



(12) 发明专利申请

(10) 申请公布号 CN 104808806 A

(43) 申请公布日 2015. 07. 29

(21) 申请号 201410041504. 8

(22) 申请日 2014. 01. 28

(71) 申请人 北京三星通信技术研究有限公司
地址 100028 北京市朝阳区太阳宫中路 12
号楼 15 层 1503
申请人 三星电子株式会社

(72) 发明人 陈又新 吴介

(74) 专利代理机构 北京市立方律师事务所
11330
代理人 郑瑜生

(51) Int. Cl.
G06F 3/023(2006. 01)

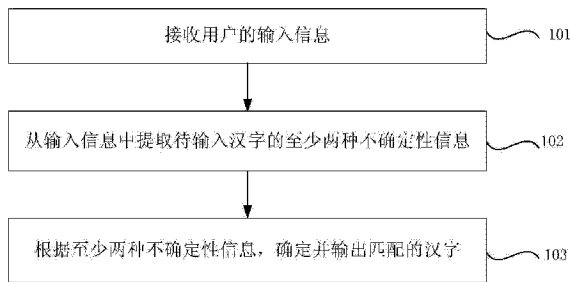
权利要求书3页 说明书15页 附图6页

(54) 发明名称

根据不确定性信息实现汉字输入的方法和装置

(57) 摘要

提供了一种根据不确定性信息实现汉字输入的方法及装置,其中,方法包括:接收用户的输入信息;从输入信息中提取待输入汉字的至少两种不确定性信息;根据至少两种不确定性信息,确定并输出匹配的汉字;其装置包括:接收模块、提取模块以及匹配模块。本发明提供的方法及装置,使得当用户对待输入汉字的读音信息或者字型信息记忆不完全时,将待输入汉字对应的字符的候选限制在一定范围内,并且结合所提取出的待输入汉字的至少两种不确定性信息,实现待输入汉字的准确输入。



1. 一种根据不确定性信息实现汉字输入的方法,其特征在于,包括以下步骤:
接收用户的输入信息;
从所述输入信息中提取待输入汉字的至少两种不确定性信息;
根据所述至少两种不确定性信息,确定并输出匹配的汉字。
2. 根据权利要求1所述的根据不确定性信息实现汉字输入的方法,其特征在于,所述不确定性信息包括以下任意一种信息:
拼音信息、声调信息、结构布局信息、偏旁部首信息、同音字信息以及位置信息。
3. 根据权利要求2所述的根据不确定性信息实现汉字输入的方法,其特征在于,通过以下任意一种方式接收所述声调信息:
接收用户手写输入所述声调信息的符号:ˊ, ˋ, ˊ 或 ˋ;
接收特定的数字或字母来表示所述声调信息。
4. 根据权利要求2所述的根据不确定性信息实现汉字输入的方法,其特征在于,通过以下任意一种方式接收所述位置信息:
接收用户输入的空格或特定的字符。
5. 根据权利要求2至4任意之一所述的根据不确定性信息实现汉字输入的方法,其特征在于,接收用户的输入信息,包括:
在输入区域检测所述用户的操作;
记录所述用户的操作输入的至少两种不确定性信息。
6. 根据权利要求5所述的根据不确定性信息实现汉字输入的方法,其特征在于,还包括:
在屏幕上显示所述至少两种不确定性信息。
7. 根据权利要求6所述的根据不确定性信息实现汉字输入的方法,其特征在于,还包括:
检测到所述用户在输入区域修改所述不确定性信息时,在屏幕上显示更新后的所述不确定性信息。
8. 根据权利要求2所述的根据不确定性信息实现汉字输入的方法,其特征在于,根据所述至少两种不确定性信息,确定并输出匹配的汉字,包括:
根据所述至少两种不确定性信息选择相应的模型;
在至少两种所述相应的模型中匹配所述待输入汉字;
对每种模型的匹配结果进行分析,确定匹配度最高的汉字为所述待输入汉字,输出所述匹配度最高的汉字。
9. 根据权利要求8所述的根据不确定性信息实现汉字输入的方法,其特征在于,还包括:
确定所述不确定性信息的优先级;
根据优先级,依次根据所述不确定性信息在相应模型中匹配所述待输入汉字。
10. 根据权利要求1所述的根据不确定性信息实现汉字输入的方法,其特征在于,还包括:
接收用户的输入,设置所述不确定性信息的优先级。
11. 一种根据不确定性信息实现汉字输入的方法,其特征在于,包括以下步骤:

接收用户的输入信息；

从所述输入信息中提取待输入汉字的不确定性信息，其中，所述不确定性信息包括以下任意一种或多种信息：

声调信息、结构布局信息、偏旁部首信息、同音字信息和位置信息；

基于已经确定的汉字和所述不确定性信息，确定并输出所述待输入汉字。

12. 一种根据不确定性信息实现汉字输入的装置，其特征在于，包括：接收模块、提取模块以及匹配模块；

所述接收模块，用于接收用户的输入信息；

所述提取模块，用于从所述输入信息中提取待输入汉字的至少两种不确定性信息；

所述匹配模块，用于根据所述至少两种不确定性信息，确定并输出匹配的汉字。

13. 根据权利要求 12 所述的根据不确定性信息实现汉字输入的装置，其特征在于，所述提取模块提取的所述不确定性信息包括以下任意一种信息：

拼音信息、声调信息、结构布局信息、偏旁部首信息、同音字信息以及位置信息。

14. 根据权利要求 13 所述的根据不确定性信息实现汉字输入的装置，其特征在于，所述接收模块用于通过以下任意一种方式接收所述声调信息：

接收用户手写输入所述声调信息的符号：ˊ，ˊ，ˋ 或 ˋ；

接收特定的数字或字母来表示所述声调信息。

15. 根据权利要求 13 所述的根据不确定性信息实现汉字输入的装置，其特征在于，所述接收模块用于通过以下任意一种方式接收所述位置信息：

接收用户输入的空格或特定的字符。

16. 根据权利要求 13 至 15 任意之一所述的根据不确定性信息实现汉字输入的装置，其特征在于，所述接收模块接收用户的输入信息，包括：

在输入区域检测所述用户的操作；

记录所述用户的操作输入的至少两种不确定性信息。

17. 根据权利要求 16 所述的根据不确定性信息实现汉字输入的装置，其特征在于，所述接收模块还用于在屏幕上显示所述至少两种不确定性信息。

18. 根据权利要求 17 所述的根据不确定性信息实现汉字输入的装置，其特征在于，所述装置还包括更新模块，

所述更新模块，用于检测到所述用户在输入区域修改所述不确定性信息时，在屏幕上显示更新后的所述不确定性信息。

19. 根据权利要求 13 所述的根据不确定性信息实现汉字输入的装置，其特征在于，根据所述至少两种不确定性信息，所述匹配模块确定并输出匹配的汉字，包括：

根据所述至少两种不确定性信息选择相应的模型；

在至少两种所述相应的模型中匹配所述待输入汉字；

对每种模型的匹配结果进行分析，确定匹配度最高的汉字为所述待输入汉字，输出所述匹配度最高的汉字。

20. 根据权利要求 19 所述的根据不确定性信息实现汉字输入的装置，其特征在于，所述匹配模块还用于：

确定所述不确定性信息的优先级；

根据优先级,依次根据所述不确定性信息在相应模型中匹配所述待输入汉字。

21. 根据权利要求 12 所述的根据不确定性信息实现汉字输入的装置,其特征在于,所述装置还包括设置模块,

所述设置模块,用于接收用户的输入,设置所述不确定性信息的优先级。

22. 一种根据不确定性信息实现汉字输入的装置,其特征在于,包括:接收模块、提取模块以及匹配模块;

所述接收模块,用于接收用户的输入信息;

所述提取模块,用于从所述输入信息中提取待输入汉字的不确定性信息,其中,所述不确定性信息包括以下任意一种或多种信息:

声调信息、结构布局信息、偏旁部首信息、同音字信息和位置信息;

所述匹配模块,用于基于已经确定的汉字和所述不确定性信息,确定并输出所述待输入汉字。

根据不确定性信息实现汉字输入的方法和装置

技术领域

[0001] 本发明涉及终端设备的技术领域,具体而言,本发明涉及根据不确定性信息实现汉字输入的方法和装置。

背景技术

[0002] 中国汉字笔画繁复,移动设备中拼音输入法的广泛使用导致人们对汉字的字形越发的不熟悉,提笔忘字的现象也越来越频繁。即便手写输入法的功能中包括了在使用者输入字形相近的字符时也能在候选中列出正确字符,但有些常用的汉字,例如犄角旮旯,饕餮等,经常让人一时之间即便大概的字形也无从想起。

[0003] 现有技术中,除了精确地写出一个汉字,还有利用辅助信息,或者通过与系统交互的先验知识,来作为汉字输入的辅助手段。

[0004] 对于那些不记得如何书写的汉字,但是人们却能够记得相应的汉字的读音和声调;或者记得相应的结构布局的类型;或者仅仅记得改字的某一个偏旁部首;或者由于香港台湾居民的拼音系统同大陆不同以及有些老年人不会拼音,只是记得相应字符的读音,只能输入相应字符的同音字。以上这些信息统称为不确定信息。

[0005] 现有的输入法往往只利用单一信息,例如拼音,或者字形来实现输入,这使得当汉字或拼音的信息记忆不完全的时候,用户无法完成相应汉字的输入。

发明内容

[0006] 本发明提供根据不确定性信息实现汉字输入的方法及装置,用于通过用户待输入汉字所提供的不确定性信息,基于上述至少两种不确定性信息来解决用户对待输入的汉字的拼音或者字型的信息记忆不完全时,无法实现汉字的准确输入的问题。

[0007] 为解决现有技术中的上述问题,本发明实施例一方面提供了一种根据不确定性信息实现汉字输入的方法,包括以下步骤:

[0008] 接收用户的输入信息;

[0009] 从所述输入信息中提取待输入汉字的至少两种不确定性信息;

[0010] 根据所述至少两种不确定性信息,确定并输出匹配的汉字。

[0011] 本发明实施例另一方面提供了一种根据不确定性信息实现汉字输入的方法,包括以下步骤:

[0012] 接收用户的输入信息;

[0013] 从所述输入信息中提取待输入汉字的不确定性信息,其中,所述不确定性信息包括以下任意一种或多种信息:

[0014] 声调信息、结构布局信息、偏旁部首信息、同音字信息和位置信息;

[0015] 基于已经确定的汉字和所述不确定性信息,确定并输出所述待输入汉字。

[0016] 本发明实施例另一方面提供了一种根据不确定性信息实现汉字输入的装置,包括:接收模块、提取模块以及匹配模块;

- [0017] 所述接收模块,用于接收用户的输入信息;
- [0018] 所述提取模块,用于从所述输入信息中提取待输入汉字的至少两种不确定性信息;
- [0019] 所述匹配模块,用于根据所述至少两种不确定性信息,确定并输出匹配的汉字。
- [0020] 本发明实施例另一方面提供了一种根据不确定性信息实现汉字输入的装置,包括:接收模块、提取模块以及匹配模块;
- [0021] 所述接收模块,用于接收用户的输入信息;
- [0022] 所述提取模块,用于从所述输入信息中提取待输入汉字的不确定性信息,其中,所述不确定性信息包括以下任意一种或多种信息:
- [0023] 声调信息、结构布局信息、偏旁部首信息、同音字信息和位置信息;
- [0024] 所述匹配模块,用于基于已经确定的汉字和所述不确定性信息,确定并输出所述待输入汉字。
- [0025] 本发明提出的上述方案,使得当用户对待输入汉字的读音信息或者字型信息记忆不完全时,将待输入汉字对应的字符的候选限制在一定范围内,并且结合所提取出的待输入汉字的至少两种不确定性信息,实现待输入汉字的准确输入。此外,本发明提出的上述方案,对现有系统的改动很小,不会影响系统的兼容性,而且实现简单、高效。
- [0026] 本发明附加的方面和优点将在下面的描述中部分给出,这些将从下面的描述中变得明显,或通过本发明的实践了解到。

附图说明

- [0027] 本发明上述的和/或附加的方面和优点从下面结合附图对实施例的描述中将变得明显和容易理解,其中:
- [0028] 图1为本发明根据不确定性信息实现汉字输入的方法一种实施例的流程图;
- [0029] 图2为本发明根据不确定性信息实现汉字输入的方法另一种实施例的流程图;
- [0030] 图3为本发明根据不确定性信息实现汉字输入的方法另一种实施例中的结构布局信息的分类类型示意图;
- [0031] 图4为本发明根据不确定性信息实现汉字输入的方法另一种实施例中方案一的流程图;
- [0032] 图5为本发明根据不确定性信息实现汉字输入的方法另一种实施例中方案四的流程图;
- [0033] 图6为本发明根据不确定性信息实现汉字输入的装置一种实施例的结构示意图;
- [0034] 图7为本发明根据不确定性信息实现汉字输入的装置另一种实施例的结构示意图;
- [0035] 图8为根据不确定性信息确定匹配度最高的汉字的示意图。

具体实施方式

[0036] 下面详细描述本发明的实施例,所述实施例的示例在附图中示出,其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的,仅用于解释本发明,而不能解释为对本发明的限制。

[0037] 本技术领域技术人员可以理解,除非特意声明,这里使用的单数形式“一”、“一个”、“所述”和“该”也可包括复数形式。应该进一步理解的是,本发明的说明书中使用的措辞“包括”是指存在所述特征、整数、步骤、操作、元件和 / 或组件,但是并不排除存在或添加一个或多个其他特征、整数、步骤、操作、元件、组件和 / 或它们的组。应该理解,当我们称元件被“连接”或“耦接”到另一元件时,它可以直接连接或耦接到其他元件,或者也可以存在中间元件。此外,这里使用的“连接”或“耦接”可以包括无线连接或无线耦接。这里使用的措辞“和 / 或”包括一个或更多个相关联的列出项的全部或任一单元和全部组合。

[0038] 本技术领域技术人员可以理解,除非另外定义,这里使用的所有术语(包括技术术语和科学术语),具有与本发明所属领域中的普通技术人员的一般理解相同的意义。还应该理解的是,诸如通用字典中定义的那些术语,应该被理解为具有与现有技术的上下文中的意义一致的意义,并且除非像这里一样被特定定义,否则不会用理想化或过于正式的含义来解释。

[0039] 本技术领域技术人员可以理解,这里所使用的“终端”、“终端设备”既包括无线信号接收器的设备,其仅具备无发射能力的无线信号接收器的设备,又包括接收和发射硬件的设备,其具有能够在双向通信链路上,进行双向通信的接收和发射硬件的设备。这种设备可以包括:蜂窝或其他通信设备,其具有单线路显示器或多线路显示器或没有多线路显示器的蜂窝或其他通信设备;PCS(Personal Communications Service,个人通信系统),其可以组合语音、数据处理、传真和 / 或数据通信能力;PDA(Personal Digital Assistant,个人数字助理),其可以包括射频接收器、寻呼机、互联网 / 内联网访问、网络浏览器、记事本、日历和 / 或 GPS(Global Positioning System,全球定位系统)接收器;常规膝上型和 / 或掌上型计算机或其他设备,其具有和 / 或包括射频接收器的常规膝上型和 / 或掌上型计算机或其他设备。这里所使用的“终端”、“终端设备”可以是便携式、可运输、安装在交通工具(航空、海运和 / 或陆地)中的,或者适合于和 / 或配置为在本地运行,和 / 或以分布形式,运行在地球和 / 或空间的任何其他位置运行。这里所使用的“终端”、“终端设备”还可以是通信终端、上网终端、音乐 / 视频播放终端,例如可以是 PDA、MID(Mobile Internet Device,移动互联网设备)和 / 或具有音乐 / 视频播放功能的移动电话,也可以是智能电视、机顶盒等设备。

[0040] 图 1 为本发明根据不确定性信息实现汉字输入的方法一种实施例的流程图。如图 1 所示,本实施例中根据不确定性信息实现汉字输入的方法流程具体包括如下步骤:

[0041] 步骤 101、接收用户的输入信息。

[0042] 作为本发明的实施例,接收用户的输入信息,包括:

[0043] 在输入区域检测用户的操作;

[0044] 记录用户的操作输入的至少两种不确定性信息。

[0045] 步骤 102、从输入信息中提取待输入汉字的至少两种不确定性信息。

[0046] 具体而言,不确定性信息包括以下任意一种信息:拼音信息、声调信息、结构布局信息、偏旁部首信息、同音字信息以及位置信息。其中,拼音信息包括拼音字符串信息和拼音声调信息,拼音字符串信息用于根据书写的笔迹特点和书写区域,分离拼音串;拼音声调信息用于根据拼音声调书写的笔迹特点和书写区域,分离拼音声调;结构布局信息用于根据结构布局信息的笔迹特点和书写区域,分离结构布局信息;偏旁部首信息用于根据结构

布局信息的笔迹特点和书写区域,分离用户可以确定的偏旁部首;同音字信息用于根据同音字书写的笔迹特点和书写区域,分离同音字信息。

[0047] 声调信息包括待输入汉字的阴平、阳平、上声或下声的信息。作为本发明的实施例,终端设备通过以下任意一种方式接收声调信息:

[0048] 接收用户手写输入声调信息的符号:ˉ,ˊ,ˇ或ˋ;

[0049] 接收特定的数字或字母来表示声调信息。

[0050] 例如,预先设置数字1,2,3,4分别表示阴平、阳平、上声或下声;或者,预先设置字母A,B,C,D分别表示阴平、阳平、上声或下声。

[0051] 作为本发明的实施例,终端设备通过以下任意一种方式接收位置信息:接收用户输入的空格或特定的字符。例如,预先设置“?”来指示位置信息。

[0052] 在实际应用中,所包含的待输入汉字的位置信息对待输入汉字有一定的定位作用,例如,当人们不知道“味同嚼蜡”中的第三个字的“嚼”字怎样写时,可以连续输入“味同?蜡”,该词语的第三个字可以用问号或者空格来代替,再结合不确定性信息中的拼音、拼音声调、结构布局、偏旁部首以及同音字等信息,以实现当用户对待输入汉字的读音信息或者字型信息记忆不完全时,仍然能够准确输入待输入汉字。

[0053] 步骤103、根据至少两种不确定性信息,确定并输出匹配的汉字。

[0054] 作为本发明的实施例,根据至少两种不确定性信息,确定并输出匹配的汉字,包括:

[0055] 确定不确定性信息的优先级;

[0056] 根据优先级,依次根据不确定性信息在相应模型中匹配待输入汉字;

[0057] 确定匹配度最高的汉字为待输入汉字,输出匹配度最高的汉字。

[0058] 在本应用场景中,通过接收用户的输入信息;从输入信息中提取待输入汉字的至少两种不确定性信息;根据至少两种不确定性信息,确定并输出匹配的汉字;使得当用户对待输入汉字的读音信息或者字型信息记忆不完全时,将待输入汉字对应的字符的候选限制在一定范围内,并且结合所提取出的待输入汉字的至少两种不确定性信息,例如,拼音信息、声调信息、结构布局信息、偏旁部首信息、同音字信息以及位置信息等信息,实现待输入汉字的准确输入。

[0059] 图2为本发明根据不确定性信息实现汉字输入的方法另一种实施例的流程图。如图2所示,本实施例中根据不确定性信息实现汉字输入的流程具体包括如下步骤:

[0060] 步骤201、接收用户的输入信息。

[0061] 作为本发明的实施例,接收用户的输入信息,包括:

[0062] 在输入区域检测用户的操作;

[0063] 记录用户的操作输入的至少两种不确定性信息。

[0064] 步骤202、接收用户的输入,设置不确定性信息的优先级。

[0065] 步骤203、从输入信息中提取待输入汉字的至少两种不确定性信息。

[0066] 具体而言,不确定性信息包括以下任意一种信息:拼音信息、声调信息、结构布局信息、偏旁部首信息、同音字信息以及位置信息。其中,拼音信息包括拼音字符串信息和拼音声调信息,拼音字符串信息用于根据书写的笔迹特点和书写区域,分离拼音串;拼音声调信息用于根据拼音声调书写的笔迹特点和书写区域,分离拼音声调;结构布局信息用于根

据结构布局信息的笔迹特点和书写区域,分离结构布局信息;偏旁部首信息用于根据结构布局信息的笔迹特点和书写区域,分离用户可以确定的偏旁部首;同音字信息用于根据同音字书写的笔迹特点和书写区域,分离同音字信息。

[0067] 图3为本发明根据不确定性信息实现汉字输入的方法另一种实施例中的结构布局信息的分类类型示意图。如图3所示,本发明所涉及的结构布局信息的分类如下所述:根据汉字的结构特点,可将汉字的结构大致划分为以下类型:独体结构、上下结构、上中下结构、左右结构、左中右结构、上单结构、下单结构、左单结构、右单结构、半包围结构、包围结构、品字形结构。

[0068] 本应用场景中所涉及的“不确定性信息组合对应的方案种类”具体如下所述:

[0069] 根据用户能够提供的不确定性信息的种类,可罗列出如下七种输入方案。输入的信息越多,系统获得的先验知识就越多,就越有可能最为快速准确地得到用户想要输入的汉字。

[0070] 为了更清晰地描述,可以将方案类型、拼音、拼音声调、结构布局、偏旁部首以及同音字之间建立映射关系,具体如表1所示。

[0071] 表1 方案类型、拼音、拼音声调、结构布局、偏旁部首以及同音字之间的映射表

[0072]

方案类型	拼音	拼音声调	结构布局	偏旁部首	同音字
方案一	√				
方案二	√	√			
方案三	√		√		
方案四	√	√	√		
方案五		√			√
方案六		√	√		√
方案七			√		√
方案八			√	√	

[0073]

[0074] 其实组合方案的种类可以更多的,上面列出的八个方案例子只是常见的用户可能会使用的组合。单一不确定性信息也可以单独构成一个方案。以单一使用结构布局信息为例,由于结构布局信息使用方框表示,容易同汉字的某些偏旁部首产生混淆,导致单独的结构布局信息被误识别成汉字,如果该汉字恰好满足上下文语义关系,则会降低系统最终给出正确汉字的准确率。同音字同理。

[0075] 偏旁部首最好能够与结构布局信息成对出现,这样会更好限制偏旁部首在字符中所处的位置,降低系统处理复杂程度。

[0076] 步骤 204、在屏幕上显示用户输入的不确定性信息。

[0077] 在实际应用中,区别于现有技术,本发明的实施例可以在屏幕上显示用户输入的不确定性信息,例如显示至少两种不确定性信息,这样明显区别于现有技术的只能提取一种不确定性信息,待输入汉字的汉字字型或拼音的信息;而本发明实施例中所提供的汉字输入的方法,可以至少提供两种以上的不确定性信息,并可以显示至少两种以上的不确定性信息,这样就可以充分的利用词语的位置信息。例如,当待输入汉字为“味同嚼蜡”时,忘记具体如何去写“嚼”字,但是,知道该字在该词语中的上下文中的位置信息,是第三个字,词语是前两个字以及顺序是“味同”,以及第四个字是“蜡”,同时,结合这个字的其它信息,例如,该字的拼音为“jiao”、该字的拼音声调为二声,该字的结构布局为左右结构,该字的偏旁部首中的右半部分的字形为“爵”,该字的同音字为“交”中的汉字之一,通过在屏幕上显示用户输入的至少两种以上的不确定性信息,就可以使得当用户对待输入汉字的读音信息或者字型信息记忆不完全时,准确便捷地输入待输入汉字。

[0078] 进一步地,检测到用户在输入区域修改不确定性信息时,在屏幕上显示更新后的不确定性信息。

[0079] 在实际应用中,区别于现有技术,本发明的实施例可以检测到用户在输入区域修改不确定性信息时,在屏幕上显示更新后的不确定性信息。这样明显区别于现有技术的只能提取一种不确定性信息,待输入汉字的汉字字型或拼音的信息;而本发明实施例中所提供的汉字输入的方法,可以至少提供两种以上的不确定性信息,并可以检测到用户在输入区域修改不确定性信息时,在屏幕上显示更新后的不确定性信息,这样就可以充分的利用词语的位置信息。例如,当待输入汉字为“味同嚼蜡”时,忘记具体如何去写“嚼”字,但是,知道该字在该词语中的上下文中的位置信息,是第三个字,词语是前两个字以及顺序是“味同”,以及第四个字是“蜡”,同时,结合这个字的其它信息,例如,该字的拼音为“jiao”、该字的拼音声调为二声,该字的结构布局为左右结构,该字的偏旁部首中的右半部分的字形为“爵”,该字的同音字为“交”中的汉字之一,通过在屏幕上显示用户输入的至少两种以上的不确定性信息,就可以使得当用户对待输入汉字的读音信息或者字型信息记忆不完全时,准确便捷地输入待输入汉字。

[0080] 具体而言,例如,当待输入汉字为“味同嚼蜡”时,忘记具体如何去写“嚼”字,但是,知道该字在该词语中的位置信息,是第三个字,词语是前两个字以及顺序是“味同”,以及第四个字是“蜡”,定位使用待输入汉字的位置信息,

[0081] 若不确定性信息为两种时,结合第三个字的拼音为“jiao”;

[0082] 若不确定性信息为三种时,结合第三个字的拼音为“jiao”,且第三个字的声调为二声;

[0083] 若不确定信息为四种时,结合第三个字的拼音为“jiao”,且第三个字的声调为二声;且该字的结构布局为左右结构。

[0084] 从上述可以看出:所包含的不确定性信息种类越多,越能够尽可能地缩小不确定性信息所确定的字符集合,从而减少该字符同上下文组合成词组的数目,进一步地减少语言模型分析时所需计算时间,提高计算精度,从而快速便捷的输入不知道具体字型或拼音的待输入汉字。

[0085] 通过检测到用户在输入区域修改不确定性信息时,并在屏幕上显示更新后的不确

定性信息,可以使得本发明实施例的汉字输入方便更具有实用性和更具有可操作性。

[0086] 步骤 205、根据至少两种不确定性信息,确定并输出匹配的汉字。

[0087] 作为本发明的实施例,根据至少两种不确定性信息,确定并输出匹配的汉字,包括:

[0088] 根据至少两种不确定性信息选择相应的模型;

[0089] 在至少两种相应的模型中匹配待输入汉字;

[0090] 对每种模型的匹配结果进行分析,确定匹配度最高的汉字为待输入汉字,输出匹配度最高的汉字。

[0091] 作为本发明的实施例,还包括:

[0092] 确定不确定性信息的优先级;

[0093] 根据优先级,依次根据不确定性信息在相应模型中匹配待输入汉字的步骤。

[0094] 作为本发明的实施例,还包括接收用户的输入,设置不确定性信息的优先级的步骤。

[0095] 在实际应用中,通过设置不确定性信息的优先级,这是明显区别于现有技术所提供的输入法的,其中,优先级会影响效率,匹配效果等等。

[0096] 具体而言,通过设置不确定信息的优先级,就可以更加快速便捷的在屏幕上找到待输入汉字。进一步而言,对于不同的汉字,各种不确定信息在匹配待输入的汉字时,重要性也不同。

[0097] 例如使用的的不确定性信息具体为三种:位置信息、拼音信息和声调信息,可以预先设置上述三种不确定性信息的优先级。具体地,设置位置信息的优先级为第一级,为最优先执行的操作;拼音信息为第二级,为次之执行的操作;声调信息为第三级,为最后执行的操作。

[0098] 当待输入汉字为“味同嚼蜡”中的“嚼”字时,根据本发明提出的技术方案,用户可以通过多种方式确定待输入的“嚼”字,例如,通过读音信息、偏旁部首、声调、结构布局及位置信息,等等。例如,当用户输入不确定信息包括:结构布局为左右结构、且右边为“爵”;声调为第二声。则这个字可能是“味同嚼蜡”中的“嚼”字,也可能是“精神嚼火”中的“嚼”字。当输入的位置信息为“味同蜡”时,根据在位置信息中的位置为第三个字判断时,仍然无法与“精神火”区分;但是根据位置信息中的信息“味同”“蜡”判断时,即立即可以将“嚼”字排除。

[0099] 进一步而言,例如,当某个用户对通过读音查找汉字非常有信心时,可以将不确定信息拼音,声调和/或同音字的优先级设置为更高;当某个用户更擅长于通过书写查找汉字时,可以将不确定信息结构布局和/或偏旁部首的优先级设置为更高。

[0100] 不同方案种类下的输入示例及处理流程:

[0101] 情景:用户想要在手写输入界面“味同嚼蜡”四个字的中文文本,但“嚼”字的具体写法忘记了。

[0102] 在方案 1 下,用户的输入场景如下所示:

[0103] 输入场景:“味同 jiao 蜡”。

[0104] 在方案 2 下,用户的输入场景如下所示:

[0105] 输入场景:“味同 jiao 蜡”,并且显示输入的 jiao 的声调为二声。

- [0106] 在方案 3 下,用户的输入场景如下所示:
- [0107] 输入场景:“味同 jiao 蜡”,并且显示输入的汉字的结构类型为左右结构。
- [0108] 在方案 4 下,用户的输入场景如下所示:
- [0109] 输入场景:“味同 jiao 蜡”,并且显示输入的汉字的结构类型为左右结构,且显示输入的 jiao 的声调为二声。
- [0110] 在方案 5 下,用户的输入场景如下所示:
- [0111] 输入场景:“味同交蜡”,并且显示输入的“交”的声调为二声。
- [0112] 在方案 6 下,用户的输入场景如下所示:
- [0113] 输入场景:“味同交蜡”,并且显示输入的汉字“交”的结构类型为左右结构,且显示输入的汉字“交”的声调为二声。
- [0114] 在方案 7 下,用户的输入场景如下所示:
- [0115] 输入场景:“味同交蜡”,并且显示输入的汉字“交”的结构类型为左右结构。
- [0116] 在方案 8 下,用户的输入场景如下所示:
- [0117] 输入场景:“味同蜡”,并且显示所缺汉字的结构类型为左右结构,且显示所缺汉字的右半部分的字形为“爵”。
- [0118] 在某些情况下,一个词语中,用户忘记写法的字符可能会超过一个。仍以“味同嚼蜡”为例,
- [0119] 在方案 9 下,用户的输入场景如下所示:
- [0120] 输入场景:“味同 jiao la”。
- [0121] 对于上述列举的各种输入场景,当按照一定的格式输入笔迹数据后,处理流程如下:
- [0122] Step1:输入模块接收用户输入的笔迹点,之后传送至预处理模块,为获得更好的识别结果而进行必要的预处理。
- [0123] Step2:切分模块确定每个字符边界。
- [0124] Step3:预处理后的笔迹数据点经由不确定性信息检测模块确定属于不确定性信息的笔迹点,将其同中文文本笔迹点分离。
- [0125] Step4:在 Step2 中得到的不确定性信息笔迹中,通过拼音字符串检测模块/拼音声调检测模块/结构布局信息检测模块/偏旁部首检测模块/同音字检测模块继续分离至少存在其中一项的不确定性信息。
- [0126] Step5:将 Step3 中分离的信息通过拼音识别模块/拼音声调识别模块/结构布局分类识别模块/偏旁部首识别模块/同音字识别模块(同中文字符识别模块)进行识别。
- [0127] Step6:识别 Step2 中分离后得到的中文文本笔迹点得到中文字符候选。
- [0128] Step7:同音字识别结果需要经过“汉字→拼音转换模块”的处理,转化成拼音字符串。
- [0129] Step8:Step4 中,识别得到的拼音字符串及 Step6 中转化得到的拼音字符串均需要经过“拼音+声调→字符转换模块”的处理,转化成中文字符候选。如果缺少声调信息,则变成拼音→字符的转换处理。
- [0130] Step9:Step4 中识别得到的结构布局信息需要经过“结构布局类别→字符转换模块”的处理,转化成中文字符候选。

[0131] Step10:字符筛选融合模块将 Step7 以及 Step8 中得到的中文字符根据偏旁部首的识别结果进行筛选,并重新打分排序,得到一组候选中文字符。

[0132] Step11:将 Step9 中产生的候选同 Step5 中产生的上下文候选组合成词组,通过语言模型计算分数,根据分数高低进行排序,输出词语 / 句子候选。

[0133] Step12:将结果输出。

[0134] 对于每个字符位置的可利用信息中至少包含了拼音或汉字二者任意其一,如上述方案 1-7 所示,除了上述处理方式之外,另一种处理方式则是将全部输入分别全部转换成拼音及汉字,在拼音模型及汉字模型下分别对该字符串进行识别打分处理,最后将分别得到的候选融合成统一候选序列输出。具体的处理流程如下:

[0135] Step21:当输入只包含输入模块接收用户输入的笔迹点,之后传送至预处理模块,为获得更好的识别结果而进行必要的预处理。

[0136] Step22:切分模块确定每个字符边界。

[0137] Step23:将 Step22 得到的结果中的每个字符转换成汉字。如果该字符位置本身书写为汉字,则识别后,根据置信度,给出汉字候选的排序。如果该字符位置本身书写为拼音,则根据拼音串的结果给出汉字候选。转换结果是保证每个字符位置对应了至少一个的汉字候选。

[0138] Step24:利用字符筛选融合模块和语言模型打分模块中的汉字语言模型打分功能对 Step23 的结果进行融合拼接,得到词语候选。

[0139] Step25:将 Step22 得到的结果中的每个字符转换成拼音。如果该字符位置本身书写为拼音,则将识别后的拼音串输出。如果该字符位置本身书写为汉字,不论该汉字为正确写法还是同音字,均识别后根据“汉字 → 拼音转换模块”将汉字候选转换成拼音串候选。转换结果是每个字符位置对应了至少一个拼音串候选。

[0140] Step26:利用语言模型打分模块中的拼音串模型打分功能对 Step25 的结果进行融合拼接,得到词语候选。

[0141] Step27:利用候选筛选融合模块,对 Step24 和 Step26 的结果进行融合。融合规则如下:

[0142] A. 若序列中有相同的候选,则该候选位置提前。

[0143] B. 统计每个候选中字符出现额度,如果某候选中出现额度高的字符个数大于其他候选,则将该候选位置提前。

[0144] C. 根据每个候选的得分,将两个候选序列中的候选融合成一个候选序列。

[0145] Step28:将融合后的结果输出。

[0146] 图 4 为本发明根据不确定性信息实现汉字输入的方法另一种实施例中方案一的流程图。如图 4 所示,本实施例中根据不确定性信息实现汉字输入的方法另一种实施例中方案一的流程具体包括如下步骤:

[0147] 接收用户输入的笔迹;

[0148] 为获得更好的识别结果,而进行必要的预处理;

[0149] 通过切分确定每个字符的边界;

[0150] 从预处理后的笔迹数据点中提取属于不确定性信息的笔迹点,将其同中文文本笔迹点分离;

[0151] 检测不确定性信息；若检测到不确定性信息，则执行相应的操作——在得到的不确定性信息笔迹中，通过拼音字符信息继续分离不确定性信息；将分离的非中文文本信息通过拼音识别模块进行识别；将识别得到的拼音字符串经过拼音→字符的转换，转化成中文字符候选；送至拼音串语言模型进行打分，根据统计的中文词条/拼音串及其出现频率，通过语言模型打分计算拼接后的词条正确与否的概率；将识别结果拼接后输出。

[0152] 若没有检测到不确定性信息，则对中文字符进行识别，即对剥离了不确定信息的中文文本字符进行识别以及对不确定信息中的同音字进行识别；送至汉字语言模型进行打分，根据统计的中文词条/拼音串及其出现频率，通过语言模型打分计算拼接后的词条正确与否的概率；将识别结果拼接后输出。

[0153] 方案一是列出的九种方案中输入形式最为简单的一个，只将拼音串作为唯一的不确定性信息。这种方案也正是现有输入模式下解决忘记汉字写法的唯一解决方案。

[0154] 图5为本发明根据不确定性信息实现汉字输入的方法另一种实施例中方方案四的流程图。如图5所示，本实施例中根据不确定性信息实现汉字输入的方法另一种实施例中方案四的流程具体包括如下步骤：

[0155] 方案四是上述九种方案中相对复杂的一种，包含了三类不确定性信息，分别为拼音串、拼音声调和结构布局。上述过程同方案一，只是加入在拼音字符串的不确定信息中，又加入了拼音声调检测信息以及结构布局信息两种不确定信息，过程如对图4的描述，具体不再赘述。

[0156] 由于加入了其余两种不确定信息中的因素，引入了对字符进行筛选融合的过程，具体如下所述：将相同输入下不同处理依据（拼音，汉字）得到的候选序列进行排序，得到统一的候选序列。

[0157] 与方案一相比，方案四所包含的信息种类更多，其不确定性信息就越多，越能够尽可能地缩小不确定性信息所确定的字符集合，从而减少该字符同上下文组合成词组的数目，进一步地减少语言模型打分所需计算时间，提高计算精度，从整体上提升系统性能。

[0158] 在上述实施例中，首先接收用户的输入信息；从输入信息中提取待输入汉字的至少两种不确定性信息；根据至少两种不确定性信息，确定并输出匹配的汉字；使得当用户对待输入汉字的读音信息或者字型信息记忆不完全时，将待输入汉字对应的字符的候选限制在一定范围内，并且结合所提取出的待输入汉字的至少两种不确定性信息，例如，拼音信息、声调信息、结构布局信息、偏旁部首信息、同音字信息以及位置信息等信息，实现待输入汉字的准确输入。

[0159] 本发明实施例另一方面提供了一种根据不确定性信息实现汉字输入的方法，包括以下步骤：

[0160] 接收用户的输入信息；

[0161] 从输入信息中提取待输入汉字的不确定性信息，其中，不确定性信息包括以下任意一种或多种信息：

[0162] 声调信息、结构布局信息、偏旁部首信息、同音字信息和位置信息；

[0163] 基于已经确定的汉字和不确定性信息，确定并输出待输入汉字。

[0164] 在本发明的实施例中，通过已经确定的汉字，例如是待输入汉字的上下文的内容，然后再结合声调信息、结构布局信息、偏旁部首信息、同音字信息和位置信息中的任意一种

或多种不确定性信息,就能够准确便捷地确定并输出待输入汉字。也就是,本发明实施例提供的方案只需要一种不确定性信息,也可实现汉字输入。

[0165] 具体地,例如,当待输入汉字为“味同嚼蜡”时,忘记具体如何去写“嚼”字,知道该词语中已经确定的汉字包括“味同”和“蜡”,结合这个字的声调信息、结构布局信息、偏旁部首信息、同音字信息和位置信息中的任意一种或多种不确定性信息,例如,该字的拼音声调为二声,或该字的结构布局为左右结构,或该字的偏旁部首中的右半部分的字形为“爵”,或该字的同音字为“交”中的汉字之一,或该字位于词语中的第三个字,就很容易确定待输入的汉字为“嚼”。也就是,通过基于已经确定的汉字(前两个字以及顺序是“味同”,以及第四个字是“蜡”)和不确定性信息(该字的拼音声调为二声,或该字的结构布局为左右结构,或该字的偏旁部首中的右半部分的字形为“爵”,或该字的同音字为“交”中的汉字之一,或该字位于词语中的第三个字),就可以准确便捷地输出待输入汉字“嚼”字。

[0166] 在本应用场景中,通过接收用户的输入信息;从输入信息中提取待输入汉字的不确定性信息;其中,不确定性信息包括以下任意一种或多种信息:声调信息、结构布局信息、偏旁部首信息、同音字信息和位置信息;基于已经确定的汉字和不确定性信息,确定并输出待输入汉字;相对于上述根据至少两种不确定性信息,确定并输出匹配的汉字的技术方案,本方案充分地利用了已经确定的待输入汉字的上下文的内容,只需要声调信息、结构布局信息、偏旁部首信息、同音字信息和位置信息中的任意一种信息,就能够准确便捷地确定并输出待输入汉字。

[0167] 由于在本发明实施例中的基于已经确定的汉字和不确定性信息,确定并输出待输入汉字的过程与上述方法实施例中已经详细描述的进行至少两种不确定性信息,确定并输出匹配的汉字的过程相类似,因此,上述根据至少两种不确定性信息,确定并输出匹配的汉字的技术方案中的实施例,也同样适用于本发明实施例中的基于已经确定的汉字和以下任意一种或多种不确定性信息,确定并输出待输入汉字的技术方案中,具体细节在此不再赘述,请参见相应部分的描述。

[0168] 图6为本发明根据不确定性信息实现汉字输入的装置600一种实施例的结构示意图。如图6所示,本实施例根据不确定性信息实现汉字输入的装置包括:接收模块610、提取模块620以及匹配模块630。

[0169] 接收模块610,用于接收用户的输入信息。

[0170] 作为本发明的实施例,接收模块接收用户的输入信息,包括:

[0171] 在输入区域检测用户的操作;

[0172] 记录用户的操作输入的至少两种不确定性信息。

[0173] 具体而言,接收模块610用于通过以下任意一种方式接收声调信息:

[0174] 接收用户手写输入声调信息的符号:ˊ,ˋ,ˊ或ˋ;

[0175] 接收特定的数字或字母来表示声调信息。

[0176] 具体而言,接收模块610用于通过以下任意一种方式接收位置信息:

[0177] 接收用户输入的空格或特定的字符。

[0178] 提取模块620,用于从输入信息中提取待输入汉字的至少两种不确定性信息。

[0179] 具体而言,提取模块620提取的不确定性信息包括以下任意一种信息:拼音信息、声调信息、结构布局信息、偏旁部首信息、同音字信息以及位置信息。其中,拼音信息包括

拼音字符串信息和拼音声调信息,拼音字符串信息用于根据书写的笔迹特点和书写区域,分离拼音串;拼音声调信息用于根据拼音声调书写的笔迹特点和书写区域,分离拼音声调;结构布局信息用于根据结构布局信息的笔迹特点和书写区域,分离结构布局信息;偏旁部首信息用于根据结构布局信息的笔迹特点和书写区域,分离用户可以确定的偏旁部首;同音字信息用于根据同音字书写的笔迹特点和书写区域,分离同音字信息。

[0180] 匹配模块 630,用于根据至少两种不确定性信息,确定并输出匹配的汉字。

[0181] 具体而言,根据至少两种不确定性信息,匹配模块 630 确定并输出匹配的汉字,包括:

[0182] 根据至少两种不确定性信息选择相应的模型;

[0183] 在至少两种相应的模型中匹配待输入汉字;

[0184] 对每种模型的匹配结果进行分析,确定匹配度最高的汉字为待输入汉字,输出匹配度最高的汉字。

[0185] 具体而言,匹配模块 630 还用于:

[0186] 确定不确定性信息的优先级;

[0187] 根据优先级,依次根据不确定性信息在相应模型中匹配待输入汉字。

[0188] 在上述实施例中,通过接收用户的输入信息;从输入信息中提取待输入汉字的至少两种不确定性信息;根据至少两种不确定性信息,确定并输出匹配的汉字;使得当用户对待输入汉字的读音信息或者字型信息记忆不完全时,将待输入汉字对应的字符的候选限制在一定范围内,并且结合所提取出的待输入汉字的至少两种不确定性信息,例如,拼音信息、声调信息、结构布局信息、偏旁部首信息、同音字信息以及位置信息等信息,实现待输入汉字的准确输入。

[0189] 图 7 为本发明根据不确定性信息实现汉字输入的装置另一种实施例的结构示意图。如图 7 所示,本实施例根据不确定性信息实现汉字输入的装置还包括:设置模块 640、更新模块 650 和存储模块 660。

[0190] 作为上述根据不确定性信息实现汉字输入的装置的实施例,设置模块 640 用于接收用户的输入,设置不确定性信息的优先级。

[0191] 进一步地,作为上述根据不确定性信息实现汉字输入的装置的实施例,接收模块 610 还用于在屏幕上显示至少两种不确定性信息。

[0192] 更进一步地,作为上述根据不确定性信息实现汉字输入的装置的实施例,更新模块 650 检测到用户在输入区域修改不确定性信息时,在屏幕上显示更新后的不确定性信息。

[0193] 进一步地,存储模块 660,用于存储用户的输入信息,(例如,用户手写输入的笔迹)、语言模型、拼音-中文字符对应词表、中文及拼音识别库等。

[0194] 除此之外,本实施例中的装置部分的处理器还具有如下功能,具体如下所述:

[0195] 对接收模块 610 接收的用户的输入信息中的笔迹进行预处理,去噪声,倾斜校正等一系列必要的预处理过程;

[0196] 对每个字符的范围进行切分识别过程;

[0197] 根据提取模块 620 所提取的不确定性信息的格式,将用户的输入信息中的不确定性信息同文本文字相区分的过程;

- [0198] 根据拼音笔迹得到拼音字符串,以识别拼音笔迹的过程;
- [0199] 根据结构布局笔迹得到结构布局分类,以识别结构布局信息;
- [0200] 识别拼音声调,识别偏旁部首,识别同音字的过程;
- [0201] 通过对剥离了不确定性信息中的中文文本字符进行识别,以及对不确定性信息中的同音字进行识别,以识别相应的中文字符的识别过程;
- [0202] 通过汉字 → 拼音的查表转换,以根据汉字的识别结果,得到相应的汉字的拼音串的过程;
- [0203] 通过拼音拼音 + 声调 → 字符的查表转换,以根据拼音串和声调组合信息同中文字符的对应关系表,得到拼音 + 声调到中文字符的转换结果的过程;
- [0204] 通过结构布局类别 → 字符的查表转换,以根据结构布局分类的识别结果,映射到对应了该分类的字符集合的过程;
- [0205] 字符筛选融合过程,具体为:将经由不同类别(拼音、结构布局、同音字)的不确定性信息推导出的字符集合进行筛选(依据是偏旁部首的识别结果,如果存在的话)以及融合(依据是按照一定的权重打分规则得到的一组按高低排列的分数),以将不同种类的不确定性信息得到的字符候选进行筛选,排序。
- [0206] 候选筛选融合过程,具体为:将不同依据下得到的大于一个候选序列按照一定规则进行排序,融合成一个最终输出的候选序列。
- [0207] 语言模型打分过程,具体为:分为对汉字语言模型进行打分的过程和对拼音串语言模型进行打分的过程两类。功能分别为:将拼音对应的中文字符候选 / 中文字符对应的拼音串候选,同前后文的中文字符 / 拼音串拼接组合,对不同的组合在此模块中进行打分,分值越高,该组合在输入候选中排名越靠前;以根据统计的中文词条 / 拼音串及其出现频率,通过模型打分计算拼接后的词条正确与否的概率。
- [0208] 在上述实施例,通过接收用户的输入信息;从输入信息中提取待输入汉字的至少两种不确定性信息;根据至少两种不确定性信息,确定并输出匹配的汉字;使得当用户对待输入汉字的读音信息或者字型信息记忆不完全时,将待输入汉字对应的字符的候选限制在一定范围内,并且结合所提取出的待输入汉字的至少两种不确定性信息,例如,拼音信息、声调信息、结构布局信息、偏旁部首信息、同音字信息以及位置信息等信息,实现待输入汉字的准确输入。
- [0209] 本发明实施例另一方面提供了一种根据不确定性信息实现汉字输入的装置,包括:接收模块、提取模块以及匹配模块;
- [0210] 接收模块,用于接收用户的输入信息;
- [0211] 提取模块,用于从输入信息中提取待输入汉字的不确定性信息,其中,不确定性信息包括以下任意一种或多种信息:
- [0212] 声调信息、结构布局信息、偏旁部首信息、同音字信息和位置信息;
- [0213] 匹配模块,用于基于已经确定的汉字和不确定性信息,确定并输出待输入汉字。
- [0214] 由于在本发明实施例中的匹配模块的基于已经确定的汉字和不确定性信息,确定并输出待输入汉字的过程与上述方法或装置实施例中已经详细描述根据至少两种不确定性信息,确定并输出匹配的汉字的过程相类似,因此,上述方法或装置实施例中根据至少两种不确定性信息,确定并输出匹配的汉字的技术方案中的实施例,也同样适用于本发明

实施例中的匹配模块的基于已经确定的汉字和以下任意一种或多种不确定性信息,确定并输出待输入汉字的技术方案中,具体细节在此不再赘述,请参见相应部分的描述。

[0215] 当每个字符位置的可利用信息中至少包含了拼音或汉字二者任意其一,图 8 为根据不确定性信息确定匹配度最高的汉字的示意图,包括如下处理过程:

[0216] 根据至少两种不确定性信息选择相应的模型;具体处理过程结合下述实例进行描述。

[0217] 接收输入信息,将“味同 jiao la”分别转换成汉字和拼音串。分别转换的目的是在尽可能地利用现有信息,分别在汉字模型和拼音模型下对不确定的输入进行建模分析,得到各自模型下的结果,并进行融合,提高结果可信度。

[0218] 显示在上述过程中所转换成的汉字或拼音串:

[0219] 具体而言,若根据“味同 jiao la”中的至少两种不确定性信息选择的是汉字语言模型,则进行以下处理流程:

[0220] 在汉字语言模型中匹配待输入汉字“味同 jiao la”;则“味同 jiao la”可能转换成的汉字为:“味同叫啦”、“妹问交蜡”以及“妹包嚼辣”;

[0221] 上述包含汉字“味同叫啦”、“妹问交蜡”以及“妹包嚼辣”的全部信息输入到汉字语言模型中进行匹配,其匹配结果为:

[0222] 可能的汉字组合为:“味同嚼蜡”、“味同角辣”、“味同交啦”、“妹问叫拉”。

[0223] 具体而言,若根据“味同 jiao la”中的至少两种不确定性信息选择的是拼音串语言模型,则进行以下处理流程:

[0224] 在拼音串语言模型中匹配待输入汉字“味同 jiao la”;则

[0225] “味同 jiao la”可能转换成的拼音串为:“wei tong jiao la”、“mei wen qiao”以及“shu bao”;

[0226] 上述包含拼音串“wei tong jiao la”、“mei wen qiao”以及“shu bao”的全部信息输入到拼音串语言模型中进行匹配,其匹配结果为:

[0227] 可能的汉字组合为:“味同嚼蜡”、“味同嚼辣”、“味同交啦”、“妹同交啦”。

[0228] 根据一定的规则融合两个候选序列,例如在两个模型下各自得到的字符串中寻找是否有相同结果,如果有,则将其列在融合序列中。如例子中所示,两列字符串中均包含“味同嚼蜡”,则将“味同嚼蜡”作为融合后结果的第一个字符串候选。同样的,两列字符串中同样均包含“味同交啦”,则将“味同交啦”作为融合后结果的第二个字符串候选。接下来两列字符串中已没有相同的结果,那么对于那些只存在于一个模型下的字符串,设计融合打分规则将字符串按照分数的高低进行排列。例如可以直接对两个模型的打分进行排序,或者根据模型参数计算加权参数将两个模型下的分数换算成统一度量下的分数,再根据分数对剩余字符串进行统一排序。得到的融合结果如下所述:“味同嚼蜡”、“味同交啦”、“味同嚼辣”以及“妹同交啦”。

[0229] 通过对上述汉字语言模型和拼音串语言模型的匹配结果进行分析,确定匹配度最高的汉字为待输入汉字,输出匹配度最高的汉字“味同嚼蜡”。

[0230] 作为本发明的实施例,在实际应用中,上述根据不确定性信息实现汉字输入的装置通常以终端设备的形式体现,终端设备包括图 6 或图 7 所示的根据不确定性信息实现汉字输入的装置。

[0231] 本发明提出的上述方案,使得当用户对待输入汉字的读音信息或者字型信息记忆不完全时,将待输入汉字对应的字符的候选限制在一定范围内,并且结合所提取出的待输入汉字的至少两种不确定性信息,实现待输入汉字的准确输入。此外,本发明提出的上述方案,对现有系统的改动很小,不会影响系统的兼容性,而且实现简单、高效。

[0232] 本技术领域技术人员可以理解,本发明包括涉及用于执行本申请中所述操作中的一项或多项的设备。这些设备可以为所需的目的而专门设计和制造,或者也可以包括通用计算机中的已知设备。这些设备具有存储在其内的计算机程序,这些计算机程序选择性地激活或重构。这样的计算机程序可以被存储在设备(例如,计算机)可读介质中或者存储在适于存储电子指令并分别耦联到总线的任何类型的介质中,所述计算机可读介质包括但不限于任何类型的盘(包括软盘、硬盘、光盘、CD-ROM、和磁光盘)、ROM (Read-Only Memory,只读存储器)、RAM (Random Access Memory,随机存储器)、EPROM (Erasable Programmable Read-Only Memory,可擦写可编程只读存储器)、EEPROM (Electrically Erasable Programmable Read-Only Memory,电可擦可编程只读存储器)、闪存、磁性卡片或光线卡片。也就是,可读介质包括由设备(例如,计算机)以能够读的形式存储或传输信息的任何介质。

[0233] 本技术领域技术人员可以理解,可以用计算机程序指令来实现这些结构图和/或框图和/或流程图中的每个框以及这些结构图和/或框图和/或流程图中的框的组合。本技术领域技术人员可以理解,可以将这些计算机程序指令提供给通用计算机、专业计算机或其他可编程数据处理方法的处理器来实现,从而通过计算机或其他可编程数据处理方法的处理器来执行本发明公开的结构图和/或框图和/或流程图的框或多个框中指定的方案。

[0234] 本技术领域技术人员可以理解,本发明中已经讨论过的各种操作、方法、流程中的步骤、措施、方案可以被交替、更改、组合或删除。进一步地,具有本发明中已经讨论过的各种操作、方法、流程中的其他步骤、措施、方案也可以被交替、更改、重排、分解、组合或删除。进一步地,现有技术中的具有与本发明中公开的各种操作、方法、流程中的步骤、措施、方案也可以被交替、更改、重排、分解、组合或删除。

[0235] 以上所述仅是本发明的部分实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

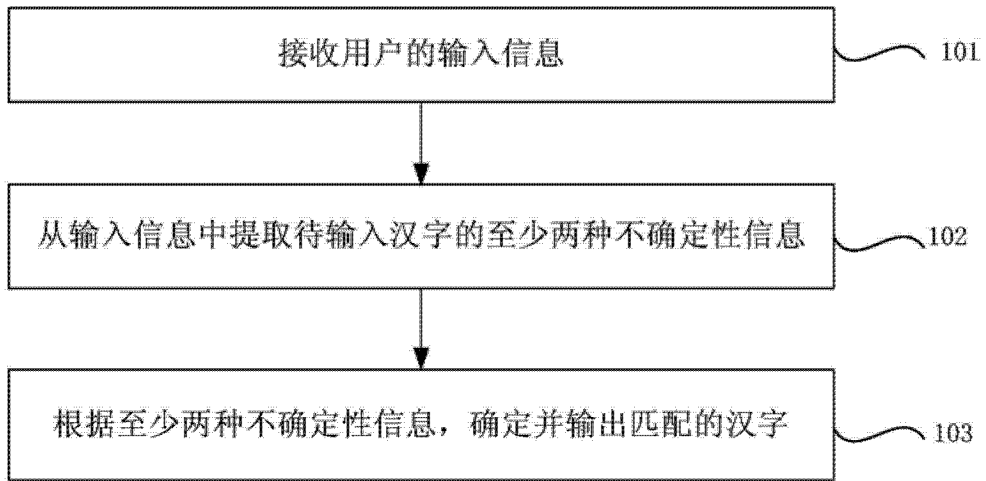


图 1

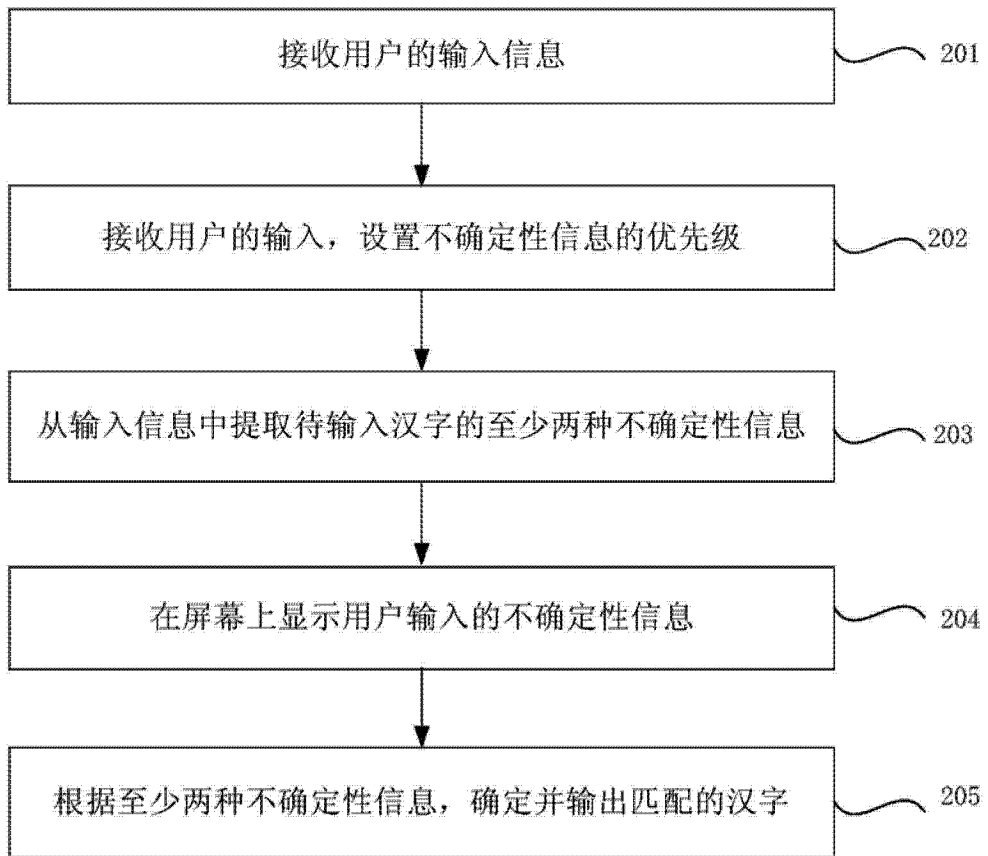


图 2

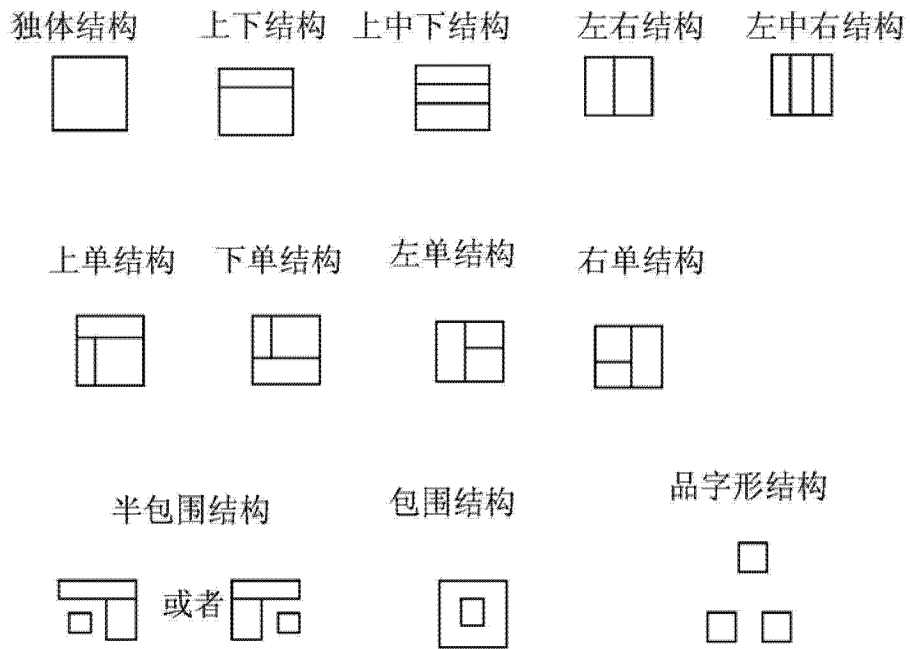


图 3

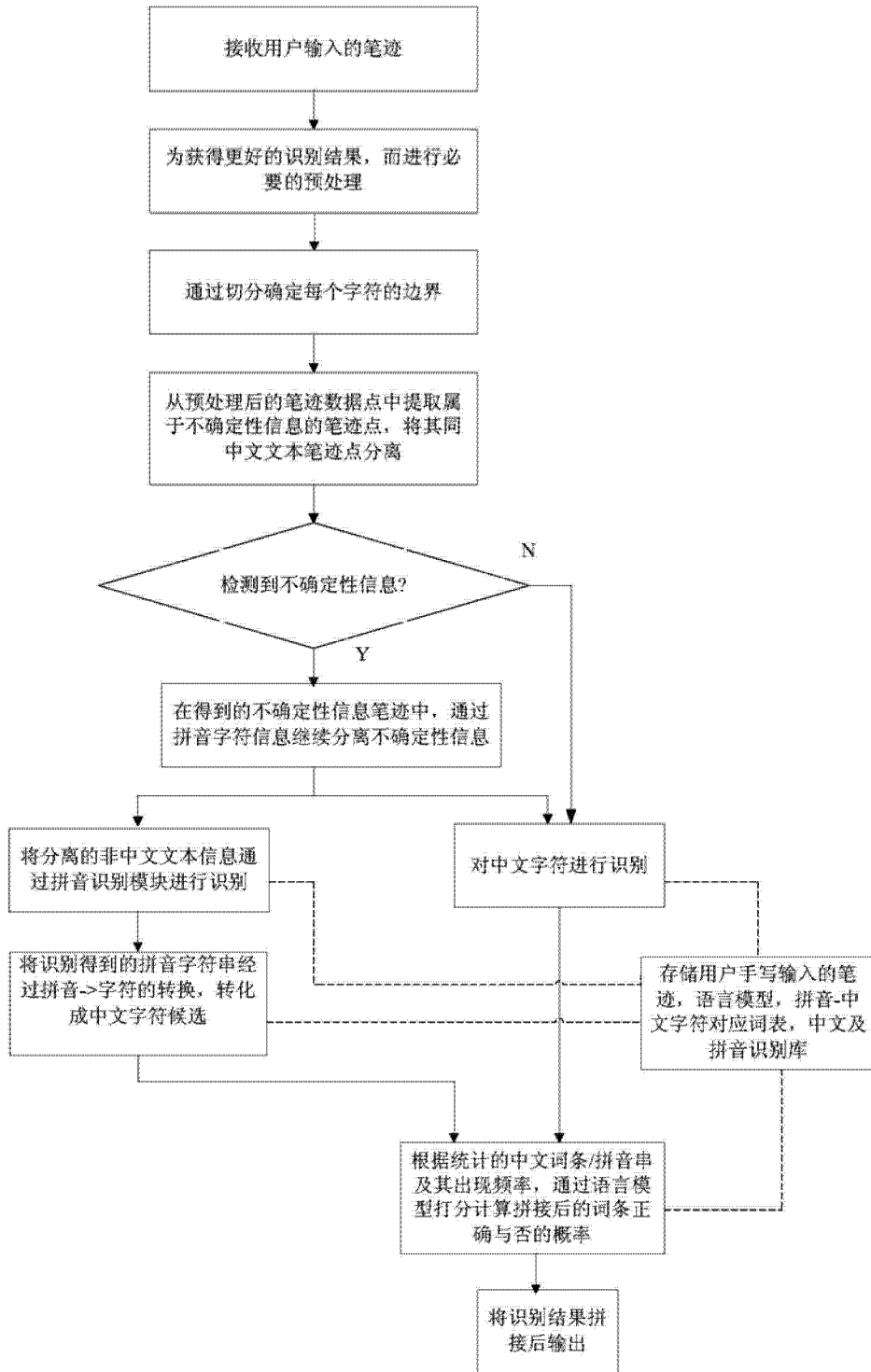


图 4

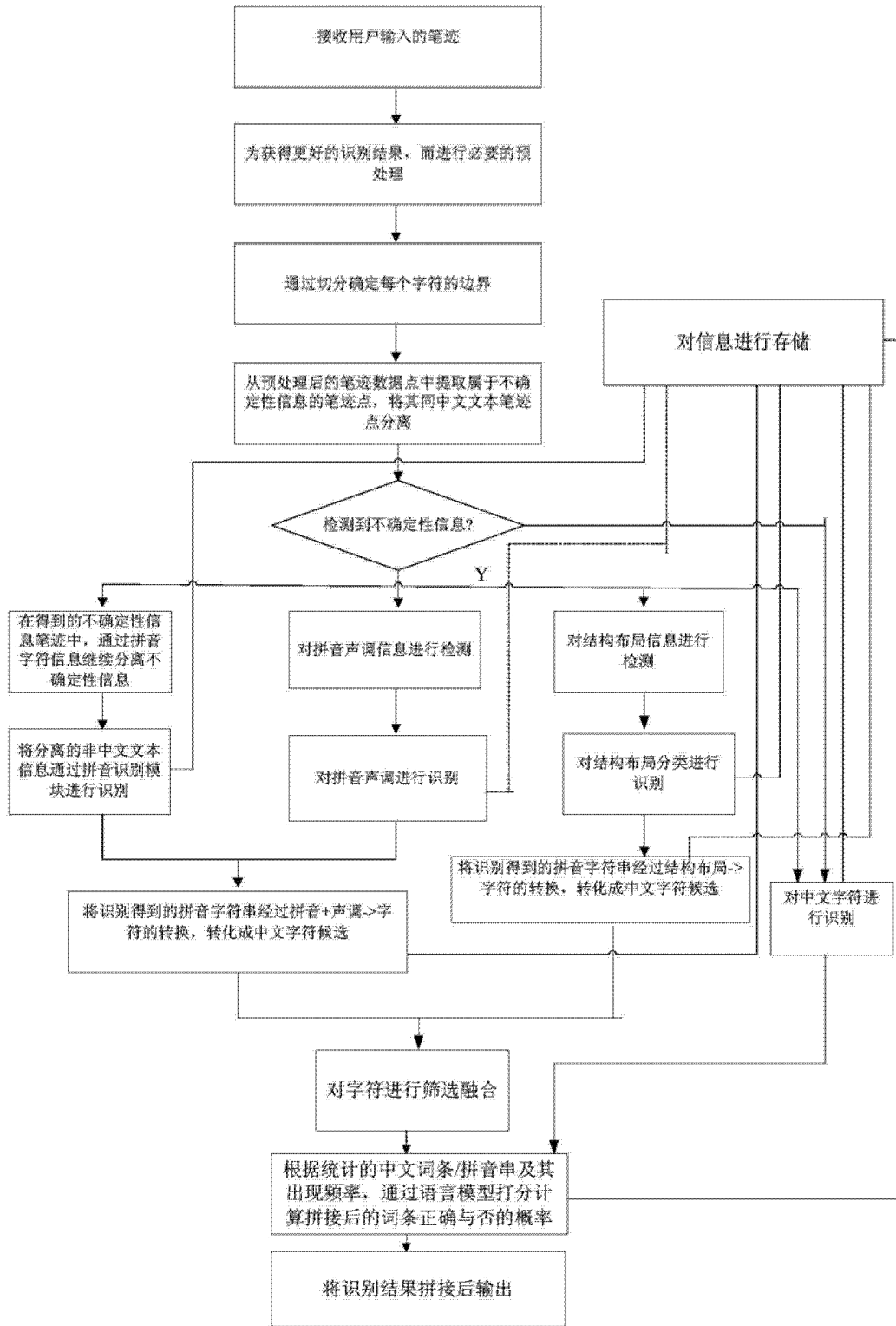


图 5

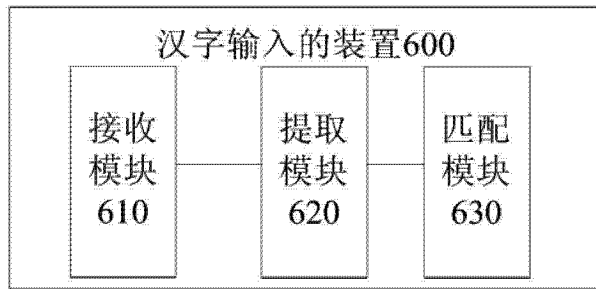


图 6

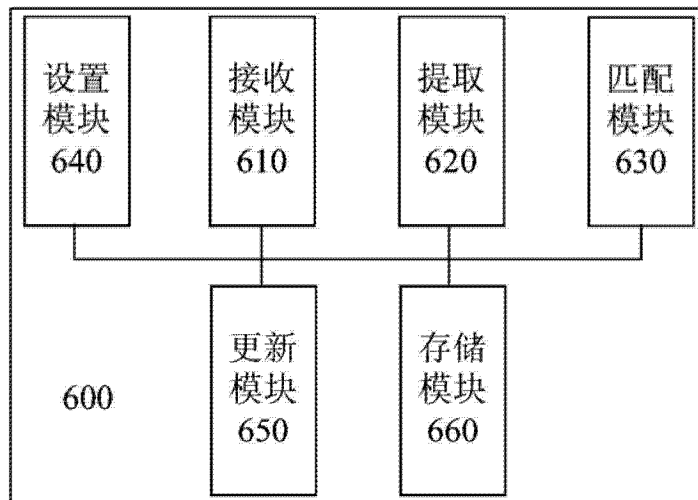


图 7

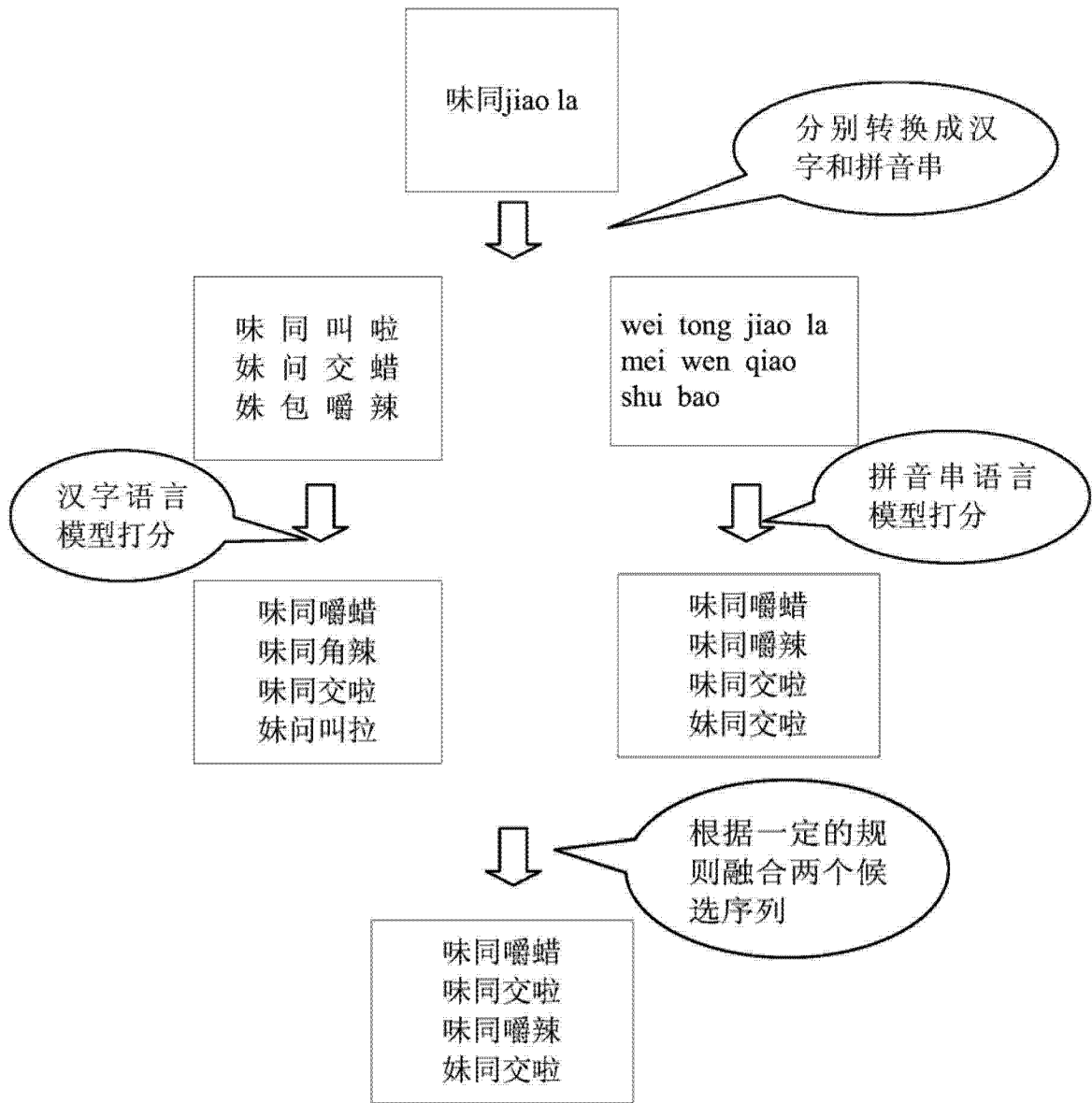


图 8