

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2014 年 4 月 3 日(03.04.2014)



(10) 国際公開番号
WO 2014/049998 A1

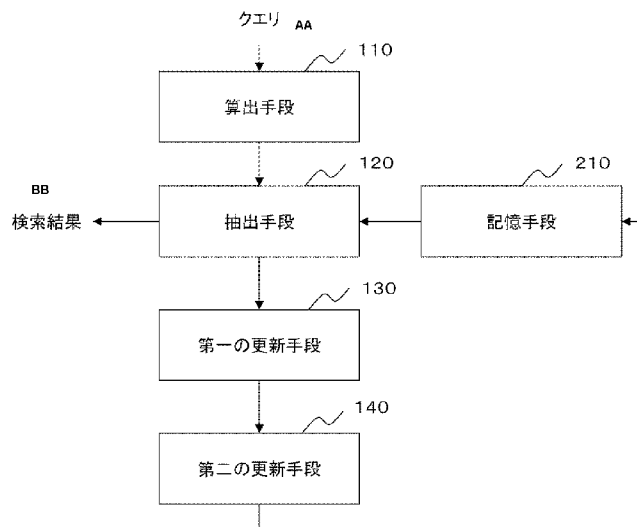
- (51) 国際特許分類:
G06F 17/30 (2006.01) *G10L 15/06* (2013.01)
G10L 15/00 (2013.01)
- (21) 国際出願番号: PCT/JP2013/005401
- (22) 国際出願日: 2013 年 9 月 12 日(12.09.2013)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2012-214952 2012 年 9 月 27 日(27.09.2012) JP
- (71) 出願人: 日本電気株式会社(NEC CORPORATION)
[JP/JP]; 〒1088001 東京都港区芝五丁目 7 番 1 号
Tokyo (JP).
- (72) 発明者: 大西 祥史(ONISHI, Yoshifumi); 〒1088001
東京都港区芝五丁目 7 番 1 号日本電気株式会
社内 Tokyo (JP).
- (74) 代理人: 下坂 直樹(SHIMOSAKA, Naoki); 〒
1088001 東京都港区芝五丁目 7 番 1 号日本電気
株式会社内 Tokyo (JP).

- (81) 指定国 (表示のない限り、全ての種類の国内保
護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA,
BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN,
CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES,
FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN,
IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS,
LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,
MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH,
PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK,
SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保
護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW,
MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシ
ア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ
(AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR,
GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT,
NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML,
MR, NE, SN, TD, TG).

[続葉有]

(54) Title: INFORMATION SEARCH SYSTEM, INFORMATION SEARCH METHOD, AND PROGRAM

(54) 発明の名称: 情報検索システム、情報検索方法およびプログラム



110 Calculation means
120 Extraction means
130 First updating means
140 Second updating means
210 Storage means
AA Query
BB Search result

(57) Abstract: The purpose of the invention is to provide an information search system, an information search method, and a program in which data pertaining to speech can be searched in an accurate manner, even when a word that does not readily appear as a recognition result is the query. The present invention is an information search system having: a calculation means for calculating a query language model, which is a language model for the inputted word or word group; an extraction means for referencing a storage means for storing the results of performing speech recognition on speech data using a speech recognition language model, and extracting, from the results, a result having a high degree of similarity with the query language model as match data; a first updating means for updating the speech recognition language model using the match data; and a second updating means for updating the results stored in the storage means using the updated speech recognition language model. The extraction means extracts, from the updated results, a result having a high degree of similarity with the query language model, and outputs a search result which is data corresponding to the extracted result.

(57) 要約:

[続葉有]

WO 2014/049998 A1



添付公開書類:

— 国際調査報告（条約第 21 条(3)）

認識結果としてあらわれにくい単語がクエリとなった場合にも、音声に関するデータを精度よく検索することが出来る、情報検索システム、情報検索方法、およびプログラムを提供すること。本発明は、入力された単語または単語集合の言語モデルであるクエリ用言語モデルを算出する算出手段と、音声認識用言語モデルを用いて音声データを音声認識した結果を記憶する記憶手段を参照して、前記結果から、前記クエリ用言語モデルとの類似度が高い結果を、適合データとして抽出する抽出手段と、前記適合データを用いて前記音声認識用言語モデルを更新する第一の更新手段と、更新された音声認識用言語モデルを用いて、前記記憶手段に記憶された結果を更新する第二の更新手段とを有し、前記抽出手段は、更新された結果から前記クエリ用言語モデルとの類似度が高い結果を抽出し、抽出された結果に対応するデータである検索結果を出力する情報検索システムである。

明 細 書

発明の名称： 情報検索システム、情報検索方法およびプログラム 技術分野

[0001] 本発明は、情報検索システム、情報検索方法およびプログラムに関し、特に、音声に関するデータを検索する、情報検索システム、情報検索方法およびプログラムに関する。

背景技術

[0002] 音声に関するデータを検索する技術の一例が、特許文献 1 に記載されている。特許文献 1 に記載された検索装置は、入力されたクエリおよび音声認識結果のテキスト間の類似度を、音声認識の信頼度を用いて算出し、類似度が高い音声認識結果を、検索結果として出力する。一般に、音声認識結果には誤認識が含まれる。この検索装置は、音声認識結果に対する信頼度を用いて、低い信頼度となる音声認識結果を検索結果から除外することにより、誤った認識結果を検索結果として出力する確率を減少させている。

先行技術文献

特許文献

[0003] 特開 2 0 1 1 - 2 4 8 1 0 7 号公報

発明の概要

発明が解決しようとする課題

[0004] 特許文献 1 に記載された技術では、音声認識結果としてあらわれにくい単語がクエリとなったときに、音声に関するデータを、精度よく検索できないという課題がある。

[0005] 例えば、音声認識において N - g r a m などの言語モデルを用いた場合、言語モデルの学習において出現頻度が低い単語は、音声認識結果にも出現しにくい。また、このような単語は、音声認識結果に出現しても言語モデルの確率値が低いため、低い信頼度となることがある。このため、このような単語に関するクエリが入力された場合、音声に関するデータを、精度よく検索

できない。

[0006] [発明の目的]

本発明の目的は、上記課題を解決し、認識結果としてあらわれにくい単語がクエリとなった場合にも、音声に関するデータを精度よく検索することが出来る、情報検索システム、情報検索方法、およびプログラムを提供することである。

課題を解決するための手段

[0007] 本発明は、入力された単語または単語集合の言語モデルであるクエリ用言語モデルを算出する算出手段と、音声認識用言語モデルを用いて音声データを音声認識した結果を記憶する記憶手段を参照して、前記結果から、前記クエリ用言語モデルとの類似度が高い結果を、適合データとして抽出する抽出手段と、前記適合データを用いて前記音声認識用言語モデルを更新する第一の更新手段と、更新された音声認識用言語モデルを用いて、前記記憶手段に記憶された結果を更新する第二の更新手段とを有し、前記抽出手段は、更新された結果から前記クエリ用言語モデルとの類似度が高い結果を抽出し、抽出された結果に対応するデータである検索結果を出力する情報検索システムである。

[0008] 本発明は、入力された単語または単語集合の言語モデルであるクエリ言語モデルを算出し、音声認識用言語モデルを用いて音声データを音声認識した結果を記憶した記憶手段を参照して、前記結果から、前記クエリ用言語モデルとの類似度が高い結果を、適合データとして抽出し、前記適合データを用いて前記音声認識用言語モデルを更新し、更新された音声認識用言語モデルを用いて、前記記憶手段に記憶された結果を更新し、更新された結果から、前記クエリ言語モデルとの類似度が高い結果を抽出し、抽出された結果に対応するデータである検索結果を出力する情報検索方法である。

[0009] 本発明は、入力された単語または単語集合の言語モデルであるクエリ言語モデルを算出するステップと、音声認識用言語モデルを用いて音声データを音声認識した結果を記憶した記憶手段を参照して、前記結果から、前記クエ

り用言語モデルとの類似度が高い結果を、適合データとして抽出するステップと、前記適合データを用いて前記音声認識用言語モデルを更新するステップと、更新された音声認識用言語モデルを用いて、前記記憶手段に記憶された結果を更新するステップと、更新された結果から、前記クエリ言語モデルとの類似度が高い結果を抽出し、抽出された結果に対応するデータである検索結果を出力するステップと、をコンピュータに実行させるプログラムである。

発明の効果

[0010] 本発明によれば、音声認識結果としてあらわれにくい単語がクエリとなった場合にも、音声に関するデータを精度よく検索することができる。

図面の簡単な説明

[0011] [図1]本発明の第一の実施の形態に係るハードウェア構成図である。

[図2]本発明の第一の実施の形態に係るブロック図である。

[図3]本発明の第一の実施の形態に係るフローチャートである。

[図4]本発明の第二の実施の形態に係るブロック図である。

[図5]本発明の第二の実施の形態に係るフローチャートである。

[図6]本発明の第三の実施の形態に係るブロック図である。

[図7]本発明の第三の実施の形態に係るフローチャートである。

[図8]本発明の第四の実施の形態に係るブロック図である。

[図9]本発明の第四の実施の形態に係るフローチャートである。

[図10]本発明の実施例に係るブロック図である。

[図11]本発明の実施例に係るフローチャートである。

[図12]本発明の情報検索システムの構成を示すブロック図である。

発明を実施するための形態

[0012] 次に、本発明の実施の形態について図面を参照して詳細に説明する。

[0013] 実施の形態 1.

図 1 は、本発明の第 1 の実施形態に係る情報検索システム 1 のハードウェア構成図である。図 1 に示すように、情報検索システム 1 は、CPU 10、

メモリ 12、ハードディスクドライブ（HDD）14、図示しないネットワークを介してデータの通信を行う通信インタフェース（IF）16、ディスプレイ等の表示装置18およびキーボードやマウス等のポインティングデバイスを含む入力装置20を有する。これらの構成要素は、バス22を通して互いに接続されており、互いにデータの入出力を行う。なお、情報検索システム1のハードウェア構成は、この構成に制限されず、適宜変更することができる。

[0014] 図2は、本発明の第1の実施の形態による情報検索システムの構成を示すブロック図である。

[0015] 図2に示すように、第1の実施の形態による情報検索システムは、算出手段110と、抽出手段120と、第一の更新手段130と、第二の更新手段140と、記憶手段210とを有する。

[0016] 記憶手段210は、音声認識用言語モデルを利用して音声データを音声認識した結果（以下、音声認識結果）を、記憶している。音声認識用言語モデルは、音声信号を単語列として認識する際に、認識されるべき単語列の制約を定義したモデルである。記憶手段210は、1つの音声データのファイルを音声認識した結果を、1つのテキストファイルの形式で記憶している。記憶手段210は、少なくとも1つ以上の音声認識結果（テキストファイル）を記憶している。

[0017] 算出手段110は、入力されたクエリに基づいて、クエリ用言語モデルを算出する。ここで、クエリとは、検索したい単語あるいは単語の集合である。

[0018] 次に、クエリ用言語モデルの算出方法の一例を示す。算出手段110は、数1によってクエリ用言語モデルを算出する。数1において、クエリ用言語モデルは、クエリの単語集合に対するユニグラム確率値 $p(w | \theta_0)$ である。ここで、 Q はクエリの単語集合、 $|Q|$ は Q の単語数、 w は単語、 θ_0 はクエリ用言語モデルのパラメータである。また、 $n(w, Q)$ は、 w が Q に含まれる単語の場合は Q に含まれる w の個数であり、 w が Q に含まれない場合

は0となる関数である。

[0019] [数1]

$$p(w|\theta_Q) = \frac{n(w, Q)}{|Q|}$$

[0020] 抽出手段120は、算出手段110で算出されたクエリ用言語モデルと、記憶手段210に記憶された各音声認識結果（各テキストファイル）との類似度を算出し、類似度が高い音声認識結果（テキストファイル）を、適合データとして抽出する。

[0021] 次に、抽出手段120の抽出方法の一例を示す。抽出手段120は、数2により、クエリ用言語モデルと音声認識結果の言語モデルのKL（カルバック・ライブラー）距離を、類似度として算出する。KL距離は、2つの、確率分布である言語モデルの差異を表す指標である。2つの言語モデルの類似度は、KL距離の値が小さいほど高い。KL（ $\theta_Q || \theta_D$ ）はKL距離、 $p(w | \theta_D)$ は、記憶手段210に記憶された、個々の音声認識結果Dの言語モデルである。

[0022] [数2]

$$KL(\theta_Q || \theta_D) = \sum_{w \in Q} p(w|\theta_Q) \ln \frac{p(w|\theta_Q)}{p(w|\theta_D)}$$

[0023] ここで、抽出手段120は、数3により、音声認識結果の言語モデル $p(w | \theta_D)$ を算出する。 $p(w | \theta_D)$ は、記憶手段210に記憶された音声認識結果の、全体集合Cの言語モデルである。ここで、 $|D|$ は音声認識結果Dの単語数であり、 μ は音声認識結果Dのユニグラム確率値と $p(w | \theta_D)$ のスムージングパラメタである。例えば μ は事前に与えられている。さらに、抽出手段120は、記憶手段210に記憶された音声認識結果全体を用いて、例えばNを3あるいは4としたN-gram確率として、 $p(w | \theta_D)$ を算出する。

[0024] [数3]

$$p(w|\theta_D) = \frac{1}{|D| + \mu} n(w, D) + \frac{\mu}{|D| + \mu} p(w|\theta_C)$$

[0025] 次に、抽出手段１２０は、例えば、算出したＫＬ距離があらかじめ定めた閾値より小さい、または閾値以下である、音声認識結果を抽出する。あるいは、抽出手段１２０は、ＫＬ距離が近い順にあらかじめ定めた個数の、音声認識結果を抽出してもよい。

[0026] 第一の更新手段１３０は、抽出手段１２０で抽出した、クエリ用言語モデルとの類似度が高い音声認識結果である適合データを利用して、音声認識用言語モデルを更新する。

[0027] 第一の更新手段１３０は、例えば、数５により音声認識用言語モデルを更新する。ここで、 $p(w|\theta_{ASR})$ は、更新される前の音声認識用言語モデルであり、 $p(w|\theta'_{ASR})$ は、更新された後の音声認識用言語モデルである。また、 $p(w|\theta_{CF})$ は、適合データの集合ＣＦの言語モデルである。 β は、更新するためのパラメータであり、例えば事前に与えられる。

[0028] [数５]

$$p(w|\theta'_{ASR}) = (1 - \beta) p(w|\theta_{CF}) + \beta p(w|\theta_{ASR})$$

[0029] 第二の更新手段１４０は、第一の更新手段１３０で更新された音声認識用言語モデルを用いて、記憶手段２１０に記憶された音声認識結果を更新する。例えば、第二の更新手段１４０は、更新された音声認識用言語モデルを用いて、音声認識結果の元データである音声データを、再度音声認識することで、記憶手段２１０に記憶された音声認識結果を更新する。

[0030] あるいは、第二の更新手段１４０は、以下の方法で結果を更新してもよい。記憶手段２１０は、更新される前の音声認識用言語モデルを利用して音声データを音声認識した結果とともに、音声認識した結果に対応する単語グラフを記憶している。または、単語グラフは、記憶手段２１０と異なる記憶手段に記憶されていてもよい。第二の更新手段１４０は、更新された音声認識用言語モデルを用いて、前記単語グラフに対する言語確率をリスコアするこ

とで、記憶手段 210 に記憶された音声認識結果を更新する。

[0031] 抽出手段 120 は、算出手段 110 で算出されたクエリ用言語モデルと、記憶手段 210 に記憶された更新された音声認識結果との類似度を算出し、類似度が高い音声認識結果を、適合データとして抽出する。

[0032] さらに、抽出手段 120 は、検索結果を出力する条件を満たす場合、抽出された音声認識結果に対応するデータの少なくとも一部を、検索結果として出力する。検索結果を出力する条件は、例えば、音声認識用言語モデルの更新、記憶手段 210 に記憶された結果の更新および適合データの抽出が、所定の回数行われたことである。また、検索結果を出力する条件は、更新された音声認識結果から抽出された音声認識結果が、更新される前の音声認識結果から抽出された音声認識結果と同じであることであってもよい。つまり、抽出される音声認識結果が変化しなくなったことである。ここで、音声認識結果に対応するデータは、音声認識結果自体であってもよい。また、音声認識結果に対応するデータは、音声認識結果の元データである音声データであってもよい。

[0033] なお、算出手段 110、抽出手段 120、第一の更新手段 130、第二の更新手段 140 の動作は、上述した例に限定されるわけではなく、適宜変更することができる。

[0034] 次に、本発明を実施するための第 1 の実施の形態の動作について詳細に説明する。

[0035] 図 3 は、第 1 の実施の形態の動作の一例を示すフローチャートである。

[0036] ステップ 101 では、算出手段 110 が、入力されたクエリに基づいて、クエリ用言語モデルを算出する。ステップ 102 では、抽出手段 120 が、算出手段 110 で算出されたクエリ用言語モデルと、記憶手段 210 に記憶された音声認識結果との類似度を算出し、類似度が高い音声認識結果を、適合データとして抽出する。ステップ 103 では、第一の更新手段 130 が、抽出手段 120 で抽出した適合データを利用して、音声認識用言語モデルを更新する。ステップ 104 では、第二の更新手段 140 が、更新された音声

認識用言語モデルを用いて、記憶手段 210 に記憶された音声認識結果を更新する。ステップ 105 では、抽出手段 120 が、算出手段 110 で算出されたクエリ用言語モデルと、記憶手段 210 に記憶された更新された音声認識結果との類似度を算出し、類似度が高い音声認識結果を、適合データとして抽出する。検索結果を出力する条件を満たさない場合、ステップ 103 に戻る。検索結果を出力する条件を満たす場合、ステップ 106 で、抽出手段 120 は、抽出された音声認識結果に対応する検索結果の少なくとも一部を出力する。

[0037] 本実施の形態によれば、クエリとして入力された単語集合と類似度が高い音声認識結果により、音声認識用言語モデルが更新される。さらに、記憶手段 210 に記憶された音声認識結果が、更新された音声認識用言語モデルによって、更新される。よって、本実施の形態による情報検索システムは、クエリに含まれた単語に対し、音声認識用言語モデルにおける確率値および音声認識結果における信頼度を適切に与えることができる。これにより、認識結果としてあらわれにくい単語がクエリとなった場合に、音声に関するデータを精度よく検索することができる。

[0038] 実施の形態 2.

図 4 は、本発明の第 2 の実施の形態による情報検索システムの構成を示すブロック図である。

[0039] 第 2 の実施の形態による情報検索システムは、第 1 の実施形態の構成要素に加えて、選別手段 150 を有する。また、本実施の形態による情報検索システムは、第 1 の実施形態の第一の更新手段 130 の代わりに、第一の更新手段 131 を有する。選別手段 150 と第一の更新手段 131 以外の構成要素については、第 1 の実施形態と同様であるため、説明を省略する。

[0040] 選別手段 150 は、適合データ間の類似度に基づいて、適合データを選別する。具体的には、選別手段 150 は、他の適合データとの類似度が低い適合データを、適合データから除外する。

[0041] 選別手段 150 は、例えば、以下のように適合データを選別する。選別手

段 150 は、適合データの集合 CF の言語モデル $p(w | \theta_{CF})$ を算出する。 $p(w | \theta_{CF})$ は、N-gram 確率値である。N は、例えば、1 または 2 である。次に、選別手段 150 は、適合データの集合 CF に含まれる適合データ F の言語モデル $p(w | \theta_F)$ を、数 6 により算出する。 $|F|$ は適合データ F の単語数であり、 σ は $p(w | \theta_{CF})$ と適合データ F のユニグラム確率値のスムージングパラメタである。 σ は事前に与えられてもよい。

[0042] [数 6]

$$p(w|\theta_F) = \frac{1}{|F| + \sigma} n(w, F) + \frac{\sigma}{|F| + \sigma} p(w|\theta_{CF})$$

[0043] 選別手段 150 は、適合データの集合 CF と適合データ F との KL 距離である $KL(\theta_{CF} || \theta_F)$ を算出し、この値が所定の値より大きい文書を除外する。KL 距離の算出方法は、数 2 と同様であるため、説明を省略する。

[0044] あるいは、選別手段 150 は、以下のように適合データを選別してもよい。選別手段 150 は、適合データの集合 CF に含まれる適合データ F1、F2 について、それぞれの言語モデルを数 6 により算出する。F1 の言語モデルは、 $P(w | \theta_{F1})$ 、F2 の言語モデルは、 $P(w | \theta_{F2})$ とする。次に、選別手段 150 は、F1 および F2 の類似度である $SKL(\theta_{F1}, \theta_{F2})$ を、数 7 により算出する。

[0045] [数 7]

$$SKL(\theta_{F1}, \theta_{F2}) = \frac{KL(\theta_{F1} || \theta_{F2}) + KL(\theta_{F2} || \theta_{F1})}{2}$$

[0046] さらに、選別手段 150 は、 $SKL(\theta_{F1}, \theta_{F2})$ に基づき、ボトムアップクラスタリングを行う。ボトムアップクラスタリングとは、距離が近い 2 つのデータを、順次階層的に指定のクラスタ数になるまでまとめあげる手法である。選別手段 150 は、主要クラスタ以外のクラスタに含まれるデータを、適合データから除外する。ここで、主要クラスタは、例えば、そのクラスタに所属する適合データの数が最大であるクラスタである。または、主要ク

ラストは、そのクラスタに所属する適合データの数が多い順に、指定した数までのクラスタでもよい。

[0047] 第一の更新手段131は、選別手段150で選別された適合データを利用して、音声認識用言語モデルを更新する。モデルを更新する方法については、第一の更新手段130と同様であるため、説明を省略する。

[0048] 図5は、第2の実施の形態の動作の一例を示すフローチャートである。ステップ101、102は、第1の実施の形態の動作と同様であるため、説明を省略する。ステップ107では、選別手段150が、適合データを選別する。ステップ113では、第一の更新手段131が、選別された適合データを用いて、音声認識結果を更新する。ステップ104～106は、第1の実施の形態の動作と同様であるため、説明を省略する。

[0049] 本実施の形態による情報検索システムは、他の適合データとの類似度が低い適合データを、適合データから除外する。よって、この情報検索システムは、適合データに誤って含まれる不適切な適合データを、適合データ間の、クエリの単語集合に含まれない単語も考慮した類似度に基づいて、除外することができる。よって、この情報検索システムは、音声認識誤りに対して、より頑健となる。

[0050] 実施の形態3.

図6は、本発明の第3の実施の形態による情報検索システムの構成を示すブロック図である。

[0051] 第3の実施の形態による情報検索システムは、第1の実施形態の構成要素に加えて、第三の更新手段160を有する。また、第1の実施形態の第一の更新手段130の代わりに、第一の更新手段132を有する。第三の更新手段160と第一の更新手段132以外の構成要素については、第1の実施形態と同様であるため、説明を省略する。

[0052] 第三の更新手段160は、抽出手段120が抽出した適合データを用いて、クエリ用言語モデルを更新する。例えば、第三の更新手段160は、数8により、クエリ用言語モデルを更新する。 $p(w | \theta_0)$ は、更新される前の

クエリ用言語モデルである。 $p(w | \theta'_q)$ は、更新された後のクエリ用言語モデルである。

[0053] [数 8]

$$p(w | \theta'_q) = (1 - \alpha) p(w | \theta_q) + \alpha p(w | \theta_{CF})$$

[0054] ここで、 $p(w | \theta_{CF})$ は、適合データの集合 C F の言語モデルであり、 α は $p(w | \theta_q)$ と $p(w | \theta_{CF})$ のスムージングパラメタである。 α は、事前に与えられてもよい。

[0055] 第一の更新手段 132 は、第三の更新手段 160 で更新されたクエリ用言語モデルを用いて、数 9 により、音声認識用言語モデルを更新する。なお、数 9 は、数 5 における $p(w | \theta_{CF})$ が、 $p(w | \theta'_q)$ に置き換わったものである。

[0056] [数 9]

$$p(w | \theta'_{ASR}) = (1 - \beta) p(w | \theta'_q) + \beta p(w | \theta_{ASR})$$

[0057] なお、クエリ用言語モデルを更新する手法は、非特許文献 1 にも記載されている。

[0058] [非特許文献 1] CC. Zhai, "Statistical Language Models for Information Retrieval A Critical Review Found", Trends Inf. Retr., Now Publishers Inc., 2008, 2, 137-213

非特許文献 1 に記載されている技術は、テキスト文書に対する検索技術の一例である。本発明における情報検索システムは、音声に関するデータを検索の対象としている。本発明における情報検索システムは、更新したクエリ用言語モデルによって、音声認識用言語モデルおよび音声認識結果を更新している。つまり、本発明における情報検索システムは、音声認識結果が音声認識に用いる言語モデルに依存して変化する性質を、利用している。

[0059] 図 7 は、第 3 の実施の形態の動作の一例を示すフローチャートである。ステップ 101、102 は、第 1 の実施の形態の動作と同様であるため、説明

を省略する。ステップ１０８では、第三の更新手段１６０が、抽出手段１２０が抽出した適合データを用いて、クエリ用言語モデルを更新する。ステップ１２３では、第一の更新手段１３２が、第三の更新手段１６０で更新されたクエリ用言語モデルを用いて、音声認識用言語モデルを更新する。ステップ１０４～１０６は、第１の実施の形態の動作と同様であるため、説明を省略する。

[0060] 本実施の形態による情報検索システムは、音声に関するデータを、精度よく検索することができる。クエリ用言語モデルが、適合データに基づいて更新される。さらに、音声認識用言語モデルも、更新されたクエリ用言語モデルによって更新される。よって、クエリ用言語モデルと音声認識用言語モデルが、整合性がとれた状態で更新されるためである。

[0061] 実施の形態４．

図８は、本発明の第４の実施の形態による情報検索システムの構成を示すブロック図である。本実施の形態は、第２の実施の形態の構成と第３の実施の形態の構成とを、組み合わせたものである。各構成要素は、第１の実施の形態～第３の実施の形態の構成要素と同様であるため、説明を省略する。

[0062] 図９は、第４の実施の形態の動作の一例を示すフローチャートである。ステップ１０１～１０８の動作は、第１の実施の形態～第３の実施の形態における対応するステップの動作と同様であるため、説明を省略する。

[0063] 本実施の形態によれば、音声に関するデータを、精度よく検索することができる。

[0064] 変形例．

図１０は、第４の実施の形態の変形例による情報検索システムの構成を示すブロック図である。

[0065] 本変形例による情報検索システムは、第４の実施の形態の構成要素に加えて、第二の記憶手段２２０、第三の記憶手段２３０、第四の記憶手段２４０を有する。

[0066] 第二の記憶手段２２０は、検索対象となる、音声データを記憶している。

- [0067] 第二の更新手段 140 は、音声認識を行う手段である。第二の更新手段 140 は、音声認識用言語モデル記憶手段 230 に記憶された音声認識用言語モデルを用いて、第二の記憶手段 220 に記憶された音声データの少なくとも一部を音声認識する。さらに、第二の更新手段 140 は、音声認識した結果を記憶手段（第一の記憶手段）210 に格納する。
- [0068] 第三の記憶手段 230 は、音声認識用言語モデルを記憶している。
- [0069] 第四の記憶手段 240 は、クエリ用言語モデルを記憶している。
- [0070] 算出手段 110 は、算出したクエリ用言語モデルを、第四の記憶手段 240 に格納する。また、第三の更新手段は、第四の記憶手段 240 に記憶されたクエリ用言語モデルを更新する。さらに、第一の更新手段は、第四の記憶手段 240 に記憶された、更新されたクエリ用言語モデルに基づいて、第三の記憶手段 230 に記憶された音声認識用言語モデルを更新する。
- [0071] その他の構成要素については、第 4 の実施の形態の構成要素と同様であるため、説明を省略する。
- [0072] 図 11 は、変形例の動作の一例を示すフローチャートである。ステップ 109 で、第二の更新手段 140 は、第三の記憶手段 230 に記憶された音声認識用言語モデルを用いて、第二の記憶手段 220 に記憶された音声データの少なくとも一部を音声認識する。さらに、ステップ 109 で、第二の更新手段 140 は、音声認識結果を第一の記憶手段 210 に格納する。ステップ 101～108 の動作は、第 1 の実施の形態～第 4 の実施の形態における対応するステップの動作と同様であるため、説明を省略する。なお、ステップ 101 は、ステップ 109 より前に動作してもよい。
- [0073] なお、上述の説明で用いた複数のフローチャートでは、複数の処理が順番に記載されているが、各実施形態で実行される処理の実行順序は、その記載の順番に制限されない。各実施形態では、図示される工程の順番を内容的に支障のない範囲で変更することができる。また、上述の各実施形態及び変形例は、内容が相反しない範囲で組み合わせることができる。
- [0074] 以上、実施形態を参照して本願発明を説明したが、本願発明は上記実施形

態に限定されるものではない。本願発明の構成や詳細には、本願発明のスコープ内で当業者が理解し得る様々な変更をすることができる。

[0075] (付記 1)

入力された単語または単語集合の言語モデルであるクエリ用言語モデルを算出する算出手段と、音声認識用言語モデルを用いて音声データを音声認識した結果を記憶する記憶手段を参照して、前記結果から、前記クエリ用言語モデルとの類似度が高い結果を、適合データとして抽出する抽出手段と、前記適合データを用いて前記音声認識用言語モデルを更新する第一の更新手段と、更新された音声認識用言語モデルを用いて、前記記憶手段に記憶された結果を更新する第二の更新手段とを有し、前記抽出手段は、更新された結果から前記クエリ用言語モデルとの類似度が高い結果を抽出し、抽出された結果に対応するデータである検索結果を出力する情報検索システム。

[0076] 図 1 2 は、本情報検索システムの構成を示すブロック図である。

[0077] (付記 2)

付記 1 に記載の情報検索システムにおいて、前記適合データの集合において、適合データ間の類似度に基づいて適合データを選別する選別手段を有し、前記第一の更新手段は、選別された適合データを用いて前記音声認識用言語モデルを更新する情報検索システム。

[0078] (付記 3)

付記 1 または 2 に記載の情報検索システムにおいて、前記適合データを用いて前記クエリ用言語モデルを更新する第三の更新手段を有し、前記第一の更新手段は、前記適合データの代わりに、更新されたクエリ用言語モデルを用いて、前記音声認識用言語モデルを更新する情報検索システム。

[0079] (付記 4)

付記 1 乃至 3 のいずれかに記載の情報検索システムにおいて、前記抽出手段は、更新された結果から抽出された結果が、更新される前の結果から抽出された結果と同じであった場合に、検索結果を出力する情報検索システム。

[0080] (付記 5)

付記 1 乃至 4 のいずれかに記載の情報検索システムにおいて、前記第二の更新手段は、更新された音声認識用言語モデルを用いて前記音声データを音声認識することで、前記結果を更新する情報検索システム。

[0081] (付記 6)

付記 1 乃至 4 のいずれかに記載の情報検索システムにおいて、前記第二の更新手段は、更新された音声認識用言語モデルを用いて、前記音声データを音声認識した結果に対応する単語グラフの言語確率をリスコアすることで、前記結果を更新する情報検索システム。

[0082] (付記 7)

入力された単語または単語集合の言語モデルであるクエリ言語モデルを算出し、音声認識用言語モデルを用いて音声データを音声認識した結果を記憶した記憶手段を参照して、前記結果から、前記クエリ用言語モデルとの類似度が高い結果を、適合データとして抽出し、前記適合データを用いて前記音声認識用言語モデルを更新し、更新された音声認識用言語モデルを用いて、前記記憶手段に記憶された結果を更新し、更新された結果から、前記クエリ言語モデルとの類似度が高い結果を抽出し、抽出された結果に対応するデータである検索結果を出力する、情報検索方法。

[0083] (付記 8)

付記 7 に記載の情報検索方法において、前記適合データの集合において、適合データ間の類似度に基づいて適合データを選別し、選別された適合データを用いて前記音声認識用言語モデルを更新する情報検索方法。

[0084] (付記 9)

入力された単語または単語集合の言語モデルであるクエリ言語モデルを算出するステップと、音声認識用言語モデルを用いて音声データを音声認識した結果を記憶した記憶手段を参照して、前記結果から、前記クエリ用言語モデルとの類似度が高い結果を、適合データとして抽出するステップと、前記適合データを用いて前記音声認識用言語モデルを更新するステップと、更新された音声認識用言語モデルを用いて、前記記憶手段に記憶された結果を更

新するステップと、更新された結果から、前記クエリ言語モデルとの類似度が高い結果を抽出し、抽出された結果に対応するデータである検索結果を出力するステップと、をコンピュータに実行させるプログラム。

[0085] (付記 10)

付記 9 に記載のプログラムにおいて、前記適合データの集合において、適合データ間の類似度に基づいて適合データを選別するステップと、選別された適合データを用いて前記音声認識用言語モデルを更新するステップと、をコンピュータに実行させるプログラム。

産業上の利用可能性

[0086] 本発明は、会話や発言が記録された音声データに対し、指定した単語あるいは単語集合に関連が大きい部分を検索できる音声検索システムなどの用途に利用することができる。

[0087] この出願は、2012年9月27日に出願された日本出願特願2012-214952を基礎とする優先権を主張し、その開示の全てをここに取り込む。

符号の説明

[0088] 1 情報検索システム

10 CPU

12 メモリ

14 HDD

16 通信IF

18 表示装置

20 入力装置

22 バス

110 算出手段

120 抽出手段

130、131、132 第一の更新手段

140 第二の更新手段

- 1 5 0 選別手段
- 1 6 0 第三の更新手段
- 2 1 0 記憶手段（第一の記憶手段）
- 2 2 0 第二の記憶手段
- 2 3 0 第三の記憶手段
- 2 4 0 第四の記憶手段

請求の範囲

- [請求項1] 入力された単語または単語集合の言語モデルであるクエリ用言語モデルを算出する算出手段と、
- 音声認識用言語モデルを用いて音声データを音声認識した結果を記憶する記憶手段を参照して、前記結果から、前記クエリ用言語モデルとの類似度が高い結果を、適合データとして抽出する抽出手段と、
- 前記適合データを用いて前記音声認識用言語モデルを更新する第一の更新手段と、
- 更新された音声認識用言語モデルを用いて、前記記憶手段に記憶された結果を更新する第二の更新手段と、を有し、
- 前記抽出手段は、更新された結果から前記クエリ用言語モデルとの類似度が高い結果を抽出し、抽出された結果に対応するデータである検索結果を出力する情報検索システム。
- [請求項2] 前記適合データの集合において、適合データ間の類似度に基づいて適合データを選別する選別手段を有し、
- 前記第一の更新手段は、選別された適合データを用いて前記音声認識用言語モデルを更新する、請求項1に記載の情報検索システム。
- [請求項3] 前記適合データを用いて前記クエリ用言語モデルを更新する第三の更新手段を有し、
- 前記第一の更新手段は、前記適合データの代わりに、更新されたクエリ用言語モデルを用いて、前記音声認識用言語モデルを更新する、請求項1または2に記載の情報検索システム。
- [請求項4] 前記抽出手段は、更新された結果から抽出された結果が、更新される前の結果から抽出された結果と同じであった場合に、検索結果を出力する、請求項1乃至3のいずれかに記載の情報検索システム。
- [請求項5] 前記第二の更新手段は、更新された音声認識用言語モデルを用いて前記音声データを音声認識することで、前記結果を更新する、請求項1乃至4のいずれかに記載の情報検索システム。

- [請求項6] 前記第二の更新手段は、更新された音声認識用言語モデルを用いて、前記音声データを音声認識した結果に対応する単語グラフの言語確率をリスクアすることで、前記結果を更新する、請求項1乃至4のいずれかに記載の情報検索システム。
- [請求項7] 入力された単語または単語集合の言語モデルであるクエリ言語モデルを算出し、
- 音声認識用言語モデルを用いて音声データを音声認識した結果を記憶した記憶手段を参照して、前記結果から、前記クエリ用言語モデルとの類似度が高い結果を、適合データとして抽出し、
- 前記適合データを用いて前記音声認識用言語モデルを更新し、
- 更新された音声認識用言語モデルを用いて、前記記憶手段に記憶された結果を更新し、
- 更新された結果から、前記クエリ言語モデルとの類似度が高い結果を抽出し、抽出された結果に対応するデータである検索結果を出力する情報検索方法。
- [請求項8] 前記適合データの集合において、適合データ間の類似度に基づいて適合データを選別し、
- 選別された適合データを用いて前記音声認識用言語モデルを更新する、請求項7に記載の情報検索方法。
- [請求項9] 入力された単語または単語集合の言語モデルであるクエリ言語モデルを算出するステップと、
- 音声認識用言語モデルを用いて音声データを音声認識した結果を記憶した記憶手段を参照して、前記結果から、前記クエリ用言語モデルとの類似度が高い結果を、適合データとして抽出するステップと、
- 前記適合データを用いて前記音声認識用言語モデルを更新するステップと、
- 更新された音声認識用言語モデルを用いて、前記記憶手段に記憶された結果を更新するステップと、

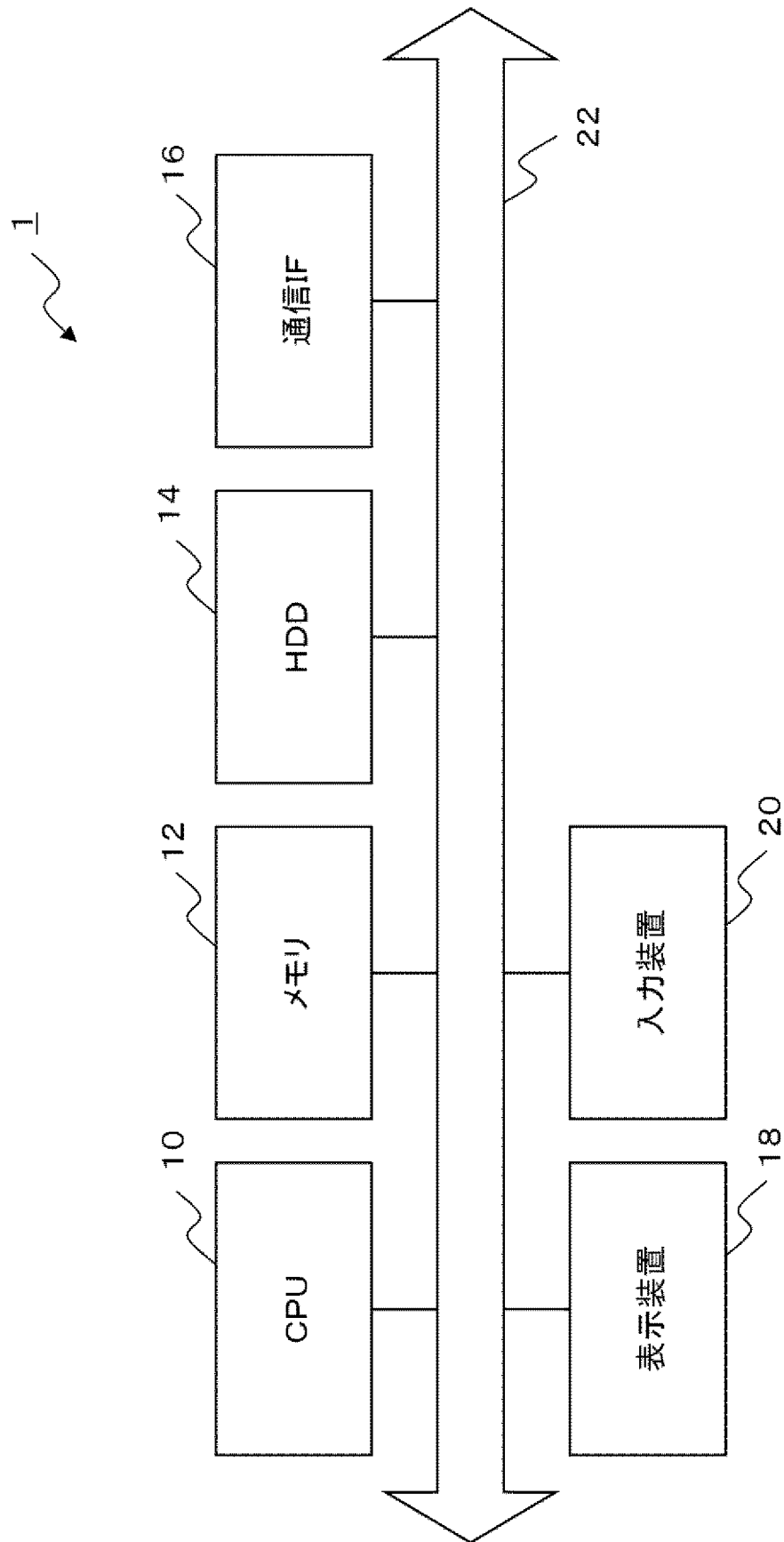
更新された結果から、前記クエリ言語モデルとの類似度が高い結果を抽出し、抽出された結果に対応するデータである検索結果を出力するステップと、をコンピュータに実行させるプログラム。

[請求項10]

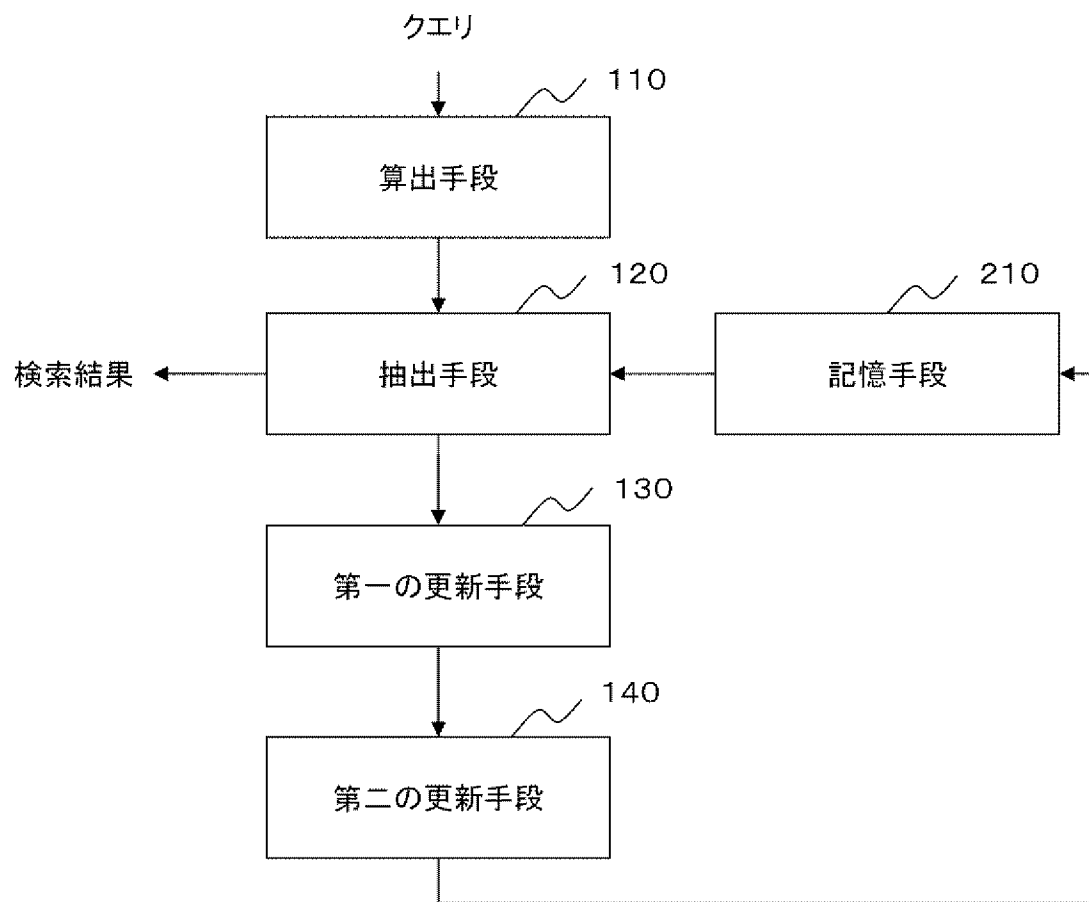
前記適合データの集合において、適合データ間の類似度に基づいて適合データを選別するステップと、

選別された適合データを用いて前記音声認識用言語モデルを更新するステップと、をコンピュータに実行させる請求項9に記載のプログラム。

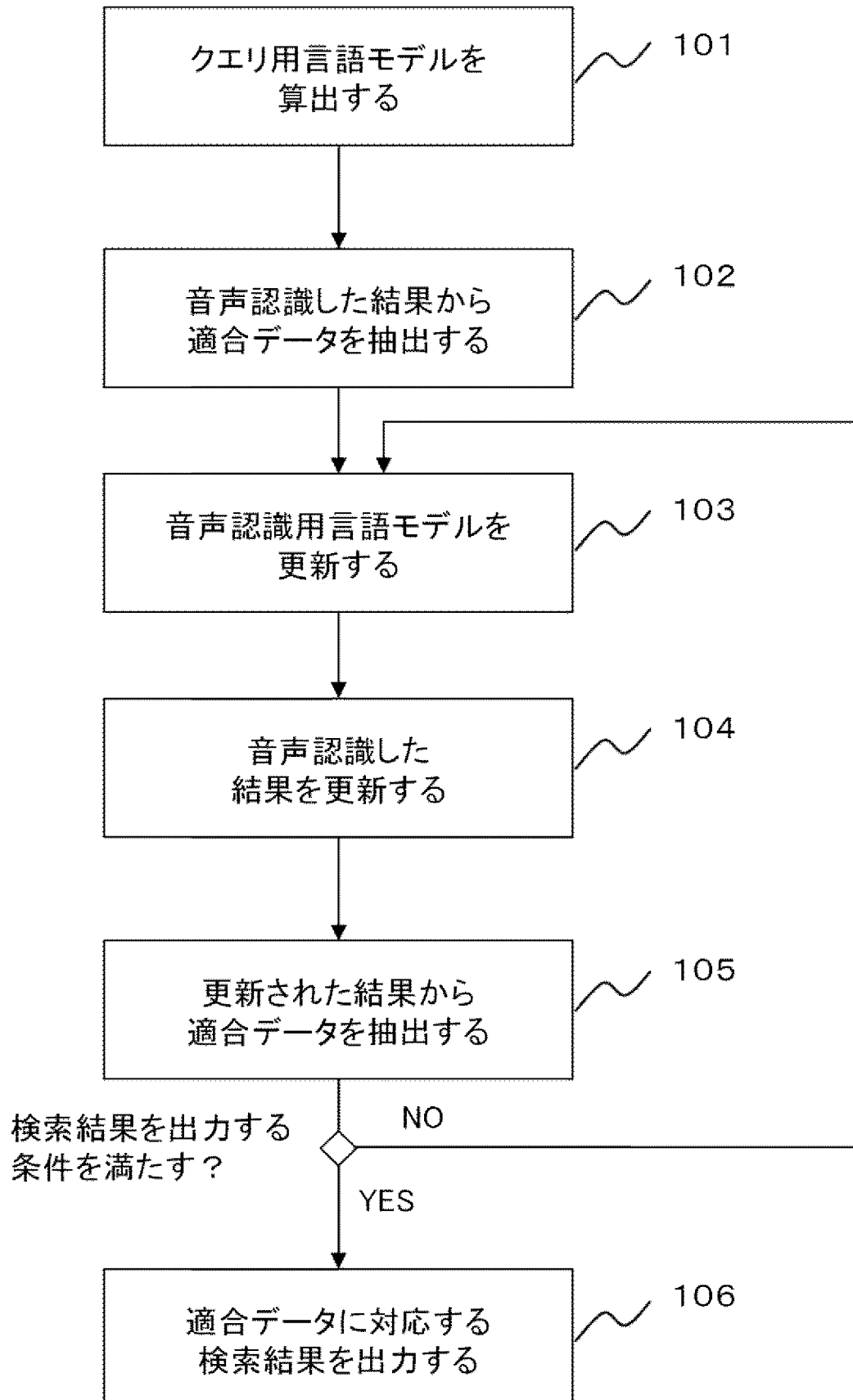
[図1]



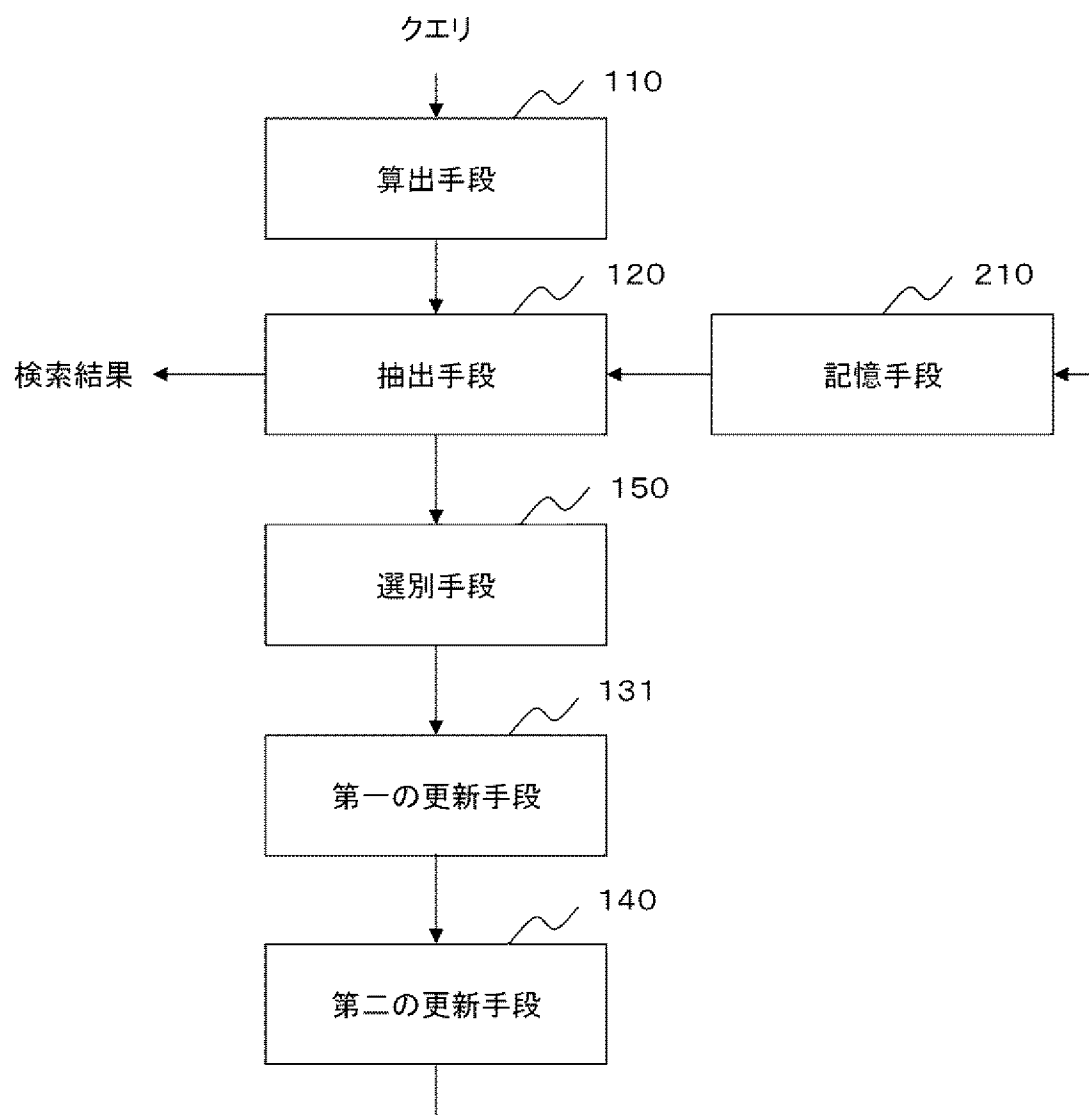
[図2]



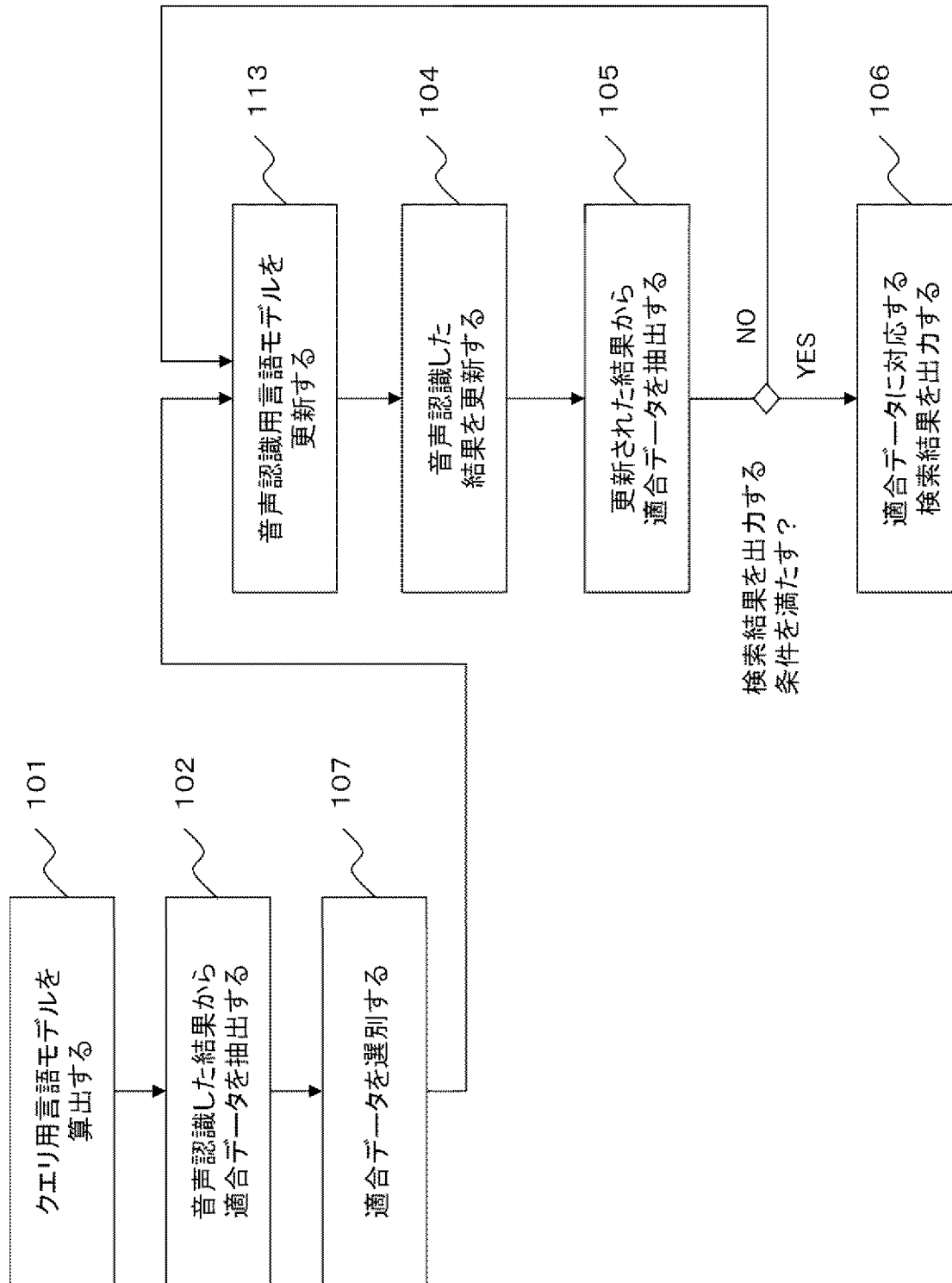
[図3]



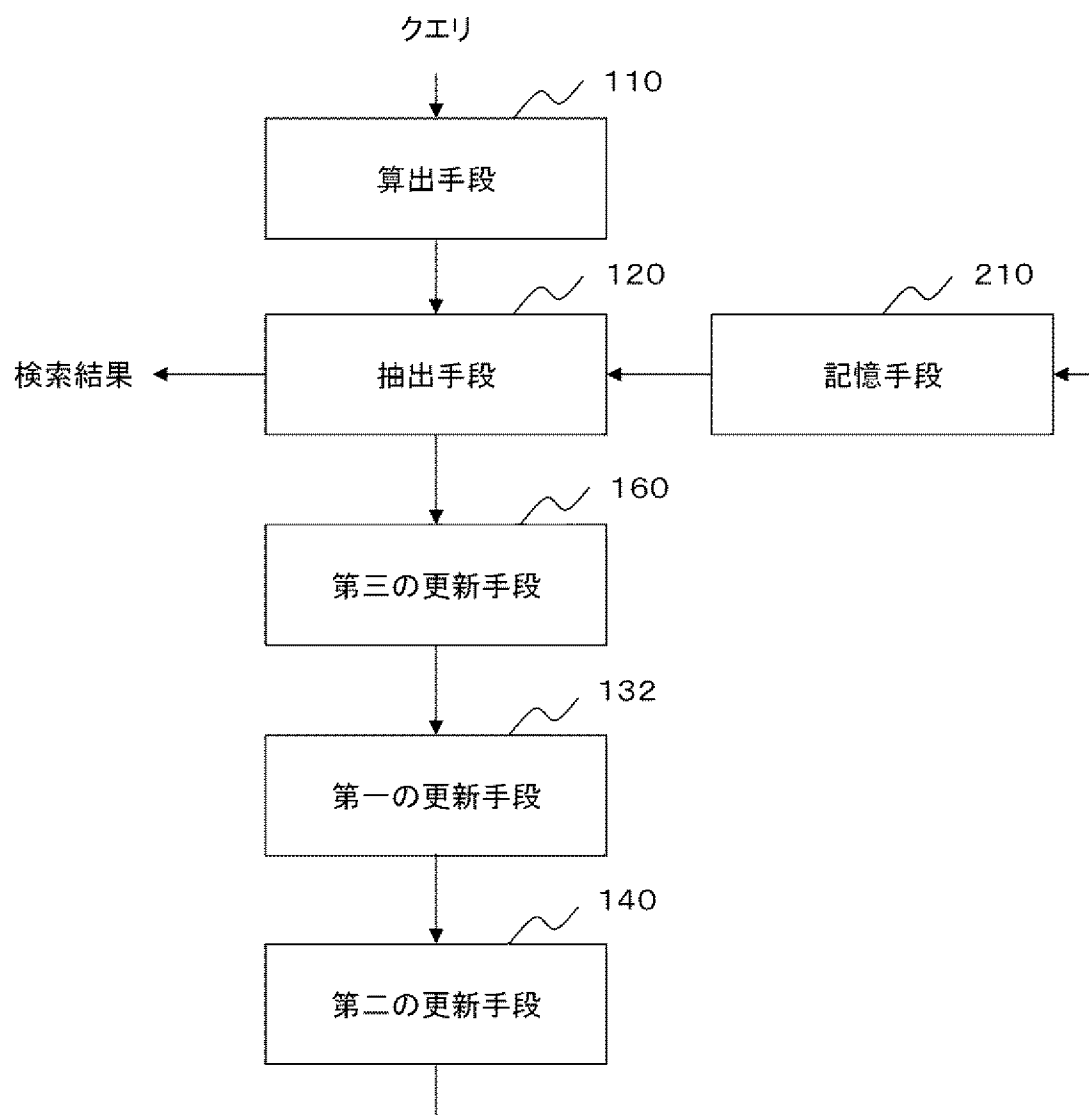
[図4]



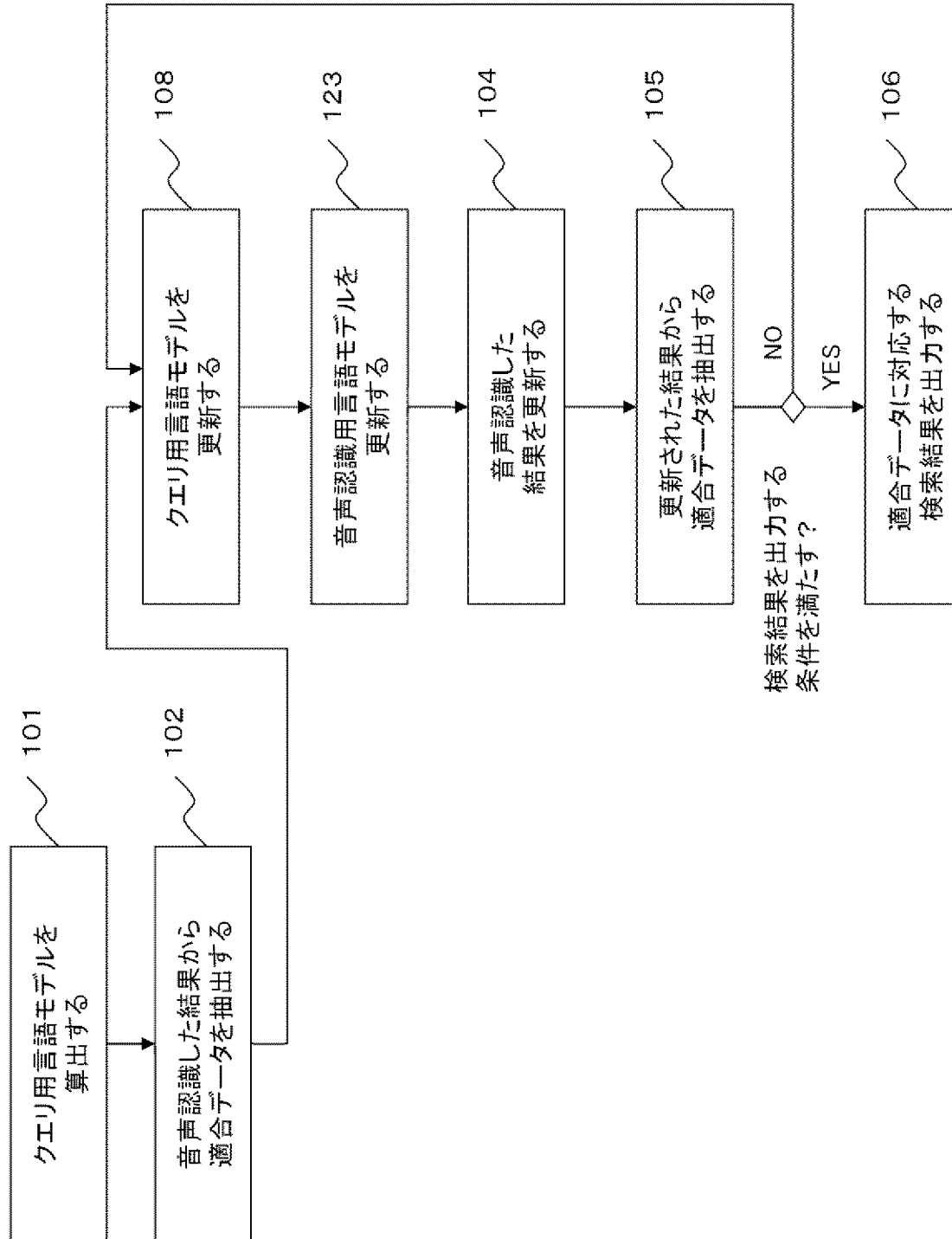
[図5]



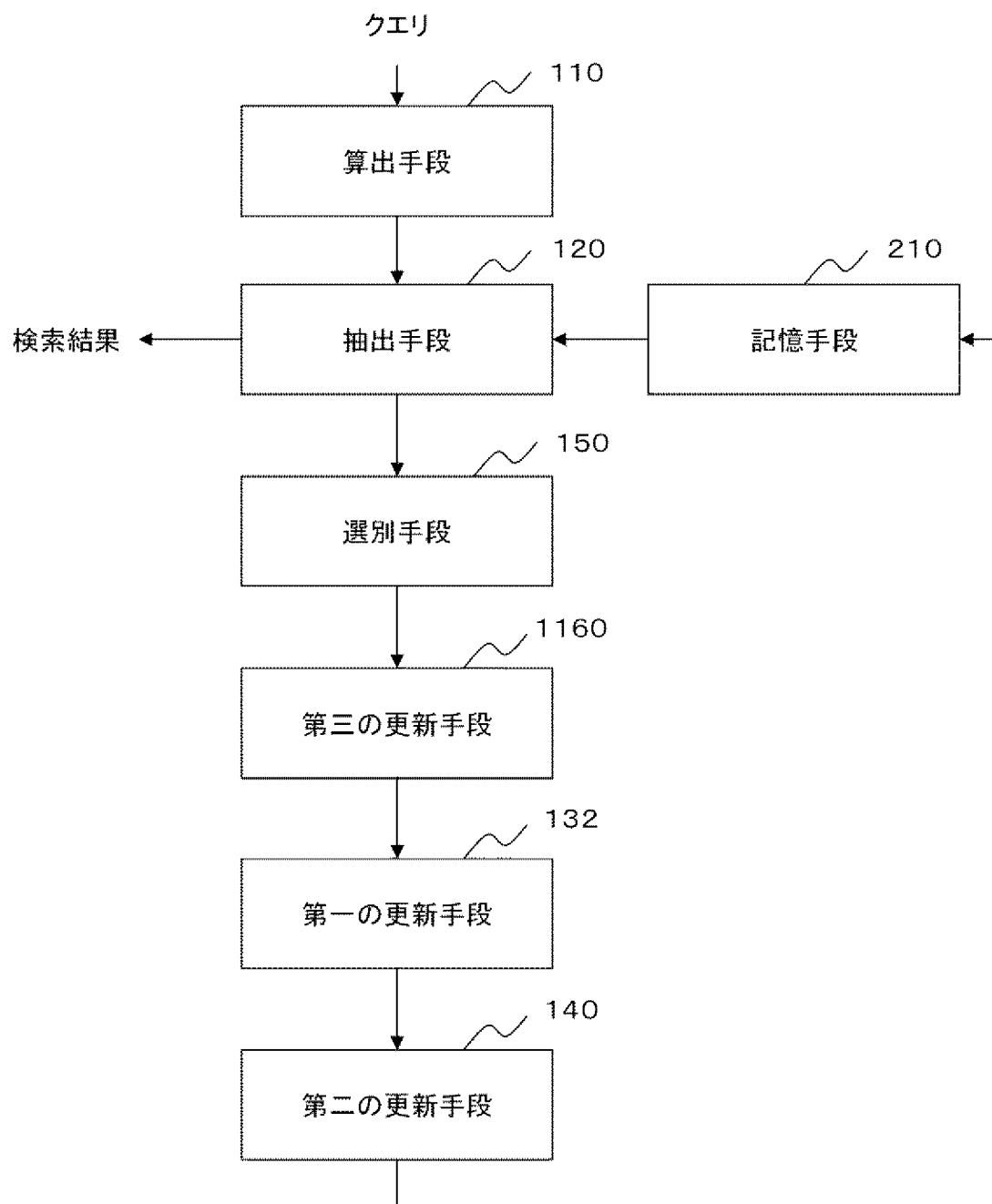
[図6]



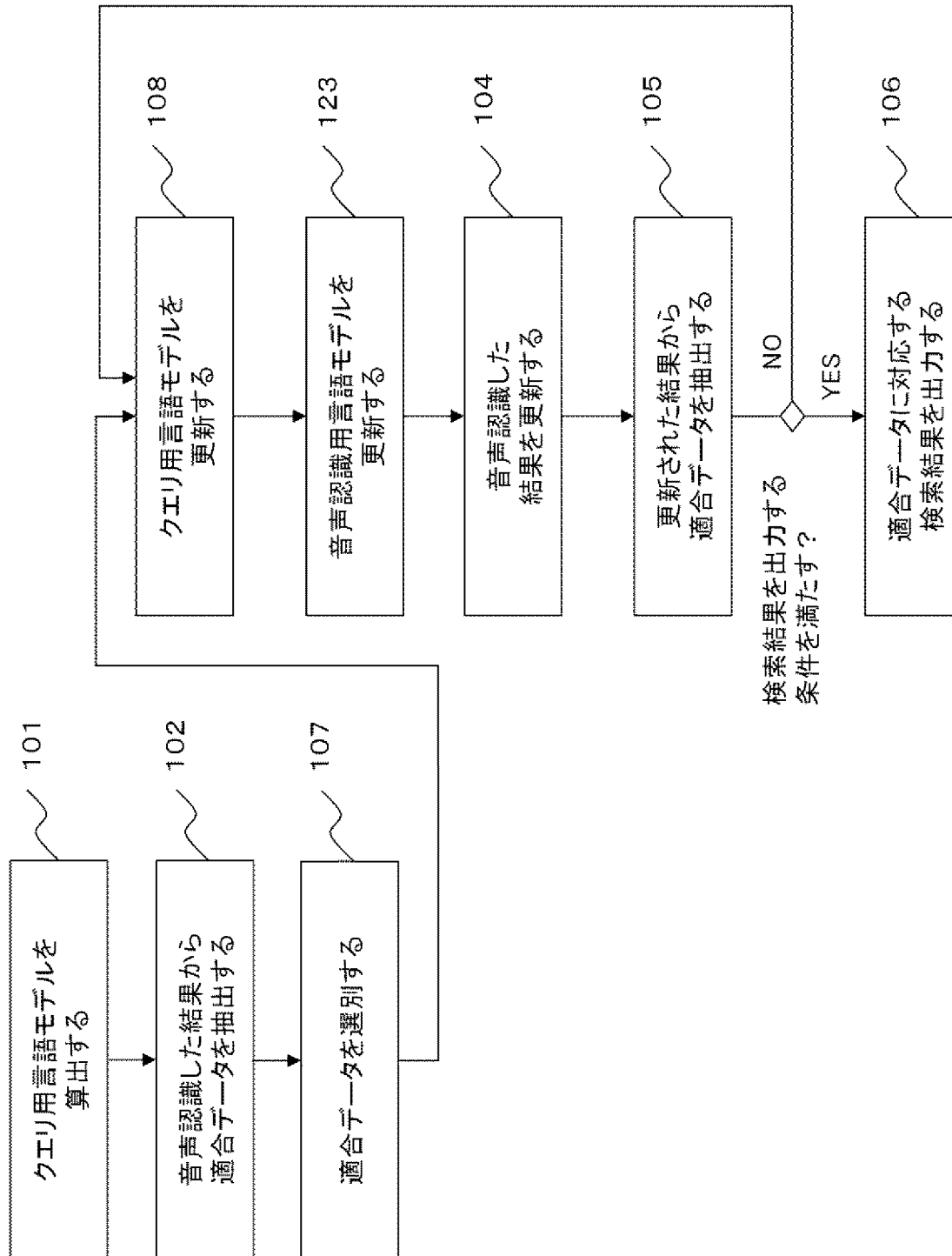
[図7]



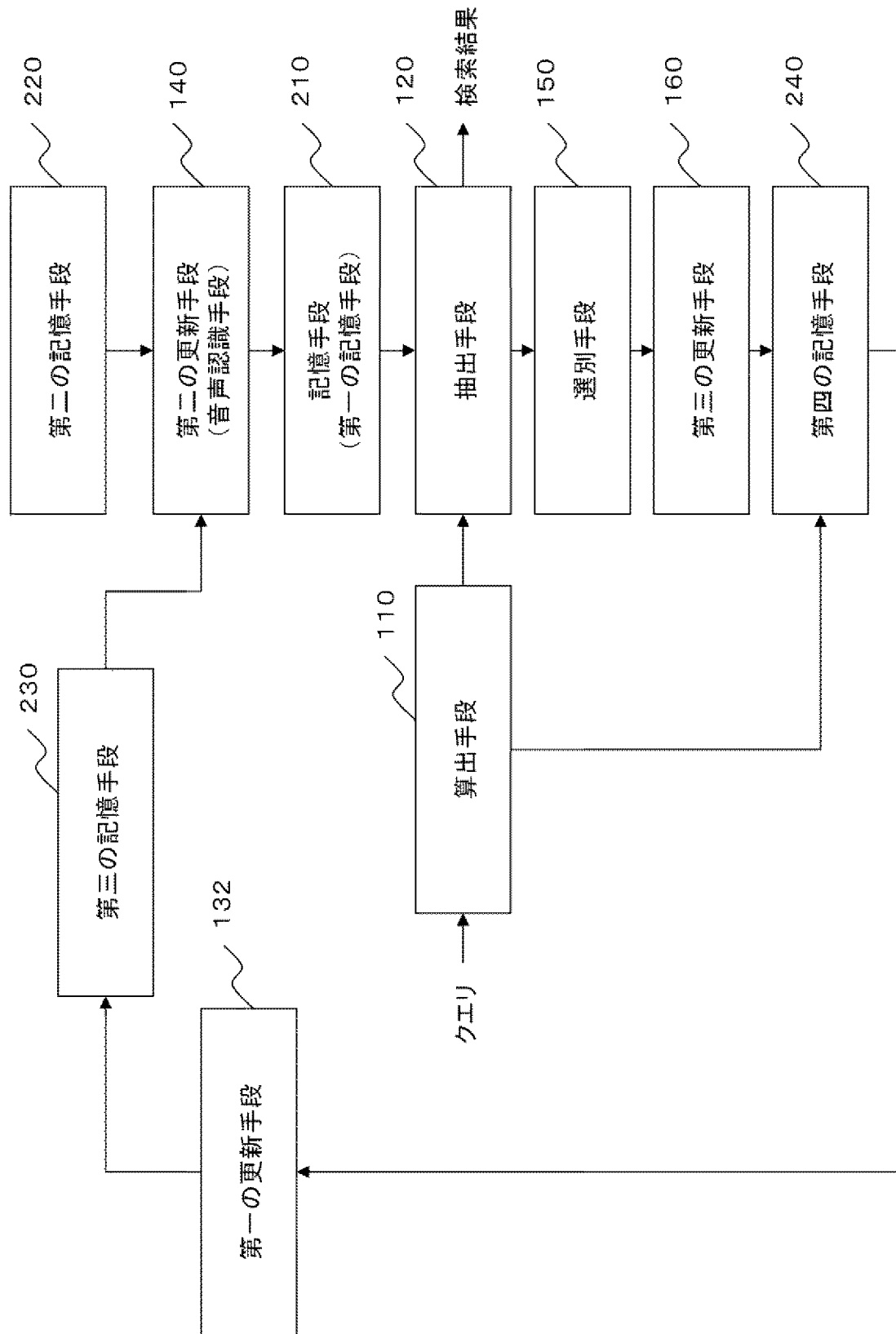
[図8]



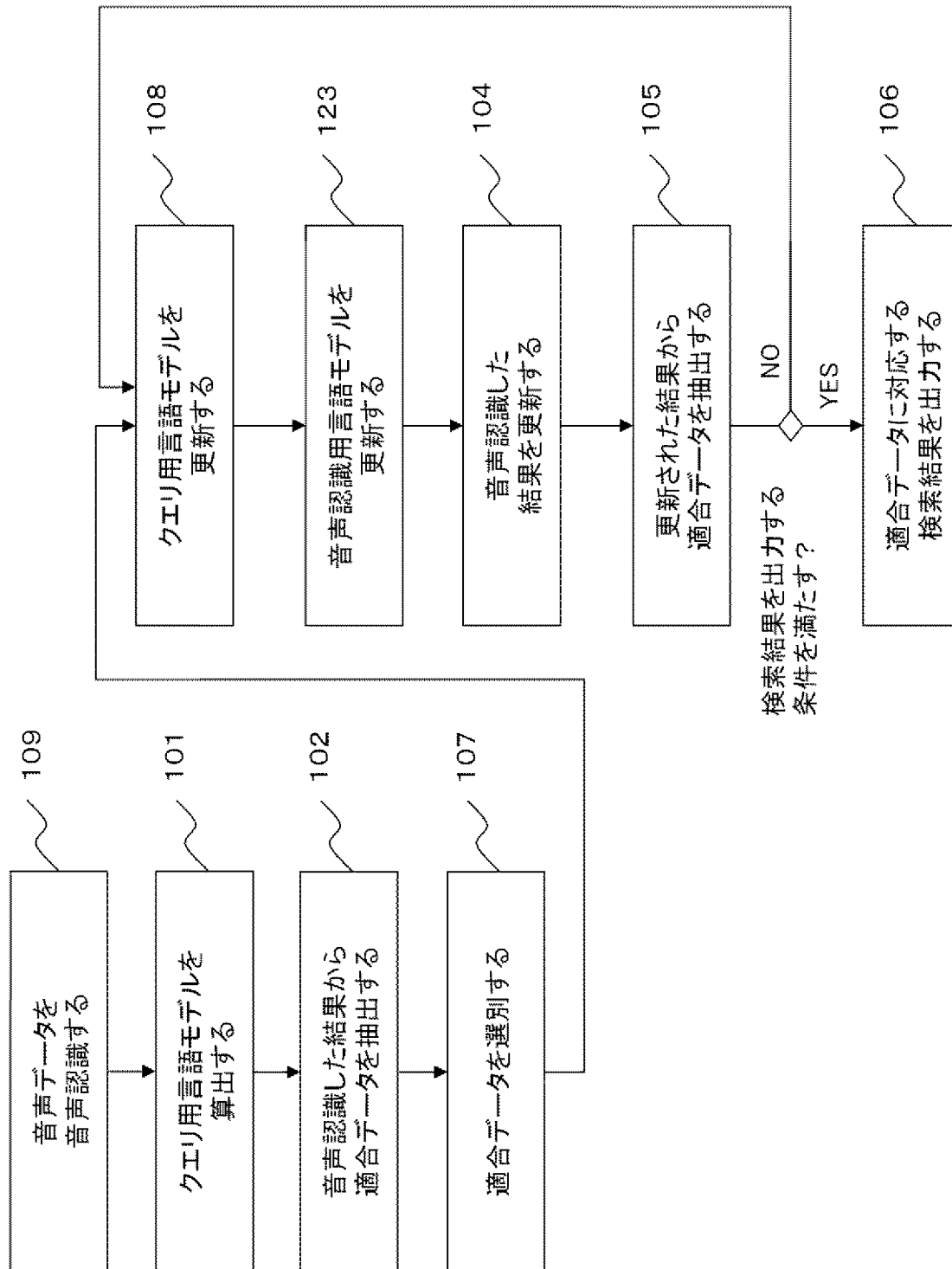
[図9]



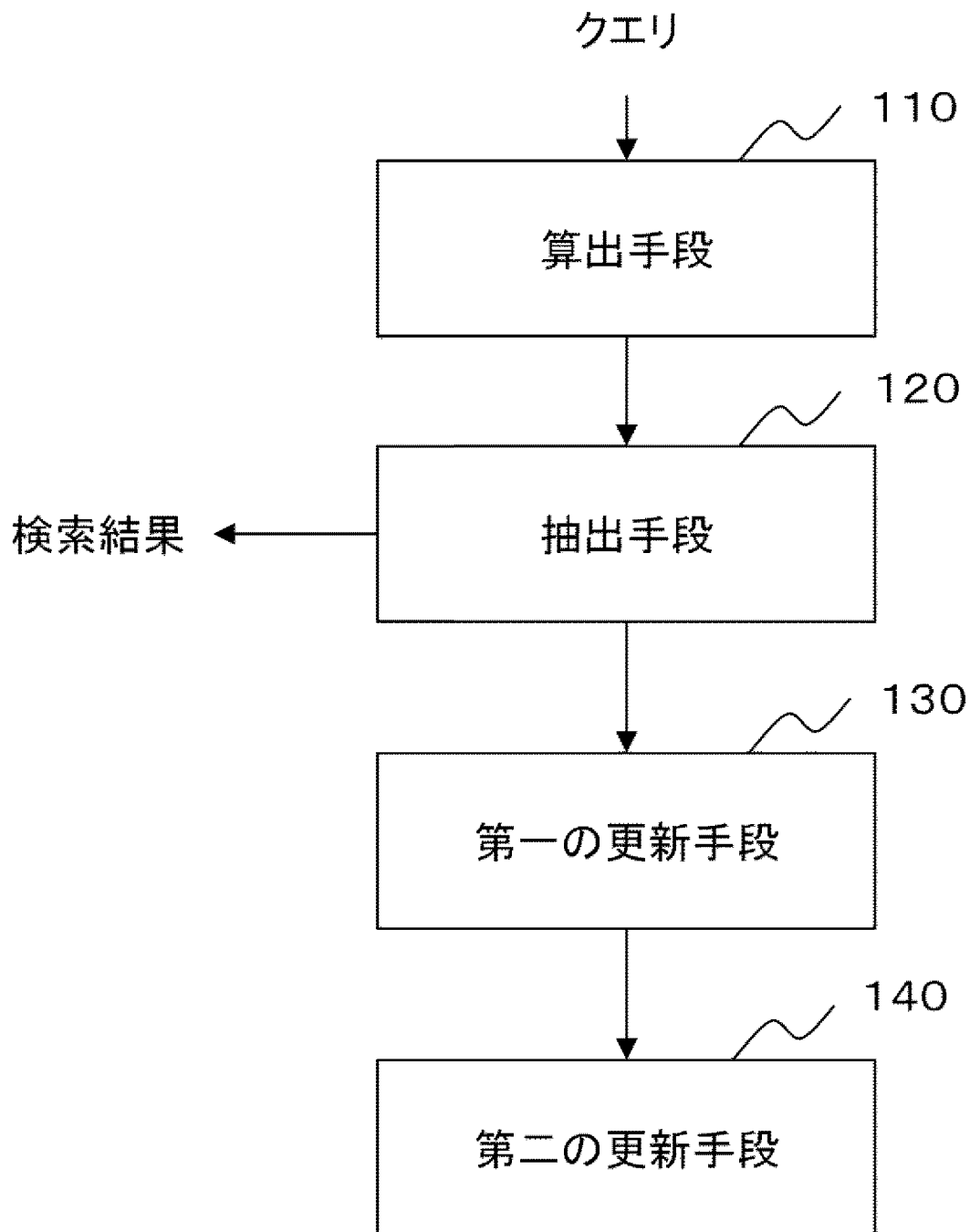
[図10]



[図11]



[図12]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/005401

A. CLASSIFICATION OF SUBJECT MATTER

G06F17/30(2006.01)i, G10L15/00(2013.01)i, G10L15/06(2013.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F17/30, G10L15/00, G10L15/06

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2013
Kokai Jitsuyo Shinan Koho	1971-2013	Toroku Jitsuyo Shinan Koho	1994-2013

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y A	JP 2003-36093 A (Japan Science and Technology Corp.), 07 February 2003 (07.02.2003), paragraphs [0010] to [0013]; fig. 1 & US 2004/0254795 A1 & WO 2003/010754 A1 & CA 2454506 A1	1, 5, 7, 9 6 2-4, 8, 10
Y	JP 2008-97082 A (Mitsubishi Electric Corp.), 24 April 2008 (24.04.2008), paragraph [0017]; fig. 2 (Family: none)	6
A	JP 2003-271629 A (National Institute of Advanced Industrial Science and Technology), 26 September 2003 (26.09.2003), entire text; all drawings (Family: none)	1-10

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
23 October, 2013 (23.10.13)

Date of mailing of the international search report
05 November, 2013 (05.11.13)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2013/005401

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2004-348552 A (Nippon Telegraph and Telephone Corp.), 09 December 2004 (09.12.2004), entire text; all drawings (Family: none)	1-10

A. 発明の属する分野の分類（国際特許分類（I P C））

Int.Cl. G06F17/30(2006.01)i, G10L15/00(2013.01)i, G10L15/06(2013.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（I P C））

Int.Cl. G06F17/30, G10L15/00, G10L15/06

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 3 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 3 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 3 年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X Y A	JP 2003-36093 A（科学技術振興事業団）2003.02.07, 【0010】 －【0013】, 第1図 & US 2004/0254795 A1 & WO 2003/010754 A1 & CA 2454506 A1	1, 5, 7, 9 6 2-4, 8, 10
Y	JP 2008-97082 A（三菱電機株式会社）2008.04.24, 【0017】, 第2図（ファミリーなし）	6
A	JP 2003-271629 A（独立行政法人産業技術総合研究所）2003.09.26, 全文, 全図（ファミリーなし）	1-10

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

2 3 . 1 0 . 2 0 1 3

国際調査報告の発送日

0 5 . 1 1 . 2 0 1 3

国際調査機関の名称及びあて先

日本国特許庁（I S A / J P）
郵便番号100-8915
東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

吉田 誠

5 M

3 6 5 9

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 9 9

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2004-348552 A (日本電信電話株式会社) 2004. 12. 09, 全文, 全図 (ファミリーなし)	1-10