段と、

前記相関度が所定の閾値以上である前記第1の単語列に対応する音声データが対応付けられた前記データを前記相関度でランキングした順位順に並べて表示させると共に、前記相関度が前記閾値未満である前記第1の単語列に対応する音声データが対応付けられた前記データを、前記データの名称の名前順、前記データが有する時間情報順、前記データのデータサイズ順、前記データの表示サイズ順のうちいずれかに従って並べて表示させるよう前記表示手段を制御する表示制御手段を有するデータ検索装置。

【発明の詳細な説明】

【技術分野】

10

【0001】

本発明は、データに付与された音声アノテーションデータを音声認識して求めた音声認識結果アノテーションデータに基づき、データを検索する装置や方法に関するものである。

【背景技術】

【0002】

昨今、デジタルカメラ等の普及が著しい。ユーザは、デジタルカメラのような携帯型撮像装置により撮像したデジタル画像を、PCやサーバ等で管理することが一般的である。例えば撮影した画像を、PCあるいはサーバ上のフォルダ内に整理したり、特定の画像を印刷し、年賀状に組み込むことが可能である。また、サーバで管理する場合は、一部の画像を他のユーザに公開することも可能である。

30

【0003】

このような作業を行う場合には、ユーザの意図する特定の画像を見つけることが必要となる。見つける対象となる画像数が少ない場合は、画像をサムネイル表示し、その一覧から目視で見つけることも可能である。しかし、対象となる画像数が何百となる場合や、対象画像群が複数フォルダに分断されて格納されている場合は、目視で見つけることは困難である。

【0004】

そこで、撮像装置上で画像に音声アノテーション（音声による注釈）を付け、検索時のその情報を使うことが行われている。例えば携帯型撮像装置により山の画像を撮像し、その画像に対して「箱根の山」と発声する。この音声データは先の画像データと対となって撮像装置内に格納された後、その画像撮像装置内あるいは画像をアップロードしたPC内で音声認識され、「はこねのやま」というテキスト情報に変換する。音声アノテーションデータがテキスト情報に変換されれば、後は一般的なテキスト検索技術で処理することができる。

40

【0005】

このような音声アノテーションを利用した先行技術に、特許文献1、特許文献2、特許文献3がある。これらの先行技術では、画像の撮像時あるいは撮像後に注釈となる音声をユーザが入力し、その音声データを既存の音声認識技術を利用して画像検索に利用している。

50

【特許文献1】特開2003-219327号公報

【特許文献2】特開2002-325225号公報

【特許文献3】特開平9-135417号公報

【発明の開示】

【発明が解決しようとする課題】

【0006】

音声認識で音声アノテーションを変換して検索する時には誤認識は避けられない。誤認識による誤りの割合が大きい場合には検索キーを正しく入れてもマッチングの相関が悪く、正しく検索されない。しかしながら、アノテーションのごく一部が誤認識により誤っていても、大部分が正解していればしばしば正しく検索することができる。

10

【0007】

ここで、検索キーとのマッチングの相関度をスコアとして検索画像をランクインして並べると、検索キーとの相関が高いアノテーションについてはランクインの上位に正しく検索されることが多いが、誤認識などにより検索キーとの相関が低い画像は、その他のアノテーションとの差がつき難くなり、急激に順位が落ちる。このように相関度が低い画像については順位順に並べるとその中から探すのが困難であり、むしろ従来のフォルダのように名前順、時間順などで整列させたり、検索結果を表示しないで検索失敗を通知するほうが好ましい。

【課題を解決するための手段】

【0008】

上記課題を解決するために、本発明のデータ検索方法は、各々のデータが所定の音声データと対応付けて蓄積されている複数のデータから、所望のデータを検索し、検索結果を表示手段に表示させるデータ検索装置のデータ検索方法であって、前記複数のデータと、前記複数のデータの各々に対応付けられた所定の音声データとを蓄積する蓄積ステップと、各々の前記音声データを音声認識することによって得られた第1の音素列を取得する取得ステップと、ユーザによる操作に応答して、検索条件に相当する検索キーを入力する入力ステップと、前記検索キーを形態素解析して単語列に分割し、さらに当該単語列に読みを付与し、第2の音素列を得る変換ステップと、各々の前記音声データから得られた前記第1の音素列について、音素マッチングを行うことにより、前記第2の音素列との相関度をそれぞれ決定する決定ステップと、前記相関度が所定の閾値以上である前記第1の音素列に対応する音声データが対応付けられた前記データを前記相関度でランクインした順位順に並べて表示させると共に、前記相関度が前記閾値未満である前記第1の音素列に対応する音声データが対応付けられた前記データを、前記データの名称の名前順、前記データが有する時間情報順、前記データのデータサイズ順、前記データの表示サイズ順のうちいずれかに従って並べて表示させるよう前記表示手段を制御する表示制御ステップを有する。また、本発明のデータ検索方法は、各々のデータが所定の音声データと対応付けて蓄積されている複数のデータから、所望のデータを検索し、検索結果を表示手段に表示させるデータ検索装置のデータ検索方法であって、前記複数のデータと、前記複数のデータの各々に対応付けられた所定の音声データとを蓄積する蓄積ステップと、各々の前記音声データを音声認識することによって得られた第1の単語列を取得する取得ステップと、ユーザによる操作に応答して、検索条件に相当する検索キーを入力する入力ステップと、前記検索キーを形態素解析して第2の単語列を得る変換ステップと、各々の前記音声データから得られた前記第1の単語列について、単語マッチングを行うことにより、前記第2の単語列との相関度をそれぞれ決定する決定ステップと、前記相関度が所定の閾値以上である前記第1の単語列に対応する音声データが対応付けられた前記データを前記相関度でランクインした順位順に並べて表示させると共に、前記相関度が前記閾値未満である前記第1の単語列に対応する音声データが対応付けられた前記データを、前記データの名称の名前順、前記データが有する時間情報順、前記データのデータサイズ順、前記データの表示サイズ順のうちいずれかに従って並べて表示させるよう前記表示手段を制御する表示制御ステップを有する。

20

30

40

50

**【発明の効果】****【0009】**

以上に述べたとおり、本発明におけるデータ検索では、データに対応する音声にデータを音声認識した結果である音声認識結果アノテーションデータと検索条件との相関をあらわすスコアによって、検索結果の表示方法を切り替えることにより、高い順位にあらわれるデータと高い順位に現れないデータを区別して探し出すことができ、ユーザの利便性向上につながる。

**【発明を実施するための最良の形態】****【0010】**

以下、図面を参照して本発明の一実施形態を詳細に説明する。

10

**【実施例】****【0011】**

図1は、本発明の一実施例に係るデータ検索装置の機能を示すブロック図である。同図において、101はデータベースである。102は、データベース101に蓄積される画像やドキュメントなどのデータである。103は、データ102に対応して音声によって注釈をつけた音声アノテーションデータである。104は、音声アノテーションデータ103を音声認識し、音素列や単語列などに変換した音声認識結果アノテーションデータである。105は、所望のデータ102を検索するために検索条件として、検索キーを入力する検索キー入力部である。106は、検索キーでマッチングを取るために、音声認識結果アノテーションデータ104と同様の書式の音素列や単語列に検索キーを変換する検索キー変換部である。

20

**【0012】**

107は、データベース101内の複数の音声認識結果アノテーションデータ104と検索キーとでマッチングを取り、それぞれの音声認識結果アノテーションデータ104に対する相関度スコアを求め、音声認識結果アノテーションデータ104に対応するデータ102を順位付けする検索部である。108は、それぞれのデータの相関度スコアに基づいてデータの表示方法を切り替える表示切り替え部である。109は、表示切り替え部108でそれぞれ指定された方法に基づいてデータ102を表示する表示部である。

**【0013】**

図1を用いて、本実施例の処理の流れを詳細に説明する。画像やドキュメントなどのデータ102には、対応する音声アノテーションデータ103及びこの音声アノテーションデータを音声認識した結果である音声認識結果アノテーションデータ104が存在する。ここで、音声認識結果アノテーションデータは、本装置に音声認識部を有しておいて作成しても構わないし、画像を撮るカメラなどの別の装置上で音声認識部を有して作成しても構わない。また、本発明においてデータ検索で使用するアノテーションデータは音声認識結果アノテーションデータ104であるため、音声アノテーションデータ103は存在しなくとも構わない。

30

**【0014】**

図2に、音声認識結果アノテーションデータ104の一例を示す。同図の201は、音声アノテーションデータ103を音声認識して音素列に変換した認識結果音素列であり、尤もらしいとされる上位5位までが順番に並んでいる。202は音声認識に用いた文法名である。本実施例では音素列に変換する文法を用いて説明するが、単語列に変換する文法を用いても良い。203は、音声に対する音素列の尤もらしさを表す認識尤度である。本実施例におけるデータ検索ではこれらの情報のうち、認識結果音素列201のみ用いるため、文法名202、認識尤度203はなくとも良い。

40

**【0015】**

まず、ユーザは検索キー入力部105に検索条件として検索キーを入力する。図3に検索キー入力部においてユーザに提示する検索キー入力ダイアログの一例を示す。このようなダイアログに、ユーザは検索したいデータに付与されている音声もしくはその一部に相当する単語、文をテキスト入力する。検索キーを入力して検索ボタンを押すと、検索キー

50

は検索キー変換部 106 にわたり、検索キーが認識結果音素列 201 と同じ形式の音素列に変換される。図 4 は検索キーを音素列に変換する様子を表す図である。検索キー「箱根の山」を、形態素解析し、単語列に分割する。さらに、単語列に読みを付与し、音素列を得る。形態素解析、読み付与の方法は一般的な自然言語処理技術を適用する。

#### 【0016】

次に、検索部 107 において検索キーの音素列と、検索対象となる全てのデータ 102 に対応する音声認識結果アノテーションデータ 104 とで音素マッチングを取り、検索キーとの相関度を表す音素正解精度を求める。マッチングは、一般的な DP マッチング法などにより取ることができる。図 5 に音素正解精度を求める様子を表す。音素正解精度は、DP マッチング等により正解数、挿入誤り数、削除誤り数、置換誤り数が得られたときに

10

、  
 $\{(\text{正解数}) - (\text{挿入誤り数}) - (\text{削除誤り数}) - (\text{置換誤り数})\} \times 100 / (\text{正解数})$

で求める。図 5 においては挿入誤りが “o” と “a” の 2箇所、そして “h” を “f” として誤った置換誤りが 1箇所であり、音素正解精度は 75 % となる。このようにして求めた音素正解精度を検索に用いるスコアとしてデータ 102 をランキングする。ここで、図 2 の音声認識結果アノテーションデータは上位 5 つの認識結果音素列が存在するが、マッチングはこれらそれぞれと行って音素正解精度を求め、最も良い音素正解精度及び認識結果音素列を採用する。ただし本発明はこれに限るものではなく、順位によって重み係数を音素正解精度にかけてから最大値を取ったり、総和を取ったりしても良い。また、音声認識結果アノテーションデータは、図 2 のように上位 N 個の認識結果を保持する形態に限らず、各音素（あるいは単語など）で構成するラティス（単語グラフ）を出力し、ラティスの始端から終端までのそれぞれのパスにおいて音素正解精度を求めても良い。

20

#### 【0017】

次に、表示切り替え部 108 において、それぞれのデータ 102 に対応する音素正解精度を閾値と比較し、音素正解精度が閾値以上のデータについては表示部 109 において音素正解精度でランキングした順位順に整列して表示され、閾値未満のデータにおいては表示部 109 の別の領域においてデータの名称の名前順、データが有する時間情報順、データのデータサイズ順、データの表示サイズ順などの前記スコア順とは異なる基準で表示する。

30

#### 【0018】

図 6 に検索データ表示の様子を示す。同図において、601 は、検索キーとの音素正解精度が閾値以上のデータ（ここでは画像）が順位順で表示される検索結果表示ウインドウである。602 は、音素正解精度が閾値未満のデータが名前順、時間順など、順位順以外の方法で表示されるデータ表示ウインドウである。図 6 においてはデータ表示ウインドウでは名前順で画像が整列されて表示されている。また、音符のボタンを押すと対応する音声アノテーションを聞くことができる。ユーザは、まず検索結果表示ウインドウ 601 に表示された音素正解精度の高いデータを見て、所望のデータが無い場合にはデータ表示ウインドウで名前順、時間順などに整列されたデータから所望のデータを探す。なお、本実施例では順位順に整列するデータとそれ以外のデータでウインドウを分けて表示したが、本発明はこれに限るものではなく、例えば同一ウインドウで領域を分けて表示しても構わない。このように、サブワード正解精度が高いデータを順位順に、その他のデータを名前順、時間順などに整列してこれらを並列に表示することで、ユーザはまず限られた数の順位順データを見て、そこに無かったら通常どおり名前、時間順で探すという併用した使い方ができるので、利便性が高まる。

40

#### 【0019】

ここで、音素正解精度とデータ検索性能との関係からの音素正解精度閾値設定方法と、閾値処理による表示方法切り替えの有効性について説明する。図 7 は、1000 データを対象に検索を行ったときに、所望する正解データの検索キーとの音素正解精度及び、正解データを音素正解精度でランキングした検索順位をプロットした散布図である。同図を見

50

ると、音素正解精度が 60 %を超えるデータについては検索順位は良い順位に集中し、順位を大きく落とすデータは無い。その一方で、音素正解精度 60 %を下回るデータは大きく順位を落とし、順位の範囲もデータによって大きく異なる。したがって、音素正解精度 60 %を超える正解データについてはロバストに上位で検索できるのに対して、音素正解精度 60 %を下回るデータに対しては大きく順位を落とし、順位の範囲もデータに大きく左右されるので検索結果として提示するのは好ましくない。

#### 【 0 0 2 0 】

そこで、この予備実験で確認した特徴を生かし、音素正解精度の閾値を 60 %に設定する。音素正解精度 60 %を超える正解データについては（実際にはシステムは正解を知らないので音素正解精度 60 %を超える全てのデータは）検索結果表示ウインドウ 601 に順位順で表示する。検索キーとの音素正解精度が閾値 60 %を超える正解については検索結果表示ウインドウ 601 においてロバストに高い順位で検索できる。音素正解精度 60 %を超えないデータについては順位順に並べてもどの順位の範囲に現れるかわからず検索効率がむしろ悪いため、データ表示ウインドウ 602 でデータの名称の名前順、データが有する時間情報順、データのデータサイズ順、データの表示サイズ順などの前記スコア順とは異なる基準で表示する。

#### 【 0 0 2 1 】

すなわち、システム設計者があらかじめ検索キーに対応する正解データが既知である検索セットを用意して図 7 のような散布図を作れば、音素正解精度と検索性能の関係がもつ図 7 のグラフ形状の特徴により適切な閾値を設定でき、検索による限られた数の順位順表示と、名前順、時間順などの通常表示とのハイブリッドなユーザ提示ができる。

#### 【 0 0 2 2 】

図 8 に、本発明のデータ検索装置を実現するハードウェア構成図を示す。同図において、801 は、データや G U I パネルなどを表示するディスプレイなどの画面表示部である。802 は、検索キーなどを入力したり G U I ボタンを押下するキーボードやマウスなどのデータ入力部である。803 は、音声アノテーションデータや警告音などの音を出力するスピーカなどの音出力部である。804 は、データベース 101 や本データ検索方法のプログラムを保持する R O M 或いはハードディスクなどの外部記憶部である。805 は、本データ検索方法のプログラム実行時に、プログラムやデータ等、一時情報を保持するための R A M である。806 は、本データ検索方法のプログラムを実行する C P U である。

#### 【 0 0 2 3 】

##### （他の実施例）

上記実施例では音素列マッチングにより音素正解精度を検索のためのスコアとして用いたが、本発明はこれに限るものではなく、例えば音素ではなく音節でのマッチングや、単語単位でのマッチングによる正解精度でも良い。また、これに音声認識で求まる認識尤度 203 を加味したり、音素間の類似度（“ p ”と“ t ”は類似度が高いなど）を用いてスコアの重み付けをしたりしても良い。また、上記実施例では、図 5 に示すとおり音素列の全体マッチングによる音素正解精度を検索のためのスコアとして用いたが、挿入誤りによるスコアの劣化を抑えるなどとしたスコアの工夫により検索キーの部分的なマッチングによって検索しても良い。この実施例では、例えば音声認識結果アノテーションデータに「箱根の山」というアノテーションが付与されているときに、部分マッチングにより「箱根」、「山」を検索キーとして検索可能となる。

#### 【 0 0 2 4 】

##### （他の実施例）

上記実施例は音素正解精度の閾値処理によって、順位順による整列とその他の整列とを切り替えて別の領域に表示するものであるが、本発明はこれに限らず、音素正解精度の閾値処理によってデータの表示方法を切り替える全ての実施形態に適用可能である。例えば、音素正解精度が閾値未満のデータは全て表示せず、音素正解精度が閾値以上のデータのみを表示させたり、音素正解精度が閾値以上のデータのみ画像を大きく表示して閾値未満のデータは小さなアイコンもしくはリンクのテキストのみを表示するといった実施形態も

10

20

30

40

50

考えられる。

【0025】

(他の実施例)

なお、本発明は、1つの機器からなる装置に適用しても、複数の機器から構成されるシステムに適用してもよい。また、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記録媒体を、システム或いは装置に供給し、そのシステム或いは装置のコンピュータ（またはCPUやMPU）が記録媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。この場合、記録媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記録した記録媒体は本発明を構成することになる。

10

【0026】

上記実施例においては、プログラムをROMに保持する場合について説明したが、これに限定されるものではなく、任意の記憶媒体を用いて実現してもよい。また、同様の動作をする回路で実現してもよい。

【0027】

プログラムコードを供給するための記録媒体としては、例えば、フロッピー（登録商標）ディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモリカード、ROMなどを用いることができる。

【0028】

また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOSなどが実際の処理の一部または全部を行ない、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

20

【0029】

更に、記録媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行ない、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【図面の簡単な説明】

30

【0030】

【図1】本発明のデータ検索装置の機能構成図である。

【図2】本発明における音声認識結果アノテーションデータの例である。

【図3】本発明の検索キー入力部における検索キー入力ダイアログの例である。

【図4】本発明の検索キー変換部における処理の様子を表す図である。

【図5】本発明の検索部における音素マッチングを表す図である。

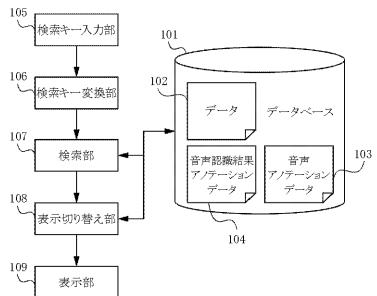
【図6】本発明の表示部における検索結果表示ウインドウとデータ表示ウインドウの一例である。

【図7】音素正解精度とデータ検索順位の関係を表す散布図である。

【図8】本発明のデータ検索装置を実現するハードウェア構成図である。

40

【図1】



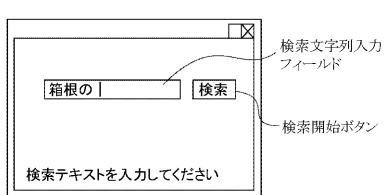
【図2】

```

string=faonenenooyamaa;type=typewriter;score=-212110;
string=haonenoyamaa;type=typewriter;score=-212144;
string=hakonenowama;type=typewriter;score=-192296;
string=hakoenoyamaa;type=typewriter;score=-202110;
string=hakonenoama;type=typewriter;score=-212144;
  
```

Diagram illustrating the recognition results. The results are shown as a list of strings with their types and scores. The list includes: string=faonenenooyamaa;type=typewriter;score=-212110, string=haonenoyamaa;type=typewriter;score=-212144, string=hakonenowama;type=typewriter;score=-192296, string=hakoenoyamaa;type=typewriter;score=-202110, and string=hakonenoama;type=typewriter;score=-212144. The results are grouped into three categories: 201 認識結果音素列 (201 Recognition Result Phoneme Sequence), 202 認識文法名 (202 Recognition Grammar Name), and 201 認識尤度 (201 Recognition Likelihood).

【図3】



【図4】

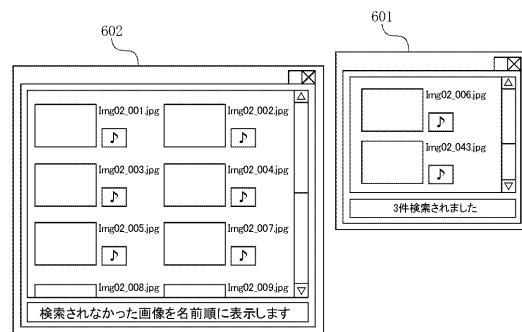


【図5】

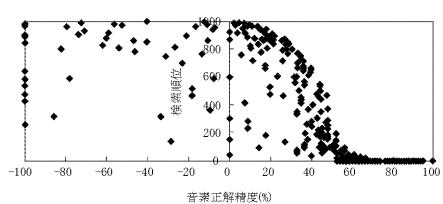
検索キー: *hakonenoyama*  
 認識結果音素列: *hakoneno(y)ama(a)*

検索キー音素数: 12  
 挿入誤り: 2  
 削除誤り: 0  
 置換誤り: 1  
 音素正解精度 =  $100 \times (12 - 2 - 0 - 1) / 12 = 75\%$

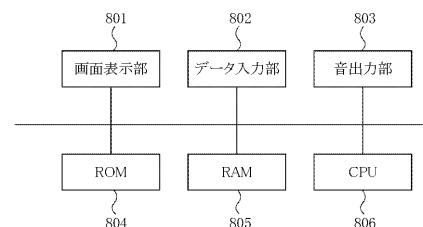
【図6】



【図7】



【図8】



---

フロントページの続き

(51)Int.Cl.

F I

H 0 4 N 5/225

F

H 0 4 N 101:00

(56)参考文献 特開平09-135417 (JP, A)

特開昭61-035498 (JP, A)

特開平03-266898 (JP, A)

特開2000-357175 (JP, A)

特開平11-175535 (JP, A)

特開平09-101970 (JP, A)

(58)調査した分野(Int.Cl., DB名)

G 0 6 F 17 / 3 0

G 1 0 L 15 / 0 0

G 1 0 L 15 / 2 2

H 0 4 N 5 / 2 2 5

H 0 4 N 101 / 0 0