

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4012143号

(P4012143)

(45) 発行日 平成19年11月21日(2007.11.21)

(24) 登録日 平成19年9月14日(2007.9.14)

(51) Int. Cl.	F I	
G06F 3/02 (2006.01)	G06F 3/02	310K
G06F 3/16 (2006.01)	G06F 3/16	320A
G06F 17/22 (2006.01)	G06F 3/16	320H
G10L 15/00 (2006.01)	G06F 17/22	503
G10L 15/24 (2006.01)	G06F 17/22	520S
請求項の数 10 (全 17 頁) 最終頁に続く		

(21) 出願番号	特願2003-418645 (P2003-418645)	(73) 特許権者	000001007
(22) 出願日	平成15年12月16日(2003.12.16)		キヤノン株式会社
(65) 公開番号	特開2005-182208 (P2005-182208A)		東京都大田区下丸子3丁目30番2号
(43) 公開日	平成17年7月7日(2005.7.7)	(74) 代理人	100076428
審査請求日	平成15年12月16日(2003.12.16)		弁理士 大塚 康德
前置審査		(74) 代理人	100112508
			弁理士 高柳 司郎
		(74) 代理人	100115071
			弁理士 大塚 康弘
		(74) 代理人	100116894
			弁理士 木村 秀二
		(74) 代理人	100130409
			弁理士 下山 治
		(74) 代理人	100134175
			弁理士 永川 行光
			最終頁に続く

(54) 【発明の名称】 情報処理装置およびデータ入力方法

(57) 【特許請求の範囲】

【請求項1】

文字列を入力する文字列入力デバイスと、音声を入力する音声入力デバイスと、文字列を表示する表示デバイスとを備える情報処理装置であって、

前記文字列入力デバイスにより入力された文字列に続く文字列候補を予測する予測手段と、

前記予測手段により予測された文字列候補に同音異表記語が含まれる場合、当該同音異表記語の各々が別のグループに分かれるように前記文字列候補を複数のグループに分類する分類手段と、

前記予測手段により予測された文字列候補を、前記分類手段により分類されたグループ毎に、前記表示デバイスに表示させる表示制御手段と、

前記表示デバイスに表示されユーザによって選択されたグループにおける文字列候補に認識対象語を絞って、前記音声入力デバイスにより入力された音声に対し音声認識を行う音声認識手段と、

前記音声認識手段による認識結果である文字列を使用文字列として確定する確定処理手段と、

を有することを特徴とする情報処理装置。

【請求項2】

使用者に認識結果の正誤を確認させるために、前記表示デバイスに表示されている文字列候補のうち、前記音声認識手段により認識結果とされた文字列を強調表示させ、この状

10

20

態で前記音声入力デバイスにより再度入力された発声に対して音声認識を実行するよう、前記表示制御手段および前記音声認識手段を制御する制御手段を更に有し、

前記確定処理手段は、前記制御手段による制御下において、前記文字列入力デバイスにおける所定のボタンが押されたことを検知した時点における認識結果を使用文字列として確定することを特徴とする請求項1に記載の情報処理装置。

【請求項3】

前記音声認識手段は、前記予測手段により予測された文字列候補に同音異表記語が含まれる場合で、かつ、少なくともいずれかの同音異表記語に別の読みがある場合には、該別の読みがある同音異表記語については、該別の読みで音声認識を行うことを特徴とする請求項1又は2に記載の情報処理装置。

10

【請求項4】

前記表示制御手段は、前記文字列候補に対応する読みを表示させることができ、前記予測手段により予測された文字列候補に同音異表記語が含まれる場合で、かつ、少なくともいずれかの同音異表記語に別の読みがある場合は、該別の読みがある同音異表記語については、該別の読みを選択して表示させることを特徴とする請求項1から3までのいずれか1項に記載の情報処理装置。

【請求項5】

文字列を入力する文字列入力デバイスと、音声を入力する音声入力デバイスと、文字列を表示する表示デバイスとを備える情報処理装置におけるデータ入力方法であって、

前記文字列入力デバイスにより入力された文字列に続く文字列候補を予測する予測ステップと、

20

前記予測ステップで予測された文字列候補に同音異表記語が含まれる場合、当該同音異表記語の各々が別のグループに分かれるように前記文字列候補を複数のグループに分類する分類ステップと、

前記予測ステップで予測された文字列候補を、前記分類ステップにより分類されたグループ毎に、前記表示デバイスに表示させる表示制御ステップと、

前記表示デバイスに表示されユーザによって選択されたグループにおける文字列候補に認識対象語を絞って、前記音声入力デバイスにより入力された音声に対し音声認識を行う音声認識ステップと、

前記音声認識ステップによる認識結果である文字列を使用文字列として確定する確定処理ステップと、

30

を有することを特徴とするデータ入力方法。

【請求項6】

使用者に認識結果の正誤を確認させるために、前記表示デバイスに表示されている文字列候補のうち、前記音声認識ステップで認識結果とされた文字列を強調表示させ、この状態で前記音声入力デバイスにより再度入力された発声に対して音声認識を実行するよう、前記表示制御ステップおよび前記音声認識ステップにおける各処理を制御する制御ステップを更に有し、

前記確定処理ステップは、前記制御ステップによる制御下において、前記文字列入力デバイスにおける所定のボタンが押されたことを検知した時点における認識結果を使用文字列として確定することを特徴とする請求項5に記載のデータ入力方法。

40

【請求項7】

前記音声認識ステップは、前記予測ステップで予測された文字列候補に同音異表記語が含まれる場合で、かつ、少なくともいずれかの同音異表記語に別の読みがある場合には、該別の読みがある同音異表記語については、該別の読みで音声認識を行うことを特徴とする請求項5又は6に記載のデータ入力方法。

【請求項8】

前記表示制御ステップは、前記文字列候補に対応する読みを表示させることができ、前記予測ステップで予測された文字列候補に同音異表記語が含まれる場合で、かつ、少なくともいずれかの同音異表記語に別の読みがある場合は、該別の読みがある同音異表記語に

50

ついては、該別の読みを選択して表示させることを特徴とする請求項 5 から 7 までのいずれか 1 項に記載のデータ入力方法。

【請求項 9】

文字列を入力する文字列入力デバイスと、音声を入力する音声入力デバイスと、文字列を表示する表示デバイスとを備えるコンピュータが実行可能なプログラムであって、

前記文字列入力デバイスにより入力された文字列に続く文字列候補を予測する予測ステップのコードと、

前記予測ステップで予測された文字列候補に同音異表記語が含まれる場合、当該同音異表記語の各々が別のグループに分かれるように前記文字列候補を複数のグループに分類する分類ステップのコードと、

10

前記予測ステップで予測された文字列候補を、前記分類ステップにより分類されたグループ毎に、前記表示デバイスに表示させる表示制御ステップのコードと、

前記表示デバイスに表示されユーザによって選択されたグループにおける文字列候補に認識対象語を絞って、前記音声入力デバイスにより入力された音声に対し音声認識を行う音声認識ステップのコードと、

前記音声認識ステップによる認識結果である文字列を使用文字列として確定する確定処理ステップのコードと、

を含むことを特徴とするプログラム。

【請求項 10】

請求項 9 に記載のプログラムを格納したコンピュータ読み取り可能な記憶媒体。

20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、キー・ボタン操作と音声認識機能とを併用したデータ入力技術に関する。

【背景技術】

【0002】

データ入力は現在の多くの情報機器を利用する上で必要となる作業である。

【0003】

我々が日常的に利用している情報機器でおそらく最も多く入力するデータは文字であり、文字の入力は一般に、キーボード操作など煩雑な入力操作を伴う。特にキーやボタンなどの数が限られた携帯電話や PDA などの小型携帯端末での文字入力は、多数のキーを備えるパーソナル・コンピュータ等に比べ、キーやボタンを押す入力操作回数が多くなる。

30

【0004】

このような煩わしい文字列入力を効率化する目的で、文字列を途中まで入力すると、入力した文字列に続く文字列の候補を予測して提示する入力予測機能（オートコンプリート機能、自動入力補完機能などともよばれる）を備えた文字入力方法が実現されている。

【0005】

入力予測機能を用いた場合、提示された文字列の候補の中に所望の文字列があれば、その文字列を選択することによって文字列の入力を完了することができる。予測性能が高ければ、全ての文字列を入力せずに文字列入力を済ませることができる。さらに候補の文字列としてかな漢字変換後の候補文字列を提示することにより、かな漢字変換の操作も省略することが可能である。

40

【0006】

このような文字入力支援のための文字列の入力予測に関する技術は多数提案されている（例えば、特許文献 1、2 および非特許文献 1 を参照。）。

【0007】

一方、上記した入力予測とは異なる文字列入力の支援技術として、音声認識がある。音声認識を用いると、基本的に文字列入力のためのキー操作が不要になるため、キー操作に不馴れな人には効率的な文字列の入力が可能である。また、音声認識はキーの数が限られた機器での文字列の入力で効果を発揮する。

50

【0008】

【特許文献1】特開平08-235318号公報

【特許文献2】特開平08-255158号公報

【非特許文献1】「POBox (Predictive Operation Based On eXample)」, URL: <http://www.csl.sony.co.jp/person/masui/OpenPOBox/index.html>

【発明の開示】

【発明が解決しようとする課題】

【0009】

入力予測技術は、近年高性能化している携帯電話やPDAなどの小型携帯端末上でも実現されており、非常に便利な機能であるが、提示された文字列の候補が複数ある場合には、所望の文字列を選択するための操作が煩雑になる場合がある。特にカーソル移動操作やスクロール操作のみで文字列を選択する場合には、所望の文字列に到達するまでカーソル位置移動の操作を繰り返す必要があり、候補が多数提示されている場合は操作回数が多くなる。

10

【0010】

一方、音声認識の技術は、近年性能の改善が進み、数万語を扱えるディクテーションソフトウェアにより、オフィスなどの比較的静かな環境では、高性能なコンピュータ上での文字列の入力が快適に行うことができるようになってきた。しかしながら、数万語を扱う音声認識では、多くの計算機リソース(CPU、メモリ)が必要なため、高性能化の進んだ最近の小型携帯端末でも快適な動作は望めない。また、現在のディクテーションソフトウェアでは背景雑音が大きい場所での認識性能は十分とは言えず、小型携帯端末がよく利用される屋外では、本来の性能を発揮できない。

20

【0011】

携帯情報端末の使用環境、リソースを考慮すると、使用者にストレスを与えることのない応答速度を実現するには、認識語数をできるだけ少なくするのが最善の方法であるが、単純に語数を少なくするだけでは認識率が落ち、修正操作なしでは所望の文字列を入力できなくなってしまう。このように現在の技術で、小型携帯端末上で音声認識だけを用いて快適な文字列の入力を行なうのはやや困難と思われる。

【0012】

また別の課題として、音声のみでは同音異表記語の識別ができないという課題もある。例えば、「きやのん」という音声入力に対して、「キヤノン」と「CANON」のいずれの表記(語義は同じ)を採るかという問題は、その音声入力からだけでは判断することができない。

30

【0013】

本発明は、このような課題に鑑みてなされたもので、文字列の予測機能と音声認識とを併用することにより、快適なデータ入力を実現することを目的としている。

【課題を解決するための手段】

【0014】

本発明の側面は、文字列を入力する文字列入力デバイスと、音声を入力する音声入力デバイスと、文字列を表示する表示デバイスとを備える情報処理装置に係り、次の構成を有することにより上記した課題を解決する。すなわち、この情報処理装置は、前記文字列入力デバイスにより入力された文字列に続く文字列候補を予測する予測手段と、前記予測手段により予測された文字列候補に同音異表記語が含まれる場合、当該同音異表記語の各々が別のグループに分かれるように前記文字列候補を複数のグループに分類する分類手段と、前記予測手段により予測された文字列候補を、前記分類手段により分類されたグループ毎に、前記表示デバイスに表示させる表示制御手段と、前記表示デバイスに表示されるユーザによって選択されたグループにおける文字列候補に認識対象語を絞って、前記音声入力デバイスにより入力された音声に対し音声認識を行う音声認識手段と、前記音声認識手段による認識結果である文字列を使用文字列として確定する確定処理手段とを有することを特徴とする。

40

50

【0015】

本発明の別の側面は、文字列を入力する文字列入力デバイスと、音声を入力する音声入力デバイスと、文字列を表示する表示デバイスとを備える情報処理装置におけるデータ入力方法に係り、前記文字列入力デバイスにより入力された文字列に続く文字列候補を予測する予測ステップと、前記予測ステップで予測された文字列候補に同音異表記語が含まれる場合、当該同音異表記語の各々が別のグループに分かれるように前記文字列候補を複数のグループに分類する分類ステップと、前記予測ステップで予測された文字列候補を、前記分類ステップにより分類されたグループ毎に、前記表示デバイスに表示させる表示制御ステップと、前記表示デバイスに表示されユーザによって選択されたグループにおける文字列候補に認識対象語を絞って、前記音声入力デバイスにより入力された音声に対し音声認識を行う音声認識ステップと、前記音声認識ステップによる認識結果である文字列を使用文字列として確定する確定処理ステップとを有することを特徴とする。

10

【発明の効果】

【0016】

本発明によれば、文字列の予測機能と音声認識とを併用することにより、快適なデータ入力が実現される。

【発明を実施するための最良の形態】

【0017】

以下、図面を参照して本発明の好適な実施形態について詳細に説明する。

【0018】

20

(実施形態1)

以下では、本発明のデータ入力装置を携帯電話機に適用した例を示す。もっとも、本発明は携帯電話機に限定されるものではなく、PDA等の携帯型端末をはじめ、パーソナルコンピュータ等のユーザからの文字入力を要するあらゆる情報処理装置に適用することが可能である。

【0019】

図1は、本実施形態における携帯電話機の構成を示すブロック図である。

【0020】

図1において、101は制御メモリ（ROM）、102は中央処理装置（CPU）、103はメモリ（RAM）、104は外部記憶装置、105はキーやボタンなどの入力装置、106は液晶モニタなどの表示装置、107は音声入力装置（マイク）、108は音声出力装置（スピーカ）、109はバスである。図示のように、外部記憶装置104には、本実施形態の携帯電話機を実現するための制御プログラム110、文字列予測のために用いる文字列予測データ209、音響モデルをはじめとする音声認識を実行する際に必要な音声認識用データ210などが記憶される。文字列予測データ209は、例えば、かな漢字変換辞書や使用者の文字入力履歴などを利用して構成される。これらの制御プログラム110やデータ209、210は、CPU102の制御の下、バス109を介してRAM103に取り込まれ、CPU102によって実行される。言うまでもないことであるが、これらは外部記憶装置104ではなくROM101に記憶してもよい。

30

【0021】

40

図2は、本実施形態における携帯電話機の文字列入力に係る処理の機能構成を示すブロック図である。

【0022】

操作入力部201は、使用者が入力装置105を用いて行なった文字入力をはじめとするボタンなどの操作を検出する。

文字列候補予測部202は、使用者が入力した文字列に続く文字列の候補を文字列予測用データ209を参照して予測する。

提示方法決定部203は、予測された文字列の候補の提示方法を決定する。

候補分類部204は決定された提示方法に従い、予測された文字列の候補を複数のグループに分類する。

50

文字列候補提示部 205 は決定された提示方法に従い文字列の候補を表示装置 106 に表示する。

別候補通知部 206 は候補を複数のグループに分けて提示する場合に、使用者に提示中の文字列候補以外に候補があることを通知する。

候補切り替え部 207 は候補を複数のグループに分けて提示する場合に、使用者が提示する文字列の候補を他のグループと切り替える操作を検出し、提示する候補のグループを切り替える。

音声認識部 208 は、文字列候補提示部 205 によって提示された文字列の候補を認識対象語とする音声認識を行なう。認識対象語の読みについては、文字列予測用データ 209 を参照して決定する。

10

音声合成部 211 は使用者へのデータの提示や各種通知を音声で行う場合に合成音を生成する。

文字列選択方法提示部 212 は、提示した文字列候補から使用者が所望の文字列を選択する方法を使用者に提示する。

選択方法決定部 213 は使用者が文字列の選択方法を選択する操作を検出し、文字列の選択方法を決定する。

制御部 200 は以上の各モジュールを制御し、文字入力に係る全体の処理を統括する。

【0023】

図 3 は入力装置 105 におけるボタン配置例を示す図である。

【0024】

20

301、302 は同心円上に配置されたボタンで、外輪の 301 は主にカーソル移動の移動方向を上下左右で指示する操作に用いられる。以降の説明では『↑』『↓』『←』『→』と表記する。内輪中央ボタン 302 は主に文字列選択の際に選択した候補を確定する操作に用いられる。以降『●』と表記する。303 乃至 306 は文字列処理の処理過程の状態の遷移によって役割が変化するボタンであり、以降、それぞれ『♪』『★』『*』『#』と表記する。

【0025】

以下、本実施形態における文字列入力処理を、図 4、5 を用いて説明する。図 4 は、本実施形態における文字列入力処理の流れを示すフローチャート、図 5 は、この文字列入力処理中における表示装置 106 の表示内容の遷移を示す図である。なお、文字列の候補の予測や音声認識については、既存の技術を使用することができるため詳細な説明は省略する。

30

【0026】

ここでは、使用者が『今日はどうもありがとう』という文字列を入力する場合について説明するが、既に使用者が『今日はどうも』までを入力し終え、続く『ありがとう』と入力しようとしている状態であると仮定する。

【0027】

『今日はどうも』まで入力した段階では、表示装置 106 の表示内容は図 5 の 510 のようになっている。

【0028】

40

使用者は、『ありがとう』と入力するため最初の文字『あ』を入力する（ステップ S401）。操作入力部 201 がこの『あ』の入力を検出すると、文字列候補予測部 202 は、文字列予測用データ 209 を参照して、『あ』に続く文字列の候補を予測する（ステップ S402）。文字列予測用データ 209 は、上述したように、例えば、ひらがなと漢字の対応関係を示したかな漢字変換辞書や使用者の過去の文字入力履歴などを利用したものである。また、携帯電話では一般に『あ』以外の『あ』行の文字『い』『う』『え』『お』を入力する場合も『あ』のボタンを複数回押して入力するため、『あ』を 1 回押した段階で、『あ』行全ての文字で始まる文字列を文字列候補として予測しても良い。

【0029】

予測した文字列候補は文字列候補提示部 205 によって出力装置 106 に提示される（

50

ステップS403)。この際、提示方法決定部203が文字列の提示順を規定してもよい。例えば、文字列予測用データ209に使用者の過去の文字入力履歴を用いる場合、これを参照して文字列の使用頻度順に表示してもよいし、あるいは最も最近に入力された文字列から順に表示してもよい。また、予測した文字列の候補数が多数の場合は、前述したような基準を用いて提示順を決め、所定の数の文字列候補だけを表示するようにしてもよい。また、別の方法として文字列候補を提示する画面領域のサイズから提示可能な文字列候補数を割り出し、その数だけ文字列候補を表示してもよい。また、ステップS403では、提示した文字列候補の選択方法と同時に、文字列選択方法提示部212により、文字列選択の方法も提示される。

【0030】

10

ここでは『あ』の入力に対して、ステップS402で、文字列候補として『アイディア』『会う』『会えない』『朝ごはん』『ありがとう』『合わない』を予測したものとして説明を続ける。また、本実施形態では、音声認識による文字列選択と、ボタン操作による文字列選択が併用できる場合を説明する。

【0031】

図5の520に、予測した文字列候補を提示した様子を示す。この例では、画面を二分割し、文字入力画面521に入力中の文字列を、文字列選択画面522に文字列候補を表示する。なお、文字列の候補の表示に伴い、文字列選択方法提示部212により文字列の選択方法を指定するための操作ボタンガイド523を表示する。また、現在の操作対象が文字入力を行なっている画面である示すため文字入力画面の背景色を変えるなどしてハイ
20 ライトする。以上のような文字列候補の提示や操作ボタンガイドの追加表示については、提示方法決定部203が制御する。

【0032】

ステップS404において、提示された文字列候補中に使用者の所望する文字列がない場合は、ステップS401に戻ってさらに続けて文字列を入力し、所望の文字列がある場合にはステップS405に進む。

【0033】

ステップS405において、使用者は所望の文字列を選択する操作へと移る。使用者が音声認識を利用する場合はステップS409へ、ボタン操作を利用する場合はステップS406へと進む。
30

【0034】

まず、ボタン操作で選択する場合について説明する。使用者はボタン操作による文字列選択を選択するための操作を行なう（ステップS406）。この操作は選択方法決定部213で検出し、以降の文字列の選択をボタン操作により実行する。本実施形態では、図3の304で示した『★』ボタンを押すことでボタン操作が選択される。図5の530に、ボタン選択を選んだ場合の表示内容を示す。530では操作対象が文字列候補を表示した画面に移ることを示すため、画面のハイライトが文字入力画面から文字列選択画面へと移り、文字選択画面の最初の文字列候補『アイディア』の位置にカーソルを表示している。また、この時点で『★』ボタンの機能を変更し、『戻る（操作対象を文字入力画面に戻す）』にする。
40

【0035】

続いて使用者はボタン操作により所望の文字列を選択する（ステップS408）。図5において、目的の文字列『ありがとう』を選択するためには、使用者は図3の301の『↓』『→』ボタンを押してカーソル位置を『ありがとう』に移動する。この間の画面の遷移を540、550に示す。

【0036】

次に、ステップS408で、使用者は使用文字列を確定する操作を行なう。550を表示している状態で使用者が302の『●』ボタンを押すと選択中の文字列『ありがとう』が使用文字列として確定される。文字列『ありがとう』を確定後の画面を560に示す。以上が予測した文字列候補をボタン操作により選択する場合の処理である。
50

【0037】

次に、ステップS405において使用者が音声認識の利用を選択した場合について説明する。使用者は音声認識による文字列選択を選択するための操作を行なう（ステップS409）。この操作は選択方法決定部213で検出し、以降の文字列の選択を音声認識により実行する。

【0038】

本実施形態では、『♪』ボタン303を押すことにより、音声認識が選択される。図5の520において、使用者が『♪』を押すと画面のハイライトが文字入力画面から文字選択画面に移る。なお、音声認識を選択した場合は、カーソル移動による文字選択を行わないため、文字列選択画面にカーソルを表示しない。次に使用者が所望の文字列『ありがとう』を発声する（ステップS410）。音声認識部208はこの使用者の発声を音声認識し（ステップS411）、得られた認識結果の文字列を使用文字列として確定する（ステップS412）。ステップS412における音声認識は、文字列候補提示部205によって提示された文字列のみを認識対象語とする。また、これらの認識対象語の読みは、文字列予測用データ209を参照して音声認識部205が決定する。音声認識により文字列が選択される場合の表示画面の遷移を図5の570から560への遷移で示す。

10

【0039】

このように、本実施形態によれば、入力装置105を用いて入力された文字に続く文字列候補が表示され、音声認識によってこの文字列候補の中から使用文字列を選択することができ、煩雑なボタン操作を大幅に減らすことができる。また、ここで行われる音声認識は表示された文字列候補だけを認識対象語とするので、この音声認識処理は少ない演算量で済む。したがって、このような音声認識を例えば小型携帯端末等で実現する場合にも、高い認識率を維持しつつ十分高速に動作させることが可能となる。

20

【0040】

(実施形態2)

上述の実施形態1では、ステップS412において、ステップS411の音声認識結果の確認をユーザが行なうことなく使用文字列を確定する例を示したが、この場合には誤って所望の文字列とは異なる文字列を音声認識した場合は誤った文字列が使用文字列として確定されてしまう。これを回避するためには、認識結果を確認するステップが必要となる。そこで本実施形態では、認識結果の確認を行ってから使用文字列を確定する例を、図6

30

【0041】

図6は、認識結果の確認を行ってから使用文字列を確定する場合の処理の流れを示すフローチャートであり、図4のフローチャートにおけるD-E間で行われるステップS410～S412に代わる処理ブロックのみを示している。その他の処理ブロックの内容は実施形態1と同様であるので説明を省略し、図4と相違するD-E間における処理部分のみを説明する。

【0042】

実施形態1における説明と同様に、使用者が所望の文字列『ありがとう』と発声すると（ステップS601）、音声認識部208はこれを音声認識し（ステップS602）、認識結果を提示する（ステップS603）。これにより使用者はその結果の正誤を判断できる（ステップS604）。認識結果が誤っている場合は、ステップS601に戻って再度所望の文字列『ありがとう』を発声する（ステップS601）。正しい認識結果が得られないかぎり、ステップS601～S604の各ステップを正しい認識結果が得られるまで繰り返し実行することになる。ステップS604において正しい認識結果が得られた場合は、使用者が文字列を確定する操作を行ない、得られた認識結果で文字列を確定する（ステップS605）。

40

【0043】

本実施形態を用いた場合の画面表示の遷移を図7に示す。

【0044】

50

図7は、実施形態1と同様に『今日はどうもありがとう』の『ありがとう』を入力する場合に、文字列選択の際の音声認識が認識誤りを起こした場合を示している。

【0045】

710は『今日はどうも』までが確定された状態で、ここで使用者が『あ』を入力すると実施形態1と同様に『あ』から予測される文字列の候補が提示される(720)。ここで、使用者が文字列選択に音声認識を利用するため『♪』ボタンを押して音声認識が開始される(730)。使用者の『ありがとう』の発声に対し、誤って『朝ごはん』が認識された場合のステップS603における表示が740である。この表示例では、提示していた文字列候補のうち認識結果に対応する文字列にカーソルを移動する(文字列にアンダーラインが引かれる)ことで認識結果を提示する例である。認識結果の提示方法は、これに限るものではなく、例えば認識結果に該当する文字列を反転表示して認識結果を提示しても良い。提示した認識結果『朝ごはん』は目的の『ありがとう』ではないので使用者は再度『ありがとう』の発声を行なう(S604、NO→S601)。二回目の発声が正しく認識され認識結果として『ありがとう』を提示した様子を750に示す(S602、S603)。認識結果が正しいので使用者は、『●』ボタンを押して文字列を確定する(S604、YES→S605)。文字列が確定すると実施形態1と同様に、表示されていた予測した文字列候補を表示する画面が消滅し、文字を入力する画面に戻る(760)。

10

【0046】

このように、本実施形態によれば文字列選択の際の音声認識が認識誤りを起こした場合に、正しく認識されるまで使用者が再度発声することができ、認識誤りの修正を容易に行うことができる。

20

【0047】

(実施形態3)

上述の実施形態では、予測した文字列候補を全て提示するか、あるいは候補が多数ある場合は所定数の文字列候補を提示するようにしていただけて、予測した文字列候補の中に同じ読みをもつ文字列候補が複数ある場合の提示方法については特に考慮されていない。本実施形態では、これらを考慮した文字列候補の提示について説明する。

【0048】

本実施形態の特徴は、予測した文字列候補が多数ある場合や、文字列候補の中に同じ読みを持つ文字列候補がある場合に、文字列候補を複数のグループに分類し、複数回に分けて候補を提示する点にある。以下、図8を用いてその処理を詳しく述べる。

30

【0049】

図8は、本実施形態における文字列候補を提示する処理の流れを示すフローチャートである。このフローチャートは図4の一部を示したもので、図4のフローチャートにおけるA-B間のステップS403、S404に代わる処理ブロックのみが示されている。その他の処理ブロックの内容は実施形態1と同様であるので説明を省略し、図4と相違するA-B間の処理部分だけを説明する。なお、本実施形態と、実施形態2で説明した処理を組み合わせられることは言うまでもない。

【0050】

まず、ステップS801では、図4のステップS402で予測した文字列候補の数が所定数Nよりも多いか否かを判断する。ここで、文字列候補数がNよりも多い場合は、文字列候補を複数に分けて提示するべくステップS803以降の処理を行なう。ここでNは一度に提示する候補数であるが、この数はあらかじめ決めておいてもいいし、予測された文字列候補の文字数および提示可能な表示領域のサイズから提示可能な候補数を文字列の予測の都度求めてもよい。

40

【0051】

候補数がN以下の場合はステップS802に進む。ここでは同音異表記語があるかどうかを判断する。例えば、文字列候補の中に『日本(にほん)』『二本(にほん)』のように読みが同じ文字列候補がある場合には、音声認識では一意に文字列を選択することができなくなるので、ステップS803以降の処理で文字列候補を複数に分けて提示する処理

50

を行なう。以上の文字列候補の提示方法の決定は提示方法決定部203で行う。なお、文字列候補数がN以下でかつ文字列候補中に同じ読みを持つ文字列が存在しない場合にはステップS808、S809に進むが、これらはそれぞれ図4のS403、S404と同様の処理であるから説明は省略する。上記したステップS801およびS802の判断処理は、提示方法決定部203の処理による。

【0052】

ステップS803では、候補分類部204により文字列候補を複数のグループに分類する。分類の際は、例えば、使用頻度の高い文字列候補から順にN個ずつを抽出してグループ分けしてもよいし、文字列候補を五十音順に並べ替えて順にN個ずつ抽出してグループ分けするようにしてもよい。ただし、同一グループ内に同じ読みの文字列候補が含まれないように分類する。また別の方法として、後段の処理で行なう音声認識の精度を高めるため、各グループ内の候補文字列の音響的な分離度が高くなるような分類基準を用いることが好ましい。

10

【0053】

次にステップS804で、使用者に提示するグループを選択する。この時、使用頻度順に分類した場合には最も使用頻度の高いグループを選択する。五十音順で分類した場合も同様に五十音順の最初のグループを選択する。音響的な分離度を基準にした場合には音響的な分離度が最も高いグループを選択する。

【0054】

続いてステップS805に進み、文字列候補提示部205が選択されたグループの文字列候補を表示装置106上に提示するとともに、別候補通知部206が提示した文字列候補以外にも文字列候補があることを使用者に対して通知する。また、実施形態1と同様に文字列の選択方法を文字列選択方法提示部212により提示する。

20

【0055】

ステップS806において、使用者は提示された文字列候補に所望の文字列があるかどうかを判断する。ここで、提示された文字列候補内に所望の文字列がある場合には図4のS405へ進み、実施形態1と同様に提示した文字列候補の中から所望の文字列を選択する操作を行なう。一方、提示された文字列候補内に所望の文字列がない場合にはステップS807において、使用者は他のグループを選択するか、図4のステップS401に戻って次の文字の入力を行う。ここで使用者が他のグループを選択した場合は、候補切り替え部207が使用者のグループ選択の操作を検出し、提示する候補を使用者が選択したグループに切り替え、S805に戻り同様の手順を繰り返す。

30

【0056】

以上の実施形態で述べた手順にしたがい、予測した文字列候補内に同じ読みを持つ文字列候補がある場合の文字列候補の提示例および提示する候補を切り替える様子を図9に示す。

【0057】

図9は、使用者が『明日の日本』と入力しようとして、『明日の』に続いて『日本』を入力するために『に』と入力した場合の文字列候補の提示例を示している。ここで、一度に提示する文字列の候補数Nは、N=8に設定されているものとする。

40

【0058】

910は『明日の』までが確定した様子を示している。続いて使用者が入力した『に』から5つの文字列候補『日本（にほん）』『二本（にほん）』『にんじん』『日程』『日時』が得られたとする（S401～S402）。N=8に設定されているため、ステップS802に移る。予測した文字列候補の中に読みが同じ『にほん』になる『日本』と『二本』の二つの文字列候補があるため（ステップS802）、『日本』と『二本』が異なるグループになるよう二つのグループを五十音順に『日時、日程、二本、にんじん』と『日本』とに分類する（ステップS803）。

【0059】

五十音順で先のグループである『日時、日程、二本、にんじん』を提示するグループと

50

して選択し（ステップS804）、使用者に選択された文字列候補を提示する（ステップS805）。同時に別候補通知部206により提示した文字列候補以外に候補があることを通知する（ステップS805）。920にその様子を示す。921の『#次の候補』は、別候補通知部206による通知の例である。

【0060】

この時点では、所望の文字列『日本』が提示されていないので、使用者は『#』を押して、他の候補を見るため別の候補を選択する（ステップS806）。使用者の候補切り替えの操作を候補切り替え部207が検出し、使用者が選択した次の候補すなわち次のグループを提示するグループ（『日本』）として選択する（ステップS804）。『日本』を使用者に提示した例が930である。提示している『日本』以外に最初に提示したグループの文字列候補があるので、921で示される『#次の候補』の他に922で示される『*前の候補』を表示することにより、他の文字列候補があることを提示する（ステップS805）。以降、実施形態1または実施形態2で示した手順により提示した文字列候補から文字列の選択・確定処理が行われる。

10

【0061】

このように、本実施形態によれば、予測した文字列候補の中に同音異表記語が含まれる場合には、同音異表記語がそれぞれ異なるグループに分かれるように文字列候補をグループ分けして、グループ毎に文字列候補を提示するようにしたので、グループ毎に認識結果を一意に定めることができ、ユーザの選択操作を省くことができる。

【0062】

20

また、この他に、上述の実施形態において、予測した文字列候補を提示する際に、予測した文字列候補中に『日本（にっぽん）』と『日本（にほん）』のように読みが異なる同一の文字列候補が複数ある場合には、提示方法決定部203において、一つだけを選択して提示するようにしてもよい。また、その際、他の文字列候補の読みとの音響的な分離度が高い読みを持つ候補を選択するようにしてもよい。例えば、『日本（にっぽん）』、『日本（にほん）』、『二本（にほん）』が予測された文字列候補であった場合、『日本（にほん）』と『二本（にほん）』が同じ読み（音響的には全く同じであるため、音響的な分離度は0）なので、『日本（にほん）』と『日本（にっぽん）』とのうちから『日本（にっぽん）』を選択し、文字列候補として『日本（にっぽん）』と『二本（にほん）』を提示するようにする。

30

【0063】

このような処理を行うことによって、ユーザの選択操作を省くことができる。

【0064】

（その他の実施形態）

以上の実施形態では、予測した文字列候補を提示する際に、文字列のみを提示したが、これに限るものではなく、文字列候補とともに各文字列候補の読みを併せて提示しても良い。また、文字列候補中にかな文字以外が含まれる場合のみ読みを提示するようにしても良い。また、かなのみの文字列候補であっても表記と読みが異なる場合は読みを提示するようにしてもよい。

【0065】

40

図10に文字列候補と読みを提示する例を示す。1001は、全ての文字列候補について読みを提示した場合の例である。1002は、文字列候補中にひらがな・カタカナ文字以外が含まれる場合だけ読みを提示した例である。例えば、文字列『アイディア』は、全てカタカナ文字で構成されるので読みを提示しない。ただし例外もある。1003は、ひらがな・カタカナのみの文字列候補であっても表記と読みが異なる場合は読みを提示した例である。例えば、文字列『キャノン』はカタカナ文字のみで構成されるが、読みが『きゃのん』なので読みを提示する。

【0066】

以上説明した実施形態では、文字列候補の提示および認識結果の提示および別候補があることの通知を表示装置106に表示する場合を説明したが、これに限るものではなく、

50

音声合成部 2 1 1 により音声合成して音声出力装置 1 0 8 から音声を用いて提示してもよい。

【0067】

また、以上説明した実施形態では、本発明の一実施形態として文字列入力に関して説明を行ったが、これに限るものでなく、例えば画像データの入力を画像に付与された名前を入力する場合など、他の形態のデータを入力する装置においても実施可能である。

【0068】

以上説明したように、本発明のデータ入力装置は、入力可能なデータの予測と音声認識を組み合わせ、予測されたデータ候補から音声認識により所望のデータを選択することにより、データの予測のみ、あるいは音声認識のみを用いた場合のデータ入力に比べ効率の良いデータ入力が可能となる。

10

【0069】

また、予測されたデータ候補内に同じ読みのデータが複数ある場合に、これらのデータを複数のグループに分けて提示し、一度に提示するデータ候補中に同じ読みを持つデータをなくすことにより、音声認識により一意に所望のデータを選択できるようになり、データ入力に音声認識を用いた際の利便性が向上する。

【0070】

(他の実施形態)

以上、本発明の実施形態を詳述したが、本発明は、前述した実施形態の機能を実現するソフトウェアのプログラムを、システムあるいは装置に直接あるいは遠隔から供給し、そのシステムあるいは装置のコンピュータがその供給されたプログラムコードを読み出して実行することによっても実現が可能である。その場合、プログラムの機能を有していれば、その形態はプログラムである必要はない。

20

【0071】

従って、本発明の機能処理をコンピュータで実現するために、そのコンピュータにインストールされるプログラムコード自体およびそのプログラムを格納した記憶媒体も本発明を構成することになる。つまり、本発明の特許請求の範囲には、本発明の機能処理を実現するためのコンピュータプログラム自体、およびそのプログラムを格納した記憶媒体も含まれる。

【0072】

その場合、プログラムの機能を有していれば、オブジェクトコード、インタプリタにより実行されるプログラム、OSに供給するスクリプトデータ等、プログラムの形態を問わない。

30

【0073】

プログラムを供給するための記憶媒体としては、例えば、フレキシブルディスク、ハードディスク、光ディスク、光磁気ディスク、MO、CD-ROM、CD-R、CD-RW、磁気テープ、不揮発性のメモ리카ード、ROM、DVD (DVD-ROM, DVD-R) などがある。

【0074】

その他、プログラムの供給方法としては、クライアントコンピュータのブラウザを用いてインターネットのホームページに接続し、そのホームページから本発明のコンピュータプログラムそのもの、もしくは圧縮され自動インストール機能を含むファイルをハードディスク等の記憶媒体にダウンロードすることによっても供給できる。また、本発明のプログラムを構成するプログラムコードを複数のファイルに分割し、それぞれのファイルを異なるホームページからダウンロードすることによっても実現可能である。つまり、本発明の機能処理をコンピュータで実現するためのプログラムファイルを複数のユーザに対してダウンロードさせるWWWサーバも、本発明のクレームに含まれるものである。

40

【0075】

また、本発明のプログラムを暗号化してCD-ROM等の記憶媒体に格納してユーザに配布し、所定の条件をクリアしたユーザに対し、インターネットを介してホームページか

50

ら暗号化を解く鍵情報をダウンロードさせ、その鍵情報を使用することにより暗号化されたプログラムを実行してコンピュータにインストールさせて実現することも可能である。

【0076】

また、コンピュータが、読み出したプログラムを実行することによって、前述した実施形態の機能が実現される他、そのプログラムの指示に基づき、コンピュータ上で稼動しているOSなどが、実際の処理の一部または全部を行い、その処理によっても前述した実施形態の機能が実現され得る。

【0077】

さらに、記憶媒体から読み出されたプログラムが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書き込まれた後、そのプログラムの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によっても前述した実施形態の機能が実現される。

10

【図面の簡単な説明】

【0078】

【図1】実施形態における携帯電話機の構成を示すブロック図である。

【図2】実施形態における携帯電話機の文字列入力に係る処理の機能構成を示すブロック図である。

【図3】実施形態における入力装置のボタン配置例を示す図である。

【図4】実施形態における文字列入力処理の流れを示すフローチャートである。

20

【図5】文字列入力処理中における表示装置の表示内容の遷移を示す図である。

【図6】認識結果の確認を行ってから文字列を確定する場合の処理の流れを示すフローチャートである。

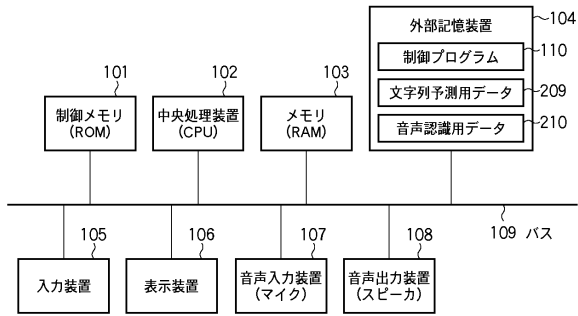
【図7】文字列選択の際の音声認識が認識誤りを起こした場合の処理を説明する図である。

【図8】実施形態3における文字列候補を提示する処理の流れを示すフローチャートである。

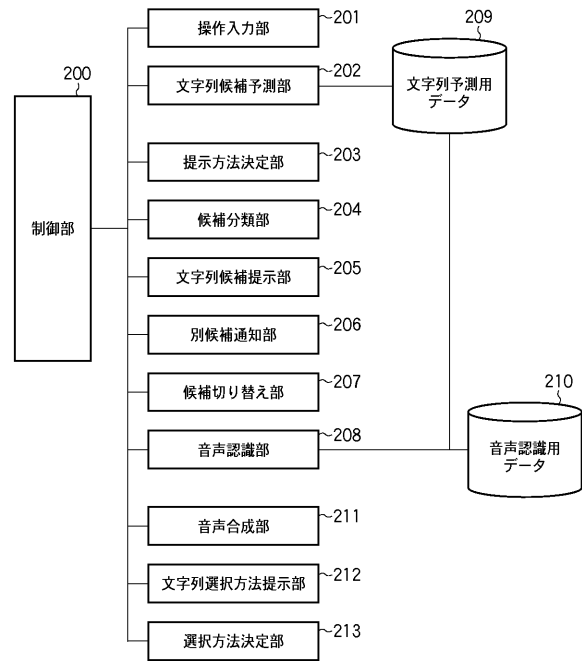
【図9】実施形態3における文字列候補の提示例を説明する図である。

【図10】文字列候補と読みを提示する例を示す図である。

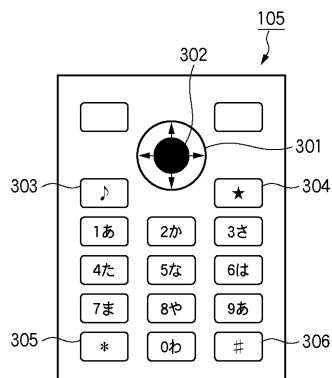
【図 1】



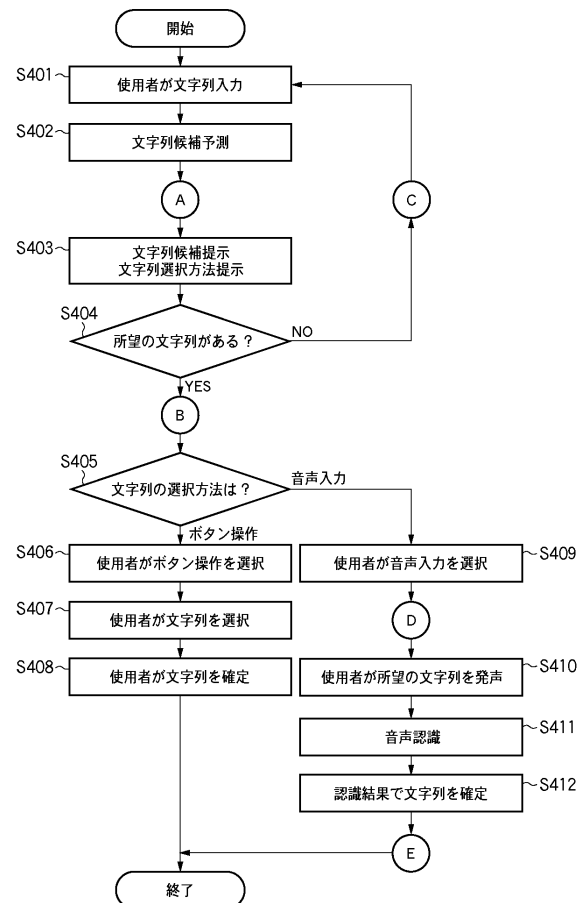
【図 2】



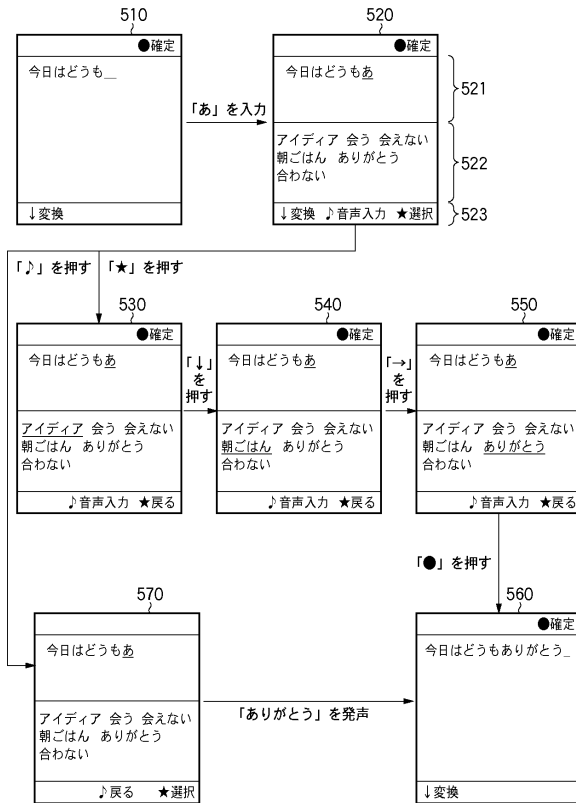
【図 3】



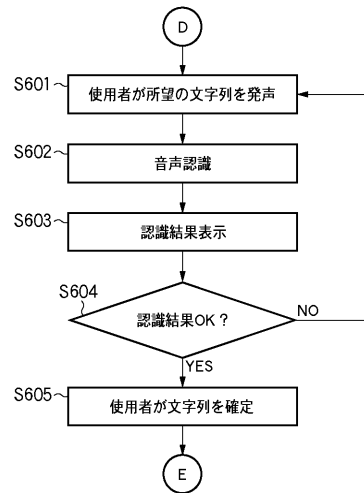
【図 4】



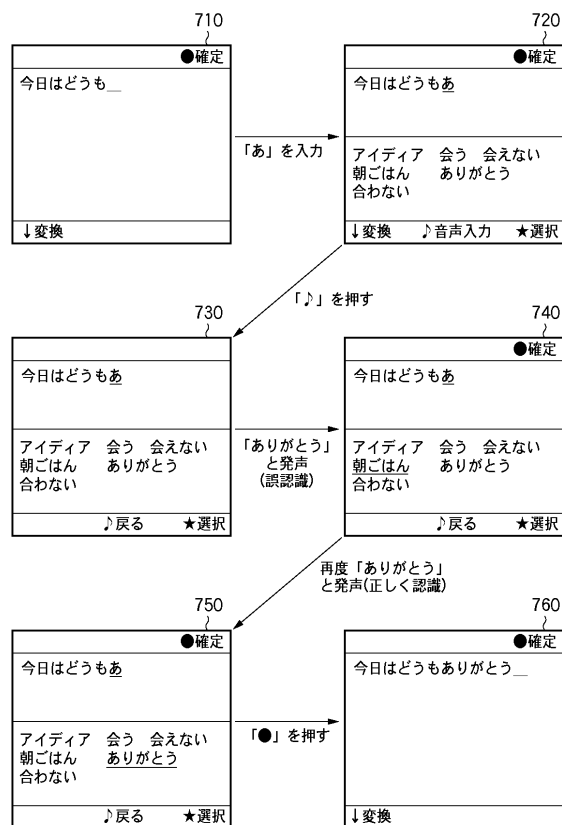
【図 5】



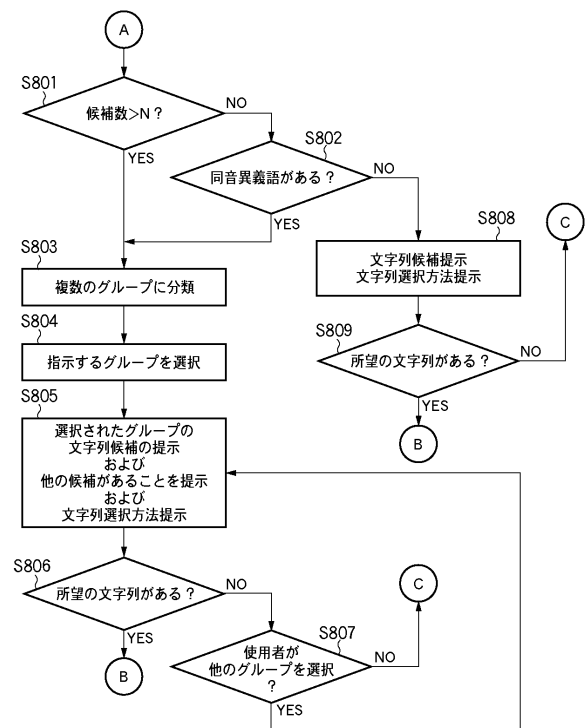
【図 6】



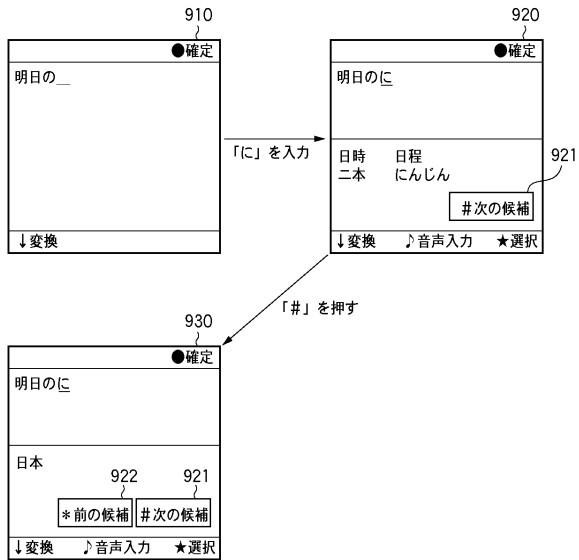
【図 7】



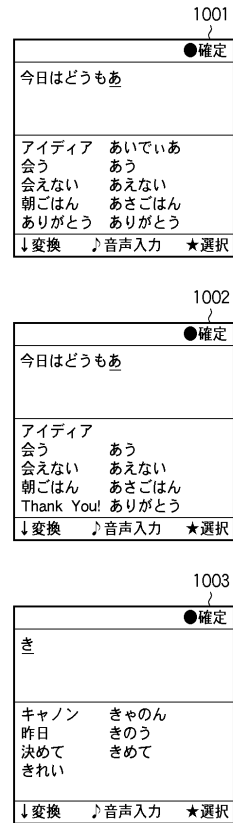
【図 8】



【図 9】



【図 10】



フロントページの続き

(51)Int.Cl. F I
G10L 15/22 (2006.01) G10L 15/00 200P
G10L 15/24 R
G10L 15/22 200V
G10L 15/22 460

(72)発明者 山本 寛樹
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

(72)発明者 小森 康弘
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 圓道 浩史

(56)参考文献 特開2001-282778 (JP, A)
特開昭62-037734 (JP, A)
特開昭60-247771 (JP, A)
特開昭61-068669 (JP, A)
特開2001-215995 (JP, A)
特開平08-202386 (JP, A)
特開平08-006940 (JP, A)

(58)調査した分野(Int.Cl., DB名)
G06F 3/02-3/027
G06F 3/16
G06F 17/22
G10L 15/00
G10L 15/22
G10L 15/24