

(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2012-43000
(P2012-43000A)

(43) 公開日 平成24年3月1日(2012.3.1)

(51) Int.Cl.	F I	テーマコード (参考)
G06F 17/21 (2006.01)	G06F 17/21 564P	5B109
G06F 17/22 (2006.01)	G06F 17/21 568A	5D015
G10L 15/00 (2006.01)	G06F 17/22 503	
	G10L 15/00 200T	

審査請求 未請求 請求項の数 11 O L (全 51 頁)

(21) 出願番号	特願2010-180944 (P2010-180944)	(71) 出願人	000002185 ソニー株式会社 東京都港区港南1丁目7番1号
(22) 出願日	平成22年8月12日 (2010.8.12)	(74) 代理人	100082131 弁理士 稲本 義雄
		(74) 代理人	100121131 弁理士 西川 孝
		(72) 発明者	前田 幸徳 東京都港区港南1丁目7番1号 ソニー株式会社内
		(72) 発明者	本田 等 東京都港区港南1丁目7番1号 ソニー株式会社内
		Fターム(参考)	5B109 KB00 RB32 RD03 5D015 LL11

(54) 【発明の名称】 検索装置、検索方法、及び、プログラム

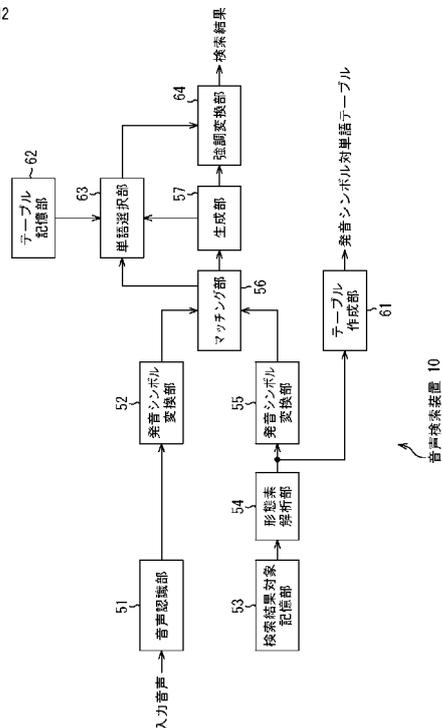
(57) 【要約】

【課題】音声検索において、利便性の高い検索結果を提供する。

【解決手段】マッチング部56は、複数の検索結果対象単語列それぞれについて、入力音声の音声認識結果とのマッチングを、発音を表す発音シンボルを用いて行い、生成部57は、そのマッチング結果に基づいて、入力音声に対する検索結果としての検索結果単語列を生成する。単語選択部63は、検索結果単語列となった検索結果対象単語列とのマッチング結果、及び、発音シンボルと単語とを対応付けた発音シンボル対単語テーブルに基づき、検索結果単語列の単語の中で強調する強調対象単語を選択し、強調変換部64は、検索結果単語列を、その検索結果単語列の単語のうちの強調対象単語を強調した強調単語列に変換する。本発明は、例えば、音声検索を行う場合に適用できる。

【選択図】図2

図2



【特許請求の範囲】

【請求項 1】

入力音声を音声認識する音声認識部と、

前記入力音声に対応する単語列の検索結果の対象となる単語列である複数の検索結果対象単語列それぞれについて、前記検索結果対象単語列の発音を表す発音シンボルの並びである検索結果対象発音シンボル列と、前記入力音声の音声認識結果の発音を表す発音シンボルの並びである認識結果発音シンボル列とを用いて、前記音声認識結果とのマッチングをとるマッチング部と、

前記検索結果対象単語列と前記音声認識結果とのマッチング結果に基づいて、前記複数の検索結果対象単語列からの、前記入力音声に対応する単語列の検索の結果である検索結果単語列を生成する生成部と、

前記検索結果単語列となった前記検索結果対象単語列との前記マッチング結果、及び、発音シンボルと単語とを対応付けた発音シンボル対単語テーブルに基づき、前記検索結果単語列の単語の中で強調する対象の強調対象単語を選択する選択部と、

前記検索結果単語列を、その検索結果単語列の単語のうちの前記強調対象単語を強調した強調単語列に変換する変換部と

を備える検索装置。

【請求項 2】

前記発音シンボル対単語テーブルには、前記検索結果対象単語列の発音シンボルと、前記検索結果対象単語列を構成する単語とが対応付けられている

請求項 1 に記載の検索装置。

【請求項 3】

前記発音シンボル対単語テーブルには、前記検索結果対象単語列の、連続する 2 個以上の発音シンボルを、単語に対応付ける対応付け単位として、前記対応付け単位と、その対応付け単位の少なくとも一部を発音シンボルに含む、前記検索結果対象単語列の単語とが対応付けられている

請求項 2 に記載の検索装置。

【請求項 4】

前記マッチング部は、

2 個以上の発音シンボルを、マッチングをとるときのマッチング単位として、前記マッチング単位で、前記音声認識結果と前記検索結果対象単語列とのマッチングを行い、

前記認識結果発音シンボル列と前記検索結果対象発音シンボル列との類似度と、

前記認識結果発音シンボル列のマッチング単位と一致する、前記検索結果対象発音シンボル列のマッチング単位であるマッチシンボルと

を、前記マッチング結果として出力する

請求項 3 に記載の検索装置。

【請求項 5】

前記選択部は、前記マッチング結果としての、前記検索結果対象発音シンボル列のマッチシンボルに基づき、前記発音シンボル対単語テーブルにおいて、前記マッチシンボルに含まれる前記対応付け単位に対応付けられている単語を、前記強調対象単語に選択する

請求項 4 に記載の検索装置。

【請求項 6】

前記マッチング単位、及び、対応付け単位は、同一の数の発音シンボルであり、

前記選択部は、前記発音シンボル対単語テーブルにおいて、前記マッチシンボルに一致する前記対応付け単位に対応付けられている単語を、前記強調対象単語に選択する

請求項 5 に記載の検索装置。

【請求項 7】

前記複数の検索結果対象単語列それぞれについて、前記検索結果対象発音シンボル列の、連続する 2 個以上の発音シンボルを、前記対応付け単位として、前記対応付け単位と、その対応付け単位の少なくとも一部を発音シンボルに含む、前記検索結果対象単語列の単

10

20

30

40

50

語とを対応付けた前記発音シンボル対単語テーブルを作成する作成部をさらに備える
請求項 6 に記載の検索装置。

【請求項 8】

前記生成部は、前記マッチング結果としての前記類似度に基づき、前記類似度が所定の
閾値以上の前記検索結果対象単語列、又は、前記類似度が上位の前記検索結果対象単語列
を、前記検索結果単語列とする

請求項 4 に記載の検索装置。

【請求項 9】

入力音声に対応する単語列を検索する検索装置が、

入力音声を音声認識する音声認識ステップと、

前記入力音声に対応する単語列の検索結果の対象となる単語列である複数の検索結果対
象単語列それぞれについて、前記検索結果対象単語列の発音を表す発音シンボルの並びで
ある検索結果対象発音シンボル列と、前記入力音声の音声認識結果の発音を表す発音シン
ボルの並びである認識結果発音シンボル列とを用いて、前記音声認識結果とのマッチング
をとるマッチングステップと、

前記検索結果対象単語列と前記音声認識結果とのマッチング結果に基づいて、前記複数
の検索結果対象単語列からの、前記入力音声に対応する単語列の検索の結果である検索結
果単語列を生成する生成ステップと、

前記検索結果単語列となった前記検索結果対象単語列との前記マッチング結果、及び、
発音シンボルと単語とを対応付けた発音シンボル対単語テーブルに基づき、前記検索結果
単語列の単語の中で強調する対象の強調対象単語を選択する選択ステップと、

前記検索結果単語列を、その検索結果単語列の単語のうちの前記強調対象単語を強調し
た強調単語列に変換する変換ステップと

を含む検索方法。

【請求項 10】

入力音声を音声認識する音声認識部と、

前記入力音声に対応する単語列の検索結果の対象となる単語列である複数の検索結果対
象単語列それぞれについて、前記検索結果対象単語列の発音を表す発音シンボルの並びで
ある検索結果対象発音シンボル列と、前記入力音声の音声認識結果の発音を表す発音シン
ボルの並びである認識結果発音シンボル列とを用いて、前記音声認識結果とのマッチング
をとるマッチング部と、

前記検索結果対象単語列と前記音声認識結果とのマッチング結果に基づいて、前記複数
の検索結果対象単語列からの、前記入力音声に対応する単語列の検索の結果である検索結
果単語列を生成する生成部と、

前記検索結果単語列となった前記検索結果対象単語列との前記マッチング結果、及び、
発音シンボルと単語とを対応付けた発音シンボル対単語テーブルに基づき、前記検索結果
単語列の単語の中で強調する対象の強調対象単語を選択する選択部と、

前記検索結果単語列を、その検索結果単語列の単語のうちの前記強調対象単語を強調し
た強調単語列に変換する変換部と

して、コンピュータを機能させるためのプログラム。

【請求項 11】

クライアント端末で入力された入力音声に対応する単語列の検索結果の対象となる単語
列である複数の検索結果対象単語列それぞれについて、前記検索結果対象単語列の発音を
表す発音シンボルの並びである検索結果対象発音シンボル列と、前記入力音声の音声認識
結果の発音を表す発音シンボルの並びである認識結果発音シンボル列とを用いて、前記音
声認識結果とのマッチングをとるマッチング部と、

前記検索結果対象単語列と前記音声認識結果とのマッチング結果に基づいて、前記複数
の検索結果対象単語列からの、前記入力音声に対応する単語列の検索の結果である検索結
果単語列を生成する生成部と、

前記検索結果単語列となった前記検索結果対象単語列との前記マッチング結果、及び、

10

20

30

40

50

発音シンボルと単語とを対応付けた発音シンボル対単語テーブルに基づき、前記検索結果単語列の単語の中で強調する対象の強調対象単語を選択する選択部と、

前記検索結果単語列を、その検索結果単語列の単語のうちの前記強調対象単語を強調した強調単語列に変換する変換部と

を備える検索装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、検索装置、検索方法、及び、プログラムに関し、特に、例えば、音声検索において、利便性の高い検索結果を提供することができるようにする検索装置、検索方法、及び、プログラムに関する。

10

【背景技術】

【0002】

ユーザから入力される音声である入力音声を用い、その音声に対応するテキスト等の単語列を検索する音声検索の方法としては、例えば、音声認識装置のみを用いる方法がある（例えば、特許文献1を参照）。

【0003】

音声認識装置のみを用いる音声検索では、音声認識装置において、あらかじめ辞書に登録された単語（語彙）の並びを、音声認識結果の対象として、入力音声の音声認識が行われ、その音声認識結果が、入力音声に対応する単語列の検索の結果である検索結果単語列として出力される。

20

【0004】

したがって、音声認識装置のみを用いる音声検索では、入力音声に対応する単語列の検索結果の対象となる単語列（以下、検索結果対象単語列ともいう）は、音声認識結果の対象である、辞書に登録された単語の並びである単語列（本明細書では、1つの単語を含む）だけであるため、ユーザの発話は、音声認識に用いられる辞書に登録された単語の並びに制限される。

【0005】

そこで、近年においては、ボイスサーチ(Voice Search)と呼ばれる音声検索の方法が提案されている。

30

【0006】

ボイスサーチでは、N-gram等の言語モデルを用いて、連続音声認識が行われ、その音声認識結果と、音声認識に用いられる辞書とは別に用意されたDB(Database)に登録されたテキストとのマッチング（DBに登録されたテキストからの、音声認識結果に対応するテキストのテキスト検索）が行われる。

【0007】

そして、そのマッチングの結果に基づき、音声認識結果にマッチする最上位の、又は、上位N位以内のテキストが、検索結果単語列として出力される。

【0008】

ボイスサーチでは、音声認識に用いられる辞書とは別に用意されたDBに登録されたテキストが、検索結果対象単語列となるため、そのDBに、多数のテキストを登録しておくことにより、その多数のテキストを、検索結果対象単語列として、音声検索を行うことができる。

40

【0009】

すなわち、ボイスサーチによれば、ユーザが、音声認識に用いられる辞書に登録された単語以外の単語を含む発話を行っても、DBに登録された検索結果対象単語列としてのテキストの範囲内で、ある程度の精度の音声検索を行うことができる。

【0010】

なお、音声及び画像が格納されたマルチメディアファイルを対象として音声認識を行うことで、マルチメディアファイルの音声から、索引付けのためのテキストを生成しておき

50

、話者識別に基づいて、マルチメディアファイルの音声を検索する方法が提案されている（例えば、特許文献2を参照）。

【先行技術文献】

【特許文献】

【0011】

【特許文献1】特開2001-242884号公報

【特許文献2】特開2000-348064号公報

【発明の概要】

【発明が解決しようとする課題】

【0012】

ところで、例えば、Web上で提供される検索サイト（Google等）では、ユーザが、テキストによりキーワードを入力すると、ブラウザにおいて、キーワードの検索結果が表示されるが、その検索結果においては、キーワードに対応する部分が、太字等で強調表示される（強調して表示される）。

【0013】

ここで、検索サイトでは、キーワードのテキスト検索が行われ、そのテキスト検索の検索結果のうちの、表記がキーワードに完全に一致するテキストが、キーワードに対応する部分として、強調表示される。

【0014】

さらに、検索サイトでは、テキスト検索の検索結果のうちの、表記がキーワードと等価なテキスト等も、キーワードに対応する部分として、強調表示される。

【0015】

具体的には、例えば、キーワード「Sony」のテキスト検索の検索結果に、表記がキーワード「Sony」と等価なテキスト「ソニー」が含まれる場合には、そのテキスト「ソニー」も、キーワードに対応する部分として、強調表示される。

【0016】

以上のように、検索結果のうちの、キーワードに対応する部分を強調表示することで、利便性の高い検索結果を提供することができる。

【0017】

すなわち、キーワードに対応する部分を強調表示した検索結果によれば、ユーザは、その検索結果が検索された理由を把握することができるので、検索結果が検索された理由が分からないことに起因するストレスを感じることを防止することができる。

【0018】

さらに、キーワードに対応する部分を強調表示した検索結果によれば、ユーザは、複数の検索結果の中から、所望の検索結果を、直感的に、迅速に見つけ出すことができる。

【0019】

ボイスサーチにおいても、検索結果対象単語列からの、入力音声の音声認識結果のテキスト検索の検索結果である検索結果単語列のうちの、表記がテキスト検索で音声認識結果に一致（完全に一致、又は、一部が一致）した部分を、入力音声に対応する部分として、強調表示することで、利便性の高い検索結果（検索結果単語列）を提供することができる。

【0020】

すなわち、例えば、ユーザが、入力音声「ローマ」を発話し、その音声認識結果「ローマ」のテキスト検索によって、テキスト「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」が、検索結果単語列として得られたとする。

【0021】

特に、検索結果単語列「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」のような文字数の多いテキストについては、入力音声に対応する部分を強調せずに表示すると、ユーザは、一見しただけでは、入力音声の音声認識結果に一致するテキスト「ローマ」を見つけることが困難なことがあり、その結果、テキスト「世界遺産都市の遺産

10

20

30

40

50

イタリアローマベネチアナポリフィレンツェ」が検索された理由が分からずに、ストレスを感じることもある。

【0022】

これに対して、検索結果単語列「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」を、「世界遺産都市の遺産イタリア”ローマ”ベネチアナポリフィレンツェ」のように、入力音声に対応する部分「ローマ」を強調して表示した場合には、ユーザは、一見しただけで、入力音声の音声認識結果に一致するテキスト「ローマ」を見つけることができ、テキスト「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」が検索された理由を容易に把握することができる。

【0023】

ところで、ボイスサーチにおいて、入力音声の音声認識結果のテキスト検索の結果得られる検索結果単語列のうちの、表記がテキスト検索で音声認識結果に一致した部分（テキスト）を、入力音声に対応する部分として、強調表示する場合には、入力音声の音声認識結果によっては、適切な強調表示が行われないことがある。

【0024】

すなわち、例えば、入力音声「都市の世界遺産」に対して、誤った音声認識結果「年の瀬解散」が得られた場合、その音声認識結果「年の瀬解散」のテキスト検索の結果得られる検索結果単語列に、入力音声に一致（完全に一致、又は、一部が一致）するテキスト「世界遺産」や「世界の都市」等が含まれていても、そのテキスト「世界遺産」や「世界の都市」（の表記）は、音声認識結果「年の瀬解散」に一致しないため、強調表示されず、利便性の高い検索結果を得ることが困難となる。

【0025】

また、音声認識結果が誤っていない場合も、入力音声に対応する部分が強調表示されないことがある。

【0026】

すなわち、例えば、ユーザが、入力音声「としのせかい」を発話した場合、音声認識結果として、「都市の世界」、及び、「年の瀬かい」のうちのいずれが得られても、音声認識結果が誤っているとは言えない。

【0027】

そして、ユーザが、「都市の世界」と発話したつもりの入力音声「としのせかい」に対し、音声認識結果「年の瀬かい」が得られた場合、音声認識結果「年の瀬かい」のテキスト検索の結果得られる検索結果単語列に、入力音声（の表記）に一致するテキスト「都市の世界」が含まれていても、そのテキスト「都市の世界」は、音声認識結果「年の瀬かい」にまったく一致しないため、強調表示されず、利便性の高い検索結果を得ることが困難となる。

【0028】

本発明は、このような状況に鑑みてなされたものであり、音声検索において、利便性の高い検索結果を提供することができるようにするものである。

【課題を解決するための手段】

【0029】

本発明の一側面の検索装置、又は、プログラムは、入力音声を音声認識する音声認識部と、前記入力音声に対応する単語列の検索結果の対象となる単語列である複数の検索結果対象単語列それぞれについて、前記検索結果対象単語列の発音を表す発音シンボルの並びである検索結果対象発音シンボル列と、前記入力音声の音声認識結果の発音を表す発音シンボルの並びである認識結果発音シンボル列とを用いて、前記音声認識結果とのマッチングをとるマッチング部と、前記検索結果対象単語列と前記音声認識結果とのマッチング結果に基づいて、前記複数の検索結果対象単語列からの、前記入力音声に対応する単語列の検索の結果である検索結果単語列を生成する生成部と、前記検索結果単語列となった前記検索結果対象単語列との前記マッチング結果、及び、発音シンボルと単語とを対応付けた発音シンボル対単語テーブルに基づき、前記検索結果単語列の単語の中で強調する対象の

10

20

30

40

50

強調対象単語を選択する選択部と、前記検索結果単語列を、その検索結果単語列の単語のうちの前記強調対象単語を強調した強調単語列に変換する変換部とを備える検索装置、又は、検索装置として、コンピュータを機能させるためのプログラムである。

【0030】

本発明の一側面の検索方法は、入力音声に対応する単語列を検索する検索装置が、入力音声を入力音声認識する音声認識ステップと、前記入力音声に対応する単語列の検索結果の対象となる単語列である複数の検索結果対象単語列それぞれについて、前記検索結果対象単語列の発音を表す発音シンボルの並びである検索結果対象発音シンボル列と、前記入力音声の音声認識結果の発音を表す発音シンボルの並びである認識結果発音シンボル列とを用いて、前記音声認識結果とのマッチングをとるマッチングステップと、前記検索結果対象単語列と前記音声認識結果とのマッチング結果に基づいて、前記複数の検索結果対象単語列からの、前記入力音声に対応する単語列の検索の結果である検索結果単語列を生成する生成ステップと、前記検索結果単語列となった前記検索結果対象単語列との前記マッチング結果、及び、発音シンボルと単語とを対応付けた発音シンボル対単語テーブルに基づき、前記検索結果単語列の単語の中で強調する対象の強調対象単語を選択する選択ステップと、前記検索結果単語列を、その検索結果単語列の単語のうちの前記強調対象単語を強調した強調単語列に変換する変換ステップとを含む検索方法である。

10

【0031】

本発明の他の一側面の検索装置は、クライアント端末で入力された入力音声に対応する単語列の検索結果の対象となる単語列である複数の検索結果対象単語列それぞれについて、前記検索結果対象単語列の発音を表す発音シンボルの並びである検索結果対象発音シンボル列と、前記入力音声の音声認識結果の発音を表す発音シンボルの並びである認識結果発音シンボル列とを用いて、前記音声認識結果とのマッチングをとるマッチング部と、前記検索結果対象単語列と前記音声認識結果とのマッチング結果に基づいて、前記複数の検索結果対象単語列からの、前記入力音声に対応する単語列の検索の結果である検索結果単語列を生成する生成部と、前記検索結果単語列となった前記検索結果対象単語列との前記マッチング結果、及び、発音シンボルと単語とを対応付けた発音シンボル対単語テーブルに基づき、前記検索結果単語列の単語の中で強調する対象の強調対象単語を選択する選択部と、前記検索結果単語列を、その検索結果単語列の単語のうちの前記強調対象単語を強調した強調単語列に変換する変換部とを備える検索装置である。

20

30

【0032】

本発明の一側面においては、入力音声に対応する単語列の検索結果の対象となる単語列である複数の検索結果対象単語列それぞれについて、前記検索結果対象単語列の発音を表す発音シンボルの並びである検索結果対象発音シンボル列と、前記入力音声の音声認識結果の発音を表す発音シンボルの並びである認識結果発音シンボル列とを用いて、前記音声認識結果とのマッチングがとられ、そのマッチング結果に基づいて、前記複数の検索結果対象単語列からの、前記入力音声に対応する単語列の検索の結果である検索結果単語列が生成される。そして、前記検索結果単語列となった前記検索結果対象単語列との前記マッチング結果、及び、発音シンボルと単語とを対応付けた発音シンボル対単語テーブルに基づき、前記検索結果単語列の単語の中で強調する対象の強調対象単語が選択され、前記検索結果単語列が、その検索結果単語列の単語のうちの前記強調対象単語を強調した強調単語列に変換される。

40

【0033】

なお、検索装置は、独立した装置であっても良いし、1つの装置を構成している内部ブロックであっても良い。

【0034】

また、プログラムは、伝送媒体を介して伝送することにより、又は、記録媒体に記録して、提供することができる。

【発明の効果】

【0035】

50

本発明の一側面によれば、音声検索において、利便性の高い検索結果を提供することができる。

【図面の簡単な説明】

【0036】

【図1】本発明を適用したレコーダの一実施の形態の構成例を示すブロック図である。

【図2】音声検索装置10の構成例を示すブロック図である。

【図3】表記シンボルを用いたマッチングで、表記が異なる音声認識結果に対して異なるマッチング結果が得られることが、音声検索の性能に有利でないことを説明する図である。

【図4】マッチング単位として、音節2連鎖を採用する場合の、発音シンボル変換部52の処理を説明する図である。 10

【図5】マッチング単位として、音節2連鎖を採用する場合の、発音シンボル変換部55の処理を説明する図である。

【図6】単語単位でのマッチング、音節単位でのマッチング、及び、音節2連鎖単位でのマッチングの結果を示す図である。

【図7】音声認識結果と検索結果対象単語列との類似度として、コサイン距離D、第1の補正距離D1、及び、第2の補正距離D2を採用した場合のマッチングのシミュレーションの結果を示す図である。

【図8】音声認識結果と検索結果対象単語列との類似度として、コサイン距離D、第1の補正距離D1、及び、第2の補正距離D2を採用した場合のマッチングの他のシミュレーションの結果を示す図である。 20

【図9】音声検索装置10の処理を説明するフローチャートである。

【図10】テーブル作成部61が作成する発音シンボル対単語テーブルの例を示す図である。

【図11】テーブル作成部61が発音シンボル対単語テーブルを作成する第1の作成方法を説明する図である。

【図12】第1の作成方法で発音シンボル対単語テーブルを作成する処理を説明するフローチャートである。

【図13】テーブル作成部61が発音シンボル対単語テーブルを作成する第2の作成方法を説明する図である。 30

【図14】第2の作成方法で発音シンボル対単語テーブルを作成する処理を説明するフローチャートである。

【図15】テーブル作成部61が作成する発音シンボル対単語テーブルの他の例を示す図である。

【図16】単語選択部63が行う処理を説明するフローチャートである。

【図17】発音シンボル対単語テーブルのさらに他の例を示す図である。

【図18】検索結果単語列の表示例を示す図である。

【図19】検索結果単語列の表示例を示す図である。

【図20】本発明を適用したコンピュータの一実施の形態の構成例を示すブロック図である。 40

【発明を実施するための形態】

【0037】

[本発明を適用したレコーダの一実施の形態]

【0038】

図1は、本発明の検索装置を適用したレコーダの一実施の形態の構成例を示すブロック図である。

【0039】

図1において、レコーダは、音声検索装置10、レコーダ機能部20、操作部31、制御部32、及び、出力I/F(Interface)33を有する。

【0040】

音声検索装置 10 には、ユーザが発話した入力音声（のデータ）が、図示せぬマイク等から供給される。

【0041】

音声検索装置 10 は、そこに供給される入力音声を用いて、例えば、ボイスサーチによる音声検索を行い、その結果得られる検索結果単語列を、出力 I/F 33 に供給する。

【0042】

レコーダ機能部 20 は、チューナ 21、記録再生部 22、及び、記録媒体 23 を有し、テレビジョン放送の番組の記録（録画）及び再生を行う。

【0043】

すなわち、チューナ 21 には、図示せぬアンテナで受信された、例えば、デジタル放送によるテレビジョン放送信号が供給される。

【0044】

チューナ 21 は、そこに供給されるテレビジョン放送信号を受信し、そのテレビジョン放送信号から所定のチャンネルのテレビジョン放送信号を抽出して、ビットストリームを復調し、記録再生部 22 に供給する。

【0045】

記録再生部 22 は、チューナ 21 から供給されるビットストリームから、EPG(Electronic Program Guide)や番組のデータ等を抽出し、出力 I/F 33 に供給する。

【0046】

また、記録再生部 22 は、EPGや番組のデータを、記録媒体 23 に記録（録画）する。

【0047】

さらに、記録再生部 22 は、記録媒体 23 から、番組のデータを再生し、出力 I/F 33 に供給する。

【0048】

記録媒体 23 は、例えば、HD(Hard Disk)や半導体メモリ等であり、記録媒体 23 には、記録再生部 22 によって、EPGや番組のデータが記録される。

【0049】

操作部 31 は、リモートコマンドや、レコーダの筐体に設けられたボタン、図示せぬディスプレイに GUI(Graphics User Interface)として表示されるボタン等で構成される。操作部 31 は、ユーザによって操作され、その操作に対応した操作信号を、制御部 32 に供給する。

【0050】

制御部 32 は、操作部 31 からの操作信号等に従い、音声検索装置 10、及び、レコーダ機能部 20 等の、レコーダを構成するブロックの制御、その他の処理を行う。

【0051】

出力 I/F 33 には、記録再生部 22 から、EPGや番組のデータが供給される。また、出力 I/F 33 には、音声検索装置 10 から、入力音声に対する音声検索の結果である検索結果単語列が供給される。

【0052】

出力部 I/F 33 は、例えば、TV(テレビジョン受像機)等の、少なくとも画像を表示することができる表示デバイスと接続されるインタフェースであり、記録再生部 22 からの EPGや番組のデータ、及び、音声検索装置 10 からの検索結果単語列を、出力部 I/F 33 に接続された、例えば、図示せぬ TV に供給して表示させる。

【0053】

[音声検索装置 10 の構成例]

【0054】

図 2 は、図 1 の音声検索装置 10 の構成例を示すブロック図である。

【0055】

図 2 では、音声検索装置 10 は、音声認識部 51、発音シンボル変換部 52、検索結果対象記憶部 53、形態素解析部 54、発音シンボル変換部 55、マッチング部 56、生成

10

20

30

40

50

部 5 7、テーブル作成部 6 1、テーブル記憶部 6 2、単語選択部 6 3、及び、強調変換部 6 4 を有し、ボイスサーチによる音声検索を行う。

【 0 0 5 6 】

ここで、従来のボイスサーチでは、音声認識結果と、検索結果対象単語列としてのテキストとのマッチングが、音声認識結果、及び、検索結果対象単語列それぞれの表記を表すシンボルである表記シンボルを用い、単語単位や、表記シンボル単位で行われる。

【 0 0 5 7 】

したがって、音声認識結果の表記シンボルに誤りがあると、マッチングにおいて、入力音声に対応する単語列とはまったく別の検索結果対象単語列が、音声認識結果にマッチし、その結果、そのような、入力音声に対応する単語列とはまったく別の検索結果対象単語列が、検索結果単語列として出力される。

10

【 0 0 5 8 】

すなわち、ユーザが、入力音声として、例えば、「としのせかい」を発話し、その音声認識結果の表記シンボル列が、例えば、「都市の世界」であった場合、単語単位のマッチングでは、音声認識結果の表記シンボル列「都市の世界」を、「都市 / の / 世界 / 」(スラッシュ(/)は、区切りを表す)のように、1 個ずつの単語に区切って、マッチングが行われ、表記シンボル単位のマッチングでは、音声認識結果の表記シンボル列「都市の世界」を、「都 / 市 / の / 世 / 界」のように、1 個ずつの表記シンボルに区切って、マッチングが行われる。

【 0 0 5 9 】

20

一方、入力音声「としのせかい」の音声認識結果の表記シンボル列が、例えば、「年の瀬かい」であった場合、単語単位のマッチングでは、音声認識結果の表記シンボル列「年の瀬かい」を、「 / 年 / の / 瀬 / かい / 」のように、1 個ずつの単語に区切って、マッチングが行われ、表記シンボル単位のマッチングでは、音声認識結果の表記シンボル列「年の瀬かい」を、「年 / の / 瀬 / か / い」のように、1 個ずつの表記シンボルに区切って、マッチングが行われる。

【 0 0 6 0 】

したがって、入力音声「としのせかい」の音声認識結果(の表記シンボル列)が、「都市の世界」である場合と、「年の瀬かい」である場合とでは、音声認識結果にマッチする検索結果対象単語列は、大きく異なり、その結果、入力音声に対応する単語列とはまったく別の検索結果対象単語列が、検索結果単語列として出力される一方、入力音声に対応する単語列が、検索結果単語列として出力されないことがある。

30

【 0 0 6 1 】

以上のように、表記シンボルを用いたマッチングは、音声認識結果との親和性が高いとはいえず、入力音声に対応する単語列が、検索結果単語列として得られないことがある。

【 0 0 6 2 】

そこで、音声検索装置 1 0 では、音声認識結果と、検索結果対象単語列とのマッチングを、音声認識結果、及び、検索結果対象単語列のそれぞれの発音を表すシンボルである発音シンボルを用いて行うことで、入力音声に対応する単語列の検索を、ロバストに行うことができるようにし、これにより、入力音声に対応する単語列が、検索結果単語列として出力されないことを防止する。

40

【 0 0 6 3 】

すなわち、図 2 の音声検索装置 1 0 において、音声認識部 5 1 には、ユーザの発話である入力音声(のデータ)が、図示せぬマイク等から供給される。

【 0 0 6 4 】

音声認識部 5 1 は、そこに供給される入力音声を音声認識し、音声認識結果(の表記シンボル)を、発音シンボル変換部 5 2 に供給する。

【 0 0 6 5 】

発音シンボル変換部 5 2 は、音声認識部 5 1 から供給される、入力音声の音声認識結果(の表記シンボル)を、その音声認識結果の発音を表す発音シンボルの並びである認識結

50

果発音シンボル列に変換し、マッチング部 5 6 に供給する。

【 0 0 6 6 】

検索結果対象記憶部 5 3 は、複数の検索結果対象単語列、すなわち、マッチング部 5 6 において、音声認識結果とのマッチングが行われ、入力音声に対応する単語列の検索の結果である検索結果単語列となり得る単語列（の表記シンボルとしてのテキスト）を記憶する。

【 0 0 6 7 】

ここで、検索結果対象記憶部 5 3 には、例えば、図 1 のレコーダの記録媒体 2 3 に記録された EPG を構成する構成要素である番組のタイトルや、出演者名、番組の内容の詳細情報等が供給されて記憶される。

【 0 0 6 8 】

また、検索結果対象記憶部 5 3 には、例えば、図 1 のレコーダの記録媒体 2 3 に録画（記録）された番組（録画番組）のメタデータである、番組のタイトルや、出演者名、詳細情報等が供給されて記憶される。

【 0 0 6 9 】

したがって、本実施の形態では、音声検索装置 1 0 において、番組のタイトルや、出演者名、詳細情報等を、検索結果対象単語列として、音声検索が行われる。

【 0 0 7 0 】

形態素解析部 5 4 は、検索結果対象記憶部 5 3 に記憶された検索結果対象単語列の形態素解析を行うことで、検索結果対象単語列を、例えば、単語（形態素）単位に分割し、検索結果対象単語列を構成する各単語（の表記）や、その単語の音韻（読み）等を、形態素解析結果として、発音シンボル変換部 5 5、及び、テーブル作成部 6 1 に供給する。

【 0 0 7 1 】

発音シンボル変換部 5 5 は、形態素解析部 5 4 から供給される形態素解析結果を用いて、検索結果対象単語列（の表記シンボル）を、その検索結果対象単語列の発音を表す発音シンボルの並びである検索結果対象発音シンボル列に変換し、検索結果対象単語列（の表記シンボル）とともに、マッチング部 5 6 に供給する。

【 0 0 7 2 】

マッチング部 5 6 は、発音シンボル変換部 5 2 からの音声認識結果の認識結果発音シンボル列と、発音シンボル変換部 5 5 からの検索対象単語列の検索結果対象発音シンボル列とを用いて、その音声認識結果と検索対象単語列とのマッチングをとり、そのマッチング結果を、生成部 5 7、及び、単語選択部 6 3 に供給する。

【 0 0 7 3 】

すなわち、マッチング部 5 6 は、検索結果対象記憶部 5 3 に記憶されたすべての検索結果対象単語列それぞれについて、入力音声の音声認識結果とのマッチングを、音声認識結果の発音シンボルと、検索結果対象単語列の発音シンボルとを用いて行う。

【 0 0 7 4 】

マッチング部 5 6 は、検索結果対象記憶部 5 3 に記憶されたすべての検索結果対象単語列それぞれについて、入力音声の音声認識結果とのマッチングをとり、そのマッチング結果を、生成部 5 7、及び、単語選択部 6 3 に供給する。

【 0 0 7 5 】

なお、認識結果発音シンボル列と検索結果対象発音シンボル列と（を用いた音声認識結果と検索結果対象単語列と）のマッチングでは、認識結果発音シンボル列と検索結果対象発音シンボル列とが類似している度合いを表す類似度が求められる。

【 0 0 7 6 】

類似度としては、例えば、ベクトル空間法のコサイン距離等を用いることができる。

【 0 0 7 7 】

ここで、マッチング部 5 6 は、例えば、2 個の発音シンボル等の 1 個以上の（連続する）発音シンボルを、マッチングをとるときのマッチング単位として、そのマッチング単位で、音声認識結果と検索結果対象単語列とのマッチングを行う。

10

20

30

40

50

【0078】

そして、マッチング部56は、認識結果発音シンボル列と検索結果対象発音シンボル列との類似度と、認識結果発音シンボル列のマッチング単位と一致する、検索結果対象発音シンボル列のマッチング単位であるマッチシンボルとを、発音シンボル変換部55からの検索結果対象単語列(の表記シンボル)とともに、マッチング結果として出力する。

【0079】

以上のようなマッチング結果のうちの、検索結果対象単語列、及び、類似度は、生成部57に供給され、検索結果対象単語列、及び、マッチシンボルは、単語選択部63に供給される。

【0080】

生成部57は、マッチング部56からのマッチング結果に基づいて、検索結果対象記憶部53に記憶された検索結果対象単語列の中からの、入力音声に対応する単語列の検索の結果である検索結果単語列を生成する。

【0081】

すなわち、例えば、生成部57は、マッチング部56からの、検索結果対象記憶部53に記憶された検索結果対象単語列すべてのマッチング結果の中から、類似度が上位の検索結果対象単語列、又は、類似度があらかじめ設定された閾値以上の検索結果対象単語列(の表記シンボル)を、入力音声に対する検索結果単語列として選択する。

【0082】

そして、生成部57は、検索結果単語列を、単語選択部63、及び、強調変換部64に供給する。

【0083】

テーブル作成部61は、形態素解析部54から供給される形態素解析結果を用いて、発音シンボルと単語とを対応付けた発音シンボル対単語テーブルを作成し、テーブル記憶部62に供給する。

【0084】

すなわち、テーブル作成部61は、検索結果対象記憶部53に記憶された検索結果対象単語列について、検索結果対象単語列の、連続する2個以上の発音シンボルを、単語に対応付ける対応付け単位として、対応付け単位と、その対応付け単位の少なくとも一部を発音シンボルに含む、検索結果対象単語列の単語とを対応付けて登録した発音シンボル対単語テーブルを作成する。

【0085】

テーブル記憶部62は、テーブル作成部61から供給される発音シンボル対単語テーブルを記憶する。

【0086】

単語選択部63は、マッチング部56から供給されるマッチング結果のうちの、生成部57から供給される検索結果単語列となった検索結果対象単語列とのマッチング結果、及び、テーブル記憶部62に記憶された発音シンボル対単語テーブルに基づき、検索結果単語列の単語の中で強調する対象の強調対象単語を選択する。

【0087】

ここで、検索結果単語列となった検索結果対象単語列とのマッチング結果には、その検索結果単語列の発音シンボルのうちの、認識結果発音シンボル列のマッチング単位と一致する発音シンボルであるマッチシンボルが含まれる。

【0088】

一方、発音シンボル対単語テーブルには、検索結果対象単語列の、連続する2個以上の発音シンボルである対応付け単位と、その検索結果対象単語列を構成する単語とがセットになるように対応付けられている。

【0089】

単語選択部63は、発音シンボル対単語テーブルに登録された、検索結果単語列(となった検索結果対象単語列)についての、対応付け単位と単語とのセット(以下、単語セッ

10

20

30

40

50

トともいう)の中で、検索結果単語列のマッチシンボルに含まれる対応付け単位を含む単語セットを検出し、その単語セットに含まれる単語を、強調対象単語に選択する。

【0090】

なお、ある検索結果単語列(となった検索結果対象単語列)について選択される強調対象単語は、その検索結果単語列に含まれる単語である。

【0091】

単語選択部63は、検索結果単語列について選択した強調対象単語を、強調変換部64に供給する。

【0092】

強調変換部64は、生成部57からの検索結果単語列を、その検索結果単語列の単語のうち、単語選択部63からの強調対象単語を強調した強調単語列に変換し、出力I/F33(図1)に供給する。

10

【0093】

すなわち、強調変換部64は、生成部57からの検索結果単語列に、その検索結果単語列について、単語選択部63から供給される強調対象単語(に一致する単語)が含まれる場合には、生成部57からの検索結果単語列のうち、強調対象単語のフォントや表示色等の表示の属性(のデータ)を、他の単語の表示の属性とは異なる属性に変更することで、生成部57からの検索結果単語列を、強調対象単語が強調表示される強調単語列に変換する。

【0094】

以上のように構成される音声検索装置10では、ユーザの発話に応じて、ボイスサーチによる音声検索の処理が行われる。

20

【0095】

すなわち、形態素解析部54は、検索結果対象記憶部53に記憶されたすべての検索結果対象単語列の形態素解析を行い、形態素解析結果を、発音シンボル変換部55と、テーブル作成部61とに供給する。

【0096】

発音シンボル変換部55は、検索結果対象記憶部53に記憶されたすべての検索結果対象単語列を、形態素解析部54からの形態素解析結果に基づいて、検索結果対象発音シンボル列に変換し、検索結果対象単語列とともに、マッチング部56に供給する。

30

【0097】

また、テーブル作成部61は、形態素解析部54から供給される形態素解析結果を用いて、検索結果対象記憶部53に記憶されたすべての検索結果対象単語列について、検索結果対象単語列の、連続する2個以上の発音シンボルである対応付け単位と、その検索結果対象単語列を構成する単語とを、単語セットとして対応付けて登録した発音シンボル対単語テーブルを作成し、テーブル記憶部62に供給して記憶させる。

【0098】

その後、ユーザが発話を行い、その発話としての入力音声は、音声認識部51に供給されると、音声認識部51は、その入力音声を音声認識し、その入力音声の音声認識結果を、発音シンボル変換部52に供給する。

40

【0099】

発音シンボル変換部52は、音声認識部51からの入力音声の音声認識結果を、認識結果発音シンボル列に変換し、マッチング部56に供給する。

【0100】

マッチング部56は、検索結果対象記憶部53に記憶されたすべての検索結果対象単語列それぞれについて、発音シンボル変換部52からの認識結果発音シンボル列と、発音シンボル変換部55からの検索結果対象発音シンボル列とを用いて、入力音声の音声認識結果とのマッチングをとり、そのマッチング結果を、生成部57と単語選択部63とに供給する。

【0101】

50

生成部 5 7 は、マッチング部 5 6 からのマッチング結果に基づいて、検索結果対象記憶部 5 3 に記憶された検索結果対象単語列の中から、入力音声に対応する単語列の検索の結果である検索結果単語列（とする検索結果対象単語列）を選択し、単語選択部 6 3 と強調変換部 6 4 とに供給する。

【0102】

単語選択部 6 3 は、マッチング部 5 6 から供給されるマッチング結果のうちの、生成部 5 7 から供給される検索結果単語列となった検索結果対象単語列とのマッチング結果、及び、テーブル記憶部 6 2 に記憶された発音シンボル対単語テーブルに基づき、検索結果単語列の単語の中で強調する対象の強調対象単語を選択し、強調変換部 6 4 に供給する。

【0103】

強調変換部 6 4 は、生成部 5 7 からの検索結果単語列を、単語選択部 6 3 からの強調対象単語が強調表示される強調単語列に変換して（出力 I/F 3 3 に）出力する。

【0104】

以上のような音声検索装置 1 0 を有する図 1 のレコーダでは、例えば、録画が行われた番組（録画番組）の中から、ユーザが所望する番組を、音声検索によって検索し、再生することができる。

【0105】

すなわち、ユーザが、再生をしようとする番組の音声検索を行うためのキーワードとして、例えば、入力音声「世界遺産」を発話すると、音声検索装置 1 0 では、検索結果対象記憶部 5 3 に検索結果対象単語列として記憶されている、録画番組のタイトル等を対象として、音声検索を行うことにより、タイトルの発音が、入力音声「世界遺産」の発音に類似する所定数の番組のタイトルが、検索結果単語列として生成され、強調単語列に変換されて、出力 I/F 3 3 に供給される。

【0106】

そして、出力 I/F 3 3 は、強調単語列（に変換された検索結果単語列）としての番組のタイトルを、再生を行う候補の番組（再生候補番組）として、（レコーダと接続された TV（テレビジョン受像機）等で）表示させる。

【0107】

このとき、番組のタイトルとしての強調単語列において、強調対象単語は、強調表示される。

【0108】

その後、ユーザが、再生候補番組の中から、再生を行う番組として、1 つの番組を、操作部 3 1 を操作すること等により選択すると、レコーダ機能部 2 0 では、その番組が再生される。

【0109】

なお、図 2 では、2 つの発音シンボル変換部 5 2 及び 5 5 が設けられているが、2 つの発音シンボル変換部 5 2 及び 5 5 は、1 つの発音シンボル変換部で兼用することが可能である。

【0110】

また、図 2 では、検索結果対象記憶部 5 3 に、検索結果対象単語列（の表記シンボルとしてのテキスト）を記憶させておくこととしたが、検索結果対象記憶部 5 3 には、検索結果対象単語列の他、その検索結果対象単語列を発音シンボルに変換した検索結果対象発音シンボル列を記憶させておくことが可能である。

【0111】

また、図 2 では、音声認識部 5 1 において、入力音声を音声認識し、音声認識結果（の表記シンボル）を出力するようにしたが、音声認識部 5 1 では、音声認識結果を発音シンボルに変換した認識結果発音シンボル列を出力するようにすることが可能である。

【0112】

この場合、音声検索装置 1 0 は、発音シンボル変換部 5 2 を設けずに構成することができる。

10

20

30

40

50

【0113】

[マッチング部 5 6 によるマッチング]

【0114】

図 3 ないし図 8 を参照して、図 2 のマッチング部 5 6 によるマッチングについて説明する。

【0115】

図 3 は、音声認識結果、及び、検索結果対象単語列のそれぞれの表記を表すシンボルである表記シンボルを用いて行われる、音声認識結果と、検索結果対象単語列としてのテキストとのマッチングを説明する図である。

【0116】

ここで、表記（シンボル）は、発音に一致しないことがある。

【0117】

具体的には、例えば、ひらがな「は」の発音（読み）は、「は」である場合と、「わ」である場合があるが、表記シンボルでは、発音の違いを表現することができない。

【0118】

また、表記シンボルでは、複数の読みがある漢字、すなわち、例えば、「市」については、その読み（発音）が「し」であるのか、又は、「いち」であるのかを、表現することができない。

【0119】

一方、例えば、表記シンボルで表された単語列「都市の世界遺産」と「年の瀬解散」とは、発音はほぼ一致するが、表記シンボルでは、「の」以外は異なる。

【0120】

このため、音声認識結果が、「都市の世界遺産」である場合と、「年の瀬解散」である場合とでは、表記シンボルを用いたマッチングでは、異なるマッチング結果が得られるが、このことは、音声検索の性能に、必ずしも有利ではない。

【0121】

すなわち、図 3 は、表記シンボルを用いたマッチングで、発音は（ほぼ）一致するか、表記が（ほとんど）異なる音声認識結果に対して異なるマッチング結果が得られることが、音声検索の性能に有利でないことを説明する図である。

【0122】

図 3 では、入力音声「都市の世界遺産」の音声認識が行われ、その入力音声「都市の世界遺産」と発音は（ほぼ）一致するが、表記が異なる、誤った音声認識結果「年の瀬解散」が得られている。

【0123】

また、図 3 では、音声認識結果「年の瀬解散」を、「年 / の / 瀬 / 解 / 散」のように、表記シンボル単位に区切って（スラッシュ（ / ）は、区切りを表す）、表記シンボル単位でのマッチングが行われている。

【0124】

さらに、図 3 では、マッチングをとる検索結果対象単語列としての、例えば、番組のタイトルとして、「世界遺産都市の遺産」、「瀬戸の歯医者さん」、及び、「衆院解散の年」の 3 つが用意されている。

【0125】

音声認識結果「年の瀬解散」と、検索結果対象単語列「世界遺産都市の遺産」とでは、表記シンボル単位では、図中、丸印を付してある 1 個の表記シンボル「の」しか一致しない。

【0126】

また、音声認識結果「年の瀬解散」と、検索結果対象単語列「瀬戸の歯医者さん」とでは、表記シンボル単位では、図中、丸印を付してある 2 個の表記シンボル「瀬」及び「の」が一致する。

【0127】

10

20

30

40

50

さらに、音声認識結果「年の瀬解散」と、検索結果対象単語列「衆院解散の年」とでは、表記シンボル単位では、図中、丸印を付してある4個の表記シンボル「解」、「散」、「の」及び「年」が一致する。

【0128】

したがって、表記シンボル単位でのマッチングにおいて求められる、音声認識結果と検索結果対象単語列との類似度としては、音声認識結果「年の瀬解散」と、検索結果対象単語列「衆院解散の年」との類似度が、最も高くなる。

【0129】

すなわち、マッチングにおいて求められる類似度として、例えば、コサイン距離を採用することとする。

【0130】

また、単語列を表すベクトルとして、例えば、単語列に存在する表記シンボルに対応するコンポーネントを1とするとともに、単語列に存在しない表記シンボルに対応するコンポーネントを0とするベクトルを採用し、2つの単語列の類似度としてのコサイン距離を、その2つの単語列を表すベクトルを用いて求めることとする。

【0131】

この場合、表記シンボル単位でのマッチングでは、音声認識結果「年の瀬解散」と、検索結果対象単語列「世界遺産都市の遺産」との類似度として、0.15が、音声認識結果「年の瀬解散」と、検索結果対象単語列「瀬戸の歯医者さん」との類似度として、0.32が、音声認識結果「年の瀬解散」と、検索結果対象単語列「衆院解散の年」との類似度として、

【0132】

したがって、例えば、マッチングの結果得られる類似度が最上位の検索結果対象単語列を、検索結果単語列とすることとすると、入力音声「都市の世界遺産」の音声認識が誤り、音声認識結果「年の瀬解散」が得られた場合には、検索結果対象単語列としての3つの番組のタイトル「世界遺産都市の遺産」、「瀬戸の歯医者さん」、及び、「衆院解散の年」のうちの、「衆院解散の年」が、検索結果単語列とされることになる。

【0133】

入力音声「都市の世界遺産」に対しては、上述の3つの番組のタイトル「世界遺産都市の遺産」、「瀬戸の歯医者さん」、及び、「衆院解散の年」のうちの、1番目の番組のタイトル「世界遺産都市の遺産」が、検索結果単語列とされることが適切である。

【0134】

しかしながら、入力音声「都市の世界遺産」が、発音では一致するが、表記が異なる「年の瀬解散」に音声認識されると、入力音声「都市の世界遺産」に対して適切な番組のタイトル「世界遺産都市の遺産」ではなく、「都市の世界遺産」とはまったく関係がないような番組のタイトル「衆院解散の年」が、検索結果単語列とされる。

【0135】

なお、入力音声「都市の世界遺産」に対して、表記が一致する「都市の世界遺産」が、音声認識結果として得られた場合には、入力音声「都市の世界遺産」に対して適切な番組のタイトル「世界遺産都市の遺産」との類似度が最上位となり、「世界遺産都市の遺産」が、検索結果単語列とされる。

【0136】

以上のように、音声認識結果が、「都市の世界遺産」である場合と、「年の瀬解散」である場合とでは、表記シンボルを用いたマッチングでは、マッチング結果（音声認識結果と、各検索結果対象単語列との類似度）が異なり、その結果、入力音声「都市の世界遺産」に対して適切な番組のタイトル「世界遺産都市の遺産」が、検索結果単語列とされる場合と、そのような適切なタイトルが検索結果単語列とされず、入力音声「都市の世界遺産」とはまったく関係がないような番組のタイトル「衆院解散の年」が、検索結果単語列とされる場合とがある。

【0137】

10

20

30

40

50

そこで、音声検索装置 10 (図 2) のマッチング部 56 では、入力音声に対して適切な番組のタイトルが、検索結果単語列として出力されないことを防止するため、発音シンボルを用いたマッチングが行われる。

【 0 1 3 8 】

ここで、発音シンボルは、例えば、音節、又は、音素の発音を表すシンボルであり、日本語については、例えば、読みを表すひらがなを採用することができる。なお、ここでは、発音シンボルとして、音節を採用することとする。

【 0 1 3 9 】

発音シンボルを用いたマッチングでは、マッチングの単位 (マッチング単位) として、音節 (の 1 つ) や、音節の 2 以上の連鎖、音素 (の 1 つ) 、音素の 2 以上の連鎖等を採用することができる。

10

【 0 1 4 0 】

なお、発音シンボルを用いたマッチングにおいて、どのようなマッチング単位を採用するかによって、マッチング結果、ひいては、音声検索の性能は異なる。

【 0 1 4 1 】

図 4 は、マッチング部 56 (図 2) でのマッチング単位として、音節 2 連鎖 (連続する 2 つの音節 (発音シンボル)) を採用する場合の、図 2 の発音シンボル変換部 52 の処理を説明する図である。

【 0 1 4 2 】

発音シンボル変換部 52 には、音声認識部 51 から、入力音声の音声認識結果 (の表記シンボル) が供給される。

20

【 0 1 4 3 】

発音シンボル変換部 52 は、音声認識部 51 から供給される音声認識結果を、発音シンボルである音節の並びに変換する。

【 0 1 4 4 】

さらに、発音シンボル変換部 52 は、音声認識結果の音節の並びの先頭から、注目する注目音節を、後方に、1音節ずつずらしていきながら、注目音節と、その注目音節の直後の音節との、連続する 2 つの音節である音節 2 連鎖を抽出し、その音節 2 連鎖の並びを、マッチング部 56 (図 2) に供給する。

【 0 1 4 5 】

ここで、音声認識結果の音節の並びも、音節 2 連鎖の並びも、音声認識結果の発音シンボルの並びである認識結果発音シンボル列であるが、特に、音節 2 連鎖の並びを、認識結果 2 連鎖シンボル列ともいう。

30

【 0 1 4 6 】

図 5 は、マッチング部 56 (図 2) でのマッチング単位として、音節 2 連鎖を採用する場合の、図 2 の発音シンボル変換部 55 の処理を説明する図である。

【 0 1 4 7 】

発音シンボル変換部 55 には、検索結果対象記憶部 53 に記憶された検索結果対象単語列としての、番組のタイトル等が、形態素解析部 54 で形態素解析されて供給される。

【 0 1 4 8 】

発音シンボル変換部 55 は、形態素解析部 54 から供給される検索結果対象単語列を、音節の並びに変換する。

40

【 0 1 4 9 】

さらに、発音シンボル変換部 55 は、検索結果対象単語列の音節の並びの先頭から、注目する注目音節を、後方に、1音節ずつずらしていきながら、注目音節と、その注目音節の直後の音節との、連続する 2 つの音節である音節 2 連鎖を抽出し、その音節 2 連鎖の並びを、マッチング部 56 (図 2) に供給する。

【 0 1 5 0 】

ここで、検索結果対象単語列の音節の並びも、音節 2 連鎖の並びも、検索結果対象単語列の発音シンボルの並びである検索結果対象発音シンボル列であるが、特に、音節 2 連鎖

50

の並びを、検索結果対象 2 連鎖シンボル列ともいう。

【 0 1 5 1 】

マッチング部 5 6 において、認識結果発音シンボル列と、検索結果対象発音シンボル列との、音節 2 連鎖の単位でのマッチングとして、認識結果 2 連鎖シンボル列と検索結果対象 2 連鎖シンボル列との類似度としての、例えば、コサイン距離を求める場合、認識結果 2 連鎖シンボル列を表すベクトルである認識結果ベクトルが求められる。

【 0 1 5 2 】

すなわち、マッチング部 5 6 は、例えば、認識結果 2 連鎖シンボル列に存在する音節 2 連鎖に対応するコンポーネントを 1 とするとともに、認識結果 2 連鎖シンボル列に存在しない音節 2 連鎖に対応するコンポーネントを 0 とするベクトルを、認識結果 2 連鎖シンボル列を表す認識結果ベクトルとして求める。

10

【 0 1 5 3 】

さらに、マッチング部 5 6 は、検索結果対象記憶部 5 3 に記憶された各検索結果対象単語列としての、例えば、番組のタイトル等についても、同様に、検索結果対象単語列の検索結果対象 2 連鎖シンボル列を表すベクトルである検索結果対象ベクトルを求める。

【 0 1 5 4 】

そして、マッチング部 5 6 は、認識結果ベクトルと、検索結果対象ベクトルとの内積を、認識結果ベクトルの大きさと検索結果対象ベクトルの大きさとの乗算値で除算した値であるコサイン距離を、音声認識結果と検索結果対象単語列との（発音シンボルの）類似度として求める、音節 2 連鎖をマッチング単位とするマッチングを行う。

20

【 0 1 5 5 】

図 6 は、単語単位でのマッチング、（ 1 つの ）音節単位でのマッチング、及び、音節 2 連鎖単位でのマッチングの結果を示す図である。

【 0 1 5 6 】

なお、図 6 では、図 3 と同様に、入力音声「都市の世界遺産」に対して、誤った音声認識結果「年の瀬解散」が得られており、検索結果対象単語列としての、例えば、番組のタイトルとして、「世界遺産都市の遺産」、「瀬戸の歯医者さん」、及び、「衆院解散の年」の 3 つが用意されている。

【 0 1 5 7 】

また、図 6 では、表記シンボルを用いての、単語単位でのマッチング、発音シンボルを用いての、音節単位でのマッチング、及び、発音シンボルを用いての、音節 2 連鎖単位でのマッチングが行われている。

30

【 0 1 5 8 】

さらに、図 6 では、音声認識結果「年の瀬解散」の単語又は発音シンボル（音節）と一致する、検索結果対象単語列の単語又は発音シンボルには、丸印を付してある。

【 0 1 5 9 】

単語単位でのマッチングでは、音声認識結果「年の瀬解散」と、検索結果対象単語列「世界遺産都市の遺産」、「瀬戸の歯医者さん」、及び、「衆院解散の年」それぞれとの類似度（コサイン距離）として、それぞれ、0.22、0.25、及び、0.75 が求められる。

【 0 1 6 0 】

したがって、例えば、マッチングの結果得られる類似度が最上位の検索結果対象単語列を、検索結果単語列として出力することとすると、入力音声「都市の世界遺産」の音声認識が誤り、音声認識結果「年の瀬解散」が得られた場合には、表記シンボルを用いての、単語単位でのマッチングでは、検索結果対象単語列としての 3 つの番組のタイトル「世界遺産都市の遺産」、「瀬戸の歯医者さん」、及び、「衆院解散の年」のうちの、音声認識結果「年の瀬解散」との類似度が 0.75 で最上位の検索結果対象単語列「衆院解散の年」が、検索結果単語列とされることになる。

40

【 0 1 6 1 】

入力音声「都市の世界遺産」に対しては、上述の 3 つの番組のタイトル「世界遺産都市の遺産」、「瀬戸の歯医者さん」、及び、「衆院解散の年」のうちの、1 番目の番組のタ

50

イトル「世界遺産都市の遺産」が、検索結果単語列とされることが適切である。

【0162】

しかしながら、入力音声「都市の世界遺産」が、発音では一致するが、表記が異なる「年の瀬解散」に音声認識されると、表記シンボルを用いての、単語単位でのマッチングでは、入力音声「都市の世界遺産」に対して適切な番組のタイトル「世界遺産都市の遺産」ではなく、「都市の世界遺産」とはまったく関係がないような番組のタイトル「衆院解散の年」が、検索結果単語列とされる。

【0163】

なお、表記シンボルを用いてのマッチングを、単語単位ではなく、表記シンボル単位で行った場合も、図3で説明したように、入力音声「都市の世界遺産」の誤った音声認識結果「年の瀬解散」に対して、入力音声「都市の世界遺産」とはまったく関係がないような番組のタイトル「衆院解散の年」が、検索結果単語列とされる。

10

【0164】

発音シンボルを用いての、音節単位のマッチングでは、音声認識結果「年の瀬解散」と、検索結果対象単語列「世界遺産都市の遺産」、「瀬戸の歯医者さん」、及び、「衆院解散の年」それぞれとの類似度として、それぞれ、0.82, 1.0、及び、0.75が求められる。

【0165】

したがって、例えば、マッチングの結果得られる類似度が最上位の検索結果対象単語列を、検索結果単語列として出力することとすると、入力音声「都市の世界遺産」の音声認識が誤り、音声認識結果「年の瀬解散」が得られた場合には、発音シンボルを用いての、音節単位でのマッチングでは、検索結果対象単語列としての3つの番組のタイトル「世界遺産都市の遺産」、「瀬戸の歯医者さん」、及び、「衆院解散の年」のうちの、音声認識結果「年の瀬解散」との類似度が1.0で最上位の検索結果対象単語列「瀬戸の歯科医さん」が、検索結果単語列とされる。

20

【0166】

すなわち、入力音声「都市の世界遺産」が、発音では一致するが、表記が異なる「年の瀬解散」に音声認識されると、発音シンボルを用いての、音節単位でのマッチングでは、入力音声「都市の世界遺産」に対して適切な番組のタイトル「世界遺産都市の遺産」ではなく、「都市の世界遺産」とはまったく関係がないような番組のタイトル「瀬戸の歯科医さん」が、検索結果単語列とされる。

30

【0167】

なお、表記シンボルを用いての、単語単位でのマッチングでは、入力音声「都市の世界遺産」に対して適切な番組のタイトル「都市の世界遺産都市の遺産」の類似度が、3つの検索結果対象単語列の中で、第3位（最下位）の値である0.22になっているが、発音シンボルを用いての、音節単位でのマッチングでは、入力音声「都市の世界遺産」に対して適切な番組のタイトル「世界遺産都市の遺産」の類似度が、3つの検索結果対象単語列の中で、第2位の値である0.82になっている。

【0168】

したがって、発音シンボルを用いての、音節単位でのマッチングは、入力音声「都市の世界遺産」に対して適切な番組のタイトル「世界遺産都市の遺産」の類似度が、表示シンボルを用いての、単語単位でのマッチングの場合よりも上位である点で、表示シンボルを用いての、単語単位でのマッチングより有効であるといえることができる。

40

【0169】

発音シンボルを用いての、音節2連鎖単位のマッチングでは、音声認識結果「年の瀬解散」と、検索結果対象単語列「世界遺産都市の遺産」、「瀬戸の歯医者さん」、及び、「衆院解散の年」それぞれとの類似度として、それぞれ、0.68, 0.43、及び、0.48が求められる。

【0170】

したがって、例えば、マッチングの結果得られる類似度が最上位の検索結果対象単語列を、検索結果単語列として出力することとすると、入力音声「都市の世界遺産」の音声認

50

識が誤り、音声認識結果「年の瀬解散」が得られた場合であっても、発音シンボルを用いての、音節2連鎖単位でのマッチングでは、検索結果対象単語列としての3つの番組のタイトル「世界遺産都市の遺産」、「瀬戸の歯医者さん」、及び、「衆院解散の年」のうちの、音声認識結果「年の瀬解散」との類似度が0.68で最上位の検索結果対象単語列、すなわち、入力音声「都市の世界遺産」に対して適切な番組のタイトル「世界遺産都市の遺産」が、検索結果単語列とされる。

【0171】

以上のように、発音シンボルを用いてのマッチングによれば、表記が音声認識結果とまったく異なる検索結果対象単語列であっても、発音が音声認識結果と類似していれば、類似度が大になるので、表記シンボルを用いてのマッチングを行う場合に比較して、入力音声に対応する単語列の検索を、ロバストに行うことができる。

10

【0172】

その結果、発音シンボルを用いてのマッチングによれば、音声認識が誤った場合でも、入力音声に対応する単語列が、検索結果単語列として出力されないことを防止（低減）することができる。

【0173】

ところで、マッチング部56（図1）において、音声認識結果（の認識結果発音シンボル列）と、検索結果対象単語列（の検索結果対象発音シンボル列）との類似度として、コサイン距離を採用する場合、例えば、上述したように、認識結果発音シンボル列に存在する音節（2連鎖）に対応するコンポーネントを1とするとともに、認識結果発音シンボル列に存在しない音節に対応するコンポーネントを0とするベクトルが、認識結果発音シンボル列を表す認識結果ベクトルとして求められる。

20

【0174】

さらに、マッチング部56では、同様にして、検索結果対象単語列の検索結果対象発音シンボル列を表す検索結果対象ベクトルが求められる。

【0175】

ここで、本実施の形態では、認識結果ベクトルのコンポーネントの値を、そのコンポーネントに対応する音節が、認識結果発音シンボル列に存在するかどうかで、1又は0とすることとするが、認識結果ベクトルのコンポーネントの値としては、そのコンポーネントに対応する音節が、認識結果発音シンボル列に出現する頻度であるtf(Term Frequency)を採用することが可能である。

30

【0176】

また、認識結果ベクトルのコンポーネントの値としては、その他、例えば、ある検索結果対象単語列には偏って出現する音節に対しては大になり、多くの検索結果対象単語列に万遍なく出現する音節に対しては小になるidf(Invert Document Frequency)や、tfとidfとの両方を加味したTF-IDFを採用することができる。

【0177】

検索結果対象ベクトルについても、同様である。

【0178】

いま、認識結果ベクトルを、 V_{UTR} と表すとともに、検索結果対象記憶部53（図1）に記憶されたi番目の検索結果対象単語列の検索結果対象ベクトルを、 $V_{TITLE}(i)$ と表すこととすると、音声認識結果と、i番目の検索結果対象単語列との類似度としてのコサイン距離Dは、式(1)に従って計算される。

40

【0179】

$$D = V_{UTR} \cdot V_{TITLE}(i) / (|V_{UTR}| |V_{TITLE}(i)|) \quad \dots (1)$$

【0180】

式(1)において、 \cdot は、内積を表し、 $|x|$ は、ベクトルxの大きさ（ノルム）を表す。したがって、コサイン距離Dは、認識結果ベクトル V_{UTR} と、検索結果対象ベクトル $V_{TITLE}(i)$ との内積 $V_{UTR} \cdot V_{TITLE}(i)$ を、認識結果ベクトル V_{UTR} の大きさ $|V_{UTR}|$ と検索結果対象ベ

50

クトル $V_{TITLE}(i)$ の大きさ $|V_{TITLE}(i)|$ との乗算値 $|V_{UTR}||V_{TITLE}(i)|$ で除算することにより求めることができる。

【0181】

コサイン距離 D は、0.0ないし1.0の範囲の値をとり、値が大きいほど、認識結果ベクトル V_{UTR} が表す認識結果発音シンボル列と、検索結果対象ベクトル $V_{TITLE}(i)$ が表す検索結果対象発音シンボル列とが類似していることを表す。

【0182】

上述したように、コサイン距離 D は、認識結果ベクトル V_{UTR} と、検索結果対象ベクトル $V_{TITLE}(i)$ との内積 $V_{UTR} \cdot V_{TITLE}(i)$ を、認識結果ベクトル V_{UTR} の大きさ $|V_{UTR}|$ と検索結果対象ベクトル $V_{TITLE}(i)$ の大きさ $|V_{TITLE}(i)|$ との乗算値で除算することにより求められるため、コサイン距離 D には、音声認識結果、及び、検索結果対象単語列の長さの相違が影響する。

10

【0183】

ここで、音声認識結果、及び、検索結果対象単語列の長さとは、音声認識結果と検索結果対象単語列とのマッチング、つまり、類似度としてのコサイン距離 D の計算を、表記シンボルを用いて、表記シンボル単位で行う場合には、音声認識結果、及び、検索結果対象単語列の表記シンボルの個数に対応する値であり、例えば、表記シンボルの個数の総和の平方根である。

【0184】

また、類似度の計算を、表記シンボルを用いて、単語単位で行う場合には、音声認識結果、及び、検索結果対象単語列の長さとは、それぞれ、音声認識結果、及び、検索結果対象単語列の単語の個数に対応する値であり、例えば、単語の個数の総和の平方根である。

20

【0185】

さらに、類似度の計算を、発音シンボルを用いて、音韻単位で行う場合には、音声認識結果、及び、検索結果対象単語列の長さとは、それぞれ、音声認識結果、及び、検索結果対象単語列の発音シンボルとしての音節の個数に対応する値であり、例えば、音節の個数の総和の平方根である。

【0186】

また、類似度の計算を、発音シンボルを用いて、音節2連鎖単位で行う場合には、音声認識結果、及び、検索結果対象単語列の長さとは、それぞれ、音声認識結果、及び、検索結果対象単語列の音節2連鎖の個数に対応する値であり、例えば、音節2連鎖の個数の総和の平方根である。

30

【0187】

いま、説明を簡単にするために、音声認識結果と検索結果対象単語列とのマッチングとしてのコサイン距離 D の計算を、表記シンボルを用いて、単語単位で行うこととすると、類似度としての式(1)のコサイン距離 D の演算は、検索結果対象ベクトル $V_{TITLE}(i)$ の大きさ $|V_{TITLE}(i)|$ による除算を含むため、例えば、音声認識結果と同一の単語列を含むが、長さ(ここでは、単語の個数)が、長い検索結果対象単語列と、短い検索結果対象単語列とでは、短い検索結果対象単語列との類似度は高くなり(コサイン距離 D が大になり)、長い検索結果対象単語列との類似度は低くなる(コサイン距離 D が小になる)傾向が強い。

40

【0188】

したがって、長さが長い検索結果対象単語列の一部が、音声認識結果として得られても、その音声認識結果と、長さが長い検索結果対象単語列との類似度が、上位にならず、そのような検索結果対象単語列が、検索結果単語列として出力されないために、入力音声に対応する単語列の検索の精度が劣化することがある。

【0189】

つまり、例えば、長いタイトルの一部が発話された場合に、その長いタイトルの類似度が、上位にならず、その長いタイトルが、検索結果単語列として出力されないことがある。

50

【0190】

また、同様の理由により、所定の検索結果対象単語列と同一の単語列を含むが、長さが、長い音声認識結果と、短い音声認識結果とでは、長い音声認識結果と所定の検索結果対象単語列との類似度は、低くなり、短い音声認識結果と所定の検索結果対象単語列との類似度は、高くなる傾向が強い。

【0191】

したがって、所定の検索結果対象単語列と同一の単語列を含むが、長さが長い音声認識結果については、その所定の検索結果対象単語列の類似度は、上位にならず、その所定の検索結果対象単語列が、検索結果単語列として出力されないために、入力音声に対応する単語列の検索の精度が劣化することがある。

10

【0192】

つまり、例えば、短いタイトルを含む長い発話がされた場合に、その短いタイトルの類似度が、上位にならず、その短いタイトルが、検索結果単語列として出力されないことがある。

【0193】

そこで、マッチング部56(図2)では、音声認識結果、及び、検索結果対象単語列の長さの相違の影響を軽減するように、コサイン距離Dを補正した補正距離を、音声認識結果と検索結果対象単語列との類似度として採用することができる。

【0194】

音声認識結果と検索結果対象単語列との類似度として、補正距離を採用する場合には、上述の音声認識結果と長い検索結果対象単語列との類似度、及び、長い音声認識結果と検索結果対象単語列との類似度が低くなることが防止され、その結果、入力音声に対応する単語列の検索を、ロバストに行うことができ、入力音声に対応する単語列の検索の精度の劣化を防止することができる。

20

【0195】

補正距離としては、第1の補正距離と、第2の補正距離とがある。

【0196】

第1の補正距離は、コサイン距離Dを求める式(1)の演算において、検索結果対象単語列の長さに比例する、検索結果対象ベクトル $V_{TITLE}(i)$ の大きさ $|V_{TITLE}(i)|$ に代えて、検索結果対象単語列の長さに比例しない値 $|V_{UTR}| \times (|V_{TITLE}(i)| / |V_{UTR}|)$ 、すなわち、認識結果ベクトル V_{UTR} の大きさ $|V_{UTR}|$ と検索結果対象ベクトル $V_{TITLE}(i)$ の大きさ $|V_{TITLE}(i)|$ との乗算値の平方根 $(|V_{TITLE}(i)||V_{UTR}|)$ を用いて求められる。

30

【0197】

ここで、コサイン距離Dを求める式(1)の演算において、検索結果対象ベクトル $V_{TITLE}(i)$ の大きさ $|V_{TITLE}(i)|$ に代えて用いる値を、以下、代用サイズ $S(i)$ ともいう。

【0198】

第1の補正距離D1は、式(2)に従って求められる。

【0199】

$$\begin{aligned}
 D1 &= V_{UTR} \cdot V_{TITLE}(i) / (|V_{UTR}|S(i)) \\
 &= V_{UTR} \cdot V_{TITLE}(i) / (|V_{UTR}||V_{UTR}| \times (|V_{TITLE}(i)| / |V_{UTR}|)) \\
 &= V_{UTR} \cdot V_{TITLE}(i) / (|V_{UTR}| (|V_{TITLE}(i)||V_{UTR}|)) \\
 &\quad \dots (2)
 \end{aligned}$$

40

【0200】

式(2)において、認識結果ベクトル V_{UTR} の大きさ $|V_{UTR}|$ と検索結果対象ベクトル $V_{TITLE}(i)$ の大きさ $|V_{TITLE}(i)|$ との乗算値の平方根 $(|V_{TITLE}(i)||V_{UTR}|)$ は、 $|V_{TITLE}(i)|$ が小さい場合、つまり、検索結果対象単語列の長さが短い場合には、 $|V_{TITLE}(i)|$ より大になり、 $|V_{TITLE}(i)|$ が大きい場合、つまり、検索結果対象単語列の長さが長い場合には、 $|V_{TITLE}(i)|$ より小になる。

【0201】

その結果、式(2)に従って求められる第1の補正距離D1は、式(1)に従って求めら

50

れるコサイン距離Dに比較して、音声認識結果の長さに対する検索結果対象単語列の長さとしての、検索結果対象ベクトル $V_{TITLE}(i)$ の大きさ $|V_{TITLE}(i)|$ の違いの影響が少ない値、すなわち、音声認識結果と検索結果対象単語列との長さの相違の影響が軽減された値となる。

【0202】

第2の補正距離は、コサイン距離Dを求める式(1)の演算において、検索結果対象単語列の長さに比例する、検索結果対象ベクトル $V_{TITLE}(i)$ の大きさ $|V_{TITLE}(i)|$ に代えて、認識結果ベクトル V_{UTR} の大きさ $|V_{UTR}|$ を、代用サイズ $S(i)$ として用いて求められる。

【0203】

したがって、第2の補正距離D2は、式(3)に従って求められる。

10

【0204】

$$D2 = V_{UTR} \cdot V_{TITLE}(i) / (|V_{UTR}|S(i)) \\ = V_{UTR} \cdot V_{TITLE}(i) / |V_{UTR}|^2$$

・・・(3)

【0205】

第2の補正距離D2は、検索結果対象ベクトル $V_{TITLE}(i)$ の大きさ $|V_{TITLE}(i)|$ を用いずに求められるので、音声認識結果の長さに対する検索結果対象単語列の長さとしての、検索結果対象ベクトル $V_{TITLE}(i)$ の大きさ $|V_{TITLE}(i)|$ の違いの影響がない値、すなわち、音声認識結果と検索結果対象単語列との長さの相違の影響が軽減(除去)された値となる。

【0206】

20

図7は、音声認識結果と検索結果対象単語列との類似度として、コサイン距離D、第1の補正距離D1、及び、第2の補正距離D2を採用した場合のマッチングのシミュレーションの結果を示す図である。

【0207】

なお、図7のシミュレーションでは、短い発話「世界遺産」に対して、正しい音声認識結果「世界遺産」が得られたこととし、検索結果対象単語列としての番組のタイトルとして、長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」と、短いタイトル「世界情勢」とを採用した。

【0208】

さらに、マッチングは、表記シンボルを用いて、単語単位で行った。

30

【0209】

また、図7では、音声認識結果「世界遺産」の単語「世界/遺産」と一致する、検索結果対象単語列としての番組のタイトルの単語には、アンダーラインを付してある。

【0210】

タイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」では、音声認識結果「世界遺産」に対して、「世界」と「遺産」との2つの単語が一致する。

【0211】

一方、タイトル「世界情勢」では、音声認識結果「世界遺産」に対して、「世界」の1つの単語だけが一致する。

【0212】

40

したがって、タイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」と、タイトル「世界情勢」とでは、音声認識結果「世界遺産」と一致する単語の数が多いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」の類似度の方が、タイトル「世界情勢」の類似度よりも上位になることが適切である。

【0213】

しかしながら、類似度として、コサイン距離Dを採用した場合、長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」の一部「世界遺産」に一致する音声認識結果「世界遺産」については、短いタイトル「世界情勢」の類似度が、0.5となり、長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」の類似度が、0.4472となって、短いタイトル「世界情勢」の類似度の方が、長いタイトル「THE世界

50

遺産都市の遺産スペシャルイタリアローマベネチア」の類似度よりも上位となる。

【0214】

すなわち、類似度として、コサイン距離Dを採用した場合には、長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」の一部「世界遺産」に一致する短い音声認識結果「世界遺産」と、その長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」との長さの相違の影響により、音声認識結果「世界遺産」に対して適切な長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」の類似度が上位にならない。

【0215】

一方、類似度として、補正距離を採用した場合、長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」の類似度の方が、短いタイトル「世界情勢」の類似度よりも上位となる。

10

【0216】

すなわち、類似度として、第1の補正距離D1を採用した場合、音声認識結果「世界遺産」については、短いタイトル「世界情勢」の類似度が、0.5となり、長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」の類似度が、0.6687となって、長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」の類似度の方が、短いタイトル「世界情勢」の類似度よりも上位となる。

【0217】

また、類似度として、第2の補正距離D2を採用した場合、音声認識結果「世界遺産」については、短いタイトル「世界情勢」の類似度が、0.5となり、長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」の類似度が、1.0となって、長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」の類似度の方が、短いタイトル「世界情勢」の類似度よりも上位となる。

20

【0218】

以上のように、類似度として、補正距離を採用した場合には、長い検索結果対象単語列の一部に一致する音声認識結果と、その長い検索結果対象単語列との長さの相違の影響が軽減され、音声認識結果「世界遺産」に対して適切な長いタイトル「THE世界遺産都市の遺産スペシャルイタリアローマベネチア」の類似度が上位になる。

【0219】

図8は、音声認識結果と検索結果対象単語列との類似度として、コサイン距離D、第1の補正距離D1、及び、第2の補正距離D2を採用した場合のマッチングの他のシミュレーションの結果を示す図である。

30

【0220】

なお、図8のシミュレーションでは、長い発話「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」に対して、正しい音声認識結果「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」が得られたこととし、検索結果対象単語列としての番組のタイトルとして、短いタイトル「世界遺産」と、長いタイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」とを採用した。

【0221】

さらに、マッチングは、表記シンボルを用いて、単語単位で行った。

40

【0222】

また、図8では、音声認識結果「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」の単語「世界/遺産/都市/の/遺産/イタリア/ローマ/ベネチア/ナポリ/フィレンツェ」と一致する、検索結果対象単語列としての番組のタイトルの単語には、アンダーラインを付してある。

【0223】

タイトル「世界遺産」では、音声認識結果「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」に対して、「世界」と「遺産」との2つの単語が一致する。

【0224】

50

一方、タイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」では、音声認識結果「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」に対して、「世界」、「遺産」、「イタリア」、及び、「フィレンツェ」の4つの単語が一致する。

【0225】

したがって、タイトル「世界遺産」と、タイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」とでは、音声認識結果「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」と一致する単語の数が多いタイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」の類似度の方が、タイトル「世界遺産」の類似度よりも上位になることが適切である。

【0226】

しかしながら、類似度として、コサイン距離Dを採用した場合、長い音声認識結果「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」については、長いタイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」の類似度が、0.4472となり、短いタイトル「世界遺産」の類似度が、0.4772となって、長いタイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」の類似度の方が、短いタイトル「世界遺産」の類似度よりも上位にならない。

【0227】

すなわち、類似度として、コサイン距離Dを採用した場合には、長い音声認識結果「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」と、短い検索結果対象単語列「世界遺産」との長さの相違の影響により、音声認識結果「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」に対して適切な長いタイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」の類似度が上位にならない。

【0228】

一方、類似度として、補正距離を採用した場合、長いタイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」の類似度の方が、短いタイトル「世界遺産」の類似度よりも上位となる。

【0229】

すなわち、類似度として、第1の補正距離D1を採用した場合、長い音声認識結果「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」については、長いタイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」の類似度が、0.4229となり、短いタイトル「世界遺産」の類似度が、0.2991となって、長いタイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」の類似度の方が、短いタイトル「世界遺産」の類似度よりも上位となる。

【0230】

また、類似度として、第2の補正距離D2を採用した場合、長い音声認識結果「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」については、長いタイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」の類似度が、0.4となり、短いタイトル「世界遺産」の類似度が、0.2となって、長いタイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」の類似度の方が、短いタイトル「世界遺産」の類似度よりも上位となる。

【0231】

以上のように、類似度として、補正距離を採用した場合には、長い音声認識結果と、短い検索結果対象単語列との長さの相違の影響が軽減され、音声認識結果「世界遺産都市の遺産イタリアローマベネチアナポリフィレンツェ」に対して適切な長いタイトル「探検ロマン世界遺産イタリアフィレンツェ歴史地区」の類似度が上位になる。

【0232】

したがって、補正距離によれば、音声認識結果と検索結果対象単語列との長さの相違の影響が軽減されることにより、入力音声に対応する単語列の検索を、ロバストに行うことができ、入力音声に対応する単語列の検索の精度の劣化を防止することができる。

【0233】

10

20

30

40

50

なお、類似度として、認識結果ベクトル V_{UTR} と検索結果対象ベクトル $V_{TITLE}(i)$ との内積 $V_{UTR} \cdot V_{TITLE}(i)$ に比例するコサイン距離 D や、第1の補正距離 $D1$ 、第2の補正距離 $D2$ 等を採用する場合、マッチング部56では、内積 $V_{UTR} \cdot V_{TITLE}(i)$ を求めるときに、認識結果発音シンボル列のマッチング単位と一致する、検索結果対象発音シンボル列のマッチング単位であるマッチシンボルを求めることができる。

【0234】

すなわち、例えば、いま、音節2連鎖（連続する2個の発音シンボル）を、マッチング単位として、マッチングを行うこととすると、そのマッチングとしての類似度の計算において、内積 $V_{UTR} \cdot V_{TITLE}(i)$ は、その内積 $V_{UTR} \cdot V_{TITLE}(i)$ の初期値を0とし、認識結果2連鎖シンボル列のマッチング単位としての音声2連鎖を、順次、注目2連鎖に選択して、検索結果対象2連鎖シンボル列の中に、注目2連鎖に一致する音節2連鎖が存在する場合には、内積 $V_{UTR} \cdot V_{TITLE}(i)$ を1だけインクリメントしていくことで求めることができる。

10

【0235】

検索結果対象2連鎖シンボル列の中に、マッチング単位である注目2連鎖に一致する音節2連鎖が存在する場合、その音節2連鎖は、マッチシンボルであるから、マッチング部56では、内積 $V_{UTR} \cdot V_{TITLE}(i)$ を求めるときに、マッチシンボル、すなわち、検索結果対象2連鎖シンボル列の中に存在する、認識結果2連鎖シンボル列の注目2連鎖に一致する音節2連鎖を求めることができる。

【0236】

ここで、マッチシンボルは、検索結果対象発音シンボル列のマッチング単位（音節2連鎖）の中で、認識結果発音シンボル列のマッチング単位と一致するマッチング単位であり、認識結果発音シンボル列に対する、検索結果対象発音シンボル列のマッチシンボルの数が多いほど、認識結果発音シンボル列と検索結果対象発音シンボル列との類似度は大になる。

20

【0237】

したがって、認識結果発音シンボル列に対する、検索結果対象発音シンボル列のマッチシンボルの数が多いほど、その検索結果対象発音シンボル列の検索結果対象単語列は、検索結果単語列として選択される可能性が高くなるので、検索結果対象単語列（の検索結果対象発音シンボル列）のマッチシンボルは、その検索結果対象単語列が、検索結果単語列として選択されることの、いわば原因（理由）となる。

30

【0238】

また、マッチシンボルは、検索結果対象単語列の中で、発音が入力音声（の音声認識結果）と一致する部分を表す発音シンボルであり、したがって、検索結果対象単語列において、発音がマッチシンボルで表される部分（を含む単語）は、入力音声に対応している（入力音声に対応する部分である）、ということが出来る。

【0239】

[音声検索装置10の処理]

【0240】

図9は、図2の音声検索装置10の処理（音声検索の処理）を説明するフローチャートである。

40

【0241】

ステップS11において、音声検索装置10は、必要な前処理を行う。

【0242】

すなわち、音声検索装置10は、入力音声が供給される前に行うことができる処理を、前処理として行う。

【0243】

ここで、前処理としては、例えば、記録媒体23（図1）に記録されたEPGを構成する構成要素である番組のタイトルや、出演者名、詳細情報等を読み出して、検索結果対象記憶部53に供給し、検索結果対象単語列として記憶させる処理がある。

【0244】

50

さらに、前処理としては、形態素解析部 5 4 及び発音シンボル変換部 5 5 において、検索結果対象記憶部 5 3 に記憶された検索結果対象単語列を、検索結果対象発音シンボル列（検索結果対象 2 連鎖シンボル列）に変換し、マッチング部 5 6 に供給する処理や、テーブル作成部 6 1 において、検索結果対象記憶部 5 3 に記憶された検索結果対象単語列について、発音シンボル対単語テーブルを作成し、テーブル記憶部 6 2 に記憶させる処理等がある。

【 0 2 4 5 】

なお、ステップ S 1 1 の前処理は、例えば、1 日ごとに、所定の時刻に行われる。あるいは、ステップ S 1 1 の前処理は、記録媒体 2 3 に録画されている録画番組が変更されたときや、記録媒体 2 3 に記録されている EPG が変更（更新）されたとき等に行われる。

10

【 0 2 4 6 】

最新の前処理の後、ユーザが発話を行い、その発話としての入力音声、音声認識部 5 1 に供給されると、ステップ S 1 2 において、音声認識部 5 1 は、そこに供給される入力音声を音声認識する。

【 0 2 4 7 】

音声認識部 5 1 が入力音声の音声認識を行うことにより得られる音声認識結果は、発音シンボル変換部 5 2 を介することにより、認識結果発音シンボル列（認識結果 2 連鎖シンボル列）となって、マッチング部 5 6 に供給される。

【 0 2 4 8 】

マッチング部 5 6 は、ステップ S 1 3 において、検索結果対象記憶部 5 3 に記憶されたすべての検索結果対象単語列それぞれについて、音声認識部 5 1 から発音シンボル変換部 5 2 を介して供給される認識結果発音シンボル列と、前処理で得られた検索結果対象単語列の検索結果対象発音シンボル列とのマッチングを、例えば、音節 2 連鎖（連続する 2 個の発音シンボル）をマッチング単位としてとり、そのマッチング結果を、生成部 5 7 と単語選択部 6 3 とに供給する。

20

【 0 2 4 9 】

すなわち、マッチング部 5 6 は、検索結果対象記憶部 5 3 に記憶された各検索結果対象単語列について、音声認識結果との類似度としての、例えば、コサイン距離、又は、補正距離等を計算し、その類似度を、検索結果対象単語列（音声認識結果との類似度が求められた検索結果対象単語列）とともに、マッチング結果として、生成部 5 7 に供給する。

30

【 0 2 5 0 】

さらに、マッチング部 5 6 は、検索結果対象記憶部 5 3 に記憶された各検索結果対象単語列について、音声認識結果との類似度としてのコサイン距離等を求めるときに、認識結果発音シンボル列のマッチング単位である音節 2 連鎖に一致する、検索結果対象発音シンボル列のマッチング単位である音節 2 連鎖を、マッチシンボルとして求め、そのマッチシンボルを、検索結果対象単語列（音声認識結果との類似度が求められた検索結果対象単語列）とともに、マッチング結果として、単語選択部 6 3 に供給する。

【 0 2 5 1 】

生成部 5 7 は、ステップ S 1 4 において、マッチング部 5 6 からのマッチング結果に基づいて、検索結果単語列を生成し、単語選択部 6 3 と強調変換部 6 4 とに供給する。

40

【 0 2 5 2 】

すなわち、生成部 5 7 は、マッチング部 5 6 からの、検索結果対象記憶部 5 3 に記憶された検索結果対象単語列すべてのマッチング結果の中から、類似度が上位の検索結果対象単語列、又は、類似度があらかじめ設定された閾値以上の検索結果対象単語列を、入力音声に対する検索結果単語列として選択し、単語選択部 6 3 と強調変換部 6 4 とに供給する。

【 0 2 5 3 】

単語選択部 6 3 は、ステップ S 1 5 において、生成部 5 7 からの検索結果単語列を、順次、注目単語列に選択する。

【 0 2 5 4 】

50

さらに、単語選択部 6 3 は、テーブル記憶部 6 2 に記憶された発音シンボル対単語テーブルに登録された、対応付け単位と単語とのセットである単語セットのうちの、注目単語列としての検索結果単語列（となった検索結果対象単語列）についての単語セットを参照し、その単語セットの中から、マッチング部 5 6 からの、注目単語列のマッチシンボルに含まれる対応付け単位（連続する 2 個以上の発音シンボル）を含む単語セットを検出する。

【 0 2 5 5 】

そして、単語選択部 6 3 は、注目単語列についての単語セットの中から検出した、注目単語列のマッチシンボルに含まれる対応付け単位を含む単語セットに含まれる単語（対応付け単位に対応付けられている単語）を、注目単語列の中で強調する対象の強調対象単語に選択し、強調変換部 6 4 に供給する。

10

【 0 2 5 6 】

強調変換部 6 4 は、ステップ S 1 6 において、生成部 5 7 からの検索結果単語列を、順次、注目単語列に選択する。

【 0 2 5 7 】

さらに、強調変換部 6 4 は、注目単語列について、単語選択部 6 3 から強調対象単語が供給される場合には、注目単語列のうちの、強調対象単語の表示の属性を変更することで、注目単語列を、そのうちの強調対象単語が強調表示される強調単語列に変換する。

【 0 2 5 8 】

強調変換部 6 4 で得られた強調単語列は、強調変換部 6 4 から出力 I/F 3 3（図 1）に供給され、その後、音声検索装置 1 0 は、音声検索の処理を終了する。

20

【 0 2 5 9 】

なお、検索結果対象単語列が、例えば、番組のタイトルや、出演者名、詳細情報である場合において、音声認識結果との類似度が上位、又は、閾値以上の検索結果対象単語列の中に、タイトル以外の、例えば、出演者名（又は詳細情報）があるときには、生成部 5 7 では、その出演者名とともに、又は、その出演者名に代えて、その出演者名をメタデータとして有する番組のタイトルを、検索結果単語列として選択することが可能である。

【 0 2 6 0 】

[発音シンボル対単語テーブルの作成]

【 0 2 6 1 】

図 1 0 は、図 2 のテーブル作成部 6 1 が作成する発音シンボル対単語テーブルの例を示す図である。

30

【 0 2 6 2 】

発音シンボル対単語テーブルには、上述したように、検索結果対象記憶部 5 3 に記憶された各検索結果対象単語列について、検索結果対象単語列の、連続する 2 個以上の発音シンボルを、単語に対応付ける対応付け単位として、対応付け単位と、その対応付け単位の少なくとも一部を発音シンボルに含む、検索結果対象単語列の単語とが対応付けられた単語セットが登録される。

【 0 2 6 3 】

ここで、例えば、いま、説明を簡単にするために、対応付け単位として、マッチング単位と同一の、連続する 2 個の発音シンボルである音節 2 連鎖を採用することとすると、例えば、検索結果対象単語列「年の瀬解散」については、図 1 0 に示す単語セットが、発音シンボル対単語テーブルに登録される。

40

【 0 2 6 4 】

すなわち、検索結果対象単語列「年の瀬解散」の発音シンボルは、「としのせかいさん」であり、検索結果対象単語列「年の瀬解散」について、対応付け単位としての音節 2 連鎖は、図 4 や図 5 で説明した、マッチング単位としての音節 2 連鎖と同様にして求めることができる。

【 0 2 6 5 】

具体的には、検索結果対象単語列「年の瀬解散」について、対応付け単位としての音節

50

2連鎖は、その発音シンボル「としのせかいさん」の先頭から、注目する注目音節（発音シンボル）を、後方に、1音節ずつずらしていきながら、注目音節と、その注目音節の直後の音節との2つの音節を抽出することで求めることができる。

【0266】

すなわち、検索結果対象単語列「年の瀬解散」については、対応付け単位としての音節2連鎖「とし」、「しの」、「のせ」、「せか」、「かい」、「いさ」、「さん」を得ることができる。

【0267】

そして、発音シンボル対単語テーブルでは、対応付け単位が、その対応付け単位の少なくとも一部を発音シンボルに含む、検索結果対象単語列の単語とが対応付けられ、単語セットとして登録される。

10

【0268】

図10では、検索結果対象単語列「年の瀬解散」についての対応付け単位「とし」が、その対応付け単位「とし」を発音シンボルに含む、検索結果対象単語列「年の瀬解散」の単語「年の瀬」と対応付けられ、単語セット[とし：年の瀬]として登録されている。

【0269】

さらに、図10では、検索結果対象単語列「年の瀬解散」についての対応付け単位「しの」が、その対応付け単位「しの」を発音シンボルに含む、検索結果対象単語列「年の瀬解散」の単語「年の瀬」と対応付けられ、単語セット[しの：年の瀬]として登録されている。

20

【0270】

また、図10では、検索結果対象単語列「年の瀬解散」についての対応付け単位「のせ」が、その対応付け単位「のせ」を発音シンボルに含む、検索結果対象単語列「年の瀬解散」の単語「年の瀬」と対応付けられ、単語セット[のせ：年の瀬]として登録されている。

【0271】

さらに、図10では、検索結果対象単語列「年の瀬解散」についての対応付け単位「せか」が、その対応付け単位「せか」の一部「せ」を発音シンボルに含む、検索結果対象単語列「年の瀬解散」の単語「年の瀬」と対応付けられ、単語セット[せか：年の瀬]として登録されている。

30

【0272】

同様に、図10では、検索結果対象単語列「年の瀬解散」についての対応付け単位「かい」、「いさ」、及び、「さん」のそれぞれが、その対応付け単位の少なくとも一部を発音シンボルに含む、検索結果対象単語列「年の瀬解散」の単語「解散」と対応付けられ、単語セット[かい：解散]、[いさ：解散]、及び、[さん：解散]として登録されている。

【0273】

生成部57において、検索結果対象単語列「年の瀬解散」が、入力音声に対する検索結果単語列として選択された場合、単語選択部63において、図10の発音シンボル対単語テーブルに登録された検索結果単語列（となった検索結果対象単語列）「年の瀬解散」についての単語セットを参照して、検索結果単語列の中で強調する対象の強調対象単語が選択される。

40

【0274】

すなわち、例えば、検索結果単語列「年の瀬解散」のマッチシンボルとして、音節2連鎖「とし」、「しの」、「のせ」、又は、「せか」が、マッチング部56から単語選択部63に供給された場合、単語選択部63では、検索結果単語列「年の瀬解散」についての単語セットのうちの、マッチシンボル「とし」、「しの」、「のせ」、又は、「せか」を含む単語セット[とし：年の瀬]、[しの：年の瀬]、[のせ：年の瀬]、又は、[せか：年の瀬]に含まれる単語「年の瀬」を、強調対象単語に選択する。

【0275】

50

さらに、例えば、検索結果単語列「年の瀬解散」のマッチシンボルとして、音節2連鎖「かい」、「いさ」、又は、「さん」が、マッチング部56から単語選択部63に供給された場合、単語選択部63では、検索結果単語列「年の瀬解散」についての単語セットのうちの、マッチシンボル「かい」、「いさ」、又は、「さん」を含む単語セット[かい：解散]、[いさ：解散]、又は、[さん：解散]に含まれる単語「解散」を、強調対象単語に選択する。

【0276】

図11は、図2のテーブル作成部61が発音シンボル対単語テーブルを作成する第1の作成方法を説明する図である。

【0277】

例えば、いま、形態素解析部54からテーブル作成部61に対して、検索結果対象単語列「世界遺産原点への回帰」の形態素解析結果が供給されたとする。

【0278】

検索結果対象単語列「世界遺産原点への回帰」の形態素解析結果には、検索結果対象単語列「世界遺産原点への回帰」を構成する単語と読み(音韻)のセット「世界 せかい」、「遺産 いさん」、「原点 げんてん」、「へ へ」、「の の」、「回帰 かいき」が含まれる。

【0279】

テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」を構成する各単語「世界」、「遺産」、「原点」、「へ」、「の」、「回帰」について、単語の読みに基づき、単語の発音を表す発音シンボル(列)を求め、さらに、単語の発音シンボルの数Sを求める。

【0280】

ここで、本実施の形態では、単語の発音シンボルは、単語の音節に等しいので、単語の発音シンボルの数Sは、単語の音節の数に等しい。

【0281】

また、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」を構成する各単語「世界」、「遺産」、「原点」、「へ」、「の」、「回帰」の発音シンボルを並べて、検索結果対象発音シンボル列「せかいいさんげんてんへのかいき」を構成する。

【0282】

そして、テーブル作成部61は、検索結果対象発音シンボル列「せかいいさんげんてんへのかいき」から、対応付け単位である音節2連鎖の系列(以下、対応付け単位系列ともいう)「せか」、「かい」、「いい」、「いさ」、「さん」、「んげ」、「げん」、「んて」、「てん」、「んへ」、「への」、「のか」、「かい」、「いき」を生成する。

【0283】

さらに、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」を構成する1番目(最初)の単語「世界」を、注目単語に選択する。

【0284】

そして、テーブル作成部61は、注目単語「世界」と、対応付け単位系列の先頭から、注目単語「世界」の発音シンボルの数Sに等しい3個の対応付け単位「せか」、「かい」、「いい」のそれぞれとを対応付けて、注目単語「世界」の発音シンボルの数Sに等しい数の単語セット[せか：世界]、[かい：世界]、[いい：世界]を構成し、検索結果対象単語列「世界遺産原点への回帰」についての単語セットとして、発音シンボル対単語テーブルに登録する。

【0285】

その後、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の対応付け単位系列から、単語セットを構成した3個の対応付け単位「せか」、「かい」、「いい」を削除することにより、新たな対応付け単位系列「いさ」、「さん」、「んげ」、「げん」、「んて」、「てん」、「んへ」、「への」、「のか」、「かい」、「いき」を生成する。

10

20

30

40

50

【0286】

さらに、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」を構成する2番目の単語「遺産」を、新たな注目単語に選択する。

【0287】

そして、テーブル作成部61は、注目単語「遺産」と、(新たな)対応付け単位系列の先頭から、注目単語「遺産」の発音シンボルの数Sに等しい3個の対応付け単位「いさ」、「さん」、「んげ」のそれぞれとを対応付けて、注目単語「遺産」の発音シンボルの数Sに等しい数の単語セット[いさ：遺産]、[さん：遺産]、[んげ：遺産]を構成し、検索結果対象単語列「世界遺産原点への回帰」についての単語セットとして、発音シンボル対単語テーブルに登録する。

10

【0288】

その後、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の対応付け単位系列から、単語セットを構成した3個の対応付け単位「いさ」、「さん」、「んげ」を削除することにより、新たな対応付け単位系列「げん」、「んて」、「てん」、「んへ」、「への」、「のか」、「かい」、「いき」を生成する。

【0289】

さらに、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」を構成する3番目の単語「原点」を、新たな注目単語に選択する。

【0290】

そして、テーブル作成部61は、注目単語「原点」と、(新たな)対応付け単位系列の先頭から、注目単語「原点」の発音シンボルの数Sに等しい4個の対応付け単位「げん」、「んて」、「てん」、「んへ」のそれぞれとを対応付けて、注目単語「原点」の発音シンボルの数Sに等しい数の単語セット[げん：原点]、[んて：原点]、[てん：原点]、[んへ：原点]を構成し、検索結果対象単語列「世界遺産原点への回帰」についての単語セットとして、発音シンボル対単語テーブルに登録する。

20

【0291】

その後、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の対応付け単位系列から、単語セットを構成した4個の対応付け単位「げん」、「んて」、「てん」、「んへ」を削除することにより、新たな対応付け単位系列「への」、「のか」、「かい」、「いき」を生成し、以下、対応付け単位系列を構成する対応付け単位がなくなるまで、同様の処理を繰り返すことで、検索結果対象単語列「世界遺産原点への回帰」についての単語セットを構成し、発音シンボル対単語テーブルに登録する。

30

【0292】

なお、対応付け単位系列を構成する対応付け単位の数が、注目単語の発音シンボルの数Sに満たない場合には、その注目単語については、対応付け単位を構成する対応付け単位の数に等しい数の単語セットが構成される。

【0293】

すなわち、検索結果対象単語列「世界遺産原点への回帰」を構成する最後の単語「回帰」が、注目単語に選択されたとき、対応付け単位系列は、対応付け単位「かい」、「いき」で構成される。

40

【0294】

したがって、対応付け単位系列を構成する対応付け単位の数は、2個であり、注目単語「回帰」の発音シンボルの数Sである3個に満たない。

【0295】

この場合、テーブル作成部61では、注目単語「回帰」と、対応付け単位系列を構成する2個の対応付け単位「かい」、「いき」のそれぞれとが対応付けられて、注目単語「回帰」の発音シンボルの数Sより少ない2個の単語セット[かい：回帰]、[いき：回帰]が構成され、検索結果対象単語列「世界遺産原点への回帰」についての単語セットとして、発音シンボル対単語テーブルに登録される。

【0296】

50

テーブル作成部 6 1 は、以上のようにして、検索結果対象単語列「世界遺産原点への回帰」の対応付け単位「せか」、「かい」、「いい」、「いさ」、「さん」、「んげ」、「げん」、「んて」、「てん」、「んへ」、「への」、「のか」、「かい」、「いき」のすべてについて、単語セットを構成し、発音シンボル対単語テーブルに登録する。

【0297】

そして、テーブル作成部 6 1 は、検索結果対象単語列「世界遺産原点への回帰」の対応付け単位の中に、同一（の発音シンボル）の対応付け単位が複数存在する場合には、その同一の複数の対応付け単位を含む複数の単語セットを、1つの単語セットにマージする。

【0298】

ここで、検索結果対象単語列「世界遺産原点への回帰」の対応付け単位「せか」、「かい」、「いい」、「いさ」、「さん」、「んげ」、「げん」、「んて」、「てん」、「んへ」、「への」、「のか」、「かい」、「いき」においては、2番目の対応付け単位「かい」と、13番目の対応付け単位「かい」とが同一になっている。

10

【0299】

この場合、2番目の対応付け単位「かい」を含む単語セット[かい：世界]と、13番目の対応付け単位「かい」を含む単語セット[かい：回帰]とは、1つの単語セット[かい：世界、回帰]にマージされる。

【0300】

その結果、対応付け単位「かい」は、単語「世界」と「回帰」との両方に対応付けられる。

20

【0301】

図12は、第1の作成方法で発音シンボル対単語テーブルを作成する処理を説明するフローチャートである。

【0302】

ステップS31において、テーブル作成部61は、検索結果対象記憶部53に記憶された検索結果対象単語列のうちの、形態素解析部54から形態素解析結果が供給される検索結果対象単語列を、注目単語列に選択し、形態素解析部54からの注目単語列の形態素解析結果を取得して、処理は、ステップS32に進む。

【0303】

ステップS32では、テーブル作成部61は、形態素解析部54からの注目単語列の形態素解析結果に基づき、注目単語列を構成する各単語の発音シンボル(列)を求め、その発音シンボルを並べることにより、注目単語列の検索結果対象発音シンボル列を求める(注目単語列を検索結果対象発音シンボル列に変換する)。

30

【0304】

さらに、テーブル作成部61は、注目単語列の検索結果対象発音シンボル列から、対応付け単位の系列である対応付け単位系列を生成し、処理は、ステップS32からステップS33に進む。

【0305】

ステップS33では、テーブル作成部61は、注目単語列を構成する単語のうちの、まだ、注目単語に選択していない、最も先頭側の単語を、注目単語に選択する。

40

【0306】

さらに、テーブル作成部61は、注目単語の形態素解析結果から、注目単語の発音シンボルの数Sを求め、処理は、ステップS33からステップS34に進む。

【0307】

ステップS34では、テーブル作成部61は、注目単語と、対応付け単位系列の先頭から、注目単語の発音シンボルの数Sだけの対応付け単位それぞれとを対応付けて、注目単語の発音シンボルの数Sに等しい数の単語セットを構成し、注目単語列についての単語セットとして、発音シンボル対単語テーブルに登録する。

【0308】

ここで、対応付け単位系列を構成する対応付け単位の数、注目単語の発音シンボルの

50

数Sに満たない場合、ステップS 3 4では、テーブル作成部6 1は、注目単語と、対応付け単位系列を構成するすべての対応付け単位のそれぞれとを対応付け、対応付け単位系列を構成する対応付け単位の数に等しい数の単語セットを構成する。

【0309】

その後、処理は、ステップS 3 4からステップS 3 5に進み、テーブル作成部6 1は、注目単語列の対応付け単位系列から、直前のステップS 3 4で単語セットを構成した対応付け単位を削除することにより、新たな対応付け単位系列を生成して、処理は、ステップS 3 6に進む。

【0310】

ステップS 3 6では、テーブル作成部6 1は、注目単語列を構成する単語の中で、まだ、注目単語に選択されていない単語があるかどうかを判定する。

10

【0311】

ステップS 3 6において、注目単語列を構成する単語の中で、まだ、注目単語に選択されていない単語があると判定された場合、処理は、ステップS 3 3に戻り、上述したように、注目単語列を構成する単語のうちの、まだ、注目単語に選択していない、最も先頭側の単語が、注目単語に選択され、以下、同様の処理が繰り返される。

【0312】

また、ステップS 3 6において、注目単語列を構成する単語の中で、注目単語に選択されていない単語がないと判定された場合、すなわち、注目単語列を構成する単語のすべてが、注目単語に選択された場合、処理は、ステップS 3 7に進み、テーブル作成部6 1は、必要に応じて、発音シンボル対単語テーブルに登録された、注目単語列についての単語セットのマージを行う。

20

【0313】

すなわち、ステップS 3 7では、発音シンボル対単語テーブルに登録された、注目単語列についての単語セットの中に、同一（の発音シンボル）の対応付け単位を含む単語セットが複数存在する場合には、テーブル作成部6 1は、その同一の対応付け単位を含む複数の単語セットを、図11で説明したように、1つの単語セットにマージする。

【0314】

そして、処理は、ステップS 3 7からステップS 3 8に進み、テーブル作成部6 1は、検索結果対象記憶部5 3に記憶された検索結果対象単語列の中で、まだ、注目単語列に選択されていない検索結果対象単語列があるかどうかを判定する。

30

【0315】

ステップS 3 8において、検索結果対象記憶部5 3に記憶された検索結果対象単語列の中で、まだ、注目単語列に選択されていない検索結果対象単語列があると判定された場合、テーブル作成部6 1は、形態素解析部5 4から形態素解析結果が供給されるのを待って、ステップS 3 1に戻り、形態素解析部5 4から形態素解析結果が供給された検索結果対象単語列を、注目単語列に選択し、以下、同様の処理が繰り返される。

【0316】

また、ステップS 3 8において、検索結果対象記憶部5 3に記憶された検索結果対象単語列の中に、注目単語列に選択されていない検索結果対象単語列がないと判定された場合、すなわち、検索結果対象記憶部5 3に記憶された検索結果対象単語列のすべてについて、単語セットが構成され、発音シンボル対単語テーブルに登録された場合、処理は、ステップS 3 9に進み、テーブル作成部6 1は、発音シンボル対単語テーブルを（テーブル記憶部6 2に）出力して、処理は終了する。

40

【0317】

図13は、図2のテーブル作成部6 1が発音シンボル対単語テーブルを作成する第2の作成方法を説明する図である。

【0318】

例えば、いま、図11の場合と同様に、形態素解析部5 4からテーブル作成部6 1に対して、検索結果対象単語列「世界遺産原点への回帰」の形態素解析結果が供給されたとす

50

る。

【0319】

テーブル作成部61は、図11の場合と同様に、検索結果対象単語列「世界遺産原点への回帰」を構成する各単語「世界」、「遺産」、「原点」、「へ」、「の」、「回帰」について、単語の読みに基づき、単語の発音を表す発音シンボル(列)を求め、その発音シンボルを並べて、検索結果対象発音シンボル列「せかいいさんげんてんへのかいき」を構成する。

【0320】

さらに、テーブル作成部61は、図11の場合と同様に、検索結果対象発音シンボル列「せかいいさんげんてんへのかいき」から、対応付け単位、ひいては、対応付け単位系列「せか」、「かい」、「いい」、「いさ」、「さん」、「んげ」、「げん」、「んて」、「てん」、「んへ」、「への」、「のか」、「かい」、「いき」を生成する。

10

【0321】

そして、テーブル作成部61は、対応付け単位系列を構成する各対応付け単位から、1番目の発音シンボルを、その対応付け単位のタグとして抽出し、各対応付け単位のタグを並べた系列である1文字タグ系列「せかいいさんげんてんへのかい」を生成する。

【0322】

その後、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」を構成する1番目の単語「世界」を、注目単語に選択する。

【0323】

さらに、テーブル作成部61は、注目単語「世界」の発音シンボル「せかい」から、まだ、注目シンボルに選択していない最も先頭側の1個の発音シンボル「せ」を、注目シンボルに選択する。

20

【0324】

そして、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の1文字タグ系列「せかいいさんげんてんへのかい」の先頭から終わりに向かって、注目シンボル「せ」に一致する最初のタグ「せ」を検出する。

【0325】

さらに、テーブル作成部61は、注目シンボル「せ」に一致する最初のタグ「せ」が付された対応付け単位「せか」と、注目単語「世界」とを対応付けて、単語セット[せか：世界]を構成し、検索結果対象単語列「世界遺産原点への回帰」についての単語セットとして、発音シンボル対単語テーブルに登録する。

30

【0326】

その後、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の1文字タグ系列「せかいいさんげんてんへのかい」から、注目単語「世界」と対応付けられた対応付け単位「せか」のタグ「せ」を削除して、新たな1文字タグ系列「かいいさんげんてんへのかい」を生成する。

【0327】

さらに、テーブル作成部61は、注目単語「世界」の発音シンボル「せかい」から、まだ、注目シンボルに選択していない最も先頭側の1個の発音シンボル「か」を、注目シンボルに選択する。

40

【0328】

そして、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の1文字タグ系列「かいいさんげんてんへのかい」の先頭から終わりに向かって、注目シンボル「か」に一致する最初のタグ「か」を検出する。

【0329】

さらに、テーブル作成部61は、注目シンボル「か」に一致する最初のタグ「か」が付された対応付け単位「かい」と、注目単語「世界」とを対応付けて、単語セット[かい：世界]を構成し、検索結果対象単語列「世界遺産原点への回帰」についての単語セットとして、発音シンボル対単語テーブルに登録する。

50

【0330】

その後、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の1文字タグ系列「かいいさんげんてんへのかい」から、注目単語「世界」と対応付けられた対応付け単位「かい」のタグ「か」を削除して、新たな1文字タグ系列「いいさんげんてんへのかい」を生成する。

【0331】

さらに、テーブル作成部61は、注目単語「世界」の発音シンボル「せかい」から、まだ、注目シンボルに選択していない最も先頭側の1個の発音シンボル「い」を、注目シンボルに選択する。

【0332】

そして、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の1文字タグ系列「いいさんげんてんへのかい」の先頭から終わりに向かって、注目シンボル「い」に一致する最初のタグ「い」を検出する。

【0333】

さらに、テーブル作成部61は、注目シンボル「い」に一致する最初のタグ「い」が付された対応付け単位「いい」と、注目単語「世界」とを対応付けて、単語セット[いい：世界]を構成し、検索結果対象単語列「世界遺産原点への回帰」についての単語セットとして、発音シンボル対単語テーブルに登録する。

【0334】

その後、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の1文字タグ系列「いいさんげんてんへのかい」から、注目単語「世界」と対応付けられた対応付け単位「いい」のタグ「い」を削除して、新たな1文字タグ系列「いさんげんてんへのかい」を生成する。

【0335】

さらに、テーブル作成部61は、上述したように、注目単語「世界」の発音シンボル「せかい」から、まだ、注目シンボルに選択していない最も先頭側の1個の発音シンボルを、注目シンボルに選択するが、注目単語「世界」の発音シンボル「せかい」には、注目シンボルに選択していない発音シンボルがない。

【0336】

注目単語の発音シンボルに、注目シンボルに選択していない発音シンボルがない場合、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」を構成する2番目の単語「遺産」を、注目単語に選択する。

【0337】

さらに、テーブル作成部61は、注目単語「遺産」の発音シンボル「いさん」から、まだ、注目シンボルに選択していない最も先頭側の1個の発音シンボル「い」を、注目シンボルに選択する。

【0338】

そして、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の1文字タグ系列「いさんげんてんへのかい」の先頭から終わりに向かって、注目シンボル「い」に一致する最初のタグ「い」を検出する。

【0339】

さらに、テーブル作成部61は、注目シンボル「い」に一致する最初のタグ「い」が付された対応付け単位「いさ」と、注目単語「遺産」とを対応付けて、単語セット[いさ：遺産]を構成し、検索結果対象単語列「世界遺産原点への回帰」についての単語セットとして、発音シンボル対単語テーブルに登録する。

【0340】

その後、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の1文字タグ系列「いさんげんてんへのかい」から、注目単語「遺産」と対応付けられた対応付け単位「いさ」のタグ「い」を削除して、新たな1文字タグ系列「さんげんてんへのかい」を生成する。

10

20

30

40

50

【0341】

その後、テーブル作成部61は、上述したように、注目単語「遺産」の発音シンボル「いさん」から、まだ、注目シンボルに選択していない最も先頭側の1個の発音シンボル「さ」を、注目シンボルに選択し、以下、1文字タグ系列を構成するタグがなくなるまで、同様の処理を繰り返すことで、検索結果対象単語列「世界遺産原点への回帰」についての単語セットを構成し、発音シンボル対単語テーブルに登録する。

【0342】

テーブル作成部61では、以上のようにして、検索結果対象単語列「世界遺産原点への回帰」の対応付け単位「せか」、「かい」、「いい」、「いさ」、「さん」、「んげ」、「げん」、「んて」、「てん」、「んへ」、「への」、「のか」、「かい」、「いき」のすべてについて、単語セットが構成され、発音シンボル対単語テーブルに登録される。

10

【0343】

そして、テーブル作成部61は、検索結果対象単語列「世界遺産原点への回帰」の対応付け単位の中に、同一（の発音シンボル）の対応付け単位が複数存在する場合には、その同一の複数の対応付け単位を含む単語セットを、図11で説明したように、1つの単語セットにマージする。

【0344】

すなわち、図13では、図11の場合と同様に、検索結果対象単語列「世界遺産原点への回帰」の対応付け単位「せか」、「かい」、「いい」、「いさ」、「さん」、「んげ」、「げん」、「んて」、「てん」、「んへ」、「への」、「のか」、「かい」、「いき」において、2番目の対応付け単位「かい」と、13番目の対応付け単位「かい」とが同一になっている。

20

【0345】

このため、2番目の対応付け単位「かい」を含む単語セット[かい：世界]と、13番目の対応付け単位「かい」を含む単語セット[かい：回帰]とは、1つの単語セット[かい：世界、回帰]にマージされ、その結果、対応付け単位「かい」は、単語「世界」と「回帰」との両方に対応付けられる。

【0346】

図14は、第2の作成方法で発音シンボル対単語テーブルを作成する処理を説明するフローチャートである。

30

【0347】

ステップS51において、テーブル作成部61は、検索結果対象記憶部53に記憶された検索結果対象単語列のうちの、形態素解析部54から形態素解析結果が供給される検索結果対象単語列を、注目単語列に選択し、形態素解析部54からの注目単語列の形態素解析結果を取得して、処理は、ステップS52に進む。

【0348】

ステップS52では、テーブル作成部61は、形態素解析部54からの注目単語列の形態素解析結果に基づき、注目単語列を構成する各単語の発音シンボル（列）を求め、その発音シンボルを並べることにより、注目単語列の検索結果対象発音シンボル列を求める（注目単語列を検索結果対象発音シンボル列に変換する）。

40

【0349】

さらに、テーブル作成部61は、注目単語列の検索結果対象発音シンボル列から、対応付け単位の系列である対応付け単位系列を生成し、処理は、ステップS52からステップS53に進む。

【0350】

ステップS53では、テーブル作成部61は、対応付け単位系列を構成する各対応付け単位から、1番目の発音シンボルを、その対応付け単位のタグとして抽出し、各対応付け単位のタグを並べた系列である1文字タグ系列を生成して、処理は、ステップS54に進む。

【0351】

50

ステップ S 5 4 では、テーブル作成部 6 1 は、注目単語列を構成する単語のうちの、まだ、注目単語に選択していない、最も先頭側の単語を、注目単語に選択し、処理は、ステップ S 5 5 に進む。

【 0 3 5 2 】

ステップ S 5 5 では、テーブル作成部 6 1 は、注目単語の発音シンボルから、まだ、注目シンボルに選択していない最も先頭側の 1 個の発音シンボルを、注目シンボルに選択して、処理は、ステップ S 5 6 に進む。

【 0 3 5 3 】

ステップ S 5 6 では、テーブル作成部 6 1 は、注目単語列の 1 文字タグ系列の先頭から終わりに向かって、注目シンボルに一致する最初のタグを検出する。

10

【 0 3 5 4 】

さらに、テーブル作成部 6 1 は、注目シンボルに一致する最初のタグが付された対応付け単位と、注目単語とを対応付けて、単語セットを構成し、注目単語列についての単語セットとして、発音シンボル対単語テーブルに登録して、処理は、ステップ S 5 6 からステップ S 5 7 に進む。

【 0 3 5 5 】

ステップ S 5 7 では、テーブル作成部 6 1 は、注目単語列の 1 文字タグ系列から、注目単語と対応付けられた対応付け単位のタグを削除して、新たな 1 文字タグ系列を生成し、処理は、ステップ S 5 8 に進む。

【 0 3 5 6 】

ステップ S 5 8 では、テーブル作成部 6 1 は、注目単語の発音シンボルの中に、注目シンボルに選択していない発音シンボルがあるかどうかを判定する。

20

【 0 3 5 7 】

ステップ S 5 8 において、注目単語の発音シンボルの中に、注目シンボルに選択していない発音シンボルがあると判定された場合、処理は、ステップ S 5 5 に戻り、テーブル作成部 6 1 は、注目単語の発音シンボルから、まだ、注目シンボルに選択していない最も先頭側の 1 個の発音シンボルを、注目シンボルに選択し、以下、同様の処理が繰り返される。

【 0 3 5 8 】

また、ステップ S 5 8 において、注目単語の発音シンボルの中に、注目シンボルに選択していない発音シンボルがないと判定された場合、処理は、ステップ S 5 9 に進み、テーブル作成部 6 1 は、注目単語列を構成する単語の中で、まだ、注目単語に選択されていない単語があるかどうかを判定する。

30

【 0 3 5 9 】

ステップ S 5 9 において、注目単語列を構成する単語の中で、まだ、注目単語に選択されていない単語があると判定された場合、処理は、ステップ S 5 4 に戻り進み、上述したように、注目単語列を構成する単語のうちの、まだ、注目単語に選択していない、最も先頭側の単語が、注目単語に選択され、以下、同様の処理が繰り返される。

【 0 3 6 0 】

また、ステップ S 5 9 において、注目単語列を構成する単語の中で、注目単語に選択されていない単語がないと判定された場合、すなわち、注目単語列を構成する単語のすべてが、注目単語に選択された場合、処理は、ステップ S 6 0 に進み、テーブル作成部 6 1 は、必要に応じて、発音シンボル対単語テーブルに登録された、注目単語列についての単語セットのマージを行う。

40

【 0 3 6 1 】

すなわち、ステップ S 6 0 では、発音シンボル対単語テーブルに登録された、注目単語列についての単語セットの中に、同一（の発音シンボル）の対応付け単位を含む単語セットが複数存在する場合には、テーブル作成部 6 1 は、その同一の対応付け単位を含む複数の単語セットを、図 1 1 及び図 1 3 で説明したように、1 つの単語セットにマージする。

【 0 3 6 2 】

50

そして、処理は、ステップ S 6 0 からステップ S 6 1 に進み、テーブル作成部 6 1 は、検索結果対象記憶部 5 3 に記憶された検索結果対象単語列の中で、まだ、注目単語列に選択されていない検索結果対象単語列があるかどうかを判定する。

【0363】

ステップ S 6 1 において、検索結果対象記憶部 5 3 に記憶された検索結果対象単語列の中で、まだ、注目単語列に選択されていない検索結果対象単語列があると判定された場合、テーブル作成部 6 1 は、形態素解析部 5 4 から形態素解析結果が供給されるのを待って、ステップ S 5 1 に戻り、形態素解析部 5 4 から形態素解析結果が供給された検索結果対象単語列を、注目単語列に選択し、以下、同様の処理が繰り返される。

【0364】

また、ステップ S 6 1 において、検索結果対象記憶部 5 3 に記憶された検索結果対象単語列の中に、注目単語列に選択されていない検索結果対象単語列がないと判定された場合、すなわち、検索結果対象記憶部 5 3 に記憶された検索結果対象単語列のすべてについて、単語セットが構成され、発音シンボル対単語テーブルに登録された場合、処理は、ステップ S 6 2 に進み、テーブル作成部 6 1 は、発音シンボル対単語テーブルを（テーブル記憶部 6 2 に）出力して、処理は終了する。

【0365】

図 1 5 は、図 2 のテーブル作成部 6 1 が作成する発音シンボル対単語テーブルの他の例を示す図である。

【0366】

図 1 0 の発音シンボル対単語テーブルでは、対応付け単位として、連続する 2 個の発音シンボルである音節 2 連鎖を採用したが、図 1 5 の発音シンボル対単語テーブルでは、対応付け単位として、連続する 3 個の発音シンボルである音節 3 連鎖（連続する 3 つの音節）が採用されている。

【0367】

図 1 5 の発音シンボル対単語テーブルには、検索結果単語列「世界遺産原点への回帰」についての単語セットが登録されており、その単語セットにおいては、音節 3 連鎖である対応付け単位と、その対応付け単位の少なくとも一部を発音シンボルに含む、検索結果対象単語列の単語とが対応付けられている。

【0368】

検索結果単語列「世界遺産原点への回帰」について、音節 3 連鎖である対応付け単位は、音節 2 連鎖の場合と同様にして求めることができる。

【0369】

すなわち、検索結果単語列「世界遺産原点への回帰」について、音節 3 連鎖である対応付け単位は、その発音シンボル列「せかいいさんげんてんへのかいき」の先頭から、注目する注目音節（発音シンボル）を、後方に、1 音節ずつずらしていきながら、注目音節と、その注目音節の直後の 2 つの音節との 3 つの音節を抽出することで求めることができる。

【0370】

ここで、検索結果単語列「世界遺産原点への回帰」については、音節 3 連鎖である対応付け単位として、「せかい」、「かいい」、「いいさ」、「いさん」、「さんげ」、「んげん」、「げんて」、「んてん」、「てんへ」、「んへの」、「へのか」、「のかい」、「かいき」を得ることができる。

【0371】

そして、検索結果単語列「世界遺産原点への回帰」についての発音シンボル対単語テーブルは、上述した第 1 の作成方法や第 2 の作成方法によって作成することができる。

【0372】

なお、対応付け単位としての発音シンボルの数 S1 は、マッチング単位としての発音シンボルの数 S2 以下である必要がある。

【0373】

10

20

30

40

50

上述したように、単語選択部 6 3 は、発音シンボル対単語テーブルに登録された、検索結果単語列（となった検索結果対象単語列）についての単語セットの中で、検索結果単語列のマッチング単位であるマッチシンボルに含まれる対応付け単位を含む単語セットを検出するので、対応付け単位としての発音シンボルの数 $S1$ が、マッチング単位としての発音シンボルの数 $S2$ よりも多い場合（ $S1 > S2$ ）には、 $S1$ 個の発音シンボルである対応付け単位が、その $S1$ 個より少ない $S2$ 個の発音シンボルであるマッチシンボルに含まれることがなく、 $S2$ 個の発音シンボルであるマッチシンボルに含まれる、 $S1$ 個の発音シンボルである対応付け単位は、存在しないからである。

【 0 3 7 4 】

[単語選択部 6 3 の処理]

10

【 0 3 7 5 】

図 1 6 は、図 9 のステップ S 1 5 で、単語選択部 6 3 が行う処理を説明するフローチャートである。

【 0 3 7 6 】

単語選択部 6 3 は、生成部 5 7 から供給される検索結果単語列（類似度が上位、又は、閾値以上の検索結果対象単語列）を、順次、注目単語列に選択し、注目単語列に対して、ステップ S 7 1 ないし S 7 3 の処理を行う。

【 0 3 7 7 】

すなわち、単語選択部 6 3 は、ステップ S 7 1 において、マッチング部 5 6 から供給されるマッチシンボルの中から、注目単語のマッチシンボル、つまり、認識結果発音シンボル列のマッチング単位と一致する、注目単語列の発音シンボル（検索結果対象発音シンボル列）のマッチング単位を取得する。

20

【 0 3 7 8 】

そして、単語選択部 6 3 は、ステップ S 7 2 において、テーブル記憶部 6 2 に記憶された発音シンボル対単語テーブルに登録された単語セットのうちの、注目単語列としての検索結果単語列（となった検索結果対象単語列）についての単語セットを参照し、その単語セットの中から、注目単語列のマッチシンボルに含まれる対応付け単位（本実施の形態では、マッチング単位と対応付け単位の発音シンボルの数が同一であるため、マッチシンボルに一致する対応付け単位）を含む単語セットを検出する。

【 0 3 7 9 】

さらに、単語選択部 6 3 は、注目単語列についての単語セットの中から検出した、注目単語列のマッチシンボルに一致する対応付け単位を含む単語セットに含まれる単語、つまり、注目単語列のマッチシンボルに一致する対応付け単位に対応付けられている単語を、注目単語列の中で強調する対象の強調対象単語に選択する。

30

【 0 3 8 0 】

そして、単語選択部 6 3 は、注目単語列のマッチシンボルのすべてに対して、強調対象単語を選択した後、ステップ S 7 3 において、その強調対象単語を、強調変換部 6 4 に出力する。

【 0 3 8 1 】

[具体例]

40

【 0 3 8 2 】

図 1 7 は、発音シンボル対単語テーブルのさらに他の例を示す図である。

【 0 3 8 3 】

図 1 7 の発音シンボル対単語テーブルでは、検索結果対象単語列「The世界遺産」、「世界遺産 原点への回帰」、及び、「速報 年の瀬に解散」のそれぞれについて、対応付け単位としての音節 2 連鎖と、その対応付け単位の少なくとも一部を発音シンボルに含む、検索結果対象単語列の単語とが対応付けられた単語セットが登録されている。

【 0 3 8 4 】

図 1 7 の発音シンボル対単語テーブルによれば、音声検索装置 1 0 において、以下のような強調単語列が得られる。

50

【0385】

すなわち、いま、ユーザが、入力音声「都市の世界遺産」を発話し、音声認識部51において、その入力音声「都市の世界遺産」に対して、正しい音声認識結果「都市の世界遺産」が得られたとする。

【0386】

さらに、入力音声「都市の世界遺産」の正しい音声認識結果「都市の世界遺産」に対して、生成部57において、検索結果対象単語列「The世界遺産」、「探検ロマン世界遺産 rome帝国植民都市」、「世界遺産100ギリシャ都市国家」、「The世界遺産 空中都市マチュピチュ」、及び、「世界遺産 古代都市テーベ」が、検索結果単語列として選択され、単語選択部63、及び、強調変換部64に供給されたこととする。

10

【0387】

いま、検索結果単語列(となった検索結果対象単語列)「The世界遺産」、「探検ロマン世界遺産 rome帝国植民都市」、「世界遺産100ギリシャ都市国家」、「The世界遺産 空中都市マチュピチュ」、及び、「世界遺産 古代都市テーベ」のうちの、例えば、「The世界遺産」を注目単語列とすると、注目単語列「The世界遺産」のマッチング単位である音節2連鎖は、「ざせ/せか/かい/いい/いさ/さん」となる(スラッシュ(/)は、マッチング単位の区切りを表す)。

【0388】

一方、入力音声「都市の世界遺産」の正しい音声認識結果「都市の世界遺産」のマッチング単位である音節2連鎖は、「とし/しの/のせ/せか/かい/いい/いさ/さん」となる。

20

【0389】

したがって、マッチング部56では、注目単語列「The世界遺産」のマッチング単位「ざせ/せか/かい/いい/いさ/さん」の中で、正しい音声認識結果「都市の世界遺産」のマッチング単位「とし/しの/のせ/せか/かい/いい/いさ/さん」と一致するマッチング単位であるマッチンボルとして、マッチング単位「せか」、「かい」、「いい」、「いさ」、「さん」が求められる。

【0390】

そして、図17の発音シンボル対単語テーブルにおいて、注目単語列「The世界遺産」については、マッチンボル「せか」に一致する対応付け単位「せか」には、単語「世界」が、マッチンボル「かい」に一致する対応付け単位「かい」には、単語「世界」が、マッチンボル「いい」に一致する対応付け単位「いい」には、単語「遺産」が、マッチンボル「いさ」に一致する対応付け単位「いさ」には、単語「遺産」が、マッチンボル「さん」に一致する対応付け単位「さん」には、単語「遺産」が、それぞれ対応付けられている。

30

【0391】

したがって、単語選択部63では、単語「世界」、及び、「遺産」のそれぞれが、強調対象単語として選択され、強調変換部64に供給される。

【0392】

その結果、強調変換部64では、注目単語列「The世界遺産」が、そのうちの強調対象単語「世界」、及び、「遺産」が強調された強調単語列「The世界遺産」に変換される。

40

【0393】

強調対象単語「世界」、及び、「遺産」が強調された強調単語列「The世界遺産」によれば、ユーザは、視覚的に、一目で、強調単語列「The世界遺産」のうちの、強調対象単語「世界」、及び、「遺産」が、入力音声「都市の世界遺産」に対応する部分であり、その強調対象単語「世界」、及び、「遺産」の発音(読み)が、入力音声「都市の世界遺産」の発音の少なくとも一部と一致したために、強調単語列「The世界遺産」が、検索結果となっていることを把握することができる。

【0394】

したがって、音声検索装置10によれば、利便性の高い検索結果を提供することができる。

50

【0395】

次に、ユーザが、入力音声「都市の世界遺産」を発話し、音声認識部51において、その入力音声「都市の世界遺産」に対して、誤った音声認識結果「年の瀬解散」が得られたとする。

【0396】

さらに、入力音声「都市の世界遺産」の誤った音声認識結果「年の瀬解散」に対して、生成部57において、検索結果対象単語列「The世界遺産」、「探検ロマン世界遺産 ローマ帝国植民都市」、「世界遺産100ギリシャ都市国家」、「The世界遺産 空中都市マチュピチュ」、及び、「世界遺産 古代都市テーベ」が、検索結果単語列として選択され、単語選択部63、及び、強調変換部64に供給されたこととする。

10

【0397】

いま、検索結果単語列(となった検索結果対象単語列)「The世界遺産」、「探検ロマン世界遺産 ローマ帝国植民都市」、「世界遺産100ギリシャ都市国家」、「The世界遺産 空中都市マチュピチュ」、及び、「世界遺産 古代都市テーベ」のうちの、例えば、「The世界遺産」を注目単語列とすると、注目単語列「The世界遺産」のマッチング単位である音節2連鎖は、上述したように、「ざせ/せか/かい/いい/いさ/さん」となる。

【0398】

一方、入力音声「都市の世界遺産」の誤った音声認識結果「年の瀬解散」のマッチング単位である音節2連鎖は、「とし/しの/のせ/せか/かい/いさ/さん」となる。

【0399】

したがって、マッチング部56では、注目単語列「The世界遺産」のマッチング単位「ざせ/せか/かい/いい/いさ/さん」の中で、誤った音声認識結果「年の瀬解散」のマッチング単位「とし/しの/のせ/せか/かい/いさ/さん」と一致するマッチング単位であるマッチシンボルとして、マッチング単位「せか」、「かい」、「いさ」、「さん」が求められる。

20

【0400】

そして、図17の発音シンボル対単語テーブルにおいて、注目単語列「The世界遺産」については、マッチシンボル「せか」に一致する対応付け単位「せか」には、単語「世界」が、マッチシンボル「かい」に一致する対応付け単位「かい」には、単語「世界」が、マッチシンボル「いさ」に一致する対応付け単位「いさ」には、単語「遺産」が、マッチシンボル「さん」に一致する対応付け単位「さん」には、単語「遺産」が、それぞれ対応付けられている。

30

【0401】

したがって、単語選択部63では、単語「世界」、及び、「遺産」のそれぞれが、強調対象単語として選択され、強調変換部64に供給される。

【0402】

その結果、強調変換部64では、注目単語列「The世界遺産」が、そのうちの強調対象単語「世界」、及び、「遺産」が強調された強調単語列「The世界遺産」に変換される。

【0403】

強調対象単語「世界」、及び、「遺産」が強調された強調単語列「The世界遺産」によれば、ユーザは、視覚的に、一目で、強調単語列「The世界遺産」のうちの、強調対象単語「世界」、及び、「遺産」が、入力音声「都市の世界遺産」に対応する部分であり、その強調対象単語「世界」、及び、「遺産」の発音が、入力音声「都市の世界遺産」の発音の少なくとも一部と一致したために、強調単語列「The世界遺産」が、検索結果となっていることを把握することができる。

40

【0404】

さらに、入力音声「都市の世界遺産」に対して、誤った音声認識結果「年の瀬解散」が得られても、発音が、入力音声「都市の世界遺産」の発音の少なくとも一部と一致する検索結果対象単語列「The世界遺産」を、検索結果(検索結果単語列)として得ることができる。

50

【0405】

したがって、音声検索装置10によれば、入力音声に対応する単語列が、音声認識が誤ることに起因して、検索結果単語列として出力されないことを防止（低減）するとともに、利便性の高い検索結果を提供することができる。

【0406】

次に、ユーザが、入力音声「都市の世界遺産」を話し、音声認識部51において、その入力音声「都市の世界遺産」に対して、誤った音声認識結果「年の瀬解散」が得られたとする。

【0407】

さらに、入力音声「都市の世界遺産」の誤った音声認識結果「年の瀬解散」に対して、生成部57において、検索結果対象単語列「世界遺産 原点への回帰」が、検索結果単語列として選択され、単語選択部63、及び、強調変換部64に供給されたこととする。

10

【0408】

いま、検索結果単語列（となった検索結果対象単語列）「世界遺産 原点への回帰」を注目単語列とすると、注目単語列「世界遺産 原点への回帰」のマッチング単位である音節2連鎖は、「せか/かい/いい/いさ/さん/んげ/げん/んて/てん/んへ/への/のか/かい/いき」となる。

【0409】

一方、入力音声「都市の世界遺産」の誤った音声認識結果「年の瀬解散」のマッチング単位である音節2連鎖は、「とし/しの/のせ/せか/かい/いさ/さん」となる。

20

【0410】

したがって、マッチング部56では、注目単語列「世界遺産 原点への回帰」のマッチング単位「せか/かい/いい/いさ/さん/んげ/げん/んて/てん/んへ/への/のか/かい/いき」の中で、誤った音声認識結果「年の瀬解散」のマッチング単位「とし/しの/のせ/せか/かい/いさ/さん」と一致するマッチング単位であるマッチシンボルとして、マッチング単位「せか」、「かい」、「いさ」、「さん」が求められる。

【0411】

そして、図17の発音シンボル対単語テーブルにおいて、注目単語列「世界遺産 原点への回帰」については、マッチシンボル「せか」に一致する対応付け単位「せか」には、単語「世界」が、マッチシンボル「かい」に一致する対応付け単位「かい」には、単語「世界」及び「回帰」の2つが、マッチシンボル「いさ」に一致する対応付け単位「いさ」には、単語「遺産」が、マッチシンボル「さん」に一致する対応付け単位「さん」には、単語「遺産」が、それぞれ対応付けられている。

30

【0412】

したがって、単語選択部63では、単語「世界」、「遺産」、及び、「回帰」のそれぞれが、強調対象単語として選択され、強調変換部64に供給される。

【0413】

その結果、強調変換部64では、注目単語列「世界遺産 原点への回帰」が、そのうちの強調対象単語「世界」、「遺産」、及び、「回帰」が強調された強調単語列「世界遺産 原点への回帰」に変換される。

40

【0414】

強調対象単語「世界」、「遺産」、及び、「回帰」が強調された強調単語列「世界遺産 原点への回帰」によれば、ユーザは、視覚的に、一目で、強調単語列「世界遺産 原点への回帰」のうちの、強調対象単語「世界」、「遺産」、及び、「回帰」が、入力音声「都市の世界遺産」に対応する部分であり、その強調対象単語「世界」、「遺産」、及び、「回帰」の発音が、入力音声「都市の世界遺産」の発音の少なくとも一部と一致したために、強調単語列「世界遺産 原点への回帰」が、検索結果となっていることを把握することができる。

【0415】

さらに、入力音声「都市の世界遺産」に対して、誤った音声認識結果「年の瀬解散」が

50

得られても、発音が、入力音声「都市の世界遺産」の発音の少なくとも一部と一致する検索結果対象単語列「世界遺産 原点への回帰」を、検索結果（検索結果単語列）として得ることができる。

【0416】

したがって、音声検索装置10によれば、入力音声に対応する単語列が、音声認識が誤ることに起因して、検索結果単語列として出力されないことを防止（低減）するとともに、利便性の高い検索結果を提供することができる。

【0417】

次に、ユーザが、入力音声「都市の世界遺産」を発話し、音声認識部51において、その入力音声「都市の世界遺産」に対して、正しい音声認識結果「都市の世界遺産」が得られたとする。

10

【0418】

さらに、入力音声「都市の世界遺産」の正しい音声認識結果「都市の世界遺産」に対して、生成部57において、検索結果対象単語列「速報 年の瀬に解散」が、検索結果単語列として選択され、単語選択部63、及び、強調変換部64に供給されたこととする。

【0419】

いま、検索結果単語列（となった検索結果対象単語列）「速報 年の瀬に解散」を注目単語列とすると、注目単語列「速報 年の瀬に解散」のマッチング単位である音節2連鎖は、「そく/くほ/ほう/うと/とし/しの/のせ/せに/にか/かい/いさ/さん」となる。

20

【0420】

一方、入力音声「都市の世界遺産」の正しい音声認識結果「都市の世界遺産」のマッチング単位である音節2連鎖は、「とし/しの/のせ/せか/かい/いい/いさ/さん」となる。

【0421】

したがって、マッチング部56では、注目単語列「速報 年の瀬に解散」のマッチング単位「そく/くほ/ほう/うと/とし/しの/のせ/せに/にか/かい/いさ/さん」の中で、正しい音声認識結果「都市の世界遺産」のマッチング単位「とし/しの/のせ/せか/かい/いい/いさ/さん」と一致するマッチング単位であるマッチシンボルとして、マッチング単位「とし」、「しの」、「のせ」、「かい」、「いさ」、「さん」が求められる。

30

【0422】

そして、図17の発音シンボル対単語テーブルにおいて、注目単語列「速報 年の瀬に解散」については、マッチシンボル「とし」に一致する対応付け単位「とし」には、単語「年の瀬」が、マッチシンボル「しの」に一致する対応付け単位「しの」には、単語「年の瀬」が、マッチシンボル「のせ」に一致する対応付け単位「のせ」には、単語「年の瀬」が、マッチシンボル「かい」に一致する対応付け単位「かい」には、単語「解散」が、マッチシンボル「いさ」に一致する対応付け単位「いさ」には、単語「解散」が、マッチシンボル「さん」に一致する対応付け単位「さん」には、単語「解散」が、それぞれ対応付けられている。

【0423】

したがって、単語選択部63では、単語「年の瀬」、及び、「解散」のそれぞれが、強調対象単語として選択され、強調変換部64に供給される。

40

【0424】

その結果、強調変換部64では、注目単語列「速報 年の瀬に解散」が、そのうちの強調対象単語「年の瀬」、及び、「解散」が強調された強調単語列「速報 年の瀬に解散」に変換される。

【0425】

強調対象単語「年の瀬」、及び、「解散」が強調された強調単語列「速報 年の瀬に解散」によれば、ユーザは、視覚的に、一目で、強調単語列「速報 年の瀬に解散」のうちの、強調対象単語「年の瀬」、及び、「解散」が、入力音声「都市の世界遺産」に対応する部分であり、その強調対象単語「年の瀬」、及び、「解散」の発音が、入力音声「都市

50

の世界遺産」の発音の少なくとも一部と一致したために、強調単語列「速報 年の瀬に解散」が、検索結果となっていることを把握することができる。

【0426】

すなわち、マッチング部56では、発音シンボルを用いたマッチングが行われるが、発音シンボルを用いたマッチングでは、音声認識結果「都市の世界遺産」とは、表記が似ていないが、発音が似ている部分がある検索結果対象単語列「速報 年の瀬に解散」の類似度が大になり、そのような検索結果対象単語列「速報 年の瀬に解散」が、検索結果単語列に選択されることがある。

【0427】

この場合、検索結果単語列(となった検索結果対象単語列)「速報 年の瀬に解散」が、何らの強調もされずに表示されると、ユーザは、入力音声「都市の世界遺産」とは、表記がまったく異なる検索結果単語列「速報 年の瀬に解散」が、入力音声「都市の世界遺産」に対して検索された理由が分からず、ストレスを感じることもある。

10

【0428】

一方、音声検索装置10では、上述したように、検索結果単語列「速報 年の瀬に解散」が、そのうちの強調対象単語「年の瀬」、及び、「解散」が強調された強調単語列「速報 年の瀬に解散」に変換されるので、検索結果単語列「速報 年の瀬に解散」の表示は、単語「年の瀬」、及び、「解散」を強調して行われる。

【0429】

この場合、ユーザは、検索結果単語列「速報 年の瀬に解散」の表示で強調されている「年の瀬」の発音「としのせ」が、入力音声「都市の世界遺産」の発音「としのせかいいさん」のうちの「としのせ」に一致し、検索結果単語列「速報 年の瀬に解散」の表示で強調されている「解散」の発音「かいさん」が、入力音声「都市の世界遺産」の発音「としのせかいいさん」のうちの「かいいさん」にほぼ一致するという理由で、検索結果単語列「速報 年の瀬に解散」が、入力音声「都市の世界遺産」に対して検索されたことを、容易に把握することができる。

20

【0430】

図18及び図19は、検索結果単語列(となった検索結果対象単語列)の表示例を示す図である。

【0431】

すなわち、図18及び図19は、入力音声「いいとも」に対して、正しい音声認識結果「いいとも」が得られた場合の検索結果単語列の表示例を示している。

30

【0432】

ここで、図18は、強調対象単語を強調していない表示例を示し、図19は、強調対象単語を強調した表示例を示している。

【0433】

図19において、強調対象単語は、矩形で囲むことにより強調されている。

【0434】

図18及び図19から、強調対象単語を強調して表示することにより、ユーザは、入力音声「いいとも」に対して、各検索結果単語列が検索された理由を、容易に把握することができ、さらに、強調対象単語に、ユーザの視線が集中しやすくなることにより、所望の検索結果である(可能性が高い)、入力音声「いいとも」に表記が一致する検索結果単語列を、迅速に見つけ出すことができることが分かる。

40

【0435】

[本発明を適用したコンピュータの説明]

【0436】

次に、上述した一連の処理は、ハードウェアにより行うこともできるし、ソフトウェアにより行うこともできる。一連の処理をソフトウェアによって行う場合には、そのソフトウェアを構成するプログラムが、汎用のコンピュータ等にインストールされる。

【0437】

50

そこで、図20は、上述した一連の処理を実行するプログラムがインストールされるコンピュータの一実施の形態の構成例を示している。

【0438】

プログラムは、コンピュータに内蔵されている記録媒体としてのハードディスク105やROM103に予め記録しておくことができる。

【0439】

あるいはまた、プログラムは、リムーバブル記録媒体111に格納（記録）しておくことができる。このようなリムーバブル記録媒体111は、いわゆるパッケージソフトウェアとして提供することができる。ここで、リムーバブル記録媒体111としては、例えば、フレキシブルディスク、CD-ROM(Compact Disc Read Only Memory)、MO(Magneto Optical)ディスク、DVD(Digital Versatile Disc)、磁気ディスク、半導体メモリ等がある。

10

【0440】

なお、プログラムは、上述したようなリムーバブル記録媒体111からコンピュータにインストールする他、通信網や放送網を介して、コンピュータにダウンロードし、内蔵するハードディスク105にインストールすることができる。すなわち、プログラムは、例えば、ダウンロードサイトから、デジタル衛星放送用の人工衛星を介して、コンピュータに無線で転送したり、LAN(Local Area Network)、インターネットといったネットワークを介して、コンピュータに有線で転送することができる。

【0441】

コンピュータは、CPU(Central Processing Unit)102を内蔵しており、CPU102には、バス101を介して、入出力インタフェース110が接続されている。

20

【0442】

CPU102は、入出力インタフェース110を介して、ユーザによって、入力部107が操作等されることにより指令が入力されると、それに従って、ROM(Read Only Memory)103に格納されているプログラムを実行する。あるいは、CPU102は、ハードディスク105に格納されたプログラムを、RAM(Random Access Memory)104にロードして実行する。

【0443】

これにより、CPU102は、上述したフローチャートにしたがった処理、あるいは上述したブロック図の構成により行われる処理を行う。そして、CPU102は、その処理結果を、必要に応じて、例えば、入出力インタフェース110を介して、出力部106から出力、あるいは、通信部108から送信、さらには、ハードディスク105に記録等させる。

30

【0444】

なお、入力部107は、キーボードや、マウス、マイク等で構成される。また、出力部106は、LCD(Liquid Crystal Display)やスピーカ等で構成される。

【0445】

ここで、本明細書において、コンピュータがプログラムに従って行う処理は、必ずしもフローチャートとして記載された順序に沿って時系列に行われる必要はない。すなわち、コンピュータがプログラムに従って行う処理は、並列的あるいは個別に実行される処理（例えば、並列処理あるいはオブジェクトによる処理）も含む。

40

【0446】

また、プログラムは、1のコンピュータ（プロセッサ）により処理されるものであっても良いし、複数のコンピュータによって分散処理されるものであっても良い。さらに、プログラムは、遠方のコンピュータに転送されて実行されるものであっても良い。

【0447】

なお、本発明の実施の形態は、上述した実施の形態に限定されるものではなく、本発明の要旨を逸脱しない範囲において種々の変更が可能である。

【0448】

すなわち、例えば、音声検索装置10（図2）は、その全部、又は、一部を、インター

50

ネット等のネットワーク上のサーバに実装することができる。

【0449】

具体的には、例えば、音声検索装置10は、その全部を、サーバに実装し、クライアント端末からの入力音声を受け付け、その入力音声に対して強調変換部64が出力する強調単語列を、音声検索の検索結果として、クライアント端末に提供することができる。

【0450】

また、例えば、音声検索装置10は、音声認識部51を、クライアント端末に実装するとともに、音声認識部51以外を、サーバに実行し、クライアント端末では、入力音声を音声認識し、その音声認識結果を、サーバに送信し、サーバでは、クライアント端末からの音声認識結果を受け付け、その音声認識結果に対して強調変換部64が出力する強調単語列を、音声検索の検索結果として、クライアント端末に提供することができる。

10

【符号の説明】

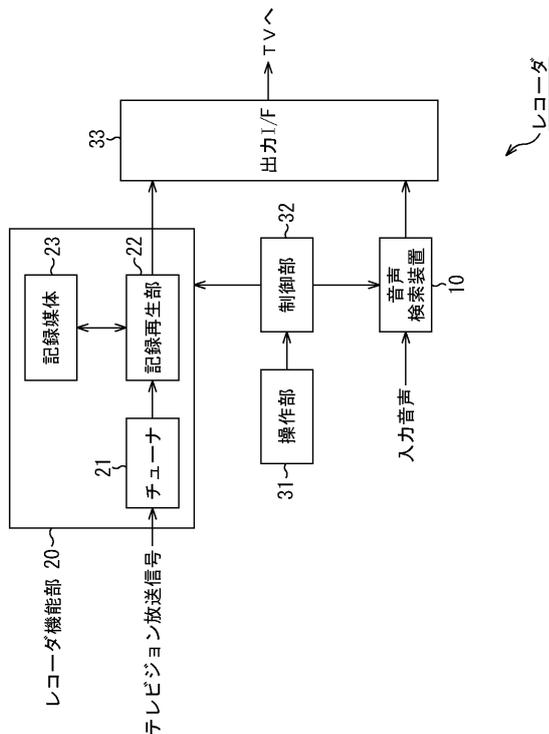
【0451】

10 音声検索装置, 20 レコーダ機能部, 21 チューナ, 22 記録再生部, 23 記録媒体, 31 操作部, 32 制御部, 33 出力I/F, 51 音声認識部, 52 発音シンボル変換部, 53 検索結果対象記憶部, 54 形態素解析部, 55 発音シンボル変換部, 56 マッチング部, 57 生成部, 61 テーブル作成部, 62 テーブル記憶部, 63 単語選択部, 64 強調変換部, 101 バス, 102 CPU, 103 ROM, 104 RAM, 105 ハードディスク, 106 出力部, 107 入力部, 108 通信部, 109 ドライブ, 110 入出力インタフェース, 111 リムーバブル記録媒体

20

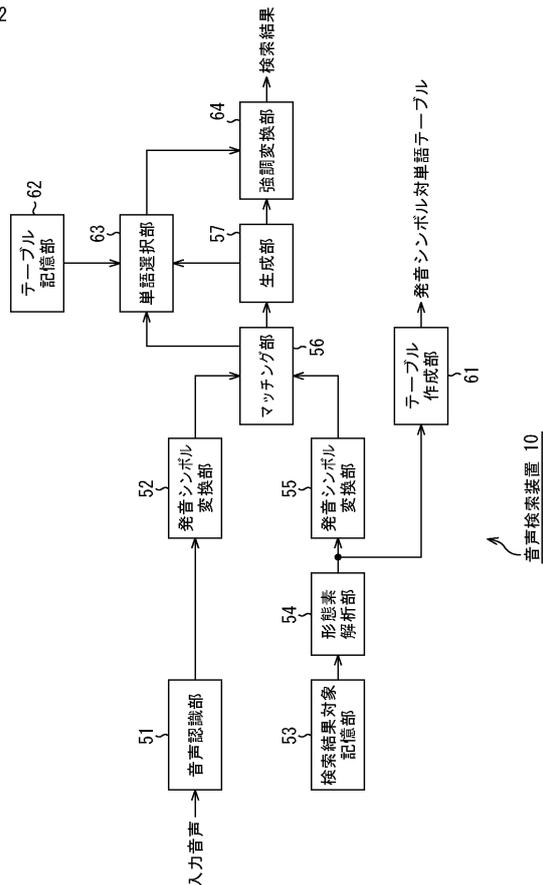
【図1】

図1



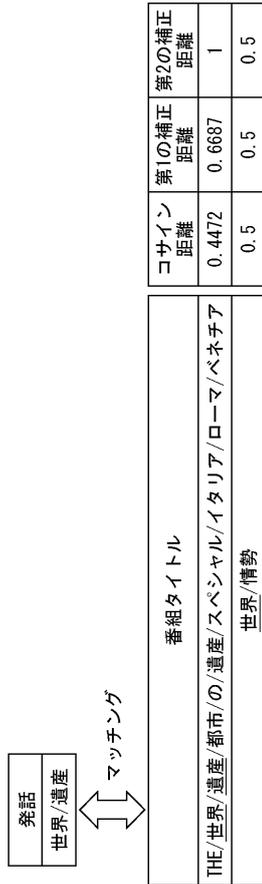
【図2】

図2



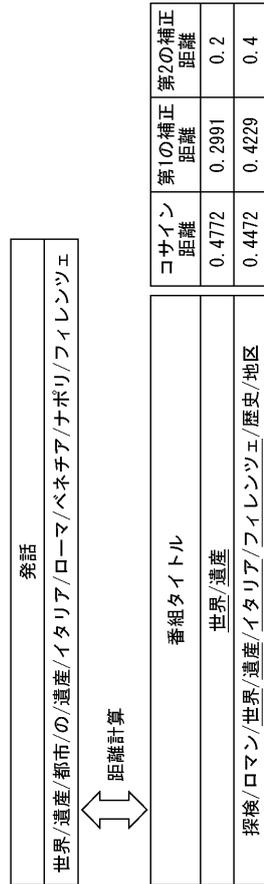
【 図 7 】

図7



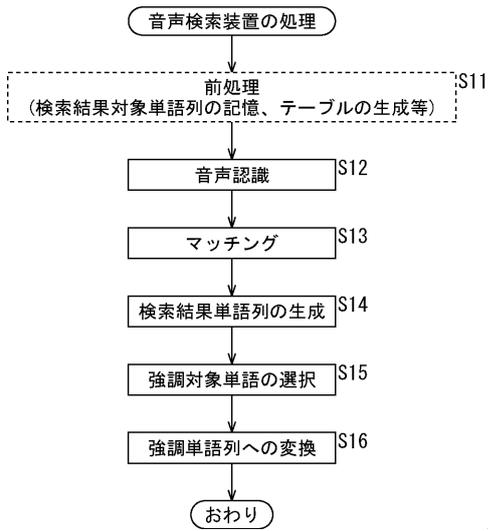
【 図 8 】

図8



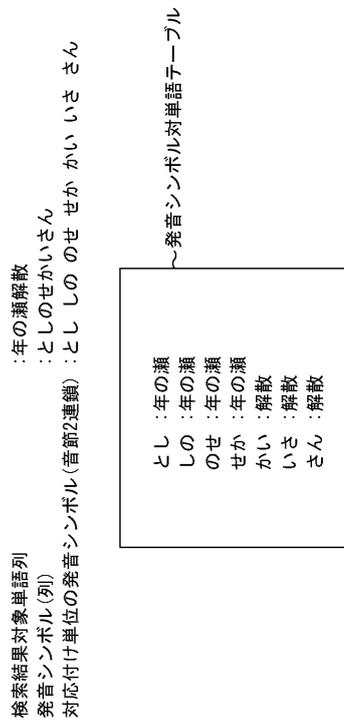
【 図 9 】

図9



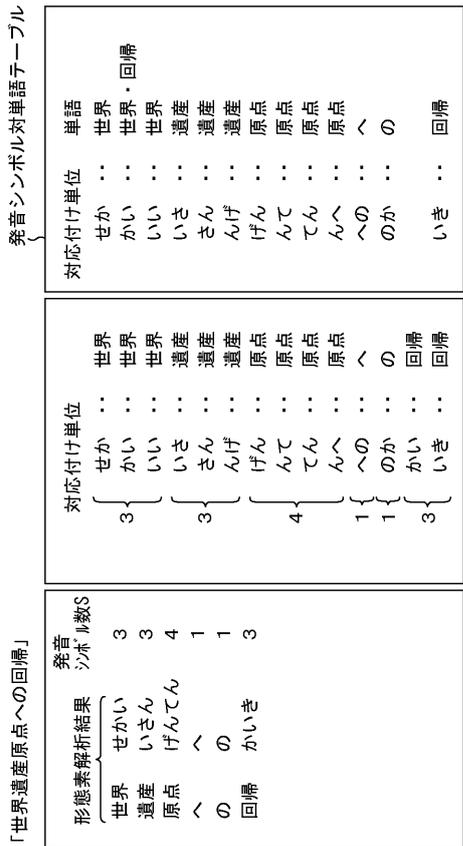
【 図 10 】

図10



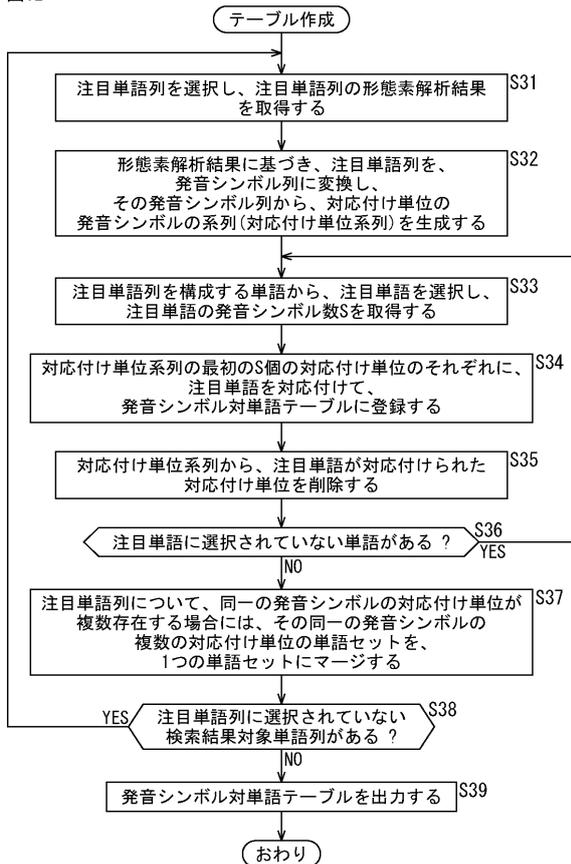
【図 1 1】

図11



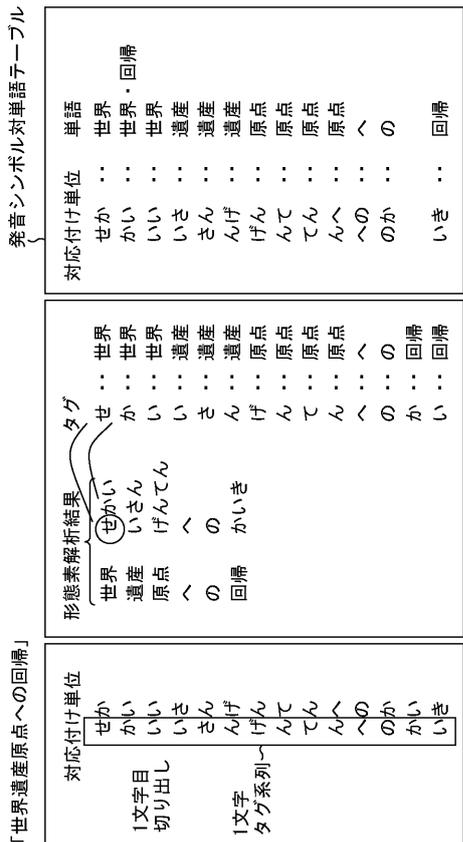
【図 1 2】

図12



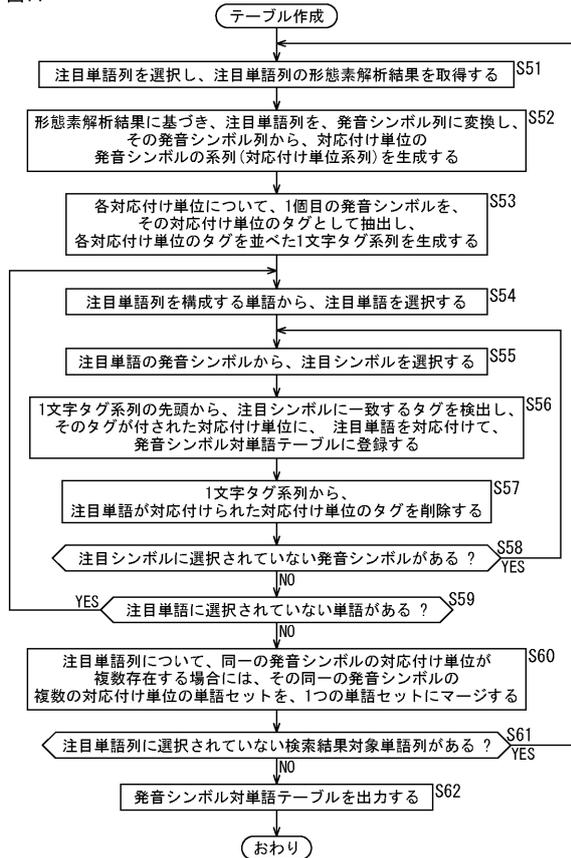
【図 1 3】

図13



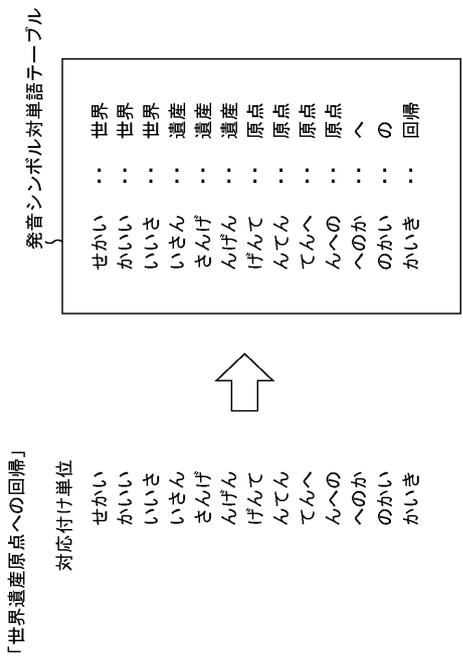
【図 1 4】

図14



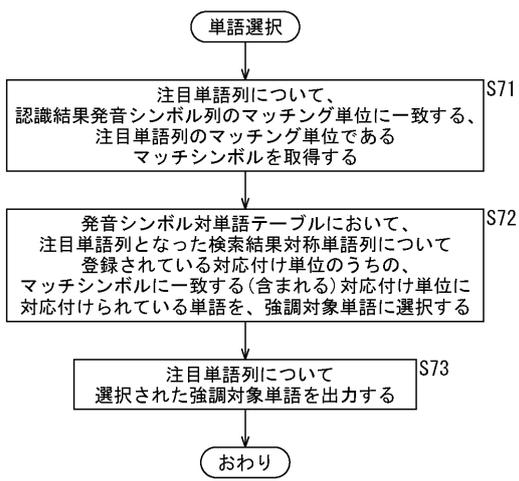
【 図 1 5 】

図15



【 図 1 6 】

図16



【 図 1 7 】

図17

発音シンボル対単語テーブル

- ・「The世界遺産」についての対応付け単位(音声2連鎖)と単語
 ざせ:The
 せか:世界
 かい:世界
 いい:遺産
 いさ:遺産
 さん:遺産
- ・「世界遺産 原点への回帰」についての対応付け単位と単語
 せか:世界
 かい:世界、回帰
 いさ:遺産
 さん:遺産
 んげ:遺産
 げん:原点
 んて:原点
 てん:原点
 んへ:原点
 への:へ
 のか:の
 いき:回帰
- ・「速報 年の瀬に解散」についての対応付け単位と単語
 そく:速報
 くほ:速報
 ほう:速報
 うと:速報
 とし:年の瀬
 しの:年の瀬
 のせ:年の瀬
 せに:年の瀬
 にか:に
 かい:解散
 いさ:解散
 さん:解散

【 図 1 8 】

図18

笑っていいとも	笑っていいとも増刊号	笑っていいともクリスマススペシャル	笑っていいとも二千八番の祭典SP	笑っていいとも二千八秋の祭典SP	土曜ドラマ赤い糸02	土曜ドラマ赤い糸03	映画連続ドラマ赤い糸の世界	カンゴロンゴ我が愛いとしの中古マンション	銀魂会わないとわからないことともある	中学生日記この愛いとしき世界へ	こころの時代宗教人生とともにもに祈りともにも彫る	獣医五十海七線の殺人カルテ	土曜ドラマ赤い糸01	はっけたいけんだいたいすきまじろう新しともだち	えいごでしゃべらないとJrかっこいいミュージシャンを目指せ	シリーズ世界遺産百代都市パレンケメケメキョー	シリーズ世界遺産百ハバナ旧市街と要塞キューパー	THE世界遺産セゴピア旧市街とローマ水道橋スベイン	アニメスポンジボブ終カーニバルーガール禁止令いとこのスタンスリー
---------	------------	-------------------	------------------	------------------	------------	------------	---------------	----------------------	--------------------	-----------------	--------------------------	---------------	------------	-------------------------	-------------------------------	------------------------	-------------------------	---------------------------	----------------------------------

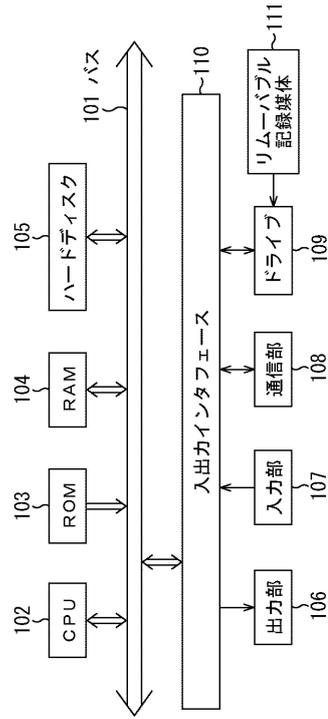
【 図 19 】

図19

笑っていいとも
 笑っていいとも増刊号
 笑っていいともクリスマススペシャル
 笑っていいとも二十八年の祭典SP
 笑っていいとも二十八年の祭典SP
 土曜ドラマ赤い糸02
 土曜ドラマ赤い糸03
 映画連続ドラマ赤い糸の世界
 カンゴロンゴ我が愛いとしの中古マンション
 鶏珍姿わないとわからないこともある
 中学生日記この愛いとしの世界へ
 こころの時代宗義人生ととも折りとに恥る
 獣医五十海七緒の殺人カルテ
 土曜ドラマ赤い糸01
 はっけんたいけんだいきまじろ断しいともだち
 えいごでしゃべらないとかっこいいミュージシャンを目指せ
 シリーズ世界遺産百代都市パレンケンメキシコ
 シリーズ世界遺産百代都市パレンケンと要塞キューバ
 TH世界遺産セゴビア旧市街とローマ水道橋スペイン
 アニメスポンジボブ終カーニバーガー禁止令とこのスタンスリー

【 図 20 】

図20



コンピュータ