

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有權機關  
國際事務局



(43) 国際公開日  
2008年2月14日 (14.02.2008)

PCT

(10) 国際公開番号  
WO 2008/018274 A1

- (51) 國際特許分類:  
G06F 17/22 (2006.01)      G06F 17/21 (2006.01)

(21) 國際出願番号: PCT/JP2007/064209

(22) 國際出願日: 2007年7月18日 (18.07.2007)

(25) 國際出願の言語: 日本語

(26) 國際公開の言語: 日本語

(30) 優先権データ:  
特願2006-218970 2006年8月10日 (10.08.2006) JP

(71) 出願人(米国を除く全ての指定国について): シャープ  
株式会社 (SHARP KABUSHIKI KAISHA) [JP/JP]; 〒  
5458522 大阪府大阪市阿倍野区長池町22番22号  
Osaka (JP).

(72) 発明者; および

(75) 発明者/出願人(米国についてのみ): 広瀬 齊志  
(HIROSE, Hitoshi). 斗谷 充宏 (HAKARIDANI, Mitsuhiro).

(74) 代理人: 特許業務法人原謙三國際特許事務所  
(HARAKENZO WORLD PATENT & TRADE-  
MARK); 〒5300041 大阪府大阪市北区天神橋2丁目  
北2番6号 大和南森町ビル Osaka (JP).

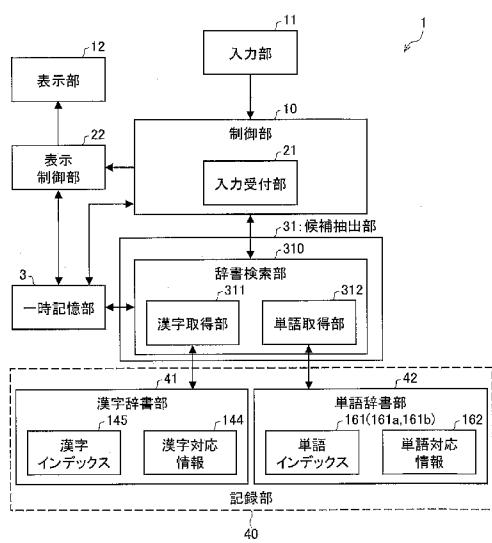
(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH,  
BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK,  
DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM,  
GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP,  
KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME,  
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,  
OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK,  
SL, SM, SV, SY, 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, LS, MW, MZ, NA, SD,  
SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY,  
KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG,

[続葉有]

**(54) Title:** CHARACTER CONVERTING DEVICE AND CHARACTER CONVERTING DEVICE CONTROL METHOD

(54) 発明の名称: 文字変換装置、文字変換装置の制御方法



- 11 INPUT UNIT
  - 12 DISPLAY UNIT
  - 22 DISPLAY CONTROL UNIT
  - 10 CONTROL UNIT
  - 21 INPUT RECEIVING UNIT
  - 3 TEMPORARY STORAGE UNIT
  - 31 CANDIDATE EXTRACTING UNIT
  - 310 DICTIONARY SEARCHING BLOCK
  - 311 KANJI-CHARACTER ACQUIRING SECTION
  - 312 WORD ACQUIRING SECTION
  - 41 KANJI-CHARACTER DICTIONARY UNIT
  - 145 KANJI-CHARACTER INDICES
  - 144 KANJI-CHARACTER CORRESPONDENCE INFORMATION
  - 40 RECORDING UNIT
  - 42 WORD DICTIONARY UNIT
  - 161 ('161a, 161b) WORD INDICES
  - 162 WORD CORRESPONDENCE INFORMATION

**(57) Abstract:** A character converting device (1) comprises a kanji character dictionary unit (41) containing a kanji character correspondence information (144) including kanji characters and their readings associated with the kanji characters and kanji character indices (145) in which the first characters of pin-ins are associated with address information on the storage locations of the kanji characters the first parts of which are the first characters in the kanji character corresponding information (144), a word dictionary unit (42) containing word correspondence information (162) including words and their readings associated with the words and word indices (161) in which the first characters of the words are associated with the address information on the storage locations of the words the first parts of which are the first characters in the word correspondence information (162), and word acquiring means which refers to the kanji character dictionary unit (41) and the word dictionary unit (42) according to the first character of the first pin-in of an input character string and acquires a word corresponding to the input character string. The time required to search for a word corresponding to the input reading can be shortened.

(57) 要約： 文字変換装置（1）は、漢字と当該漢字の読みとが対応付けられた漢字対応情報（144）および、ピンインの先頭文字と該先頭文字を先頭とする漢字の、漢字対応情報（144）での格納先のアドレス情報が対応付けられた漢字インデックス（145）を有する漢字辞書部（41）と、単語と該単語の読みとが対応付けられた単語対応情報（162）および、単語の先頭となる先頭文字と、該先頭文字を先頭に持つ単語の、単語対応情報（162）での格納先のアドレス情報とが対応付けられた単語インデックス（161）を有する単語辞書と、入力文字列の先頭ピンインの先頭文字に基づいて、漢字辞書部（41）と単語辞書部（42）とを参照し、入力文字列に該当する単語を取得する単語取得手段と、

W を備えている。よって、入力された読みに対する単語の検索時間を短縮できる。



CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, 添付公開書類:  
IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, — 國際調査報告書  
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, — 補正書  
ML, MR, NE, SN, TD, TG).

## 明細書

文字変換装置、文字変換装置の制御方法

### 技術分野

[0001] 本発明は、ユーザが、キーボードやタッチパネル、タブレット、マウス等のポインティングデバイスを使って文字あるいは文字列を入力するための装置であり、変換指示を受け付けて、入力された文字あるいは文字列に応じて、変換された文字あるいは文字列を選択可能に表示し、この表示された文字あるいは文字列をユーザが選択することで所望の文字あるいは文字列の入力が行われる文字変換装置に関するものである。

### 背景技術

[0002] 日本語や中国語などの漢字や単語を、ユーザがキーボードを用いて情報機器へ文字入力し表示させるには、漢字や単語に対する読み(表音文字)を入力し、変換指示操作を行うことで実現されている。読みを漢字に変換するには、読みと漢字や単語とを関連付けた辞書データが必要であり、辞書データの語彙数の量が情報機器(入力システム)の性能の一面となっている。

[0003] しかし、語彙数が多くなると、情報機器では、入力された読みに一致する漢字を辞書データから検索する時間が増加する。そのため、ユーザが読みを入力し変換を指示してから、変換結果を表示するまでのレスポンスが悪くなり、結果的に情報機器の使いにくさに繋がる。

[0004] そこで、漢字や単語の検索時間を削減するための工夫が行われている。例えば、辞書中の単語について、読みの先頭文字毎にその先頭文字の読みで始まる単語の読みの最大文字数を索引情報中に持たせることで、入力文字列に対する検索対象の制限を行い、検索に関する処理量の削減を行う変換装置がある(例えば、特許文献1参照)。

特許文献1:日本国公開特許公報特開平6-149787公報(1994年5月31日公開)  
発明の開示

- [0005] しかしながら、上記従来の装置では、入力された読みの文字数が、索引情報に記録されている入力読み文字列の先頭文字の読みで始まる単語の読みの最大文字数以内であれば、その読みの範囲の全辞書データと一致しているか否かをくまなくチェックする必要がある。そのため、変換できる語彙数を増やすために辞書データに収録する語彙数が増加すると、先頭読み文字毎の漢字、単語数が多くなり、前記条件の時の処理時間が増加する。結局、表示出力までのレスポンスが低下してしまい、情報機器は使いにくいものとなってしまう。
- [0006] 本発明は、上記の問題点に鑑みてなされたものであり、その目的は、入力された表音文字に対する単語の検索時間を短縮できる文字変換装置、文字入力方法、文字入力プログラムおよびコンピュータ読取可能な記録媒体を実現することにある。
- [0007] 本発明に係る文字変換装置は、上記課題を解決するために、入力された1以上の表音文字よりなる入力文字列を変換文字に変換し、当該変換文字からなる単語を決定できる文字変換装置において、  
(a-1) 変換文字を識別するための変換文字識別情報と、該変換文字の表音を示す変換文字表音情報とが対応付けられた変換文字対応情報および、(a-2) 入力される表音文字と、該表音文字を上記変換文字表音情報の先頭として持つ上記変換文字識別情報の、上記変換文字対応情報での格納先を示す変換文字格納先情報とが対応付けられた変換文字インデックス、を有する変換文字辞書と、  
(b-1) 単語を識別するための単語識別情報と、該単語の表音を示す単語表音情報とが対応付けられた単語対応情報および、(b-2) 単語の先頭となる先頭変換文字を識別するための先頭変換文字識別情報と、該先頭変換文字識別情報が示す先頭変換文字を先頭に持つ単語の、上記単語対応情報での格納先を示す単語格納先情報とが対応付けられた単語インデックス、を有する単語辞書と、  
(c) 入力された上記入力文字列の先頭の表音文字を基に、上記変換文字辞書を参照し、該入力文字列の先頭の表音文字を上記変換文字表音情報の先頭として持つ変換文字識別情報を取得する先頭変換文字取得手段と、  
(d) 取得された上記先頭変換文字識別情報を基に、上記単語辞書を参照し、該先頭変換文字識別情報が示す変換文字を先頭に持つ単語を取得する単語取得手段

と、を備えていることを特徴としている。

- [0008] 上記構成によると、初めに先頭変換文字取得手段が、変換文字辞書を参照に、入力文字列の先頭の表音文字から、変換文字を検索し、次に、単語取得手段が、単語辞書を参照に、先頭変換文字取得手段が取得した変換文字を先頭に持つ単語を取得することができる。よって、単語の検索処理に係る時間を短縮できる。つまり、入力文字列と表音(読み)情報との比較を行う回数が制限できるので、応答速度の速い文字変換装置を提供することができる。
- [0009] ここで、変換文字とは、変換された後の文字を示している。変換文字が漢字であるとすると、表音文字(読み)の先頭文字は次のようにになる。例えば、中国語で、「
- [0010] [表1]

## 毎年

- [0011] (以下では「外字1」とする)」の読みを表記すると「meinian」であり、「m」が読みの先頭文字にあたる。また、日本語で、「毎年」の読みを表記すると、「まいとし」なので、先頭文字は「ま」である。

- [0012] また、先頭変換文字は、例えば「外字1」では、「

- [0013] [表2]

## 每

- [0014] (以下では「外字2」とする)」に当たる。変換文字表音情報は、例えば「外字1」では、「meinian」である。

- [0015] なお、変換文字識別情報は、変換後の文字である変換文字を識別できる情報であればどのようなものでもよく、変換文字そのものであってもよい。例えば、変換文字が漢字である場合、漢字そのものであってもよいし、漢字コードであってもよい。これは、単語識別情報と、先頭変換文字識別情報とについても同様である。

- [0016] 本発明のさらに他の目的、特徴、および優れた点は、以下に示す記載によって十分わかるであろう。また、本発明の利益は、添付図面を参照した次の説明で明白になるであろう。

## 図面の簡単な説明

[0017] [図1]本発明の一実施形態を示すものであり、文字変換装置の要部構成を示すブロック図である。

[図2]上記文字変換装置が行う文字の変換についての処理のフローを示す図である。

[図3]本発明の実施例に係る漢字インデックスを説明する図である。

[図4]本発明の実施例に係る漢字対応情報を説明する図である。

[図5]本発明の実施例に係る1次インデックスを説明する図である。

[図6]本発明の実施例に係る2次インデックスを説明する図である。

[図7]本発明の実施例に係る単語対応情報を説明する図である。

[図8]本発明の実施例に係るピンイン辞書インデックスを説明する図である。

[図9]本発明の実施例に係るピンイン辞書を説明する図である。

[図10]本発明の実施例に係るピンイン分割処理例を表した図である。

[図11]本発明の実施例に係るピンイン分割処理において、確定したピンインを除くピンインの分割処理例を表した図である。

[図12]本発明の実施形態に係る文字変換装置の要部構成を示すブロック図である。

[図13]本発明の文字変換装置を適用したマルチファンクションプリンタの概略構成を示すブロック図である。

[図14]文字変換装置の操作パネルに表示される、文字入力処理時の操作画面の一例を示す図である。

[図15]文字変換装置の候補抽出部が候補漢字を抽出するときに参照する、漢字辞書記録部に記録されている漢字辞書の例を示す図である。

[図16]本発明の他の実施形態に係る候補熟語を抽出する文字変換装置の要部構成を示すブロック図である。

[図17]文字変換装置の操作パネルの表示部に表示される候補熟語の例を示す図である。

[図18]熟語辞書記録部が記録する熟語辞書の例を示す図である。

[図19]本発明の他の実施形態に係る文字変換装置の処理の流れを示すフローチャートである。

[図20]操作パネル上の候補文字表示領域および確定文字表示領域の表示例を示す図である。

[図21]本発明の他の実施形態に係る学習機能を備えた文字変換装置の要部構成を示すブロック図である。

[図22]候補順序決定部32のさらに詳細な構成を示すブロック図である。

[図23]連結度記録部に記録される、漢字(熟語)ごとの直前(前方)文字列との連結度を示した連結度テーブルの例を示す図である。

[図24]文字変換装置の候補順序決定部が表示順序を決定するときの、候補の分類手順を示す図である。

[図25]候補ごとに算出した表示順序の評価値を記憶するデータ構造の一例を示す図である。

[図26]候補順序決定要素記録部に記録される優先度情報の一例を示す図である。

[図27]操作パネルの文字入力枠に筆記された漢字の例、および、該漢字の座標情報の例を示す図である。

[図28]操作パネルの表示例を示す図であり、図28の(a)は、表示一体型タブレットの文字入力枠に筆記された漢字の一例を示す図であり、図28の(b)は、候補文字表示領域に表示された認識文字候補の一例を示す図であり、図28の(c)は、候補文字表示領域に表示された候補熟語の一例を示す図である。

[図29]筆記データ「外字2」の認識結果の一例を示す図である。

[図30]本発明の他の実施形態に係る繁体字を入力する文字変換装置の要部構成を示すブロック図である。

[図31]本発明の他の実施形態における、繁体字入力のための注音文字入力処理時の操作画面の一例を示す図である。

[図32]本発明の他の実施形態における、文字変換装置の候補抽出部が候補漢字を抽出するときに参照する、漢字辞書記録部に記録されている漢字辞書の一例を示す図である。

[図33]本発明の他の実施形態における、文字変換装置の熟語辞書記録部が記録する熟語辞書の一例を示す図である。

[図34]本発明の他の実施形態における、文字変換装置の連結度記録部に記録される、漢字(熟語)ごとの直前(前方)文字列との連結度を示した連結度テーブルの例を示す図である。

[図35]操作パネルの表示部に表示される操作画面の一例を示す図である。

[図36]操作パネルの表示部に表示される操作画面の一例を示す図である。

[図37]操作パネルの表示部に表示される操作画面の一例を示す図である。

[図38]操作パネルの表示部に表示される操作画面の一例を示す図である。

## 発明を実施するための最良の形態

[0018] 本発明の一実施形態(実施形態1～4)について図12～38に基づいて説明すると以下の通りである。本実施形態では、一例として、入力装置と出力装置とが一体になっているタッチパネルを備えたマルチファンクションプリンタ(複写機)に、本発明の文字変換装置を適用した場合について説明する。しかし、本実施形態で説明する複写機は一例であって、本発明の文字変換装置を、携帯電話、携帯端末、パソコン、リモコンなどに適用することも可能である。

[0019] [背景技術と問題点]

従来の文字変換装置において表意文字を得ようとすると、表音文字よりも操作が繁雑になってしまふ。表意文字の文字種は膨大な量であり、それらに1つ1つキーを割り当てるなど現実的ではない。さらに、表意文字は同音異義語が多く存在することが一般的であり、複数の表音文字の組み合わせによっても、それに対応する表意文字の候補は膨大な量となる。その結果、膨大な量の候補から適切な表意文字を選び出す作業は、ユーザに多くの手間を強いることになる。

[0020] この問題点は、日本語の文字入力システムに限って生じるものではなく、同様の表意文字入力システムであれば、同様に生じるものである。例えば、中国語はピンインを利用(中国語の特徴である四声は省略)して入力する。漢字を入力するには、次のような手順で行う。入力したい漢字のピンイン表記(アルファベットの文字列)をキーボードから入力し、そのピンインに対応する漢字への変換操作を行って所望の漢字を得る。上記問題点を解決するために、所望の表意文字を得るために表意文字の入力操作を簡素化することが可能な文字変換装置について説明する。

- [0021] また、本実施形態では、ピンイン表記(アルファベット列)から中国語(簡体字)を入力可能な文字変換装置について説明するが、本発明の文字変換装置は中国語に限定されず、種々の表音文字から所望の表意文字を得られるものである。なお、上記ピンイン表記について、中国語には本来、4種の異なる発音を表す四声が存在するが、説明の簡略化のために四声の概念を省略して説明する。
- [0022] [実施形態1]  
(マルチファンクションプリンタの構成)  
図13は、本発明の文字変換装置1を適用したマルチファンクションプリンタの概略構成を示すブロック図である。
- [0023] マルチファンクションプリンタは、図13に示すとおり、制御部10、操作パネル(タブレット)2、一時記憶部3、記録部40、画像読み取りユニット4、画像転写ユニット5、紙送りユニット6を備えている。なお、ここで説明するマルチファンクションプリンタの構成は、本発明を実施するための一例として示すものであって、本発明の範囲を限定するものではない。
- [0024] 制御部10は、マルチファンクションプリンタを統括制御するものである。制御部10は、記録部40に記録されている各種プログラムを読み出して、複写機としての機能を果たす、画像読み取りユニット4、画像転写ユニット5、および、紙送りユニット6などを制御し、複写などの処理を行う。また、本発明に係る文字入力処理を行うための各部もこの制御部10に含まれている。
- [0025] 記録部40は、操作パネル2、複写機として動作するための上述の各種ユニットの制御プログラム、および、文字変換装置1の制御プログラムを記録するものであり、ROM(read only memory)などの不揮発性記憶素子で実現される。
- [0026] 一時記憶部3は、プログラムを処理するための作業領域、読み取り画像の一時記憶に利用されるほか、本実施形態では、文字入力処理時の学習文字(列)などの一時記憶に利用される。RAM(random access memory)などの揮発性メモリで実現される。
- [0027] 操作パネル2は、ユーザがマルチファンクションプリンタに対して各種処理の実行を指示するためのものである。例えば、表示装置と入力装置とが一体に形成されたタッチパネルや、表示一体型タブレットで実現される。さらに、表示一体型タブレットを用

いて筆記された文字を認識する手書き文字認識機能を備えていてもよい。また、ユーザが表示一体型タブレットの指示した位置に基づいて、その指示された文字、文字列、その他機能キーなどの情報を検出する機能を備えている。

- [0028] ユーザは、操作パネル2に表示された、アルファベット(表音文字)に対応する各キーにペンや指などで軽く触れることにより、所望のアルファベットをマルチファンクションプリンタに入力することが可能となる。
- [0029] あるいは、上記マルチファンクションプリンタは、操作パネル2に加えて、キーボード(第1入力部／第3入力部)7(例えば、PCで使用する汎用キーボード)を備えていてもよい。この場合、操作パネル2の表示されるキー、および、キーボード7のキーのいずれを用いてもアルファベットを入力することが可能となる。表音文字を入力するためのキーを備えた入力部(第1入力部)を、操作パネル2およびキーボード7の両方で同時に実現することもできるし、キーボード7のみで実現することもできる。あるいは、表音文字の入力を操作パネル2で行うか、キーボード7で行うかを切り替えられる構成としてもよい。
- [0030] 以下、上記文字入力処理を行う本発明の文字変換装置1について、さらに詳細に説明する。

[0031] (文字変換装置の構成)

図12は、本実施形態に係る文字変換装置1の要部構成を示すブロック図である。文字変換装置1は、図12に示すとおり、操作パネル2、制御部10、記録部40、および、一時記憶部3を備えた構成となっている。

- [0032] 操作パネル2は、上述したように、表示装置と入力装置とが一体に形成されたタッチパネルで実現されている。
- [0033] 表示部12は、ユーザに各種情報を提供するものである。図14は、操作パネル2に表示される文字入力処理時の操作画面の一例を示す図である。図14に示すとおり、表示部12には、ユーザが文字を入力するためのキーボードを表示する文字入力領域121、入力の候補として挙げられている文字(または、文字列)を表示する候補文字(列)表示領域122、および、すでに入力が完了し確定している入力確定文字を表示する確定文字(列)表示領域123が表示されている。入力部11は、ユーザが中国

語発音表記のピンイン(表音文字列／表音情報)を入力するためのものであり、文字入力領域121に表示されたアルファベットのキーがユーザにより押下されると、パネル上のその押下位置に対応するアルファベットの信号が、文字変換装置1の入出力制御部20に入力される。また、入力部11は、ユーザが入力を確定させたい文字を選択するためのものもある。候補文字表示領域122に表示されている漢字(表意文字／変換文字)が押下されると、その押下位置に表示されている漢字の信号が、入出力制御部20に入力され、該漢字の入力が確定する。

- [0034] 制御部10は、入出力制御部20および文字入力制御部30を制御するものである。入出力制御部20は、操作パネル2と制御部10との間でやりとりされる信号の入出力を制御するものであり、入力受付部21、表示制御部(表示制御手段)22、および、候補選択受付部23を含んでいる。文字入力制御部30は、文字入力処理を実行するものであり、候補抽出部(文字候補抽出手段／文字列候補抽出手段)31、候補順序決定部(候補順序決定手段)32、および、文字確定部(文字入力確定手段)33を含んでいる。なお、制御部10内の各部は、コンピュータの中央演算装置(CPU)およびCPUによって実行されるソフトウェアにより実現される。このソフトウェアも、上述の記録部40に格納され、一時記憶部3に読み出されてCPUにより実行される。
- [0035] 入出力制御部20の入力受付部21は、入力部11から入力された文字(アルファベットなど)の信号を受け付けて、候補抽出部31に供給するものである。表示制御部22は、表示部12の文字入力領域121、候補文字表示領域122、および、確定文字表示領域123の各々に表示すべき情報が表示されるよう表示部12を制御するものである。候補選択受付部23は、入力部11から入力された、候補文字表示領域122内の選択された候補漢字の信号、および、該選択された候補漢字の入力を確定する指示信号を受け付けて、文字確定部33に供給するものである。
- [0036] 文字入力制御部30の候補抽出部31は、入力された文字に基づいて、候補となる漢字(または)単語を記録部40に記憶している各辞書部から取得するものである。本実施形態では、候補抽出部31は、入力されたアルファベット(列)に基づいて、候補となる漢字を記録部40の漢字辞書部(文字記録部)41から抽出する。
- [0037] より詳細には、まず、入力されたアルファベット(列)に対応するピンインを特定して、

特定した各ピンインに対応する漢字を、候補漢字として抽出する。候補抽出部31は、抽出した候補漢字を、内部バッファ(不図示)に記憶し、候補順序決定部32および表示制御部22に候補漢字の表示を指示する。

- [0038] 候補抽出部31が参照する漢字辞書部41には、アルファベット(列)－ピンイン－漢字が対応付けて記録されている。漢字辞書部41の詳細は後述する。
- [0039] 候補順序決定部32は、候補抽出部31が抽出した候補漢字の表示順序を、所定の条件に基づいて決定するものである。表示順序とは、図14に示すように、候補文字表示領域122に候補漢字を表示する順序のことである。抽出されたすべての候補漢字が候補文字表示領域122内に一度に表示できる場合、順序が上の候補漢字ほど、候補文字表示領域122の上部に表示されるようにし、ユーザが選択しやすいようすればよい。また、候補漢字が一度に表示できる数を上回る場合は、下位に続く順序の候補漢字を表示するようページを切り換えるようにしてもよい。この場合でも、順序が上の候補漢字ほど初めのページに表示されるようにして、上位の候補漢字ほどユーザに選ばれやすいようにしておく。したがって、候補順序決定部32は、抽出された候補漢字を、所定の条件に基づいてユーザに選択されやすいものから順に表示される順序を決定する。どのような条件に基づいて候補漢字の順序を決定するのかについては後に詳しく説明する。
- [0040] 文字確定部33は、入力部11より入力されたユーザが選択した候補漢字の信号、および、該選択された候補漢字の入力を確定する指示信号を、候補選択受付部23より受け取り、ユーザにより選択された上記候補漢字の入力を確定させるものである。まず、選択された候補漢字は、確定前の選択されている状態を示す下線とともに、確定文字表示領域123に表示させる。さらに、入力確定の指示信号に応じて文字確定部33が確定した候補漢字は、入力確定文字として下線が付されない状態で、表示制御部22を介して確定文字表示領域123に表示される。
- [0041] (漢字辞書)
- 次に、漢字辞書部41に記録される漢字辞書について説明する。図15は、候補抽出部31が候補漢字を抽出するときに参照する、漢字辞書部41に記録されている漢字辞書の例を示す図である。

[0042] 図15に示すとおり、本実施形態における漢字辞書140は、入力受付部21が受け付けたアルファベット(列)とピンインとの対応付けを記録した、アルファベット－ピンイン対応表(A／P対応表141)、および、ピンインと漢字との対応付けを記録した、ピンイン－漢字対応表(P／C対応表142)とを含んでいる。これにより、入力されたアルファベット(列)に対応する候補漢字を直接抽出することが可能となる。

[0043] 具体的には、例えば、ユーザが「外字2」(ピンインは「mei」という漢字を入力したために、入力部11を用いて、まず「m」のキーを押下したとする。入力受付部21より「m」の信号が供給されると、候補抽出部31は、まず、漢字辞書140のA／P対応表141から、「m」と前方一致するピンイン(図15の例では、「ma」と「mei」)を特定し、該ピンインに対応する漢字が記録されている、P／C対応表142のアドレスを検出する(「cp\_ml」と「cp\_mi」)。

[0044] 続いて、P／C対応表142の「cp\_ml」と「cp\_mi」とに基づいて、候補抽出部31は、「m」と前方一致のピンインに対応する漢字(

[0045] [表3]

妈

[0046] 、麻、没、…「外字2」、

[0047] [表4]

浼

[0048] …)を抽出する。抽出された漢字は、上述したとおり、候補漢字として候補順序決定部32に供給される。ここで、ユーザによりさらに「mei」と入力された場合には、候補抽出部31は、「mei」と前方一致する漢字に絞り込んで、漢字辞書部41から候補漢字を抽出する。

[0049] 上記構成によれば、候補抽出部31は、ユーザが入力したアルファベットを受け付けて、そのアルファベットを含む、ピンインを特定する。候補抽出部31は、特定したピンインに基づいて、ユーザが入力したいと所望する漢字の候補を漢字辞書部41から抽出することができる。

[0050] 候補順序決定部32は、上記候補抽出部31が抽出した漢字の候補を所定の条件

に基づいて並べ替え、候補の表示順序を決定する。表示制御部22は、漢字の候補が、上記候補順序決定部32が決定した表示順序で表示されるよう表示部12を制御する。

- [0051] 最後に、ユーザが表示された入力したい候補漢字の中から、所望の漢字を指定すると、文字確定部33は、指定された漢字の入力を確定させる。
- [0052] これにより、所望の漢字を得るための入力操作を簡素化することが可能となる。
- [0053] また、候補抽出部31が抽出する候補となる漢字は、ユーザのアルファベット入力操作により、外部記憶装置(図示せず)への記憶および外部記憶装置からの読み込みが行えるようになっていてもよく、読み込まれた漢字を、候補を表示する際に利用するようにしてもよい。
- [0054] (変形例1－候補単語の出力)  
さらに、本発明の文字変換装置1は、某漢字が入力確定文字として選択されたのち、該漢字から始まる1つの意味を持つ1以上の漢字からなる漢字列(以下、単語(表意文字列)と称する)を候補単語として、候補文字表示領域122(図14)に表示することが好ましい。これにより、ユーザは、上記某漢字とともに単語を形成し得る漢字を候補文字表示領域122において得ることができるので、続く漢字のピンインを別途入力しなくても、所望の漢字を得ることが可能となる。したがって、所望の漢字を得るための入力操作を簡素化することが可能となる。
- [0055] 図16は、候補単語を出力する文字変換装置1の要部構成を示すブロック図である。なお、図16の各構成要素に付された符号は、図12の各構成要素に付された符号に対応しており、同じ符号は、同じ構成要素を示している。したがって、上述の各実施形態すでに説明した構成要素についての説明は繰り返さない。図16に示す文字変換装置1において、図12と異なる点は、文字確定部33から候補抽出部31に制御信号が伝達される点と、記録部40に単語辞書部(文字列記録部)42が含まれている点である。
- [0056] 文字確定部33は、ユーザにより選択された漢字(または、入力確定文字)が候補選択受付部23を介して入力されると、続いて、候補抽出部31に対して、当該選択されたから始まる単語を抽出する指示を送る。

- [0057] 候補抽出部31は、上記の指示に基づいて、上記選択された漢字から始まる単語を、候補単語として単語辞書部42から抽出する。図17は、操作パネル2の表示部12に表示される候補単語の例を示す図である。図17に示す例では、漢字「外字2」が選択されているときに、「外字2」から始まる単語が、候補単語として表示部12の候補文字表示領域122に表示されている。
- [0058] 候補抽出部31が参照する単語辞書部42には、漢字—単語が対応付けて記録されている。図18は、単語辞書部42が記録する単語辞書の例を示す図である。図18に示すとおり、本実施形態における単語辞書は、漢字辞書部41に記録されている漢字と、該漢字から始まる単語との対応付けを記録した、漢字—単語対応表(C/I対応表143)を含んでいる。これにより、選択された漢字に対応する候補単語をさらに抽出することが可能となる。
- [0059] 具体的には、例えば、ユーザにより、漢字「外字2」が選択されると、候補抽出部31は、P/C対応表142(図15)の「外字2」のアドレスを検出する('wp\_mi+j')。続いて、単語辞書部42を参照し、アドレス「wp\_mi+j」に基づいて、「外字2」から始まる単語(「外字2」月、「外字2」天...)を抽出する。抽出された単語は、上述の構成と同様に、候補単語として候補順序決定部32に供給され、候補文字表示領域122に表示される(図17)。
- [0060] ある漢字が入力されると、その漢字の次に続く文字(ユーザが所望する漢字)は、何の脈絡もない漢字よりも、その漢字から始まる単語を構成する漢字である可能性が高い。したがって、入力が確定された漢字から始まる単語を候補として表示すれば、ユーザは、その単語の候補の中から所望の単語を選択するのみで、別途ピンインを入力することなく、中国語を入力することが可能となる。これにより、ユーザは、ピンインの入力に係る操作数について、さらに少ない操作数で、所望の漢字の入力を得ることが可能となる。
- [0061] (フローチャート)
- 次に、本発明の文字変換装置1(図16)の処理の流れを、図19に示すフローチャートと、図20に示す操作パネル2上の候補文字表示領域122および確定文字表示領域123の表示例に基づき説明する。以下では、ユーザが、「外字2」天到郊外散

歩」を入力しようとした場合について説明する。

- [0062] ユーザが操作パネル2の入力部11からアルファベット(漢字「外字2」(mei)のピンインの先頭アルファベット「m」)を入力すると(S201のYES)、入力受付部21は、入力されたアルファベット「m」の信号を、候補抽出部31に供給する。次に、候補抽出部31は、入力されたアルファベットに対応するピンインを特定し、特定したピンインに對応する漢字を、候補漢字として漢字辞書部41から抽出する(S202)。
- [0063] 候補順序決定部32は、所定の条件に基づいて上記抽出された候補漢字の表示順序を決定する(詳細は後述)(S203)。続いて、表示制御部22は、候補順序決定部32が決定した表示順序でn個の候補漢字を表示するよう表示位置を定め、上記候補漢字を表示部12に出力し、ユーザに提示する(S204)。図20の(a)に、n=5として、漢字候補を5個づつ表示する場合の表示例を示す。
- [0064] ここで、ユーザにより、1度目に候補文字表示領域122に表示された漢字以外の次候補漢字を表示する指示(図20の(a)に所望の漢字「外字2」がないので、例えば、「次候補表示」のボタンを押すなどする)が入力されると(S205のNO、S206のYES)、S204に戻り、表示制御部22は、次の候補漢字のリストを表示する(図20の(b))。さらに、次候補漢字の表示指示も、候補漢字の選択指示も入力されず(S205、S206においてNO)、新たにアルファベットが入力された場合には、S202に戻り、候補抽出部31は、これまでに入力されたアルファベット列に対応するピンインをさらに絞り込んで特定し、特定したピンインに對応する漢字を、候補漢字として抽出する。
- [0065] 一方、ユーザが、入力部11を介して、候補文字表示領域122に表示された所望の漢字「外字2」(図20の(b)の“3011”)を選択すると(S205のYES)、候補選択受付部23は、選択された候補漢字「外字2」の信号を受け付けて、文字確定部33に供給する。
- [0066] 文字確定部33は、選択された「外字2」を表示するよう表示制御部22に指示とともに、候補抽出部31に対して、次に「外字2」から始まる単語を抽出するよう指示する。候補抽出部31は、続いて、「外字2」から始まる単語を候補単語として単語辞書部42から抽出する(S207)(図20の(c))。
- [0067] 候補順序決定部32は、所定の条件に基づいて上記抽出された候補単語の表示順

序を決定する(詳細は後述)(S208)。続いて、表示制御部22は、候補順序決定部32が決定した表示順序にて上記候補単語を表示部12に出力し、ユーザに提示する(S209)。

- [0068] ここで、ユーザにより、次候補単語を表示する指示が入力されると(S210のNO、S211のYES)、S209に戻り、表示制御部22は、次の候補単語のリストを表示する。さらに、次候補単語の表示指示も、候補単語の選択指示も入力されず(S210、S211においてNO)、新たにアルファベットが入力された場合には、S202に戻り、候補抽出部31は、これまでに入力されたアルファベットに対応するピンインを特定し、特定したピンインに対応する漢字を、候補漢字として抽出する。
- [0069] 一方、ユーザが、入力部11を介して、候補文字表示領域122に表示された所望の単語「外字2」天(図20の(c)の“020”)を選択すると(S210のYES)、候補選択受付部23は、選択された候補単語「外字2」天の信号を受け付けて、文字確定部33に供給する。
- [0070] 文字確定部33は、選択された候補単語を表示するよう表示制御部22に指示し、表示制御部22は、選択された候補単語を表示部12に表示する(S212)(図20の(d))。同時に、候補抽出部31に対して、次に該候補単語から始まる単語を抽出するよう指示する。文字の入力を終了する指示が入力されなければ(S213のNO)、候補抽出部31は、S207に戻り、上記選択された候補単語から始まる単語を候補単語として単語辞書部42から抽出する。
- [0071] ユーザの次の所望の漢字「到」が候補漢字として表示されない場合、ユーザは、「到」(dao)の先頭ピンイン「d」を入力することができる(S211のNO、S201のYES)(図20の(e))。以降同様の処理を繰り返し、所望の漢字を入力することが可能となる。
- [0072] 選択された候補単語が選択されたのち(「散歩」まで入力が確定されたのち)、文字の入力を終了する指示が入力されれば(S213のYES)、処理を終了する。
- [0073] (変形例2—学習機能)  
さらに、本発明の文字変換装置1は、ユーザの選択により入力が確定された漢字(単語)を記録する構成を備えていることが好ましい。これにより、文字変換装置1は、過去に入力した漢字、つまり、ユーザが使用したことのある漢字の記録に基づいて、

候補漢字を表示することができる。過去に使用された漢字は、ユーザによって再度使用される可能性は大きい。したがって、そのような使用履歴が反映された候補漢字の一覧を用いることにより、ユーザは少ない操作数で所望の漢字を得ることが可能となる。

- [0074] 図21は、学習機能を備えた文字変換装置1の要部構成を示すブロック図である。なお、図21の各構成要素に付された符号は、図12の各構成要素に付された符号に対応しており、同じ符号は、同じ構成要素を示している。したがって、上述の各実施形態すでに説明した構成要素についての説明は繰り返さない。図21に示す文字変換装置1において、図12と異なる点は、文字入力制御部30が、さらに、学習制御部(学習制御手段)34を有している点と、記録部40に学習辞書記録部(学習語記録部)43およびペア学習辞書記録部(学習語記録部)44が含まれている点である。
- [0075] 学習制御部34は、ユーザの指示により、文字変換装置1に入力(使用)された漢字の記録(学習)を制御するものである。具体的には、ユーザにより使用された漢字を学習漢字として、学習辞書記録部43に格納したり(以下、単独学習(機能)と称す)、または、ユーザにより使用された漢字を、直前に入力されている漢字と関連付けてペア学習辞書記録部44に格納したり(以下、ペア学習(機能)と称す)する。ユーザにより過去に入力されたことのある漢字は、ユーザによって再度使用される可能性が大きいので、そのような学習漢字を候補漢字として利用することにより、ユーザがより少ない操作数で所望の漢字を得る確率が高まる。
- [0076] 学習辞書記録部43は、ユーザにより過去に使用された漢字(入力確定文字)を単独学習漢字として記録するものである。学習辞書記録部43に記録される学習辞書には、その単独学習漢字が記録された日時の情報や、使用頻度の情報(過去に何回使用されたか、所定期間内に何回使用されたかなど)が、漢字ごとに対応付けて記録されていてもよい。
- [0077] 学習制御部34は、文字確定部33から、ユーザにより選択された候補漢字の信号を受け取ると、一時記憶部3に該選択された候補漢字を書き込む。学習制御部34は、文字確定部33が確定指示信号に応じて入力を確定させたとき、文字確定部33より当該入力確定文字の単独学習処理の指示を受け付ける。学習制御部34は、一時記

憶部3から、上記候補漢字を内部バッファに読み出し、一時記憶部3から読み出された候補漢字の削除および内容の更新を行う。そして、内部バッファに記憶した候補漢字を単独学習漢字として学習辞書記録部43に記憶する。

- [0078] ペア学習辞書記録部44は、ユーザにより過去に使用された漢字を、当該漢字がどの漢字に続いて確定されたのかを示す情報(直前漢字)と関連付けて、ペア学習漢字として記録するものである。直前漢字は、すでに学習辞書記録部43において単独学習漢字として記録されている。ペア学習辞書記録部44に記録される学習辞書には、そのペア学習漢字ごとに、上記直前漢字が関連付けて記録される。この漢字のペアごとに、記録された日時の情報や、使用頻度の情報を対応付けて記録してもよい。
- [0079] 学習制御部34は、文字確定部33より入力確定文字のペア学習処理の指示を受け付ける。学習制御部34は、一時記憶部3から、上記入力確定文字と、その直前漢字とを内部バッファに読み出し、一時記憶部3から読み出された文字列の削除および内容の更新を行う。そして、内部バッファに記憶した入力確定文字を直前漢字と関連付けて、ペア学習漢字としてペア学習辞書記録部44に記憶する。
- [0080] なお、上記では、学習辞書記録部43およびペア学習辞書記録部44において、各学習辞書が漢字ごとに記録される構成について説明したが、これに限定されない。上記構成に加えて、単語を辞書に記録する構成としてもよい。すなわち、学習制御部34の単独学習機能により、使用された単語が、学習辞書記録部43に記録されるようにもよい。例えば、「給水」という単語が使用された場合、「給」という漢字に加え、「給水」という単語が、単独学習漢字(単語)として学習辞書記録部43に記録される。
- [0081] また、学習制御部34のペア学習機能によれば、「大学」「教授」という単語が続けて入力された場合、漢字「教」を、「学」および／または「大学」の直前漢字(単語)と関連付けて、ペア学習漢字として記録することもできる。さらに、単語「教授」を、「学」および／または「大学」の直前漢字(単語)と関連付けて、ペア学習単語として記録することもできる。
- [0082] (候補順序決定要素)  
次に、文字変換装置1の候補順序決定部32(図12、16、21)が候補(候補漢字／

候補単語)の表示順序を決定するための所定の条件(以下、候補順序決定要素と称する)について説明する。

- [0083] 上述したとおり、操作パネル2に設けられた表示部12において、一度により多くの候補を表示できる方が、ユーザがより早く所望の漢字を得る可能性が高くなるというメリットがある。しかも、操作パネル2をタッチパネルなどの表示／入力一体型の入出力デバイスで実現した場合、ユーザが所望の漢字を得るための操作数を格段に減らすことが可能である。しかし、一方で、多くの候補を大きな表示画面に一覧表示しても、その中から所望の候補を探す負担が逆に大きくなるという問題が生じる。そこで、限られた表示領域に限られた数の候補を表示する際、ユーザが所望する(選択する)可能性の大きい漢字(単語)から順に表示されることが望まれる。
- [0084] 本発明では、ユーザの嗜好に合致した、より精度の高い表示順序で候補が表示されるよう、表示する候補の順序を決定する。これにより、所望の漢字を得るための入力操作を簡素化することが可能な文字変換装置を実現することができる。
- [0085] 以下では、候補単語の表示順序を決定する方法について説明するが、候補漢字／候補単語のいずれの表示順序も同じ方法にて決定される。
- [0086] 本実施形態では、候補順序決定部32は、以下の7つの候補順序決定要素に基づいて、候補の表示順序を決定する。
- [0087] (要素A)出現頻度  
出現頻度とは、ある言語の文字(列)において、どの程度出現するのかの指標となる情報である。出現頻度が高ければ高いほど、その言語使用環境下においてよく使用されているということになるので、出現頻度の高い候補ほど表示順序を上位にする。出現頻度は、既知の統計データを用いた固定的な情報を用いればよい。
- [0088] (要素B)ピンイン長の一致／不一致  
入力されたアルファベット(列)の文字数と、対応する候補のピンインの文字数が一致するか否かに基づいて表示順序を決定する。ピンイン長が、入力されたピンイン文字数と一致する候補の表示順序を上位にする。
- [0089] (要素C)直前文字列との連続性  
その時点までにすでに入力が確定している入力確定文字および／または現時点で

選択中の文字(列)の直後に続く可能性(直前文字列との連続性)の大きい候補ほど表示順序を上位にする。あるいは、過去の入力履歴より、上記直前文字列に連続して入力されたことの有無を判定して、連続して入力された履歴が残っている候補の表示順序を上位にする。

[0090] (要素D) 前方文字列との関連性

その時点までにすでに入力が確定している入力確定文字(列)に関連し、当該文字(列)の後方に入力される可能性(前方文字列との関連性)の大きい候補ほど表示順序を上位にする。

[0091] (要素E) 単独学習漢字／単語

学習辞書記録部43に記録されている単独学習漢字(単語)であるか否かによって、表示順序を決定する。過去に入力(使用)されたことのある漢字(単語)であれば、再び、使用される可能性は大きい。したがって、単独学習漢字(単語)として記録されている候補の表示順序を上位にする。

[0092] (要素F) ペア学習漢字／単語

ペア学習辞書記録部44に記録されているペア学習漢字(単語)であるか否かによって、表示順序を決定する。過去に入力(使用)されたことのある漢字(単語)の組み合わせであれば、再び、その組み合わせで使用される可能性は大きい。したがって、ペア学習漢字(単語)として記録されている候補の表示順序を上位にする。

[0093] (要素G) 学習時期

候補漢字(単語)に一致する学習漢字(単語)が、学習辞書記録部43および／またはペア学習辞書記録部44に記録された時期に基づいて表示順序を決定する。記録された時期が新しいものほど、ユーザにより選択される可能性が大きいので、最近記録された漢字(単語)の候補ほど表示順序を上位にする。

[0094] 以上の各要素を条件にすることにより、ユーザに選択される可能性の大きい候補ほど、優先して候補文字表示領域122(図14)に表示させることができるので、ユーザは、少ないキー操作数で、所望の漢字を得ることが可能となる。

[0095] 上述したとおり、候補順序決定要素とは、候補順序決定部32が、候補漢字／候補単語の表示順序を決定するために必要な情報である。より具体的には、候補順序決

定部32が、候補漢字／候補単語がユーザに選択される可能性の大きさを示す評価値を算出するために使用する情報である。候補(候補漢字／候補単語)ごとに候補順序決定要素の情報が対応づけられて記録部40の各(記録)部に記憶されており、候補順序決定部32は、記録部40の各部を参照して、候補ごとに候補順序決定要素を特定することが可能となる。さらに、記録部40には、候補順序決定部32が特定した候補順序決定要素からどのようにして上記評価値を算出するのかを規定した評価値算出ルールが記憶されていてもよい。これにより、候補順序決定部32は、特定した候補順序決定要素と評価値算出ルールとにしたがって、候補ごとに評価値を求め、候補の表示順序を決定することが可能となる。候補順序決定部32の評価値算出方法は、後述する。

[0096] (候補順序決定部の構成)

図22は、候補順序決定部32のさらに詳細な構成を示すブロック図である。候補順序決定部32は、学習漢字判定部131、ピンイン長判定部132、連結度判定部133、出現頻度判定部134、および、学習時期判定部135を備えている。記録部40(図12、16、21)は、上述の各記憶部に加えて、出現頻度記録部45および連結度記録部(関連度記録部)46を含んでいる。一時記憶部3は、内部に、抽出候補一時記憶部51、評価値一時記憶部52、および、表示順序一時記憶部53を含んでいる。

[0097] 一時記憶部3内の抽出候補一時記憶部51は、候補抽出部31が抽出した候補(以下、抽出候補と称する)としての漢字および／または単語を一時的に保持するものである。候補順序決定部32は、抽出候補一時記憶部51から抽出候補を読み出して、各抽出候補の表示順序を決定する処理を行う。

[0098] 評価値一時記憶部52は、候補順序決定要素と評価値算出ルールに基づいて算出された、上記各抽出候補に付与される評価値を一時的に保持するものである。表示順序は、最終的に算出された評価値を基に決定されてもよい。

[0099] 表示順序一時記憶部53は、上記各抽出候補に対して決定された表示順序の情報を一時的に保持するものである。候補順序決定部32は、決定した表示順序を表示順序一時記憶部53に格納する。表示順序一時記憶部53の表示順序の情報は、抽出候補を表示部12に表示する際に表示制御部22によって参照される。

- [0100] 候補順序決定部32の学習漢字判定部131は、抽出候補が、学習漢字(単語)であるか否か、および、それが単独学習であるかペア学習であるかを判定するものである。具体的には、抽出された候補と一致する漢字(単語)が学習辞書記録部43内に記録されている場合は、単独学習漢字(単語)であると判定し、ペア学習辞書記録部44内に記録されている場合は、ペア漢字(単語)であると判定する。学習漢字判定部131は、判定結果を評価値一時記憶部52の評価値または表示順序一時記憶部53の表示順序に反映させる。これにより、要素Eおよび／または要素Fに基づく表示順序を決定することができる。
- [0101] ピンイン長判定部132は、入力受付部21(図12、16、21)に入力されたアルファベット(列)の文字数と、対応する候補のピンインの文字数が一致するか否かを判定するものである。例えば、漢字「外字2」(ピンイン:mei)を得るために、「me」まで入力していたとする。この時点では、入力文字数は2で、対応する候補漢字「外字2」(mei)のピンイン長は3であるため、ピンイン長判定部132は、ピンイン長不一致と判定する。一方、ユーザによって「mei」まで入力された時点では、ピンイン長が一致すると判定し、候補漢字「外字2」の表示順序を上がるよう、評価値一時記憶部52の評価値または表示順序一時記憶部53の表示順序を変更する。これにより、要素Bに基づく表示順序を決定することができる。
- [0102] 連結度判定部133は、その時点までにすでに入力が確定している入力確定文字および／または現時点で選択中の文字(漢字／単語)とのつながり具合、つまり、連結度に応じて次にユーザが所望する文字を予測して候補に挙げる判定を行うものである。連結度判定部133は、連結度記録部46に記録されている連結度テーブル150を参照して、確定あるいは選択された前方文字列および／または直前文字列との連結度を判定する。前方文字列および／または直前文字列との連結度が高いものほど、ユーザが所望する文字である確率が高いとして、その表示順序を上位にする。これにより、要素Cおよび／または要素Dに基づく表示順序を決定することができる。連結度記録部46の連結度テーブル150および連結度判定部133の詳細については後述する。
- [0103] 出現頻度判定部134は、出現頻度記録部45に記録される、既知の統計データに

基づいて、候補漢字(単語)が出現する(使用される)頻度を判定するものである。出現頻度が高いものほど、その文字がよく使用されていることになるので、ユーザにより選択される可能性も大きい。したがって、出現頻度の高い候補ほど表示順序を上位にする。これにより、要素Aに基づく表示順序を決定することができる。

- [0104] 学習時期判定部135は、候補漢字(単語)に一致する学習漢字(単語)が、学習辞書記録部43および／またはペア学習辞書記録部44に記録された時期を判定するものである。学習漢字(単語)が記録された時期が新しいものほど、ユーザにより選択される可能性も大きい。したがって、学習時期が最近の漢字(単語)ほど上位になるよう候補の表示順序を決定する。これにより、要素Gに基づく表示順序を決定することができる。
- [0105] 上記構成によれば、候補順序決定部32は、上述のさまざまな条件に基づいて、候補の表示順序を決定することができる。
- [0106] これにより、ユーザの希望をより正確に反映した表示順序で候補を表示することが可能となり、結果として、表音文字の入力に係る操作数について、より少ない操作数で、所望の表意文字の入力を得ることが可能となる。
- [0107] (連結度の判定)
- 次に、連結度判定部133の動作について、さらに詳細に説明する。
- [0108] 図23は、連結度記録部46に記録される、漢字(単語)ごとの直前(前方)文字列との連結度を示した連結度テーブル150の例を示す図である。図23に示すとおり、連結度テーブル150は、漢字(単語)ごとに格納されている直前文字列との連続性情報のアドレスおよび前方文字列との関連性情報のアドレスを検出するための連結度対応表151と、漢字(単語)ごとに直前文字列との連続性を格納する連続性テーブル152と、漢字(単語)ごとに前方文字列との関連性を格納する前方関連性テーブル153とを含んでいる。
- [0109] 例えば、「交通」という単語がユーザにより選択されると、連結度判定部133は、直前文字列としての「交通」のあとに連続する確率が高い漢字(単語)を連結度テーブル150から検索する。連結度判定部133は、連結度対応表151を参照し、直前文字列「交通」との連続性情報が、連続性テーブル152のアドレス「npi+1」に格納されて

いることを検知する。

- [0110] 次に、連結度判定部133は、連續性テーブル152のアドレス「npi+1」を参照し、「規則(連續性21)」、「大学(16)」、「工具(15)」…の順に、連續性が高いことを検出する。最後に、上述の各候補単語の表示順序を、「規則」「大学」「工具」…の順にすることを決定する。
- [0111] また、単語「上海」の入力が確定されたのちに、「交通」が選択されると、連結度判定部133は、前方文字列としての「上海」の後ろに関連して入力される確率が高い漢字(単語)を連結度テーブル150から検索する。連結度判定部133は、連結度対応表151を参照し、前方文字列「上海」との関連性情報が、前方関連性テーブル153のアドレス「ppi+s」に格納されていることを検知する。
- [0112] 次に、連結度判定部133は、前方関連性テーブル153のアドレス「ppi+s」を参照し、「大学(関連性10)」、「規則(5)」、「警察(3)」…の順に、関連性が高いことを検出する。最後に、上述の各候補単語の表示順序を、「大学」「規則」「警察」…の順にすることを決定する。
- [0113] これにより、ユーザがよく入力する単語の並びにしたがって、候補漢字(単語)の表示順序を決定することができるので、ユーザの嗜好を反映した表示順序でもって候補漢字・単語を表示させることができる。結果として、ユーザのキー操作回数を減らし、所望の漢字を得るためのユーザの入力操作を簡素化することが可能となる。
- [0114] (連結度の判定一変形例)  
表示順序の決定を、要素C(直前文字列との連續性)および要素D(前方文字列との関連性)の両方を用いて実行する場合、要素Dに基づく判定結果を優先させてもよい。この場合、上述の例で言えば、「交通」が選択された場合に、候補単語「規則」を常に優先して表示するのではなく、前方に「上海」が入力されている場合には、候補単語「大学」を優先して表示することができる。
- [0115] ユーザが「上海交通大学」を頻繁に入力している文字変換装置1において、「上海交通大学」を入力するときの文字変換装置1の動作について説明する。ユーザが、「上」(shang)の入力のために、ピンインの先頭文字「s」を入力すると、上述の手順で、「上」を候補漢字として表示する。ユーザによって「上」が選択されると、「上」で始ま

る単語を候補単語として表示する。

- [0116] ユーザによって、「上海」が選択されると、確定文字表示領域123(図14)上に単語「上海」が表示する。同時に、文字変換装置1の候補抽出部31は、入力確定文字「上海」に続いて入力される候補単語を抽出する。候補順序決定部32は、「上海」に続く確率の高い単語を上位に表示するために、図22の連結度記録部46を参照し、候補単語の表示順序を決定する。
- [0117] まず、連結度判定部133は、連結度対応表151(図23)に基づいて、連続性テーブル152のアドレス「npi+s」を参照し、単語候補を、単語「上海」に直接連続する確率(連結度)の高い順に並べる。すなわち、図23に示す例では、「大学(15)」「交通(13)」「警察(12)」「規則(8)」…の順に候補が求まる。続いて、連結度判定部133は、入力確定文字「上海」の前方に位置する単語との関連性に基づいて、「上海」のあとに続く候補単語の表示順序を並べ替える。しかし、ここでは、「上海」の前方に位置する漢字・単語がないため、前方文字列との関連性による表示順序の変更は実行されない。
- [0118] 上記手順で表示された候補単語の中から、ユーザによって「交通」が選択されると、文字変換装置1は、続いて、入力確定文字「交通」に直接連続する候補単語を上位に表示するための表示順序を決定する。
- [0119] 連結度判定部133は、連続性テーブル152のアドレス「npi+1」を参照し、連結度に基づいて「規則(21)」「大学(16)」「工具(15)」「警察(14)」…の順に候補を求める。次に、前方関連性テーブル153のアドレス「ppi+s」を参照して、「交通」の前方に位置する漢字・単語(ここでは、「上海」)との関連性(関連度)に基づいて、「交通」のあとに続く候補単語の表示順序を並べ替える。より具体的には、図23に示す例では、前方関連性テーブル153によれば、候補単語の「上海」との関連度は、順に、「大学(10)」「規則(5)」「警察(3)」「工具(2)」…となっている。
- [0120] そこで、連結度判定部133は、上記候補単語の関連度を優先し、関連度の順に表示順序を変更する。この場合、「上海交通」に続く、候補単語の表示順序は、「規則」「大学」「工具」「警察」…の順から、「大学」「規則」「警察」「工具」…の順に変更される。あるいは、連結度および関連度の合計に基づいて、表示順序を変

更してもよい。また、連結度および関連度の平均値、最大値に基づいて表示順序を変更してもよい。例えば、連結度と関連度との合計を評価値として算出することを規定した評価値算出ルールが、記録部40に記憶されており、候補順序決定部32は、上記評価値算出ルールにしたがって、連結度および関連度の合計を評価値として算出し、その評価値に基づいて表示順序を決定することができる。

[0121] 以上のことから、ユーザの嗜好に沿った、より精度の高い表示順序でもって候補漢字・単語を表示させることができるので、ユーザのキー操作回数を減らすことができる。結果として、所望の漢字を得るためのユーザの入力操作を簡素化することが可能となる。

[0122] (表示順序決定手順)

本実施形態では、上述の各候補順序決定要素を組み合わせて、抽出された候補の表示順序を決定する。以下にその一例について説明するが、本発明は、以下で説明する組み合わせに限定されず、上述の各要素を適宜組み合わせることにより表示順序を決定することが可能である。

[0123] 図24は、候補順序決定部32が表示順序を決定するときの、候補の分類手順を示す図である。候補順序決定部32は、まず、所定の条件(上記要素A～Gおよびその組み合わせ)によって、表示順序ごとにいくつかのグループに候補を分類し、そのグループに内において、最終的な表示順序を決定する。

[0124] 候補抽出部31によって抽出された候補単語Gは、まず、一時記憶部3の抽出候補一時記憶部51(図22)に格納される。格納された候補単語Gに対し、学習漢字判定部131(図22)は、学習単語であるか否かの判定を行う(S301)。また、ピンイン長判定部132は、候補単語Gが抽出された時点で入力されているピンイン長(アルファベットの文字数)と、候補単語Gの各ピンイン長との比較を行う(S302)。

[0125] ここで、候補単語Gのうち、学習単語であり、かつ、ピンイン長も一致すると判定された候補単語は、表示順序が最上位のグループAに分類される。学習単語であるが、ピンイン長は一致しないと判定された候補単語は、グループAに次いで上位のグループBに分類される。学習単語ではないが、ピンイン長は一致すると判定された候補単語は、グループB下位のグループCに分類される。学習単語でもなく、ピンイン長も

一致しないと判定された候補単語は、最下位のグループDに分類される。

- [0126] グループA、Bに分類された、学習単語である候補単語に対し、学習漢字判定部131は、さらに、単独学習単語であるかペア学習単語であるかの判定を行う(S303)。また、連結度判定部133は、直前文字列との連続性の有無についての判定を行う(S304)。
- [0127] ここで、グループAの候補単語のうち、ペア学習単語であって、直前文字列との連続性ありと判定された候補単語は、最上位のグループA1に分類される。ペア学習単語であるが、直前文字列との連続性なしと判定された候補単語は、上位のグループA2に分類される。ペア学習単語でないが、直前文字列との連続性ありと判定された候補単語は、下位のグループA3に分類される。ペア学習単語でなく、直前文字列との連続性もないと判定された候補単語は、最下位のグループA4に分類される。グループBの候補単語についても同様に分類される。
- [0128] 一方、グループC、Dに分類された候補単語に対しては、連結度判定部133が直前文字列との連続性の有無を判定して(S305)、上位のグループC(D)1と下位のグループC(D)2とに分類する。
- [0129] 最終的に、全候補単語(G)は、上位のグループから順に、グループA1、グループA2、グループA3、グループA4、グループB1、グループB2、グループB3、グループB4、グループC1、グループC2、グループD1、グループD2、の小グループに分類される。グループA、Bの各小グループは、さらに学習時期(要素G)によって順序が決定され(最近学習された漢字ほど上位)、グループC、Dの各小グループは、さらに出現頻度(要素A)によって順序が決定される(一般によく使われる単語ほど上位)。
- [0130] なお、上述の実施形態では、学習された候補漢字・単語を、所定の条件で(すなわち、新しく学習された順に)並べ替えると説明したが、候補として抽出された候補漢字の後に、過去に学習された候補単語として並べてもよい。あるいは、過去に学習された単語が候補漢字より優先して表示されるように並べてもよい。また、候補漢字を含む過去に使用された単語が複数ある場合は、文字数が少ない順に並べてもよい。あるいは、時間的な経過として、学習された順もしくはその逆順に並べてもよい。

[0131] (変形例3－四声の入力)

本実施形態では、四声の区別を無視してアルファベット(列)とピンインとを対応させる構成について説明したが、本発明の文字変換装置1は上記構成に限定されない。四声の区別を行って、アルファベット(列)とピンインとを対応させる構成も可能である。例えば、四声コード(例えば、PCでの入力では一声から四声を1から4、軽声を5で表現している)を含めて、図15に示すA／P対応表141を作成すればよい。より具体的には、図15に示すA／P対応表141およびP／C対応表142において、ピンイン「shang」は、漢字「上」と関連付けられていることがわかる。ここで、「上」の四声ありのピンインは「三声のshang」であるため、A／P対応表141の「cp\_si」に対応する「shang」を「shang3」と記述することで、四声ありのピンイン「shang3」と「上」とを関連付けることができる。

[0132] [実施形態2]

本発明の文字変換装置1は、さらに、上述の候補順序決定要素に優先度の情報を付与し、候補漢字(単語)の表示順序を決定するときの条件(評価値算出ルール)を、ユーザが任意に設定できるように構成してもよい。表示順序を決定するときの条件をユーザが任意に設定できるので、文字変換装置1を備えたマルチファンクションプリンタにおいて、ユーザの嗜好に合致した表示順序で候補を表示することが可能となる。結果として、所望の漢字を得るための入力操作を簡素化することができる。

[0133] (文字変換装置の構成)

本実施形態における文字変換装置1は、上述の実施形態1の構成に加えて、文字入力制御部30の内部に、候補順序決定要素設定部(候補順序決定要素設定手段)35を備えており、記録部40の内部に、さらに候補順序決定要素記録部47を含んでいる(図12)。

[0134] 本実施形態では、図22に示す評価値一時記憶部52に、候補ごとに算出した評価値を一時的に記憶し、その評価値に基づいて表示順序を決定する。すなわち、候補順序決定要素設定部35によって設定された、優先度の高い候補順序決定要素ほど、上記評価値の算出により大きな影響を与える。候補順序決定要素設定部35が設定した候補順序決定要素およびその優先度は、候補順序決定要素記録部47に記

録される。

- [0135] まず、評価値一時記憶部52に記憶される候補ごとの評価値のデータ構造について説明する。
- [0136] 図25は、候補ごとに算出した表示順序の評価値を記憶するデータ構造の一例を示す図である。本実施形態では、候補ごとの評価値を16ビットで表現する。すなわち、図25に示す16の領域に、ある候補漢字(単語)の特性(要素A～G)を示す「0」または「1」の値を記憶する。本実施形態では、ビット位置は、評価値の高低に関連があり、上位ビットになるにつれて、評価値の上位桁を表すようになっている。よって、優先度の高い候補順序決定要素に関する判定結果は、上位ビットに反映させる。すなわち、評価値算出ルールは、候補漢字(単語)の候補順序決定要素要素から「0」または「1」の値をどのように求めるかを規定するとともに、そのように求まった値を、評価値のどのビット位置に反映させるかを規定している。
- [0137] このようなデータ構造によれば、文字変換装置1が候補ごとの評価値を算出するときの、評価値に与える要素の重要度をビット位置によって調節することが可能となる。つまり、重要な要素ほど上位位置に割り当てているので、ビットのON「1」／OFF「0」操作を行うことのみで、候補の表示順序の評価値を容易に適切に数値化することができる。また、各要素に対して個別に変数を割り当てる、表示順序を決定するために確保すべき評価値のデータサイズが大きくなる。よって、各要素による判定結果を1ビットに割り当てていることにより、データサイズの削減を実現することができる。
- [0138] 候補順序決定要素記録部47は、候補の表示順序を決定するための候補順序決定要素の優先度情報を記録するためのものである。図26の(a)および図26の(b)は、候補順序決定要素記録部47に記録される優先度情報の一例を示す図である。
- [0139] 優先度情報は、候補順序決定要素ごとに、その候補順序決定要素に関する判定結果を、評価値のデータ構造のどのビット位置に反映させるかを示すビット位置情報を含んでいる。図26の(a)に示す例では、要素Bの判定結果を図25に示すビット位置12の領域に格納するよう設定されている。すなわち、ピンイン長が一致した場合は、ビット位置12の領域に「1」が、一致しなければ「0」が評価値として記憶される。高いビット位置を割り当てられている候補順序決定要素ほど、優先度が高く、候補の評価

値を定めるのに与える影響が大きい。

- [0140] なお、本実施形態では、評価値を16ビットで表現し、ビット位置0～12の領域を用いて評価値を算出しているが、本発明はこの例に限定されない。候補順序決定要素数を減らしたり、新たな要素の追加または各要素の評価値としてのレンジ幅の変更をしたりする場合は、それらが十分に納まるビット長にすることも可能である。例えば、図26の(a)および図26の(b)に示す例において、出現頻度を8ビットで表現しているが、出現頻度の差を明確にするために16ビット確保する、あるいは、出現頻度の差を重視しないために4ビットに縮小することなどが可能である。また、空きのビット位置13～15を用いてもよい。
- [0141] 候補順序決定要素設定部35は、入力部11(図12)を介してユーザが指定する候補順序決定要素の優先度を設定し、優先度情報を候補順序決定要素記録部47に格納するものである。
- [0142] より具体的には、例えば、ユーザが候補順序決定要素の優先度設定メニューを、操作パネル2を用いて文字変換装置1に指示すると、操作パネル2の表示部12は、候補順序決定要素の優先順位の指定を促す操作画面を表示する。例えば、「(各要素(A～G)を)優先したい順に指示してください。」などのメッセージとともに、優先度を設定できる候補順序決定要素一覧を表示すればよい。ユーザが、優先したい順に表示されている各要素を指示すると、決定するか否かを問うメッセージの表示を行う。例えば、「変更を決定する場合は、決定キーを押してください。」などのメッセージを表示すればよい。
- [0143] ユーザが操作パネル2の決定キーを押すと、候補順序決定要素設定部35は、入力された指示信号に応じて、低いビット位置から、ユーザが指定した優先度が低い候補順序決定要素の領域として順に割り当て、各候補順序決定要素の判定結果を格納するビット位置の情報を変更する。
- [0144] 例えば、ユーザが、図26の(a)に示した例の状態から、要素Fのペア学習漢字／単語を最も優先度の高い候補順序決定要素として、順に、要素E、要素F、要素B、要素C、要素D、要素A／Gを指定したとすると、優先度情報は、図26の(b)に示すとおりに変更される。図26の(b)に示す例では、要素Fの判定結果が、ビット位置12の領

域に格納されることになる。

- [0145] 上記構成によれば、候補順序決定部32の各部(図22)は、候補順序決定要素記録部47が指定する候補順序決定要素の優先度情報に基づいて、判定結果に係る評価値を指定された評価値一時記憶部52のビット位置に格納する。
- [0146] これにより、表示順序を決定するための評価値を、ユーザが優先させたい候補順序決定要素をもとに算出することが可能となるので、よりユーザの嗜好に合致した表示順序で候補を表示させることが可能となる。
- [0147] なお、この候補順序決定要素の優先度情報は、候補漢字と候補単語とで別に設定できる構成としてもよいし、候補漢字・候補単語の両方の表示順序に同様に適用する1つの優先度情報を設定できる構成としてもよい。候補漢字と候補単語とで別に設定できる構成とした場合、優先度の変更操作を、1回の優先度設定メニュー呼び出し操作によって、候補漢字、候補単語の順に候補順序決定要素の優先度変更操作を行えるようにしてもよいし、候補漢字用、候補単語用それぞれの優先度設定メニュー呼び出し操作および優先度変更操作を行えるようにしてもよい。
- [0148] また、設定終了キーを設け、設定終了キーを押した時点までの内容に関して、所望の優先度情報の変更を行うようにしてもよい。例えば、候補順序決定要素の優先度の変更途中で設定終了キーを押した場合には、優先度の変更内容を反映しないようにし、候補順序決定要素の変更だけを反映さえるなどすることが可能である。
- [0149] 次に、図22に示す候補順序決定部32の各部の判定結果に基づいて、図25に示す評価値を算出する方法についてより詳細に説明する。なお、以下で説明する例では、優先度は、図26の(a)に示すとおりに設定されているものとする。
- [0150] まず、候補抽出部31が抽出した候補は、一旦抽出候補一時記憶部51に格納される。評価値一時記憶部52において、格納された候補ごとに、その評価値を表す16ビットの領域(図25)が確保される。
- [0151] 次に、候補が学習漢字(単語)である場合には、学習時期判定部135が、そうでない場合には、出現頻度判定部134が、学習時期または出現頻度をそれぞれ算出し、評価値算出ルールに基づいて、算出結果をビット位置0～7(図25)に格納する。
- [0152] 出現頻度は、あらかじめ大量の新聞や雑誌などの様々な文書文字データについて

統計的に文字および単語ごとに求めた使用頻度であり、頻度が大きいほど一般的によく使われる文字・単語であるとする。本実施形態では、出現頻度の領域を8ビット確保しているので、出現頻度に基づいて、評価値算出ルールにしたがい、候補の評価値を0～255の間で定めることができる。

- [0153] 学習時期についても同様に、最近学習されたものほど評価値が高くなるよう0～255の間で定めることができる。学習時期とは、候補漢字(単語)がいつ学習されたかを表す。評価値算出ルールについて具体的に説明すれば以下のとおりである。例えば、学習時期をX、学習できる最大個数をNとし、学習できる漢字・単語の学習した時点をn、現時点の学習できる時点をmとすると任意の漢字・単語の学習時期Xは、以下の式で求まる。

$m - n \geq 0$  のとき

$$X = N + m - n - 1 \cdots \text{(式1-1)}$$

$m - n < 0$  のとき

$$X = n - m - 1 \cdots \text{(式1-2)}$$

ここで、学習時期Xにおいて、Xの値が小さいほど過去に学習されたものとする。

- [0154] なお、上記式のXの値域が、学習時期の評価値として確保されているビット長の範囲(本実施形態では8ビット)を越える場合は、Xの値をビット長の範囲への数値変換や、ビット長の変更を行えばよい。

- [0155] ビット位置8～12には、それぞれ、要素D、C、F、E、Bについての判定結果が格納される。学習漢字判定部131、ピンイン長判定部132、および、連結度判定部133は、候補の特性が条件を満足していると判定した場合には「1」、そうでない場合には「0」を指定されたビット位置にセットする。例えば、連結度判定部133は、候補がすでに入力されている前方文字列と関連性ありと判定した場合には、ビット位置8に「1」をセットする。

- [0156] すべての候補について表示順序の評価値を算出すると、候補順序決定部32は、候補を評価値一時記憶部52に記憶される評価値の大きい順に並べ替え、最終的な表示順序を表示順序一時記憶部53に格納する。最後に、表示制御部22(図12)は、表示順序一時記憶部53に記憶されている上記表示順序に基づいて、候補を表示

部12に表示する。

[0157] [実施形態3]

上述の各実施形態では、ピンインを構成するアルファベットキーの入力による中国語入力の説明を行った。しかし、本発明の文字変換装置1はこれに限定されない。操作パネル2の入力部11(図12)が、手書き文字の入力機能を備え、入出力制御部20が手書き文字を認識することにより、中国語を入力する文字変換装置1を構成することも可能である。

[0158] また、手書き文字を入力するための入力部(筆記文字入力部)11は、例えば、表示一体型タブレットとしての操作パネル2の表示領域に手書き文字入力のための枠を表示することにより実現することができる。

[0159] 文字変換装置1は、入出力制御部20内部に、上記文字入力枠にユーザが筆記した文字の筆記情報(座標データ)を認識するための文字認識部(文字認識手段)(図示せず)を備える。また、記録部40に、上記文字認識部で認識された入力文字パターンとマッチングするために用いる認識辞書(図示せず)をさらに含む。そして、表示制御部22が、入力文字の認識結果を出力するように構成されている。

[0160] ユーザは、操作パネル2の文字入力枠に所望の文字を筆記すると、操作パネル2は筆記された文字の座標情報60(図27)を検出し、入出力制御部20の内部バッファに記憶する。文字認識部は、内部バッファに記憶している座標情報と認識辞書の辞書パターンとマッチング(例えば、DPマッチング)を行い、入力パターンと辞書パターンの類似度(例えば距離)を求め、辞書パターンに対応する文字コードと類似度を関連付けて内部バッファに記憶する。文字認識部は、入力パターンと認識辞書の各パターンとのマッチングを行うと、内部バッファの認識結果を類似度の高い順に並べ替える。そして、認識結果を候補漢字(単語)として候補抽出部31に供給する。

[0161] 以下、手書き文字認識機能を備えた文字変換装置1の動作を具体的に説明する。

[0162] ユーザが、表示一体型タブレットの文字入力枠に「外字2」を筆記すると(図28の(a))、操作パネル2は「外字2」の座標データを検出し、内部バッファに記憶する(図27)。文字認識部は、内部バッファの「外字2」の座標データと認識辞書の各辞書パターンとDPマッチングを行い、距離を計算し、距離の小さい順に並べ替える(図29)。入

出力制御部20は、認識結果を表示制御部22に転送する。

- [0163] 表示制御部22は、認識結果に基づいて、認識文字候補となるn個の文字(ここでは、漢字)の表示位置を求め、表示位置の情報および認識文字候補n個を操作パネル2の表示部12に表示する(図28の(b))。操作パネル2は、表示された認識文字候補と表示位置を関連付けて内部バッファに記憶する。
- [0164] ユーザは、「外字2」が表示されていることを確認すると、操作パネル2上の「外字2」を指定する。操作パネル2は指示された位置を検出し、入出力制御部20の候補選択受付部23に位置情報を転送する。そして、候補選択受付部23で「外字2」が指示されたことを検出し、「外字2」から始まる単語(「外字2」天」...)が表示される(図28の(c))。このときの動作は、すでに説明しているので、ここでは繰り返さない。
- [0165] これにより、ユーザが所望の漢字に対応するピンインを把握していない場合でも、所望の漢字を入力することが可能となり、所望の漢字を得るための入力操作を簡素化することができる。
- [0166] なお、手書き文字を認識するための、文字変換装置1の上記文字認識部の機能について、マルチファンクションプリンタを例に挙げて説明したが、これに限定されない。表示一体型のタッチパネルを備えた機器であればいずれも、上記文字認識部を適用することが可能である。したがって、筆記文字を入力することにより所望の漢字・単語を入力することが可能な本発明の文字変換装置を、上記マルチファンクションプリンタ以外の機器においても実現することが可能である。
- [0167] [実施形態4]  
(繁体字の入力)  
上述の実施形態では、ピンイン表記から中国語(簡体字)を入力可能な文字変換装置について説明した。しかしながら、本発明の文字変換装置1は上記構成に限定されない。
- [0168] 例えば、本発明の文字変換装置1を、主に台湾で使用される繁体字を入力するための文字変換装置として実現することも可能である。
- [0169] 本実施形態では、以下、表音文字としての注音文字(注音の字母)の入力を受け付けて、表意文字としての繁体字を、候補漢字(単語)としてユーザに提示する文字変

換装置1について説明する。

- [0170] 図30は、本実施形態に係る文字変換装置1の要部構成を示すブロック図である。なお、図30の各構成要素に付された符号は、図12・図22の各構成要素に付された符号に対応しており、同じ符号は、同じ構成要素を示している。したがって、上述の各実施形態すでに説明した構成要素についての説明は繰り返さない。
- [0171] 図30に示す文字変換装置1において、図12・図22と異なる点は、入力部11が、注音文字を入力するための、各注音文字に対応付けられたキーを備えている点と、入力受付部21が、上記入力された注音文字を受け付けて候補抽出部31に供給するための注音文字入力受付部24を備えている点と、漢字辞書部41において、注音文字－注音、および、注音－繁体字の対応関係を記録した対応表が含まれている点である。
- [0172] 図31は、本実施形態における、繁体字入力のための注音文字入力処理時の操作画面の一例を示す図である。本実施形態における表示一体型の操作パネル2では、アルファベットを表示するほかに、注音文字を表示してもよい(文字入力領域121)。繁体字の入力では、アルファベットではなく注音文字を入力することが一般的である。図31に示すとおり、キーボードのキーに注音文字が割り当てられており、これにより、繁体字に対応した注音を入力することができる。
- [0173] 図32は、本実施形態における、漢字辞書部41(図30)に記録される、漢字辞書(繁体字)の例を示す図である。
- [0174] 図33は、本実施形態における、単語辞書部42に記録される、単語辞書の例を示す図である。
- [0175] 図34は、本実施形態における、連結度記録部46に記録される、連結度テーブル150aの例を示す図である。
- [0176] 上記構成によれば、中国語(簡体字)を入力する場合と同様に、所望の表意文字(台湾で主に用いられる繁体字)を得るための、表音文字(注音の字母)を入力する操作を簡素化することが可能な文字変換装置を実現することが可能となる。
- [0177] [学習機能について]  
(学習漢字・単語における表示順序決定方法)

学習漢字・単語の表示順序は、例えば以下のような計算によって、学習された順序(学習順位Yと称する)を算出することにより、決定することもできる。学習順位Yの算出は、図22に示す学習時期判定部135が実行すればよい。

$$Y = \text{学習番号} - \text{学習カウンター} - 1 \cdots \text{(式2)}$$

学習番号は、新たに学習される漢字に割り当てるべきユニークな番号を表しており、この番号が、学習カウンタとして学習漢字とともに記録されることにより、その漢字がいつ学習されたかを知ることができるものである。図21の学習辞書記録部43(または、ペア学習辞書記録部44)には、500個の学習漢字(単語)を記録することが可能であるとすると、学習番号は、0~499までを用意すればよい。

[0178] 上記式2の計算により、Yがマイナスになる場合は、学習番号の上限値(ここでは、500程度)を加算する。

[0179] さらに、学習時期判定部135は、候補の並び替えの手間を省くために、候補ごとに求まつた学習順位Yをまるめる処理を行ってもよい。例えば、以下の手順にしたがつて、学習順位をまるめることができる。

(イ)  $0 \leq Y \leq 9$  のとき、 $Y \leftarrow Y$ (そのまま)

(ロ)  $10 \leq Y \leq 29$  のとき、 $Y \leftarrow (Y - 10) / 2 + 10$

(ハ)  $30 \leq Y \leq 109$  のとき、 $Y \leftarrow (Y - 30) / 8 + 20$

(二)  $Y \geq 110$  のとき、 $Y \leftarrow (Y - 110) / 32 + 30$ 、さらに、 $Y > 39$  のとき、 $Y \leftarrow 39$

以上のように、上記式2を用いれば、学習漢字(単語)を学習された順、あるいは、学習された時期が新しい順に効率よく並べ替えることができるるので、学習時期に応じて候補としての表示順序を決定することが可能となる。

[0180] (学習カウンタ付与方法)

次に、図21に示す学習制御部34が制御する学習番号(初期値=0)および学習カウンタ(初期値=0)について、より具体的に説明する。学習辞書記録部43に500個の学習漢字を記録できるものとし、学習辞書記録部43における学習漢字の記録位置(アドレス)の初期値をID=0とする。

[0181] 学習番号=0のときに漢字「学」を学習すると、学習制御部34は、学習辞書記録部43に学習されている学習漢字のうち、現在の学習番号に等しい学習カウンタを持つ

学習漢字を検索する。検索された学習漢字のアドレスIDの漢字を削除し、新規の学習漢字をセットする。現在の学習番号に等しい学習カウンタを持つ学習漢字が無い場合は、IDの位置に学習漢字と学習カウンタを共に記録する。そして、学習番号に1を加算する。ここでは、最初の学習であるため、アドレス「ID=0」の位置に、学習漢字「学」を学習カウンタ=0とともに記録し、学習番号に1加算して、学習番号=1とする。

[0182] この動作を繰り返して、学習番号=499のときに「文」を学習すると、学習制御部34は、学習辞書記録部43のアドレス「k」の位置に、学習漢字「文」を、学習カウンタ499とともに記録し、学習番号=0に更新する(学習番号499に1を加算した結果、学習番号が500以上になる場合は、500を減算して学習番号を0に戻す)。

[0183] さらに、新しく漢字(例えば「大」)が学習されると、学習制御部34は、新たな学習漢字「大」を、学習カウンタ0とともに記録されている「学」に上書きする形で記録する。このとき「大」には、学習カウンタ0が関連付けて記録される。続いて、上記と同様学習番号=2に更新する。

[0184] (学習辞書のデータ構造)

次に、学習辞書記録部43(および、ペア学習辞書記録部44)に記録される学習単語のデータ構造について説明する。

[0185] 学習辞書記録部43に記録される学習単語には、その単語を構成する漢字の文字コードから生成されるグループ識別子(GId値)が割り当てられており、学習単語はGId値と関連付けて記録される。このGId値を用いることにより、学習辞書記録部43から目的の学習単語を検索するための処理効率が向上する。

[0186] GId値の生成方法について、具体例を挙げて説明する。なお、以下の説明では、漢字を一意に特定するための漢字コードおよびGId値をともに16ビットで表すものとする。本実施形態では、一例として、GId値は、当該単語の文字数を示す上位4ビットと、単語を構成する各漢字の漢字コードの演算により求まった数値を示す下位12ビットとで構成されるものとするが、本発明の文字変換装置1で利用されるGId値のデータ構造はこれに限定されない。

[0187] 例えば、「上海」という単語の入力が確定されたとする。このとき、学習制御部34(図

21)は、まず「上海」という単語のGId値を生成する。GId値を生成するために、単語「上海」を構成する各漢字の漢字コードの下位8ビットを、各漢字を表す数値(「上」=「009F(16進数)」、「海」=「00A3(16進数)」)として用いるものとする。例えば、GId値の下位12ビットを、上記漢字コード16ビットの下位8ビット同士の加算により求めるると、学習制御部34は、下位12ビットとして、 $9F + A3 = 「141(16進数)」$ を算出する。

[0188] 次に、「上海」は2文字であるので、この単語が2文字で構成されていることを示す「2(16進数)」を上位4ビットとして算出する。以上のようにして、単語「上海」のGId値を「2141(16進数)」として算出する。

[0189] 上述のようにして生成されたGId値は、どのような文字数からなる単語であっても(ただし、上述の例のように文字数を4ビットで表す場合、最大8文字までを単語として学習するものとする)、固定長(例えば、16ビット)に要約されることになる。したがって、固定長で構成されるGId値によって、目的の単語を絞り込んでから、検索処理を実行することができるため、学習単語の検索効率を向上させることが可能となる。

[0190] [操作パネルについて]

操作パネル2に表示される文字入力処理時の操作画面は、図14や図17で示す例に限定されない。操作パネル2に表示される操作画面の他の例について説明する。

[0191] 操作パネル2に表示される、ピンインを入力するためのアルファベットキーは、図35に示すように、PCのキーボードを模したキー配列としてもよいし、図36に示すように、入力文字をアルファベット順にm×nのキー配列に当てはめるようにしてもよい。

[0192] また、候補漢字を表示するための候補文字表示領域122を確定文字表示領域123に重畳させて表示してもよい(図37の(a)～(c))。

[0193] さらに、候補漢字を選択後、選択した漢字から始まる候補単語を表示するための第2の候補文字表示領域122'を、候補文字表示領域122および確定文字表示領域123に重畳させて表示してもよいし(図38の(a))、候補文字表示領域122と並べて表示してもよい(図38の(b))。

[0194] [実施形態5]

本発明の一実施形態について図1～図11に基づいて説明すると以下の通りである

。ここでは、漢字(変換文字、変換された後の文字)の読み(表音文字)を入力して漢字を含む単語を決定する装置とするが、これに限定されない。つまり、以下で説明する実施の形態は本発明を実施するための一例であって、本発明を実現する上で制限されるものではない。

- [0195] 図1は、本実施の形態に係る文字変換装置1の構成を示すブロック図である。文字変換装置1は、入力部11、辞書検索部310、表示制御部22、漢字辞書部(変換文字辞書)41、単語辞書部(単語辞書)42、一時記憶部3、制御部10、表示部12を備えている。なお、図1の各構成要素に付された符号は、図12・図16・図21・図22・図30の各構成要素に付された符号に対応しており、同じ符号は、同じ構成要素を示している。なお、図1には示されていない図12・図16・図21・図22・図30の各部を、図1の文字変換装置1が備えていてもよい。
- [0196] 入力部11は、キーボード等からなり、ユーザが所望の入力文字列の読み(表音文字)、後段で説明する選択した漢字、選択した単語等を、入力することができるよう構成されている。また、入力部11は、ユーザが文字変換装置1への各種指示を入力するように構成されている。なお、入力部11と表示部12とを兼用して、操作パネル2(タッチパネル)として設けられていてもよい。
- [0197] 表示制御部22は、入力部11から入力された読み及び読みを基に、辞書検索部310が検索した漢字、単語及びそれらの候補をユーザが選択可能に表示部12に表示する。表示部12は、表示制御部22の制御の下、各種データの表示を行う。
- [0198] 本実施形態では、候補抽出部31は、辞書検索部310を備えている。辞書検索部310は、記録部40に記憶されている各辞書部(漢字辞書部41および単語辞書部42)を検索し、候補となる漢字(または単語)を検出する。本実施形態では、辞書検索部310は、さらに、漢字取得部(先頭変換文字取得手段)311と単語取得部(単語取得手段)312とを備えている。
- [0199] 漢字取得部311は、入力受付部(受付手段)21が受け付けた読みの文字列を入力とし、漢字辞書部41を検索して所望の条件に一致した漢字(あるいは、以下で説明する漢字情報)を検索結果として一時記憶部3に記憶する。さらに単語取得部312は、漢字取得部311の検索結果及び入力受付部21が受け付けた読みの文字列を入

力とし、単語辞書部42を検索して所望の条件に一致した単語(あるいは、以下で説明する単語情報)を検索結果として一時記憶部3に記憶する。漢字取得および単語取得の詳細は後述する。

- [0200] 漢字辞書部41は、読みと漢字とを対応付けた漢字対応情報144を保持している。さらに、読みの先頭文字と該先頭文字を先頭とする漢字の、漢字対応情報144での格納先のアドレス情報を対応付けた漢字インデックス(漢字インデックステーブル)145を保持している。ここでは、漢字インデックス145は、漢字の読みの順に並べられているものとする。
- [0201] 単語辞書部42は、単語と該単語の読みとが対応付けられた単語対応情報162を保持している。さらに、単語の先頭となりうる先頭文字と、該先頭文字を先頭に持つ単語の、単語対応情報162での格納先のアドレス情報とが対応付けられた単語インデックス(単語インデックステーブル)161を保持している。単語対応情報162と単語インデックス161とは、単語の先頭文字毎にグループ化されている。
- [0202] また、漢字対応情報144は、漢字に対応付けて補助情報を格納している。また、単語対応情報162は、単語に対応付けて補助情報を格納している。これら、補助情報は、例えば、読みの文字数(例えば、上述の各実施形態におけるピンイン長(要素B))(検索補助情報)や出現頻度(上述の要素A)(表示補助情報)等が挙げられる。上記補助情報は、読みの文字数及び出現頻度を所定の値域、例えば、1から100の範囲に変換した数値とする。以下では、漢字情報とは、漢字とその漢字に対応付けられた情報(読み、補助情報)とを指すものとする。また、単語情報とは、単語とその単語に対応付けられた情報(読み、補助情報)とを指すものとする。しかし、補助情報は、もちろん、上記に限定されず、これ以外の情報であってもよい。例えば、上述の各実施形態で述べた候補順序決定要素(要素C～要素G)を補助情報として、候補(漢字／単語)に対応付けて記憶しておいてもよい。
- [0203] 制御部10は、文字変換装置1における各種構成の動作を統括的に制御する。例えば、入力部11から入力された文字列の文字数をカウントする。また、辞書検索部310、表示制御部22の制御を行う。また、入力受付部21を備えおり、ユーザが入力部11から入力した文字列を受け付けるようになっている。

- [0204] 次に、文字変換装置1の処理動作について、図2に示すフローチャートを用いて説明する。
- [0205] ユーザが入力部11から入力したい文字列の先頭文字の読みとして、その読みの先頭文字を含む1以上の読み文字列を入力部11から入力すると、入力受付部21は入力された読み文字列を制御部10に転送する。
- [0206] 制御部10は入力部11から転送された読み文字列の文字数をカウントし、読み文字列とその文字数を一時記憶部3に記憶する。そして、制御部10は辞書検索部310に漢字検索を指示する(S101)。
- [0207] 次に、辞書検索部310の漢字取得部311は、一時記憶部3から読み文字列と読み文字列の文字数を辞書検索部310の内部バッファ(図示せず)に読み込む。漢字取得部311は、漢字インデックス145を参照し、読み文字列の先頭文字を基に、漢字対応情報144内での検索すべき検索対象範囲の終了位置のアドレス(iNum)を求めて内部バッファに保持する(S102)。さらに、漢字対応情報144内での検索すべき検索対象範囲の開始位置のアドレス(iStart)を求めて内部バッファに保持する(S103)。
- [0208] そして、漢字取得部311は、辞書検索部310の内部バッファに保持したアドレスを基に、漢字対応情報144を参照して漢字情報を読み込み、漢字対応情報144の読み情報と入力文字列の読みとを比較する(S104)。入力文字列の読みと、漢字対応情報144での読みが一致している場合は(S104でYES)、漢字取得部311は、読みが一致の条件を満足した漢字情報を一時記憶部3に記憶する(S105)。
- [0209] そして、iStartを更新する(S106)。その後、漢字取得部311は内部バッファに保持しているiNumとS106で更新されたiStartとを比較する(S107)。比較した結果、検索する漢字について、iStartがiNum以上になっていると(S107においてNO)、制御部10に漢字検索が終了したことを知らせる。
- [0210] なお、漢字検索の具体的な説明は、実施例を用いて後述する。
- [0211] 制御部10は、辞書検索部310から漢字検索が終了したことの通知を受けると、一時記憶部3に記憶されている漢字検索結果を所定の順に並べ替え(S108)、表示位置の情報を附加して一時記憶部3に記憶する。加えて、表示制御部22は、表示部1

2に漢字検索結果の表示を行う(S109)。なお、上記所定の順に並べ替える際、補助情報の出現頻度の順に並べ替えててもよいし、上述の各実施形態で説明したとおり、候補順序決定要素(および評価値算出ルール)に基づいて、候補順序決定部32が評価値を算出し、それに基づいて並べ替えててもよい。

- [0212] 次に、ユーザの指示入力があるかを確認する(S110)。ユーザは、表示部12に表示された漢字検索結果を目視して、所望の漢字を選択し、入力部11を用いてその漢字が表示されている位置を指示入力する。このようなユーザの指示入力があると(S110にてYES)、入力部11は指示位置の位置情報を制御部10の入力受付部21に送信する。
- [0213] 制御部10は、入力受付部21で受け付けた位置情報と、一時記憶部3に記憶している漢字検索結果の表示位置の情報を比較し、入力された漢字を検出する(S111)。さらに、制御部10は、ユーザにより指定された漢字を辞書検索部310に転送し、単語検索を行うことを指示する。なお、このとき、候補選択受付部23が、ユーザにより選択された漢字を検出し、学習制御手段が、それを学習漢字として、学習辞書記録部43またはペア学習辞書記録部44に記憶してもよい。
- [0214] 辞書検索部310の単語取得部312は、単語検索の指示を受けると、単語インデックス161を参照し、取得した漢字の識別情報(文字コード)を基に、単語対応情報162内での検索すべき検索対象範囲の終了位置のアドレス(iNum)を求めて内部バッファに保持する(S112)。さらに、単語対応情報162内で検索すべき検索対象範囲の開始位置のアドレス(iStart)を求めて内部バッファに保持する(S113)。
- [0215] そして、単語取得部312は、辞書検索部310の内部バッファに保持したアドレスを基に、単語対応情報162を参照して単語情報を読み込み、所定の条件に一致する単語情報を一時記憶部3に記憶する(S114)。
- [0216] その後、単語取得部312は、検索対象の単語についてのiStartを更新し(S114a)、内部バッファに保持しているiNumとS114aで更新したiStartとを比較する(S115)。比較した結果、検索対象の単語について、iStartがiNum以上になっていると(S115においてNO)、制御部10に単語検索が終了したことを知らせる。
- [0217] なお、単語検索の具体的な説明は実施例を用いて後述する。

- [0218] そして、制御部10は、辞書検索部310から単語検索が終了したことの通知を受けると、一時記憶部3に記憶されている単語検索結果を所定の順に並べ替え(S116)、表示位置の情報を付加して一時記憶部3に記憶する。加えて、表示制御部22は、表示部12にて表示部12に単語検索結果の表示を行う(S117)。なお、上記所定の順に並べ替える際、補助情報の出現頻度の順に並べ替えてもよいし、上述の各実施形態で説明したとおり、候補順序決定要素(および評価値算出ルール)に基づいて、候補順序決定部32が評価値を算出し、それに基づいて並べ替えてもよい。
- [0219] さらに、ユーザが表示部12に表示された単語検索結果から所望の単語を目視して、所望の単語を選択し、その単語が表示されている位置を指示入力する。指示入力があると、入力部11は入力された位置の位置情報を制御部10の入力受付部21に送信する。なお、このとき、候補選択受付部23が、ユーザにより選択された漢字を検出し、学習制御手段が、それを学習漢字として、学習辞書記録部43またはペア学習辞書記録部44に記憶してもよい。
- [0220] 制御部10は、入力受付部21で受け付けた位置情報と、一時記憶部3に記憶している単語検索結果の表示位置の情報と比較し、ユーザにより指定された単語を検索する。そして表示制御部22は表示部12にて検索された単語を表示する。
- [0221] つまり、本発明に係る文字変換装置では、漢字辞書部41及び単語辞書部42に、インデックスと漢字対応情報144及び単語対応情報162を持たせ、同一先頭漢字の単語をグループ化し、漢字対応情報及び単語対応情報に補助情報を持たせ、補助情報を基に検索対象の漢字及び単語を制限することで、検索時間の短縮を行えるようにしている。さらに、検索された候補となる漢字(単語)の表示順序は、その候補の補助情報(候補順序決定要素)に基づいて、候補順序決定部32によって、ユーザによって選択される可能性が高いものほど優先して表示されるように決定される。したがって、ユーザが所望する順序に並べ替えられた候補の漢字(単語)を、より短い時間で表示することができる。結果として、ユーザの利便性を向上させることが可能となる。
- [0222] なお、以下の実施例5-1、5-2では、上記説明した文字変換装置1を、ピンイン(表音文字)を入力して漢字(変換文字)に変換する中国語変換装置として説明する。

この場合、以下で詳細に説明するが、漢字インデックス145、漢字対応情報144、単語インデックス161、単語対応情報162は、次のようにになっているものとする。

- [0223] 漢字インデックス145は、図3に示すようになっており、ピンインの先頭文字(先頭の表音文字)と、このピンインの先頭文字を漢字のピンイン(表音情報)の先頭として持つ漢字の、漢字対応情報144での漢字アドレス(格納先情報)とが対応付けられている。図3で、漢字インデックス145の左の記載(cidx\_m1、cidx\_m2等)は、漢字インデックス145内でのアドレスである。具体的には、cidx\_m1とは、「読みの先頭がmであり、1声」の漢字インデックス145内でのアドレスである。
- [0224] また、漢字対応情報144は、図4に示すように、漢字とそのピンインと、補助情報(図4では、検索条件および頻度情報)とが対応付けられている。図4の漢字対応情報144の左の記載(cp\_m1、cp\_m2等)は、漢字対応情報144内でのアドレスである。
- [0225] また、単語インデックス161は、図5に示す1次インデックス(単語1次インデックステーブル)161aおよび図6に示す2次インデックス(単語2次インデックステーブル)161bを有している。図5に示すように、1次インデックス161aは、漢字とそのコードと、その漢字を先頭として持つ単語の2次インデックス161bでのアドレスが対応付けられている。2次インデックス161bは、漢字と第2文字目の漢字の先頭ピンインと、その単語対応情報162でのアドレスが対応付けられている。図6の2次インデックス161bの左の記載(widx\_di、widx\_gi等)は、2次インデックス161b内でのアドレスである。
- [0226] さらに、単語対応情報162は、図7に示すように、単語とその読みのピンインが対応付けられている。図7の単語対応情報162の左の記載(wp\_di、wp\_gi等)は、単語対応情報162内でのアドレスである。なお、漢字対応情報144にて、漢字コードとピンインとが対応している場合には、1次インデックス161aは必ずしも必要ではない。
- [0227] (実施例5-1)
- 本実施例では、上記説明した文字変換装置1を、ピンインを入力して漢字に変換する中国語変換装置として説明するが、これに限定されず、例えばひらがな(あるいはカタカナ)を入力して漢字を含む単語を決定する装置であってもよい。
- [0228] 具体的に、ユーザが「外字1」を入力するために、「外字2」の中国読みであるピンインの「mei」を入力した場合について説明する。

[0229] 初めに、漢字検索の処理について説明する。なお、入力読みと漢字情報の読みとが一致する条件は、入力読みが漢字情報の読みに前方一致であることとする。

[0230] (漢字検索)

辞書検索部310の漢字取得部311は、一時記憶部3から読み文字列「mei」を辞書検索部310の内部バッファに読み込む。そして、辞書検索部310は読み文字の先頭文字「m」を基に以下の式で漢字インデックス145の「m」の範囲のアドレスを求め、それらのアドレスに記述している漢字アドレスを内部バッファに記憶する。

[0231] アドレスは、

(1)入力読みに四声コードがある場合

先頭アドレス=(入力読みの文字コード-ベースの読みの文字コード)

×漢字インデックスの単位サイズ×5

+ (入力読みの四声コード-1) ×漢字インデックスの単位サイズ

終了アドレス=(入力読みの文字コード-ベースの読みの文字コード)

×漢字インデックスの単位サイズ×5

+ 入力読みの四声コード×漢字インデックスの単位サイズ

(2)入力読みに四声コードが無い場合

先頭アドレス=(入力読みの文字コード-ベースの読みの文字コード)

×漢字インデックスの単位サイズ×5

終了アドレス=(入力読みの文字コード-ベースの読みの文字コード+1)

×漢字インデックスの単位サイズ×5

として決定される。

[0232] 上記アドレスを決定する、「5」を掛けているのは、次の理由からである。中国語の発音は、ピンインや注音のように読みを表す部分と発声の上げ下げ(調子)を表す声調(四声)の組み合せである。声調が5種類(1~4声と軽声)があるので、5としている。

[0233] ピンインによる中国語入力では、ピンインはアルファベットで表現されるので、ベースの読みの文字コードは「a(=0x61(アスキーコード))」とし、注音による中国語入力では、ベースの読みの文字コードは「

[0234] [表5]

ウ

[0235] (=0xa373)」とすればよい。

[0236] 漢字インデックス(漢字インデックテーブル)41の単位サイズは、単語インデックス(単語インデックテーブル)51をピンインの先頭文字のアルファベット順とし、同じピンインの先頭文字は四声の1から5の順に並ぶように構成していれば、漢字アドレスのバイト長となる。例えば、本実施として動作している装置では4バイトとなっている。

[0237] ここでは、四声コードはないので、「m(=0x6d)」の範囲のアドレスは(2)式により、先頭アドレスが「(0x6d-0x61)×漢字インデックスの単位サイズ×5」(つまり、図3のcidx\_m1)、終了アドレスが「(0x6d-0x61+1)×漢字インデックスの単位サイズ×5」(つまり、図3のcidx\_n1)によって計算され、各アドレスが指示する漢字インデックスに記述している漢字アドレス(cp\_m1、及びcp\_n1)を内部バッファに記憶する。図3のcidx\_m1は、読みの先頭がmで1声のもの、の漢字インデックス145内のアドレスを表している。それに対応するcp\_m1は、読みの先頭がmで1声の読みの漢字対応情報144のアドレスを表している。cidx\_m5、cp\_m5は読みの先頭がmで軽声に対応している。

[0238] なお、 $(0x6d - 0x61) = 12$ であり、 $(0x6d - 0x61 + 1) = 13$ となる。

[0239] 辞書検索部310は、検索対象のアドレスcpiをcp\_m1とする。次に辞書検索部310は、漢字対応情報144からアドレスcpiの漢字情報を読み込み、入力ピンイン数を基に検索条件を判定する。その結果、検索条件を満足していれば、入力読みと漢字情報の読みとの比較を行い、入力読みが漢字情報の読みに前方一致であれば一致したものとして一時記憶部3に該漢字情報を記憶し、cpiのみ更新する。

[0240] 図4に示す漢字対応情報144の例では、検索条件を以下のようにしている。

検索条件は、

0:無条件に検索対象とする

1:入力ピンイン数が1文字であれば、検索対象としない

2:入力ピンイン数が2文字以下であれば、検索対象としない

3:入力ピンイン数が3文字以下であれば、検索対象としない

とする。

- [0241] 入力ピンインが「mei」の場合、入力ピンイン数が3であるため、検索条件が0、1、2に対応している漢字が、検索処理の対象となる。
- [0242] 上記処理は、cpiがcp\_n1に等しくなるまで処理を継続し、cpiがcp\_n1と等しくなれば、一時記憶部3に記憶している漢字情報を補助情報のひとつである出現頻度で並べ換えを行い、漢字検索処理を終了する。
- [0243] なお、上記説明では、四声コードの入力が行われない場合で説明したが、四声コード付きで「mei3」が入力された場合は、「m」の第3声のアドレスが上記(1)式により計算され、先頭アドレスが「(0x6d-0x61)×漢字インデックスの単位サイズ×5+(3-1)×漢字インデックスの単位サイズ」(つまり、図3のcidx\_m3)、終了アドレスが「(0x6d-0x61)×漢字インデックスの単位サイズ×5+3×漢字インデックの単位サイズ」(つまり、図3のcidx\_m4)となり、各アドレスが指示示す漢字インデックスに記述している漢字アドレス(cp\_m3、及びcp\_m4)を内部バッファに記憶する。
- [0244] なお、漢字アドレスを求めた後の処理は上記検索処理と同様であるため説明を省略する。
- [0245] (単語検索)
- 以下に単語検索の説明を行う。具体的に、ユーザが「外字1」の中国読みであるピンインの「mei」を入力し、漢字検索結果が一時記憶部3に保持され、表示制御部22表示装置に表示した後に漢字検索結果から「外字2」を指示入力したものとして説明を行う。
- [0246] 単語辞書部42の1次インデックス161aは、図5に示すように、漢字コードの昇順に記憶されているものとする。
- [0247] 制御部10は、ユーザの指示位置(X1、Y1)と一時記憶部3に記憶している漢字検索結果の表示位置(xi\_min、yi\_min、xi\_max、yi\_max)(i=1~n、nは候補数)を比較し、 $xi_{min} \leq X1 \leq xi_{max}$ かつ $yi_{min} \leq Y1 \leq yi_{max}$ が成立するときに、i番目の候補が選択されたものと判断し、対応する漢字を辞書検索部310に転送し、単語検索を行うことを指示する。
- [0248] 辞書検索部310は、単語検索の指示を受けると、転送された漢字のコードに一致する単語インデックス161を検索し、単語対応情報162での検索対象の範囲を求め

る。具体的には、辞書検索部310は、転送された「外字2」について、1次インデックス161aを2分探索し、「外字2」と一致する漢字コードC3BFを取得する。そして、漢字コードC3BFに対応する2次インデックス161b内のアドレス「wdx\_mi」を用いて、2次インデックス161bを参照し、「外字2」が先頭の単語の単語対応情報162におけるアドレス(単語アドレス)「wp\_mi」を、「外字2」が先頭の単語対応情報162での開始アドレスとして内部バッファに記憶する。さらに、辞書検索部310は、1次インデックス161aにおいてC3BFの次の漢字コードC3C0(図5中の

[0249] [表6]

## 冒

[0250] (以下では「外字6」とする))の2次インデックスのアドレス「wdx\_mi+1」を用いて、2次インデックス161bを参照し、「外字6」が先頭の単語の単語対応情報162におけるアドレス(単語アドレス)「wp\_mi+1」を、「外字2」が先頭の単語対応情報162での終了アドレスとして内部バッファに記憶する。

[0251] その後、辞書検索部310の単語取得部312は、「wp\_mi」から「wp\_mi+1」までの単語情報を読み出し、単語情報の漢字列を検索結果として、一時記憶部3に記憶する。そして、候補順序決定部32が、一時記憶部3の単語情報(候補単語)の補助情報(図示せず)である出現頻度の高い順に並べ替えを行い、単語検索処理を終了する。なお、この並べ替えは単なる例示であり、他の方法で並べ替えてもよい。例えば、候補順序決定要素(要素A～要素G)に基づいて、表示順序を決定してもよい。最後に、上記したように、検索され並べ替えられた単語を、表示部12に表示する。

[0252] (実施例5-2)

本実施例5-2では、ユーザが単語を入力するために、単語の読みを入力した場合の動作を説明する。ここで、実施例5-1では、入力したい文字列の先頭文字の読みを入力するのに対して、本実施例5-2では、入力したい単語の読みを入力するとの違いがある。しかし、入力部11から入力された読み文字列を制御部10の入力受付部21が受け付ける処理は実施例5-1と同様となる。そのため、本実施例ではユーザが「外字1」を入力するために「外字1」の中国読みであるピンインの「meinian」を入力した後の処理を説明する。

- [0253] なお、制御部10は、入力読みの文字列のピンインを分割する機能を備えているものとする。入力読み文字列のピンインの分割処理は後述する。
- [0254] 制御部10は、入力部11から読み文字列を「meinian」を受けると、読み文字列をピンインに分割し、分割結果として、「mei」と「nian」とでピンインに分割されたものが2つあることを示す分割数(あるいは、分割位置が2つあることを示す分割位置数)2を一時記憶部3に記憶する。
- [0255] 辞書検索部310の漢字取得部311は、一時記憶部3から読み文字列「mei」を読み込み、辞書検索部310の内部バッファに記憶する。漢字取得部311は、「mei」の先頭文字「m」を基に前記(2)式を用いて、漢字インデックス145を参照して、「m」の範囲のアドレスを求め、それらのアドレスに記述している漢字アドレスを内部バッファに記憶する。
- [0256] ここでも四声コードがないので、先頭アドレスが「(0x6d－0x61) × 漢字インデックスの単位サイズ × 5」(つまり、図3のcidx\_m1)、終了アドレスが「(0x6d－0x61 + 1) × 漢字インデックスの単位サイズ × 5」(つまり、図3のcidx\_n1)によって計算され、漢字辞書部41の漢字インデックス145を参照し、各アドレスが指し示す漢字アドレス(cp\_m1、及びcp\_n1)を内部バッファに記憶する。
- [0257] 辞書検索部310の漢字取得部311は、検索対象のアドレスcpiをcp\_m1とする。次に漢字取得部311は、アドレスcpiの漢字情報を読み込み、漢字対応情報144を参照し、入力ピンイン数を元に検索条件を判定する。その結果、検索条件を満足していれば、入力読みと漢字情報の読みとを比較する。
- [0258] 入力読みが漢字情報の読みと完全一致であれば、一時記憶部3に該漢字情報と完全一致であること(例えば、1)を対応付けて記憶して、cpiを更新する。
- [0259] 入力読みが漢字情報の読みに前方一致であれば一致したものとして一時記憶部3に該漢字情報と前方一致であること(例えば、0)を対応付けて記憶して、cpiを更新する。
- [0260] 完全一致あるいは前方一致でなければ、cpiのみ更新する。
- [0261] そして、cpiがcp\_n1と等しくなれば、候補順序決定部32が、一時記憶部3に記憶している漢字情報のうち、完全一致の漢字情報を出現頻度で並べ換えを行い、さら

に、前方一致の漢字情報を完全一致の漢字情報より候補順位が下位となるように出現頻度で並べ替えを行い、漢字検索処理を終了する。候補順序決定部32は、さらに、他の補助情報(例えば、候補順序決定要素(要素A～要素G))に基づいて、表示順序を決定してもよい。

- [0262] 上記処理にて、読み文字列「mei」に対して、「mei」に完全一致の漢字情報として「外字2」の漢字情報が検索結果として一時記憶部3に記憶されている。
- [0263] 次に単語検索処理の説明を行う。
- [0264] 上記の漢字検索後に表示された漢字検索候補からユーザが「外字2」を指示したものとして説明する。なお、ユーザが「外字2」の指示に対する候補選択の処理は、実施例5-1に記載の処理と同様であるため説明を省略する。
- [0265] また、1次インデックス161a、2次インデックス161b、及び単語対応情報162は、漢字コードの昇順に並んでいるものとする。
- [0266] 辞書検索部310の単語取得部312は、単語検索の指示を受けると、単語インデックス161の1次インデックスを参照して、転送された漢字のコードに一致する第1のインデックス情報を取得する。例えば、1次インデックスでは、単語の先頭漢字の漢字コード順に並んでおり、単語取得部312は、2分探索により転送された漢字のコードと一致する第1のインデックス情報を検索して、転送された漢字コードに一致する第1のインデックス情報と、1次インデックス内で並べられている次の漢字コードの第1のインデックス情報を、内部バッファに記憶する。
- [0267] ここでは、ユーザは「外字2」を指示しているので、1次インデックス161aにおいて、漢字コードに一致する2次インデックスアドレスはwdx\_mi、次の漢字コードの2次インデックスアドレスはwdx\_mi+1であり、これらが内部バッファに記憶される。
- [0268] 辞書検索部310の単語取得部312は、一時記憶部3から分割した読み文字列の2番目以降の文字列を内部バッファに読み込む。辞書検索部310は、1番目の分割した読みで検索された漢字(先頭漢字、1文字目漢字)と、2番目の分割された読みの先頭ピンイン文字とで単語対応情報162の検索範囲を絞るために、1次インデックス161aに付加されている2次インデックス161bのアドレスを基に、2番目の分割された読みの先頭文字「n」と一致するものを、2次インデックス161bにて検索し、内部バッ

ファに検索された2次インデックス161bと、2次インデックス161b内で並べられている次の第2のインデックス情報を記憶する。ここでは、1文字目の漢字「外字2」と2番目の分割読みの先頭文字「n」とに一致する2次インデックス情報wp\_mi、および、2次インデックス内で並べられている次の2次インデックス情報wp\_mi+k、が内部バッファに記憶される。

- [0269] 辞書検索部310の単語取得部312は、検索対象のアドレスwpiを第2インデックス情報に付加されている単語情報のアドレスwp\_miとする。「wp\_mi」から「wp\_mi+k」までの単語情報を読み出しここで、単語対応情報162での単語情報の読みは、単語の各漢字とを対応付けるため、例えば、「外字1」の読みを「MeiNian」のように単語の各漢字に対応する読みの先頭部分を大文字表記しているものとする。
- [0270] 次に辞書検索部310の単語取得部312は、アドレスwpiの単語情報を読み込み、単語対応情報162を参照して、単語対応情報162の漢字の個数と入力読みの分割数を比較する。単語情報の漢字の個数が入力読みの分割数以下であれば、単語情報の各漢字に対応する読みと分割読みを先頭から比較し、読みが完全一致の漢字数と読みが前方一致の漢字数をカウントする。
- [0271] 「単語情報の漢字の個数」と「読みが完全一致の漢字数+読みが前方一致の漢字数」が一致すれば、一時記憶部3に単語情報と読みが完全一致の漢字数と読みが前方一致の漢字数を記憶し、wpiを更新する。「単語情報の漢字の個数」と「読みが完全一致の漢字数+読みが前方一致の漢字数」が一致しなければ、wpiを更新する。
- [0272] そして、 $wpi < wp\_mi + k$ であれば処理を継続し、 $wpi = wp\_mi + k$ となれば、処理を終了する。
- [0273] 辞書検索部310は、一時記憶部3に記憶している単語情報において、以下の条件で順位付けを行う。
- (i) 読みの分割数と「読みが完全一致の漢字数+読みが前方一致の漢字数」とが等しい単語情報
- ・「読みが完全一致の漢字数」が多いほど上位候補とする
  - ・「読みが完全一致の漢字数」が同じものは、出現頻度で順位付ける
- (ii) 読みの分割数と「読みが完全一致の漢字数+読みが前方一致の漢字数」とが

## 等しくない単語情報

- ・「読みが完全一致の漢字数」が多いほど上位候補とする
- ・「読みが完全一致の漢字数」が同じものは、出現頻度で順位付ける

最終的にユーザが表示された検索結果から所望の文字列を選択することで、入力が行われるが、上記(i)の条件に対応する検索結果を選択すると、新たに読み文字列の入力を行うことになるが、(ii)の条件に対応する検索結果を選択した場合は、入力した読み文字列の一部に対応するため、選択した文字列に対応する読み文字列を除く読み文字列を制御部10の入力読みとして上記処理を繰り返すことができる。

- [0274] 上記説明では、単語検索で単語情報の漢字の個数と「読みが完全一致の漢字数 + 読みが前方一致の漢字数」が一致したときのみ単語検索結果としている。しかし、単語情報の漢字の個数が「読みが完全一致の漢字数 + 読みが前方一致の漢字数」より多い場合も単語検索結果としてもよい。このような構成にすることで、所望の単語の先頭から数文字分の読みだけで入力することが可能となる。
- [0275] また、上記説明では、表示された漢字検索結果から所望の漢字をユーザが選択することで単語検索を開始している。しかし、漢字検索が終了した後で、漢字検索結果を表示せずに候補順位1位の漢字について単語検索を開始し、単語検索が完了したときに検索結果を表示してもよい。あるいは、漢字検索が終了した後で、漢字検索結果を表示せずに候補の一部もしくは全部について単語検索を開始し、単語検索が完了したときに検索結果を表示してもよい。
- [0276] 次に、入力読みのピンイン(n文字)の分割処理について説明する。ここで分割の単位を音節と呼ぶ。
- [0277] 分割処理は、以下のようを行う。
- (a) 分割位置iを1にする
  - (b) 先頭文字からi文字目を第1音節とし、i+1文字目から韻母、声母、声母+韻母として成立する最長のピンインを求め、第2音節とする
  - (c) i文字目の分割位置の評価値(以下、位置評価値と称する)として、第1音節長 + 第2音節長を保持する
  - (d) 分割位置iに1を加算し(i=i+1)、i>nであれば(e)、そうでなければ(b)へ処理

を移す

(e)最大位置評価値を有するものとして処理された分割位置の個数により、最終的な分割位置を確定する

(e-1)最大位置評価値を有するものとして処理された分割位置の個数が1の場合

最大位置評価値を有するものとして処理された分割位置を最終的な分割位置とする

(e-2)最大位置評価値を有するものとして処理された分割位置が複数個ある場合

(1)1音節だけで構成されている

(2)第1音節、第2音節とも2文字以上

(3)第1音節の文字数が第2音節よりも多い

の順に、上位の条件を有する分割位置を最終的な分割位置とする

具体例として、「meinian」の分割処理について説明する。入力ピンインの一部或いは全部がピンインとして成立するか否かは、ピンイン辞書に登録されているピンインと比較し、一致するピンインがあればピンインとして成立する可能性があるとする。ピンイン辞書は、図9に示すように、ピンインとして成立する可能性のある長さ1から7のピンインが登録されている。

[0278] ここで、入力ピンインとは、ピンインとして入力された文字列であり、

(1)ピンインとして登録されている文字あるいは文字列そのもの

(2)ピンインとして登録されている文字あるいは文字列が組み合わされた文字列

(3)(1)(2)以外の文字列(間違い入力の場合)

の3種類がある。

[0279] なお、本実施形態5において、分割処理時に算出される位置評価値は、入力ピンインを2音節に区切るための区切り位置を決定する場合に、その位置で音節が区切られる尤もらしさを数値化したものである。この位置評価値は、音節区切り位置ごとに算出される。したがって、上述の実施形態1～4で用いられた、候補漢字／候補単語ごとに算出される評価値(候補漢字／候補単語がユーザに選択される可能性の大きさを示す評価値)とは区別される。

[0280] 入力ピンイン「meinian」の分割位置iを1とする。先頭文字「m」を用いて、図8で示

すピンイン辞書インデックス(ピンイン辞書インデックステーブル)を参照し、先頭文字が「m」で1文字のピンイン辞書アドレスpr\_m1及び2文字のピンイン辞書アドレスpr\_m2を読み込む。そして、ピンイン辞書のpr\_m1からpr\_m2までに記憶されているピンインと比較する。この場合、「m」は存在するので、ピンインとして成立する。

- [0281] 次に、「m」を除く「einian」を第2音節とし、先頭文字が「e」で6文字のピンイン辞書アドレスpr\_e6及び7文字のピンイン辞書アドレスpr\_e7を読み込み、pr\_e6からpr\_e7までに記憶されているピンインと比較する。一致するピンインがあれば、1文字目の位置評価値が7となるが、一致するピンインが無い場合は、第2音節の最終文字から1文字減らして新しい第2音節とし、同様にピンイン辞書と比較し、一致するピンインがあれば、そのときの第2音節の長さ+第1音節の長さ(ここでは1)を分割位置1の位置評価値とする。一致するピンインが無い場合には、さらに、第2音節の最終文字から1文字減らして、ピンイン辞書と比較するという処理を行い、これを、位置評価値が出るまで繰り返す。

- [0282] 上記処理が終わると、入力ピンイン「meinian」の分割位置iを*i=i+1 (=2)*とする。そこで、先頭文字から2文字「me」を元にピンイン辞書インデックステーブルから「m」が先頭の2文字のピンイン辞書アドレスpr\_m2からpr\_m3までに記憶されているピンインと比較する。「me」が存在すれば、第1音節を「me」とする。次に「me」を除く「inian」を第2音節とし、同様にピンイン辞書に一致するピンインがあるかを検索する。先頭から切り出し位置7までの処理を行い、最大位置評価値の分割位置の評価を行う。図10では、最大位置評価値が7であり、最大位置評価値が7となる分割位置の個数が1なので、最終分割位置を3と確定する。なお、図10では、分割位置が3のものについての、「分割位置数2」とは、「mainian」は、第3文字iと第4文字nの間に1つの目の分割位置があり、第7文字nとその後(文字はなし)の間に2つ目の分割位置があるとここでは判断されているので、よって、分割位置が2つあることを示している。

- [0283] 次に、分割位置までのピンインを除く「nian」について、上記と同様の処理を繰り返し、入力ピンインの分割位置を求める。「nian」では、図11に示すように、分割位置が求まり、位置評価値4の分割位置は2個があるので、1音節で構成されている「nian」が

選択され、分割位置4が求まる。

- [0284] 分割位置までのピンインを除くと、残りのピンインが無いので、入力ピンインの先頭位置からの位置に補正(分割位置3、7)し、処理を終了する。
- [0285] なお、以上では、ピンインに対する分割処理について説明したが、注音についても同様に注音テーブルを構成することで、上記処理にて分割を行うことができる。ピンイン辞書のインデックステーブルは、ピンインの最大長が7であるため、各ピンインの先頭文字について、長さ1から7までの7個単位で構成している。他方で、注音の最大長は、4であるため、各注音の先頭文字について、長さ1から4までの4個単位で構成すればよい。
- [0286] 以上で説明したように本願発明では、同一先頭漢字の単語をグループ化し、漢字辞書及び単語辞書に入力ピンインに応じて検索対象を制限するための情報を持っている。具体的には、漢字辞書は、ピンインの長さに応じた検索対象か否かの情報を持っている。単語辞書は、インデックス情報に2文字目漢字の読みの先頭文字を持っている、ピンイン列(漢字n文字分)を分割し、漢字検索された1文字目漢字と2文字目の先頭読みから単語検索対象を制限できる。このような構成であるため、検索条件に一致するものを検索するので、語彙数を増加しても、入力文字列と読みの情報の比較を行う回数が制限され、応答速度の速い文字列入力装置を提供することができる。読みの入力が入力したい漢字の2文字目以降の読みを入力すると、さらに単語の検索対象を絞ることができるので、検索処理の速い文字変換装置を提供することができる。
- [0287] さらに、上述の検索処理によって求まった漢字(単語)の候補は、候補ごとに対応付けられている補助情報(あるいは、候補順序決定要素)に基づいて、ユーザによって選択される可能性が高いものほど優先して表示されるような順序に並べ替えて表示するので、候補を表示するまでの応答時間を短縮できる上に、ユーザが所望する順序で候補を表示することができるので、ユーザの利便性をさらに向上させることが可能となる。
- [0288] 最後に、文字変換装置1の各ブロック、特に候補順序決定部32、辞書検索部310、表示制御部22、および制御部10は、ハードウェアロジックによって構成してもよい

し、次のようにCPUを用いてソフトウェアによって実現してもよい。具体的には、記録部40、漢字辞書部41、単語辞書部42、一時記憶部3はメモリ、ハードディスクなどの記憶装置によって実現され、候補順序決定部32、辞書検索部310、表示制御部22、制御部10はコンピュータの中央演算装置(CPU)及びCPUに実行されるソフトウェアにより実現される。このソフトウェアも記憶装置に格納され、メモリに読み出されてCPUにより実行される。

[0289] つまり、文字変換装置1は、各機能を実現する制御プログラムの命令を実行するCPU(central processing unit)、上記プログラムを格納したROM(read only memory)、上記プログラムを展開するRAM(random access memory)、上記プログラムおよび各種データを格納するメモリ等の記憶装置(記録媒体)などを備えている。そして、本発明の目的は、上述した機能を実現するソフトウェアである文字変換装置1の制御プログラムのプログラムコード(実行形式プログラム、中間コードプログラム、ソースプログラム)をコンピュータで読み取り可能に記録した記録媒体を、上記文字変換装置1に供給し、そのコンピュータ(またはCPUやMPU)が記録媒体に記録されているプログラムコードを読み出し実行することによっても、達成可能である。

[0290] 上記記録媒体としては、例えば、磁気テープやカセットテープ等のテープ系、フロッピー(登録商標)ディスク／ハードディスク等の磁気ディスクやCD-ROM／MO／MD／DVD／CD-R等の光ディスクを含むディスク系、ICカード(メモリカードを含む)／光カード等のカード系、あるいはマスクROM／EPROM／EEPROM／フラッシュROM等の半導体メモリ系などを用いることができる。

[0291] また、文字変換装置1を通信ネットワークと接続可能に構成し、上記プログラムコードを通信ネットワークを介して供給してもよい。この通信ネットワークとしては、特に限定されず、例えば、インターネット、イントラネット、エキストラネット、LAN、ISDN、VAN、CATV通信網、仮想専用網(virtual private network)、電話回線網、移動体通信網、衛星通信網等が利用可能である。また、通信ネットワークを構成する伝送媒体としては、特に限定されず、例えば、IEEE1394、USB、電力線搬送、ケーブルTV回線、電話線、ADSL回線等の有線でも、IrDAやリモコンのような赤外線、Bluetooth(登録商標)、802.11無線、HDR、携帯電話網、衛星回線、地上波デジタル網

等の無線でも利用可能である。なお、本発明は、上記プログラムコードが電子的な伝送で具現化された、搬送波に埋め込まれたコンピュータデータ信号の形態でも実現され得る。

[0292] 以上のように、本発明に係る文字変換装置は、入力された1以上の表音文字よりなる入力文字列を変換文字に変換し、当該変換文字からなる単語を決定できる文字変換装置において、

(a-1) 変換文字を識別するための変換文字識別情報と、該変換文字の表音を示す変換文字表音情報とが対応付けられた変換文字対応情報および、(a-2) 入力される表音文字と、該表音文字を上記変換文字表音情報の先頭として持つ上記変換文字識別情報の、上記変換文字対応情報での格納先を示す変換文字格納先情報とが対応付けられた変換文字インデックス、を有する変換文字辞書と、

(b-1) 単語を識別するための単語識別情報と、該単語の表音を示す単語表音情報とが対応付けられた単語対応情報および、(b-2) 単語の先頭となる先頭変換文字を識別するための先頭変換文字識別情報と、該先頭変換文字識別情報が示す先頭変換文字を先頭に持つ単語の、上記単語対応情報での格納先を示す単語格納先情報とが対応付けられた単語インデックス、を有する単語辞書と、

(c) 入力された上記入力文字列の先頭の表音文字を基に、上記変換文字辞書を参照し、該入力文字列の先頭の表音文字を上記変換文字表音情報の先頭として持つ変換文字識別情報を取得する先頭変換文字取得手段と、

(d) 取得された上記先頭変換文字識別情報を基に、上記単語辞書を参照し、該先頭変換文字識別情報が示す変換文字を先頭に持つ単語を取得する単語取得手段と、

を備えている。

[0293] また、本発明に係る文字変換装置は、上記構成に加え、上記先頭変換文字取得手段が取得した先頭変換文字識別情報が示す先頭変換文字を、ユーザが選択可能に表示部に表示する表示制御手段と、上記ユーザにより選択された先頭変換文字を受け付ける受付手段と、を備え、上記単語取得手段は、上記選択された先頭変換文字の先頭変換文字識別情報に基づいて、単語を取得する処理を行ってもよい。

- [0294] 上記構成によると、ユーザが選択した先頭変換文字に基づいて、つまり、先頭変換文字を絞って、単語を取得することができる。
- [0295] また、本発明に係る文字変換装置は、上記構成に加え、上記変換文字対応情報では、さらに、上記変換文字識別情報に、該変換文字識別情報が示す変換文字についての表示補助情報が対応付けられており、上記表示制御手段は、上記表示補助情報に基づき、先頭変換文字を並べ替えて、ユーザが選択可能に表示部に表示してもよい。
- [0296] 上記構成によると、表示補助情報に基づき、先頭変換文字を並べ替えて、ユーザが選択可能に表示部に表示することができる。表示補助情報とは、例えば、その漢字の出現頻度の情報であってもよい。表示補助情報はこれ以外の情報であってもよく、表示制御手段が、先頭変換文字を並べ替えることができる情報であればよい。
- [0297] また、本発明に係る文字変換装置は、上記構成に加え、上記単語辞書の単語インデックスは、単語の先頭となる先頭変換文字毎に、上記先頭変換文字識別情報と、単語の2文字目の先頭の表音文字である2文字目先頭表音文字との組み合わせに対して、上記単語対応情報における格納先が割り振られていてもよい。
- [0298] 上記構成によると、単語取得手段は、先頭変換文字識別情報と、単語の2文字目の先頭の表音文字である2文字目先頭表音文字との組み合わせから、単語を検索することができる。よって、検索範囲をより絞ることができ、より素早い応答が可能となる。
- [0299] また、本発明に係る文字変換装置は、上記構成に加え、上記変換文字対応情報では、さらに、上記変換文字識別情報に、該変換文字識別情報が示す変換文字についての取得補助情報が対応付けられており、上記先頭変換文字取得手段は、上記取得補助情報を基に上記変換文字識別情報を取得してもよい。
- [0300] 上記構成によると、先頭変換文字取得手段は、変換文字識別情報が示す変換文字についての取得補助情報を基に、変換文字識別情報を取得することができる。取得補助情報とは、先頭変換文字取得手段の変換文字辞書を用いた検索を補助する情報であればよく、例えば先頭変換文字の表音文字の長さ等である。
- [0301] また、本発明に係る文字変換装置は、上記構成に加え、上記取得補助情報は、以

以下の4つの情報に分類されていてもよい。

- (1) 変換文字識別情報が示す変換文字が、上記入力文字列に対して、無条件に取得対象となるという情報。
- (2) 変換文字識別情報が示す変換文字が、上記入力文字列が1文字では、取得対象外になるという情報。
- (3) 変換文字識別情報が示す変換文字が、上記入力文字列が2文字以下では、取得対象外になるという情報。
- (4) 変換文字識別情報が示す変換文字が、上記入力文字列が3文字以下では、取得対象外になるという情報。

[0302] 上記構成によると、先頭変換文字取得手段が、すばやく、効果的に先頭変換文字を取得することができる。

[0303] また、本発明に係る文字変換装置は、上記構成に加え、上記単語対応情報では、単語の先頭の変換文字が同一のものが1グループとしてまとめられていてもよい。

[0304] 上記構成によると、単語取得手段は、グループとしてまとめられた単語について検索できるので、すばやく、効果的に単語を取得することができる。

[0305] さらに、各上記グループ内で、単語は、単語の先頭の表音文字の所定の配列順に並べられていてもよい。

[0306] 上記構成によると、単語取得手段は、単語の先頭の表音文字の所定の配列順に並べられてグループ内で単語を検索できるので、よりすばやく、単語を取得することができる。ここで、単語の先頭の表音文字の所定の配列順とは、例えば、単語の読みがピンインであればアルファベット順、ひらがなであれば50音順、等である。

[0307] また、本発明に係る文字変換装置は、上記構成に加え、上記変換文字は漢字であり、上記変換文字対応情報では、漢字の先頭の表音文字が同一、かつ、声調コードが同一の、上記変換文字識別情報である漢字識別情報がグループとしてまとめられていてもよい。

[0308] 上記構成によると、先頭変換文字取得手段は、グループ内で求めたい先頭変換文字について検索すればよいので、応答速度が速くなる。

[0309] また、中国語変換装置として用いられる上記文字変換装置の、上記変換文字対応

情報では、各上記グループで、上記漢字識別情報が、漢字の先頭の表音文字の所定の配列順に並べてられていてよい。

- [0310] 漢字の先頭の表音文字の所定の配列順に並べてあるので、先頭変換文字取得手段は、より素早く先頭変換文字を取得することができる。
- [0311] また、中国語変換装置として用いられる上記いずれかの文字変換装置の、前記変換文字インデックスでは、上記漢字の先頭の表音文字が所定の配列順となっており、各上記グループの上記変換文字対応情報での格納先を示す情報を上記変換文字格納情報として、上記漢字の先頭の表音文字に対応付けられていてよい。
- [0312] 上記構成によると、変換文字インデックスでは、漢字の先頭の表音文字が所定の配列順となっており、各上記グループの変換文字対応情報での格納先を示す情報を上記変換文字格納情報として、漢字の先頭の表音文字に対応付けられている。よって、先頭変換文字取得手段は、グループの変換文字対応情報での格納先から、グループを絞って、より応答速度を速めて、検索することができる。
- [0313] ここで、中国語変換装置として用いられる上記何れかの文字変換装置では、上記漢字は簡体字であり、かつ、上記表音はピンインで表現されていてよい。
- [0314] 上記構成によると、ユーザが1文字以上のピンイン列を入力すると、ピンイン列の先頭文字から先頭漢字の候補を絞り込み、先頭漢字の候補を基に検索対象の単語を制限できるので、辞書検索処理にかかる時間を短縮できる。それゆえ、応答速度の速い中国語入力装置を提供できる。また、入力したい文字列の先頭文字のピンインだけを入力しても連続して2文字目以降のピンインを入力しても所望の文字列を入力することができる文字変換装置を提供できる。また一方では、入力したい文字列の先頭文字のピンインが不完全であり、2文字目以降のピンインを入力しても所望の文字列を入力することができる中国語入力装置を提供できる  
あるいは、中国語変換装置として用いられる上記何れかの文字変換装置では、上記漢字は繁体字であり、かつ、上記表音は注音で表現されていてよい。
- [0315] 上記構成によると、ユーザが1文字以上の注音文字列を入力すると、注音文字列の先頭文字から先頭漢字の候補を絞り込み、先頭漢字の候補を基に検索対象の単語を制限できるので、辞書検索処理にかかる時間を短縮できる。それゆえ、応答速度の

速い中国語入力装置を提供できる。また、入力したい文字列の先頭文字の注音文字だけを入力しても、連続して2文字目以降の注音文字を入力しても、所望の文字列を入力することができる中国語入力装置を提供することができる。他方で、入力したい文字列の先頭文字の注音が不完全であり、2文字目以降の注音を入力しても所望の文字列を入力することができる中国語入力装置を提供できる。

- [0316] また、本発明に係る文字変換装置の制御方法は、入力された1以上の表音文字よりなる入力文字列を変換文字に変換し、当該変換文字からなる単語を決定することができる文字変換装置の制御方法において、(a-1) 変換文字を識別するための変換文字識別情報と、該変換文字の表音を示す変換文字表音情報とが対応付けられた変換文字対応情報および、(a-2) 入力される表音文字と、該表音文字を上記変換文字表音情報の先頭として持つ上記変換文字識別情報の、上記変換文字対応情報での格納先を示す変換文字格納先情報とが対応付けられた変換文字インデックス、を有する変換文字辞書と、(b-1) 単語を識別するための単語識別情報と、該単語の表音を示す単語表音情報とが対応付けられた単語対応情報および、(b-2) 単語の先頭となる変換文字を識別するための先頭変換文字識別情報と、該先頭変換文字識別情報が示す変換文字を先頭に持つ単語の、上記単語対応情報での格納先を示す単語格納先情報とが対応付けられた単語インデックス、を有する単語辞書と、を用いて、

入力された上記入力文字列の先頭の表音文字を基に、上記変換文字辞書を参照し、該入力文字列の先頭の表音文字を上記変換文字表音情報の先頭として持つ変換文字識別情報を取得する先頭変換文字取得ステップと、

取得された上記先頭変換文字識別情報を基に、上記単語辞書を参照し、該先頭変換文字識別情報が示す変換文字を先頭に持つ単語を取得する単語取得ステップと、

を含む処理を行う。

- [0317] 上記方法によると、上記文字変換装置と同様に、単語の検索処理に係る時間を短縮することができる。

- [0318] また、本発明に係る文字変換装置は、コンピュータによって実現してもよく、この場

合には、コンピュータを上記何れかの文字変換装置における上記各手段として動作させることにより上記文字変換装置をコンピュータにて実現させる文字変換プログラム、及びその文字変換プログラムを記録したコンピュータ読み取り可能な記録媒体も、本発明の範疇に入る。

[0319] これらの構成によれば、文字変換プログラムを、コンピュータに読み取り実行することによって、上記文字変換装置と同一の作用効果を実現することができる。

[0320] [補足事項]

本発明に係る文字変換装置は、入力された表音文字または表音文字列を対応する表意文字に変換して、ユーザが選択する表意文字または表意文字列を入力する文字変換装置において、上記表音文字と、該表音文字を含む表音文字列に対応する表意文字とを関連付けて記録する文字記録部と、ユーザにより入力された表音文字を受け付けて、該表音文字に関連付けられた表意文字を、表意文字の候補として上記文字記録部から抽出する文字候補抽出手段と、上記候補抽出手段が抽出した候補の表示順序を、所定の条件に基づき決定する候補順序決定手段と、上記候補順序決定手段が決定した表示順序に基づいて候補を表示するようデータを出力する表示制御手段とを備えていることを特徴としている。

[0321] 上記構成によれば、文字候補抽出手段は、ユーザが入力した表音文字を受け付けて、その表音文字を含む、表音文字列を特定する。表音文字列は、表意文字と対応している。表音文字と表音文字列との対応関係、および、表音文字列と表意文字との対応関係は、文字記録部に記録されている。したがって、文字候補抽出手段は、ユーザから入力を受け付けた表音文字に基づいて、ユーザが入力したいと所望する表意文字の候補を上記文字記録部から抽出することができる。

[0322] 続いて、候補順序決定手段は、上記文字候補抽出手段が抽出した表意文字の候補を所定の条件に基づいて並べ替え、候補の表示順序を決定する。表示制御手段は、表意文字の候補が、上記候補順序決定手段が決定した表示順序で表示されるよう候補のデータを出力する。

[0323] これにより、ユーザは、表示された候補の中から所望の表意文字を指定するのみで、該表意文字を入力することができる。

- [0324] 以上のことから、ユーザは、表音文字の入力に係る操作数について、より少ない操作数で、所望の表意文字の入力を得ることが可能となる。
- [0325] より具体的には、例えば、表音文字がアルファベットであり、表意文字が中国語の漢字であるとする。すなわち、アルファベット文字列は、ピンインであり、1つのピンインには1または複数の漢字が対応している。上記文字記録部には、アルファベット－ピンイン、ピンイン－漢字の対応関係が記録されている。
- [0326] したがって、アルファベットが1文字入力されると、文字候補抽出手段は、文字記録部を参照して、当該アルファベットから始まるピンインを特定し、特定したピンインに対応する1または複数の漢字を候補として抽出する。
- [0327] 上記候補として抽出された漢字は、ユーザに提示される。ユーザは、提示された候補から所望の漢字を指定するだけでよい。つまり、アルファベットを入力したのち、それに対応するピンインを選択する操作を必要とすることなく、所望の漢字を指定して得ることが可能となる。
- [0328] 以上のことから、所望の表意文字(漢字)を得るための、表音文字(ピンイン)の入力操作を簡素化することが可能な文字変換装置を提供することができる。
- [0329] なお、中国語の発音には四声があるために、1つのピンイン表記に対して4種類の発音が存在する。この四声を区別する情報を加えて、対応する漢字を記録してもよい。例えば、上記文字記録部に、四声の区別を格納するためのフィールドを用意し、ユーザが入力する四声の区別に基づいて、対応する漢字をさらに絞り込めば、上記文字候補抽出手段の抽出処理を効率よく行うことができる。
- [0330] あるいは、例えば、表音文字が注音の字母であり、表意文字が繁体字であってよい。すなわち、表音文字列は注音であり、1つの注音には1または複数の繁体字が対応している。上記文字記録部には、注音の字母－注音、注音－繁体字の対応関係が記録されていればよい。
- [0331] 上記文字変換装置は、さらに、上記表意文字と、該表意文字を含む表意文字列とを関連付けて記録する文字列記録部と、ユーザにより選択された表意文字を受け付けて、該表意文字に関連付けられた表意文字列を、表意文字列の候補として上記文字列記録部から抽出する文字列候補抽出手段とを備えていてよい。

- [0332] あるいは、本発明の文字変換装置は、入力された表音文字または表音文字列を、対応する表意文字に変換して、ユーザが選択する表意文字を入力する文字変換装置において、上記表意文字と、該表意文字を含む表意文字列とを関連付けて記録する文字列記録部と、ユーザにより選択された表意文字を受け付けて、該表意文字に関連付けられている表意文字列を、表意文字列の候補として上記文字列記録部から抽出する文字列候補抽出手段と、上記文字列候補抽出手段が抽出した候補の表示順序を、所定の条件に基づき決定する候補順序決定手段と、上記候補順序決定手段が決定した表示順序に基づいて候補を表示するようデータを出力する表示制御手段とを備えていることを特徴としている。
- [0333] 上記構成によれば、文字列記録部には、表意文字(例えば、漢字)と該表意文字を含む表意文字列(例えば、当該漢字を先頭にした熟語、など)との対応関係が記録されている。したがって、上記文字列候補抽出手段は、表意文字がユーザにより選択されたのに応じて、該表意文字に関連付けられた表意文字列を、候補として上記文字列記録部から抽出する。
- [0334] 例えば、中国語の例を用いて説明する。ある漢字が入力されると、その漢字の次に続く文字(ユーザが所望する漢字)は、何の脈絡もない漢字よりも、その漢字から始まる熟語を構成する漢字である可能性が高い。したがって、入力が確定された漢字から始まる熟語を候補として表示すれば、ユーザは、その熟語の候補の中から所望の熟語を選択するのみで、別途ピンインを入力することなく、中国語を入力することが可能となる。
- [0335] これにより、ユーザは、表音文字の入力に係る操作数について、さらに少ない操作数で、所望の表意文字の入力を得ることが可能となる。
- [0336] さらに、上記文字変換装置は、ユーザにより筆記された表意文字を入力する筆記文字入力部をさらに備え、上記筆記文字入力部に入力された表意文字を認識し、該表意文字の情報を上記文字列候補抽出手段に供給する文字認識手段を備えていることが好ましい。
- [0337] これにより、所望の表意文字に対応する表音文字(列)が明瞭でない場合でも、所望の表意文字の入力を得ることが可能となる。

- [0338] さらに、上記文字変換装置の文字候補抽出手段は、ユーザにより入力された表音文字を受け付けると、連続して受け付けた表音文字列に対応する表意文字を、候補として抽出することが好ましい。
- [0339] 上記構成によれば、表音文字がユーザによって入力されるたびに、特定するべき表音文字列の候補数が絞られる。よって、文字候補抽出手段は、対応する表意文字の候補数を大幅に絞り込んで抽出することができ、表示制御手段はその絞り込まれた表意文字の候補を出力することができる。
- [0340] 以上のことから、表音文字がユーザによって入力されるごとに、表示された表意文字の候補が、ユーザが所望する表意文字を含んでいる確率が上がるよう、文字変換装置を構成することができる。結果として、所望の表意文字を得るための、表音文字の入力操作を簡素化することが可能な文字変換装置を提供することができる。
- [0341] さらに、上記文字変換装置の上記候補順序決定手段は、上記文字列候補抽出手段により抽出された表意文字列の候補が、上記文字候補抽出手段により抽出された表意文字の候補よりも優先して表示されるよう表示順序を決定することが好ましい。
- [0342] 上記構成によれば、文字候補抽出手段が文字記録部より抽出した表意文字の候補より、文字列候補抽出手段が文字列記録部より抽出した表意文字列の候補の方が先に表示されるよう、候補順序決定手段が候補の表示順序を決定する。
- [0343] 例えば、日本語や中国語の漢字入力において、何の脈絡もない漢字よりも、すでに入力されている漢字に続いて熟語(表意文字列)を構成するような漢字が続く(ユーザが所望する)可能性が高い。したがって、例えば、入力が確定された漢字から始まる熟語を候補として表示すれば、ユーザは、漢字(表意文字)の候補から表示された場合に比べて、より早く、少ない操作数で所望の漢字を入力することが可能となる。
- [0344] 上記文字変換装置は、さらに、ユーザにより選択された表意文字を学習文字として学習語記録部に記録する学習制御手段を備え、上記候補順序決定手段は、上記文字候補抽出手段が抽出した候補が上記学習語記録部に記録された学習文字である場合に、該候補が優先して表示されるよう表示順序を決定してもよい。
- [0345] あるいは、ユーザにより選択された表意文字列を学習文字列として学習語記録部に記録する学習制御手段を備え、上記候補順序決定手段は、上記文字列候補抽出

手段が抽出した候補が上記学習語記録部に記録された学習文字列である場合に、該候補が優先して表示されるよう表示順序を決定してもよい。

- [0346] 上記構成によれば、学習制御手段は、文字変換装置において入力が確定された(ユーザが所望した)表意文字または表意文字列を学習文字(列)として学習語記録部に記録させる。
- [0347] そして、上記候補抽出手段が候補を文字(列)記録部より抽出したのち、候補順序決定手段は、その抽出された候補のうち、上記学習語記録部に記録されているものを学習文字として判定することができる。候補が学習文字であるということは、すなわち、ユーザが過去に所望した表意文字(列)であって、学習文字でない候補よりも、ユーザに再び所望される可能性が高いことを意味する。
- [0348] 以上のことから、学習文字である候補の表示順序を上位にすることにより、ユーザがより早く、少ない操作数で所望の漢字を入力することが可能な文字変換装置を実現することができる。
- [0349] あるいは、上記文字変換装置は、さらに、ユーザにより選択された表意文字を、該表意文字の直前に入力されている直前文字または直前文字列と関連付けて、ペア学習文字として学習語記録部に記録する学習制御手段を備え、上記候補順序決定手段は、上記文字候補抽出手段が抽出した候補が、直前文字または直前文字列と関連付けられて上記学習語記録部に記録されたペア学習文字である場合に、該候補が優先して表示されるよう表示順序を決定してもよい。
- [0350] もしくは、ユーザにより選択された表意文字列を、該表意文字列の直前に入力されている直前文字または直前文字列と関連付けて、ペア学習文字列として学習語記録部に記録する学習制御手段を備え、上記候補順序決定手段は、上記文字列候補抽出手段が抽出した候補が、直前文字または直前文字列と関連付けられて上記学習語記録部に記録されたペア学習文字列である場合に、該候補が優先して表示されるよう表示順序を決定してもよい。
- [0351] 上記構成によれば、学習制御手段は、文字変換装置において入力が確定された(ユーザが所望した)表意文字または表意文字列を学習文字(列)として学習語記録部に記録させる際、その学習文字(列)の直前にすでに入力が確定している文字(列)

) (以下、直前文字(列)と称す)と関連付けて、ペア学習文字(列)として記録させる。

- [0352] 上記ペア学習文字(列)は、上記関連付けられた直前文字(列)に続いて、ユーザに所望された履歴があるということになる。したがって、上記直前文字(列)が入力された場合には、上記ペア学習文字(列)がユーザに所望される可能性が高い。
- [0353] 上記候補順序決定手段は、候補抽出手段が抽出した候補が、ペア学習文字(列)である場合には、その候補の表示順序を上位に変更する。
- [0354] これにより、ユーザがより早く、少ない操作数で所望の漢字を入力することが可能な文字変換装置を実現することができる。
- [0355] さらに、上記文字変換装置の上記学習制御手段は、学習語としての、上記学習文字または学習文字列、もしくは、上記ペア学習文字またはペア学習文字列を分類するためのグループ識別子を生成し、該グループ識別子を各学習語に関連付けて上記学習語記録部に記録し、上記グループ識別子は、上記学習語を構成する表意文字にあらかじめ割り当てられた、所定の長さの文字コードに基づいて生成されことが好ましい。
- [0356] 上記構成によれば、学習制御手段は、学習語記録部に記録される学習語(学習文字または学習文字列、もしくは、ペア学習文字またはペア学習文字列)に対して、それを複数のグループに分類するためのグループ識別子を生成して割り当てる。学習制御手段は、上記学習語を構成する各表意文字にあらかじめ割り当てられた、文字コードを基に求められる数値を用いて上記グループ識別子を生成する。
- [0357] このようにして生成されたグループ識別子は、どのような学習語に割り当てるものであっても固定長となる。すなわち、具体例を挙げると、2文字の熟語からなる学習語であっても、6文字の熟語からなる学習語であっても、それらには同サイズのグループ識別子が割り当たられる。
- [0358] 上記学習語記録部に記録される学習語の数が、膨大な量になればなるほど、また、それが熟語の場合、文字数が多くなればなるほど、データサイズが大きくなる。つまり、候補が学習語であるか否かを判定するために上記学習語記録部から合致する学習語を検索するのに膨大な時間がかかり、文字入力の処理効率が低下するという問題が発生する。

- [0359] しかし、上述したとおり、各学習文字に固定長のグループ識別子を割り当てれば、検索において、入力文字(列)と学習語が一致あるいは同一グループの文字(列)であるかを判定する時間は、学習語がどのような文字数からなっていようと関係なく一律となる。以上のことから、学習語の増加に対して検索にかかる時間が比例的に増加することを抑制することができ、処理効率の低下への影響が小さくすることができる。
- [0360] 上記文字変換装置は、さらに、ユーザにより選択された表意文字または表意文字列と、該表意文字または表意文字列より前に入力が確定している前方文字または前方文字列との関連度合いを記録する関連度記録部を備え、上記候補順序決定手段は、上記候補が、上記前方文字または前方文字列との関連度が高いものほど優先して表示されるよう表示順序を決定してもよい。
- [0361] 上記構成によれば、候補順序決定手段は、前方文字(列)との関連度合いに基づいて、候補の表示順序を決定する。上記関連度合いとは、上記文字入力確定手段が入力を確定した表意文字または表意文字列と、該表意文字または表意文字列より前に入力が確定されている前方文字または前方文字列との関連性のことである。
- [0362] より具体的には、例えば、「上海交通大学」という熟語が頻繁に入力される場合について説明する。「交通」に続いて「規則」などの別の熟語が連続して入力される可能性が高い場合であっても、それよりも前方に「上海」がある場合は、候補順序決定手段は、「上海」と「大学」との関連性を考慮して、「大学」を候補の上位に変更する。
- [0363] これにより、候補順序決定手段は、直前の文字(列)との連続性のみならず、それよりも前方の文字(列)との関連性を考慮して、候補となる表意文字(列)の表示順序を決定するので、よりユーザの希望を反映した順序で候補を提示することが可能となる。
- [0364] 結果として、ユーザは、表音文字の入力に係る操作数について、より少ない操作数で、所望の表意文字の入力を得ることが可能となる。
- [0365] さらに、上記文字変換装置は、上記候補順序決定手段が候補の表示順序を決定する際に参照する所定の条件であって、ユーザにより指定された条件を設定する候補順序決定要素設定手段を備え、上記候補順序決定手段は、上記候補順序決定要素設定手段が設定した条件に基づき、候補の表示順序を決定することが好ましい

。

- [0366] 上記構成によれば、上記候補順序決定手段が候補の表示順序を決定する際に参考する所定の条件が複数ある場合に、どのような条件に基づいて表示順序を決定するのかを、ユーザの指示に応じて、候補順序決定要素設定手段が設定する。
- [0367] 上記所定の条件の例としてはいくつか考えられる。例えば、上述したとおり、(1)候補が文字(漢字)であるか、文字列(熟語)であるか、(2)候補が学習文字であるか否か、(3)候補がペア学習文字であるか否か、(4)いつの時期に学習されたのか、(5)一般的にどのくらい頻繁に使用されるのか、(6)現在読みとして入力された表音文字長は、候補の表音文字長と一致しているか否か(例えば、「技術」を入力したい場合に、「ぎじ(2文字)」と「ぎじゅつ(4文字)」とは表音文字長が不一致である)、(7)直前文字列との連続性はあるか、(8)前方文字列との関連性はあるか、(9)当該ユーザによってどのくらいの頻度で使用されているのか、など、さまざまな条件が考えられる。
- [0368] ここで、どのような条件に基づいて(あるいは、どのような条件を重要視して)候補の表示順序を決定するのかということを、ユーザが任意に設定できれば、候補順序決定手段が決定する候補の表示順序は、ユーザの希望をより正確に反映したものとなる。
- [0369] 結果として、ユーザは、表音文字の入力に係る操作数について、より少ない操作数で、所望の表意文字の入力を得ることが可能となる。
- [0370] 本発明に係る複写機は、上述の文字変換装置と、各表音文字に対応するキーを備えてユーザにより押下されたキーに対応する表音文字を上記文字変換装置に供給する第1入力部と、ユーザが選択した候補を上記文字変換装置に入力する第2入力部と、上記文字候補抽出手段または上記文字列候補抽出手段が抽出する候補を、上記候補順序決定手段が定めた表示順序にしたがって表示する表示部とを備え、上記第1入力部および第2入力部と、上記表示部とは、互いに一体に形成されたタッチパネルまたはタブレットであることを特徴としている。
- [0371] これにより、ユーザにより選択された表音文字や表意文字を上記文字変換装置に入力するための第1および第2入力部と、候補としての表意文字(列)をユーザに提

示するための表示部とを一体に備えたタブレットを有する複写機において、ユーザは、表音文字の入力に係る操作数について、より少ない操作数で、所望の表意文字の入力を得ることが可能となる。

- [0372] 上記複写機は、ユーザにより筆記された表意文字を入力する筆記文字入力部をさらに備え、上記筆記文字入力部に入力された表意文字を認識し、該表意文字の情報を上記文字変換装置に供給する文字認識手段を備えていることが好ましい。
- [0373] これにより、所望の表意文字に対応する表音文字(列)が明瞭でない場合でも、所望の表意文字の入力を得ることが可能な複写機を実現することが可能となる。
- [0374] 本発明に係る文字入力方法は、入力された表音文字または表音文字列を対応する表意文字に変換して、ユーザが選択する表意文字または表意文字列を入力する文字変換装置における文字入力方法において、ユーザにより入力された表音文字を受け付けて、該表音文字に関連付けられている表意文字を、上記表音文字と、該表音文字を含む表音文字列に対応する表意文字とを関連付けて記録する文字記録部から、表意文字の候補として抽出する第1ステップと、上記第1ステップにて抽出した候補の表示順序を、所定の条件に基づき決定する第2ステップと、上記第2ステップにて決定した表示順序に基づいて候補を表示するようデータを出力するする第3ステップとを含むことを特徴としている。
- [0375] 上記方法によれば、第1ステップにて、ユーザ入力した表音文字を含む表音文字列と対応する表意文字の候補が抽出され、第2ステップにて、抽出された候補の表示順序が決定される。このとき、表示順序は、例えば、上述したような所定の条件に基づいて決定されればよい。そして、第3ステップにて、決定された表示順序に基づいて表意文字の候補が出力される。
- [0376] これにより、ユーザにより上記候補の中から所望の表意文字が指定されると、その指定された表意文字の入力が確定される。
- [0377] 以上のことから、所望の表意文字(漢字)を得るための、表音文字(ピンイン)の入力操作を簡素化することが可能な文字変換装置を提供することができる。
- [0378] なお、上記文字変換装置および複写機は、コンピュータによって実現してもよく、この場合には、コンピュータを上記各手段として動作させることにより上記文字変換装

置または複写機をコンピュータにて実現させる制御プログラム、およびそれを記録したコンピュータ読み取り可能な記録媒体も、本発明の範疇に入る。

[0379] [補足事項2]

本発明に係る文字変換装置は、入力された表音文字または表音文字列を対応する表意文字に変換して、ユーザが選択する表意文字または表意文字列を入力する文字変換装置において、上記表音文字と、該表音文字を含む表音文字列に対応する表意文字とを関連付けて記録する文字記録部と、ユーザにより入力された表音文字を受け付けて、該表音文字に関連付けられた表意文字を、表意文字の候補として上記文字記録部から抽出する文字候補抽出手段と、上記候補抽出手段が抽出した候補の表示順序を決定する候補順序決定手段と、上記候補順序決定手段が決定した表示順序に基づいて候補を表示するようデータを出力する表示制御手段とを備え、過去に入力された表意文字または表意文字列と、該表意文字または表意文字列の直前よりさらに前方に入力された前方文字または前方文字列とを対応付けて、上記前方文字または前方文字列の後方に該表意文字または表意文字列が入力される可能性の高さを示す関連度が関連度記録部に記憶されており、上記候補順序決定手段は、上記関連度記録部から、上記候補の前方に入力されている前方文字または前方文字列と、各候補との関連度を検出し、該関連度が高い候補ほど優先して表示されるよう表示順序を決定することを特徴としている。

[0380] あるいは、本発明の文字変換装置は、入力された表音文字または表音文字列を対応する表意文字に変換して、ユーザが選択する表意文字を入力する文字変換装置において、上記表意文字と、該表意文字を含む表意文字列とを関連付けて記録する文字列記録部と、ユーザにより選択された表意文字を受け付けて、該表意文字に関連付けられている表意文字列を、表意文字列の候補として上記文字列記録部から抽出する文字列候補抽出手段と、上記文字列候補抽出手段が抽出した候補の表示順序を決定する候補順序決定手段と、上記候補順序決定手段が決定した表示順序に基づいて候補を表示するようデータを出力する表示制御手段とを備え、過去に入力された表意文字または表意文字列と、該表意文字または表意文字列の直前よりさらに前方に入力された前方文字または前方文字列とを対応付けて、上記前方文字ま

たは前方文字列の後方に該表意文字または表意文字列が入力される可能性の高さを示す関連度が関連度記録部に記憶されており、上記候補順序決定手段は、上記関連度記録部から、上記候補の前方に入力されている前方文字または前方文字列と、各候補との関連度を検出し、該関連度が高い候補ほど優先して表示されるよう表示順序を決定することを特徴としている。

- [0381] これにより、ユーザは、表音文字の入力に係る操作数について、さらに少ない操作数で、所望の表意文字の入力を得ることが可能となる。
- [0382] さらに、過去に入力された表意文字または表意文字列と、該表意文字または表意文字列の直前に入力された直前文字または直前文字列とを対応付けて、上記直前文字または直前文字列の直後に該表意文字または表意文字列が入力される可能性の高さを示す連續度が連續度記録部に記憶されており、上記候補順序決定手段は、上記前方文字または前方文字列との関連度、および、上記直前文字または直前文字列との連續度が高い候補ほど優先して表示されるよう表示順序を決定することが好ましい。
- [0383] さらに、上記文字変換装置の文字候補抽出手段は、ユーザにより入力された表音文字を受け付ける度に、該表音文字をすでに受け付けた表音文字列の末尾に連結し、該連結によって得られた表音文字列に対応する表意文字を候補として抽出することが好ましい。
- [0384] さらに、上記文字変換装置の上記学習制御手段は、上記学習語記録部に記録する対象となる学習語としての、上記学習文字または学習文字列、もしくは、上記ペア学習文字またはペア学習文字列を分類するためのグループ識別子を生成し、該グループ識別子を各学習語に関連付けて上記学習語記録部に記録し、上記グループ識別子は、上記学習語を構成する表意文字にあらかじめ割り当てられた、所定の長さの文字コードに基づいて生成されることが好ましい。
- [0385] さらに、上記文字変換装置は、上記候補順序決定手段が候補の表示順序を決定する際に参照する所定の条件である、上記候補がユーザに選択される可能性の大きさを示す評価値の算出に使用可能な複数種類の要素から、ユーザにより指定された要素を設定する候補順序決定要素設定手段を備え、上記各候補に対応付けて上記

関連度とは異なる要素が要素記録部に記憶されており、上記候補順序決定手段は、上記関連度と、上記候補順序決定要素設定手段によって設定された上記要素記録部に記憶されている要素に基づき、候補の表示順序を決定することが好ましい。

- [0386] 本発明に係る文字入力方法は、入力された表音文字または表音文字列を対応する表意文字に変換して、ユーザが選択する表意文字または表意文字列を入力する文字変換装置における文字入力方法において、上記文字変換装置は、文字候補抽出手段と、候補順序決定手段と、表示制御手段とを備えており、上記文字候補抽出手段が、ユーザにより入力された表音文字を受け付けて、該表音文字に関連付けられている表意文字を、上記表音文字と、該表音文字を含む表音文字列に対応する表意文字とを関連付けて記録する文字記録部から、表意文字の候補として抽出する第1ステップと、上記候補順序決定手段が、上記第1ステップにて抽出した候補の表示順序を決定する第2ステップと、上記表示制御手段が、上記第2ステップにて決定した表示順序に基づいて候補を表示するようデータを出力する第3ステップとを含み、上記第2ステップは、過去に入力された表意文字または表意文字列と、該表意文字または表意文字列の直前よりさらに前方に入力された前方文字または前方文字列とを対応付けて、上記前方文字または前方文字列の後方に該表意文字または表意文字列が入力される可能性の高さを示す関連度を記憶する関連度記録部から、上記候補の前方に入力されている前方文字または前方文字列と、各候補との関連度を検出し、該関連度が高い候補ほど優先して表示されるよう表示順序を決定するステップを含むことを特徴としている。
- [0387] 以上のことから、所望の表意文字(漢字)を得るための、表音文字(ピンイン)の入力操作を簡素化することが可能な文字変換装置を提供することができる。
- [0388] 本発明に係る文字入力方法は、入力された表音文字または表音文字列を対応する表意文字に変換して、ユーザが選択する表意文字を入力する文字変換装置における文字入力方法において、上記文字変換装置は、文字候補抽出手段と、候補順序決定手段と、表示制御手段とを備えており、上記文字候補抽出手段が、ユーザにより選択された表意文字を受け付けて、該表意文字に関連付けられている表意文字列を、上記表意文字と該表意文字を含む表意文字列とを関連付けて記録する文字

列記録部から、表意文字列の候補として抽出する第1ステップと、上記候補順序決定手段が、上記第1ステップにて抽出した候補の表示順序を決定する第2ステップと、上記表示制御手段が、上記第2ステップにて決定した表示順序に基づいて候補を表示するようデータを出力する第3ステップとを含み、上記第2ステップは、過去に入力された表意文字または表意文字列と、該表意文字または表意文字列の直前よりさらに前方に入力された前方文字または前方文字列とを対応付けて、上記前方文字または前方文字列の後方に該表意文字または表意文字列が入力される可能性の高さを示す関連度を記憶する関連度記録部から、上記候補の前方に入力されている前方文字または前方文字列と、各候補との関連度を検出し、該関連度が高い候補ほど優先して表示されるよう表示順序を決定するステップを含むことを特徴としている。

[0389] 本発明は上述した各実施形態、各実施例に限定されるものではなく、請求項に示した範囲で種々の変更が可能である。すなわち、請求項に示した範囲で適宜変更した技術的手段を組み合わせて得られる実施形態についても本発明の技術的範囲に含まれる。

### 産業上の利用可能性

[0390] 本発明は、読み(表音文字)を変換する、特に読みを漢字に変換する各種装置に適用することができる。

## 請求の範囲

- [1] 入力された1以上の表音文字よりなる入力文字列を変換文字に変換し、当該変換文字からなる単語を決定できる文字変換装置において、  
変換文字を識別するための変換文字識別情報と、該変換文字の表音を示す変換文字表音情報とが対応付けられた変換文字対応情報、および、入力される表音文字と、該表音文字を上記変換文字表音情報の先頭として持つ上記変換文字識別情報の、上記変換文字対応情報での格納先を示す変換文字格納先情報とが対応付けられた変換文字インデックス、を有する変換文字辞書と、  
単語を識別するための単語識別情報と、該単語の表音を示す単語表音情報とが対応付けられた単語対応情報、および、単語の先頭となる変換文字を識別するための先頭変換文字識別情報と、該先頭変換文字識別情報が示す変換文字を先頭に持つ単語の、上記単語対応情報での格納先を示す単語格納先情報とが対応付けられた単語インデックス、を有する単語辞書と、  
入力された上記入力文字列の先頭の表音文字を基に、上記変換文字辞書を参照し、該入力文字列の先頭の表音文字を上記変換文字表音情報の先頭として持つ変換文字識別情報を取得する先頭変換文字取得手段と、  
取得された上記先頭変換文字識別情報を基に、上記単語辞書を参照し、該先頭変換文字識別情報が示す変換文字を先頭に持つ単語を取得する単語取得手段と、を備えていることを特徴とする、文字変換装置。
- [2] 上記先頭変換文字取得手段が取得した先頭変換文字識別情報が示す先頭変換文字を、ユーザが選択可能に表示部に表示する表示制御手段と、  
上記ユーザにより選択された先頭変換文字を受け付ける受付手段と、を備え、  
上記単語取得手段は、上記選択された先頭変換文字の先頭変換文字識別情報に基づいて、単語を取得する処理を行うことを特徴とする請求項1に記載の文字変換装置。
- [3] 上記変換文字対応情報では、さらに、上記変換文字識別情報に、該変換文字識別情報が示す変換文字についての表示補助情報が対応付けられており、  
上記表示制御手段は、上記表示補助情報に基づき、先頭変換文字を並べ替えて

、ユーザが選択可能に表示部に表示することを特徴とする請求項2に記載の文字変換装置。

- [4] 上記単語辞書の単語インデックスは、  
単語の先頭となる先頭変換文字毎に、上記先頭変換文字識別情報と、単語の2文字目の先頭の表音文字である2文字目先頭表音文字との組み合わせに対して、上記単語対応情報における格納先が割り振られていることを特徴とする請求項1～3の何れか1項に記載の文字変換装置。
- [5] 上記変換文字対応情報では、さらに、上記変換文字識別情報に、該変換文字識別情報が示す変換文字についての取得補助情報が対応付けられており、  
上記先頭変換文字取得手段は、上記取得補助情報を基に上記変換文字識別情報を取り得することを特徴とする請求項1～4の何れか1項に記載の文字変換装置。
- [6] 上記取得補助情報は、以下の4つの情報に分類されることを特徴とする請求項5に記載の文字変換装置。
  - (1) 変換文字識別情報が示す変換文字が、上記入力文字列に対して、無条件に取り得対象となるという情報
  - (2) 変換文字識別情報が示す変換文字が、上記入力文字列が1文字では、取り得対象外になるという情報
  - (3) 変換文字識別情報が示す変換文字が、上記入力文字列が2文字以下では、取り得対象外になるという情報
  - (4) 変換文字識別情報が示す変換文字が、上記入力文字列が3文字以下では、取り得対象外になるという情報
- [7] 上記単語対応情報では、単語の先頭の変換文字が同一のものが1グループとしてまとめられていることを特徴とする請求項1～6の何れか1項に記載の文字変換装置。
- [8] さらに、各上記グループ内で、単語は、単語の先頭の表音文字の所定の配列順に並べられていることを特徴とする請求項7に記載の文字変換装置。
- [9] 上記変換文字は漢字であり、  
上記変換文字対応情報では、漢字の先頭の表音文字が同一、かつ、声調コードが

同一の、上記変換文字識別情報である漢字識別情報がグループとしてまとめられており、

中国語変換装置として用いられることを特徴とする請求項1～8の何れか1項に記載の文字変換装置。

- [10] 上記変換文字対応情報では、各上記グループで、上記漢字識別情報が、漢字の先頭の表音文字の所定の配列順に並べてられていることを特徴とする請求項9に記載の文字変換装置。
- [11] 前記変換文字インデックスでは、上記漢字の先頭の表音文字が所定の配列順となっており、各上記グループの上記変換文字対応情報での格納先を示す情報を上記変換文字格納情報として、上記漢字の先頭の表音文字に対応付けられていることを特徴とする請求項9または10に記載の文字変換装置。
- [12] 上記漢字は簡体字であり、かつ、上記表音はピンインで表現される、中国語変換装置として用いられることを特徴とする請求項9～11の何れか1項に記載の文字変換装置。
- [13] 上記漢字は繁体字であり、かつ、上記表音は注音で表現される、中国語変換装置として用いられることを特徴とする請求項9～11の何れか1項に記載の文字変換装置。  
。
- [14] 入力された1以上の表音文字よりなる入力文字列を変換文字に変換し、当該変換文字からなる単語を決定することができる文字変換装置の制御方法において、  
変換文字を識別するための変換文字識別情報と、該変換文字の表音を示す変換文字表音情報とが対応付けられた変換文字対応情報、および、入力される表音文字と、該表音文字を上記変換文字表音情報の先頭として持つ上記変換文字識別情報の、上記変換文字対応情報での格納先を示す変換文字格納先情報とが対応付けられた変換文字インデックス、を有する変換文字辞書と、  
単語を識別するための単語識別情報と、該単語の表音を示す単語表音情報とが対応付けられた単語対応情報、および、単語の先頭となる変換文字を識別するための先頭変換文字識別情報と、該先頭変換文字識別情報が示す変換文字を先頭に持つ単語の、上記単語対応情報での格納先を示す単語格納先情報とが対応付けら

れた単語インデックス、を有する単語辞書と、  
を用いて、

入力された上記入力文字列の先頭の表音文字を基に、上記変換文字辞書を参照し、該入力文字列の先頭の表音文字を上記変換文字表音情報の先頭として持つ変換文字識別情報を取得する先頭変換文字取得ステップと、

取得された上記先頭変換文字識別情報を基に、上記単語辞書を参照し、該先頭変換文字識別情報が示す変換文字を先頭に持つ単語を取得する単語取得ステップと、

を含む処理を行うことを特徴とする、文字変換装置の制御方法。

- [15] 請求項1～13のいずれか1項に記載の文字変換装置の上記各手段としてコンピュータを機能させるための文字変換プログラム。
- [16] 請求項15に記載の制御プログラムを記録した、コンピュータ読取可能な記録媒体。

## 補正書の請求の範囲

### [2007年12月18日（18. 12. 2007）国際事務局受理]

された単語インデックス、を有する単語辞書と、

を用いて、

入力された上記入力文字列の先頭の表音文字を基に、上記変換文字辞書を参照し、該入力文字列の先頭の表音文字を上記変換文字表音情報の先頭として持つ変換文字識別情報を取得する先頭変換文字取得ステップと、

取得された上記先頭変換文字識別情報を基に、上記単語辞書を参照し、該先頭変換文字識別情報が示す変換文字を先頭に持つ単語を取得する単語取得ステップと、

を含む処理を行うことを特徴とする、文字変換装置の制御方法。

15. (補正後) 請求項1～13および17～22のいずれか1項に記載の文字変換装置の各手段としてコンピュータを機能させるための文字変換プログラム。

16. 請求項15に記載の制御プログラムを記録した、コンピュータ読取可能な記録媒体。

17. (追加) 上記変換文字識別情報には、該変換文字識別情報が示す変換文字についての表示補助情報が対応付けられており、

上記先頭変換文字取得手段が取得した先頭変換文字識別情報が示す先頭変換文字の表示順序を、上記表示補助情報に基づき決定する順序決定手段と、

上記順序決定手段が決定した表示順序に基づいて上記先頭変換文字をユーザが選択可能に表示部に表示する表示制御手段とを備えていることを特徴とする請求項1に記載の文字変換装置。

18. (追加) ユーザにより選択された上記先頭変換文字を学習文字として学習語記録部に記録する学習制御手段を備え、

上記順序決定手段は、先頭変換文字が上記学習語記録部に記録された学習文字であるか否かを示す表示補助情報に基づいて、上記先頭変換文字取得手段が取得した先頭変換文字が学習文字であるか否かを判断して、学習文字である上記先頭変換文字が他の先頭変換文字に優先して表示されるように表示順序を決定することを特徴とする請求項17

に記載の文字変換装置。

19. (追加) 上記学習語記録部には、上記先頭変換文字が学習文字として上記学習制御手段によって記録された時期を示す学習時期が、上記表示補助情報として各先頭変換文字に対応付けて記憶されており、

上記順序決定手段は、上記学習時期が新しい先頭変換文字ほど優先して表示されるように表示順序を決定することを特徴とする請求項18に記載の文字変換装置。

20. (追加) 上記単語識別情報には、該単語識別情報が示す単語についての表示補助情報が対応付けられており、

上記単語取得手段が取得した単語の表示順序を、上記表示補助情報に基づき決定する順序決定手段と、

上記順序決定手段が決定した表示順序に基づいて上記単語をユーザが選択可能に表示部に表示する表示制御手段とを備えていることを特徴とする請求項1に記載の文字変換装置。

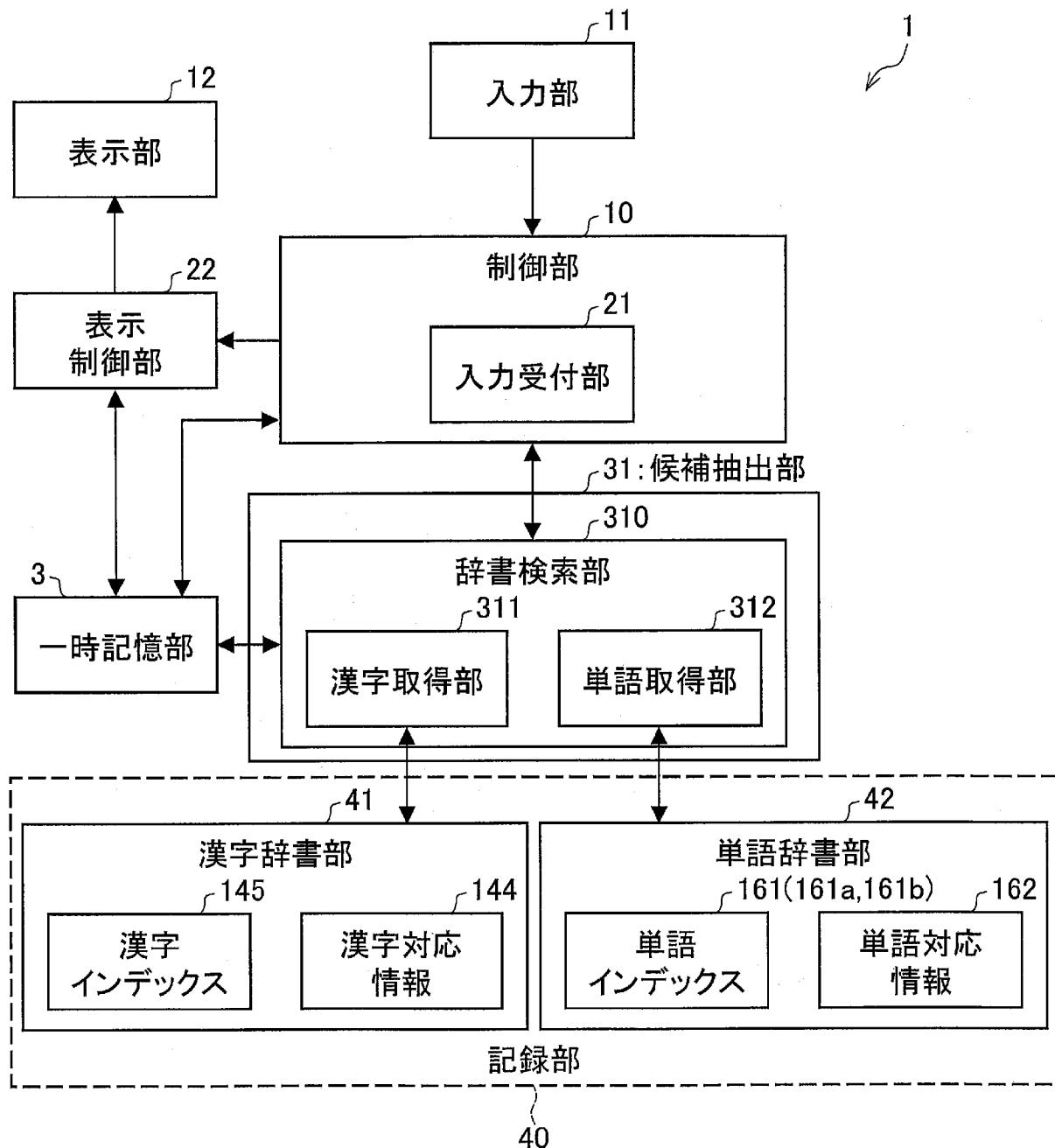
21. (追加) ユーザにより選択された上記単語を学習単語として学習語記録部に記録する学習制御手段を備え、

上記順序決定手段は、単語が上記学習語記録部に記録された学習単語であるか不かを示す表示補助情報に基づいて、上記単語取得手段が取得した単語が学習単語であるか否かを判断して、学習単語である上記単語が他の単語に優先して表示されるように表示順序を決定することを特徴とする請求項20に記載の文字変換装置。

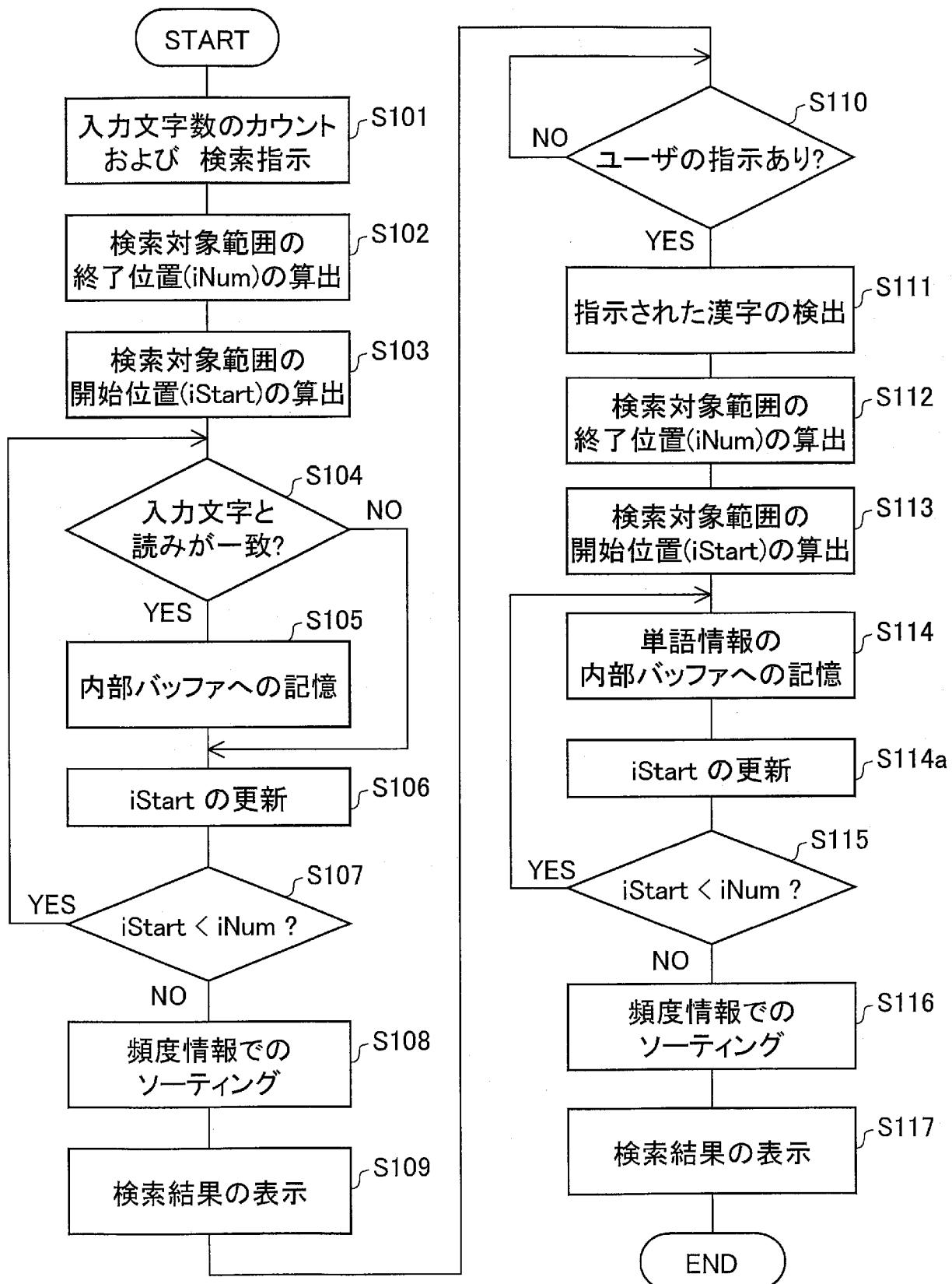
22. (追加) 上記学習語記録部には、上記単語が学習単語として上記学習制御手段によって記録された時期を示す学習時期が、上記表示補助情報として各単語に対応付けて記憶されており、

上記順序決定手段は、上記学習時期が新しい単語ほど優先して表示されるように表示順序を決定することを特徴とする請求項21に記載の文字変換装置。

[図1]



[図2]



[図3]

145  
↙

ピンイン 先頭文字	漢字 アドレス
:	:
:	:
cidx_m1	cp_m1
cidx_m2	cp_m2
cidx_m3	cp_m3
cidx_m4	cp_m4
cidx_m5	cp_m5
cidx_n1	cp_n1
cidx_n2	cp_n2
:	cp_n3
:	cp_n4
:	cp_n5
cidx_o1	cp_o1
:	:
:	:

[図4]

144

↓

	漢字	漢字ピンイン	検索条件	頻度情報
cp_m1	:	:	:	:
	妈	ma1	0	1 7 3
	麻	ma1	2	1 0 6
	:	:	:	:
	蒙	meng1	3	0
	:	:	:	:
cp_m2	麻	ma2	0	1 5 6
	吗	ma2	2	1 0 2
	:	:	:	:
	没	mei2	1	2 0 7
	糜	mei2	1	0
	煤	mei2	0	1 3 5
cp_m3	蒙	meng2	0	1 6 2
	:	:	:	:
	糜	mi2	0	8 0
	:	:	:	:
	吗	ma3	0	1 7 5
	:	:	:	:
cp_m4	每	mei3	0	1 8 4
	浼	mei3	0	1 9
	:	:	:	:
	:	:	:	:
	妹	mei4	0	1 6 3
	媚	mei4	0	1 2 4
cp_m5	:	:	:	:
	吗	ma5	2	8 8
cp_n1	:	:	:	:
	拈	:	0	1 0 0
cp_n2	:	:	:	:
	年	:	0	2 1 6
	粘	:	0	1 1 9
	:	:	:	:

[図5]

161a



漢字	漢字コード	2次インデックス内 漢字アドレス
啊	B0A1	:
阿	B0A2	:
:	:	:
大	B4F3	widx_di
:	:	:
规	B9E6	widx_gi
:	:	:
交	BDBB	widx_ji
:	:	:
每	C3BF	widx_mi
冒	C3C0	widx_mi+1
:	:	:
妹	C3C3	widx_mi+j
:	:	:
拈	C4E9	widx_ni
年	C4EA	widx_nj
碾	C4EB	widx_nk
:	:	:
:	:	:
上	C9CF	widx_si
:	:	:
:	:	:
:	:	:

[図6]

161b  
↙

先頭漢字	2文字目漢字 先頭ピンイン	単語 アドレス
:	:	:
:	:	:
widx_di	大 x	wp_di
:	:	:
widx_gi	規 z	wp_gi
:	:	:
widx_ji	交 t	wp_ji
:	:	:
widx_mi	每 n	wp_mi
:	t	wp_mi+k
:	:	:
widx_mi+1	冒 c	wp_mi+l
:	:	:
widx_mi+j	妹 m	wp_mi+j
:	:	:
:	:	:
widx_ni	拈 :	wp_ni
:	:	:
widx_nj	年 :	wp_nj
:	:	:
:	:	:
widx_si	上 h	wp_si+l
:	h	wp_si+2
:	:	:
:	m	wp_si+m
	:	:

[図7]

162

	単語	単語ピンイン
wp_di	大学	DaXue
wp_gi	规则	GuiZe
wp_ji	交通 警察	JiaoTong JingCha
wp_mi	每年	MeiNian
wp_mi+k	每天	MeiTian
wp_mi+l	冒充	MaoCong
wp_si	上海 上海市	ShangHai ShangHaiShi

[図8]

ピンイン 先頭文字	ピンイン 辞書アドレス
a	pr_a1
:	pr_a2
:	:
:	pr_a7
b	pr_b1
:	:
:	pr_17
m	pr_m1
:	pr_m2
:	pr_m3
:	pr_m4
:	pr_m5
:	pr_m6
:	pr_m7
n	pr_n1
:	:
:	:
:	pr_z7
:	pr_end

[図9]

	ピンイン長	ピンイン
pr_a1	1	a
pr_a2	2	a 1
:	:	:
pr_b1	1	b
:	:	:
:	:	:
pr_m1	1	m
pr_m2	2	m 1
	2	m 2
:	:	:
	2	m u
pr_m3	3	m a 1
:	:	:
:	:	:
:	:	:
pr_m4	4	m a i 1
:	:	:
:	:	:
:	:	:
:	:	:
:	:	:
:	:	:
pr_z7	7	z h u a n g 1
:		z h u a n g 2
:		z h u a n g 3
:		z h u a n g 4
		z h u a n g 5
pr_end		

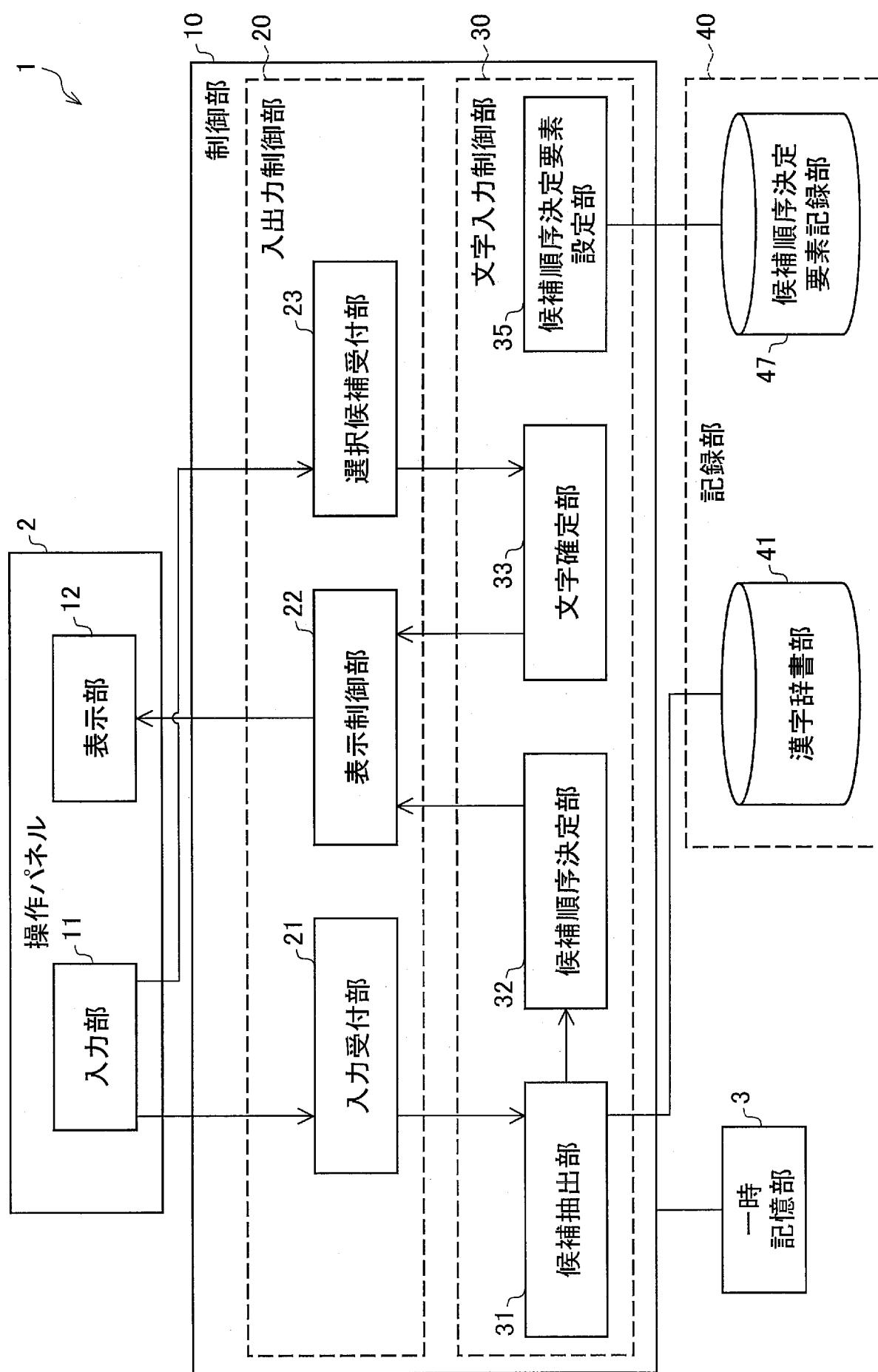
[図10]

分割位置	音節区切り 位置評価値	分割位置数	第1音節	第2音節
1	4	2	m	ein
2	2	1	me	なし
3	7	2	mei	nian
4	4	1	mein	なし

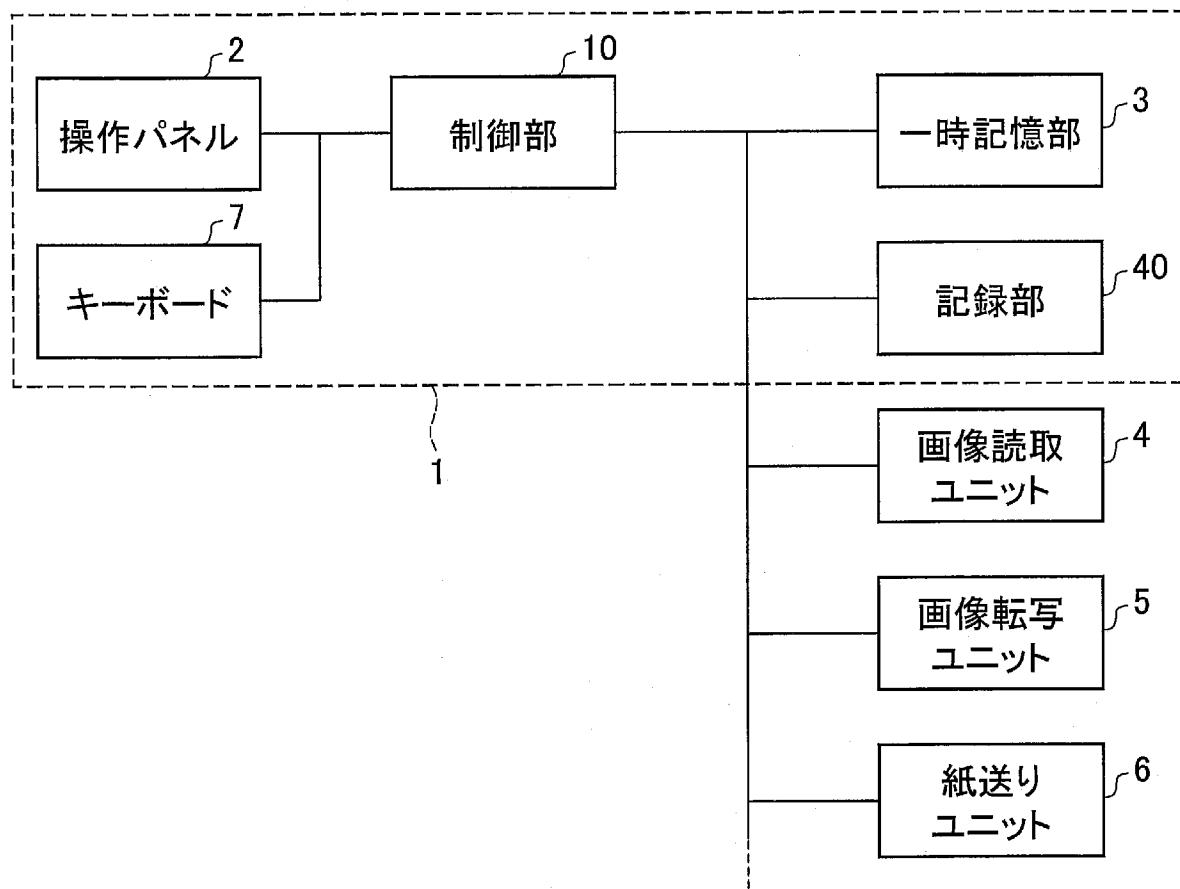
[図11]

分割位置	音節区切り 位置評価値	分割位置数	第1音節	第2音節
1	1	1	n	なし
2	4	2	ni	an
4	4	1	nian	なし

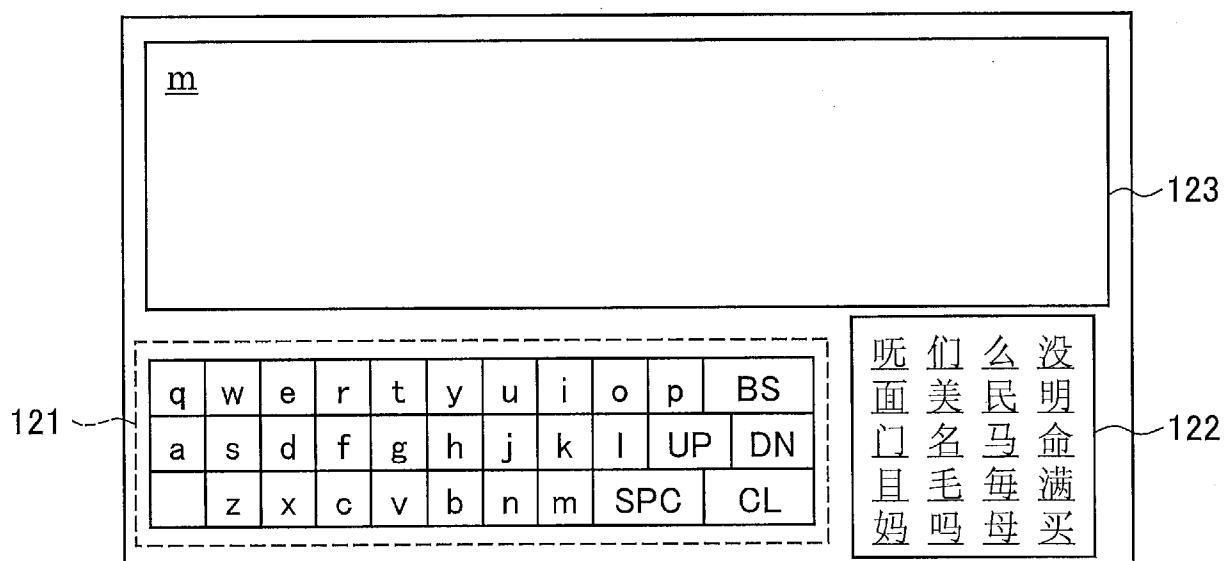
[図12]



[図13]



[図14]



[図15]

140

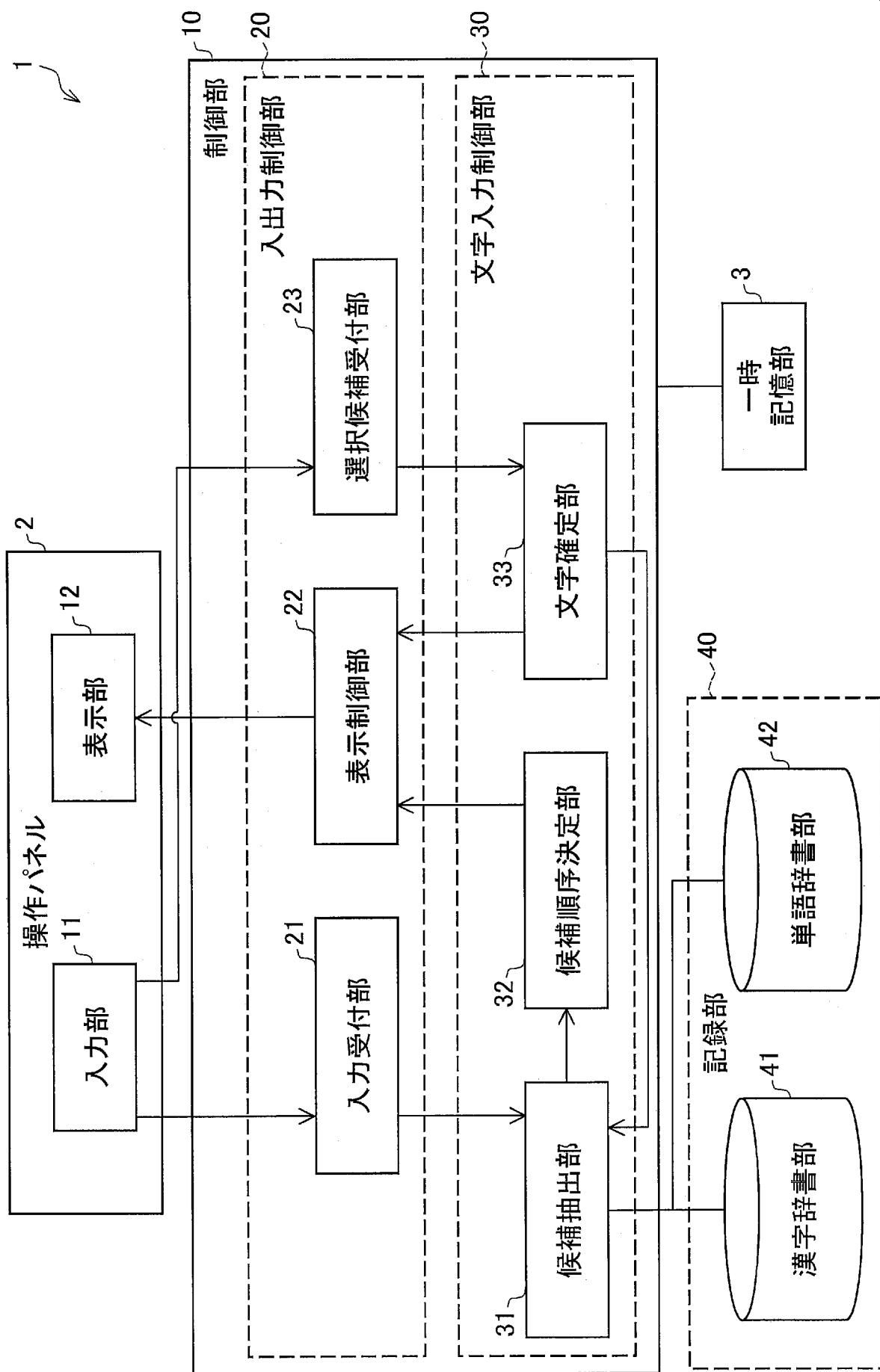
ピンイン	アドレス
:	:
ma	cp_mi1
:	:
mei	cp_mi
:	:
:	:
:	:
shang	cp_si
:	:
:	:

141

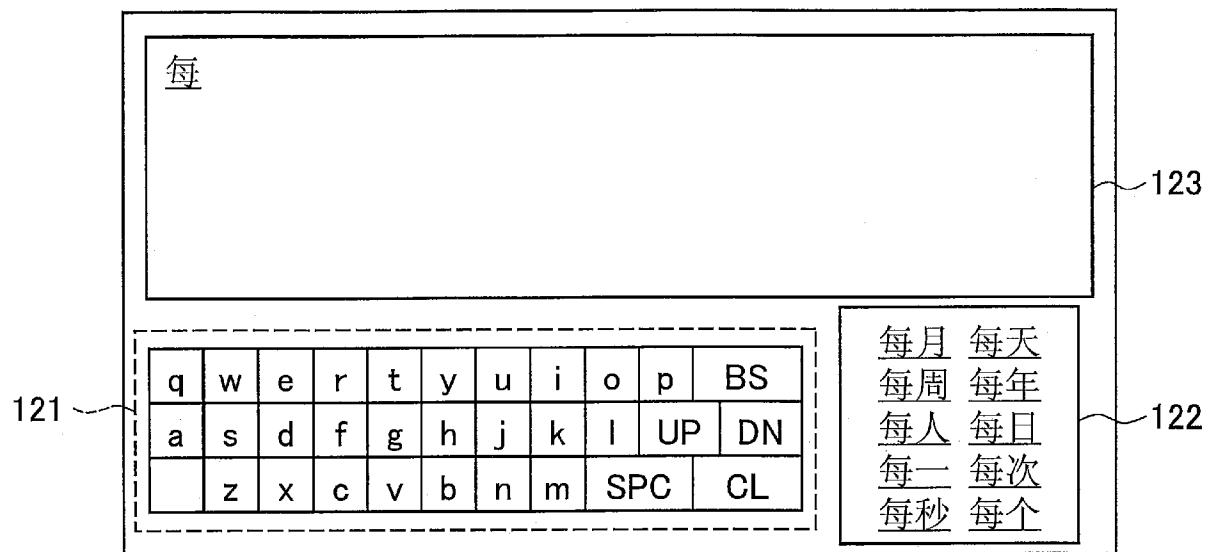
漢字	アドレス
妈	wp_mi
麻	wp_mi+1
:	:
没	wp_mi
糜	wp_mi+1
煤	wp_mi+2
:	:
每	wp_mi+j
浼	wp_mi+j+1
:	:
：	:
上	wp_si
:	:

142

[図16]



[図17]

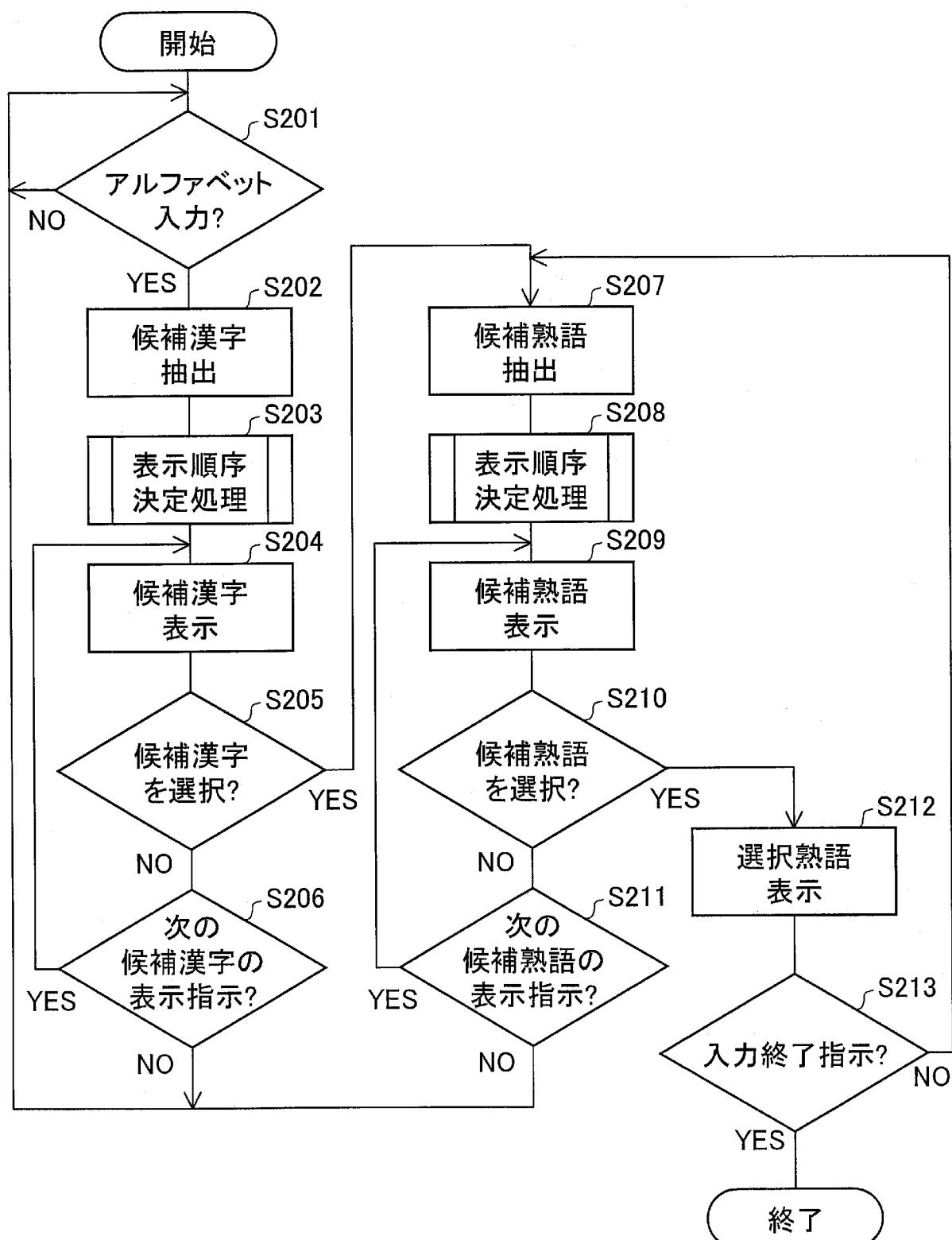


[図18]

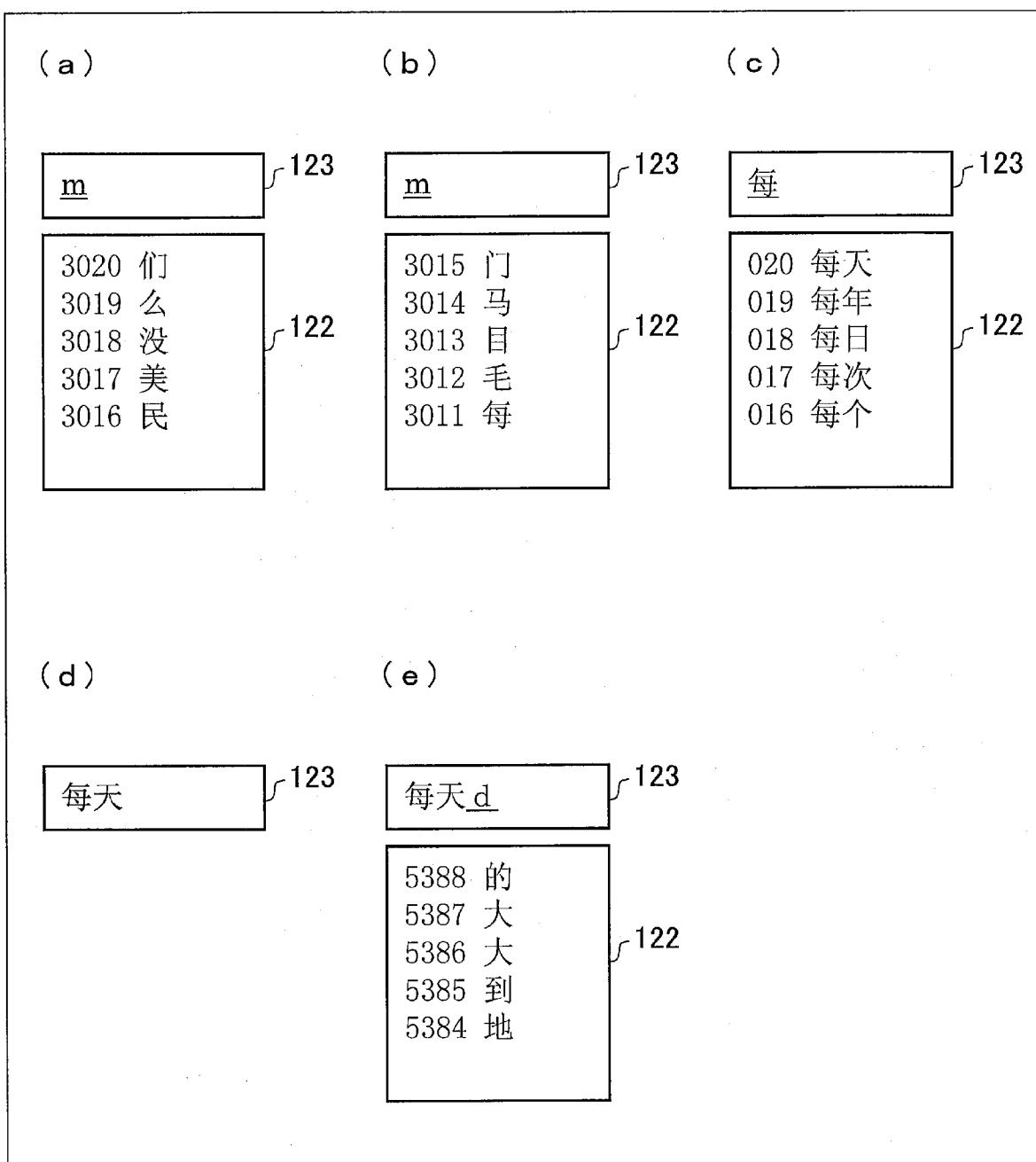
143

	熟語	アドレス
wp_mi+j	毎	np[i]+m
	毎月	np[i]+m+1
	每天	np[i]+m+2
	:	:
wp_si	上	np[i]+s-1
	上海	np[i]+s
	上海市	np[i]+s+1
	:	:
	:	:

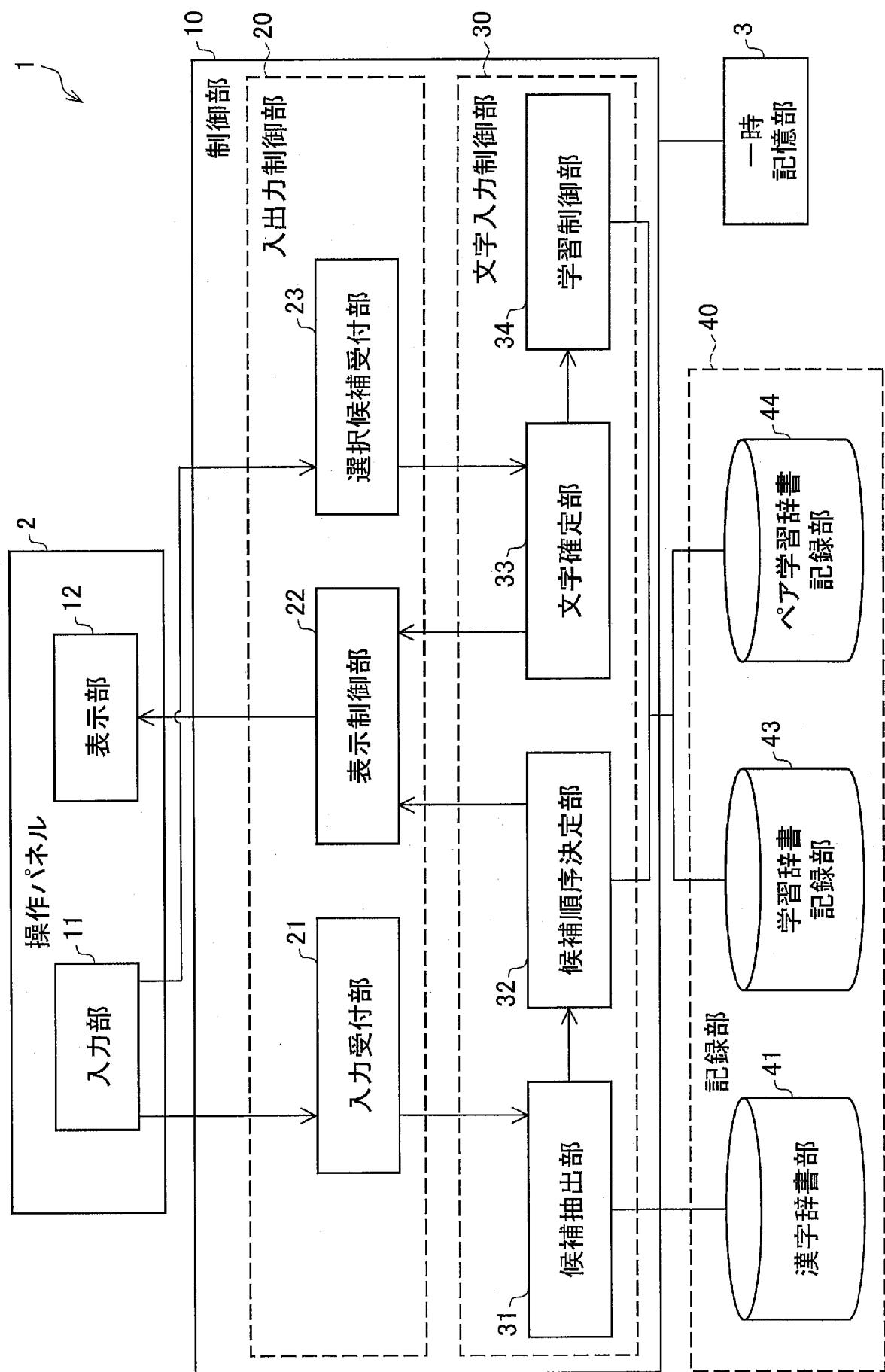
[図19]



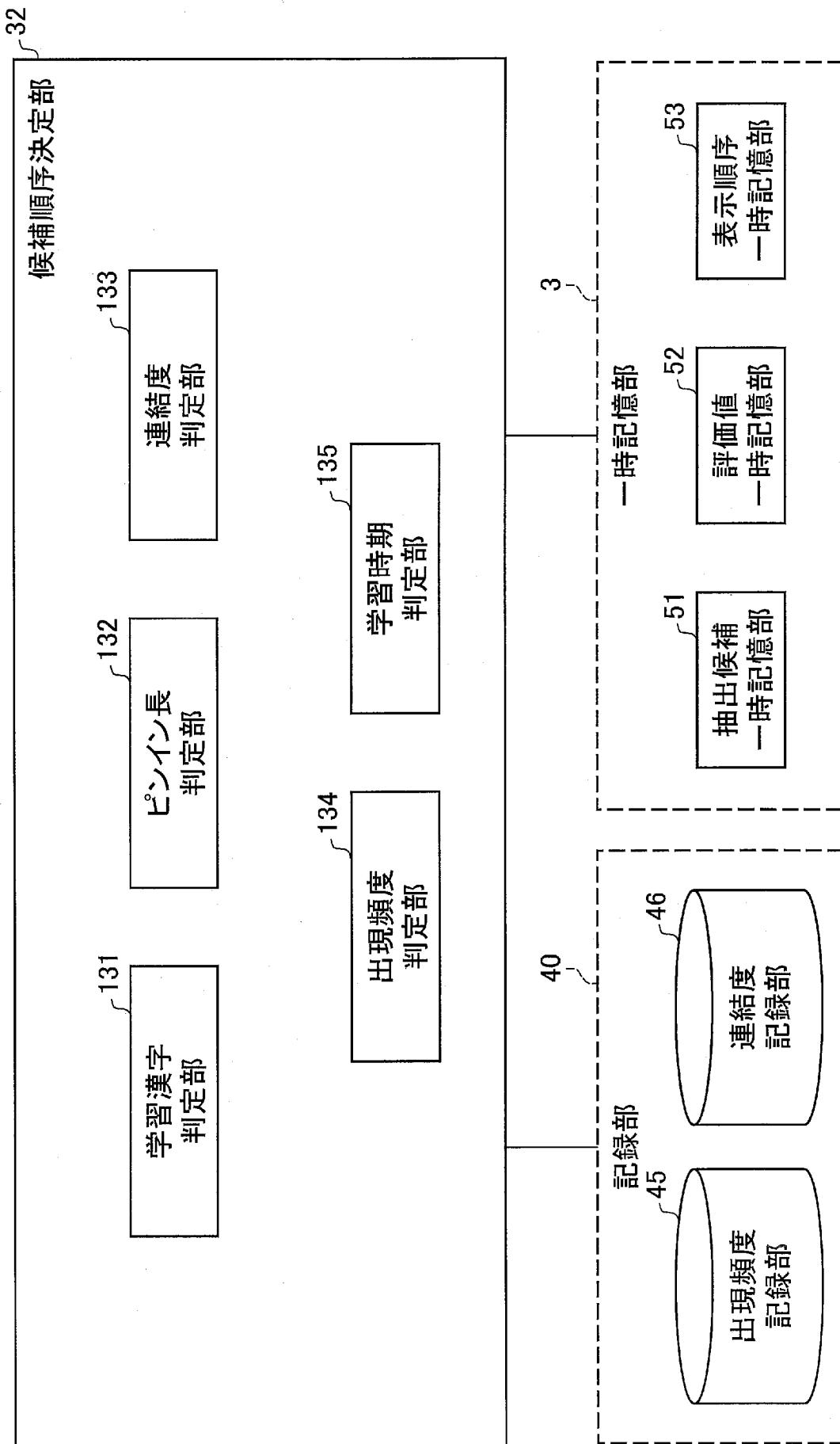
[図20]



[図21]



[図22]



[図23]

150 ↗

	漢字・熟語	連続性アドレス	関連性アドレス
wp_di	大学	: npi+j	: ppi+j
wp_gi	规则	: npi+k	: ppi+k
wp_ji	交通 警察	: npi+l npi+l+1	: ppi+l ppi+l+1
wp_ti	工具	: npi+t	: ppi+t
wp_mi+1	每	: npi+m	: ppi+m
wp_mi+2	每天	: npi+m+1	: ppi+m+1
wp_si	上 上海 上海市	: npi+s-1 npi+s npi+s+1	: ppi+s-1 ppi+s ppi+s+1
		:	:
		:	:

~151

152 ↗

	标志	大学	工具	规则	交通	警察	上海	信号灯
:	:							
npi+l	交通	12	16	15	21	1	14	1
:	:							
npi+j	每							
:	:							
:	:							
npi+s-1	上							
npi+s	上海	5	15	6	8	13	12	1
npi+s+1	上海市							4
:	:							
:	:							

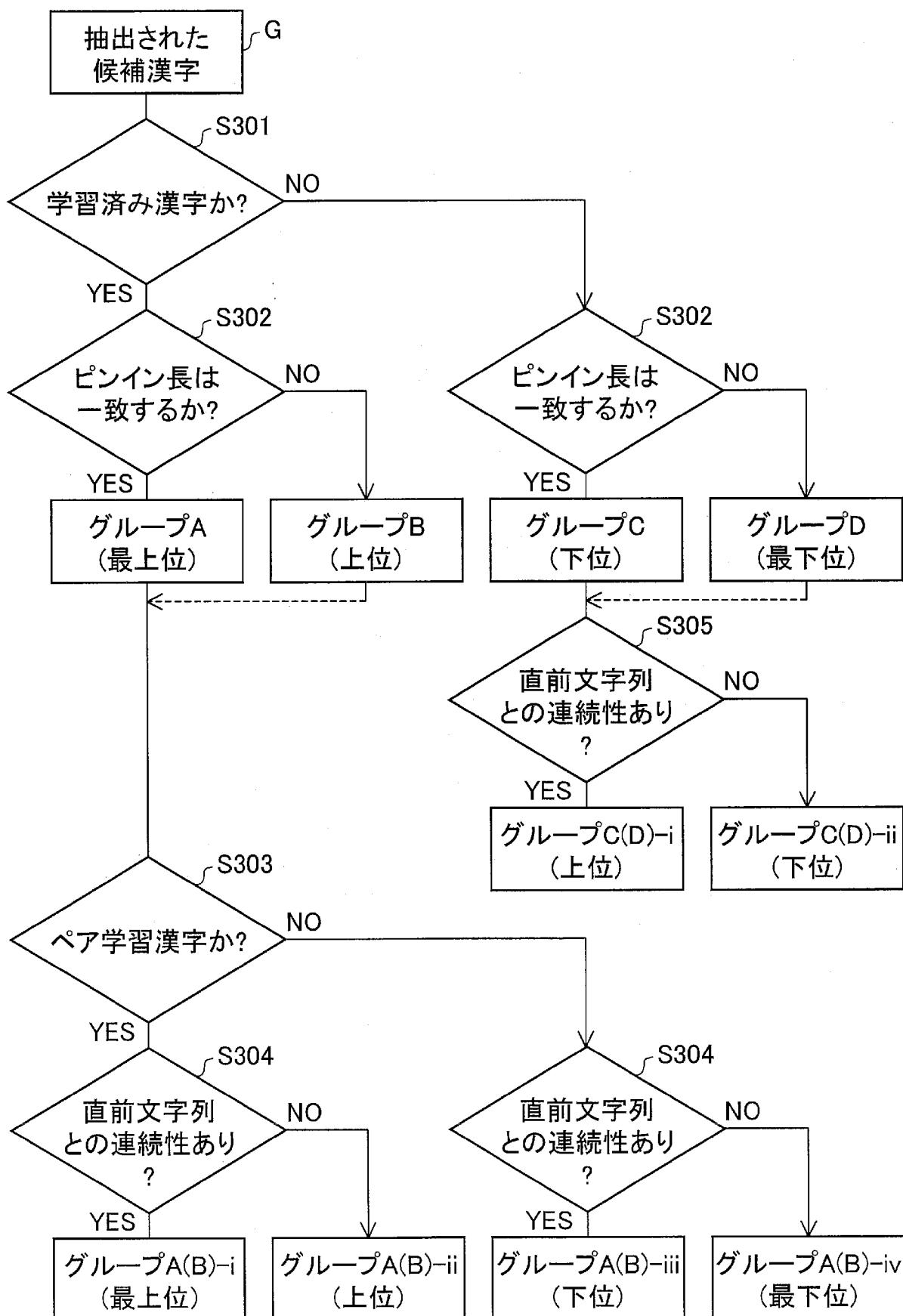
152

153 ↗

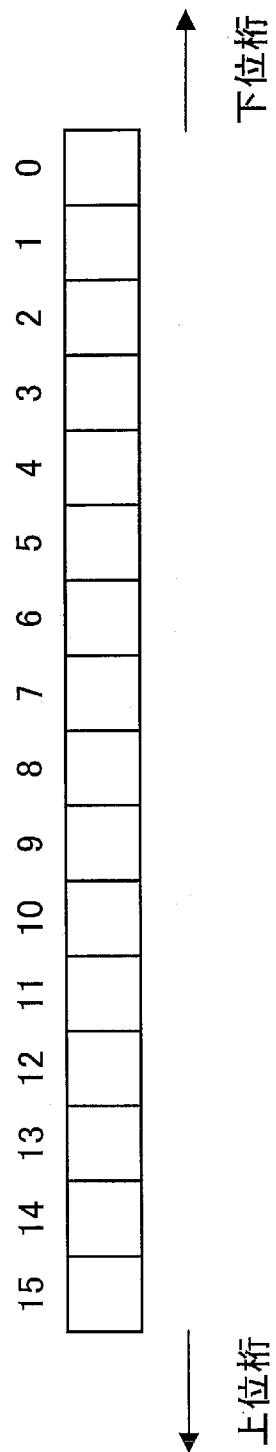
	标志	大学	工具	规则	交通	警察	上海	信号灯
:	:							
ppi	标志	1	6	2	4	8	5	4
:	:							5
ppi+j	大学	1	1	3	3	7	6	10
:	:							2
ppi+k-1	工具	1	2	1	1	2	2	5
ppi+k	规则	5	10	6	1	7	5	8
:	:							5
ppi+l	交通	1	2	2	3	1	5	5
ppi+l+1	警察	1	5	1	5	8	1	8
:	:							2
ppi+s	上海	1	10	2	5	1	3	1
:	:							1
ppi+t	信号灯	1	3	1	2	5	5	5
:	:							1

153

[図24]



[図25]



[図26]

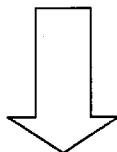
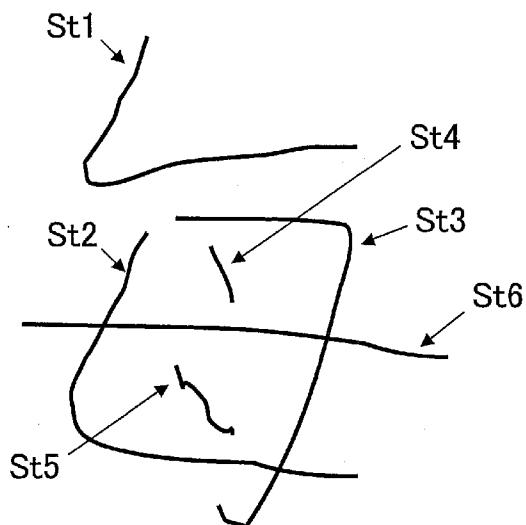
( a )

ピンイン長一致/不一致(要素B)	12,12
単独学習漢字/熟語(要素E)	11,11
ペア学習漢字/熟語(要素F)	10,10
直前文字列連續性(要素C)	9,9
前方文字列関連性(要素D)	8,8
出現頻度/学習時期(要素A/G)	0,7

( b )

ピンイン長一致/不一致(要素B)	10,10
単独学習漢字/熟語(要素E)	11,11
ペア学習漢字/熟語(要素F)	12,12
直前文字列連續性(要素C)	9,9
前方文字列関連性(要素D)	8,8
出現頻度/学習時期(要素A/G)	0,7

[図27]



60

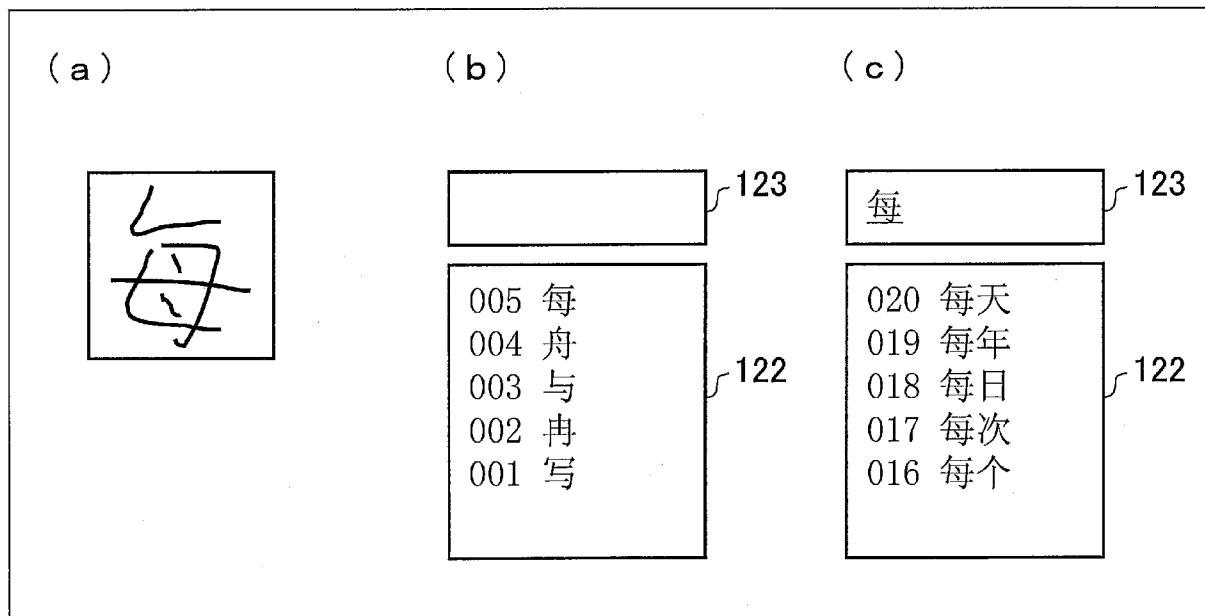
		$P_1$											
on	$x_1, y_1$	$t_1$	.....		on	$x_{k-1}, y_{k-1}$	$t_{k-1}$	off	-1,-1	$t_k$			
on	$x_{k+1}, y_{k+1}$	$t_{k+1}$	.....		on	$x_{l-1}, y_{l-1}$	$t_{l-1}$	off	-1,-1	$t_l$	on	$x_{l+1}, y_{l+1}$	$t_{l+1}$
.....													
		on	$x_{m-1}, y_{m-1}$	$t_{m-1}$	off	-1,-1	$t_m$	on	$x_{m+1}, y_{m+1}$	$t_{m+1}$			
		on	$x_{n-1}, y_{n-1}$	$t_{n-1}$	off	-1,-1	$t_n$				$P_n$		

St1:  $P_0 \sim P_{k-1}$ St2:  $P_k \sim P_{l-1}$ St3:  $P_l \sim P_{m-1}$ 

:

:

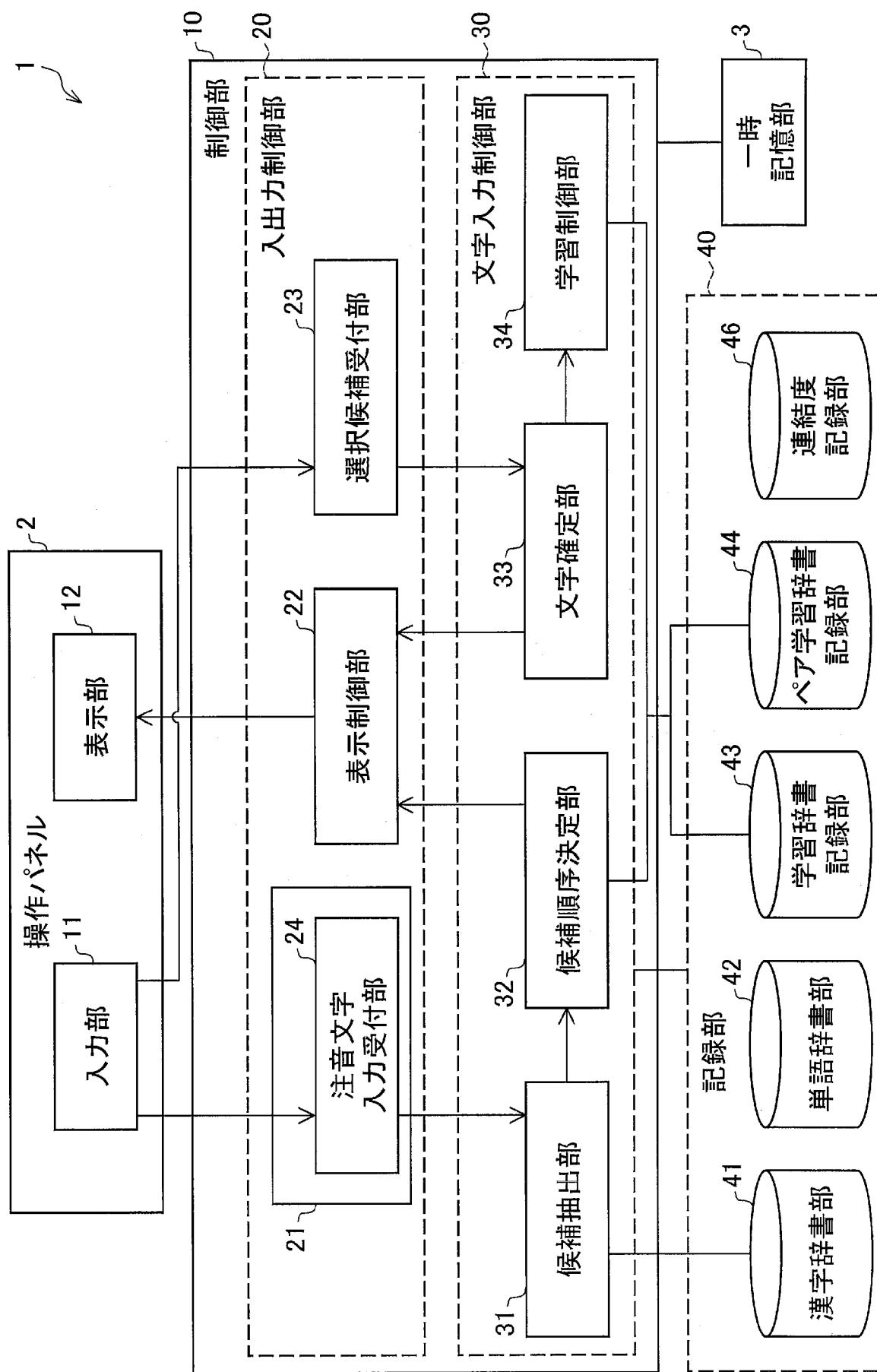
[図28]



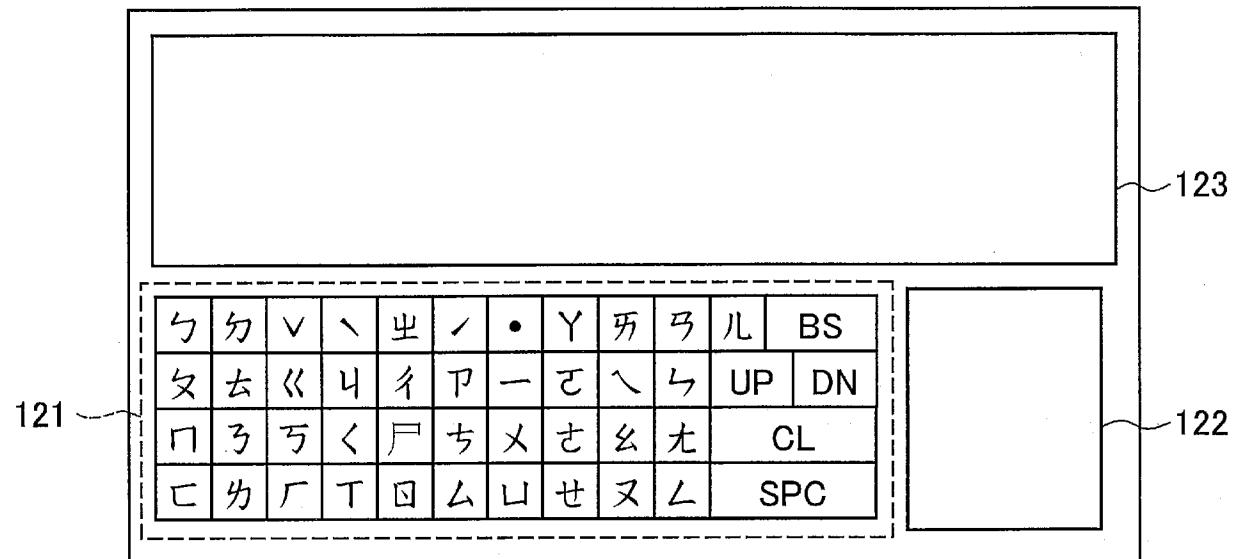
[図29]

每	10
舟	25
与	35
冉	40
写	50
:	:
:	:

[図30]



[図31]



[図32]

注音字母	アドレス	漢字	アドレス
:	:	:	:
ㄇ	cp_mi	：	：
:	cp_mi	：	：
ㄇ	cp_mi	没	wp_mi
ㄇ	cp_mi	糜	wp_mi+1
ㄇ	cp_mi	煤	wp_mi+2
：	：	：	：
：	：	：	：
：	：	：	：
ㄇ	cp_si	每	wp_mi+j
ㄇ	cp_si	浼	wp_mi+j+1
：	：	：	：
：	：	：	：
：	：	上	wp_si

[図33]

	熟語	アドレス
wp_mi+j	毎	npi+m
	毎月	npi+m+1
	每天	npi+m+2
	:	:
wp_si	上	npi+s-1
	上海	npi+s
	上海市	npi+s+1
	:	:
	:	:

[図34]

	漢字・熟語	連続性アドレス	関連性アドレス
wp_di	大學	npi+j	ppi+j
wp_gi	規則	npi+k	ppi+k
wp_ji	交通警察	npi+l npi+l+1	ppi+l ppi+l+1
wp_ti	工具	npi+t	ppi+t
wp_mi+1	每	npi+m	ppi+m
wp_mi+2	每天	npi+m+1	ppi+m+1
wp_si	上海	npi+s-1	ppi+s-1
	上海市	npi+s	ppi+s
		npi+s+1	ppi+s+1
		⋮	⋮
		⋮	⋮

150a

~151a

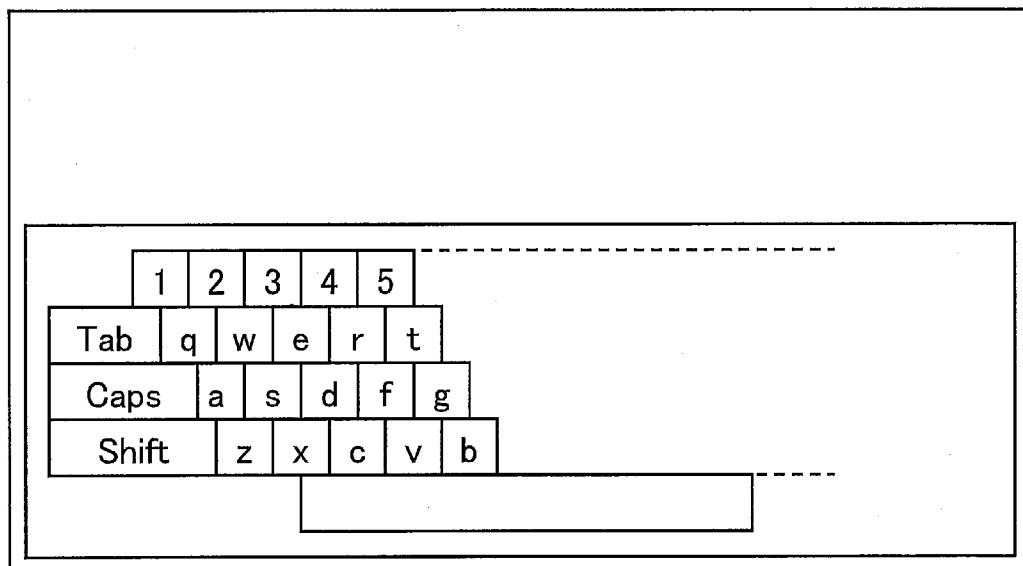
	標志	大學	工具	規則	交通	警察	上海	信号燈
:	:							
npi	交通	12	14	15	15	1	14	1
:	:							
npi+j	每							
:	:							
:	:							
npi+s-1	上							
npi+s	上海	5	13	6	8	13	12	1
npi+s+1	上海市							4
:	:							
:	:							

152a

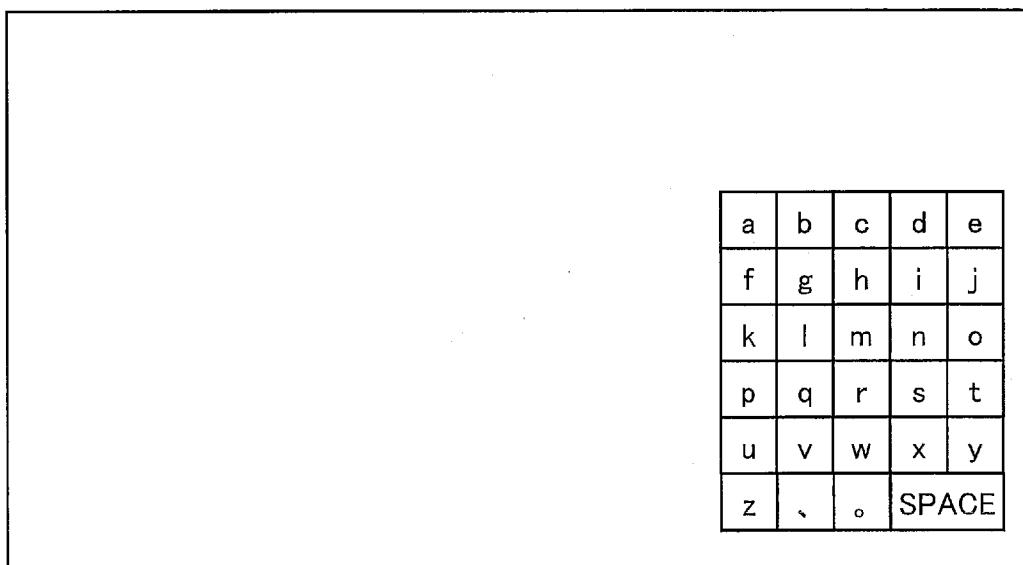
	標志	大學	工具	規則	交通	警察	上海	信号燈
:	:							
ppi	標志	1	6	2	4	8	5	4
:	:							5
ppi+j	大學	1	1	3	3	7	6	10
:	:							2
ppi+k-1	工具	1	2	1	1	2	2	5
ppi+k	規則	5	10	6	1	7	5	8
:	:							5
ppi+l	交通	1	2	2	3	1	5	5
ppi+l+1	警察	1	5	1	5	8	1	8
:	:							2
ppi+s	上海	1	2	2	2	1	3	1
:	:							1
ppi+t	信号燈	1	3	1	2	5	5	5
:	:							1

153a

[図35]

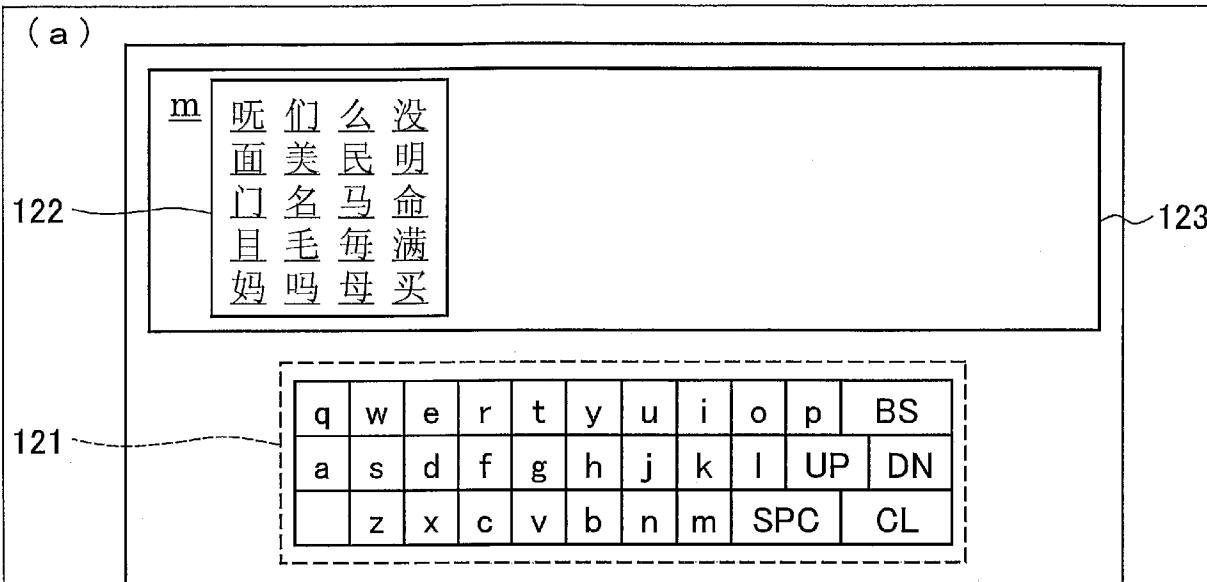


[図36]

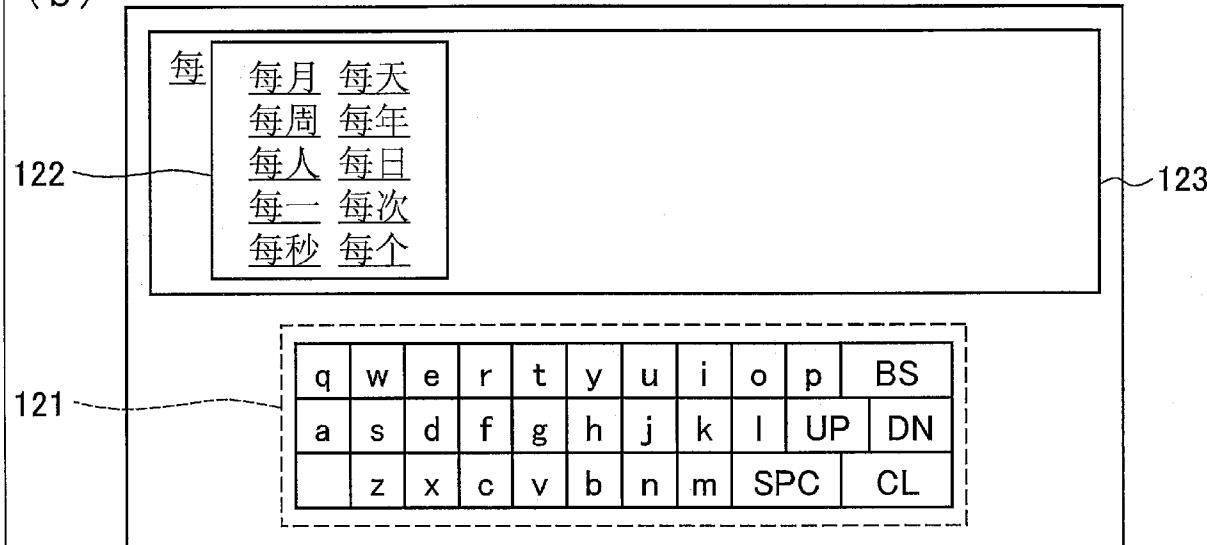


[図37]

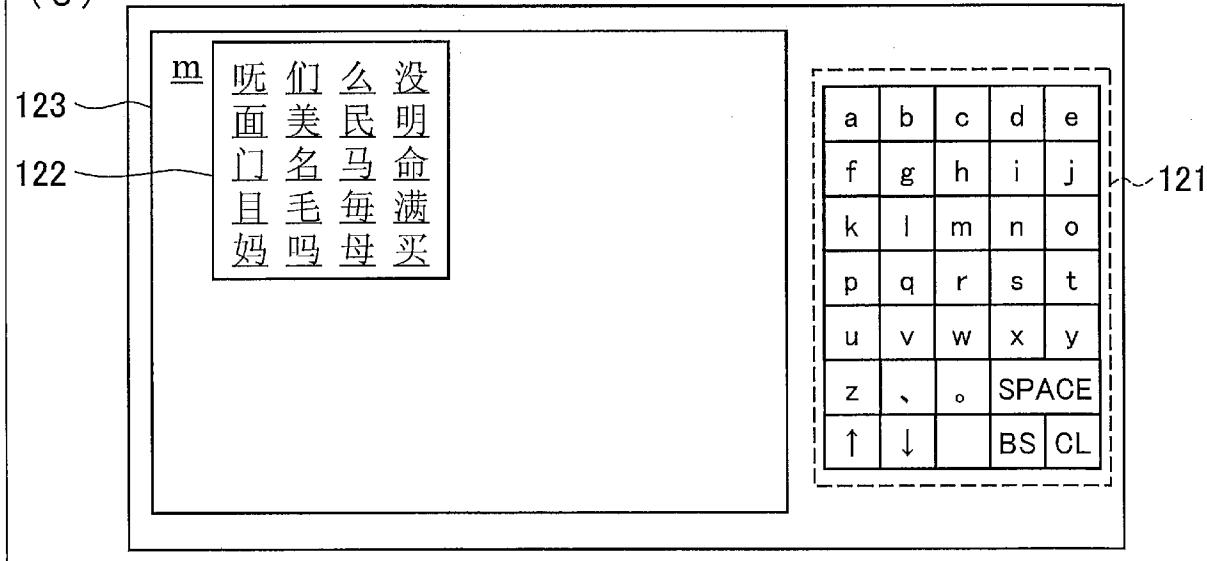
(a)



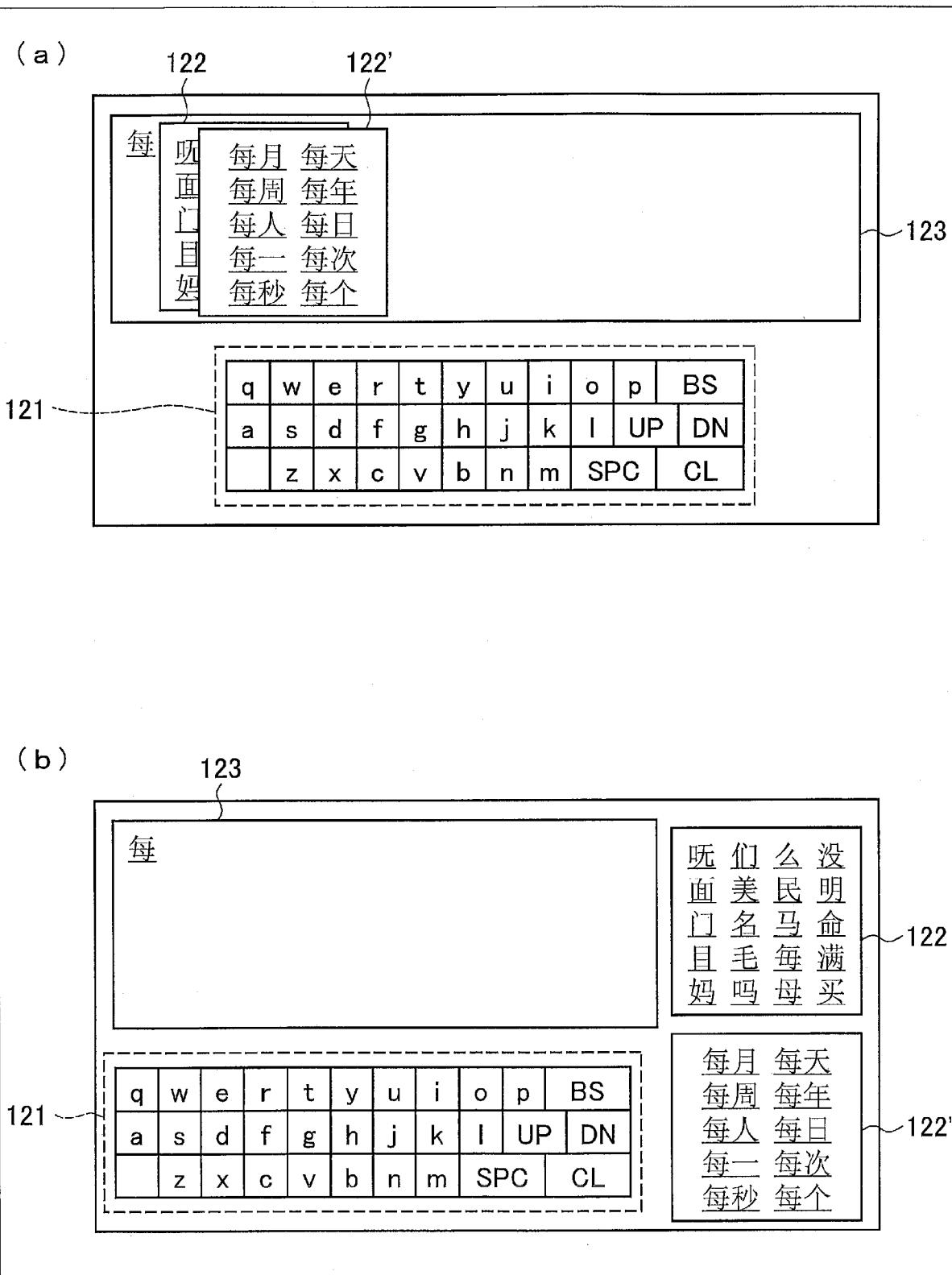
(b)



(c)



[図38]



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2007/064209

**A. CLASSIFICATION OF SUBJECT MATTER**

G06F17/22 (2006.01)i, G06F17/21 (2006.01)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)

G06F17/20-17/261

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	1996-2007
Kokai Jitsuyo Shinan Koho	1971-2007	Toroku Jitsuyo Shinan Koho	1994-2007

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.
A	JP 4-225415 A (Mitsubishi Electric Corp.) , 14 August, 1992 (14.08.92) , Full text; all drawings (Family: none)	1-16
A	JP 2006-50160 A (Sharp Corp.) , 16 February, 2006 (16.02.06) , Par. No. [0106]; Fig. 39 (Family: none)	1-16

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

12 October, 2007 (12.10.07)

Date of mailing of the international search report

23 October, 2007 (23.10.07)

Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Faxsimile No.

Telephone No.

**INTERNATIONAL SEARCH REPORT**

International application No.

PCT/JP2007/064209

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CD-ROM of the specification and drawings annexed to the request of Japanese Utility Model Application No. 88662/1991 (Laid-open No. 40946/1993) (Casio Computer Co., Ltd.), 01 June, 1993 (01.06.93), Full text; all drawings (Family: none)	1-16
A	JP 2006-155322 A (NEC Fielding Ltd.), 15 June, 2006 (15.06.06), Full text; all drawings (Family: none)	1-16
E, A	JP 2007-188449 A (Sharp Corp.), 26 July, 2007 (26.07.07), Full text; all drawings (Family: none)	1-16
A	JP 2004-5379 A (Fujitsu Ltd.), 08 January, 2004 (08.01.04), Par. Nos. [0053] to [0054] & US 2003/233615 A1 & EP 1359516 A2 & CN 1452083 A	6

## A. 発明の属する分野の分類(国際特許分類(IPC))

Int.Cl. G06F17/22 (2006.01)i, G06F17/21 (2006.01)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(IPC))

Int.Cl. G06F17/20-17/261

## 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	1922-1996年
日本国公開実用新案公報	1971-2007年
日本国実用新案登録公報	1996-2007年
日本国登録実用新案公報	1994-2007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 請求の範囲の番号
A	JP 4-225415 A (三菱電機株式会社) 1992.08.14, 全文, 全図 (ファミリーなし)	1-16
A	JP 2006-50160 A (シャープ株式会社) 2006.02.16, 段落 0106, 図 39 (ファミリーなし)	1-16

 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」国際出願目前で、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

12. 10. 2007

## 国際調査報告の発送日

23. 10. 2007

## 国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

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)

成瀬 博之

5M

9192

電話番号 03-3581-1101 内線 3599

C (続き) . 関連すると認められる文献		関連する 請求の範囲の番号
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	
A	日本国実用新案登録出願 3-88662 号 (日本国実用新案登録出願公開 5-40946 号) の願書に添付した明細書及び図面の内容を記録した CD-ROM (カシオ計算機株式会社) 1993. 06. 01, 全文, 全図 (ファミリーなし)	1-16
A	JP 2006-155322 A (NECフィールディング株式会社) 2006. 06. 15, 全文, 全図 (ファミリーなし)	1-16
EA	JP 2007-188449 A (シャープ株式会社) 2007. 07. 26, 全文, 全図 (ファミリーなし)	1-16
A	JP 2004-5379 A (富士通株式会社) 2004. 01. 08, 段落 53-54 & US 2003/233615 A1 & EP 1359516 A2 & CN 1452083 A	6