

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5372148号  
(P5372148)

(45) 発行日 平成25年12月18日(2013.12.18)

(24) 登録日 平成25年9月27日(2013.9.27)

(51) Int.Cl. F I  
**G06F 17/30 (2006.01)** G O 6 F 17/30 3 8 O E  
 G O 6 F 17/30 1 1 O G

請求項の数 15 (全 23 頁)

(21) 出願番号	特願2011-516899 (P2011-516899)	(73) 特許権者	309027126
(86) (22) 出願日	平成21年7月6日(2009.7.6)		ニュアンス コミュニケーションズ, イン
(65) 公表番号	特表2011-527058 (P2011-527058A)		コーポレイテッド
(43) 公表日	平成23年10月20日(2011.10.20)		アメリカ合衆国 マサチューセッツ州 O
(86) 国際出願番号	PCT/US2009/049730		1 8 0 3 パーリントン ワン・ウェイサ
(87) 国際公開番号	W02010/003155		イド・ロード (番地なし)
(87) 国際公開日	平成22年1月7日(2010.1.7)	(74) 代理人	100078282
審査請求日	平成23年2月25日(2011.2.25)		弁理士 山本 秀策
(31) 優先権主張番号	61/078, 299	(74) 代理人	100062409
(32) 優先日	平成20年7月3日(2008.7.3)		弁理士 安村 高明
(33) 優先権主張国	米国 (US)	(74) 代理人	100113413
(31) 優先権主張番号	61/078, 293		弁理士 森下 夏樹
(32) 優先日	平成20年7月3日(2008.7.3)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 モバイルデバイス上で日本語テキストを処理する方法およびシステム

(57) 【特許請求の範囲】

【請求項1】

モバイルデバイスを介して複数の項目を検索するためのコンピュータ実装方法であって、該複数の項目のうちの個別の項目は、日本語テキスト部分によって特徴付けられ、該コンピュータ実装方法は、

日本語テキストによって特徴付けられる項目を識別するためにモバイルデバイス上で検索クエリを受信することと、

該受信された検索クエリに基づいて一式の検索語を生成することと、

テキスト情報を格納する1つ以上のテキストフィールドを検索することにより、該検索語に一致する複数の一致項目を見出すことと、

該複数の一致項目の各々について、

各一致項目が配置されている該テキストフィールド内の1つ以上の自然な開始点を決定することと、

自然な開始点に対する該一致項目の距離を決定することと、

各一致項目と自然な開始点との間の距離に基づいて順序付けられる一致項目のリストを提供することと

を含む、コンピュータ実装方法。

【請求項2】

前記一式の一致項目は、メディアファイル、アドレス帳入力、文書ファイル、またはアプリケーションのうちの少なくとも1つを含む、請求項1に記載のコンピュータ実装方法

。

## 【請求項 3】

前記自然な開始点を決定することは、  
前記日本語テキスト部分内の句読点を識別することと、  
該識別された句読点に近接する該自然な開始点の位置を特定することと  
を含む、請求項 1 に記載のコンピュータ実装方法。

## 【請求項 4】

前記自然な開始点を決定することは、  
前記日本語テキスト部分内の文字体系の変化を識別することと、  
該識別された文字体系の変化における該自然な開始点の位置を特定することと  
を含む、請求項 1 に記載のコンピュータ実装方法。

10

## 【請求項 5】

前記自然な開始点を決定することは、  
前記日本語テキスト部分内の電話番号の形式を有する文字列を識別することと、  
該識別された文字列に近接する該自然な開始点の位置を特定することと  
を含む、請求項 1 に記載のコンピュータ実装方法。

## 【請求項 6】

モバイルデバイスから複数の項目を検索するためのシステムであって、該システムは、ユーザ検索クエリに基づいて一式の検索語を生成するように構成されている変換コンポーネントと、

20

テキスト情報を格納する 1 つ以上のテキストフィールドを検索することにより、該検索語に一致する複数の一致項目を見い出すように構成されている検索コンポーネントと、各一致項目が配置されている該テキストフィールド内の 1 つ以上の自然な開始点を決定するように構成されている開始点決定コンポーネントと、

自然な開始点に対する該一致項目の距離を計算するように構成されている距離計算コンポーネントと、

各一致項目と自然な開始点との間の距離に基づいて順序付けられる一致項目のリストを表示するように構成されているディスプレイと

を備える、システム。

## 【請求項 7】

前記複数の一致項目は、メディアファイル、アドレス帳入力、文書ファイル、画像ファイル、またはアプリケーションのうちの少なくとも 1 つを含む、請求項 6 に記載のシステム。

30

## 【請求項 8】

前記開始点決定コンポーネントは、  
前記テキストフィールド内で句読点を識別することと、  
前記識別された句読点に近接する前記自然な開始点の位置を特定することと  
によって、前記自然な開始点を決定するように構成されている、請求項 6 に記載のシステム。

## 【請求項 9】

前記開始点決定コンポーネントは、  
前記テキストフィールド内の文字体系の変化を識別することと、  
前記識別された文字体系の変化における前記自然な開始点の位置を特定することと  
によって、前記自然な開始点を決定するように構成されている、請求項 6 に記載のシステム。

40

## 【請求項 10】

前記開始点決定コンポーネントは、  
前記テキストフィールド内の電話番号を示す文字列を識別することと、  
前記識別された文字列に近接する前記自然な開始点の位置を特定することと  
によって、前記自然な開始点を決定するように構成されている、請求項 6 に記載のシス

50

テム。

【請求項 1 1】

モバイルデバイスを通じてアクセス可能な一式の項目の中で検索することを方法によって行うように、該モバイルデバイスのプロセッサを制御するための命令を含むコンピュータ読み取り可能な記憶媒体であって、該一式の項目のうちの個別の項目は、日本語テキストによって特徴付けられ、

該方法は、

日本語テキストによって特徴付けられる項目を識別するように該モバイルデバイス上で検索クエリを受信することと、

該受信された検索クエリに基づいて一式の検索語を生成することと、

テキスト情報を格納する1つ以上のテキストフィールドを検索することにより、該検索語に一致する複数の一致項目を見出すことと、

該複数の一致項目の各々について、

各一致項目が配置されている該テキストフィールド内の1つ以上の自然な開始点を決定することと、

自然な開始点に対する該一致項目の距離を決定することと、

各一致項目と自然な開始点との間の距離に基づいて順序付けられる一致項目のリストを提供することと

を含む、コンピュータ読み取り可能な記憶媒体。

【請求項 1 2】

前記一式の一致項目は、メディアファイル、アドレス帳入力、文書ファイル、またはアプリケーションのうちの少なくとも1つを含む、請求項 1 1 に記載のコンピュータ読み取り可能な記憶媒体。

【請求項 1 3】

前記自然な開始点を決定することは、

前記日本語テキスト部分内の句読点を識別することと、

該識別された句読点に近接する該自然な開始点の位置を特定することと

を含む、請求項 1 1 に記載のコンピュータ読み取り可能な記憶媒体。

【請求項 1 4】

前記自然な開始点を決定することは、

前記日本語テキスト部分内の文字体系の変化を識別することと、

該識別された文字体系の変化における前記自然な開始点の位置を特定することと

を含む、請求項 1 1 に記載のコンピュータ読み取り可能な記憶媒体。

【請求項 1 5】

前記自然な開始点を決定することは、

前記日本語テキスト部分内の電話番号の形式を有する文字列を識別することと、

該識別された文字列に近接する前記自然な開始点の位置を特定することと

を含む、請求項 1 1 に記載のコンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

(関連出願に対する相互参照)

本願は、共に2008年7月3日に出願された、米国仮出願第61/078,293号(名称「IMPROVED METHOD FOR SEARCHING JAPANESE TEXT USING A MOBILE DEVICE」)および米国仮出願第61/078,299号(名称「IMPROVED METHOD OF WORD SELECTION FOR JAPANESE TEXT ENTRY ON A MOBILE DEVICE」)の利益を主張する。

【背景技術】

## 【 0 0 0 2 】

(背景)

表記される日本語は、概して、いくつかの異なる文字セットからの文字の組み合わせである。特に、日本人は、表語文字表記体系、2つの個別の表音文字体系に加えて、ラテン文字、アラビア数字、および他の言語から取り入れられた他の記号を使用する。ひらがなおよびカタカナと呼ばれる2つの生来の文字体系は、文字(かなと呼ばれる)を使用して音節を表す。ひらがなおよびカタカナは、合計でおよそ90文字である。漢字と呼ばれる文字セットは、単語または単語の部分を表す、何千もの表語文字から成る。

## 【 0 0 0 3 】

図1は、日本語テキストを表すための3つの主要体系を示す。概して、日本語は、見出し語で表記され、その例は、図1の右側の列に示される。見出し語とは、漢字、かな、ラテン文字、アラビア数字、記号、および句読点を含む、上記の文字体系のうちのいずれかからの文字を有するテキストを言う。日本語テキストは、典型的に単語の境界を区切るためにスペースを使用しない。

10

## 【 0 0 0 4 】

漢字は、非常に大量の文字セットで、約何万もの文字に及ぶ。したがって、演算装置に日本語テキストを入力するためのシステムは、概して、ラテン文字(ローマ字と呼ばれる)またはかなを入力として受信し、入力を見出し語に変換する。図1の左側の列に示されるように、ローマ字は、ラテン文字を使用する日本語の音標表現である。ローマ字で表記した日本語は読むことが困難であるため、ローマ字は概して入力だけに使用される。例えば、ローマ字は、典型的に、QWERTYレイアウトを有するキーボード上で使用される。

20

## 【 0 0 0 5 】

図1の中央の列は、「読むための」の日本語の用語である、「読み」の例である。読みとは、かな文字を使用する日本語テキストの音標表現を言う。かなは、一般的には、12キーのキーボードを有するモバイルデバイス上で使用されるが、QWERTYキーボードを使用してテキストを入力するためにも使用される場合がある。12キーのレイアウトでは、キーボードは通常キーあたり5つのかなを装備する。ユーザは、所望のかなが表示されるまで、選択されたキーを複数回たたくことにより、5つのかなから特定の文字を選択することができる。例えば、図1の中央の列に表示される読みは、5つの異なる一式のキーを押すことにより入力され得る、5つの個別のかなを有する。

30

## 【 0 0 0 6 】

日本語テキストを入力するためのシステムは、ローマ字、読み、および見出し語の間を変換するように、変換エンジンを提供する。概して、単一の読みに変換する、多数の異なるローマ字が存在し得る。しかしながら、ローマ字から読みへの音訳方法はかなりよく定義されているため、入力システムは、ローマ字から読みに容易に変換することができる。例えば、図1の左側の一式の矢印は、左側の列内の3つのローマ字単語が、中央の列内の単一の読みにマップすることを示す。入力システムの中には、頻度分析等の曖昧性解消方法を使用して、ローマ字の中の一般的なユーザエラーを修正することが可能なものもある。

40

## 【 0 0 0 7 】

対照的に、読みと見出し語との間には、多対多の関係が存在する。図1の中央の列から図1の右側の列への矢印により示されるように、中央の列内の読みは、少なくとも5つの異なる見出し語に変換され得る。可能な見出し語は、かな、漢字、およびアラビア数字を含む、いくつかの文字セットからの文字を含む。加えて、図1は、3つの可能な読みは、右側の列の一番下の単一の見出し語にマップできることを示す。概して、1つの読みに対して、一致し得る見出し語は少なくとも2~4つ存在するだろうが、多数の潜在的な一致が存在するかもしれない。

【発明の概要】

【発明が解決しようとする課題】

50

## 【0008】

表記される日本語の複雑性は、セルラフォン、スマートフォン、ポータブルメディアプレーヤ、ポータブル電子メールデバイス、ポータブルゲーム機等のモバイルデバイス上で使用される場合、これらの装置はしばしばユーザ入力のために数字キーボードまたは低減キーボードを使用するため、特に問題となる。これらの入力コンポーネントを使用して日本語テキストを入力することは、複雑で、非常に時間がかかり得る。これらの入力方法を使用してテキストを検索することは同様に問題であり得る。このように、モバイルデバイスにおいて日本語テキストを入力し、モバイルデバイス上で特定のテキストを検索するプロセスを簡素化できるシステムを有することは有用である。

## 【課題を解決するための手段】

## 【0009】

本明細書において、日本語テキスト等の複雑な言語テキストを処理するための方法およびシステムが開示される。以下の詳細説明は、本発明の多様な実施形態の完全な理解および有効な説明のための具体的な詳細を提供する。当業者は、しかしながら、これらの詳細の多くを提供しなくても本発明が実践され得ることを理解するであろう。加えて、一部の周知の構造または機能は、多様な実施形態の関連説明を不必要に不明瞭にすることがないように、詳細に表示または説明されない場合がある。以下に提示される説明において使用される用語は、本発明の所与の具体的な実施形態の詳細説明に関連して使用されているが、その広義の妥当な様式において解釈されるように意図される。

本願発明は、例えば、以下の項目を提供する。

## (項目1)

モバイルデバイスを介して複数の項目を検索するためのコンピュータ実装方法であって、該複数の項目のうちの個別の項目は、日本語テキスト部分の特徴とし、該コンピュータ実装方法は、

日本語テキストにより特徴付けられる項目を識別するためにモバイルデバイス上で検索クエリを受信することと、

該受信した検索クエリに基づいてテキスト検索語を生成することと、

該テキスト検索語に基づいて、一式の項目から複数の一致項目を決定することであって、該複数の一致項目の各々は、該テキスト検索語に対応する一致テキストを有する日本語テキスト部分を含む、ことと、

該複数の一致項目について、

該日本語テキスト部分内の開始点を決定することと、

該開始点に対する該一致テキストの位置を決定することと、

該開始点に対する該決定した位置に基づいて、該複数の一致項目の中の該一致項目の優先順位を決定することと、

該決定した優先順位に基づいて順序付けられる一致項目のリストを提供することと

を含む、コンピュータ実装方法。

## (項目2)

前記一式の項目は、メディアファイル、アドレス帳入力、文書ファイル、またはアプリケーションのうちの少なくとも1つを含む、項目1に記載のコンピュータ実装方法。

## (項目3)

前記開始点を決定することは、

前記日本語テキスト部分内の句読点を識別することと、

該識別した句読点に近接する該開始点の位置を特定することと

を含む、項目1に記載のコンピュータ実装方法。

## (項目4)

前記開始点を決定することは、

前記日本語テキスト部分内の文字体系の変化を識別することと、

該識別した文字体系の変化における該開始点の位置を特定することと

を含む、項目1に記載のコンピュータ実装方法。

(項目5)

前記開始点を決定することは、  
前記日本語テキスト部分内の電話番号の形式を有する文字列を識別することと、  
該識別した文字列に近接する該開始点の位置を特定することと  
を含む、項目1に記載のコンピュータ実装方法。

(項目6)

前記優先順位を決定することは、  
前記一致テキストから該一致テキストの前の最も近い開始点までの文字数距離を計算することと、  
該計算した文字数距離に基づいて該優先順位を決定することと  
を含む、項目1に記載のコンピュータ実装方法。

10

(項目7)

前記優先順位を決定することはさらに、前記一致テキストが前記日本語テキスト部分の先頭にある場合、一致項目により高い優先度を割り当てることを含む、項目1に記載のコンピュータ実装方法。

(項目8)

前記開始点を決定することはさらに、前記日本語テキスト部分内で第1の開始点および第2の開始点を決定することを含み、前記優先順位を決定することはさらに、前記一致テキストが該第1の開始点と該第2の開始点との間のテキスト全てを含む場合、前記一致項目により高い優先度を割り当てることを含む、項目1に記載の方法。

20

(項目9)

モバイルデバイスから複数の項目を検索するためのシステムであって、  
ユーザ検索クエリに基づいて検索語を生成するように構成される変換コンポーネントと

該生成した検索語に基づいて、該モバイルデバイスを介してアクセス可能な複数の一致項目の位置を特定するように構成される検索コンポーネントであって、該複数の一致項目の各々は、該検索語に対応する一致日本語テキストを含むテキストフィールドを含む、検索コンポーネントと、

該複数の一致項目の各々の該テキストフィールド内の開始点を決定するように構成される開始点決定コンポーネントと、

30

該決定した開始点と、該複数の一致項目の各々に対する該一致日本語テキストとの間の距離を計算するように構成される距離計算コンポーネントと、

該計算した距離に基づいて該複数の一致項目の順序を決定し、該決定した順序に基づいて該複数の一致項目の少なくともいくつかをユーザに出力するように構成される、順序付けコンポーネントと

を備える、システム。

(項目10)

前記複数の一致項目は、メディアファイル、アドレス帳入力、文書ファイル、画像ファイル、またはアプリケーションのうちの少なくとも1つを含む、項目9に記載のシステム

40

(項目11)

前記開始点決定コンポーネントは、  
前記テキストフィールド内で句読点を識別することと、  
前記識別した句読点に近接する前記開始点の位置を特定することと  
によって、前記開始点を決定するように構成される、項目9に記載のシステム。

(項目12)

前記開始点決定コンポーネントは、  
前記テキストフィールド内の文字体系の変化を識別することと、  
前記識別した文字体系の変化における前記開始点の位置を特定することと  
によって、前記開始点を決定するように構成される、項目9に記載のシステム。

50

(項目 13)

前記開始点決定コンポーネントは、  
前記テキストフィールド内の電話番号を示す文字列を識別することと、  
前記識別した文字列に近接する前記開始点の位置を特定することと  
によって、前記開始点を決定するように構成される、項目 9 に記載のシステム。

(項目 14)

前記順序付けコンポーネントは、  
前記一致テキストから該一致テキストの前の最も近い開始点までの文字数距離を計算することと、  
該計算した文字数距離に基づいて前記優先順位を決定することと  
によって、該順位を決定するように構成される、項目 9 に記載のシステム。

10

(項目 15)

前記順位を決定することはさらに、前記一致する日本語テキストが前記テキストフィールドの先頭にある場合、一致項目により高い優先度を割り当てることを含む、項目 9 に記載のシステム。

(項目 16)

モバイルデバイスを通じてアクセス可能な一式の項目の中で検索するように、該モバイルデバイスのプロセッサを制御するための命令を含む、コンピュータ読み取り可能記憶媒体であって、該一式の項目の個別の項目は、日本語テキストを特徴とし、  
日本語テキストにより特徴付けられる項目を識別するように該モバイルデバイス上で検索クエリを受信することと、

20

該受信した検索クエリに基づいてテキスト検索語を生成することと、  
該テキスト検索語に基づいて、該一式の項目から複数の一致項目を決定することであって、該複数の一致項目の各々は、該テキスト検索語に対応する一致テキストを有する日本語テキスト部分を含む、ことと、

該複数の一致項目の各々について、  
該日本語テキスト部分内の開始点を決定することと、  
該開始点に対する該一致テキストの位置を決定することと、  
該開始点に対する該決定した位置に基づいて、該複数の一致項目の中の該一致項目の優先順位を決定することと、

30

該決定した優先順位に基づいて順位付けられる、一致項目のリストを提供することと  
を含む方法による、コンピュータ読み取り可能記憶媒体。

(項目 17)

前記一式の項目は、メディアファイル、アドレス帳入力、文書ファイル、またはアプリケーションのうちの少なくとも 1 つを含む、項目 16 に記載のコンピュータ読み取り可能記憶媒体。

(項目 18)

前記開始点を決定することは、  
前記日本語テキスト部分内の句読点を識別することと、  
該識別した句読点に近接する該開始点の位置を特定することと  
を含む、項目 16 に記載のコンピュータ読み取り可能記憶媒体。

40

(項目 19)

前記開始点を決定することは、  
前記日本語テキスト部分内の文字体系の変化を識別することと、  
該識別した文字体系の変化における前記開始点の位置を特定することと  
を含む、項目 16 に記載のコンピュータ読み取り可能記憶媒体。

(項目 20)

前記開始点を決定することは、  
前記日本語テキスト部分内の電話番号の形式を有する文字列を識別することと、  
該識別した文字列に近接する前記開始点の位置を特定することと

50

を含む、項目 1 6 に記載のコンピュータ読み取り可能記憶媒体。

(項目 2 1)

前記優先順位を決定することは、

前記一致テキストから該一致テキストの前の最も近い開始点までの文字数距離を計算することと、

前記計算した文字数距離に基づいて前記優先順位を決定することと

を含む、項目 1 6 に記載のコンピュータ読み取り可能記憶媒体。

(項目 2 2)

前記優先順位を決定することはさらに、前記一致テキストが前記日本語テキスト部分の先頭にある場合、一致項目により高い優先度を割り当てることを含む、項目 1 6 に記載のコンピュータ読み取り可能記憶媒体。

10

(項目 2 3)

前記開始点を決定することはさらに、前記日本語テキスト部分内で第 1 の開始点および第 2 の開始点を決定することを含み、前記優先順位を決定することはさらに、前記一致テキストが前記第 1 の開始点と前記第 2 の開始点との間のテキストの全てを含む場合、前記一致項目により高い優先度を割り当てることを含む、項目 1 6 に記載のコンピュータ読み取り可能記憶媒体。

【図面の簡単な説明】

【0010】

【図 1】図 1 は、日本語テキストを表すための先行技術を示す。

20

【図 2】図 2 は、日本語テキストを処理するための好適なモバイルデバイスの正面図である。

【図 3】図 3 は、モバイルデバイスが動作する代表的環境のネットワーク図である。

【図 4】図 4 は、モバイルデバイスの例示的な設計図を示す高レベルのブロック図である。

【図 5】図 5 は、予測テキスト入力システムを使用して、日本語のテキストを入力する 3 つの段階を表す図である。

【図 6】図 6 は、単一行の見出し語を使用して、予測テキスト入力システムの結果を表す、代表的なユーザインターフェースである。

【図 7】図 7 は、日本語のための予測テキスト入力システムの論理的ブロック図である。

30

【図 8】図 8 は、予測テキスト入力システムにより実行されるプロセスのフローチャートである。

【図 9】図 9 は、日本語テキストを検索するように構成される検索システムによる、モバイルデバイス上での検索の結果を表す、代表的なユーザインターフェースである。

【図 10】図 10 は、モバイルデバイス上で日本語テキストを検索するための検索システムの論理的ブロック図である。

【図 11】図 11 は、検索システムにより実行されるプロセスのフローチャートである。

【発明を実施するための形態】

【0011】

1. 代表的なモバイルデバイスおよびワイヤレス環境

40

図 2 は、日本語テキストを処理するための好適なモバイルデバイス 200 の正面図である。図 2 に示されるように、モバイルデバイス 200 は、筐体 201、複数の押しボタン 202、方向キーパッド 204 (例えば、5 方向キー) マイク 205、スピーカー 206、および筐体 201 により支えられるディスプレイ 210 を含むことができる。モバイルデバイス 200 はまた、PDA 電話、セルラフォン、スマートフォン、ポータブルメディアプレーヤ、ポータブルゲーム機、ポータブル電子メールデバイス (例えば、Blackberry)、または他のモバイル通信デバイスに概して見られる、他のマイク、送受信機、写真センサ、および/または他の演算コンポーネントも含むことができる。

【0012】

ディスプレイ 210 は、液晶ディスプレイ (LCD)、電子インクディスプレイ、およ

50

び/またはユーザインターフェースを提示するように構成される他の好適な種類のディスプレイを含む。モバイルデバイス200はまた、ユーザからの入力を受信するように構成されるタッチセンサコンポーネント209も含むことができる。例えば、タッチセンサコンポーネント209は、抵抗性、容量性、赤外線、弾性表面波(SAQ)、および/または別の種類のタッチ画面を含むことができる。タッチセンサコンポーネント209は、ディスプレイ210と一体化されてもよく、またはディスプレイ210から独立していてもよい。示される実施形態において、タッチセンサコンポーネント209およびディスプレイ210は、概して、同様な大きさのアクセス領域を有する。他の実施形態において、タッチセンサコンポーネント209およびディスプレイ210は、異なる大きさのアクセス領域を有することができる。例えば、タッチセンサコンポーネント209は、ディスプレイ210の境界を超えて拡張するアクセス領域を有することができる。モバイルデバイス200はまた、ユーザからテキストまたは数字の入力を受信することが可能である、12キーの数字キーパッド212も含む。代替として、モバイルデバイス200は、ユーザ入力を受信するための完全なQWERTYキーボードを含むことができる。ハードウェアキーボードまたはキーボードの代わりに、またはこれに加えて、モバイルデバイス200はまた、ユーザにタッチセンサコンポーネント209を経由してテキストまたは数字の入力を提供することを可能にするように、ディスプレイ210上にソフトウェアキーボードまたはキーパッドを提供することもできる。

10

#### 【0013】

図3は、モバイルデバイスが動作する代表的環境300のネットワーク図である。複数のモバイルデバイス200は、ワイヤレスネットワークによりカバーされる領域でローミングする。モバイルデバイスは、例えば、セルラフォン、PDA電話、スマートフォン、ポータブルメディアプレーヤ、ポータブルゲーム機、ポータブル電子メールデバイス(例えば、BlackBerry)または他のモバイルインターネットデバイスである。モバイルデバイス200は、ワイヤレス接続306を経由して、送受信器310と通信する。ワイヤレス接続306は、デジタルデータを伝送するための任意のワイヤレスプロトコルを使用して実装され得る。例えば、接続は、GSM、UMTSまたはCDMA2000等のセルラーネットワークプロトコル、またはWiMax(IEEE802.16)、WiFi(IEEE802.11)またはBluetooth等の非セルラーネットワークプロトコルを使用し得る。これらのモバイルデバイスではワイヤレス接続が最も一般的であるが、デバイスはイーサネット(登録商標)等の有線接続を使用して通信することもできる。

20

30

#### 【0014】

送受信機310は、ワイヤレスネットワークのためにバックホールを提供する、1つ以上のネットワークに接続される。例えば、送受信機310は、モバイルネットワークと遠隔の電話機316との間に接続を提供する、公衆交換電話網(PSTN)312に接続されてもよい。モバイルデバイス200のユーザが音声通話呼を行う際、送受信機310は、ワイヤレスネットワークの音声バックホール(図示せず)を経由して、呼をPSTN312に送る。PSTN312は、次いで、呼を遠隔の電話機316に自動的に接続する。遠隔の電話機316が別のモバイルデバイスである場合、呼は、第2のワイヤレスネットワークバックホールを経由して別の送受信機に送られる。

40

#### 【0015】

送受信機310はまた、リモートサービス318または他のデバイスへのパケットベース接続を提供する、1つ以上のパケットベースネットワーク314にも接続される。モバイルデバイス200から送受信機310に伝送されるデータは、ワイヤレスネットワークのデータバックホール(図示せず)を経由して、パケットベースネットワーク314(例えば、インターネット)に送られる。パケットベースネットワーク314は、電子メールサーバ320、ウェブサーバ322、およびインスタントメッセージサーバ324等のリモートサービス318にワイヤレスネットワークを接続する。もちろん、リモートサービス318は、ファイル転送プロトコル(FTP)サーバまたはストリーミングメディア

50

サーバ等の、インターネットまたは他のネットワーク上で利用可能な任意の他のアプリケーションを含むことができる。

【0016】

図4は、モバイルデバイス200の例示的な設計図を示す高レベルのブロック図である。モバイルデバイス200は、相互接続406に連結されるプロセッサ402およびメモリ404を含む。図4に示される相互接続406は、適切なブリッジ、アダプタ、またはコントローラにより接続される、任意の1つ以上の個別の物理的バス、ポイントツーポイント接続、または両方を表す抽象的概念である。プロセッサ402は、モバイルデバイス200の中央処理装置(CPU)を含むことができ、従って、ソフトウェアまたはファームウェアを実行することにより、モバイルデバイス200の全体的な動作を制御することができる。プロセッサ402は、1つ以上のプログラマブル汎用または特殊マイクロプロセッサ、デジタルシグナルプロセッサ(DSP)、プログラマブルコントローラ、特定用途向け集積回路(ASIC)、プログラマブルロジックデバイス(PLD)等、またはこのようなデバイスの組み合わせであってもよく、または含むことができる。

10

【0017】

メモリ404は、任意の形式の固定または取り外し可能ランダムアクセスメモリ(RAM)、読み取り専用メモリ(ROM)、フラッシュメモリ等、またはこのようなデバイスの組み合わせを表す。プロセッサにより実行されるソフトウェアまたはファームウェアは、記憶領域410および/またはメモリ404に格納されてもよく、典型的には、オペレーティングシステム408ならびに1つ以上のアプリケーション418を含む。ソフトウェアまたはオペレーティングシステムにより利用されるデータ414もまた、記憶領域またはメモリに格納される。記憶領域410は、フラッシュメモリ、ハードドライブ、または他の大容量記憶デバイスであってもよい。

20

【0018】

モバイルデバイス200は、ユーザに装置の制御を可能にする、入力装置412を含む。入力装置412は、キーボード、トラックパッド、タッチセンサ画面、または他の標準の電子入力装置を含むことができる。モバイルデバイス200はまた、ディスプレイ210(図2)等のユーザインターフェースを表示するために好適な表示装置414も含む。ワイヤレス通信モジュール416は、短距離または長距離ワイヤレスプロトコルを使用して、ネットワーク上からリモートデバイスと通信する能力をモバイルデバイス200に提供する。

30

【0019】

2. 予測テキスト入力

日本語のモバイルデバイスのための予測テキスト入力を提供するためのシステムおよび方法が開示される(以下、「テキスト入力システム」または「システム」と参照される)。以下に詳細が説明されるように、数字キーパッドを有する日本語のモバイルデバイスのユーザの場合、テキスト入力は、概して2段階のプロセスである。第1段階において、モバイルデバイスは、ユーザ入力を1つ以上の読みに変換し、これがユーザに表示される。第2段階において、モバイルデバイスは、選択された読みに対応する見出し語のリストを表示する。ユーザは、次いで、第2のリストから、所望の見出し語を選択する。本明細書において開示されるテキスト入力システムは、このプロセスを単一段階に短縮する。ユーザ入力を受信した後、テキスト入力システムは、受信した入力に対応する全ての読みを決定する。テキスト入力システムは、次いで、可能な読みの全てに対応する、一式の一致見出し語を決定し、一式の見出し語の部分または全てをユーザに表示する。テキスト入力システムは、対応する読みに応じて、見出し語をグループ化することができる。代替として、システムは、ユーザが選択する可能性がより高い見出し語の予想に基づいた順序で見出し語を表示してもよく、可能性が高い一致は、可能性が低い一致よりもリストの上位に表示される。システムはまた、最も可能性が高い見出し語だけを表示し、可能性が低い結果を隠すように構成されてもよい。

40

【0020】

50

コンピュータシステムに日本語テキストを入力するための明示的なローマ字方法では、ユーザは、Q W E R T Yキーボード上でローマ字を使用して日本語を入力する。システムは、次いで、ローマ字をかなに自動的に変換し、その後、変換エンジンは、かなを見出し語に自動的に変換することができる。明示的な読み入力方法では、ユーザは、およそ50文字のかな文字を装備するQ W E R T Yキーボード上で個別のかなを選択する。明示的な読み方法は、電話機ではまれであるが、他の家庭用電子機器では一般的である。低減キーボードを有する携帯電話機または他の装置上では、ユーザは、上記のマルチタップ方法を使用してテキストを入力することができる。その場合、ユーザは、所望のかなを入力するために、かなあたり1回から5回まで単一のキーをたたいて、かなのリストを反復させる。これらの方法の各々の場合、システムは、入力されたかなに対して、可能性が高い見出し語変換のリストを表示する。ユーザは、次いで、リストから所望の見出し語を選択することができる。

10

#### 【0021】

ユーザはまた、Massachusetts州BurlingtonのNuance CommunicationsからライセンスされるT9システム等の予測入力システムを使用してテキストを入力してもよい。予測入力システムは、部分入力に基づいて完全な単語を予想することにより、入力を簡素化する。12キーのキーパッド(モバイルデバイスのような)を装備するモバイルデバイスは、マルチタップ方法に加えて、日本語のためのT9システムに対応することができる。予測入力システムを使用する場合、ユーザは、読みの中のかなあたり1つのキーを入力する。日本語のT9エンジンは、一致する読みを活用変化または組み合わせるように、単語リストおよび文法の組み合わせを使用する。プロセスでは、所望の見出し語を予想しようとする。しかしながら、変換プロセスは、複数の可能性を生成し、不明瞭になる場合がある。多数の可能な一致が存在する場合、ユーザは、所望の読みを選択してから、選択した読みに一致するように所望の見出し語を選択しなければならない。

20

#### 【0022】

図5は、T9システムを使用する日本語テキスト入力の2段階のプロセスで使用され、かつ本明細書において開示されるテキスト入力システムの一段階のプロセスにおいて使用されるような代表的テキストデータを表すチャート500である。図5の列505は、特定の一式のキーを押した結果として生成される読みの例示的リストを示す。上記のように、読みは、可能な一致を予測するように、単語リストおよび文法の組み合わせを使用して生成される。一部の読みは、つづり修正または単語補完を使用して生成されてもよく、つまり、つづり修正は、誤って入力される文字を修正するために使用されてもよく、単語補完は、その最初の文字に基づいて完全な単語を提供するために使用されてもよい。読みのリストはまた、その地域をつづりから標準的な日本語のつづりを生成することにより、つづりにおける地域的な差を修正するように構成されてもよい。リスト上の読みは、読みがユーザの入力に一致する可能性に応じて順序付けられてもよい。つまり、列505の第1の読みは、統計的に、ユーザの入力に一致する可能性が最も高い場合があり、列505の最後の読みは、ユーザの入力に一致する可能性が最も低い場合がある。図5の列510は、生成された読みに等しいローマ字を示し、一方で、列515は、読みに関連する見出し語を示す。図5に示されるように、特定の読みには、一致する可能性のある見出し語が不定数ある。読みリストと同様に、見出し語もまた各見出し語が選択される可能性に応じて、順序付けられてもよい。つまり、列515内の各リストの第1の見出し語は、統計的に、ユーザの入力に一致する可能性が最も高い場合があり、列515内の各リストの最後の見出し語は、ユーザの入力に一致する可能性が最も低い場合がある。

30

40

#### 【0023】

T9システムの2段階プロセスを使用すると、日本語テキストを入力するユーザには、最初に、列505から選択された読みのリストが提示される。ユーザが、表示された選択肢から読みを選択したら、T9システムは、選択された読みに関連する見出し語(列515内に含まれるような)のリストを表示する。ユーザは、次いで、表示されたリストから

50

、所望の見出し語を選択する。ユーザが見出し語を選択する前に最初に読みを選択することに伴う問題は、所望の見出し語を入力するために、ユーザに2段階を完了することを要求することである。2段階のプロセスは、ユーザが長いメッセージを入力しようとする場合、時間がかかる可能性がある。したがって、所望のテキストを入力するために必要な動作の数を削減する、日本語テキストを入力するための方法を提供することは有用であろう。

#### 【0024】

図6は、単一行の見出し語を使用して、予測テキスト入力システムの結果を表示する、代表的なユーザインターフェース600である。示されるインターフェース600では、T9システムに関して検討された2段階のプロセスは、ユーザに表示される単一の組み合わせリストを使用することによって、1段階のプロセスに崩壊される。図6に示されるように、見出し語の単一のリスト605は、テキスト入力システムによりユーザに表示される。一式の見出し語は、対応する読み別にグループ化される（グループ化された一式の見出し語は、図中、明確化のために丸で囲まれている）。このように、インターフェースに示される最初の4つの可能性（丸で囲まれたセット610として示される）は、ローマ字「h o u t a i」に関連する。次の5つの見出し語（丸で囲まれたセット615として示される）は、ローマ字「j o u t a i」に関連し、次の2つの見出し語（丸で囲まれたセット620として示される）は、ローマ字「k o u t a i」に関連する。追加の見出し語のグループ化は、リスト605で、表示画面上の左から右へと続く。表示されたインターフェースを使用して、ユーザは、最初に対応する読みを選択する必要なく、表示されたリストから所望の見出し語を選択することができる。

#### 【0025】

図6には単一のリストが水平方向に表示されているが、リストは縦方向に表示されてもよく、またはユーザに組み合わせリストをスクロールすることを可能にするように、スクロール機能を有してもよいことが理解されるであろう。例えば、各セットは、ディスプレイ上の異なる行に表示されてもよく、ユーザは、セットリスト内をスクロールすることが可能になってよい。

#### 【0026】

表示される見出し語の各グループに対して、テキスト入力システムは、対応する見出し語の全て、または対応する見出し語のサブセットを表示することができる。例えば、セット610の内容は、チャート500の行520から選択される。セット610は、列515から選択される、関連の見出し語のうち2つを含む。セット615の内容は、チャート500の行525から選択される。セット615は、ローマ字「j o u t a i」に関連する、列515から選択されるような見出し語のうち4つを含む。セット620の内容は、チャート500の行530から選択される。セット615は、列515から選択されるような見出し語のうち2つを含む。ユーザへの合図として、テキスト入力システムはまた、最も可能性が高いローマ字および/または読みも表示することができる。例えば、セット610は、列510から選択されたローマ字「h o u t a i」を含み、その次に、列505から選択された関連の読みが続く。

#### 【0027】

使用可能な見出し語のうちサブセットが表示される際、テキスト入力システムは、表示される見出し語がユーザによって選択される可能性に基づいて、サブセットを選択することができる。組み合わせリストはまた、選択される可能性に基づいた優先順位で、部分または全ての使用可能な見出し語を表示することもできる。例えば、テキスト入力システムは、リストの先頭に可能性が高い一致を配置し（読み順にグループ化される）、最後に残りの一致を配置することにより（全ての読みで選択の可能性順にグループ化される）、組み合わせリスト605を生成することができる。代替として、テキスト入力システムは、可能な見出し語の全リストに基づいて（つまり、つづり修正、地域修正、または単語補完に基づいて含まれる単語を含む）可能性が高い一致を表示するが、ユーザの入力に正確に一致する読みを有する残りの見出し語だけを表示してもよい。

## 【 0 0 2 8 】

組み合わせリストに表示される見出し語は、いくつかの因子に基づいて順序付けられてもよく、この因子は（特定の順序ではない）以下を含む。

- ・読みリスト内のインデックス（例えば、システムは、ユーザの入力に一致する可能性が高い読みの見出し語を多く表示する場合がある）。
- ・見出し語リストのインデックス（例えば、システムは、任意の特定の読みに関連する見出し語の数を限定して表示する場合がある）。
- ・キーシーケンスが有効なローマ字であるかどうか。
- ・読みが単語リストにあるかどうか（例えば、システムは、システムの単語リストまたは辞書で検出されない読みの見出し語を表示しない場合がある）。
- ・読みが地域修正に基づいて生成されたかどうか。
- ・読みがつづり修正に基づいて生成されたかどうか。
- ・読みが単語補完に基づいて生成されたかどうか。

組み合わせリスト 6 0 5 を生成するために、システムは、各使用可能な見出し語に対する上記の因子のうち 1 つ以上に、数値を割り当てることができる。数値は、各因子が見出し語により満たされるかどうかに基づいていてもよく、または、数値は、見出し語に対する因子の実際の値に基づいていてもよい（例えば、インデックス値に基づく因子の場合）。各因子は、因子の認識される重要度に応じて加重されてもよく、各見出し語に対する総合的な関連点数は、全ての関連する因子の加重された数値を合計することにより計算される。システムは、次いで、関連点数をしきい関連値に比較することにより、組み合わせリストに対して可能性が高い見出し語を決定することができる。システムは、読みに応じて（図 6 に示されるように）グループ内に可能性が高い見出し語を含む組み合わせリストを表示する。上記のように、残りの見出し語は、次いで、組み合わせリスト内の、可能性が高い見出し語が表示された後に表示される。代替として、組み合わせリストの項目は、総合的な関連点数により順序付け（つまり、順位付け）されてもよい。

## 【 0 0 2 9 】

図 7 は、モバイルデバイス 2 0 0 上に実装されてもよい、テキスト入力システム 7 0 0 の論理的ブロック図である。システムの態様は、専用ハードウェア回路、プログラマブル回路、またはこれらの組み合わせとして実装されてもよい。本明細書において更に詳細が検討されるように、テキスト入力システム 7 0 0 は、システムの機能を促進するように、いくつかのモジュールを含む。多様なモジュールは、単一の装置に存在すると説明されるが、モジュールは必ずしも物理的に共存する必要はない。一部の実施形態において、多様なモジュールは、複数の物理的装置に分散され得て、モジュールにより実装される機能性は、リモートサービスを呼び出すことによって提供されてもよい。同様に、データ構造は、モバイルストレージまたはリモートストレージに格納され得て、1 つ以上の物理的装置内に分散される。プログラマブル実装を想定すると、このシステムの機能性をサポートするためのコードは、光学式ドライブ、フラッシュメモリ、またはハードドライブ等のコンピュータ読み取り可能媒体上に格納されてもよい。当業者は、これらの個々のコンポーネントおよびサブコンポーネントのうち少なくとも一部は、特定用途向け集積回路（ASIC）、プログラマブルロジックデバイス（PLD）、またはソフトウェアおよび/またはファームウェアで構成される汎用プロセッサを使用して、実装されてもよいことを理解するであろう。

## 【 0 0 3 0 】

図 7 に示されるように、テキスト入力システム 7 0 0 は、図 2 に示されるキーパッド 2 1 2 等の入力コンポーネント 7 0 2 を介して、ユーザ入力を受信する。上記のように、キーボードまたはキーパッドは、ハードウェアキーパッド 2 1 2 として、またはタッチセンサコンポーネント 2 0 9 を介して使用される表示キーパッドとして、実装されてもよい。テキスト入力システム 7 0 0 は、ディスプレイ 2 1 0 等の表示コンポーネント 7 0 4 を介して、ユーザに見出し語の順序付けリストを出力する。システム 7 0 0 は、テキスト入力システムの動作に関連する構成およびデータを格納するように構成される、記憶コンポー

10

20

30

40

50

ネット706にアクセスすることができる。

【0031】

テキスト入力システム700は、入力コンポーネント702からユーザのキーストロークを受信するように構成される、読み変換コンポーネント710を含み、受信されたキーストロークに基づいて、一式の可能な読み変換を決定する。一式の可能な読み変換は、受信されたキーストロークを一式の可能な読みに翻訳するように、記憶コンポーネント706に格納される読み参照テーブルを使用して決定されてもよい。テキスト入力システム700はまた、読み変換コンポーネント710により生成される一式の可能な読みに対応する見出し語のリストを決定するように構成される、見出し語参照コンポーネント712も含む。このために、見出し語参照コンポーネント712は、記憶コンポーネント706に格納される1つ以上の辞書を使用することができる。見出し語参照コンポーネントはまた、見出し語のリストを生成するために、つづり修正および地域修正も実行することができる。このように、見出し語参照コンポーネント712は、正確な一致を決定することに加えて、各読みに近似した一致を検索することができる。

10

【0032】

テキスト入力システム700はまた、ユーザに表示するための見出し語リストの順序付けまたはグループ化を決定するように構成される、順序付けコンポーネント714も含むことができる。このために、順序付けコンポーネント714は、見出し語の各々について関連点数を決定するための上記の因子（例えば、読みリストのインデックス、見出し語リストのインデックス等）を決定するように構成される、基準コンポーネント716と相互作用する。次いで、順序付けコンポーネント716は、関連点数に基づいて見出し語の順序付けリストを生成する。順序付けコンポーネント716は、最も関連性の高い見出し語だけが表示されるように、表示コンポーネント704に提供される見出し語の数を制限することができる。

20

【0033】

図8は、テキスト入力システム700により実行されるプロセス800のフローチャートである。プロセスはブロック802で開始し、テキスト入力システムは、入力コンポーネント702から入力を受信する。入力は、1つ以上の不明瞭なキーストロークの形式であってもよい。ブロック804で、テキスト入力システムは、受信したキーストロークに対応する一式の読みを決定する。一式の読みを決定する際、システムは、同一ではないが類似の入力シーケンスに対応する読みを決定することによって、つづり修正を実行しようと試行することができる。システムはまた、入力シーケンスで開始する可能な単語を予測することによって、読みを決定することもできる。

30

【0034】

次いで、プロセスはブロック806に進み、テキスト入力システムは、ステップ804において決定された読みに一致する、一式の見出し語を識別する。上記のように、システムは、読みに基づいてインデックスが付けられた1つ以上の辞書を検索することによって、一致見出し語を決定することができる。一部の実施形態においては、一式の見出し語は、検索に使用されている読みに正確に対応する見出し語だけを含む。他の実施形態においては、システムはまた、特定の読みで始まる、または特定の読みを含む見出し語も呼び出す。

40

【0035】

次いで、プロセスはブロック808に進み、システムは、一式の見出し語に対する順序を決定する。上記のように、システムは、見出し語の関連性を順位付けるために、見出し語の各々に対して関連点数を計算することができる。最高の関連点数を有する見出し語は、リスト内で昇格されてもよく、最低の関連点数を有する見出し語は、リスト内で降格されてもよい。次いで、システムはブロック810に進み、順序付けられた見出し語リストをユーザに表示する。これによって、ユーザは、最小限の努力で所望の見出し語を迅速かつ容易に選択することが可能になる。

【0036】

50

### 3. テキスト検索

モバイルデバイス上で日本語テキストを入力することに加えて、ユーザはまた、モバイルデバイス上で特定のテキストを検索かつ検出することを望む場合がある。ユーザに特定のテキストの場所をより容易に特定することを可能にするように、モバイルデバイスを介して日本語テキストを検索するためのシステムおよび方法が開示される（以下、「検索システム」または「システム」と参照される）。検索システムは、モバイルデバイス上のキーボードまたはキーボードを経由してユーザ入力を受信し、入力を一式的検索語に変換する。一部の実施形態において、システムは、上記のテキスト入力システムを使用して、入力を見出し語に変換する。しかしながら、特定のシーケンスを選択するために見出し語のリストをユーザに提供する代わりに、システムは、生成したリストを一式的検索語として使用する。検索語を生成した後、システムは、一致項目を検出するために、モバイルデバイスによりアクセス可能な項目の中でテキストフィールドを検索する。次いで、システムは、各一致項目のテキストフィールド内で1つ以上の自然な開始点を決定する。以下に詳細を説明するように、開始点は、テキストフィールドの先頭、および文字セット内の句読点または変化の位置を含むことができる。開始点を決定した後、システムは、各一致項目の一致テキストと自然な開始点との間の距離を決定する。システムは、次いで、計算された距離、および一致の配列、項目の種類、項目がその前に使用された回数等の他の要素に基づいて、順番付けた一式的検索結果を提供する。一部の実施形態において、システムは、複数の検索語を使用して結果のリストを生成する。次いで、複数の検索語の各々に対する距離および他の因子を組み合わせることによって順序付けが決定される。

【0037】

図9は、日本語テキストを検索するように構成される検索システムにより、モバイルデバイス上で検索の結果を表示する、代表的なユーザインターフェース900である。検索システムは、モバイルデバイスによりアクセス可能な項目を検出するために使用されてもよい。これらの項目は、モバイルデバイス上にローカルに、またはネットワーク接続を経由してアクセス可能なリモートストレージに格納されてもよい。本明細書に使用される場合、「項目」とは、デバイス機能、アプリケーション、またはデータ（アドレス帳入力と、ファイルと、文書と、音楽ファイル、画像ファイル、ビデオファイル等のメディアファイルと、等を含む）等のモバイルデバイスに関連するデータオブジェクトである。個別の項目は、検索に使用されてもよい、1つ以上のテキストフィールドを有することができる。本明細書に使用される場合、「テキストフィールド」とは、特定のテキスト情報を格納するために割り当てられるスペースである。例えば、音楽ファイルは、曲名、アーティスト、またはアルバムを格納するための複数のテキストフィールドを有することができる。同様に、アドレス帳入力は、名前、電話番号、または電子メールアドレスを格納するための複数のテキストフィールドを有することができる。テキストフィールドは、ファイルの一部として、または個別のインデックスに格納されてもよい。

【0038】

図9に示される実施例において、ユーザは、モバイルデバイス上の「5」および「6」のキーを選択したことになっている。キーの選択は、テキスト入力領域905の表示「56」によって反映される。「5」および「6」のキーを選択することによって、ユーザは、検索システムに、「5」および「6」のキーに関連する文字の組み合わせを検索するように命令したことになる。各キーに関連する文字は、キー上の数字の上の位置915のキー上に反映される。したがって、「5」および「6」のキーに関連する文字は、「ko」、「km」およびリスト上でハイライトされている第2の項目のような多様なかな入力を含む。ユーザインターフェース上の結果領域910に示されるように、検索システムは、5つの一致項目を返しており、一致した文字の組み合わせは表示された項目内でハイライトされている。5つの項目は、多様な種類の日本語文字ならびにラテン文字を含む。各項目は、項目の種類を示す、先行アイコン920によって識別される。画面上の項目925および930は、アドレス帳からの名前である。これらの項目の右側にある文字は、左側の漢字に対する読みを示す。項目935および940は音楽ファイルで、項目945はユ

ーザにより使用され得るデバイス機能（例えば、ブックマーク）である。図9に示されるように、2つの文字に対する一致は、各検索結果内の任意の位置で検出されてもよい。

【0039】

日本語の構造は、日本語テキストを検索する上で追加の問題を提示する。例えば、複数の文字体系を使用することに加えて、日本語テキストはしばしば、1つの単語の終了および別の単語の先頭のスペースまたは他の標識を欠く。本明細書に開示される検索システムは、単語、文、またはグループの自然な開始点を検出するために検索されるテキストをセグメント化することによって、一致および検索結果の表示を改善する。次いで、システムは、自然な開始点で発生する一致を、より離れて発生する一致よりも高く順位付ける。

【0040】

英語テキストの場合、自然な開始点は、概して、文の先頭、空白の後、または句読点の後に配置される。日本語テキストの場合、検索システムは以下の技法のうちの1つ以上を使用して、自然な開始点を識別する。

- ・日本語表記において、単語リストおよび文法規則を使用する特殊なアルゴリズム（「セグメント化エンジン」と呼ばれる）が、自然な開始点を推論するために使用され得る。

- ・句読点、または2つの文字体系の間の変化（例えば、かな文字と漢字との間、または漢字とアラビア数字との間）等、自然な開始点を識別するために単純なパターンが使用され得る。例えば、〈缶ビール、たこ焼き〉という語句には、「缶ビール」および「たこ焼き」の単語を明示的に区分する、カンマが存在する。自然な開始点を識別するために単純なパターンを使用すると、日本語の文に存在する可能性がある自然な開始点全てのサブセットを識別することができるだけでなく、演算リソースが限定されたモバイルデバイス上に実装するコストが軽減される。

- ・電話番号は、自然な開始点の別の例である。電話番号は各国において既定の形式を有し、検索システムが使用して開始点を決定することができる。例えば、(206)234-5678のような米国の電話番号の場合、電話番号の中の数字ではない文字は、自然な開始点を決定するために使用され得る。このように、「234」および「456」の検索はどちらも電話番号には一致するが、「234」の検索の一致は、番号の自然な開始点で発生するため、より有意であると考えられる。

【0041】

一式の一致が検出された後、検索システムは一式の一致を返し、多様な因子を使用して、検索結果の順序を決定する。例えば、システムは、自然な開始点からの距離の順序で一致項目を表示するように構成されてもよい。この順序付け方法論は、システムにより使用されて、図9に示される検索結果を生成する。図9の項目935において、入力検索語は、単語の先頭、つまり、自然な開始点からゼロの距離にある文字に一致した。第2の一致項目（項目925）は、単語の先頭で自然な開始点から1文字の距離を有する。同様に、第3、第4および第5の項目（それぞれ、項目940、945および930）は、それぞれ、自然な開始点から2、3、および4文字の距離を有する。一致文字列内の自然な開始点から一致文字までの距離に応じて検索結果を順序付けることによって、本明細書に開示される検索システムは、ユーザに対して、検索結果リストの上位に潜在的に関連性が高い検索結果を提示することが可能である。

【0042】

自然な開始点からの距離に加えて、システムは、検索結果を順序付ける場合に以下（特定の順序ではない）を含む、他の因子を考慮することができる。

- ・一致がフィールドの開始に合致するかどうか（例えば、システムは、フィールド内の自然な開始点にある一致よりも、フィールドの開始にある一致のほうがより関連性が高いと考える場合がある）。

- ・一致が単語の開始に合致するかどうか。

- ・一致した項目の種類（例えば、項目は電話番号または曲名かどうか）。

- ・一致のいずれかが、一次フィールドまたは二次フィールドにあるかどうか（例えば、システムは、会社名または市町村の一致よりも、連絡先の名前の一致のほうがより関連性が

10

20

30

40

50

高いと考える場合がある)。

- ・ 検索語が、自然な開始点と次の隣接の自然な開始点との間のテキストの全てに一致したかどうか、または開始点の間のテキストの部分だけに一致したかどうか。
- ・ 一致した項目が以前に使用されたことがあるかどうか(つまり、一致した項目は、以前、検索結果からユーザにより選択されたかどうか)。
- ・ 一致した項目が使用された回数(つまり、一致した項目が検索結果からユーザによってこれまでに選択された回数)。

検索結果の順序を決定するために、検索結果の各項目について、検索システムは、各因子が検索結果により満たされるかどうかに基づいて、上記の因子のうちの1つ以上に数値を割り当てることができる。各因子は、因子の認識される重要度に応じて加重されてもよく、各項目に対する総合的な関連点数は、全ての関連する因子の加重された数値を合計することにより計算される。検索結果の項目は、次いで、総合的な関連点数別にリスト(つまり、順位付け)される。

#### 【0043】

システムはまた、複数の検索語を使用して同時に検索することが可能であってもよい。複数語の検索において、システムは、加重因子を組み合わせて、組み合わせ点数に基づいて並び替えるように構成されてもよい。組み合わせ点数は、検索語の点数の合計、加重された可能性の積(または対数の合計として)等のいくつかの方法を使用して、または特殊な条件ロジックを備える比較器を使用して、計算され得る。特殊比較器の実施例として、2つの結果を返す2つの語に対する検索を検討する。第1の結果では、両方の語は、自然な開始点から1文字離れている。第2の結果では、一方の語は自然な開始点に合致し、他方は自然な開始点から3文字離れている。システムが、自然な開始点からの距離だけに基づいて結果を順序付けるように構成されている場合、第1の結果は第2よりも距離の和が小さいため、第1の結果を第2の結果の前に順序付ける。そうではなく、システムが合致を優先するように構成されている場合、語のうちの1つは開始点に合致していたため、第2の結果を第1の結果の前に順序付ける。

#### 【0044】

図10は、モバイルデバイス上で日本語テキストを検索するための検索システム1000の論理的ブロック図である。システム1000は、入力コンポーネント702を介してユーザ入力を受信し、表示コンポーネント704を介して検索結果の順序付けリストを出力し、記憶コンポーネント706からデータを呼び出す。これらのコンポーネントの各々は、動作において、図7の上記のコンポーネントに対応する。記憶コンポーネント706は、ユーザ入力を日本語に変換するために使用される辞書を含むことに加えて、モバイルデバイス上に格納される項目のデータベースまたはインデックスも含むことができる。上記のように、これらの項目は、例えば、オーディオファイル、ビデオファイル、アドレス帳入力、ブックマーク、または他のアプリケーション、機能、またはデータファイルであってもよく、検索システムにより検索され得る1つ以上のテキストフィールドを有する。

#### 【0045】

検索システム1000は、ユーザ入力(入力コンポーネント702から受信される)を一式の見出し語検索語に変換するように構成される、変換コンポーネント1010を含む。変換コンポーネント1010は、一式の検索語を生成するように、上記で検討したテキスト入力システムのプロセスに類似のプロセスを使用することができる。概して、検索語のリストは、ユーザ入力に対応する全ての見出し語を含む。

#### 【0046】

検索システム1000はまた、変換コンポーネント1010により生成される検索語に基づいて、モバイルデバイスまたはモバイルデバイスによりアクセス可能な遠隔場所を検索するように構成される、検索コンポーネント1012も含む。検索は、記憶コンポーネント706により格納される項目のこれまでに生成されたデータベースまたはインデックスを検索することを含むことができる。概して、検索コンポーネント1012は、モバイルデバイス上の項目のテキストフィールド内の任意の場所で一致テキストを検索する(つ

10

20

30

40

50

まり、検索語の発生個所)。検索コンポーネント1012は、次いで、検索語に対応する一致項目のリストを生成する。

【0047】

検索システム1000はまた、項目のテキストフィールド内で1つ以上の自然な開始点を決定するために、検索結果の各々を処理するように構成される、開始点決定コンポーネント1014も含む。上記のように、システムは、テキスト内の句読点または文字セットの変化を検出する等、開始点を決定するための多様な方法を使用することができる。次いで、開始点情報は、自然の開始点から各一致テキストまでの距離を決定するように構成される、距離計算コンポーネント1016により使用される。一部の実施形態において、距離は、一致テキストの開始と、一致テキストの開始前に発生する最も近い開始点との間の文字数に等しい。他の実施形態において、距離は、一致テキストの開始からいずれかの方向における最も近い開始点までの文字数である。計算された距離は、計算された距離に基づいて検索結果を順序付け、表示コンポーネント704を介して順序付けられた検索結果をユーザに提供するように構成される、順序付けコンポーネント1018により使用される。順序付けコンポーネント1018はまた、検索結果の順序を決定するために、上記の追加の因子を使用することもできる。

10

【0048】

図11は、検索システム1000により実行されるプロセス1100のフローチャートである。プロセスはブロック1102で開始し、システムはユーザ入力を受信する。ユーザ入力は、ハードウェアキーパッドまたはキーボードを経由して、またはソフトウェア表示キーパッドまたはキーボードを経由して提供されてもよい。ブロック1104で、検索システムは、ユーザ入力を1つ以上の検索語に変換する。ユーザ入力からテキスト検索語への変換は、上記に開示される予測テキスト入力方法に類似のプロセスを使用して実行されてもよい。つまり、検索システムは、受信した入力を1つ以上の読みに変換し、読みを使用して、一式の対応する見出し語を決定することができる。全ての可能な読みに対応する一式の見出し語は、次いで、検索システムにより一式の検索語として使用される。一式の検索後を決定した後、プロセスはブロック1106に進み、検索システムは、決定された一式の検索語に対応する一式の検索結果を生成する。一部の実装において、システムは、モバイルデバイスおよび検索時に関連の遠隔場所を直接検索して一致項目を検出する。他の実施形態において、システムは、項目のデータベースまたは他のそれまでに生成されたインデックスを使用して検索を実行する。インデックスは、項目に関連する1つ以上のテキストフィールドの内容等、各項目の情報を含む。例えば、システムは、モバイルデバイス上またはモバイルデバイスによりアクセス可能な遠隔場所に格納されたメディアファイルのタイトルまたは詳細情報を格納するインデックスに依存してもよい。

20

30

【0049】

プロセスは次いでブロック1108に進み、検索システムは、上記の方法を使用して、一致項目の各々のテキストフィールド内の1つ以上の自然な開始点を決定する。ブロック1110で、検索システムは、上記のように、各一致項目の一致テキストと開始点との間の距離を決定する。ブロック1112で、検索システムは、上記のように、計算された距離および他の因子を使用して、一式の順序付けられた検索結果を生成する。ブロック1114で、システムは、ユーザに表示するための順序付けられた結果を提供する。一致したテキスト内の自然な開始点に依存する順序で検索結果をユーザに提示することによって、ユーザは、モバイルデバイス上またはモバイルデバイスを介してアクセス可能な項目の位置を迅速かつ容易に特定することが可能である。

40

【0050】

テキスト入力および検索システムは日本語のコンテキストにおいて上記で説明されるが、システムはそれには限定されない。当業者は、類似のシステムは、中国語または韓国語等、複雑な表記文字を使用する他の言語におけるテキスト入力および検索に使用され得ることを理解するであろう。特に、本システムは、漢字およびハングル(韓国文字)を組み合わせたテキストをしばしば含む、韓国語で有用であろう。

50

【 0 0 5 1 】

前述のように、本発明の具体的な実施形態は、説明の目的で本明細書に説明されたが、本発明から逸脱することなく、多様な変形が行われてもよいことが理解されるであろう。したがって、本発明は、添付の請求項による以外には限定されない。

【 図 1 】

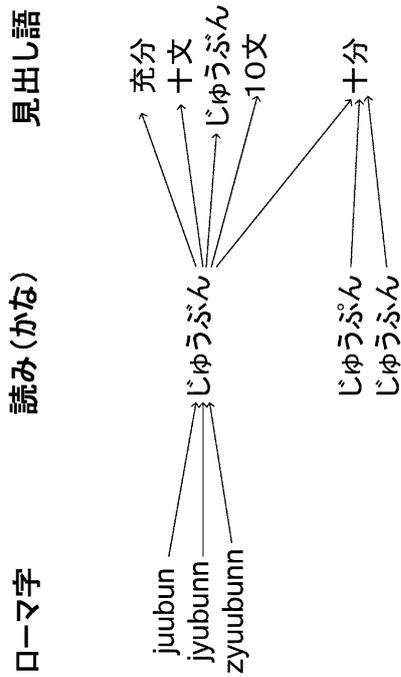


FIG. 1 (先行技術)

【 図 2 】

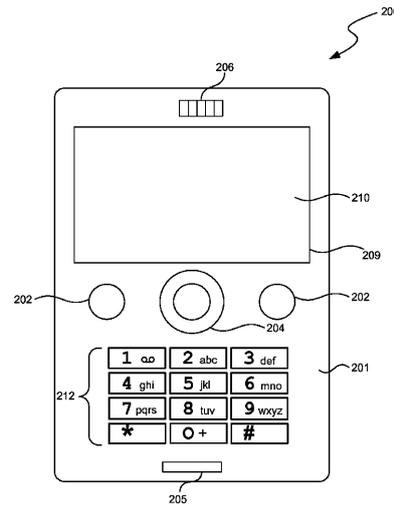


FIG. 2



【図7】

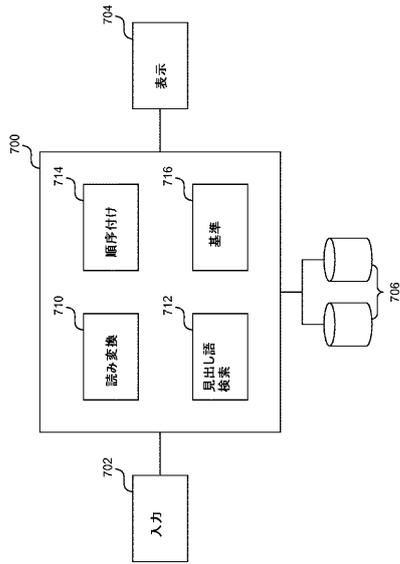


FIG. 7

【図8】

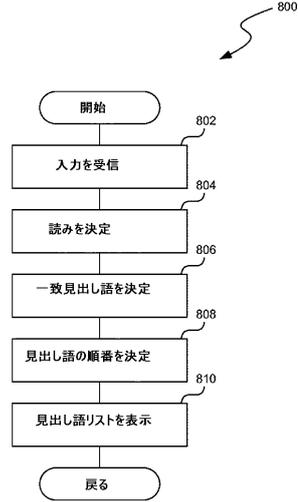


FIG. 8

【図9】

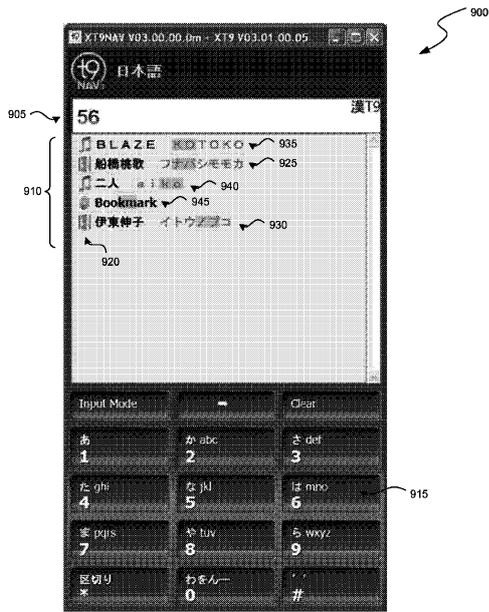


FIG. 9

【図10】

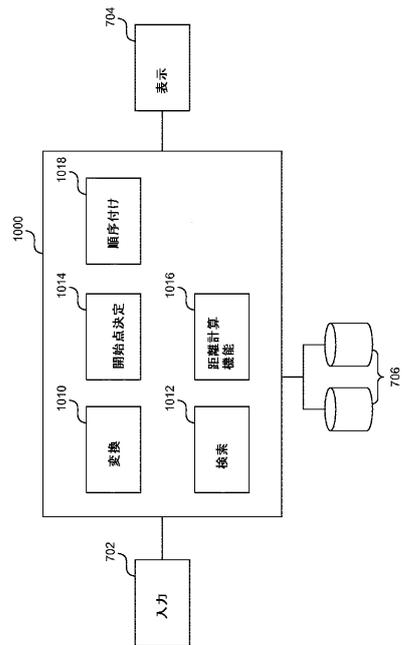


FIG. 10

【図 11】

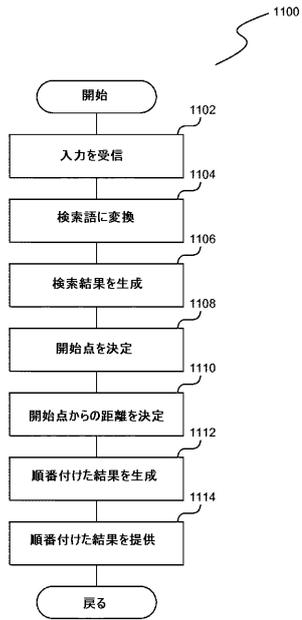


FIG. 11

## フロントページの続き

- (72)発明者 アンルー, エルランド  
アメリカ合衆国 マサチューセッツ 01803, バーリントン, ウェイサイド ロード 1
- (72)発明者 マーシャル, ケビン  
アメリカ合衆国 マサチューセッツ 01803, バーリントン, ウェイサイド ロード 1
- (72)発明者 ウァデル, ゴードン  
アメリカ合衆国 マサチューセッツ 01803, バーリントン, ウェイサイド ロード 1
- (72)発明者 ステインメッツ, ウィリアム ヘンリー  
アメリカ合衆国 マサチューセッツ 01803, バーリントン, ウェイサイド ロード 1
- (72)発明者 石附 支  
アメリカ合衆国 マサチューセッツ 01803, バーリントン, ウェイサイド ロード 1

審査官 齊藤 貴孝

- (56)参考文献 米国特許出願公開第2007/0118533 (US, A1)  
特開2002-175330 (JP, A)  
国際公開第2007/076269 (WO, A1)  
特開平09-054781 (JP, A)  
特開2007-102685 (JP, A)  
特表2010-508592 (JP, A)  
特表2009-520283 (JP, A)  
特開平05-233704 (JP, A)  
特開2001-325252 (JP, A)

- (58)調査した分野(Int.Cl., DB名)  
G06F 17/30