

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2014 年 7 月 3 日 (03.07.2014)



(10) 国际公布号
WO 2014/101577 A1

- (51) 国际专利分类号:
G06F 3/023 (2006.01)
- (21) 国际申请号: PCT/CN2013/086968
- (22) 国际申请日: 2013 年 11 月 12 日 (12.11.2013)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201210592774.9 2012 年 12 月 31 日 (31.12.2012) CN
- (71) 申请人: 百度在线网络技术(北京)有限公司
(BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) [CN/CN]; 中国北京市海淀区上地十街 10 号百度大厦, Beijing 100085 (CN)。
- (72) 发明人: 陆阳阳 (LU, Yangyang); 中国北京市海淀区上地十街 10 号百度大厦, Beijing 100085 (CN)。
孟可丰 (MENG, Kefeng); 中国北京市海淀区上地十街 10 号百度大厦, Beijing 100085 (CN)。
- (74) 代理人: 北京汉昊知识产权代理事务所(普通合伙) (HANHOW INTELLECTUAL PROPERTY); 中国北京市东长安街 1 号东方广场西 1 区 1111 室, Beijing 100738 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。
- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO,

[见续页]

(54) Title: METHOD AND DEVICE USED FOR PROVIDING INPUT CANDIDATE ITEM CORRESPONDING TO INPUT CHARACTER STRING

(54) 发明名称: 一种用于提供输入字符串所对应的输入候选项的方法与设备

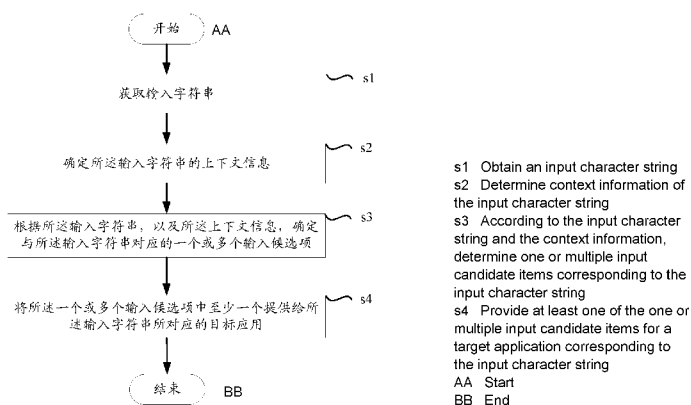


图 4 / Fig.4

(57) Abstract: A method and a device used for providing an input candidate item corresponding to an input character string. The method comprises: an input device obtaining an input character string; determining context information of the input character string; according to the input character string and the context information, determining one or multiple input candidate items corresponding to the input character string; and providing at least one of the one or multiple input candidate items for a target application corresponding to the input character string. Compared with the prior art, the method and the device determines the one or multiple input candidate items corresponding to the input character string according to the context information of the input character string, and provides at least one of the one or multiple input candidate items for the target application corresponding to the input character string, which therefore improves input flexibility, matches users' input requirement, improves input efficiency of an input method, and improves user experience.

(57) 摘要:

[见续页]



WO 2014/101577 A1



RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, **本国际公布:**

CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, — 包括国际检索报告(条约第 21 条(3))。
TG)。

一种用于提供输入字符串所对应的输入候选项的方法与设备。该方法包括：输入设备获取输入字符串；确定所述输入字符串的上下文信息；根据所述输入字符串，以及所述上下文信息，确定与所述输入字符串对应的一个或多个输入候选项；将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。与现有技术相比，该方法和设备通过根据输入字符串的上下文信息，确定与所述输入字符串对应的一个或多个输入候选项；并将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用；从而提高了输入灵活度，与用户的输入需求相匹配，还提高了输入法的输入效率，改善了用户体验。

一种用于提供输入字符串所对应的输入候选项的方法与设备

技术领域

5 本发明涉及互联网技术领域，尤其涉及一种用于提供输入字符串所对应的输入候选项的技术。

背景技术

10 对于中文输入而言，存在大量的同音异义词，而实际用户在输入过程中仅会选择其中的一个，当前的输入法将多个输入候选项提供给用户，由用户根据需求进行选择，当输入候选项较多时，用户进行选择操作则会导致输入的缓慢低效。考虑到文字的输入间具有一定的语义或逻辑关系，如何利用所述语义或逻辑关系，对输入候选项进行处理，从而提升用户的使用体验，成为本领域技术人员亟需解决的一个
15 问题。

发明内容

 本发明的目的是提供一种用于提供输入字符串所对应的输入候选项的方法与设备。

20 根据本发明的一个方面，提供了一种用于提供输入字符串所对应的输入候选项的方法，其中，该方法包括以下步骤：

- a 获取输入字符串；
- b 确定所述输入字符串的上下文信息；
- c 根据所述输入字符串，以及所述上下文信息，确定与所述输入
25 字符串对应的一个或多个输入候选项；
- d 将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。

 根据本发明的另一方面，还提供了一种用于提供输入字符串所对

应的输入候选项的输入设备，其中，该设备包括：

获取装置，用于获取输入字符串；

上下文确定装置，用于确定所述输入字符串的上下文信息；

候选项确定装置，用于根据所述输入字符串，以及所述上下文信息，确定与所述输入字符串对应的一个或多个输入候选项；

提供装置，用于将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。

与现有技术相比，本发明通过根据输入字符串的上下文信息，确定与所述输入字符串对应的一个或多个输入候选项；并将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用；从而提高了输入灵活度，与用户的输入需求相匹配，还提高了输入法的输入效率，改善了用户体验。

而且，本发明还可以通过对与所述输入字符串对应的一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选项；进一步地，还可以根据所述初步输入候选项与所述上下文信息对应的搭配相关度信息，对所述一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选项；更进一步地，还可以根据对文本信息的分析处理，获得搭配频度满足一定阈值的分词搭配关系，从而建立或更新所述词语搭配数据库；从而提高了输入字符串与输入候选项的匹配准确度，与用户的输入需求相匹配，提高了输入法的输入效率，改善了用户体验。

而且，本发明还可以确定包含所述输入字符串的增补字符串，并根据所述增补字符串，获得与所述增补字符串对应的一个或多个增补输入候选项，通过对所述增补输入候选项进行筛选处理，以获得所述一个或多个输入候选项；从而提高了输入灵活度，提高了输入字符串与输入候选项的匹配准确度，与用户的输入需求相匹配，提高了输入法的输入效率，改善了用户体验。

而且，本发明还可以所述输入字符串所对应的相关文本进行自然语言处理，从所述相关文本提取对应的焦点词语，以作为所述上下文；从而提高了输入字符串与输入候选项的匹配准确度，与用户的输入需求

相匹配，提高了输入法的输入效率，改善了用户体验。

而且，本发明还可以检测所述输入字符串是否超出预定的长度阈值，若所述输入字符串超出所述长度阈值，根据所述输入字符串，以及所述上下文信息，迭代确定与所述输入字符串对应的一个或多个输入候
5 选项；进一步地，通过检测所述输入字符串在输入词库中对应的一个或多个抽样输入候选项中至少一个中的字数是超出预定的字数阈值，以确定所述输入字符串是否超出预定的长度阈值；进一步地，当所述输入字符串超出所述长度阈值，还可以确定部分输入候选项与其余输入候选项，并根据所述一个或多个部分输入候选项，以及所述一个或多个其余
10 输入候选项，确定所述一个或多个输入候选项；从而提高了输入灵活度，提高了输入字符串与输入候选项的匹配准确度，与用户的输入需求相匹配，提高了输入法的输入效率，改善了用户体验。

附图说明

15 通过阅读参照以下附图所作的对非限制性实施例所作的详细描述，本发明的其它特征、目的和优点将会变得更明显：

图1示出根据本发明一个方面的一种用于提供输入字符串所对应的输入候选项的输入设备示意图；

图2示出根据本发明一个优选实施例的一种用于提供输入字符串
20 所对应的输入候选项的输入设备示意图；

图3示出根据本发明另一个优选实施例的一种用于提供输入字符串所对应的输入候选项的输入设备示意图；

图4示出根据本发明另一个方面的一种用于提供输入字符串所对应的输入候选项的方法流程图；

25 图5示出根据本发明一个优选实施例的一种用于提供输入字符串所对应的输入候选项的方法流程图；

图6示出根据本发明另一个优选实施例的一种用于提供输入字符串所对应的输入候选项的方法流程图。

附图中相同或相似的附图标记代表相同或相似的部件。

具体实施方式

下面结合附图对本发明作进一步详细描述。

图 1 示出根据本发明一个方面的一种用于提供输入字符串所对应的输入候选项的输入设备示意图；其中，所述输入设备包括获取装置 11、上下文确定装置 12、候选项确定装置 13、提供装置 14。具体地，获取装置 11 获取输入字符串；上下文确定装置 12 确定所述输入字符串的上下文信息；候选项确定装置 13 根据所述输入字符串，以及所述上下文信息，确定与所述输入字符串对应的一个或多个输入候选项；提供装置 14 将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。在此，输入设备包括但不限于网络设备、用户设备、或网络设备与用户设备通过网络相集成所构成的设备。其中，所述网络设备其包括但不限于计算机、网络主机、单个网络服务器、多个网络服务器集或多个服务器构成的云；在此，云基于云计算 (Cloud Computing) 的大量计算机或网络服务器构成，其中，云计算是分布式计算的一种，由一群松散耦合的计算机集组成的一个虚拟超级计算机。所述用户设备其包括但不限于任何一种可与用户通过键盘、遥控器、触摸板、或声控设备进行人机交互的电子产品，例如计算机、智能手机、PDA、游戏机、或 IPTV 等。所述网络包括但不限于互联网、广域网、城域网、局域网、VPN 网络、无线自组织网络 (Ad Hoc 网络) 等。本领域技术人员应能理解，其他的输入设备同样适用于本发明，也应包含在本发明保护范围以内，并在此以引用方式包含于此。

上述各装置之间是持续不断工作的，在此，本领域技术人员应理解“持续”是指上述各装置分别实时地或者按照设定的或实时调整的工作模式要求，进行输入字符串的获取、上下文信息的确定、输入候选项的确定、输入候选项的提供等，直至输入设备停止获取输入字符串。

获取装置 11 获取输入字符串。具体地，所述获取装置 11 通过从

输入框中接收用户或其他应用等通过输入、选择、点击等操作所输入的字符串；或者根据各种通信协议(Communications Protocol)，通过各种数据传输接口，获取从其他应用的输入框中的字符串；其中，所述输入框如短信编辑框、内容输入框等。其中，所述输入字符串包括但不限于通过如全拼或简拼等方式所输入的拼音编码字符等。

上下文确定装置 12 确定所述输入字符串的上下文信息。具体地，所述上下文确定装置 12 通过根据获取装置 11 所获取的输入字符串，利用当前所输入的字符串所属的文字等，将所述当前输入字符串所属的文字与一个或多个词库进行匹配，获取与所述输入字符串相关联的一种或多种匹配信息，并作为所述输入字符串的上下文信息；例如，当输入“huacao”时，先确定所述输入字符串所属的文字可能为“花草”，根据“花草”在所属词库中的匹配信息，确定其下文信息可能为“茶”等；或者，所述上下文确定装置 12 通过根据与所述输入字符串相对应的历史记录信息，对所述输入字符串的上下文信息进行确定，例如，所述输入字符串前刚刚上屏的前一个输入字符串，则可作为所述输入字符串的前文信息。在此，所述上下文信息包括但不限于与所述输入字符串相对应的上文信息或下文信息；例如，当所述输入为“shishi”时，当前次输入所上屏的词组为“客观”时，则所述“shishi”对应的上文信息为“客观”；若屏幕上存在一个词组“计划”，且输入“shishi”时，光标位于“计划”前，则其所对应的下文信息为“计划”等。

候选项确定装置 13 根据所述输入字符串，以及所述上下文信息，确定与所述输入字符串对应的一个或多个输入候选项。具体地，候选项确定装置 13 通过根据获取装置 11 所获取的输入字符串，以及上下文确定装置 12 所确定的上下文信息，通过将与所述输入字符串相对应的一个或多个上下文信息直接作为一个或多个输入候选项，或者将所述输入字符串相对应的上下文信息进行处理后作为输入候选项的方式，确定与所述输入字符串对应的一个或多个输入候选项。例如，继上例，当所述输入字符串为“shishi”时，其所对应的上文信息可

能为“客观”，“我”，“谈论”等，当所述上文信息为“客观”时，通过利用将所述输入字符串，以及所述上下文信息进行匹配等方式，确定输入候选项为“事实”、“实时”、“实施”等；当所述上文信息为“我”时，确定输入候选项为“试试”、“实施”、“试试”等；当所述
5 上文信息为“谈论”时，确定输入候选项为“时事”、“事实”、“世事”等。

提供装置 14 将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。具体地，提供装置 14 通过直接提供，或根据各种通信协议(Communications Protocol)，通过各种数据传输
10 接口，通过例如重新排序或筛选等方式，将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。其中，所述目标应用包括但不限于与所述获取装置 11 所对应的输入框相对应的相同应用，或者与所述获取装置 11 所对应的输入框相关联的其他应用等，如直接利用所述输入候选项进行搜索等。

15 优选地，所述候选项确定装置 13 还可以根据所述输入字符串，以及所述上下文信息，确定包含所述输入字符串的增补字符串，其中，所述增补字符串还包括所述上下文信息所对应的字符串；根据所述增补字符串，在输入词库中进行匹配查询，以获得与所述增补字符串对应的一个或多个增补输入候选项；根据所述上下文信息，对所述一个或多个增
20 补输入候选项进行筛选处理，以获得所述一个或多个输入候选项。具体地，所述候选项定装置 13 还可以根据所述输入字符串，以及所述上下文信息，通过将所述输入字符串结合所述上下文信息与例如输入词库进行匹配等方式，确定包含所述输入字符串的增补字符串，例如，当输入字符串为“xuxu”时，当上下文信息为“形态”，则确定所述增补字符串为“rusheng”；或者，当输入字符串为“tianan”，上下文信息为“北
25 京”，则确定所述增补字符串为“men”，在此，所述增补字符串包括但不限于通过匹配等方式，确定一个或多个通过拼接而成的多段字符串。所述候选项定装置 13 通过将所述增补字符串在输入词库中进行匹配查询，以获得与所述增补字符串对应的一个或多个增补输入候选项；继前

例，所述增补字符串为“rusheng”时所对应的增补输入候选项如“如生”、“入声”、“入胜”等，所述增补字符串为“men”时，所对应的增补输入候选项如“天安门”、“门”等。或者，所述增补输入候选项还可以直接包括所述上下文信息，如直接生成“形态栩栩如生”、“北京天安门”等。根据所述上下文信息，对所述一个或多个增补输入候选项进行筛选处理，继前例，根据上下文信息，分别确定输入候选项为“栩栩如生”、“天安门”从而获得所述一个或多个输入候选项；在此，若所述增补输入候选项中包含上下文信息，则将所述增补输入候选项中的上下文信息删除后，获得所述一个或多个输入候选项，如所述增补输入候选项“北京天安门”中的“北京”删除，获得所述一个或多个输入候选项“天安门”。

优选地，所述上下文确定装置 12 还可以通过对所述输入字符串所对应的相关文本进行自然语言处理，从所述相关文本提取对应的焦点词语，以作为所述上下文，其中，所述焦点词语邻近所述输入字符串所对应的输入位置。具体地，所述上下文确定装置 12 还可以通过分词或语义分析等方式，对所述输入字符串所对应的相关文本进行自然语言处理，其中，所述相关文本包括但不限于所述输入字符串本身所对应的文本，或者所述输入字符串所对应的输入位置前后已经上屏的相关文本；通过对所述相关文本中经由自然语言处理，提取对应的焦点词语以作为所述上下文。例如，当所述输入字符串为“shishi”时，其对应的相关文本为“上海地铁一期工程”，且所述输入字符串位于“上海地铁一期工程”之后，通过对所述相关文本进行自然语言处理，提取焦点词语为“工程”，并将“工程”作为上下文。进一步地，因此，当输入字符串为“shishi”时，该输入设备根据所述上下文“工程”，确定输入候选项为“实施”、“实时”、“试试”等。

图 2 示出根据本发明一个优选实施例的一种用于提供输入字符串所对应的输入候选项的输入设备示意图；其中，所述输入设备包括获取装置 11'、上下文确定装置 12'、候选项确定装置 13'、提供装置 14'，在此，所述候选项确定装置 13'包括匹配单元 131'和处理单元 132'。

具体地，获取装置 11' 获取输入字符串；上下文确定装置 12' 确定所述输入字符串的上下文信息；匹配单元 131' 根据所述输入字符串，在输入词库中进行匹配查询，以获得与所述输入字符串对应的一个或多个初步输入候选项；处理单元 132' 根据所述上下文信息，对所述一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选项；提供装置 14' 将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。其中，所述获取装置 11'、上下文确定装置 12'、提供装置 14' 与图 1 所示对应装置相同或基本相同，故此处不再赘述，并通过引用的方式包含于此。

上述各装置之间是持续不断工作的，在此，本领域技术人员应理解“持续”是指上述各装置分别实时地或者按照设定的或实时调整的工作模式要求，进行输入字符串的获取、上下文信息的确定、初步输入候选项的获取、输入候选项的确定、输入候选项的提供等，直至输入设备停止获取输入字符串。

匹配单元 131' 根据所述输入字符串，在输入词库中进行匹配查询，以获得与所述输入字符串对应的一个或多个初步输入候选项。具体地，所述匹配单元 131' 通过获取所述输入字符串，将所述输入字符串与输入词库中的词进行匹配，以获得与所述输入字符串对应的一个或多个初步输入候选项。例如，当输入字符串为“shishi”时，根据在输入词库中进行匹配查询后，获取“实时”、“试试”、“实施”、“事实”、“适时”、“逝世”等多个初步输入候选项。

处理单元 132' 根据所述上下文信息，对所述一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选项。具体地，处理单元 132' 根据所述输入字符串的上下文信息，通过根据自然语言模型或词语搭配数据库等，确定上下文信息与初步输入候选项的搭配概率，并根据所述搭配概率对所述初步输入候选项进行筛选或排序，以获得输入候选项。例如，继前例，当上下文信息为“客观”，则经处理后，将初步输入候选项进行重新排序，顺序变为“事实”、“实时”、“试试”、“实施”、“适时”、“逝世”等。

优选地，所述处理单元 132'还可以根据所述上下文信息，在词语搭配数据库中进行匹配查询，以确定所述初步输入候选项与所述上下文信息对应的搭配相关度信息；根据所述搭配相关度信息，对所述一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选项。具体地，所述处理单元 132'还可以根据所述上下文信息，在词语搭配数据库中进行匹配查询，通过根据所述词语搭配数据库中的一个或多个词条与所述上下文的匹配关系等，确定所述初步输入候选项与所述上下文信息对应的搭配相关度信息；例如若或为 bull 变量，则所述搭配相关度信息表示如是否存在搭配；或为连续变量，如所述搭配相关度信息表示搭配概率为多少；在此，所述搭配相关度信息可根据如词语搭配数据库或其他相关数据库中的上下文与初步输入候选项的组合频率等通过机器学习获得。根据所述搭配相关度信息，通过对所述一个或多个初步输入候选项进行如筛选或排序处理，获得所述一个或多个输入候选项。

更优选地，所述输入设备还包括文本处理装置（未示出）和建立装置（未示出）；其中，文本处理装置对一个或多个文本信息进行分析处理，以获得搭配频度满足一定阈值的分词搭配关系；建立装置根据所述分词搭配关系，建立或更新所述词语搭配数据库。具体地，文本处理装置例如通过对各类日志、各类文章等进行获取，并对所获取文本信息进行如先分词，后统计相邻词间的共现频率等分析方法，当所述共现频率超过一定阈值时，则判断所述相邻词间存在一定的搭配关系，从而获得搭配频度满足一定阈值的分词搭配关系；建立装置根据所述分词搭配关系，建立所述词语搭配数据库，或对现有词语搭配数据库进行更新等，在此，所述词语搭配数据库中包括但不限于分词搭配关系以及分词搭配关系之间的搭配概率信息；并且，所述词语搭配数据库中可包含现有固定词语，如成语、俗语等，同时经过所述更新，可以将由于现有网络应用的快速发展所产生的新的词语及表述进行包含，从而有效提高用户的输入效率。

图 3 示出根据本发明另一个优选实施例的一种用于提供输入字符串所对应的输入候选项的输入设备示意图；其中，所述输入设备包括

获取装置 11”、上下文确定装置 12”、候选项确定装置 13”、提供装置 14”，在此，所述候选项确定装置 13”包括检测单元 133”和迭代单元 134”。具体地，获取装置 11”获取输入字符串；上下文确定装置 12”确定所述输入字符串的上下文信息；检测单元 133”检测所述输入字符串是否超出预定的长度阈值；若所述输入字符串超出所述长度阈值，迭代单元 134”根据所述输入字符串，以及所述上下文信息，迭代确定与
5 所述输入字符串对应的一个或多个输入候选项；提供装置 14”将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。其中，所述获取装置 11”、上下文确定装置 12”、提供装置 14”与图 1 所示对应装置相同或基本相同，故此处不再赘述，并通过引用的方式包含于此。

上述各装置之间是持续不断工作的，在此，本领域技术人员应理解“持续”是指上述各装置分别实时地或者按照设定的或实时调整的工作模式要求，进行输入字符串的获取、上下文信息的确定、长度阈值的检测、输入候选项的确定、输入候选项的提供等，直至输入设备
15 停止获取输入字符串。

检测单元 133”检测所述输入字符串是否超出预定的长度阈值。具体地，所述检测单元 133”通过基于预置的或用户自行设置的长度阈值，对所述输入字符串的长度进行检测，检测所述输入字符串的长度是否超
20 出了预定的长度阈值。例如，当预定的字符长度阈值为 N，当所述输入字符串的长度为 N+1 时，则检测单元 133”检测所述输入字符串超出了预定的长度阈值。

若所述输入字符串超出所述长度阈值，迭代单元 134”根据所述输入字符串，以及所述上下文信息，迭代确定与所述输入字符串对应的一个或多个输入候选项。具体地，若所述输入字符串超出所述长度阈值，迭代单元 134”通过根据所述输入字符串，以及所述上下文信息，通过例如根据所述长度阈值，对所述输入字符串进行分段等处理，确定与第
25 一个长度阈值的输入字符串所对应的所述输入候选项，并将所确定的输入候选项再次作为上下文信息，从而迭代确定下一个长度阈值的输

入字符串所对应的输入候选项等的方式，确定与所述输入字符串对应的一个或多个输入候选项。

优选地，所述检测单元 133”还可以获取所述输入字符串在输入词库中对应的一个或多个抽样输入候选项；检测所述一个或多个抽样输入候选项中至少一个中的字数是超出预定的字数阈值，以确定所述输入字符串是否超出预定的长度阈值。具体地，所述检测单元 133”还可以通过例如将所述输入字符串直接在所述输入词库中进行匹配，获取与所述全部输入字符串所对应的一个或多个输入候选项，并从中随机抽取或指定抽取（如选择第一个输入候选项等）方式，获取所述输入字符串在输入词库中对应的一个或多个抽样输入候选项；通过对所述一个或多个抽样输入候选项进行检测，如所述抽样输入候选项中的其中至少一个字数超出预定的字数阈值，则确定所述输入字符串是否超出预定的长度阈值。

优选地，所述迭代单元 134”还可以当所述输入字符串超出所述长度阈值，根据所述上下文信息，以及所述输入字符串中与所述上下文信息相邻的部分输入字符串，确定与所述部分输入字符串对应的一个或多个部分输入候选项；根据所述部分输入字符串，以及所述输入字符串中其余输入字符串，确定与所述其余输入字符串对应的一个或多个其余输入候选项，其中，所述部分输入候选项被用作所述其余输入字符串的上下文信息；根据所述一个或多个部分输入候选项，以及所述一个或多个其余输入候选项，确定所述一个或多个输入候选项。

具体地，当所述输入字符串超出所述长度阈值，根据所述上下文信息，通过例如分词等处理，对所述输入字符串进行分词或分段等，从而获取所述输入字符串中与所述上下文信息相邻的部分输入字符串，通过将所述部分输入字符串结合所述上下文信息，确定与所述部分输入字符串对应的一个或多个部分输入候选项；在此，所述确定方法与图 1 中所述候选项确定装置 13 中的确定方法相同或相似，故此处不再赘述，并通过引用的方式包含于此。所述迭代单元 134”将所述部分输入候选项作为其余输入字符串的上下文信息，通过迭代的方式，确定与所述其

余输入字符串相对应的一个或多个其余输入候选项；迭代后通过例如对字符串之间的关系，对所述部分输入候选项与其余输入候选项进行拼接等方式，确定一个或多个输入候选项。

图4示出根据本发明另一个方面的一种用于提供输入字符串所对应的输入候选项的方法流程图。具体地，在步骤s1中，输入设备获取输入字符串；在步骤s2中，输入设备确定所述输入字符串的上下文信息；在步骤s3中，输入设备根据所述输入字符串，以及所述上下文信息，确定与所述输入字符串对应的一个或多个输入候选项；在步骤s4中，输入设备将所述一个或多个输入候选项中至少一个提供
10 给所述输入字符串所对应的目标应用。

上述各步骤之间是持续不断工作的，在此，本领域技术人员应理解“持续”是指上述各步骤分别实时地或者按照设定的或实时调整的工作模式要求，进行输入字符串的获取、上下文信息的确定、输入候选项的确定、输入候选项的提供等，直至输入设备停止获取输入字符
15 串。

在步骤s1中，输入设备获取输入字符串。具体地，在步骤s1中，输入设备通过从输入框中接收用户或其他应用等通过输入、选择、点击等操作所输入的字符串；或者根据各种通信协议(Communications Protocol)，通过各种数据传输接口，获取从其他应用的输入框中的字
20 符串；其中，所述输入框如短信编辑框、内容输入框等。其中，所述输入字符串包括但不限于通过如全拼或简拼等方式所输入的拼音编码字符等。

在步骤s2中，输入设备确定所述输入字符串的上下文信息。具体地，在步骤s2中，输入设备通过根据步骤s1所获取的输入字符串，
25 利用当前所输入的字符串所属的文字等，将所述当前输入字符串所属的文字与一个或多个词库进行匹配，获取与所述输入字符串相关联的一种或多种匹配信息，并作为所述输入字符串的上下文信息；例如，当输入“huacao”时，先确定所述输入字符串所属的文字可能为“花草”，根据“花草”在所属词库中的匹配信息，确定其下文信息可能

为“茶”等；或者，在步骤 s2 中，输入设备通过根据与所述输入字符串相对应的历史记录信息，对所述输入字符串的上下文信息进行确定，例如，所述输入字符串前刚刚上屏的前一个输入字符串，则可作为所述输入字符串的前文信息。在此，所述上下文信息包括但不限于与所述输入字符串相对应的上文信息或下文信息；例如，当所述输入为“shishi”时，当前次输入所上屏的词组为“客观”时，则所述“shishi”对应的上文信息为“客观”；若屏幕上存在一个词组“计划”，且输入“shishi”时，光标位于“计划”前，则其所对应的下文信息为“计划”等。

在步骤 s3 中，输入设备根据所述输入字符串，以及所述上下文信息，确定与所述输入字符串对应的一个或多个输入候选项。具体地，在步骤 s3 中，输入设备通过根据步骤 s1 中所获取的输入字符串，以及步骤 s1 中所确定的上下文信息，通过将与所述输入字符串相对应的一个或多个上下文信息直接作为一个或多个输入候选项，或者将所述输入字符串相对应的上下文信息进行处理后作为输入候选项的方式，确定与所述输入字符串对应的一个或多个输入候选项。例如，继上例，当所述输入字符串为“shishi”时，其所对应的上文信息可能为“客观”，“我”，“谈论”等，当所述上文信息为“客观”时，通过利用将所述输入字符串，以及所述上下文信息进行匹配等方式，确定输入候选项为“事实”、“实时”、“实施”等；当所述上文信息为“我”时，确定输入候选项为“试试”、“实施”、“试试”等；当所述上文信息为“谈论”时，确定输入候选项为“时事”、“事实”、“世事”等。

在步骤 s4 中，输入设备将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。具体地，在步骤 s4 中，输入设备通过直接提供，或根据各种通信协议 (Communications Protocol)，通过各种数据传输接口，通过例如重新排序或筛选等方式，将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。其中，所述目标应用包括但不限于与所述步骤 s1 中所对应的输入框相对应的相同应用，或者与所述步骤 s1 中所对应

的输入框相关联的其他应用等，如直接利用所述输入候选项进行搜索等。

优选地，在步骤 s3 中，输入设备还可以根据所述输入字符串，以及所述上下文信息，确定包含所述输入字符串的增补字符串，其中，所述增补字符串还包括所述上下文信息所对应的字符串；根据所述增补字符串，在输入词库中进行匹配查询，以获得与所述增补字符串对应的一个或多个增补输入候选项；根据所述上下文信息，对所述一个或多个增补输入候选项进行筛选处理，以获得所述一个或多个输入候选项。具体地，在步骤 s3 中，输入设备还可以根据所述输入字符串，以及所述上下文信息，通过将所述输入字符串结合所述上下文信息与例如输入词库进行匹配等方式，确定包含所述输入字符串的增补字符串，例如，当输入字符串为“xuxu”时，当上下文信息为“形态”，则确定所述增补字符串为“rusheng”；或者，当输入字符串为“tianan”，上下文信息为“北京”，则确定所述增补字符串为“men”，在此，所述增补字符串包括但不限于通过匹配等方式，确定一个或多个通过拼接而成的多段字符串。在步骤 s3 中，输入设备通过将所述增补字符串在输入词库中进行匹配查询，以获得与所述增补字符串对应的一个或多个增补输入候选项；继前例，所述增补字符串为“rusheng”时所对应的增补输入候选项如“如生”、“入声”、“入胜”等，所述增补字符串为“men”时，所对应的增补输入候选项如“天安门”、“门”等。或者，所述增补输入候选项还可以直接包括所述上下文信息，如直接生成“形态栩栩如生”、“北京天安门”等。根据所述上下文信息，对所述一个或多个增补输入候选项进行筛选处理，继前例，根据上下文信息，分别确定输入候选项为“栩栩如生”、“天安门”从而获得所述一个或多个输入候选项；在此，若所述增补输入候选项中包含上下文信息，则将所述增补输入候选项中的上下文信息删除后，获得所述一个或多个输入候选项，如所述增补输入候选项“北京天安门”中的“北京”删除，获得所述一个或多个输入候选项“天安门”。

优选地，在步骤 s2 中，输入设备还可以通过对所述输入字符串所

对应的相关文本进行自然语言处理，从所述相关文本提取对应的焦点词语，以作为所述上下文，其中，所述焦点词语邻近所述输入字符串所对应的输入位置。具体地，在步骤 s2 中，输入设备还可以通过分词或语义分析等方式，对所述输入字符串所对应的相关文本进行自然语言处理，其中，所述相关文本包括但不限于所述输入字符串本身所对应的文本，或者所述输入字符串所对应的输入位置前后已经上屏的相关文本；通过对所述相关文本中经由自然语言处理，提取对应的焦点词语以作为所述上下文。例如，当所述输入字符串为“shishi”时，其对应的相关文本为“上海地铁一期工程”，且所述输入字符串位于“上海地铁一期工程”之后，通过对所述相关文本进行自然语言处理，提取焦点词语为“工程”，并将“工程”作为上下文。进一步地，因此，当输入字符串为“shishi”时，该输入设备根据所述上下文“工程”，确定输入候选项为“实施”、“实时”、“试试”等。

图 5 示出根据本发明一个优选实施例的一种用于提供输入字符串所对应的输入候选项的方法流程图。具体地，在步骤 s1' 中，输入设备获取输入字符串；在步骤 s2' 中，输入设备确定所述输入字符串的上下文信息；在步骤 s31' 中，输入设备根据所述输入字符串，在输入词库中进行匹配查询，以获得与所述输入字符串对应的一个或多个初步输入候选项；在步骤 s32' 中，输入设备根据所述上下文信息，对所述一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选项；在步骤 s4' 中，输入设备将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。其中，所述步骤 s1'、步骤 s2'、步骤 s4' 与图 4 所示对应步骤相同或基本相同，故此处不再赘述，并通过引用的方式包含于此。

上述各步骤之间是持续不断工作的，在此，本领域技术人员应理解“持续”是指上述各步骤分别实时地或者按照设定的或实时调整的工作模式要求，进行输入字符串的获取、上下文信息的确定、初步输入候选项的获取、输入候选项的确定、输入候选项的提供等，直至输入设备停止获取输入字符串。

在步骤 s31' 中，输入设备根据所述输入字符串，在输入词库中进行匹配查询，以获得与所述输入字符串对应的一个或多个初步输入候选项。具体地，在步骤 s31' 中，输入设备通过获取所述输入字符串，将所述输入字符串与输入词库中的词进行匹配，以获得与所述输入字符串对应
5 的一个或多个初步输入候选项。例如，当输入字符串为“shishi”时，根据在输入词库中进行匹配查询后，获取“实时”、“试试”、“实施”、“事实”、“适时”、“逝世”等多个初步输入候选项。

在步骤 s32' 中，输入设备根据所述上下文信息，对所述一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选项。具体
10 地，在步骤 s32' 中，输入设备根据所述输入字符串的上下文信息，通过根据自然语言模型或词语搭配数据库等，确定上下文信息与初步输入候选项的搭配概率，并根据所述搭配概率对所述初步输入候选项进行筛选或排序，以获得输入候选项。例如，继前例，当上下文信息为“客观”，则经处理后，将初步输入候选项进行重新排序，顺序变为“事实”、“实
15 时”、“试试”、“实施”、“适时”、“逝世”等。

优选地，在步骤 s32' 中，输入设备还可以根据所述上下文信息，在词语搭配数据库中进行匹配查询，以确定所述初步输入候选项与所述上下文信息对应的搭配相关度信息；根据所述搭配相关度信息，对所述一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选
20 项。具体地，在步骤 s32' 中，输入设备还可以根据所述上下文信息，在词语搭配数据库中进行匹配查询，通过根据所述词语搭配数据库中的一个或多个词条与所述上下文的匹配关系等，确定所述初步输入候选项与
所述上下文信息对应的搭配相关度信息；例如若为 bull 变量，则所述搭配相关度信息表示如是否存在搭配；或为连续变量，如所述搭配相关
25 度信息表示搭配概率为多少；在此，所述搭配相关度信息可根据如词语搭配数据库或其他相关数据库中的上下文与初步输入候选项的组合频率等通过机器学习获得。根据所述搭配相关度信息，通过对所述一个或多个初步输入候选项进行如筛选或排序处理，获得所述一个或多个输入候选项。

更优选地，所述方法还包括步骤 s5'（未示出）和步骤 s6'（未示出）；其中，在步骤 s5' 中，输入设备对一个或多个文本信息进行分析处理，以获得搭配频度满足一定阈值的分词搭配关系；在步骤 s6' 中，输入设备根据所述分词搭配关系，建立或更新所述词语搭配数据库。具体地，在步骤 s5' 中，输入设备例如通过对各类日志、各类文章等进行获取，并对所获取文本信息进行如先分词，后统计相邻词间的共现频率等分析方法，当所述共现频率超过一定阈值时，则判断所述相邻词间存在一定的搭配关系，从而获得搭配频度满足一定阈值的分词搭配关系；在步骤 s6' 中，输入设备根据所述分词搭配关系，建立所述词语搭配数据库，或对现有词语搭配数据库进行更新等，在此，所述词语搭配数据库中包括但不限于分词搭配关系以及分词搭配关系之间的搭配概率信息；并且，所述词语搭配数据库中可包含现有固定词语，如成语、俗语等，同时经过所述更新，可以将由于现有网络应用的快速发展所产生的新的词语及表述进行包含，从而有效提高用户的输入效率。

图 6 示出根据本发明另一个优选实施例的一种用于提供输入字符串所对应的输入候选项的方法流程图。具体地，在步骤 s1'' 中，输入设备获取输入字符串；在步骤 s2'' 中，输入设备确定所述输入字符串的上下文信息；在步骤 s3'' 中，输入设备检测所述输入字符串是否超出预定的长度阈值；若所述输入字符串超出所述长度阈值，在步骤 s4'' 中，输入设备根据所述输入字符串，以及所述上下文信息，迭代确定与所述输入字符串对应的一个或多个输入候选项；在步骤 s4'' 中，输入设备将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。其中，所述步骤 s1''、步骤 s2''、步骤 s4'' 与图 4 所示对应步骤相同或基本相同，故此处不再赘述，并通过引用的方式包含于此。

上述各步骤之间是持续不断工作的，在此，本领域技术人员应理解“持续”是指上述各步骤分别实时地或者按照设定的或实时调整的工作模式要求，进行输入字符串的获取、上下文信息的确定、长度阈值的检测、输入候选项的确定、输入候选项的提供等，直至输入设备

停止获取输入字符串。

在步骤 s33”中，输入设备检测所述输入字符串是否超出预定的长度阈值。具体地，在步骤 s33”中，输入设备通过基于预置的或用户自行设置的长度阈值，对所述输入字符串的长度进行检测，检测所述输入字符串的长度是否超出了预定的长度阈值。例如，当预定的字符长度阈值为 N，当所述输入字符串的长度为 N+1 时，则在步骤 s33”中，输入设备检测所述输入字符串超出了预定的长度阈值。

若所述输入字符串超出所述长度阈值，在步骤 s34”中，输入设备根据所述输入字符串，以及所述上下文信息，迭代确定与所述输入字符串对应的一个或多个输入候选项。具体地，若所述输入字符串超出所述长度阈值，在步骤 s34”中，输入设备通过根据所述输入字符串，以及所述上下文信息，通过例如根据所述长度阈值，对所述输入字符串进行分段等处理，确定与第一个长度阈值的输入字符串所对应的所述输入候选项，并将所确定的输入候选项再次作为上下文信息，从而迭代确定下一个长度阈值的输入字符串所对应的输入候选项等的方式，确定与所述输入字符串对应的一个或多个输入候选项。

优选地，在步骤 s33”中，输入设备还可以获取所述输入字符串在输入词库中对应的一个或多个抽样输入候选项；检测所述一个或多个抽样输入候选项中至少一个中的字数是超出预定的字数阈值，以确定所述输入字符串是否超出预定的长度阈值。具体地，在步骤 s33”中，输入设备还可以通过例如将所述输入字符串直接在所述输入词库中进行匹配，获取与所述全部输入字符串所对应的一个或多个输入候选项，并从中随机抽取或指定抽取（如选择第一个输入候选项等）方式，获取所述输入字符串在输入词库中对应的一个或多个抽样输入候选项；通过对所述一个或多个抽样输入候选项进行检测，如所述抽样输入候选项中的其中至少一个字数超出预定的字数阈值，则确定所述输入字符串是否超出预定的长度阈值。

优选地，在步骤 s34”中，输入设备还可以当所述输入字符串超出所述长度阈值，根据所述上下文信息，以及所述输入字符串中与所述上

下文信息相邻的部分输入字符串，确定与所述部分输入字符串对应的一个或多个部分输入候选项；根据所述部分输入字符串，以及所述输入字符串中其余输入字符串，确定与所述其余输入字符串对应的一个或多个其余输入候选项，其中，所述部分输入候选项被用作所述其余输入字符串的上下文信息；根据所述一个或多个部分输入候选项，以及所述一个或多个其余输入候选项，确定所述一个或多个输入候选项。

具体地，当所述输入字符串超出所述长度阈值，根据所述上下文信息，通过例如分词等处理，对所述输入字符串进行分词或分段等，从而获取所述输入字符串中与所述上下文信息相邻的部分输入字符串，通过将所述部分输入字符串结合所述上下文信息，确定与所述部分输入字符串对应的一个或多个部分输入候选项；在此，所述确定方法与图 4 中所述步骤 s3 中的确定方法相同或相似，故此处不再赘述，并通过引用的方式包含于此。在步骤 s34”中，输入设备将所述部分输入候选项作为其余输入字符串的上下文信息，通过迭代的方式，确定与所述其余输入字符串相对应的一个或多个其余输入候选项；迭代后通过例如对字符串之间的关系，对所述部分输入候选项与其余输入候选项进行拼接等方式，确定一个或多个输入候选项。

对于本领域技术人员而言，显然本发明不限于上述示范性实施例的细节，而且在不背离本发明的精神或基本特征的情况下，能够以其他的具體形式实现本发明。因此，无论从哪一点来看，均应将实施例看作是示范性的，而且是非限制性的，本发明的范围由所附权利要求而不是上述说明限定，因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。此外，显然“包括”一词不排除其他单元或步骤，单数不排除复数。装置权利要求中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第一，第二等词语用来表示名称，而并不表示任何特定的顺序。

权 利 要 求 书

1. 一种用于提供输入字符串所对应的输入候选项的方法，其中，该方法包括以下步骤：

- 5 a 获取输入字符串；
 b 确定所述输入字符串的上下文信息；
 c 根据所述输入字符串，以及所述上下文信息，确定与所述输入字符串对应的一个或多个输入候选项；
 d 将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。

2. 根据权利要求 1 所述的方法，其中，所述步骤 c 包括：

- 根据所述输入字符串，在输入词库中进行匹配查询，以获得与所述输入字符串对应的一个或多个初步输入候选项；
 c1 根据所述上下文信息，对所述一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选项。

3. 根据权利要求 2 所述的方法，其中，所述步骤 c1 包括：

- 根据所述上下文信息，在词语搭配数据库中进行匹配查询，以确定所述初步输入候选项与所述上下文信息对应的搭配相关度信息；
 - 根据所述搭配相关度信息，对所述一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选项。

4. 根据权利要求 3 所述的方法，其中，该方法还包括：

- 对一个或多个文本信息进行分析处理，以获得搭配频度满足一定阈值的分词搭配关系；
 - 根据所述分词搭配关系，建立或更新所述词语搭配数据库。

5. 根据权利要求 1 所述的方法，其中，所述步骤 c 包括：

- 根据所述输入字符串，以及所述上下文信息，确定包含所述输入字符串的增补字符串，其中，所述增补字符串还包括所述上下文信息所对应的字符串；
 - 根据所述增补字符串，在输入词库中进行匹配查询，以获得与所

述增补字符串对应的一个或多个增补输入候选项;

- 根据所述上下文信息, 对所述一个或多个增补输入候选项进行筛选处理, 以获得所述一个或多个输入候选项。

5 6. 根据权利要求 1 至 5 中任一项所述的方法, 其中, 所述步骤 b 包括:

- 通过对所述输入字符串所对应的相关文本进行自然语言处理, 从所述相关文本提取对应的焦点词语, 以作为所述上下文, 其中, 所述焦点词语邻近所述输入字符串所对应的输入位置。

10 7. 根据权利要求 1 至 6 中任一项所述的方法, 其中, 所述步骤 c 包括:

c2 检测所述输入字符串是否超出预定的长度阈值;

c3 若所述输入字符串超出所述长度阈值, 根据所述输入字符串, 以及所述上下文信息, 迭代确定与所述输入字符串对应的一个或多个输入候选项。

15 8. 根据权利要求 7 所述的方法, 其中, 所述步骤 c2 包括:

- 获取所述输入字符串在输入词库中对应的一个或多个抽样输入候选项;

- 检测所述一个或多个抽样输入候选项中至少一个中的字数是超出预定的字数阈值, 以确定所述输入字符串是否超出预定的长度阈值。

20 9. 根据权利要求 7 或 8 所述的方法, 其中, 所述步骤 c3 包括:

- 当所述输入字符串超出所述长度阈值, 根据所述上下文信息, 以及所述输入字符串中与所述上下文信息相邻的部分输入字符串, 确定与所述部分输入字符串对应的一个或多个部分输入候选项;

25 - 根据所述部分输入字符串, 以及所述输入字符串中其余输入字符串, 确定与所述其余输入字符串对应的一个或多个其余输入候选项, 其中, 所述部分输入候选项被用作所述其余输入字符串的上下文信息;

- 根据所述一个或多个部分输入候选项, 以及所述一个或多个其余输入候选项, 确定所述一个或多个输入候选项。

10. 一种用于提供输入字符串所对应的输入候选项的输入设备, 其

中，该设备包括：

获取装置，用于获取输入字符串；

上下文确定装置，用于确定所述输入字符串的上下文信息；

候选项确定装置，用于根据所述输入字符串，以及所述上下文信息，

5 确定与所述输入字符串对应的一个或多个输入候选项；

提供装置，用于将所述一个或多个输入候选项中至少一个提供给所述输入字符串所对应的目标应用。

11. 根据权利要求 10 所述的输入设备，其中，所述候选项确定装置包括：

10 匹配单元，用于根据所述输入字符串，在输入词库中进行匹配查询，以获得与所述输入字符串对应的一个或多个初步输入候选项；

处理单元，用于根据所述上下文信息，对所述一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选项。

12. 根据权利要求 11 所述的输入设备，其中，所述处理单元用于：

15 - 根据所述上下文信息，在词语搭配数据库中进行匹配查询，以确定所述初步输入候选项与所述上下文信息对应的搭配相关度信息；

- 根据所述搭配相关度信息，对所述一个或多个初步输入候选项进行处理，以获得所述一个或多个输入候选项。

13. 根据权利要求 12 所述的输入设备，其中，该设备还包括：

20 文本处理装置，用于对一个或多个文本信息进行分析处理，以获得搭配频度满足一定阈值的分词搭配关系；

建立装置，用于根据所述分词搭配关系，建立或更新所述词语搭配数据库。

14. 根据权利要求 10 所述的输入设备，其中，所述候选项确定装置
25 用于：

- 根据所述输入字符串，以及所述上下文信息，确定包含所述输入字符串的增补字符串，其中，所述增补字符串还包括所述上下文信息所对应的字符串；

- 根据所述增补字符串，在输入词库中进行匹配查询，以获得与所

述增补字符串对应的一个或多个增补输入候选项;

- 根据所述上下文信息, 对所述一个或多个增补输入候选项进行筛选处理, 以获得所述一个或多个输入候选项。

15 15. 根据权利要求 10 至 14 中任一项所述的输入设备, 其中, 所述上下文确定装置用于:

- 通过对所述输入字符串所对应的相关文本进行自然语言处理, 从所述相关文本提取对应的焦点词语, 以作为所述上下文, 其中, 所述焦点词语邻近所述输入字符串所对应的输入位置。

10 16. 根据权利要求 10 至 15 中任一项所述的输入设备, 其中, 所述候选项确定装置包括:

检测单元, 用于检测所述输入字符串是否超出预定的长度阈值;

迭代单元, 用于若所述输入字符串超出所述长度阈值, 根据所述输入字符串, 以及所述上下文信息, 迭代确定与所述输入字符串对应的一个或多个输入候选项。

15 17. 根据权利要求 16 所述的输入设备, 其中, 所述检测单元用于:

- 获取所述输入字符串在输入词库中对应的一个或多个抽样输入候选项;

- 检测所述一个或多个抽样输入候选项中至少一个中的字数是超出预定的字数阈值, 以确定所述输入字符串是否超出预定的长度阈值。

20 18. 根据权利要求 16 或 17 所述的输入设备, 其中, 所述迭代单元用于:

- 当所述输入字符串超出所述长度阈值, 根据所述上下文信息, 以及所述输入字符串中与所述上下文信息相邻的部分输入字符串, 确定与所述部分输入字符串对应的一个或多个部分输入候选项;

25 - 根据所述部分输入字符串, 以及所述输入字符串中其余输入字符串, 确定与所述其余输入字符串对应的一个或多个其余输入候选项, 其中, 所述部分输入候选项被用作所述其余输入字符串的上下文信息;

- 根据所述一个或多个部分输入候选项, 以及所述一个或多个其余输入候选项, 确定所述一个或多个输入候选项。

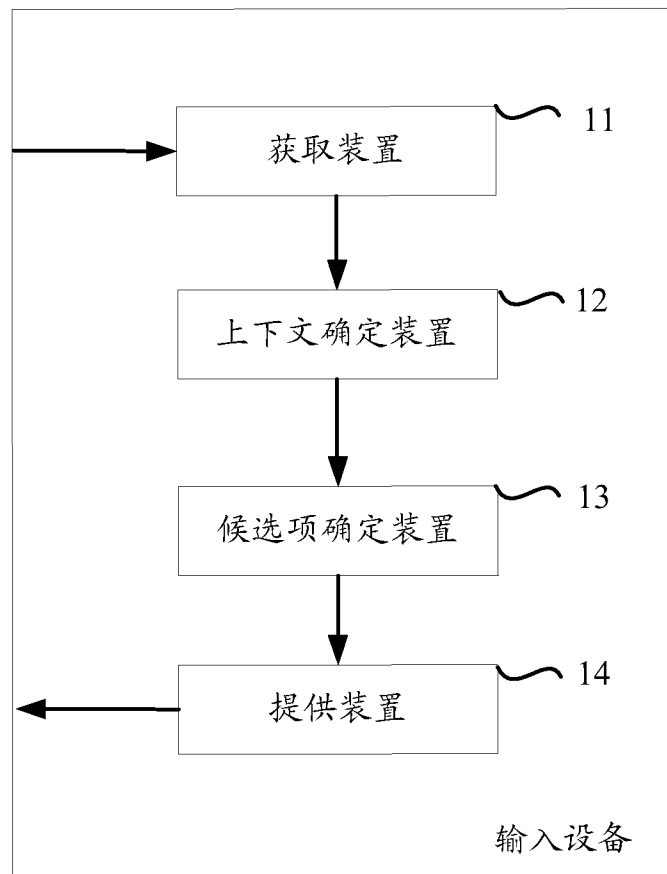


图1

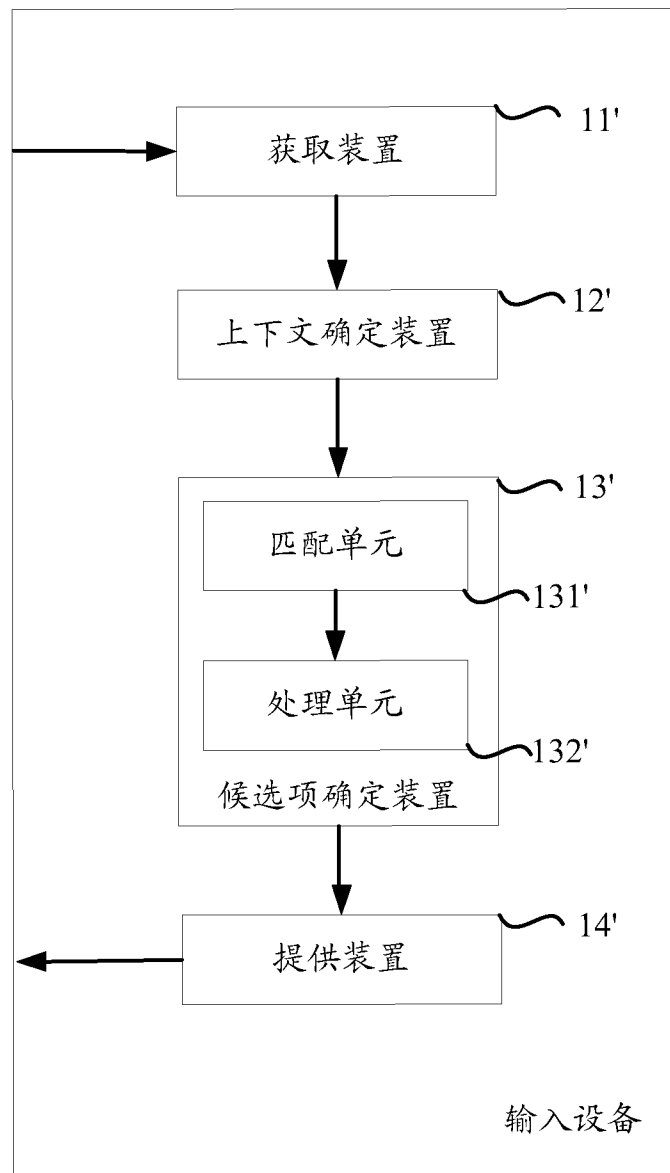


图2

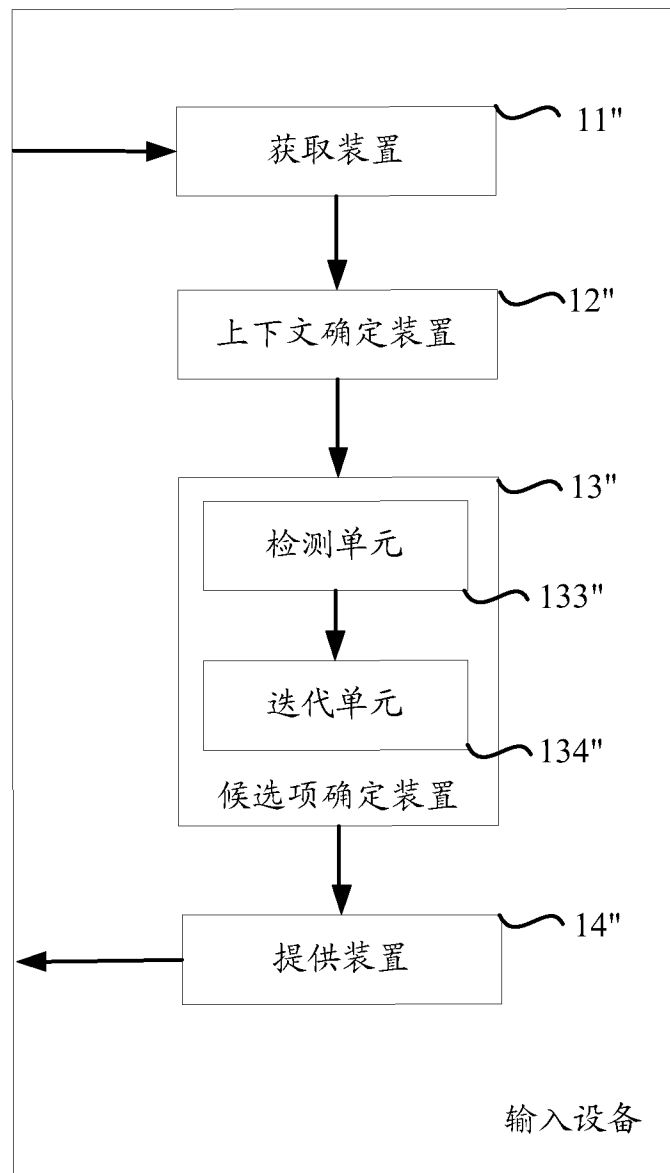


图 3

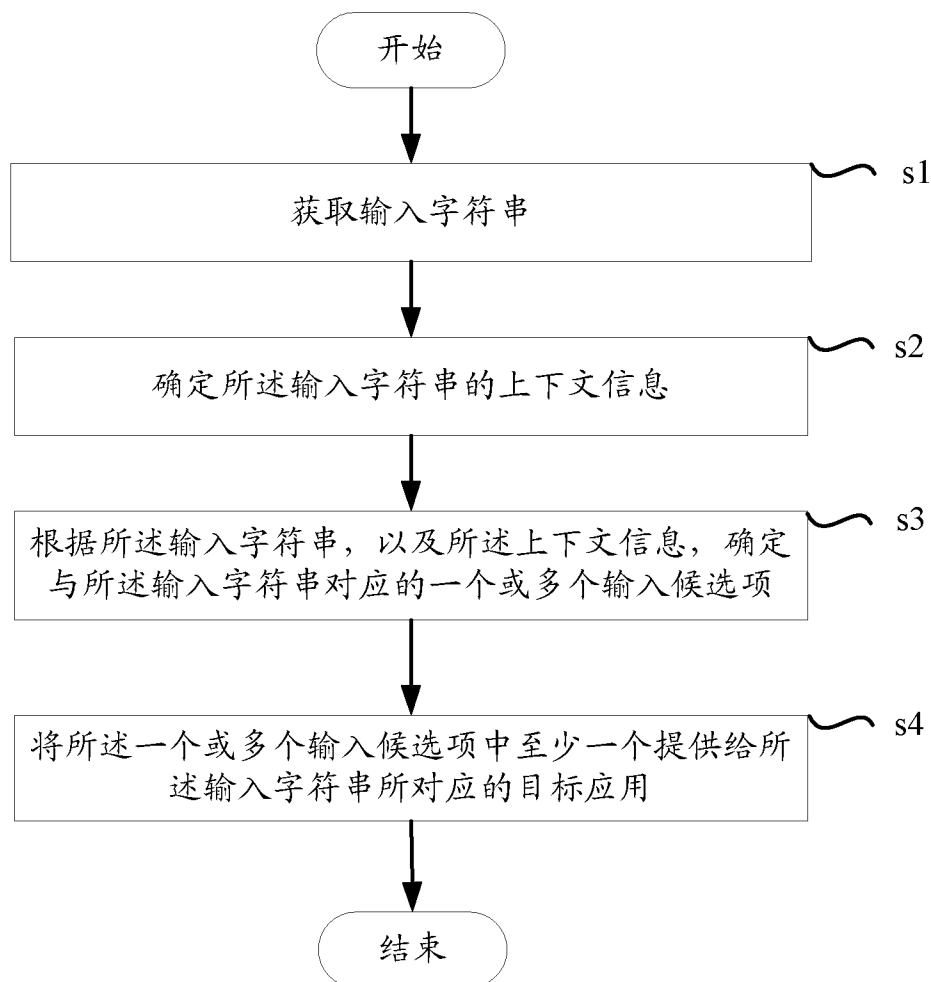


图 4

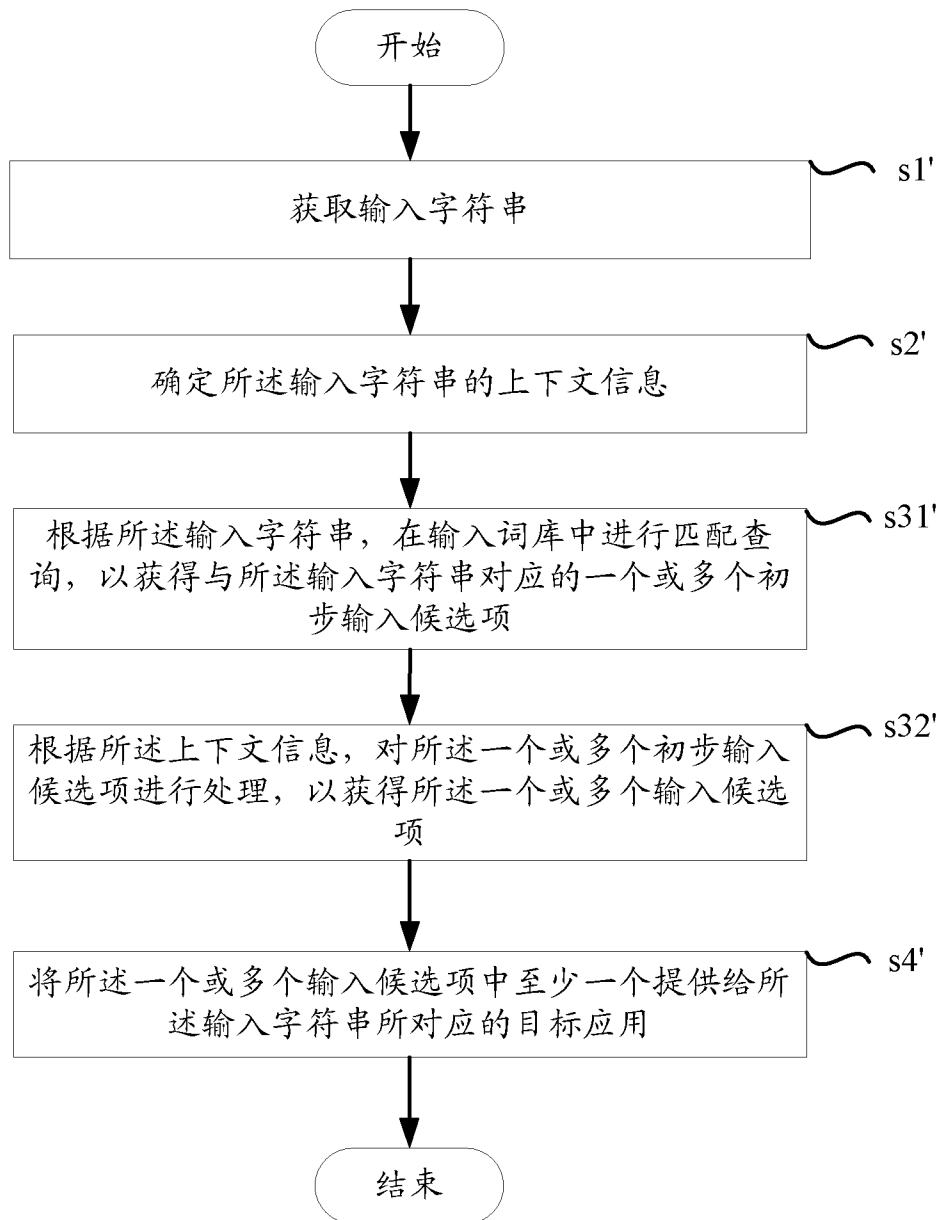


图5

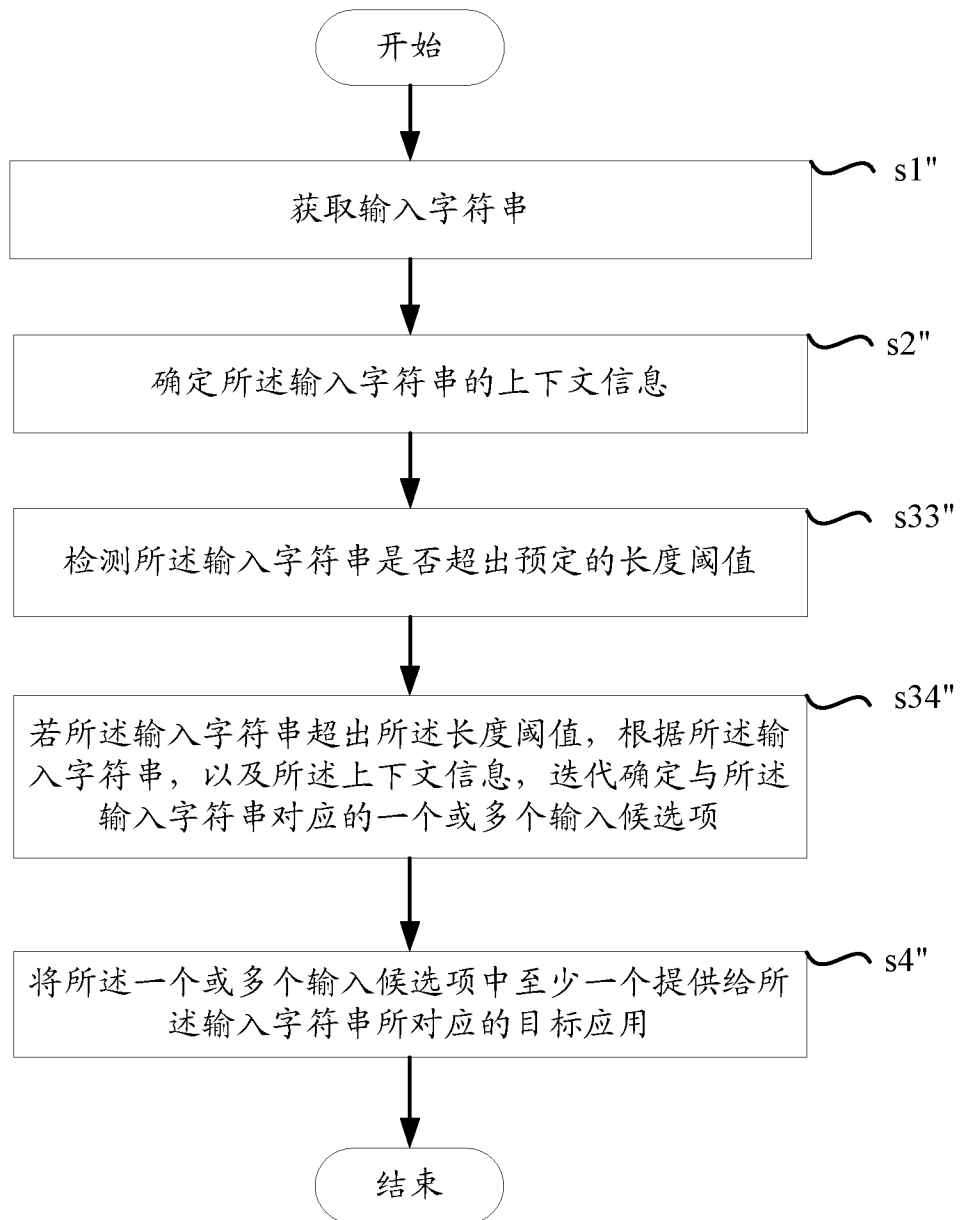


图6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/086968

A. CLASSIFICATION OF SUBJECT MATTER

G06F 3/023 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G06F 3/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) WPI, EPODOC, CNPAT, CNKI: INPUT, CHARACTER, CANDIDATE, CONTEXT, PROVIDE, DETERMINE, DATABASE, MATCH, RELATE, RELEVANCE, SEMANT+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 101246410 A, (GOOGLE INC.), 20 August 2008 (20.08.2008), claims 14 and 23	1, 10
A	CN 102346743 A, (TENCENT TECHNOLOGY SHENZHEN CO., LTD.), 08 February 2012 (08.02.2012), the whole document	1-18
P,X	CN 103076892 A, (BAIDU ON-LINE NETWORK TECHNOLOGY CO., LTD.), 01 May 2013 (01.05.2013), claims 1-18	1-18

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

11 February 2014 (11.02.2014)

Date of mailing of the international search report

27 February 2014 (27.02.2014)

Name and mailing address of the ISA
State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao
Haidian District, Beijing 100088, China
Facsimile No. (86-10) 62019451

Authorized officer

LI, Fei

Telephone No. (86-10) 62411696

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/086968

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 101246410 A	20.08.2008	US 2008193015 A1	14.08.2008
		US 8028230 B2	27.09.2011
		WO 2008100951 A2	21.08.2008
		WO 2008100951 A3	19.03.2009
		TW 200900968 A	01.01.2009
		KR 20090109585 A	20.10.2009
		JP 2010518534 A	27.05.2010
CN 102346743 A	08.02.2012	None	
CN 103076892 A	01.05.2013	None	

国际检索报告

国际申请号

PCT/CN2013/086968

A. 主题的分类

G06F 3/023 (2006.01) i

按照国际专利分类(IPC)或者同时按照国家分类和 IPC 两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: G06F 3/-

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词 (如使用))

WPI, EPODOC, CNPAT, CNKI: 输入、字符、候选、上下文、语义、提供、确定、词库、匹配、查询、相关度; INPUT, CHARACTER, CANDIDATE, CONTEXT, PROVIDE, DETERMINE, DATABASE, MATCH, RELATE, RELEVANCE

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 101246410 A, (谷歌股份有限公司), 20.8 月 2008(20.08.2008), 权利要求 14 和 23	1、10
A	CN 102346743 A, (腾讯科技(深圳)有限公司), 08.2 月 2012(08.02.2012), 全文	1-18
P, X	CN 103076892 A, (百度在线网络技术(北京)有限公司), 01.5 月 2013(01.05.2013), 权利要求 1-18	1-18



其余文件在 C 栏的续页中列出。



见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

11.2 月 2014 (11.02.2014)

国际检索报告邮寄日期

27.2 月 2014 (27.02.2014)

ISA/CN 的名称和邮寄地址:

中华人民共和国国家知识产权局

中国北京市海淀区蓟门桥西土城路 6 号 100088

传真号: (86-10)62019451

受权官员

李菲

电话号码: (86-10) **62411696**

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2013/086968

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN 101246410 A	20.08.2008	US 2008193015 A1	14.08.2008
		US 8028230 B2	27.09.2011
		WO 2008100951 A2	21.08.2008
		WO 2008100951 A3	19.03.2009
		TW 200900968 A	01.01.2009
		KR 20090109585 A	20.10.2009
		JP 2010518534 A	27.05.2010
CN 102346743 A	08.02.2012	无	
CN 103076892 A	01.05.2013	无	