



(12) 发明专利

(10) 授权公告号 CN 107589855 B

(45) 授权公告日 2021. 05. 28

(21) 申请号 201710495840.3

(22) 申请日 2012.05.29

(65) 同一申请的已公布的文献号
申请公布号 CN 107589855 A

(43) 申请公布日 2018.01.16

(62) 分案原申请数据
201210172759.9 2012.05.29

(73) 专利权人 阿里巴巴集团控股有限公司
地址 英属开曼群岛大开曼资本大厦一座四
层847号邮箱

(72) 发明人 傅茂建

(74) 专利代理机构 北京润泽恒知识产权代理有
限公司 11319
代理人 刘祥景

(51) Int.Cl.
G06F 3/023 (2006.01)
G06F 16/00 (2019.01)
审查员 毛元超

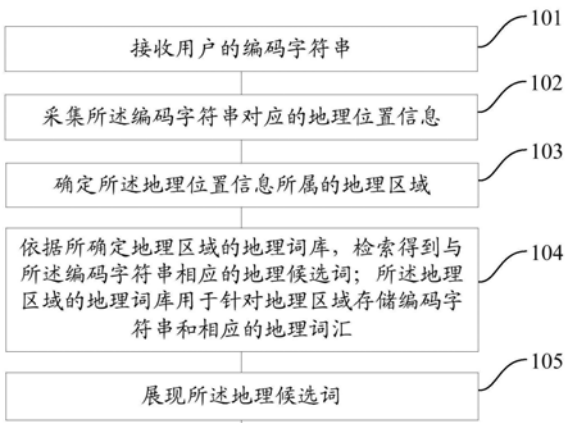
权利要求书2页 说明书15页 附图3页

(54) 发明名称

一种根据地理位置推荐候选词的方法和装置

(57) 摘要

本申请提供了一种根据地理位置推荐候选词的方法和装置,其中的方法具体包括:接收用户的编码字符串;采集所述编码字符串对应的地理位置信息;确定所述地理位置信息所属的地理区域;依据所确定地理区域的地理词库,检索得到与所述编码字符串相应的地理候选词;所述地理区域的地理词库用于针对地理区域存储编码字符串和相应的地理词汇;展现所述地理候选词。本申请能够减轻输入的繁琐程度,且能够提高输入法的智能性。



1. 一种根据地理位置推荐候选词的方法,其特征在于,包括:

接收用户的编码字符串;

采集所述编码字符串对应的地理位置信息;

在所述编码字符串对应的应用程序环境信息与地理位置的相关度符合预设的相关度条件的情况下,依据所述地理位置信息所对应地理区域的地理词库,检索得到与所述编码字符串和所述地理位置信息相应的地理候选词;所述地理区域的地理词库用于存储编码字符串、相应地理区域内的地理位置信息和相应的地理词汇;所述地理词汇包括:相应地理区域内的历史输入词汇;所述地理词汇包括:未命中相应地理区域的地理词库、且地理位置属性符合预设的地理位置属性条件的上屏字符;所述应用程序环境信息包括:应用程序的标识、或应用程序的类别;

展现所述地理候选词。

2. 如权利要求1所述的方法,其特征在于,通过如下步骤构建所述地理区域的地理词库:

对地理区域内的地理词汇的使用频率进行统计,收集统计结果符合预设的使用频率条件的地理词汇;

将所收集的地理词汇及相应的地理位置信息和编码字符串储存至所述地理区域的地理词库。

3. 如权利要求2所述的方法,其特征在于,通过如下步骤获取所述地理区域内的地理词汇:

依据地理位置类别从互联网信息中收集所述地理区域内的地理词汇;和/或

在用户输入触发条件词汇并上屏时,记录上屏的触发条件词汇;所述触发条件词汇为依据地理位置类别预设的所述地理区域内的地理词汇;和/或

将地理区域内的历史输入词汇作为所述地理区域内的地理词汇。

4. 如权利要求1至3中任一所述的方法,其特征在于,所述方法还包括:

当用户上屏的字符未命中所述地理区域的地理词库时,分析未命中的字符对应的地理位置属性;

在分析得到的地理位置属性符合预设的地理位置属性条件时,依据所述未命中的字符,对所述地理区域的地理词库进行更新。

5. 如权利要求1所述的方法,其特征在于,还包括:

获取所述编码字符串对应的应用程序环境信息。

6. 如权利要求1至3中任一所述的方法,其特征在于,还包括:

依据所述编码字符串,在普通词库中进行检索,得到相应的普通候选词;

对所述普通候选词和所述地理候选词进行展现。

7. 如权利要求6所述的方法,所述对所述普通候选词和所述地理候选词进行展现,进一步包括:

依据候选词的排序因子,对所述普通候选词和地理候选词进行排序;所述排序因子包括如下因子中的一项或多项:候选词与所述地理位置信息的匹配度、候选词的使用频率、搜索排序策略和应用程序环境信息与地理位置的相关度;

依据排序结果,对所述普通候选词和地理候选词进行展现。

8. 一种根据地理位置推荐候选词的装置,其特征在于,包括:

接口模块,用于接收用户的编码字符串;

采集模块,用于采集所述编码字符串对应的地理位置信息;

地理区域确定模块,用于确定所述地理位置信息所属的地理区域;

第一检索模块,用于在所述编码字符串对应的应用程序环境信息与地理位置的相关度符合预设的相关度条件的情况下,依据所述地理位置信息所对应地理区域的地理词库,检索得到与所述编码字符串和所述地理位置信息相应的地理候选词;所述地理区域的地理词库用于存储编码字符串、相应地理区域内的地理位置信息和相应的地理词汇;所述地理词汇包括:相应地理区域内的历史输入词汇;所述地理词汇包括:未命中相应地理区域的地理词库、且地理位置属性符合预设的地理位置属性条件的上屏字符;所述应用程序环境信息包括:应用程序的标识、或应用程序的类别;及

展现模块,用于展现所述地理候选词。

9. 如权利要求8所述的装置,其特征在于,还包括:

用于构建地理区域的地理词库的地理词库构建模块,所述地理词库构建模块包括:

统计子模块,用于对地理区域内的地理词汇的使用频率进行统计;

收集子模块,用于收集统计结果符合预设的使用频率条件的地理词汇;及

储存子模块,用于将所收集的地理词汇及相应的编码字符串和地理位置信息储存至所述地理区域的地理词库。

10. 如权利要求9所述的装置,其特征在于,还包括:

用于获取所述地理区域内的地理词汇的第一获取模块,所述第一获取模块包括:

互联网搜集子模块,用于依据地理位置类别从互联网信息中搜集该地理区域内的地理词汇;和/或

触发获取子模块,用于在用户输入触发条件词汇并上屏时,记录上屏的触发条件词汇;所述触发条件词汇为依据地理位置类别预设的该地理区域内的地理词汇和/或

历史输入词汇获取子模块,用于将地理区域内的历史输入词汇作为所述地理区域内的地理词汇。

11. 如权利要求8至10中任一所述的装置,其特征在于,所述装置还包括:

地理位置属性分析模块,用于当用户上屏的字符未命中所述地理区域的地理词库时,分析未命中的字符对应的地理位置属性;

地理词汇更新模块,用于在分析得到的地理位置属性符合预设的地理位置属性条件时,依据所述未命中的字符,对所述地理区域的地理词库进行更新。

一种根据地理位置推荐候选词的方法和装置

技术领域

[0001] 本申请涉及输入法技术领域,特别是涉及一种根据地理位置推荐候选词的方法和装置。

背景技术

[0002] 目前随着计算机技术以及互联网技术的普及与发展,一方面,用于实现计算机文字输入的输入法越来越深入到人们的生活中,另一方面,用户对于输入法的智能性要求也越来越高。

[0003] 为提高智能性,现有技术一些输入法能够支持地名和地方方言的输入,其构建地方语言包(地名+地方方言),如果用户下载地方语言包到本地,则使用其中的地名或地方方言。例如,杭州的用户下载了四川语言包,则在用户输入编码字符串时,如果该编码字符串命中四川语言包中的地名或四川方言,则会优先展现该命中的地名或方言候选词,供用户选择。

[0004] 现有技术一些输入法能够满足地名或地方方言的快速输入,但随着移动通讯技术的发展和移动设备的普及,在有些情况下用户所处的地理位置对用户的输入需求也有很大的影响,尤其当用户所使用的移动设备经常移动时,这种影响尤为显著。例如,当用户到一个商场逛街时,很有可能告知朋友该商场的名称;又如,当用户在一个餐馆吃饭时,很有可能通过发微博介绍该餐馆,或者,当用户在一个旅游城市度假时,由于对该旅游景点不熟悉,可能需要在地图中搜索该旅游城市周边的商圈、餐饮、酒店、超市、电影院、景点、学校、银行等服务信息。

[0005] 在上述的应用场景中,显然现有的输入法并不能智能地向用户提供这些与地理位置相关的输入内容。例如,用户下载了广东方言包,其在杭州市的江干区富春路701号万象城逛街时,输入“zs j”的本意是获得候选词“张生记”,但现有技术展现地名“珠三角”对用户的输入需求毫无益处。此时,如果用户需要获得候选词“张生记”,则需要通过全拼输入“zhangshengji”才能更准确地获得想要的候选词,这无疑增加了拼音输入的繁琐程度。

[0006] 因此,现有技术不能智能地了解用户的输入需求,不能总是在输入过程中出现用户想要的候选词,增加了输入的繁琐程度,这对输入法的智能性提出了更高的要求。

[0007] 总之,需要本领域技术人员迫切解决的一个技术问题就是:如何能够减轻输入的繁琐程度,以及提高输入法的智能性。

发明内容

[0008] 本申请所要解决的技术问题是提供一种根据地理位置推荐候选词的方法和装置,能够减轻输入的繁琐程度,且能够提高输入法的智能性。

[0009] 为了解决上述问题,本申请公开了一种根据地理位置推荐候选词的方法,包括:

[0010] 接收用户的编码字符串;

[0011] 采集所述编码字符串对应的地理位置信息;

- [0012] 确定所述地理位置信息所属的地理区域；
- [0013] 依据所确定地理区域的地理词库，检索得到与所述编码字符串相应的地理候选词；所述地理区域的地理词库用于针对地理区域存储编码字符串和相应的地理词汇；
- [0014] 展现所述地理候选词。
- [0015] 优选的，通过如下步骤构建所述地理区域的地理词库：
- [0016] 对地理区域内的地理词汇的使用频率进行统计，收集统计结果符合相应的预设的使用频率条件的地理词汇；
- [0017] 将所收集的地理词汇及相应的编码字符串储存至该地理区域的地理词库。
- [0018] 优选的，通过如下步骤获取该地理区域内的地理词汇：
- [0019] 依据地理位置类别从互联网信息中搜集该地理区域内的地理词汇；和/或
- [0020] 在用户输入触发条件词汇并上屏时，记录上屏的触发条件词汇；所述触发条件词汇为依据地理位置类别预设的该地理区域内的地理词汇。
- [0021] 优选的，所述地理区域的地理词库还用于存储编码字符串、相应地理区域内的地理位置信息和相应的地理词汇；所述方法还包括：
- [0022] 依据所确定地理区域的地理词库，检索得到与所述编码字符串和所述地理位置信息相应的地理候选词。
- [0023] 优选的，所述方法还包括：
- [0024] 获取所述编码字符串对应的应用程序环境信息；
- [0025] 则所述依据所确定地理区域的地理词库，检索得到与所述编码字符串相应的地理候选词的步骤进一步包括：在所述应用程序环境信息与地理位置的相关度符合预设的相关度条件时，依据所确定地理区域的地理词库，检索得到与所述编码字符串相应的地理候选词。
- [0026] 优选的，所述方法还包括：
- [0027] 依据所述编码字符串，在普通词库中进行检索，得到相应的普通候选词；
- [0028] 所述展现所述地理候选词的步骤为，在展现所述普通候选词的同时，展现所述地理候选词。
- [0029] 优选的，所述展现所述地理候选词的步骤进一步包括：
- [0030] 依据候选词的排序因子，对所述普通候选词和地理候选词进行排序；所述排序因子包括如下因子中的一项或多项：候选词与所述地理位置信息的匹配度、候选词的使用频率和搜索排序策略；
- [0031] 依据排序结果，对所述普通候选词和地理候选词进行展现。
- [0032] 优选的，所述排序因子还包括所述应用程序环境信息与地理位置的相关度。
- [0033] 优选的，所述依据所确定地理区域的地理词库，检索得到与所述编码字符串相应的地理候选词的步骤进一步包括：
- [0034] 从客户端的所确定地理区域的地理词库中检索得到与所述编码字符串相应的地理候选词；和/或
- [0035] 将所述编码字符串和所述地理位置信息发送至服务器端，由服务器端从所确定地理区域的地理词库中检索得到与所述编码字符串相应的地理候选词。
- [0036] 优选的，所述方法还包括：

[0037] 当用户上屏的字符未命中所确定地理区域的地理词库时,分析该未命中的字符的地理位置属性;

[0038] 在分析得到的地理位置属性符合预设的地理位置属性条件时,将该未命中的字符作为地理词汇同步至所有客户端的相应地理区域的地理词库。

[0039] 另一方面,本申请还公开了一种根据地理位置推荐候选词的装置,包括:

[0040] 接口模块,用于接收用户的编码字符串;

[0041] 采集模块,用于采集所述编码字符串对应的地理位置信息;

[0042] 地理区域确定模块,用于确定所述地理位置信息所属的地理区域;

[0043] 第一检索模块,用于依据所确定地理区域的地理词库,检索得到与所述编码字符串相应的地理候选词;所述地理区域的地理词库用于针对地理区域存储编码字符串和相应的地理词汇;及

[0044] 展现模块,用于展现所述地理候选词。

[0045] 优选的,所述装置还包括:

[0046] 用于构建所述地理区域的地理词库的地理词库构建模块,所述地理词库构建模块包括:

[0047] 统计子模块,用于对地理区域内的地理词汇的使用频率进行统计;

[0048] 收集子模块,用于收集统计结果符合相应的预设的使用频率条件的地理词汇;及

[0049] 储存子模块,用于将所收集的地理词汇及相应的编码字符串储存至该地理区域的地理词库。

[0050] 优选的,所述装置还包括:

[0051] 用于获取该地理区域内的地理词汇的第一获取模块,所述第一获取模块包括:

[0052] 互联网搜集子模块,用于依据地理位置类别从互联网信息中搜集该地理区域内的地理词汇;和/或

[0053] 触发获取子模块,用于在用户输入触发条件词汇并上屏时,记录上屏的触发条件词汇;所述触发条件词汇为依据地理位置类别预设的该地理区域内的地理词汇。

[0054] 优选的,所述地理区域的地理词库还用于存储编码字符串、相应地理区域内的地理位置信息和相应的地理词汇;所述装置还包括:

[0055] 第二检索模块,用于依据所确定地理区域的地理词库,检索得到与所述编码字符串和所述地理位置信息相应的地理候选词。

[0056] 优选的,所述装置还包括:

[0057] 应用程序环境获取模块,用于获取所述编码字符串对应的应用程序环境信息;

[0058] 则所述第一检索模块,具体用于在所述应用程序环境信息与地理位置的相关度符合预设的相关度条件时,依据所确定地理区域的地理词库,检索得到与所述编码字符串和所述地理位置信息相应的地理候选词。

[0059] 优选的,所述装置还包括:

[0060] 第三检索模块,用于依据所述编码字符串,在普通词库中进行检索,得到相应的普通候选词;

[0061] 则所述展现模块,具体用于在展现所述普通候选词的同时,展现所述地理候选词。

[0062] 优选的,所述展现模块进一步包括:

[0063] 排序子模块,用于依据候选词的排序因子,对所述普通候选词和地理候选词进行排序;所述排序因子包括如下因子中的一项或多项:候选词与所述地理位置信息的匹配度、候选词的使用频率和搜索排序策略;

[0064] 次序展现子模块,用于依据排序结果,对所述普通候选词和地理候选词进行展现。

[0065] 优选的,所述排序因子还包括所述应用程序环境信息与地理位置的相关度。

[0066] 优选的,所述第一检索模块进一步包括:

[0067] 客户端检索子模块,用于从客户端的所确定地理区域的地理词库中检索得到与所述编码字符串和所述地理位置信息相应的地理候选词;和/或

[0068] 服务器端检索子模块,用于将所述编码字符串和所述地理位置信息发送至服务器端,由服务器端从所确定地理区域的地理词库中检索得到与所述编码字符串和所述地理位置信息相应的地理候选词。

[0069] 优选的,所述装置还包括:

[0070] 分析模块,用于当用户上屏的字符未命中所确定地理区域的地理词库时,分析该未命中的字符的地理位置属性;

[0071] 同步模块,用于在分析得到的地理位置属性符合预设的地理位置属性条件时,将该未命中的字符作为地理词汇同步至所有客户端的相应地理区域的地理词库。

[0072] 与现有技术相比,本申请具有以下优点:

[0073] 本申请依据输入的编码字符串和编码字符串对应的地理位置信息对候选词的综合影响,向用户推荐契合所述编码字符串和所述地理位置信息的地理候选词;并且,由于所述地理候选词为依据所述地理位置信息所属的地理区域的地理词库检索得到的,所述地理区域的地理词库用于存储编码字符串和相应的地理词汇,因此,本申请能够依据地理位置信息推荐用户所处地理区域内的地理候选词,也即能够在输入过程中出现用户想要的候选词,能够提高输入法的智能性,降低用户输入的繁琐程度。

[0074] 另外,本申请还可以在所述应用程序环境信息与地理位置的相关度符合预设的相关度条件时,才认为用户存在对地理候选词的输入需求,才会激活输入法的向用户推荐用户所处地理区域内的地理候选词的功能;相对于现有技术,本申请增加了用户输入需求的判断,且能够依据地理位置信息推荐用户所处地理区域内的地理候选词,也即能够在输入过程中出现用户想要的候选词,能够提高输入法的智能性。

附图说明

[0075] 图1是本申请一种根据地理位置推荐候选词的方法实施例1的流程图;

[0076] 图2是本申请一种根据地理位置推荐候选词的方法实施例2的流程图;

[0077] 图3是本申请一种根据地理位置推荐候选词的方法实施例3的流程图;

[0078] 图4是本申请一种根据地理位置推荐候选词的装置实施例的结构图。

具体实施方式

[0079] 为使本申请的上述目的、特征和优点能够更加明显易懂,下面结合附图和具体实施方式对本申请作进一步详细的说明。

[0080] 参照图1,其示出了本申请一种根据地理位置推荐候选词的方法实施例1的流程

图,具体可以包括:

[0081] 步骤101、接收用户的编码字符串;

[0082] 步骤102、采集所述编码字符串对应的地理位置信息;

[0083] 步骤103、确定所述地理位置信息所属的地理区域;

[0084] 步骤104、依据所确定地理区域的地理词库,检索得到与所述编码字符串相应的地理候选词;所述地理区域的地理词库用于针对地理区域存储编码字符串和相应的地理词汇;

[0085] 步骤105、展现所述地理候选词。

[0086] 本申请可以应用于各种用户需要与地理位置相关的输入内容的场景。

[0087] 例如,当用户到北京东城区的一个商场逛街时,需要告知朋友该商场的名称,于是在短信窗口中输入编码字符串“dfxtd”,则本申请可以依据地理位置信息自动推荐用户所处地理区域内的地理词汇“东方新天地”。

[0088] 又如,当用户在其在杭州市的江干区富春路701号万象城逛街时,看到了“张生记”这个餐馆,想搜索该餐馆的评价信息,于是在大众点评网上输入“zsj”,则本申请可以依据地理位置信息自动推荐用户所处地理区域内的地理词汇“张生记”。

[0089] 或者,当用户在丽江旅游时,需要查询某景点的位置,于是在地图中搜索输入“lshh”,则本申请可以依据地理位置信息自动推荐用户所处地理区域内的地理词汇“拉市海”。

[0090] 在输入过程中,用户希望输入法能够理解他们想要输入的内容,需要输入法能够根据他们的输入信息提供最想要的候选词,这是用户的输入需求。

[0091] 现有技术中将候选词局限在地名或方言这两个方面,且未考虑用户所处的地理位置对用户的输入需求产生的影响,故不能智能地了解用户的输入需求,不能总是在输入过程中出现用户想要的候选词。

[0092] 为了智能地了解用户的输入需求,本申请依据输入的编码字符串和编码字符串对应的地理位置信息对候选词的综合影响,向用户推荐契合所述编码字符串和所述地理位置信息的地理候选词;并且,由于所述地理候选词为依据所述地理位置信息所属的地理区域的地理词库检索得到的,所述地理区域的地理词库用于存储编码字符串和相应的地理词汇,因此,本申请能够依据地理位置信息推荐用户所处地理区域内的地理候选词,也即能够在输入过程中出现用户想要的候选词,能够提高输入法的智能性。

[0093] 编码字符串对应的地理位置信息主要是指用户在输入编码字符串时所处的地理位置信息。在实际中,可以针对一个用户的编码字符串,依据其IP(网络之间互连的协议, Internet Protocol)地址、移动设备GPS(全球定位系统, Global Positioning System)或移动网络采集得到相应的地理位置信息。本申请对具体的采集所述编码字符串对应的地理位置信息的方法不加以限制。

[0094] 关于如何确定所述地理位置信息所属的地理区域。在本申请的一种应用示例中,可以预存地理区域的地理位置信息范围(经纬度范围),然后将所述地理位置信息与预存地理区域的地理位置范围进行匹配,如果

[0095] 当前地理位置信息经度大于预存地理区域的GPS经度起始值

[0096] 并且当前地理位置信息经度小于预存地理区域的GPS经度截止值

[0097] 并且当前地理位置信息纬度大于预存地理区域的GPS纬度起始值

[0098] 并且当前地理位置信息纬度小于预存地理区域的GPS纬度截止值

[0099] 则可以确定当前地理位置信息所述的地理区域为匹配成功的该预存地理区域。

[0100] 在本申请的另一种应用示例中,可以依据所述地理位置信息去各种地理服务网站查询得到所属的地理区域。总之,本申请对具体的确定所述地理位置信息所属的地理区域的方法不加以限制。

[0101] 本领域技术人员可以根据实际需要调整所述地理区域的细粒度,例如,对于城市而言,所述地理区域的细粒度具体可以包括省、市、区、街道、社区等等;对于农村而言,所述地理区域的细粒度具体可以包括省、市、县、乡镇、村等等。

[0102] 本申请实施例中,地理区域的地理词库用于针对地理区域存储编码字符串和相应的地理词汇。故在本申请的一些实施例中,可以建立编码字符串与地理词汇之间的映射关系,并保存至相应的地理词库。

[0103] 参照表1,示出了本申请一种地理区域的地理词库的存储结构示例,其描述的是杭州市城区的地理词库,相应的存储结构具体可以包括编码字符串和地理词汇。

[0104] 表1

[0105]	编码字符串	地理词汇
	lft	雷峰塔
	zs j	张生记
	tb	淘宝
	wx j	王星记
	zxq	张小泉

[0106] 在本申请的一种应用示例中,假设用户输入的编码字符串为“zs j”,且依据用户所处的地理位置信息确定用户处于杭州市城区,则可以使用杭州市城区的地理词库,检索得到与所述编码字符串相应的地理候选词“张生记”。故相对于用户使用现有技术,需要输入“zhangshengji”才能获得候选词“张生记”,本申请能够有效减轻用户输入的繁琐度。

[0107] 在本申请的一种优选实施例中,可以通过如下步骤构建所述地理区域的地理词库:

[0108] 步骤A1、对地理区域内的地理词汇的使用频率进行统计,收集统计结果符合相应的预设的使用频率条件的地理词汇;

[0109] 步骤A2、将所收集的地理词汇及相应的编码字符串储存至该地理区域的地理词库。

[0110] 通常来说,地理区域内的地理词汇的使用频率越高,则可以认为该地理词汇对应的人物或事物名称因与该地理区域的相关性越高而被广泛使用,甚至在某些情况下该地理词汇对应的人物或事物名称为该地理区域的标志之一。如前述应用场景中“东方新天地”与“北京市东城区”、“张生记”与“杭州市”、“拉市海”与“丽江市”的相关性较高,又如,“金簋簋街”与“北京市簋街”的相关性较高等等。

[0111] 故本优选实施例收集使用频率符合相应的预设的使用频率条件的地理词汇,并保存至相应地理区域的地理词库。这里相应的预设的使用频率条件是指与统计对象(本申请中为地理区域内的地理词汇)、具体采用的统计方法、和/或统计环境等相适应的条件。通常

来说,可以将统计对象的使用次数与统计环境内所有统计对象的使用次数的比值作为相应的使用频率。

[0112] 在本申请的一种优选实施例中,本申请可以提供如下获取该地理区域内的地理词汇的方案:

[0113] 获取方案A1、依据地理位置类别从互联网信息中搜集该地理区域内的地理词汇;和/或

[0114] 获取方案A2、在用户输入触发条件词汇并上屏时,记录上屏的触发条件词汇;所述触发条件词汇为依据地理位置类别预设的该地理区域内的地理词汇。

[0115] 本领域技术人员可以根据实际需要设置所述地理位置类别,如在本申请的一种应用示例中,所述地理位置类别可以包括如下类别中的一项或多项:商圈、餐饮、酒店、超市、电影院、景点、学校、银行和球馆。当然,该应用示例并不作为本申请地理位置类别的应用限制。

[0116] 假设欲获取“中关村”区域内的“餐饮”类别下的地理词汇和相应的地理位置信息,则获取方案A1可以在“大众点评网”和/或“QQ美食”等网站上搜集“中关村”区域内的餐馆,并进行汇总处理;假设欲获取“北京市东城区”内的“商圈”类别下的地理词汇,则获取方案A1可以在与商圈相关的网站或地图搜索服务网站中搜集“北京市东城区”内的商圈,并进行汇总处理等等。

[0117] 在获取方案A2的具体实现中,可以在输入法客户端植入一些记录触发条件来进行记录触发,例如,客户端设置一个用于存储触发条件词汇的模块,当用户输入上屏的词汇命中该模块时,会触发输入法去记录上屏的词汇。

[0118] 需要说明的是,本领域技术人员可以根据实际需求,依据地理位置类别预设该地理区域内的地理词汇作为所述触发条件词汇。本申请对具体的所述触发条件词汇的预设方法不加以限制。

[0119] 获取方案A2依据输入法用户的历史输入词汇获取该地理区域内的地理词汇,使得该地理区域的地理词库中存储的地理词汇能够符合该地理区域的大部分用户的使用习惯,并且,最终向用户推荐的地理候选词也能够符合该地理区域的大部分用户的使用习惯,从而能够进一步提高输入法的智能性。

[0120] 例如,张三、李四、王五、赵六均为牡丹园区域的住户,其中,张三、李四、王五均曾经在牡丹园区域内输入了编码字符串“jbchw”并上屏了“江边城外”,但是赵六的用户词库中没有“江边城外”这个词汇,那么,在赵六在牡丹园区域内输入编码字符串“jbchw”时,本申请也可以向赵六推荐“江边城外”这个候选词。

[0121] 需要说明的是,在使用获取方案A1和/或获取方案A2时,使用频率的统计环境可以包括获取方案A1对应的一项或多项地理位置类别所对应的环境、获取方案A2对应的一项或多项地理位置类别所对应的环境、或者获取方案A1和2对应的一项或多项地理位置类别所对应的环境。

[0122] 例如,使用获取方案A2获取了“江边城外”牡丹园区域内的这个地理词汇,则在统计“江边城外”的使用频率时,可以将输入法用户对“江边城外”的次数作为分子,可以将获取方案A2获取的所有地理位置类别中的一项或多项对应地理词汇的次数作为分母,也可以将获取方案A1和获取方案A2共同获取的所有地理位置类别中的一项或多项对应地理词汇

的次数作为分母,本申请对具体的使用频率的统计环境不加以限制。

[0123] 另外,与采用的统计方式相应,可以采用相应的预设的使用频率条件来判定使用频率的高低,例如,一种预设的使用频率条件的示例为“使用频率大于0.05”等等。

[0124] 在本申请的另一种优选实施例中,所述依据所确定地理区域的地理词库,检索得到与所述编码字符串相应的地理候选词的步骤可以进一步包括:

[0125] 从客户端的所确定地理区域的地理词库中检索得到与所述编码字符串相应的地理候选词;和/或

[0126] 将所述编码字符串和所述地理位置信息发送至服务器端,由服务器端从所确定地理区域的地理词库中检索得到与所述编码字符串相应的地理候选词。

[0127] 在本申请的再一种优选实施例中,所述方法还可以包括:

[0128] 当用户上屏的字符未命中所确定地理区域的地理词库时,分析该未命中的字符的地理位置属性;

[0129] 在分析得到的地理位置属性符合预设的地理位置属性条件时,将该未命中的字符作为地理词汇同步至所有客户端的相应地理区域的地理词库。

[0130] 在此提供一种分析该未命中的字符的地理位置属性的方案:采集该未命中的字符相应的地理位置信息,确定该地理位置信息所对应的地理位置范围,并统计该未命中的字符在相应地理位置范围内出现的频度,如果统计的频度符合预设的地理位置属性条件(例如,大于某个阈值0.2),则可以认为该未命中的字符为地理词汇,于是同步至所有客户端的相应地理区域的地理词库。

[0131] 在具体实现中,可由客户端或服务器端执行所述分析该未命中的字符的地理位置属性的操作。例如,在客户端发现用户上屏的字符未命中所确定地理区域的地理词库时,可以选择自身执行所述分析该未命中的字符的地理位置属性的操作,并且,在分析得到的地理位置属性符合预设的地理位置属性条件时,将该未命中的字符作为地理词汇上传至服务器端,由服务器端执行相应的同步操作;或者,在客户端发现用户上屏的字符未命中所确定地理区域的地理词库时,可以选择将该未命中的字符上传至服务器端,由服务器端执行相应的分析和同步操作。总之,本申请旨在使得所有的客户端使用最新的地理区域的地理词库,使得地理区域的地理词库能够赶上某个地理区域内的人物或事物名称的变化,例如,某个地理区域内新开了一家餐馆“外婆家”,则来该地理区域内的用户都可以享受到与该餐馆相应的地理词汇的推荐服务,让这些用户及时感受到输入法的智能性。

[0132] 在本申请的一种优选实施例中,所述地理区域的地理词库还可用于存储编码字符串、相应地理区域内的地理位置信息和相应的地理词汇;则所述方法还可以包括:

[0133] 依据所确定地理区域的地理词库,检索得到与所述编码字符串和所述地理位置信息相应的地理候选词。

[0134] 参照表2,示出了本申请一种地理区域的地理词库的存储结构示例,其描述的是杭州市区的地理词库,相应的存储结构具体可以包括编码字符串、地理词汇、GPS经度起始、GPS经度截止、GPS纬度起始和GPS纬度截止等字段,其中,GPS经度起始、GPS经度截止、GPS纬度起始和GPS纬度截止组成的GPS经纬度范围用于表示地理位置信息。

[0135] 表2

[0136]

编码字符串	地理词汇	GPS 经度起始	GPS 经度截止	GPS 纬度起始	GPS 纬度截止
lft	雷峰塔	东经 120.193235	东经 120.19435	北纬 30.26123	北纬 30.26234
zsj	张生记	东经 120.194235	东经 120.19535	北纬 30.26323	北纬 30.26434
tb	淘宝	东经 120.196235	东经 120.19735	北纬 30.26623	北纬 30.26734

[0137] 在本申请的一种应用示例中,假设用户输入的编码字符串为“zsj”,该编码字符串对应的地理位置信息为“东经120.19424,北纬30.26383”,且确定该地理位置信息属于杭州市区,则可以使用杭州市区的地理词库,检索得到与所述编码字符串和所述地理位置信息相应的地理候选词的过程具体可以包括:分别将编码字符串“zsj”和地理位置信息“东经120.19424,北纬30.26383”与地理词库中的相应字段进行匹配,其中,地理位置信息的匹配具体可以为:

[0138] 当前地理位置信息经度大于地理词库中的GPS经度起始值

[0139] 并且当前地理位置信息经度小于地理词库中的GPS经度截止值

[0140] 并且当前地理位置信息纬度大于地理词库中的GPS纬度起始值

[0141] 并且当前地理位置信息纬度小于地理词库中的GPS纬度截止值

[0142] 根据上面示例,可以判断当前地理位置信息(东经120.194245,北纬30.26383)在地理词汇“张生记”的经纬度范围内:

[0143] 东经120.19423<东经120.19424<东经120.19535

[0144] 北纬30.26323<北纬30.26383<北纬30.26423

[0145] 由于编码字符串“zsj”也与地理词库中的编码字符串相匹配,于是,地理词汇“张生记”被命中。故相对于用户使用现有技术,需要输入“zhangshengji”才能获得候选词“张生记”,本申请能够有效减轻用户输入的繁琐度。

[0146] 在本申请的一种优选实施例中,可以通过如下步骤构建所述地理区域的地理词库:

[0147] 步骤B1、对地理区域内的地理词汇的使用频率进行统计,收集统计结果符合相应的预设的使用频率条件的地理词汇;

[0148] 步骤B2、将所收集的地理词汇及相应的地理位置信息和编码字符串储存至该地理区域的地理词库。

[0149] 在本申请的一种优选实施例中,本申请可以提供如下获取该地理区域内的地理词汇和相应的地理位置信息的方案:

[0150] 获取方案B1、依据地理位置类别从互联网信息中搜集该地理区域内的地理词汇和相应的地理位置信息;和/或

[0151] 获取方案B2、在用户输入触发条件词汇并上屏时,采集所述触发条件词汇的地理位置信息,记录上屏的触发条件词汇;所述触发条件词汇为依据地理位置类别预设的该地

理区域内的地理词汇。

[0152] 以上对用于针对地理区域存储编码字符串和相应的地理词汇、及用于存储编码字符串、相应地理区域内的地理位置信息和相应的地理词汇的两种地理区域的地理词库进行了介绍；二者的主要不同在于，前者以编码字符串作为索引，后者以编码字符串和地理位置信息作为索引；由于二者在原理上具有相似之处，故相似部分相互参照即可。

[0153] 另外，本领域技术人员可以根据使用两种地理区域的地理词库中的任一，且可以分别设置相应的地理区域的细粒度，本申请对具体的使用方式不加以限制。

[0154] 参照图2，其示出了本申请一种根据地理位置推荐候选词的方法实施例2的流程图，具体可以包括：

[0155] 步骤201、接收用户的编码字符串；

[0156] 步骤202、采集所述编码字符串对应的地理位置信息；

[0157] 步骤203、确定所述地理位置信息所属的地理区域；

[0158] 步骤204、获取所述编码字符串对应的应用程序环境信息；

[0159] 步骤205、在所述应用程序环境信息与地理位置的相关度符合预设的相关度条件时，依据所确定地理区域的地理词库，检索得到与所述编码字符串相应的地理候选词；所述地理区域的地理词库用于针对地理区域存储编码字符串和相应的地理词汇；

[0160] 步骤206、展现所述地理候选词。

[0161] 为了智能地了解用户的输入需求，本申请在适当的场合下才会向用户推荐用户所处地理区域内的地理候选词，具体而言，在所述应用程序环境信息与地理位置的相关度符合预设的相关度条件时，才会激活输入法的向用户推荐用户所处地理区域内的地理候选词的功能；在所述应用程序环境信息与地理位置的相关度不符合预设的相关度条件时，则不会激活上述推荐功能。

[0162] 宿主程序和寄宿程序为本技术领域内的概念，在此进行原理性的说明。其中，宿主程序可用于在计算机环境下，提供寄宿程序赖以生存的软件环境也即宿主环境(host environment)，同时，宿主程序还可以动态加载寄宿程序提供的DLL(动态链接库,Dynamic Link Library)，以动态加载外部功能。在实际中，寄宿程序提供的DLL可以包括输入法提供的输入法编辑器链接库(如存放在c:\windows\system32下的SogouPy.ime)等等。

[0163] 在Windows、Unix、Linux等操作系统中，作为宿主程序，每个应用程序都可以选择是否打开输入法，并在打开输入法时向输入法传递一个程序句柄；对于输入法而言，可以根据该程序句柄获取当前应用程序的应用程序环境信息。例如，输入法在运行中，调用GetModuleFilename发现程序路径名“C:\ProgramFiles\MicrosoftOffice\OFFICE11\WINWORD.EXE”，即可确定对应的应用程序环境信息为“WinWord.exe”，也即当前编码串是在word中输入的。

[0164] 在实际应用中，本领域技术人员可以收集一些需要使用输入法的应用程序，并预置应用程序与地理位置的相关度。这里，应用程序与地理位置的相关度主要用于表示该应用程序中的文字输入与地理位置有关的可能性。如果应用程序与地理位置的相关度较低，则认为该应用程序中的文字输入与地理位置的有关的的可能性较低，从而认为此时用户对地理词汇的输入需求的可能性较小，故此时向用户推荐用户所处地理区域内的地理词汇的意义不大；反之，如果应用程序与地理位置的相关度较高，则认为该应用程序中的文字输入与

地理位置的有关的可能性较高,从而认为此时用户对地理词汇的输入需求的可能性较大,故此向用户推荐用户所处地理区域内的地理词汇能够在输入过程中出现用户想要的候选词,从而能够提高输入法的智能性。

[0165] 关于如何预置应用程序与地理位置的相关度,在具体实现中,可以对需要使用输入法的应用程序进行分类。例如,在本申请的一种应用示例中,需要使用输入法的应用程序的类别具体可以包括:地图应用、周边搜索、手机浏览器、普通搜索、微博、购物类、即时通讯窗口、游戏、操作系统工具等等。

[0166] 参照表3,示出了本申请一种需要使用输入法的应用程序的类别与相应的地理位置的相关度的示例。

[0167] 表3

[0168]	需要使用输入法的应用程序的类别与 地理位置的相关度	需要使用输入法的应用程序的类别
	高	地图应用、周边搜索
	中	机浏览器、普通搜索、微博、购物类、 即时通讯窗口
	低	游戏、操作系统工具

[0169] 在实际应用中,可以采用预设的相关度条件来判定应用程序与地理位置的相关度的高低。所述预设的相关度条件可以用量化条件,也可以用等级条件。例如,一种量化条件的示例为“应用程序与地理位置的相关度大于0.3”;又如,一种等级条件的示例为“应用程序与地理位置的相关度为中或高”。

[0170] 综上,本实施例只有在所述应用程序环境信息与地理位置的相关度符合预设的相关度条件时,才认为用户存在对地理候选词的输入需求,才会激活输入法的向用户推荐用户所处地理区域内的地理候选词的功能;相对于现有技术,本申请增加了用户输入需求的判断,且能够依据地理位置信息推荐用户所处地理区域内的地理候选词,也即能够在输入过程中出现用户想要的候选词,能够提高输入法的智能性。

[0171] 参照图3,其示出了本申请一种根据地理位置推荐候选词的方法实施例3的流程图,具体可以包括:

[0172] 步骤301、接收用户的编码字符串;

[0173] 步骤302、采集所述编码字符串对应的地理位置信息;

[0174] 步骤303、确定所述地理位置信息所属的地理区域;

[0175] 步骤304、依据所确定地理区域的地理词库,检索得到与所述编码字符串相应的地理候选词;所述地理区域的地理词库用于存储编码字符串和相应的地理词汇;

[0176] 步骤305、依据所述编码字符串,在普通词库中进行检索,得到相应的普通候选词;

[0177] 步骤306、在展现所述普通候选词的同时,展现所述地理候选词。

[0178] 相对于实施例1,实施例3可以在展现所述普通候选词的同时,展现所述地理候选词。

[0179] 在本申请的另一优选实施例中,所述展现所述地理候选词的步骤可以进一步包

括：

[0180] 依据候选词的排序因子,对所述普通候选词和地理候选词进行排序;所述排序因子具体可以包括如下因子中的一项或多项:候选词与所述地理位置信息的匹配度、候选词的使用频率和搜索排序策略;

[0181] 依据排序结果,对所述普通候选词和地理候选词进行展现。

[0182] 候选词是依据地理词库或普通词库检索得到的,对于依据地理词库检索得到的地理候选词,其与所述地理位置信息的匹配度就是地理词库中存储的地理位置信息与所述地理位置信息的匹配度(一般为80%-100%之间),对于依据普通词库检索得到的普通候选词,可以认为其与所述地理位置信息的匹配度为0。

[0183] 如果在联网或者连接服务器输入,客户端或服务器端可以将这些候选词输入至搜索引擎中进行搜索,把搜索结果数目多的候选词排到前面,这也是搜索排序策略的应用。

[0184] 由于候选词的使用频率作为排序因子为本技术领域内的公知技术,故在此不作赘述。

[0185] 在本申请的一种优选实施例中,在能够获取所述编码字符串对应的应用程序环境信息以及所述应用程序环境信息与地理位置的相关度的情况下,所述排序因子还可以包括所述应用程序环境信息与地理位置的相关度。

[0186] 需要说明的是,在使用多种排序因子时,可以分别赋予每种排序因子一定的权重,然后依据这多种排序因子的加权结果对各候选词进行排序。

[0187] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述的部分,可以参见其他实施例的相关描述即可。

[0188] 并且,本领域技术人员易于想到的是:上述实施例1—实施例3任意组合应用都是可行的,故上述实施例1—实施例3之间的任意组合都是本发明的实施方案,但是由于篇幅限制,本说明书在此就不一一详述了。

[0189] 为使本领域技术人员更好地理解本申请,以下通过一个具体的例子来说明本申请一种根据地理位置推荐候选词的方法流程,该方法流程具体可以包括:

[0190] 步骤S1、接收用户的编码字符串;

[0191] 步骤S2、采集所述编码字符串对应的地理位置信息;

[0192] 随着智能手机、平板电脑的逐步普及,移动设备的操作系统一般均具备获取所在地理位置的API(应用程序编程接口,Application Programming Interface),本申请可以直接使用常见操作系统提供的获取地理位置功能。

[0193] 例如,本申请的一种应用示例中,操作系统可以首先判断GPS模块是否存在或者是否开启,如果开启正常,则会直接开启输入法的向用户推荐用户所处地理区域内的地理候选词的功能并继续获取当前地理位置信息,如果开启不正常,则会返回正常的输入法功能。

[0194] 其中,获取当前地理位置信息的操作具体可以包括:打开操作系统的位置管理服务,对所述位置管理服务进行初始化,获得GPS位置信息,通过间隔时间(如10秒)的监听,不断更新以获得当前最准确的GPS位置信息。

[0195] 步骤S3、确定所述地理位置信息所属的地理区域;

[0196] 步骤S4、获取所述编码字符串对应的应用程序环境信息;

[0197] 步骤S5、判断所述应用程序环境信息与地理位置的相关度是否符合预设的相关度

条件,若是,则执行步骤S6,否则执行步骤S8;

[0198] 步骤S6、依据所确定地理区域的地理词库,检索得到与所述编码字符串和所述地理位置信息相应的地理候选词;

[0199] 步骤S7、展现所述地理候选词;

[0200] 步骤S8、依据所述编码字符串,在普通词库中进行检索,得到相应的普通候选词;

[0201] 步骤S9、展现所述普通候选词。

[0202] 与前述方法实施例相应,本申请还提供了一种根据地理位置推荐候选词的装置实施例,参照图4所示的结构图,具体可以包括:

[0203] 接口模块401,用于接收用户的编码字符串;

[0204] 采集模块402,用于采集所述编码字符串对应的地理位置信息;

[0205] 地理区域确定模块403,用于确定所述地理位置信息所属的地理区域;

[0206] 第一检索模块404,用于依据所确定地理区域的地理词库,检索得到与所述编码字符串相应的地理候选词;所述地理区域的地理词库用于针对地理区域存储编码字符串和相应的地理词汇;及

[0207] 展现模块405,用于展现所述地理候选词。

[0208] 在本申请的一种优选实施例中,所述装置还可以包括:用于构建所述地理区域的地理词库的地理词库构建模块,所述地理词库构建模块可以进一步包括:

[0209] 统计子模块,用于对地理区域内的地理词汇的使用频率进行统计;

[0210] 收集子模块,用于收集统计结果符合相应的预设的使用频率条件的地理词汇;及

[0211] 储存子模块,用于将所收集的地理词汇及相应的地理位置信息和编码字符串储存至该地理区域的地理词库。

[0212] 在本申请的一种优选实施例中,所述装置还可以包括:用于获取该地理区域内的地理词汇的第一获取模块,所述第一获取模块可以进一步包括:

[0213] 互联网搜集子模块,用于依据地理位置类别从互联网信息中搜集该地理区域内的地理词汇;和/或

[0214] 触发获取子模块,用于在用户输入触发条件词汇并上屏时,记录上屏的触发条件词汇;所述触发条件词汇为依据地理位置类别预设的该地理区域内的地理词汇。

[0215] 在本申请的一种优选实施例中,所述地理区域的地理词库还用于存储编码字符串、相应地理区域内的地理位置信息和相应的地理词汇;则所述装置还可以包括:

[0216] 第二检索模块,用于依据所确定地理区域的地理词库,检索得到与所述编码字符串和所述地理位置信息相应的地理候选词。

[0217] 在本申请的一种优选实施例中,所述装置还可以包括:

[0218] 应用程序环境获取模块,用于获取所述编码字符串对应的应用程序环境信息;

[0219] 则所述第一检索模块404,可具体用于在所述应用程序环境信息与地理位置的相关度符合预设的相关度条件时,依据所确定地理区域的地理词库,检索得到与所述编码字符串和所述地理位置信息相应的地理候选词。

[0220] 在本申请的一种优选实施例中,所述装置还可以包括:

[0221] 第三检索模块,用于依据所述编码字符串,在普通词库中进行检索,得到相应的普通候选词;

[0222] 则所述展现模块405,可具体用于在展现所述普通候选词的同时,展现所述地理候选词。

[0223] 在本申请的一种优选实施例中,所述装置还可以包括:排序子模块,用于依据候选词的排序因子,对所述普通候选词和地理候选词进行排序;所述排序因子具体可以包括如下因子中的一项或多项:候选词与所述地理位置信息的匹配度、候选词的使用频率和搜索排序策略;

[0224] 次序展现子模块,用于依据排序结果,对所述普通候选词和地理候选词进行展现。

[0225] 在本申请的一种优选实施例中,所述排序因子还可以包括所述应用程序环境信息与地理位置的相关度。

[0226] 在本申请的一种优选实施例中,所述第一检索模块可以进一步包括:

[0227] 客户端检索子模块,用于从客户端的所确定地理区域的地理词库中检索得到与所述编码字符串和所述地理位置信息相应的地理候选词;和/或

[0228] 服务器端检索子模块,用于将所述编码字符串和所述地理位置信息发送至服务器端,由服务器端从所确定地理区域的地理词库中检索得到与所述编码字符串和所述地理位置信息相应的地理候选词。

[0229] 在本申请的一种优选实施例中,所述装置还可以包括:

[0230] 分析模块,用于当用户上屏的字符未命中所确定地理区域的地理词库时,分析该未命中的字符的地理位置属性;

[0231] 同步模块,用于在分析得到的地理位置属性符合预设的地理位置属性条件时,将该未命中的字符作为地理词汇同步至所有客户端的相应地理区域的地理词库。

[0232] 在本申请的一种优选实施例中,所述采集模块402,可具体用于依据用户的IP地址、移动设备GPS或移动网络采集所述编码字符串对应的地理位置信息。

[0233] 对于装置实施例而言,由于其与方法实施例基本相似,所以描述的比较简单,相关之处参见方法实施例的部分说明即可。

[0234] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0235] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0236] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0237] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0238] 尽管已描述了本申请的优选实施例,但本领域内的技术人员一旦得知了基本创造性概念,则可对这些实施例做出另外的变更和修改。所以,所附权利要求意欲解释为包括优选实施例以及落入本申请范围的所有变更和修改。

[0239] 本说明书中的各个实施例均采用递进的方式描述,每个实施例重点说明的都是与其他实施例的不同之处,各个实施例之间相同相似的部分互相参见即可。

[0240] 以上对本申请所提供的一种根据地理位置推荐候选词的方法和装置,进行了详细介绍,本文中应用了具体个例对本申请的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本申请的方法及其核心思想;同时,对于本领域的一般技术人员,依据本申请的思想,在具体实施方式及应用范围上均会有改变之处,综上所述,本说明书内容不应理解为对本申请的限制。

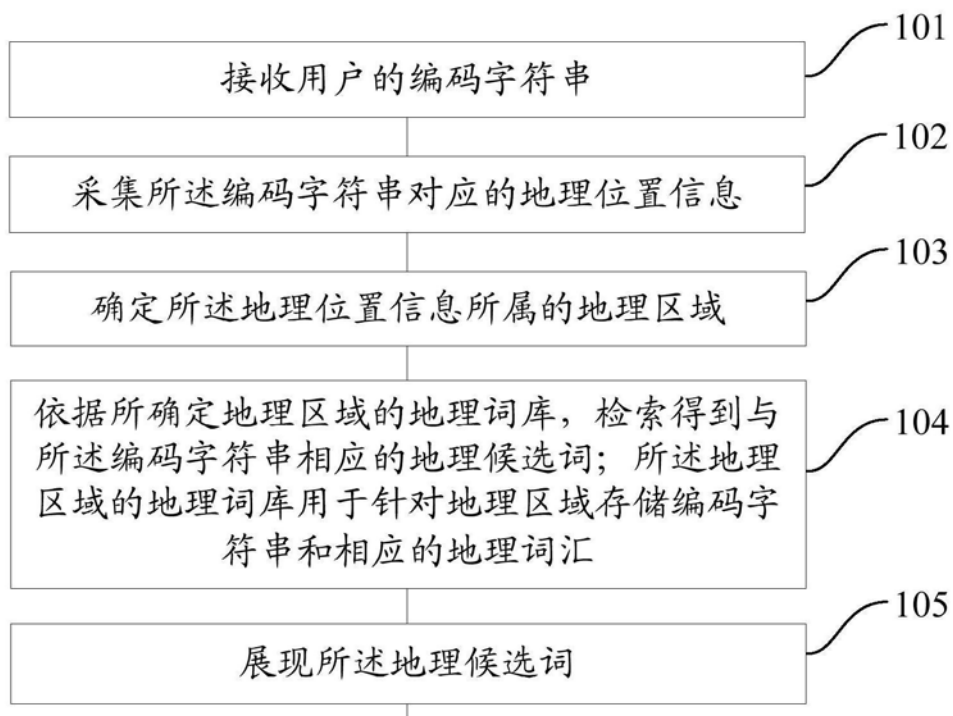


图1

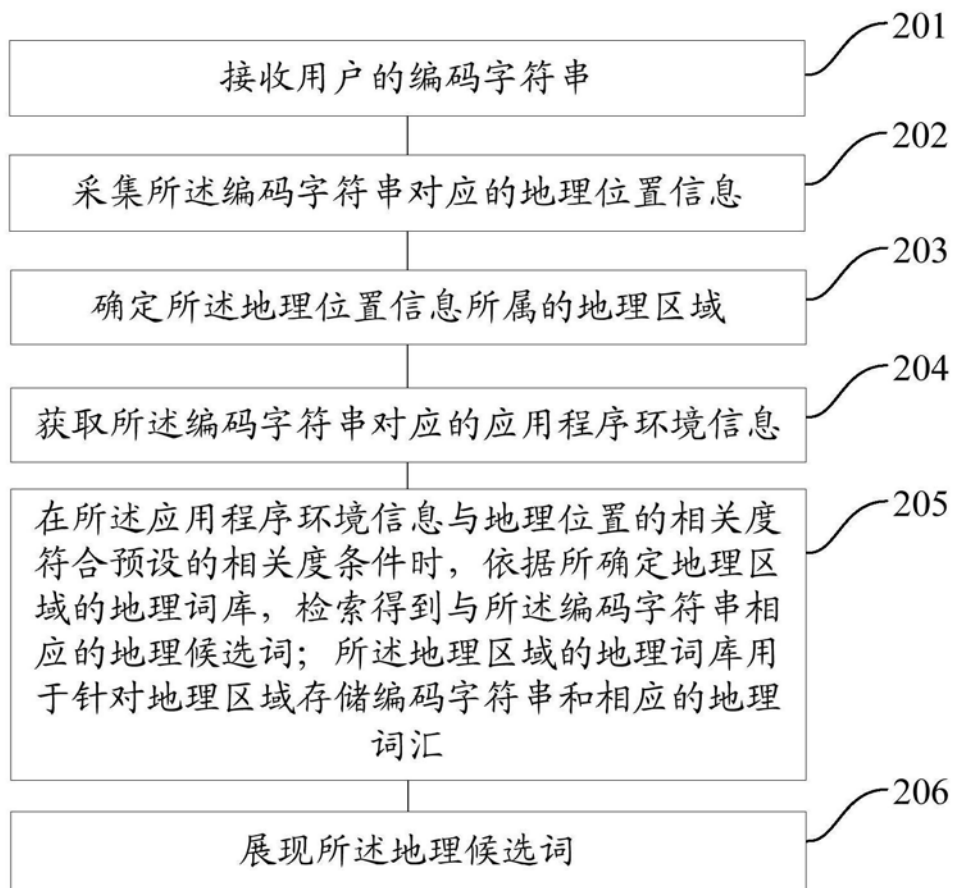


图2

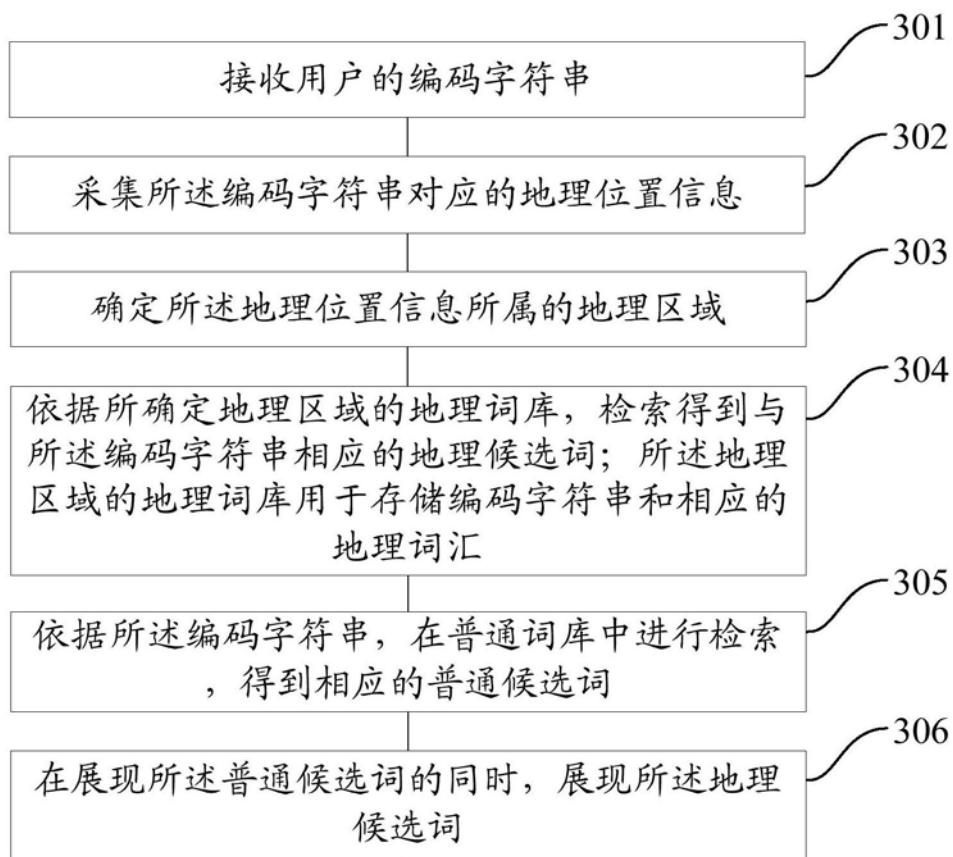


图3

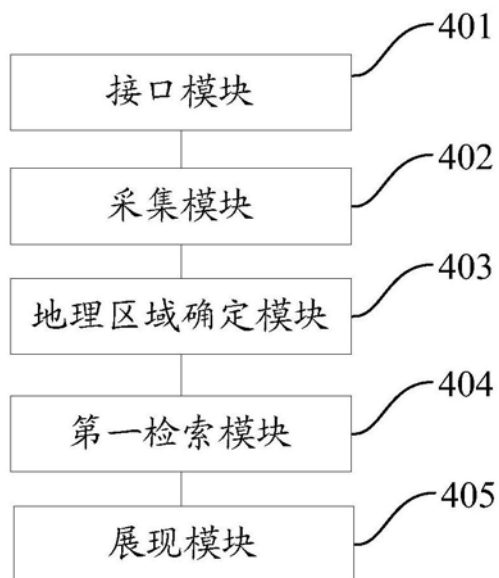


图4