

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3795692号

(P3795692)

(45) 発行日 平成18年7月12日(2006.7.12)

(24) 登録日 平成18年4月21日(2006.4.21)

(51) Int. Cl.		F I		
G 0 6 F	17/22	(2006.01)	G O 6 F	17/22 5 0 3
G 1 0 L	15/26	(2006.01)	G O 6 F	17/22 5 1 2 G
			G O 6 F	17/22 5 2 6 E
			G 1 0 L	15/26

請求項の数 3 (全 17 頁)

(21) 出願番号	特願平11-34934	(73) 特許権者	500046438
(22) 出願日	平成11年2月12日(1999.2.12)		マイクロソフト コーポレーション
(65) 公開番号	特開2000-242638(P2000-242638A)		アメリカ合衆国 ワシントン州 9805
(43) 公開日	平成12年9月8日(2000.9.8)		2-6399 レッドモンド ワン マイ
審査請求日	平成11年2月15日(1999.2.15)		クロソフト ウェイ
審査番号	不服2002-14165(P2002-14165/J1)	(74) 代理人	100077481
審査請求日	平成14年7月25日(2002.7.25)		弁理士 谷 義一
		(74) 代理人	100088915
			弁理士 阿部 和夫
		(74) 代理人	100129171
			弁理士 柿沼 健一
		(72) 発明者	鹿子木 宏明
			東京都調布市調布ヶ丘1丁目18番1号
			マイクロソフト株式会社 マイクロソフト
			調布技術センター内
			最終頁に続く

(54) 【発明の名称】 文字処理装置および方法

(57) 【特許請求の範囲】

【請求項1】

入力 of 音声 を 音声認識 し、音声認識結果を未確定のかな混じり漢字文字列として表示し、当該表示された文字列を文字処理した後、確定文字列を出力する文字処理装置において、

変換対象の未確定のかな混じり漢字文字列を指定する変換対象指定手段と、

当該指定された未確定のかな混じり漢字文字列に対して変換可能な1以上の文字列候補を漢字変換辞書から読み出して表示する表示手段と、

当該表示された文字列候補の1つを選択して読みの表示を指示する読み表示指示手段と

、
該読み表示指示手段の読みの表示の指示に応じて、前記読み表示指示手段により選択された文字列についての1以上の読みの候補を漢字変換辞書から読み出して表示する読み表示手段と、

前記読み表示手段により表示された1以上の読みの候補の1つを選択する第1選択手段と、

当該選択された読みの候補に対応する1以上のかな混じり漢字文字列を変換候補として漢字変換辞書から読み出して表示する変換候補表示手段と、

当該表示された1以上のかな混じり漢字文字列の1つを選択する第2選択手段と、

当該選択されたかな混じり漢字文字列で、前記変換対象指定手段により指定された未確定のかな混じり漢字文字列を置換する置換手段と

10

20

をさらに具えたことを特徴とする文字処理装置。

【請求項 2】

請求項 1 に記載の文字処理装置において、前記読み表示手段は 1 以上の読みに加えて、前記変換対象指定手段により指定された未確定のかな混じり漢字文字列をも表示することを特徴とする文字処理装置。

【請求項 3】

請求項 1 に記載の文字処理装置において、該文字処理装置は、再変換後の指定文字列について他の文字列に変換する機能を有し、前記音声認識で得られる 1 以上の認識候補を記憶しておく記憶手段と、前記指定文字列と対応する 1 以上の認識候補を前記記憶手段から読み出し、変換候補として表示する変換候補表示手段とを具えたことを特徴とする文字処理装置。

10

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、入力した文字を処理する文字処理装置および方法に関し、より詳しくは、音声認識により得られる文字列を処理する文字処理装置および方法に関する。

【0002】

【従来の技術】

従来この種の代表的な文字処理装置の処理例を図 1 ~ 図 3 を参照して説明する。図 1 ~ 図 3 は音声入力および入力音声の認識結果を文字処理するための表示内容を示す。図 1 (A) において、文書処理画面 1 においてユーザはアイコン 1 A をマウスにより指定して文字処理装置を起動する。文字入力装置はパーソナルコンピュータが文字処理プログラムを実行することにより実現される。

20

【0003】

文字処理装置の起動により表示画面には図 1 (B) のように文字入力用画面 2 が現れる。ユーザがメニュー項目 2 A を操作して音声入力モードに切り替えると、第 1 図 (C) に示すように音声入力用のウィンドウ (アプレット) 3 が表示される。ここで、ボタン 3 A をユーザがマウスにより操作するとマイクロフォンからの音声入力が可能となる。たとえば、ユーザがマイクロホンから「これはマイクのテストです」と音声で入力すると、コンピュータ内に搭載された音声認識プログラムにより入力音声は音声認識される。最近では連続の音声を認識することが可能となってきている。

30

【0004】

入力音声信号は音素と呼ばれる音韻よりも短い非常に短い時間長さに区切られ、信号の特徴成分が抽出される。この特徴成分と、あらかじめコンピュータに用意してある特徴成分 (音素の内容が判明している) とを比較し、特徴成分が類似するものの音素内容を検出することで、音素の音声認識を行う。認識の結果として、得られた音素を時系列的に結合すると、音韻 (母音、子音) の文字が得られる。この文字を結合し、文法的に区切られた文字列は音声認識用の辞書、すなわち、音韻 (読みを表す文字列) と、その読みに対応するかな漢字混じり文字列 (漢字のみを含む) を記載した辞書に基づき、かな漢字混じり単語に変換され、図 2 (D) の表示画面 1 内に表示される。

40

【0005】

なお、音声認識用の辞書に上記音声信号の特徴成分を単語単位で含め、単語単位で音声認識するような音声認識方法も提案されている。

【0006】

音声認識では誤認識が生じることがあるので、通常、文字処理装置は、音声認識結果 4 内の文字を削除したり、新しい文字をキーボード入力したり、他の同音異義語を選択するなどの文字処理を行うことができる。

【0007】

図 2 (D) の表示画面 1 中で、ユーザが文字カーソルを文字「を」に合わせ、削除 (Delete) ボタンを操作すると、文字カーソルで指示した文字 (「を」) を表示画面 1 上から削除

50

することができる（図2の（F）参照）。

【0008】

さらに認識結果4の中の文字列「マイク」を文字カーソルにより指定し、変換ボタンを操作すると、図2（E）に示すように、指定された文字列と発音が類似する単語、すなわち、音声認識で、類似度が高かった1以上の音声認識候補がプルダウンメニュー5の形態で表示される。ユーザはプルダウンメニュー5の中の所望の単語を選択すると、これまでに表示されていた単語が選択された単語と置換される。

【0009】

このようにして、不要な文字を削除したり（文字の挿入もある）、単語の置換を行った上で、図3（G）のEnterボタン8をユーザが操作すると、音声入力された文字列が確定し、ワードプロソフト等入力された文字列を使用して情報処理を行うアプリケーションに確定した文字列が引き渡される。

10

【0010】

一般的に文字処理される文字列にはアンダーライン等が付されて表示され、確定後の文字列はアンダーラインが取り除かれる（図3（H）参照）。

【0011】

【発明が解決しようとする課題】

音声入力による文字入力方法は、ユーザがキーボードで文字入力を行わなくてもよいという利点ある反面、誤認識結果を修正するためにはキーボード、あるいはマウスによるユーザの操作が不可欠である。このため、キーボードあるいはマウスの操作をできるだけ、

20

【0012】

また、音声入力した文字列を音声合成し、出力することで、合成音声と表示の文字列とを比較することで、音声認識結果の誤りを検出する装置が提案されているが、この装置は音だけでは、誤認識部分を検出することができないという解決すべき課題もあった。

【0013】

さらに本願出願は本願に先立って、キーボードから入力した文字列と音声で入力した文字列とを一括して文字処理できるようにした文字処理装置を提案しているが、キーボード入力と、音声入力を併用する場合には、打鍵の音がマイクロホンから混入し、音声の入力と誤認識されてしまうという改善すべき課題があった。

30

【0014】

そこで、本発明の第1の目的は、従来よりも操作性を改善した文字処理装置および方法を提供することにある。

【0015】

本発明の第2の目的は、音だけで音声認識結果の誤り部分を検出することができる文字処理装置を提供することにある。

【0016】

本発明の第3の目的は、キーの打鍵による音声の誤認識を阻止することが可能な文字処理装置を提供することにある。

【0017】

40

【課題を解決するための手段】

このような目的を達成するために、請求項1の発明は、入力の音声を音声認識し、音声認識結果を未確定のかな混じり漢字文字列として表示し、当該表示された文字列を文字処理した後、確定文字列を出力する文字処理装置において、変換対象の未確定のかな混じり漢字文字列を指定する変換対象指定手段と、当該指定された未確定のかな混じり漢字文字列に対して変換可能な1以上の文字列候補を漢字変換辞書から読み出して表示する表示手段と、当該表示された文字列候補の1つを選択して読みの表示を指示する読み表示指示手段と、該読み表示指示手段の読みの表示の指示に応じて、前記読み表示指示手段により選択された文字列についての1以上の読みの候補を漢字変換辞書から読み出して表示する読み表示手段と、前記読み表示手段により表示された1以上の読みの候補の1つを選択する

50

第1選択手段と、当該選択された読みの候補に対応する1以上のかな混じり漢字文字列を変換候補として漢字変換辞書から読み出して表示する変換候補表示手段と、当該表示された1以上のかな混じり漢字文字列の1つを選択する第2選択手段と、当該選択されたかな混じり漢字文字列で、前記変換対象指定手段により指定された未確定のかな混じり漢字文字列を置換する置換手段とをさらに具えたことを特徴とする。

【0018】

請求項2の発明は、請求項1に記載の文字処理装置において、前記読み表示手段は1以上の読みに加えて、前記変換対象指定手段により指定された未確定のかな混じり漢字文字列をも表示することを特徴とする。

【0019】

請求項3の発明は、請求項1に記載の文字処理装置において、該文字処理装置は、再変換後の指定文字列について他の文字列に変換する機能を有し、前記音声認識で得られる1以上の認識候補を記憶しておく記憶手段と、前記指定文字列と対応する1以上の認識候補を前記記憶手段から読み出し、変換候補として表示する変換候補表示手段とを具えたことを特徴とする。

【0037】

【発明の実施の形態】

以下、図面を参照して本発明の実施形態を詳細に説明する。

【0038】

最初に本発明に係わる文字処理方法を説明する。本実施形態の文字処理方法には以下の新規機能を有するので、各機能について順に説明していく。

【0039】

(1) 必要に応じて文字処理および確定処理を行えばよい音声連続入力機能
 (2) 単語指定するだけで、置換候補を自動表示する機能
 (3) 置換候補に対する読みの表示および読みからかな漢字への逆変換機能
 (4) 文字列を確定した後、再変換のために読みの類似候補(音声認識候補)を表示する機能

(5) 確定された文章を合成音声で再生出力(読み上げと称する)するとともに、合成音声と、入力音声とを一緒に出力する読み上げ・プレイバック同時機能

(6) 音声入力についての許可/禁止処理

(A) 音声連続入力機能

従来では、音声認識結果として得られるかな漢字混じり文については、その音声認識結果が正しいものであってもユーザは確定の操作を行わなければならない。この操作を改善したものが音声連続入力機能である。

【0040】

ユーザが連続的に音声入力を行っていくと、表示画面には、音声認識結果が図4(A)のように表示されていく。本実施形態の特徴は、認識結果が所定長さを超えようとする、先に入力された音声の認識結果、図4の例では、「日本××新聞社は臨時国会閉幕に」が自動的に確定される。残りの文字列は確定前の文字処理可能な状態のままとなる(図4(B)参照)。したがって、音声認識が正しい限り、ユーザは連続的に音声入力を続ければよく、最後に、1回だけ、確定操作を行えばよい。従来では、連続的に音声入力を行うと、その長さ(たとえば、3文節)に制限があるので、たとえば、3文節ごとに確定操作を行わなければならないの対し、本実施形態では、最後の1回だけとなる。

【0041】

また、入力音声をすべて自動確定せず、後の音声認識結果は未確定としているので、誤認識結果を修正することも可能となる。

【0042】

(B) 置換候補自動表示機能

従来では、確定前に音声認識結果の文字列を部分的に置換する場合には、置換すべき文字列を指定し、置換候補(いわゆる変換候補)の表示を特定のボタンで指示しなければなら

10

20

30

40

50

なかったのに対し、本実施形態では、置換すべき文字列を図5(C)のように文字カーソル(図中の|記号)で指定するだけで、自動的に置換候補が表示される。ユーザが図5(D)のように置換候補を選択し、置換処理を行うことができる。また、ユーザが、この後、文字カーソルを図6(E)に示すように認識結果の最後尾に合わせると、文字処理可能状態から音声入力可能状態となる。なお、表示画面に表示される文字列は音声認識において、類似度が高い1以上の音声認識候補の中のもっとも類似する文字列が使用されるが、類似度が高い1以上の音声認識候補は保存され、この置換候補表示機能における表示に使用される

(C)読み候補の表示および逆読み機能(上述の(3)の機能)

本実施形態では、音声認識結果を未確定状態としたまま、キーボードから入力した読みをも上記音声認識結果とともに確定対象の文字列として扱うことができる。このために、本実施形態では、音声入力され、表示されていて未確定の状態の文字列と、キーボードから入力した読みの文字列とを一括して文字処理することができる。このように2種の文字列が混在する状態で、かつ、未確定の状態での任意の文字列を範囲指定し、変換の指示をユーザが与えると、範囲指定された文字列について、漢字変換用辞書が参照され、範囲指定された文字列と対応する1以上の文字列が漢字変換用辞書から抽出され、漢字変換候補として図7の(A-1)の符号5のように表示される。

【0043】

なお、ここで、表示される変換候補は上記(B)の機能で表示される候補(音声認識の類似度の高い音声認識候補)とは異なる点に注意されたい。上記漢字変換用辞書は、よみ、かな漢字混じり文字列についての変換前の文字列と、この文字列に対応する変換後のかな漢字文字列を複数組記載した辞書である。変換前の文字列としては上述のよみを表す文字列、文法的、意味的にまちがった誤記の文字列が含まれる。また、変換後の文字列にはその1以上の読みが付加されている。

【0044】

文字カーソル指定により表示された図7の(A-1)のプルダウンメニューの中から図7の音声認識候補の変換候補「失わ」をユーザが選択すると、表示画面中の指定文字列がこの文字列と置換されることは述べたが、本実施形態では確定操作を行わない限り、漢字変換辞書から得られる他の文字列候補に置換できる。

【0045】

確定のためにユーザが未確定の文字列を従来と同様、文節単位で区切る。区切られた文字列の任意の文字列を指定して変換キーを操作すると、図7(B)のようにプルダウンメニュー5-1の形態で、漢字変換用辞書から得られる置換候補が(この場合には1つ)表示される。

【0046】

ユーザが置換候補を選択した上で、読みの表示を指示するとこの文字列のすべての文字列と読みとが図7のプルダウンメニュー5-2に表示される。このプルダウンメニュー5-2は、ユーザにより選択された変換候補「失わせて」に対応し、漢字変換用辞書に記載されている1以上の読みの候補を選択目的で表示する。なお、このプルダウンメニュー5-2の中に、選択された置換候補「失わせて」も含まれていることに注意されたい。

【0047】

プルダウンメニュー5-2は読み(あるいは置換候補)から、かな混じり漢字文字列を表示させることができる。ユーザがプルダウンメニューの中の所望の文字列、たとえば、「ひろうせて」を選択すると図7の(A-2)のようにこの読みを持つかな混じり漢字文字列が漢字変換用辞書から抽出され、その置換候補も表示される(図7(A-2)参照、また、プルダウンメニュー5-2の候補「いやおうせて」をユーザが選択するとその文字列に対応した、読み候補を表示するためのプルダウンメニューが図8の(C)のように表示される。

【0048】

以上、説明したように、(C)の機能では、表示画面で範囲指定した文字列に対応し、か

10

20

30

40

50

な漢字変換用辞書に記載されている文字列をプルダウンメニュー 5 に表示する。これにより、キーボードから入力された文字列を漢字等に変換して、文字列を確定することができる。また、プルダウンメニュー 5 に所望の文字列がない場合は読みについてのプルダウンメニュー 5 - 2 を介して、確定すべき文字列を検索することができる。

【 0 0 4 9 】**(D) 確定後の読みの類似候補の表示機能**

確定処理を行うと、確定された文字列は、通常、アプリケーションに引き渡される。なお、音声認識機能を有しない従来例においては、確定後の文字列を選択するとかな漢字変換時に表示された変換候補を再表示して、他の変換候補と置換する機能を有する文字処理装置が提案されている。

10

【 0 0 5 0 】

しかしながら、音声認識機能を有する従来例では、文字列確定後に変換候補を再表示する機能は有していない。そこで、本実施形態では、文字変換（置換）処理に使用された類似単語候補、すなわち、図 7 の（ A - 2 ）で表示される読みの類似文字列候補で、音声認識時に音声認識候補として使用された文字列を確定後の再変換処理においても表示する。図 9 の表示がこの処理内容を示す。ユーザにより確定された文字列「失わせて」を選択し、再変換を不図示のメニューで指示すると、文字カーソルの位置指定において表示されたものと同じプルダウンメニュー 6 が表示される。

【 0 0 5 1 】

なお、図 9 の例では、選択候補は 1 つであるが、複数ある場合には、その複数の文字列が選択のために表示される。このメニュー 6 においても読みが表示され、読みの項目を選択すると読みを記載したプルダウンメニューが表示される。

20

【 0 0 5 2 】**(E) プレイバック・読み上げ機能**

音声認識された文字列を音声合成し、合成音声を出力する読み上げ機能を有する装置や、入力音声をそのまま再生（プレイバック）するプレイバック機能を有する装置は従来から知られているが、上記双方の機能を有する装置はいまだない。本実施形態では音声認識された未確定文字列について読み上げと、プレイバック機能を同時に起動することができる。この 2 つの機能により、ユーザは目視に頼らず音だけで、音声認識結果の誤り部分を検出することができる。

30

【 0 0 5 3 】**(F) 音声入力についての許可 / 禁止処理**

本実施形態では、キーボード 1 2 から入力が行われている場合には、マイクロホン 1 0 から音声の入力があっても、その入力を文字認識プログラムでは受け付けない。これによりキーボード 1 2 の打鍵の音をマイクロホン 1 0 が拾って、音声認識において誤認識を生じるトラブルを回避する。

【 0 0 5 4 】

以上の機能を実現するためのシステム構成を図 1 0 および図 1 1 を参照して説明する。図 1 0 は文字処理装置を搭載した情報処理装置、たとえば、パーソナルコンピュータの外観を示す。

40

【 0 0 5 5 】

図 1 0 において、1 0 は音声を入力するためのマイクロホン、1 1 は音声入力され、音声認識された文字列、文字処理された後の確定文字列を表示するためのディスプレイである。1 2 は文字処理対象の文字列を入力するとともに、ディスプレイ 1 1 に表示された文字カーソルを移動し、文字位置を指定するためのカーソルキー（矢印キーとも呼ばれる）を有するキーボードである。

【 0 0 5 6 】

1 3 は表示画面の位置指定を行うためのポインティングデバイスであり、本実施形態ではマウスを使用する。

【 0 0 5 7 】

50

図 1 1 は情報処理装置の代表的な内部構成を示す。図 1 1 において、CPU 2 0 1 は HDD (ハードディスク) 2 0 3 に保存されている文字処理プログラムを実行することにより文字処理装置として機能する。なお、CPU 2 0 1 は装置全体をもオペレーティングシステムにしたがって制御する。

【 0 0 5 8 】

システムメモリ 2 0 2 は CPU 2 0 1 の演算処理で使用するデータ、ディスプレイ 1 1 に表示するデータ、文字処理に関わるデータの記憶に使用される。さらに、システム制御に必要なデータをもシステムメモリ 2 0 2 に記憶される。

【 0 0 5 9 】

HDD 2 0 3 には上述のオペレーティングシステム、文字処理プログラムの他、音声認識に使用する音声認識用辞書、音声認識プログラム、漢字変換用辞書を保存する。

10

【 0 0 6 0 】

なお、HDD 2 0 3 に保存される上述のデータは不図示の CDROM ドライブを介して CDROM から実装することができる。また、フロッピーディスクから上記データを実装する場合にはフロッピーディスクドライブが使用される。

【 0 0 6 1 】

2 0 4、2 0 6、2 0 8 ~ 2 1 0 は後述の回路とバス 2 1 1 を接続する入出力インターフェース (I/O) である。I/O 2 0 4 は CPU 2 0 1 から出力される合成音声信号および音声信号をデジタル/アナログ変換器 2 0 5 に転送する。デジタル信号からアナログ信号に変換された合成音声信号および音声信号が左および右のスピーカから再生される。

20

【 0 0 6 2 】

マイクロホン 1 0 から入力された音声はアナログ音声信号に変換され、アナログデジタル変換器 2 0 7 によりデジタル形態の音声信号に変換され、I/O 2 0 6 を介して CPU 2 0 1 に引き渡される。

【 0 0 6 3 】

I/O 2 0 8 はキーボード 1 2 からの入力情報を CPU 2 0 1 に転送する。I/O 2 0 9 はディスプレイ 1 1 で表示するイメージデータを転送する。

【 0 0 6 4 】

以上の回路を有する文字処理システムの動作説明に先立って、本実施形態の処理データの流れを図 1 2 を参照して説明しておく。

30

【 0 0 6 5 】

マイクロホン 1 0 から入力された音声は音声認識プログラム 1 0 0 により音声認識される。また、入力音声は記憶装置 (システムメモリ 2 0 2) 1 0 5 に記憶される。本実施形態の音声認識プログラムは隠れマルコフモデルとよばれる音声の特徴を使用して音声認識を行う。音声認識の結果としては入力音声 (たとえば、文節) に対して類似度が高い 1 以上のかな漢字混じり文字列が得られる。この中のもっとも類似度が高い文字列が文字処理プログラム 1 0 2 を介して表示される。類似度の高い文字列は音声認識用辞書から従来と同様の音声認識処理において抽出された後、文字処理プログラム 1 0 2 に引き渡される。

【 0 0 6 6 】

キーボード 1 2 から入力された読みの文字列は文字処理プログラム 1 0 2 に引き渡され、かな混じり漢字 (漢字のみの文字列、カタカナの文字列をも含む) に変換される。なお、音声で入力された文字列と、キー入力された文字列とを結合した文字列をも 1 つの文字列としてかな混じり漢字の変換の対象とすることができる。

40

【 0 0 6 7 】

文字処理プログラムはユーザの文字カーソル操作、マウス操作、その他、挿入、削除等のキー操作を受け付け、操作に応じて上述した文字処理方法にしたがって、文字処理を行う。読み上げ・プレイバック機能が指示された場合には、記憶装置 1 0 5 に記憶された原音声データ (入力の音声信号) と文字処理プログラムにより音声合成された合成音声データは左右のスピーカ 1 0 6 から別個に同時に再生される。

【 0 0 6 8 】

50

文字処理プログラム102により文字処理（漢字変換、挿入、削除）されて確定された文字列は文書処理プログラム（ワープロソフト）104等文字列を入力とするアプリケーションに引き渡される。なお、文書処理プログラムによる処理中で範囲指定された確定文字列について再変換の指示がユーザにより行われた場合には、文字処理プログラム102に再変換の対象の文字列が引き渡され、変換候補が表示画面に表示される。ユーザの候補選択の後、確定された文字列は文書処理プログラム104に引き渡されて、表示画面に表示される。

【0069】

図12の回路の動作、主にCPU201の処理内容を図13～図18のフローチャートを参照して説明する。これらのフローチャートに示す処理手順は文字処理プログラムとして、CPU201の実行可能なプログラム言語の形態でHDD203にあらかじめ保存されている。CPU201の実行時にはHDD203からシステムメモリ202にロードされる。

10

【0070】

図13は文字処理プログラムのメイン処理手順を示す。図13の処理手順は入力イベント、すなわち、マイクロホン10からの音声入力、キーボード12からのキー入力、あるいはマウス13のいずれからか情報入力があった場合に実行される。

【0071】

以下に述べる音声許可フラグは音声の入力が許可されているときにオン、音声の入力が許可されていない（禁止）場合にオフとなる。

20

【0072】

マイクロホン10から音声入力があると、CPU201は音声許可フラグがオンであることを確認した上で、音声認識プログラムを起動し、入力音声データを引き渡す。これにより、従来と同様の音声認識用辞書を使用した音声認識処理が行われる（ステップS11）。音声の認識結果、この場合、文節単位での発声、すなわち、読みの類似度の高い1以上のかな混じり漢字文字列が後述の選択候補の表示のためにシステムメモリ202に記憶され、もっとも類似度の高いかな混じり漢字文字列がディスプレイ11に表示される（ステップS12）。なお、このとき、従来と同様、未確定を表す点線が表示の文字列に付される。

【0073】

一方、キーボード12からの入力があることがステップS20で検出されると、システムメモリ202内の文字処理用の専用領域に入力情報が記憶されるとともに、ディスプレイ11にも表示される（ステップS21）。マウス13の操作により、文字処理関連の指示が行われた場合には、その指示に応じた文字処理が行われる（ステップS31）。種類ごとの文字処理については後述する。

30

【0074】

図14は音声連続機能を実現するための処理手順を示す。音声認識結果をシステムメモリ202に記憶し、ディスプレイ11に表示する際に（図13のステップS11）、図14の処理手順が実行される。CPU201はシステムメモリ202内の文字処理用専用領域（以下、ワークエリアと略称する）に記憶されており、未確定の文字数を計数する。

40

【0075】

計数した文字数（あるいは文節数）、（ただし、音声認識されたばかりの文字列でワークエリアに記憶されていない文字列を含む）があらかじめ定めた文字数（あるいは文節数）を超えている場合には、時間的に早く記憶されている1文節の文字列（あるいは所定文字数の文字列）、すなわち、先頭部の文字列を未確定から確定に変更する（ステップS100～S101）。本実施例では上記ワークエリアから表示用の文字記憶領域に自動確定の文字列を削除する。次にワークエリア内の未確定の文字数と文字認識された文字列の文字数の合計を計算し（ステップS102）、再び、あらかじめ定めた文字数との比較を行う（ステップS100）。以下、ステップS100～S102のループ処理を繰り返す。これにより、今、音声認識された文字列をワークエリアに記憶しても、オーバーフロー（記

50

憶容量を越えること)を避けることができる。

【0076】

ステップS100の判定がNOとなったときに、CPU201は音声認識された文字列をワークエリアに記憶するとともに、新たに記憶された文字列をこれまでに記憶されている文字列(未確定文字列と確定文字列)とともに表示する(ステップS110 S120)。

【0077】

以上の処理をCPU201が実行することにより、ワークエリア内に記憶される未確定文字数は一定の値を超えることはなく、また、図4に示すように先に音声で入力された文字列は文節単位で確定される。以上が音声入力の場合の処理であるが、キーボード12から連続的に文字列が入力された場合にも同様の処理を行うことができる。

10

【0078】

なお、音声認識結果として得られた1以上の類似文字列のうち、もっとも類似する文字列がワークエリアに記憶され、1以上の類似文字列全体は、ワークエリアに記憶されたもっとも類似する文字列と関連付けられてシステムメモリ202の別領域に記憶される。後述するが、別領域に記憶された類似文字列は文字カーソルの位置指定に応じた置換候補の表示に使用される。

【0079】

図15は文字カーソルにより未確定文字列の中の特定文字が指定された場合の処理内容を示す。この処理手順は、文字カーソルがキーボード12のカーソルキーもしくはマウス13により移動されたことをCPU201が検出したときに、CPU201は文字処理関連の指示ありと判断し、図13の処理手順をステップS10~S30 S31へと進め、図15の処理手順を実行する。図15において、CPU201はカーソルキーの位置が未確定文字列中に位置したことを検出すると、指定された文字位置を含む文字列と対応する置換候補をシステムメモリ202から読み出して、図5の(C)のようにプルダウンメニューの形態で表示する(ステップS200 S210 S220 S230)。ユーザが置換候補を選択すると、CPU201は選択された候補と、現在、未確定文字列として表示されている文字列とを置換する(ステップS250)。

20

【0080】

本実施形態では、文字カーソルキーの位置指定による置換候補の自動表示と関連させて、次のような機能を用意している。ユーザが、文字カーソルにより未確定文字列を指定した状態で、キーボード12の特定のキー(たとえば、ENTERキー)を操作すると、その操作が、ステップS220で検出され、文字カーソルは未確定文字列の最後尾に移動し、以後、音声を使用した文字入力が許可された状態(音声許可フラグオン)となる。

30

【0081】

次に、置換候補の読みの表示についての処理を図16を参照して説明する。

【0082】

図16の処理手順は未確定の文字列を確定するために、未確定の文字列を文節単位に区切り、区切られた文字列の中の特定の文節をユーザが指定して変換の指示を行ったときに実行される。図16において、CPU201はユーザが変換対象の区切り文字列の指定を受け付け、キーボード12の変換キーを操作すると、CPU201は指定された文字列について関連する1以上の文字列、すなわち、同音異義語や、カタカナだけの文字列、文法的に正しい文字列を漢字変換用辞書から取得する。このとき、1以上の読みも共に取得される(ステップS300 S310 S320)。

40

【0083】

取得された文字列(置換候補)は図7の(B)のように表示される(ステップS330)。ユーザが候補選択を行った上で読みの表示の指示をマウス13により行くと、CPU201は選択された文字列についての読み候補(漢字変換用辞書から取得済み)を図7(B)の符号5-2のように表示する(ステップS340 S350 S370、なお、選択されたかな混じり漢字を含む)。

50

【 0 0 8 4 】

この読み候補をユーザが選択すると、CPU 201は同じ読みを持つ文字列（読みから逆引きの文字列）を漢字変換用辞書から抽出し、表示する（ステップS 374、図7（A-2）参照）。ユーザが表示された文字列候補の中から所望の文字列を選択すると選択された文字列による置換が行われる（ステップS 376 S 78）。以上説明したようにこの処理手順では、読みからかな漢字混じり文字列を選択することも可能となる。

【 0 0 8 5 】

確定後の再変換処理において、漢字変換用辞書のかわりに、システムメモリ202に記憶しておいた類似文字列候補（音声認識候補）を使用する点は表示する候補が異なるだけであり、詳細な処理手順の説明を要しないであろう。

10

【 0 0 8 6 】

次に、プレイバック・読み上げ処理を図17を参照して説明する。図17の処理手順は、図1（C）の音声アプレット3中に新たに設けられたプレイバック・読み上げ指示ボタンあるいは従来からあるプレイバックまたは読み上げボタンが操作されたときに実行される。図17において、CPU 201は操作ボタンの種類をステップS 400～S 420で判別する。プレイバックボタンが操作された場合には手順はステップS 400 S 401へと進み、従来と同様のプレイバック処理（システムメモリ202に保存されている入力音声再生する処理）。

【 0 0 8 7 】

また、読み上げボタンが操作された場合には、手順はステップS 410 S 411 S 412と進み、従来と同様、現在、表示されている未確定文字列について、CPU 201は音声合成を行って、合成音声をスピーカ106から再生出力する。

20

【 0 0 8 8 】

一方、プレイバック・再生ボタンが操作された場合には、読み上げのための音声合成データをたとえば、左のスピーカ、入力の原音声を右のスピーカからというように再生出力のスピーカを分けて原音声と、合成音声を出力する（ステップS 420 S 430 S 44）。

【 0 0 8 9 】

次に音声入力とキーボード入力の制御について図18を参照して説明する。

【 0 0 9 0 】

図18の処理手順は一定周期割り込みで実行される。CPU 201はキーボード12用のI/O 208に保持されているキーコード信号を読み出し、キー入力の有無を判定する。キー入力がある場合には、キーコード信号が保持されていないので、これによりステップS 500の判定結果はYESとなり、手順はステップS 501へと進む。

30

【 0 0 9 1 】

このステップで、CPU 501はタイマーの計数値をゼロにリセットして、音声許可フラグをオフし、音声の入力を禁止する。タイマーは1回のキーの押下時間よりやや大きい非常に時間でカウントアップするものを使用する。連続的にキー入力がある場合には、ステップS 500 S 501 S 502の手順が繰り返し実行され、タイマーはカウントアップする前に再起動される。

40

【 0 0 9 2 】

一方、キー入力がなくなると、ステップS 100の判定結果がYES判定のとなる。したがって、最後のキー入力で計時を開始したタイマーがカウントアップするまでは、音声許可フラグはオフ状態を続け、最後のキー入力処理が終了するまでは音声入力処理が禁止される。キー入力がなく、かつ、タイマーがカウントアップすると、手順は、ステップS 500 S 510 S 511へと進み、このステップで、音声許可フラグがオンとなり音声の入力が許可される。この音声許可フラグのオンオフにしたがって。CPU 201は図13の音声入力、認識処理（ステップS 11）を行う。

【 0 0 9 3 】

上述の実施形態の他に次の形態を実施できる。

50

【 0 0 9 4 】

1) 上述の実施形態では文字カーソルの文字列指定では、音声認識で得られた1以上の認識候補を置換候補として表示し、確定のための変換の指示があったときには漢字変換用辞書に記載された変換候補を表示している。しかしながら、文字カーソルの指定や変換の指示で表示する候補については、音声認識で得られた認識候補と、漢字変換用辞書から得られる変換候補を混在されて表示し、ユーザが選択することも可能である。

【 0 0 9 5 】

2) 上述の実施形態では図9に示すように読みを表示しているが、この読みは漢字変換用変換辞書から取得してもよいし、音声認識用辞書から取得するようにしてもよい。さらには双方の辞書から取得するようにしてもよいこと勿論である。

10

【 0 0 9 6 】

3) 本実施形態の漢字変換用辞書は、ワープロなどで使用される判事変換用辞書と異なり、読みに加えて誤記のかな漢字文字列をも変換前の文字列としている。このため、本実施形態の漢字変換用辞書は置換用辞書と考えることができる。置換の対象の文字列としては、上述の読み、文法的に間違っただけの文字列(誤記、仮名遣いの異なる文字列、スペルの誤りのある英文字列をいわゆる置換辞書に記載しておくことができる。

【 0 0 9 7 】

4) 上述の実施形態では、パーソナルコンピュータに文字処理装置を搭載する例を説明したが、音声で文字を入力する各種の装置に対して本発明を適用することができる。

【 0 0 9 8 】

20

【 発明の効果 】

以上、説明したように、請求項1の発明では、未確定のかな混じり漢字文字列を変換する場合に、変換候補の読みを知ることができるので、操作者が選択しようとしている文字列が、操作者自身が望む文字列であるか否かを確認することができる。

さらに読みからかな混じり漢字文字列を逆引きすることができるので、操作者が選択することができる文字列の範囲を多種、多岐とすることができる。

【 0 0 9 9 】

請求項2の発明では読みに加えて、変換候補の文字列も表示されるので、読みとかな漢字文字列の対比が可能となる。

【 0 1 0 0 】

30

従来の再変換機能では、漢字変換辞書の変換候補が表示されていたのに対し請求項3の発明では、音声認識候補が表示されるので、入力音声に類似した文字列を再変換することができる。

【 図面の簡単な説明 】

【 図 1 】 (A) ~ (C) は音声入力のための操作の推移を示す説明図である。

【 図 2 】 (D) ~ (F) は音声入力のための操作の推移を示す説明図である。

【 図 3 】 (G) および (H) は音声入力のための操作の推移を示す説明図である。

【 図 4 】 (A) および (B) は連続音声についての音声認識結果の処理の変化を示す説明図である。

【 図 5 】 (C) および (D) は置換候補の表示の一例を示す説明図である。

40

【 図 6 】 (E) は文字カーソルの移動結果を示す説明図である。

【 図 7 】 (A - 1)、(A - 2)、(B) は置換候補の表示例を示す説明図である。

【 図 8 】 (C) は読みから逆引きした置換候補の表示例を示す説明図である。

【 図 9 】 再変換時の置換候補の表示例を示す説明図である。

【 図 1 0 】 音声処理装置を有する情報処理装置の外観を示す構成図である。

【 図 1 1 】 図 1 0 の情報処理装置の内部構成を示すブロック図である。

【 図 1 2 】 本発明実施例のデータの流れを示すブロック図である。

【 図 1 3 】 本発明実施形態のメイン処理手順を示すフローチャートである。

【 図 1 4 】 入力文字の記憶、表示のための処理内容を示すフローチャートである。

【 図 1 5 】 文字カーソルの指定に伴う置換候補の表示処理を示すフローチャートである。

50

【図16】本発明実施形態の文字列置換、読み候補の表示、読みからの逆引きのための処理を示すフローチャートである。

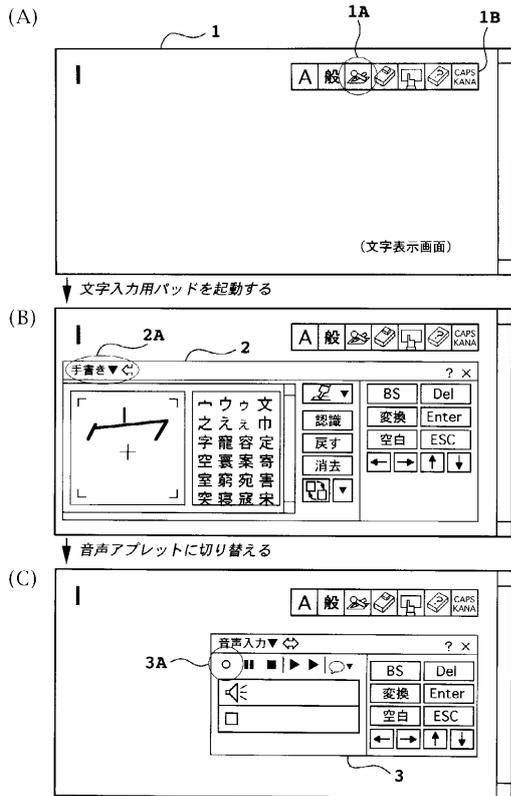
【図17】プレイバック・読み上げの処理内容を示すフローチャートである。

【図18】キー入力に関連した音声入力制御の内容を示すフローチャートである。

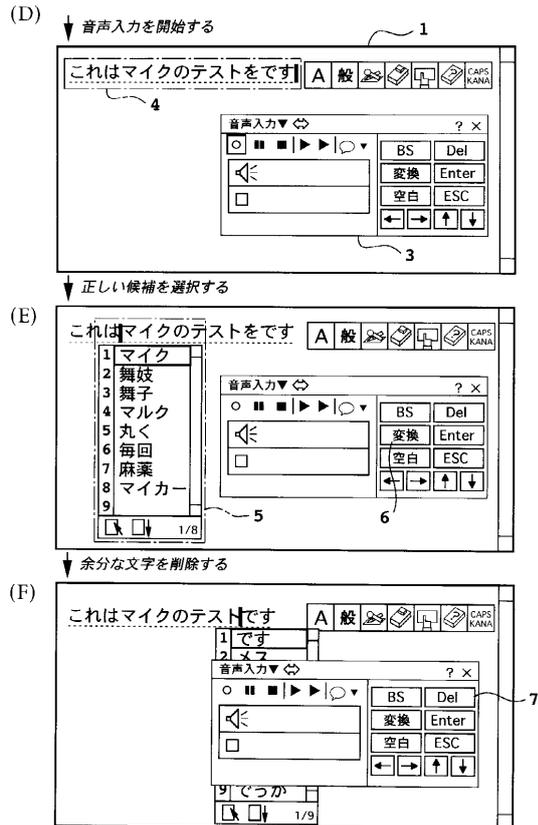
【符号の説明】

- 1 1 ディスプレイ
- 1 2 キーボード
- 1 3 マウス
- 2 0 1 CPU
- 2 0 2 システムメモリ
- 2 0 3 HDD

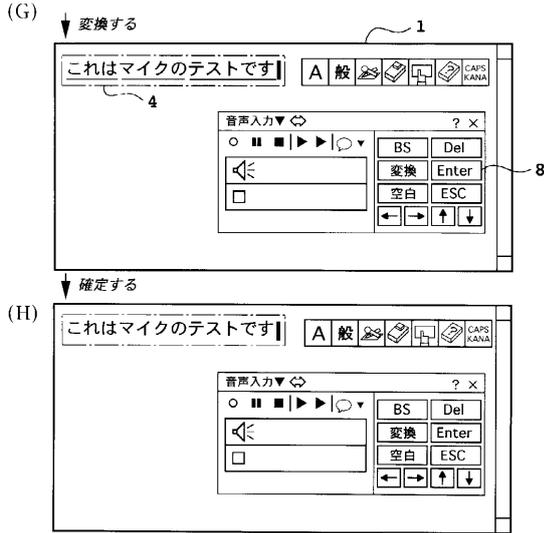
【図1】



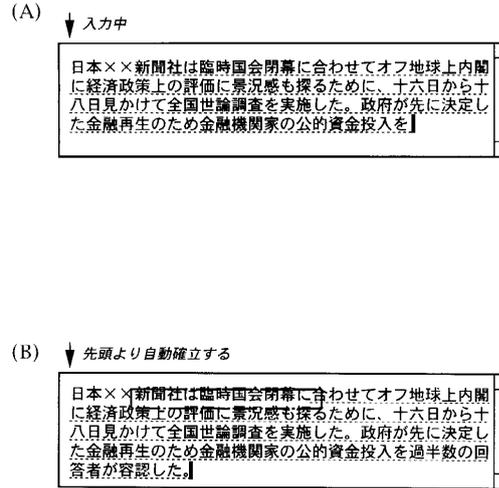
【図2】



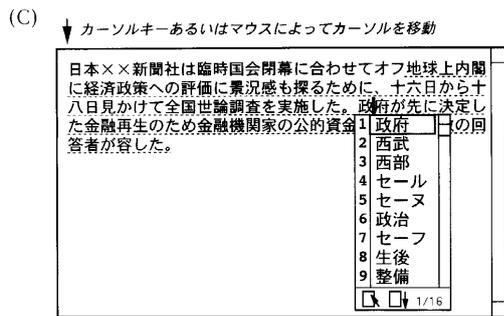
【図 3】



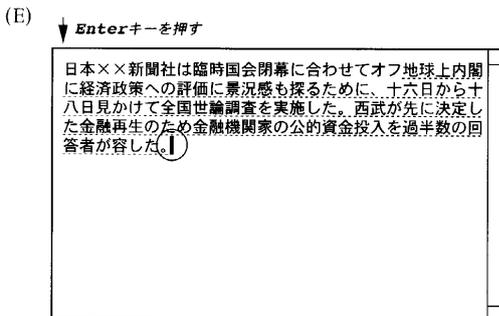
【図 4】



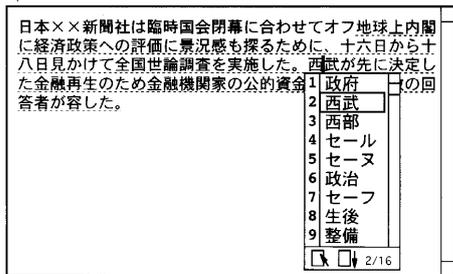
【図 5】



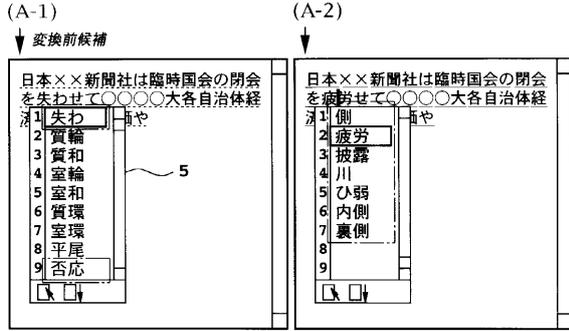
【図 6】



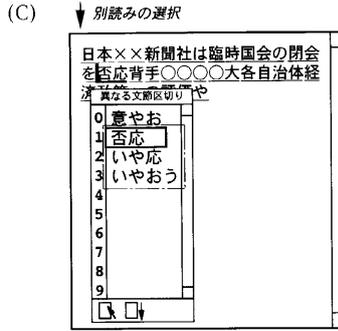
(D) ↓ 変換キーあるいはスペースキーを押す



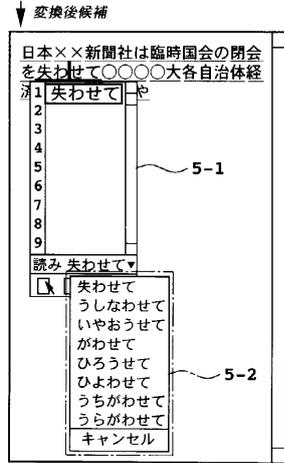
【 図 7 】



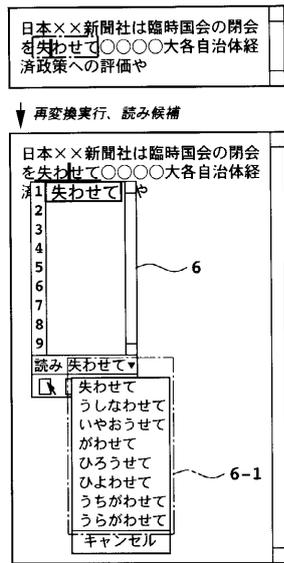
【 図 8 】



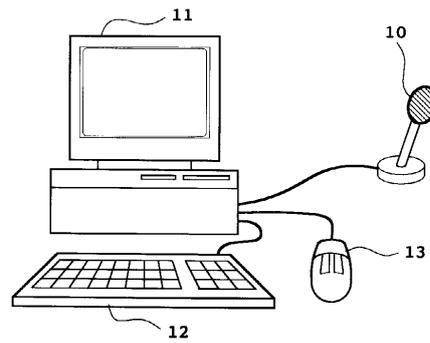
(B)



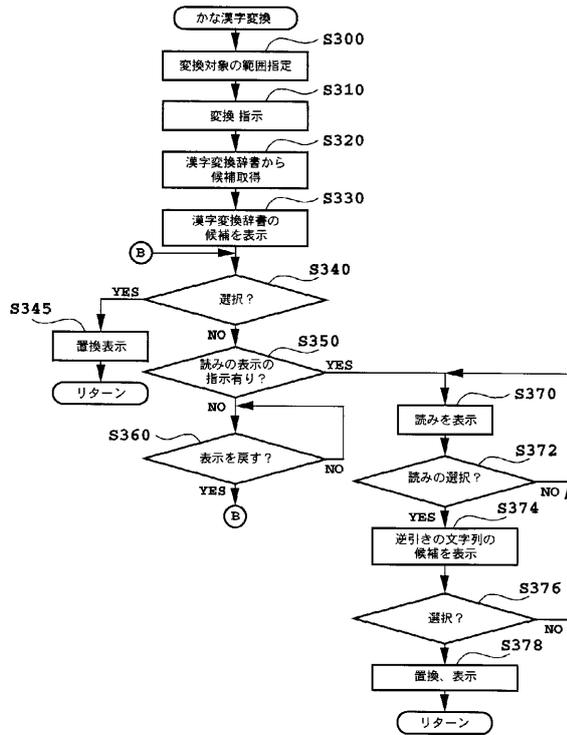
【 図 9 】



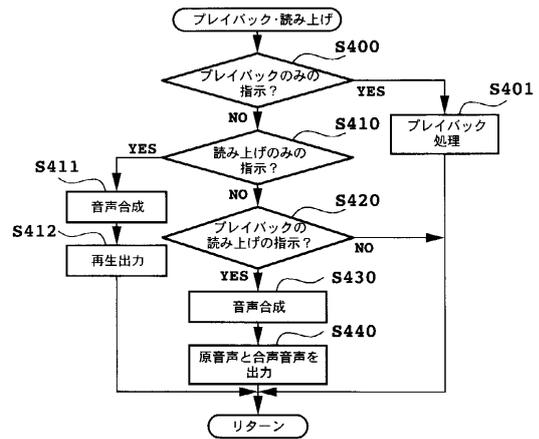
【 図 10 】



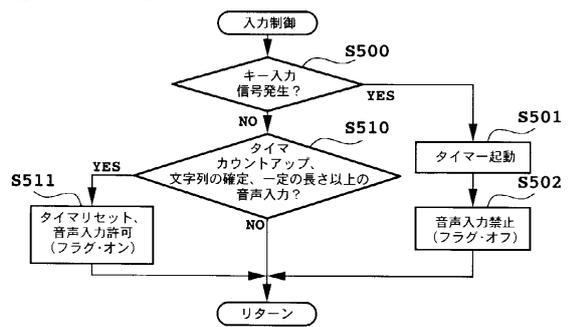
【図16】



【図17】



【図18】



フロントページの続き

(72)発明者 藤村 武志
東京都調布市調布ヶ丘1丁目18番1号 マイクロソフト株式会社 マイクロソフト調布技術センター内

(72)発明者 只埜 正義
東京都調布市調布ヶ丘1丁目18番1号 マイクロソフト株式会社 マイクロソフト調布技術センター内

合議体

審判長 藤内 光武

審判官 野崎 大進

審判官 右田 勝則

(56)参考文献 特開平10-143503(JP,A)
特開昭62-266649(JP,A)
特開昭61-125677(JP,A)
特開平8-202708(JP,A)
特開昭61-125676(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F17/22

G10L15/26