



(12)发明专利申请

(10)申请公布号 CN 108351876 A

(43)申请公布日 2018.07.31

(21)申请号 201580084742.X

(51)Int.Cl.

(22)申请日 2015.09.22

G06F 17/30(2006.01)

(85)PCT国际申请进入国家阶段日

2018.05.22

(86)PCT国际申请的申请数据

PCT/CN2015/090237 2015.09.22

(87)PCT国际申请的公布数据

W02017/049454 EN 2017.03.30

(71)申请人 纽昂斯通讯公司

地址 美国马萨诸塞州

(72)发明人 韩客松 陈岳峰 徐燃

(74)专利代理机构 中国国际贸易促进委员会专

利商标事务所 11038

代理人 周衡威

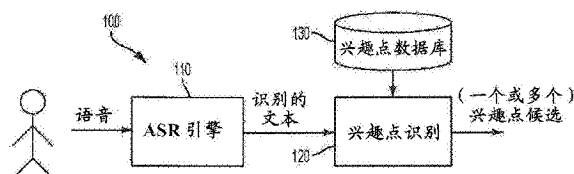
权利要求书4页 说明书21页 附图5页

(54)发明名称

用于兴趣点识别的系统和方法

(57)摘要

提供了一种包括至少一个处理器和至少一个计算机可读存储介质的系统。所述至少一个计算机可读存储介质可以存储多个兴趣点段索引。所述至少一个计算机可读存储介质可以进一步存储将所述至少一个处理器编程为执行以下动作的指令：将第一文本段与存储在所述至少一个计算机可读存储介质中的第一兴趣点段索引匹配；将第二文本段与存储在所述至少一个计算机可读存储介质中的第二兴趣点段索引匹配；并且使用第一兴趣点段索引和第二兴趣点段索引来标识与第一文本段和第二文本段这二者匹配的一个或多个候选兴趣点条目。



1. 一种系统,包括:

至少一个处理器;以及

存储多个兴趣点段索引的至少一个计算机可读存储介质,其中,所述至少一个计算机可读存储介质进一步存储将所述至少一个处理器编程为执行以下动作的指令:

将第一文本段与存储在所述至少一个计算机可读存储介质中的第一兴趣点段索引匹配;

将第二文本段与存储在所述至少一个计算机可读存储介质中的第二兴趣点段索引匹配;并且

使用第一兴趣点段索引和第二兴趣点段索引来标识与第一文本段和第二文本段这二者都匹配的一个或多个候选兴趣点条目。

2. 根据权利要求1所述的系统,其中,所述至少一个处理器被编程为:

使用第一兴趣点段索引来标识与第一文本段匹配的一个或多个兴趣点条目的第一集合;

使用第二兴趣点段索引来标识与第二文本段匹配的一个或多个兴趣点条目的第二集合;并且

将在第一集合和第二集合这二者中都出现的一个或多个兴趣点条目标识为所述一个或多个候选兴趣点条目。

3. 根据权利要求1所述的系统,其中,所述至少一个计算机可读存储介质进一步存储语言模型,所述语言模型包括与多个兴趣点段相关的统计信息,并且其中,所述至少一个处理器被进一步编程为:

使用所述语言模型来从输入音频信号识别第一文本段和第二文本段。

4. 根据权利要求3所述的系统,其中:

所述第一文本段包括所述多个兴趣点段中的第一兴趣点段,所述第一兴趣点段对应于所述第一兴趣点段索引;并且

所述第二文本段包括所述多个兴趣点段中的第二兴趣点段,所述第二兴趣点段对应于所述第二兴趣点段索引。

5. 根据权利要求1所述的系统,其中,所述至少一个处理器被进一步编程为:

将第一得分与第一候选兴趣点条目相关联,所述第一得分指示第一候选兴趣点条目与第一文本段和第二文本段的相似性程度;

将第二得分与第二候选兴趣点条目相关联,所述第二得分指示第二候选兴趣点条目与第一文本段和第二文本段的相似性程度;并且

至少部分地基于第一得分和第二得分来对第一候选兴趣点条目和第二候选兴趣点条目进行排名。

6. 根据权利要求5所述的系统,其中,所述至少一个处理器被进一步编程为:

至少部分地通过在文本上将第一文本段和第二文本段与第一候选兴趣点条目的兴趣点名称进行比较来生成文本得分;

至少部分地通过将第一文本段和第二文本段的音标表示与第一候选兴趣点条目的兴趣点名称的音标表示进行比较来生成读音得分;并且

将第一得分生成为文本得分和读音得分的加权和。

7. 根据权利要求1所述的系统, 其中, 以编码的形式存储所述多个兴趣点段索引, 并且其中, 所述至少一个处理器被进一步编程为:

在使用第一兴趣点段索引和第二兴趣点段索引标识与第一文本段和第二文本段这两者都匹配的一个或多个候选兴趣点条目之前, 对第一兴趣点段索引和第二兴趣点段索引进行解码。

8. 一种由包括至少一个处理器和存储多个兴趣点段索引的至少一个计算机可读存储介质的系统来执行的方法, 所述方法包括以下动作:

将第一文本段与存储在所述至少一个计算机可读存储介质中的第一兴趣点段索引匹配;

将第二文本段与存储在所述至少一个计算机可读存储介质中的第二兴趣点段索引匹配; 并且

使用第一兴趣点段索引和第二兴趣点段索引来标识与第一文本段和第二文本段这两者都匹配的一个或多个候选兴趣点条目。

9. 根据权利要求8所述的方法, 其中, 使用第一兴趣点段索引和第二兴趣点段索引来标识一个或多个候选兴趣点条目的动作包括以下动作:

使用第一兴趣点段索引来标识与第一文本段匹配的一个或多个兴趣点条目的第一集合;

使用第二兴趣点段索引来标识与第二文本段匹配的一个或多个兴趣点条目的第二集合; 并且

将在第一集合和第二集合这两者中都出现的一个或多个兴趣点条目标识为所述一个或多个候选兴趣点条目。

10. 根据权利要求8所述的方法, 其中, 所述至少一个计算机可读存储介质进一步存储语言模型, 所述语言模型包括与多个兴趣点段相关的统计信息, 并且其中, 所述方法进一步包括以下动作:

使用所述语言模型来从输入音频信号识别第一文本段和第二文本段。

11. 根据权利要求10所述的方法, 其中:

所述第一文本段包括所述多个兴趣点段中的第一兴趣点段, 所述第一兴趣点段对应于所述第一兴趣点段索引; 并且

所述第二文本段包括所述多个兴趣点段中的第二兴趣点段, 所述第二兴趣点段对应于所述第二兴趣点段索引。

12. 根据权利要求8所述的方法, 进一步包括以下动作:

将第一得分与第一候选兴趣点条目相关联, 所述第一得分指示第一候选兴趣点条目与第一文本段和第二文本段的相似性程度;

将第二得分与第二候选兴趣点条目相关联, 所述第二得分指示第二候选兴趣点条目与第一文本段和第二文本段的相似性程度; 并且

至少部分地基于第一得分和第二得分来对第一候选兴趣点条目和第二候选兴趣点条目进行排名。

13. 根据权利要求12所述的方法, 进一步包括以下动作:

至少部分地通过在文本上将第一文本段和第二文本段与第一候选兴趣点条目的兴趣

点名称进行比较来生成文本得分；

至少部分地通过将第一文本段和第二文本段的音标表示与第一候选兴趣点条目的兴趣点名称的音标表示进行比较来生成读音得分；并且

将第一得分生成为文本得分和读音得分的加权和。

14. 根据权利要求8所述的方法，其中，以编码的形式存储所述多个兴趣点段索引，并且其中，所述方法包括以下动作：

在使用第一兴趣点段索引和第二兴趣点段索引标识与第一文本段和第二文本段这两者都匹配的一个或多个候选兴趣点条目之前，对第一兴趣点段索引和第二兴趣点段索引进行解码。

15. 存储多个兴趣点段索引的至少一个计算机可读存储介质，所述至少一个计算机可读存储介质进一步存储将至少一个处理器编程为执行包括以下动作的方法的指令：

将第一文本段与存储在所述至少一个计算机可读存储介质中的第一兴趣点段索引匹配；

将第二文本段与存储在所述至少一个计算机可读存储介质中的第二兴趣点段索引匹配；并且

使用第一兴趣点段索引和第二兴趣点段索引来标识与第一文本段和第二文本段这两者都匹配的一个或多个候选兴趣点条目。

16. 根据权利要求15所述的至少一个计算机可读存储介质，其中，使用第一兴趣点段索引和第二兴趣点段索引来标识一个或多个候选兴趣点条目的动作包括以下动作：

使用第一兴趣点段索引来标识与第一文本段匹配的一个或多个兴趣点条目的第一集合；

使用第二兴趣点段索引来标识与第二文本段匹配的一个或多个兴趣点条目的第二集合；并且

将在第一集合和第二集合这二者中都出现的一个或多个兴趣点条目标识为所述一个或多个候选兴趣点条目。

17. 根据权利要求15所述的至少一个计算机可读存储介质，进一步存储语言模型，所述语言模型包括与多个兴趣点段相关的统计信息，其中，所述方法进一步包括以下动作：

使用所述语言模型来从输入音频信号识别第一文本段和第二文本段。

18. 根据权利要求17所述的至少一个计算机可读存储介质，其中：

所述第一文本段包括所述多个兴趣点段中的第一兴趣点段，所述第一兴趣点段对应于所述第一兴趣点段索引；并且

所述第二文本段包括所述多个兴趣点段中的第二兴趣点段，所述第二兴趣点段对应于所述第二兴趣点段索引。

19. 根据权利要求15所述的至少一个计算机可读存储介质，其中，所述方法进一步包括以下动作：

将第一得分与第一候选兴趣点条目相关联，所述第一得分指示第一候选兴趣点条目与第一文本段和第二文本段的相似性程度；

将第二得分与第二候选兴趣点条目相关联，所述第二得分指示第二候选兴趣点条目与第一文本段和第二文本段的相似性程度；并且

至少部分地基于第一得分和第二得分来对第一候选兴趣点条目和第二候选兴趣点条目进行排名。

20. 根据权利要求19所述的至少一个计算机可读存储介质,其中,所述方法进一步包括以下动作:

至少部分地通过在文本上将第一文本段和第二文本段与第一候选兴趣点条目的兴趣点名称进行比较来生成文本得分;

至少部分地通过将第一文本段和第二文本段的音标表示与第一候选兴趣点条目的兴趣点名称的音标表示进行比较来生成读音得分;并且

将第一得分生成为文本得分和读音得分的加权和。

用于兴趣点识别的系统和方法

背景技术

[0001] 一些导航系统(比如用在移动设备(例如,智能电话、平板计算机等)上的导航应用或车载导航系统)包括兴趣点类集。兴趣点(POI) 可以是用户可能希望导航到的任何地点。兴趣点的示例包括但不限于,餐馆、酒店、零售店、机场、火车站、公园、博物馆、加油站、工厂等。

[0002] 一些导航系统使得用户可以使用语音来搜索兴趣点。例如,用户可以说“Logan International Airport”。语音信号可以被麦克风捕捉,并且被导航系统处理,例如,通过将语音信号与兴趣点数据库中的条目匹配来进行处理。导航系统可以提示用户确认标识出的兴趣点确实是用户想要的,并且可以对该兴趣点设置路线。

发明内容

[0003] 本公开的各方面涉及用于兴趣点识别的系统和方法。

[0004] 根据一些实施例,提供了一种系统,该系统包括至少一个处理器和存储多个兴趣点段索引的至少一个计算机可读存储介质,其中,所述至少一个计算机可读存储介质进一步存储将所述至少一个处理器编程为执行以下动作的指令:将第一文本段与存储在所述至少一个计算机可读存储介质中的第一兴趣点段索引匹配;将第二文本段与存储在所述至少一个计算机可读存储介质中的第二兴趣点段索引匹配;并且使用第一兴趣点段索引和第二兴趣点段索引来标识与第一文本段和第二文本段这二者都匹配的一个或多个候选兴趣点条目。

[0005] 根据一些实施例,一种方法由包括至少一个处理器和存储多个兴趣点段索引的至少一个计算机可读存储介质的系统来执行,所述方法包括以下动作:将第一文本段与存储在所述至少一个计算机可读存储介质中的第一兴趣点段索引匹配;将第二文本段与存储在所述至少一个计算机可读存储介质中的第二兴趣点段索引匹配;并且使用第一兴趣点段索引和第二兴趣点段索引来标识与第一文本段和第二文本段这二者都匹配的一个或多个候选兴趣点条目。

[0006] 根据一些实施例,提供了存储多个兴趣点段索引的至少一个计算机可读存储介质,所述至少一个计算机可读存储介质进一步存储将至少一个处理器编程为执行包括以下动作的方法的指令:将第一文本段与存储在所述至少一个计算机可读存储介质中的第一兴趣点段索引匹配;将第二文本段与存储在所述至少一个计算机可读存储介质中的第二兴趣点段索引匹配;并且使用第一兴趣点段索引和第二兴趣点段索引来标识与第一文本段和第二文本段这二者都匹配的一个或多个候选兴趣点条目。

附图说明

[0007] 将参照附图来描述本公开的各种方面和实施例。

[0008] 图1示出根据一些实施例的说明性兴趣点识别系统100。

[0009] 图2示出根据一些实施例的说明性语音识别系统200。

[0010] 图3示出根据一些实施例的可以用于从未分段的兴趣点数据库构建编有索引(indexed)的兴趣点数据库的说明性过程300。

[0011] 图4示出根据一些实施例的说明性兴趣点识别系统400。

[0012] 图5示出根据一些实施例的用于将输入文本与一个或多个候选兴趣点条目匹配的说明性过程500。

[0013] 图6示意性地示出在其上可以实现本公开的一个或多个方面的说明性计算机1000。

具体实施方式

[0014] 本公开的各方面涉及用于兴趣点识别的技术。例如,提供了用于从由用户提供给导航系统的输入来识别兴趣点的技术。在一些实施例中,用户输入可以经由语音提供。然而,应意识到,本文所描述的技术不限于与任何特定类型的输入一起使用,因为在一些实施例中,可以使用这些技术中的一种或多种来对非语音输入(例如,手写、打字的文本等)进行处理。

[0015] 一些导航系统使用客户端-服务器架构。例如,客户端设备(例如,智能电话、由制造商合并到车辆的仪表板中的计算设备、由消费者安装在车辆中的计算设备等)可以捕捉用户输入,并且基于用户输入将请求发送到服务器计算机。服务器计算机可以对该请求进行处理,并且将响应提供给客户端设备,客户端设备可以继而基于从服务器计算机接收的响应将输出呈递给用户。

[0016] 相比之下,一些导航系统能够在不与任何服务器计算机通信的情况下执行兴趣点识别。例如,车载导航系统可以具有兴趣点条目的本地存储器,并且可以能够在本地执行自动语音识别(ASR)处理。

[0017] 客户端-服务器架构可以提供一些优点。例如,与客户端设备相比,服务器计算机可以访问更多资源,比如存储和/或处理能力。因此,服务器计算机可以能够执行更鲁棒(robust)的识别处理(例如,通过应用更复杂的语音识别技术和/或在更大的兴趣点数据库中搜索匹配)。然而,发明人已经认识和意识到,许多用户可能更喜欢本地解决方案。作为一个例子,由于隐私顾虑,一些用户可能更喜欢不将搜索项发送到服务器计算机。作为另一个例子,在网络连接不可用或质量低下的情况下(例如,当用户正在开车通过农村地区或隧道时),基于云的解决方案可能变得不可用。

[0018] 因此,在一些实施例中,可以提供不依赖于与任何服务器计算机的通信的兴趣点识别系统。例如,可以提供使用较少的存储和/或处理能力的改进的兴趣点识别技术。在一些实施例中,改进技术与常规技术相比可以使用少大约60%的存储。然而,应意识到,与服务器计算机的通信不一定排除在外,因为在一些实施例中,兴趣点识别系统可以以不同模式工作,比如在线模式和离线模式,在在线模式下,兴趣点系统发送请求,并且从服务器计算机接收对应的响应,在离线模式下,兴趣点识别系统在本地执行兴趣点识别。在一些实施例中,离线模式与在线模式相比可以提供少大约40%的延时。

[0019] 发明人已经认识到,一些国家或区域可以具有许多个兴趣点。例如,根据一些地图数据提供商,中国具有超过2-3千万个兴趣点。因此,如果每个兴趣点名称被看作可识别单词,则可能存在超过2-3千万个可识别单词。发明人已经认识和意识到,这样的大的词汇量

可以负面地影响兴趣点识别系统的性能,尤其是当在资源受约束的环境(例如,具有有限的处理器速度、存储器大小、存储器速度、高速缓存大小等)中操作时,该环境常见于移动设备(比如智能电话或车辆中的车载计算机)上。因此,在一些实施例,提供了用于高效地存储和搜索兴趣点条目的技术。

[0020] 发明人已经还认识和意识到现有兴趣点识别方法的一些缺点。例如,当用户以与兴趣点识别系统中表示兴趣点的方式不同的方式来标识兴趣点时,一些兴趣点识别系统可能性能不佳。作为例子,兴趣点识别系统可以包括由数据提供商(例如,专业地图提供商)编译和维护的兴趣点类集。在这样的类集中,波士顿的洛根机场(Logan Airport)可以在兴趣点条目中表示为“Boston Logan International Airport”。然而,用户在请求兴趣点信息时可能没有说出全名。例如,用户可能简单地说“Logan Airport”或“Boston Logan”。作为另一个例子,用户可能胡乱拼凑兴趣点名称中的单词(例如,因为用户不能记住或不确切地知道该名称在兴趣点条目中是如何表示的)。例如,不是说“the Mall at Chestnut Hill”(其可能是官方名称),用户而是可能说“Chestnut Hill Mall”。因为用户输入不是完全匹配兴趣点条目,所以即使当请求的兴趣点信息存在于系统中时,系统也可能无法返回该信息。因此,在一些实施例,可以提供对部分输入和/或错误输入更鲁棒的兴趣点识别系统。

[0021] 根据一些实施例,可以提供兴趣点条目类集,其中,每个兴趣点名称可以是分段的。例如,不是将整个短语“Boston Logan International Airport”存储为兴趣点名称,而是可以对短语进行分段,并且可以将所得到的段(例如,“Boston”|“Logan”|“International”|“Airport”)存储在兴趣点条目中。

[0022] 可以以任何合适的方式对兴趣点名称进行分段。例如,在单词边界由空格指示的语言(例如,英语、西班牙语、德语、法语等)中,可以仅基于找到空格的地方来对兴趣点名称进行分段。可替代地或另外地,可以使用粒度更细的或更粗的分段。作为一个例子,可以对复合词(例如,“airport”)进行分段,使得每个组成部分在单独的段(例如,“air”|“port”)中。作为另一个例子,可以将两个或更多个单词的搭配(例如,“opera house”)保存在一个段中。

[0023] 在单词边界没有明确指示的语言(例如,中文、日语、韩语、泰语等)中,可以适用合适的分段工具来对兴趣点名称进行分段。例如,可以将兴趣点名称“上海浦东国际机场”(“Shanghai Pudong International Airport”)分段为“上海”|“浦东”|“国际”|“机场”(“Shanghai”|“Pudong”|“International”|“Airport”)。

[0024] 根据一些实施例,兴趣点识别系统可以以编码的形式存储兴趣点名称的段。例如,条目“Boston City Hall”可以存储为<A,B,C>,其中,A、B和C分别是用于“Boston”、“City”和“Hall”的编码。以这种方式,兴趣点条目类集中的“Boston”的每一次出现都可以替换为编码A。同样地,“City”(分别地,“Hall”)的每一次出现都可以替换为B(分别地,C)。

[0025] 在一些实施例中,可以使用可变长度编码方法(例如,哈夫曼(Huffman)码),其中,更频繁出现的段可以具有比不太频繁出现的段短的编码。例如,单词“Boston”在兴趣点名称类集中可能频繁出现,可以使用短的位串作为用于“Boston”的编码。另一方面,单词“Logan”在兴趣点名称类集中可能很少出现,可以使用长的位串作为用于“Logan”的编码。如果使用可变长度编码方法来生成“Boston”的短编码A,则每次用编码A替换单词“Boston”

可以带来一定量的存储减少。因为“Boston”在兴趣点条目类集中频繁出现,所以可以通过累积许多少量减少来实现显著的整体节省。此外,通过将较短的编码分配给更频繁出现的段并且将较长的编码分配给不太频繁出现的段,通过更频繁出现的段实现的存储减少可以比抵消通过不太频繁出现的段引发的存储增加要更多。然而,应意识到,本公开的各方面不限于可变长度编码的使用,或者根本不限于任何编码。

[0026] 根据一些实施例,可以提供用于构建用于兴趣点识别的语言模型的技术。例如,语言模型可以包括用于将概率分配给单词序列的信息,其中,单词可以是兴趣点名称的一段,并且无需是整个兴趣点名称。语言模型可以是任何合适的类型,包括但不限于统计语法、n元模型等。

[0027] 在一些实施例中,可以使用分段的兴趣点名称的类集来训练语言模型。例如,兴趣点名称“Boston Logan International Airport”可以作为由单词“Boston”、“Logan”、“International”和“Airport”组成的训练句子而被处理。可以基于类集中的分段的兴趣点名称来计算转移概率(例如,观察到单词“Airport”跟随序列“Boston”、“Logan”、“International”的概率)。

[0028] 在一些实施例中,可以使用分段的兴趣点名称来创建用于自动语音识别(ASR)的上下文。例如,可以用读音信息来增强使用分段的兴趣点名称的类集训练的语言模型,以创建ASR上下文。在一些实施例中,ASR上下文可以将语言模型中的单词与读音信息相关联。例如,取决于第一个单词前面或后面的第二个单词,第一个单词的读音可以是不同的。作为一个例子,单词“Quincy”可以与两个不同的读音 /□kwnsi/和/□kwnzi/相关联。当后面是单词“Massachusetts”时,单词“Quincy”可能趋向于读作/□kwnzi/。相比之下,当后面是单词“Illinois”时,单词“Quincy”可能趋向于读作/□kwnsi/。可以使用录制音频的语料库来训练转移概率(例如,假定后面的单词是“Illinois”,单词“Quincy”读作/□kwnsi/的概率),或者可以从已建立的读音信息源获得转移概率。

[0029] 根据一些实施例,可以为兴趣点名称的段创建索引。例如,索引可以指示在其中找到该特定段的一个或多个兴趣点条目。作为例子,兴趣点名称类集可以包括以下条目:

[0030] 1.Boston City Hall

[0031] 2.Faneuil Hall

[0032] 3.Symphony Hall

[0033] 4.Boston Common

[0034] 在该例子中,可以将单词“Boston”的索引创建为指示“Boston”出现在条目1和4中。类似地,可以将单词“Hall”的索引创建为指示“Hall”出现在条目1-3中。如下面所说明的,可以使用这样的索引来促进兴趣点识别(例如,以改进对于部分输入和/或错误输入的鲁棒性)。

[0035] 在一些实施例中,兴趣点识别系统可以使用兴趣点段的索引来执行识别处理。例如,对于每个被识别的段,系统可以检索对应的索引,并且使用该索引来标识被识别的段出现在其中的兴趣点条目。因此,可以获得一个或多个兴趣点条目集合,其中,每个集合包括一个或多个兴趣点条目,并且对应于被识别的段。然后可以通过取这些组的交集来获得一个或多个候选兴趣点条目。

[0036] 作为例子,用户可能说“City Hall”,其可以被分割为两个单词序列<“City”,

“Hall”>。参照上面的例子,单词“City”的索引可以指示“City”出现在条目1中,而单词“Hall”的索引可以指示“Hall”出现在条目1-3中。通过取集合{1}和{1,2,3}的交集,系统可以确定条目1是候选匹配。以这种方式,部分输入(例如,“City Hall”,而不是全名“Boston City Hall”)可以被正确地识别。此外,即使段是由用户按不同的次序输入的(例如,“City Hall Boston”,而不是“Boston City Hall”),识别结果也会是相同的,因为集合交集运算是可交换的(communative)并且结合的(associative)。

[0037] 应意识到,上面介绍并且在下面更详细讨论的技术可以以许多方式中的任何一种方式实现,因为所公开的技术不限于任何特定的实现方式。附图中所示的并且在本文中所描述的例子仅仅是出于说明性的目的提供。

[0038] 图1示出了根据一些实施例的说明性兴趣点识别系统100。在该例子中,兴趣点识别系统100包括自动语音识别(ASR)引擎110、兴趣点识别组件120和兴趣点数据库130。

[0039] 说明性兴趣点识别系统100可以以任何合适的方式实现,例如,通过使用可执行指令编程的至少一个处理器和/或使用专门的硬件来实现。在一些实施例中,说明性兴趣点识别系统100可以在一个或多个车载设备(比如工厂安装的车载计算机)上实现。可替代地或另外地,所述一个或多个设备可以包括售后设备,或者仅仅是用户带来的移动设备。

[0040] 发明人已经认识和意识到,说明性兴趣点识别系统100可以在资源约束环境中实现。例如,在其上可以实现说明性兴趣点识别系统100的设备可以具有有大约1千兆字节(gigabyte)、2千兆字节、5千兆字节、10千兆字节、20千兆字节、50千兆字节、100千兆字节……的容量的存储器,并且可以具有有大约500兆赫、800兆赫、1千兆赫、2千兆赫、5千兆赫、10千兆赫、20千兆赫、50千兆赫、100千兆赫……的速度的存储器。然而,发明人已经认识和意识到,处理器和/或存储器可以不完全分配给识别处理,而是可以还用于其他功能,比如音乐回放、电话、全球定位系统(GPS)等。例如,在大约1千兆字节的存储器可用的情况下,只有大约300至400兆字节可以用于识别处理。在即将出现的资源密集的特征(例如,自动驾驶)的情况下,即使存储器大小为100千兆字节或更大和/或处理器速度为100千兆赫或更大,兴趣点条目的高效存储和搜索也会是有利的。

[0041] 在一些实施例中,ASR引擎110可以从用户接收语音输入。例如,用户可能说“浦东上海机场”(“Pudong Shanghai Airport”)。ASR引擎110可以对语音输入执行识别处理,并且将识别的文本输出到兴趣点识别组件120。在一些实施例中,ASR引擎110输出的识别的文本可以在提供给兴趣点识别组件120之前进行处理,例如,以移除无关的单词,比如“我想去(I want to go to)”、“我们正在去(We are going to)”、“导航到(Navigate to)”等。然而,这不是必需的,因为在一些实施例中,ASR引擎110可以被配置为从语音输入提取兴趣点名称,并且ASR引擎110输出的识别的文本可以直接提供给兴趣点识别组件120。

[0042] 在一些实施例中,兴趣点识别组件120可以在兴趣点数据库130中搜索与识别的文本匹配的一个或多个条目。发明人已经认识和意识到,在一些情况下,ASR引擎110输出的识别的文本可以是用户所说的查询的错误转录和/或不完整转录。因此,兴趣点识别组件120可能不能标识出兴趣点数据库130中的匹配条目。下面结合图4-图5来描述用于处理这样的错误的说明性技术。

[0043] 在一些实施例中,兴趣点识别组件120可以将识别的文本分段为输入段以促进对于兴趣点数据库130中的一个或多个匹配条目的搜索。例如,可以将识别的文本“浦东上海

机场”(“Pudong Shanghai Airport”) 分段为输入段“浦东”|“上海”|“机场”| (“Pudong”|“Shanghai”|“Airport”|)。对于适当的语言可以使用任何合适的分段技术,因为本公开的各方面不限于任何特定分段技术的使用。

[0044] 在一些实施例中,存储在兴趣点数据库130中的兴趣点名称可能已经进行了分段,例如通过使用与兴趣点识别组件120对识别的文本进行分段所用的技术类似的技术进行分段。除了分段的兴趣点名称之外,兴趣点数据库130可以存储至少一个段的索引,这些段出现在存储在兴趣点数据库130中的至少一个兴趣点名称中。

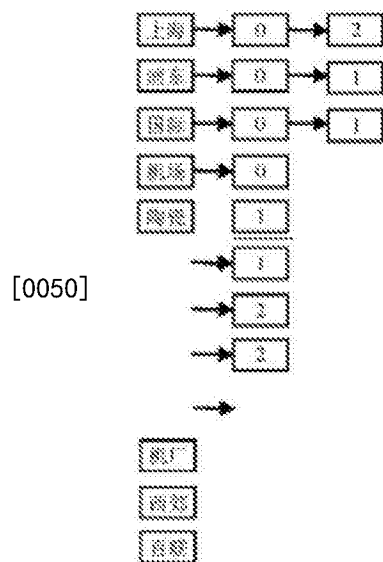
[0045] 例如,在一些实施例中,兴趣点数据库130可以包括以下说明性兴趣点条目。

[0046] -条目0:上海浦东国际机场(Shanghai Pudong International Airport)

[0047] -条目1:浦东国际陶瓷机厂(Pudong International Ceramic Factory)

[0048] -条目2:上海西郊百联(Shanghai Western Brilliance¹)

[0049] 在一些实施例中,可以将说明性条目分别分段为例如“上海|浦东|国际|机场”(“Shanghai|Pudong|International|Airport”)、“浦东|国际|陶瓷|机厂”(“Pudong|International|Ceramic|Factory”)、“上海|西郊|百联”(“Shanghai|Western|Brilliance”)。可以为每个段创建和存储索引,例如,以促进搜索。在一些实施例中,以下说明性索引可以存储在兴趣点数据库130中。



¹ 西郊百联是中国上海著名的购物中心

[0051] 在一些实施例中,索引的头节点可以是出现在存储在兴趣点数据库130中的至少一个兴趣点名称中的段,并且其余的节点可以记录该段出现在其中的条目。例如,上面的第一个说明性索引对应于单词“上海”(“Shanghai”),并且指示该单词出现在条目0和条目2中。

[0052] 在一些实施例中,存储在兴趣点数据库130中的索引可以根据一些合适的次序来排序。作为一个例子,每个头节点中的兴趣点名称段可以编码为数字,并且索引可以被排序为使得这些编码按照上升的次序或下降的次序。作为另一个例子,兴趣点名称段可以不被编码,并且索引可以被排序为使得兴趣点名称段按照字典式排序。例如,中文中的字符可以首先按读音(例如,基于拼音按字母顺序地)排序,然后按每个字符中的笔画的数量排序,或

者反过来。具有多个字符的段可以被排序为字符序列,其中第一个字符是最重要的。也可以使用另一合适的排序,因为本公开的各方面不限于任何特定排序的使用。

[0053] 发明人已经认识和意识到,对存储在兴趣点数据库130中的索引进行排序可以促进搜索。例如,给定输入段(例如,“浦东”或“Pudong”),可以使用高效的搜索算法(例如,二分搜索)来快速地标识具有与输入段匹配的头节点的索引(例如,上面的列表中的第二个说明性索引),并且继而可以使用该索引来标识输入段出现在其中的兴趣点条目(例如,条目0和条目1)。

[0054] 在一些实施例中,兴趣点识别组件120可以搜索存储在兴趣点数据库130中的索引,以标识从ASR引擎110输出的识别的文本获得的每个输入段的至少一个匹配索引。例如,输入段“浦东”(“Pudong”)、“上海”(“Shanghai”)和“机场”(“Airport”)可以分别与上面的列表中的第二个索引、第一个索引和第四个索引匹配。兴趣点识别组件120可以从兴趣点数据库130检索这些索引,并且使用这些索引来确定一个或多个候选兴趣点条目。

[0055] 1) 上面的列表中的第二个索引<“浦东”(“Pudong”),0,1> 可以指示目标兴趣点条目是条目0或是条目1,因为“浦东”(“Pudong”) 只出现在这些条目中。

[0056] 2) 上面的列表中的第一个索引<“上海”(“Shanghai”),0,2> 可以指示目标兴趣点条目是条目0或条目2,因为“上海”(“Shanghai”) 只出现在这些条目中。

[0057] 3) 上面的列表中的第四个索引<“机场”(“Airport”),0>可以指示目标兴趣点条目必须是条目0,因为“机场”(“Airport”) 只出现在条目0中。

[0058] 以这种方式,兴趣点识别组件120可以获得一个或多个兴趣点条目集合,每个集合包括一个或多个兴趣点条目,并且对应于输入段。例如,兴趣点识别组件120可以使用索引<“浦东”(“Pudong”), 0,1>来标识集合{条目0,条目1},该集合对应于输入段“浦东”(“Pudong”)。类似地,兴趣点识别组件120可以使用索引<“上海”(“Shanghai”),0,2>来标识集合{条目0,条目2},该集合对应于输入段“上海”(“Shanghai”),并且兴趣点识别组件120可以使用索引<“机场”(“Airport”),0>来标识集合{条目0},该集合对应于输入段“机场”(“Airport”)。

[0059] 在一些实施例中,兴趣点识别组件120可以取兴趣点条目集合的交集来确定一个或多个候选兴趣点条目。例如,兴趣点识别组件120 可以取集合{条目0,条目1}、{条目0,条目2}和{条目0}的交集,这些集合分别是基于输入段“浦东”(“Pudong”)、“上海”(“Shanghai”)和“机场”(“Airport”)获得的。这些集合的交集可以只包括一个条目,即,条目0,并且该条目可以作为兴趣点识别结果返回。该结果可以提供给用户进行确认,和/或提供给导航系统以使得导航系统可以相应地设置路线。

[0060] 在一些实施例中,兴趣点识别组件120可以不对从ASR引擎110 输出的识别的文本获得的每一个输入段来检索对应的索引。例如,在上面的例子中,索引<“浦东”(“Pudong”),0,1>和<“上海”(“Shanghai”),0,2>可以足以使候选兴趣点条目池变窄到一个候选,即,条目0。因此,兴趣点识别组件120可以停止,而不检索“机场”(“Airport”)的索引,这可以改进兴趣点识别系统100的响应时间。

[0061] 发明人已经认识和意识到,上述说明性技术对用户所做的一些类型的错误可以是鲁棒的。例如,在上面的例子中,用户提供了不完整的兴趣点名称,其中单词“国际”(“International”)缺失,因为整个兴趣点名称为“上海浦东国际机场”(“Shanghai Pudong

International Airport”)。此外,用户记录了两个段,即,“浦东上海”(“Pudong Shanghai”),与“上海浦东”(“Shanghai Pudong”)相反。尽管有这些错误,兴趣点识别组件120仍可能能够正确地将语音输入与兴趣点条目“上海浦东国际机场”(“Shanghai Pudong International Airport”)匹配。在一些实施例中,通过使用本文所描述的某些技术,错误率可以降低多于50%。

[0062] 尽管各种例子在上面是结合图1描述的,但是应意识到,这样的例子仅仅是出于例示说明的目的提供的。例如,本公开的各方面不限于使用两个分开的组件来实现语音识别和兴趣点识别,因为在一些实施例中,单个组件可以执行两个功能。

[0063] 图2示出了根据一些实施例的说明性语音识别系统200。在该例子中,语音识别系统200包括自动语音识别(ASR)引擎210,ASR引擎210可以被配置为使用语言模型240和/或ASR上下文215来执行语音识别处理。在一些实施例中,自动语音识别(ASR)引擎210 可以用在兴趣点识别系统(例如,图1所示的说明性兴趣点识别系统 100)中。

[0064] 说明性语音识别系统200可以以任何合适的方式实现,例如,使用用可执行指令编程的至少一个处理器和/或使用专门的硬件来实现。在一些实施例中,说明性语音识别系统200可以在车载设备上实现。该设备可以是工厂安装的车载计算机。可替代地或另外地,该设备可以是售后设备,或者仅仅是用户带来的移动设备。

[0065] 在一些实施例中,语言模型240和ASR上下文215中的一个或两个可以使用分段的兴趣点数据库230来构建,分段的兴趣点数据库230 又可以使用未分段的兴趣点数据库220来构建。在一些实施例中,未分段的兴趣点数据库220和/或分段的兴趣点数据库230可以存储在实现语音识别系统200的设备外部的地点处,或者可以在用于生成语言模型240和/或ASR上下文215之后根本不存储。作为结果,语音识别系统200使用的存储量可以减少。

[0066] 发明人已经认识和意识到,可以使用分段来缩小用于语音和/或兴趣点识别系统的词汇量。例如,在像中国的大国家中,可能存在超过 2-3千万个兴趣点。如果每个兴趣点名称都被看作可识别单词,则可能存在超过2-3千万个可识别单词。发明人已经认识和意识到,每个兴趣点名称可以是一个或多个段的组合,并且可能存在数量少得多的可能的段(例如,一百万或两百万不同的段)。因此,通过将每个段(而不是整个兴趣点名称)看作可识别词,词汇量可以显著地减少(例如,从数千万个单词减少到几百万个单词)。

[0067] 在一些实施例中,可以基于一个或多个因素(比如与语音和/或兴趣点识别系统预期在其中操作的环境相关联的约束)的任何合适组合来标识期望的词汇量。这样的约束的例子包括但不限于处理器速度、存储器大小、存储器速度等。一旦被标识,可以通过调整分段的粒度水平来实现期望的词汇量。例如,在一些实施例中,可以使用迭代过程,其中,在每次迭代中,可以使用某个粒度水平来对兴趣点名称进行分段,并且根据所得到的词汇量是太大还是太小,可以提高或降低粒度水平。可以重复这样的迭代,直到实现期望的词汇量为止。

[0068] 在单词边界由空格指示的语言(例如,英语、西班牙语、德语、法语等)中,可以仅基于找到空格的地方来对兴趣点名称进行分段。可替代地或另外地,可以使用粒度更细的或更粗的分段,例如用来实现如上所述的期望的词汇量。作为一个例子,可以对复合词(例如,“airport”)进行分段,使得每个组成部分在单独的段(例如,“air” | “port”)中。作为另一个例子,可以将两个或更多个单词的搭配(例如,“opera house”)保持在一个段中。

[0069] 在单词边界没有明确指示的语言(例如,中文、日语、韩语、泰语等)中,可以适用合适的分段工具来对兴趣点名称进行分段。例如,可以将兴趣点名称“上海浦东国际机场”(“Shanghai Pudong International Airport”)分段为“上海”|“浦东”|“国际”|“机场”(“Shanghai”|“Pudong”|“International”|“Airport”)。

[0070] 可以使用任何一种或多种合适的技术来执行分段。例如,在一些实施例中,可以使用交互式过程来训练分段模型,该分段模型可以是基于条件随机场(CRF)、隐马尔可夫模型(HMM)等的分段模型。例如,可以使用加标签训练集合来构建分段模型,该分段模型然后可以用于对未加标签数据的集合进行分段。一个或多个错误可以由人标记,并且用于改动分段模型。可以重复该过程,直到实现某个程度的准确度。

[0071] 在一些实施例中,加标签训练集合可以包括被分割为分别加有标签“Beginning”(开始)、“Middle”(中间)和“End”(结束)的三个段的兴趣点名称。例如,可以将兴趣点名称“上海浦东国际机场”(“Shanghai Pudong International Airport”)分段为“上海”|“浦东”|“国际”|“机场”(“Shanghai”|“Pudong”|“International”|“Airport”),其中,段“上海”(“Shanghai”)可以加标签“Beginning”,段“浦东”和“国际”(“Pudong”和“International”)可以加标签“Middle”,段“机场”(“Airport”)可以加标签“End”。然而,应意识到,本公开的各方面不限于任何特定的标签集合的使用或任何特定水平的分段粒度。

[0072] 再次参照图2,在一些实施例中,可以使用合适的分段模型来对未分段的兴趣点数据库220中的兴趣点名称进行分段,并且可以使用所得的分段的兴趣点数据库230来构建语言模型240。语言模型240可以包括指示在分段的兴趣点数据库230中观察到某些段序列的频率的统计信息。例如,搭配“上海”|“浦东”(“Shanghai”|“Pudong”)可以比搭配“上海”|“西郊”(“Shanghai”|“Western”)更频繁地在分段的兴趣点数据库230中出现。作为结果,假定前一个段是“上海”(“Shanghai”),则语言模型240可以将高于“西郊”(“Western”)的转移概率分配给“浦东”(“Pudong”)。

[0073] 在一些实施例中,可以使用分段的兴趣点数据库230和/或语言模型240来构建ASR上下文215。例如,可以为语言模型240增补读音信息来创建ASR上下文215。可替代地或另外地,可以使用来自分段的兴趣点数据库230的一个或多个兴趣点名称、连同相关联的读音信息来创建ASR上下文215。ASR上下文215可以是基于语法的上下文或另一合适类型的上下文。

[0074] 在一些实施例中,ASR上下文215可以包括指示单词读音可以如何取决于周围单词而不同地发音的音标(phonetic)转移概率。例如,单词“Quincy”可以与两个不同的读音/□kwnsi/和/□kwnzi/相关联。当后面是单词“Massachusetts”时,单词“Quincy”可能趋向于读作 /□kwnzi/。相比之下,当后面是单词“Illinois”时,单词“Quincy”可能趋向于读作/□kwnsi/。因此,ASR上下文215可以根据哪个单词被发现在“Quincy”的后面(例如,“Massachusetts”和“Illinois”)来将不同的概率与“Quincy”的不同读音相关联。可以使用录制音频的语料库来训练这样的音标转移概率,或者可以从已建立的读音信息源获得这样的音标转移概率。

[0075] 发明人已经认识和意识到使用分段的兴趣点数据库创建用于语音识别的语言模型和/或ASR上下文的各种优点。例如,如上面所讨论的,使用兴趣点名称的段(而不是整个兴趣点名称)创建的语言模型和/或ASR上下文可以具有缩小的词汇量,并且这样可以占据

较少的存储。此外,使用从兴趣点名称创建的语言模型和/或ASR上下文(而不是通用的语言模型和/或ASR上下文)可以改进语音识别准确度(例如,通过消除作为可能的识别结果的、与兴趣点识别系统交互的用户不太可能说的单词序列)。然而,应意识到,本公开的各方面不限于使用(分段的或未分段的)兴趣点数据库来创建语言模型或ASR上下文。

[0076] 在一些实施例中,ASR引擎210可以使用语言模型240和/或ASR上下文215来对从用户捕捉的语音进行处理。例如,ASR引擎210可以使用语言模型240和/或ASR上下文215来将语音输入与最有可能的声音序列匹配,并且与最有可能的声音序列相对应的单词序列可以作为识别的文本被输出。在一些实施例中,ASR引擎210可以输出包括分别与n个最有可能的声音序列相对应的n个单词序列的n个最佳结果,并且每个这样的单词序列可以与指示对应的声音序列与语音输入的匹配程度的置信得分相关联。

[0077] 尽管各种例子在上面是结合图2描述的,但是应意识到这样的例子仅仅是出于示例说明的目的提供的。例如,本公开的各方面不限于将ASR上下文实现为与ASR引擎分开的模块,因为在一些实施例中 ASR上下文可以合并到ASR引擎中。此外,在一些实施例中,结合图2描述的技术中的一种或多种可以用于识别除了兴趣点查询之外的语音输入。例如,可以对除了兴趣点名称之外的术语(例如,医学术语)的数据库进行分段,并且使用该数据库来创建语言模型和/或ASR上下文。更进一步地,在一些实施例中,兴趣点数据库的分段、语言模型的创建和/或ASR上下文的创建可以由与执行语音和/或兴趣点识别的系统不同的系统来执行。例如,兴趣点数据库的分段、语言模型的创建和/或ASR上下文的创建可以由兴趣点识别软件的供应商执行,并且分段的兴趣点数据库、语言模型和/或ASR上下文可以加载到执行语音和/或兴趣点识别的系统(例如,集成到车辆中的计算机或移动电话)中。

[0078] 图3示出了根据一些实施例的可以用于从未分段的兴趣点数据库构建编有索引的兴趣点数据库的说明性过程300。例如,过程300可以用于构建图1所示的说明性兴趣点数据库130。在一些实施例中,过程300可以在离线阶段期间例如由兴趣点识别软件的供应商执行。所得的编有索引的兴趣点数据库可以加载到用于兴趣点识别的设备(例如,集成到车辆中的计算机或移动电话)中。

[0079] 在动作310,可以从未分段的兴趣点数据库(比如图2所示的未分段的兴趣点数据库220)检索一个或多个兴趣点名称。可以使用任何一种或多种合适的技术(包括但不限于上面结合图2描述的那些)来对这一个或多个兴趣点名称进行分段。在一些实施例中,可以对未分段的兴趣点数据库中的所有兴趣点名称进行分段。然而,这不是必需的,因为在一些实施例中,可以对一些兴趣点名称(例如,没有超过某个阈值长度的兴趣点名称)不进行分段。

[0080] 在一些实施例中,可以将分段的兴趣点名称存储在分段的兴趣点数据库(比如图2所示的说明性兴趣点数据库230)中。这样的分段的兴趣点数据库可以既用于生成图1所示的说明性兴趣点数据库130(其用于执行兴趣点识别),又用于生成图2所示的说明性语言模型240和/或说明性ASR上下文215(其用于执行语音识别)。然而,应意识到,本公开的各方面不限于使用相同的分段的兴趣点数据库来进行语音识别和兴趣点识别。在一些实施例中,可以使用通用语言模型和/或通用ASR上下文来执行语音识别。

[0081] 在动作320,可以为出现在至少一个兴趣点名称中的段生成索引,例如如上面结合图1描述的那样。例如,在一些实施例中,未分段的兴趣点数据库可以包括以下说明性兴趣

点条目。

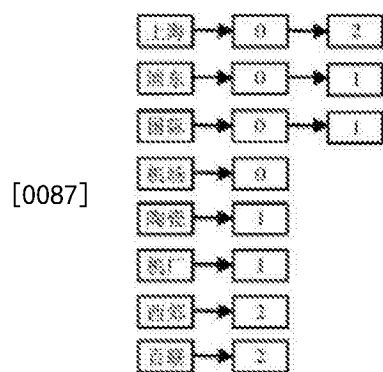
[0082] 一条目0:上海浦东国际机场(Shanghai Pudong International Airport)

[0083] 一条目1:浦东国际陶瓷机厂(Pudong International Ceramic Factory)

[0084] 一条目2:上海西郊百联(Shanghai Western Brilliance)

[0085] 可以将这些条目分别分段为例如“上海|浦东|国际|机场 (“Shanghai|Pudong|International|Airport”)”、“浦东|国际|陶瓷|机厂” (“Pudong|International|Ceramic|Factory”)、以及“上海|西郊|百联” (“Shanghai|Western|Brilliance”)。

[0086] 在一些实施例中,如下所示,可以为每个段创建索引。每个索引可以包括具有一个或多个节点的列表。对应的段(例如,“上海”或“Shanghai”)可以存储在头节点处,而每个其余的节点可以存储该段出现在其中的兴趣点条目(例如,对于段“上海”或“Shanghai”,为条目0和条目2)的标识符。然而,应意识到,本公开的各方面不限于将索引存储为列表,因为另一类型的数据结构(例如,二叉树)也可以用于存储指示对应段出现在其中的一个或多个兴趣点条目的信息。



[0088] 在动作330,可以对一个或多个索引进行编码,例如,以减少用于存储这一个或多个索引的空间量。发明人已经认识和意识到,可变长度编码可以用于实现显著的存储节省。例如,在一些实施例中,较短的编码可以用于在许多兴趣点条目中出现的段,因此具有大索引,而较长的编码可以用于只在一个或几个条目中出现的段,因此具有小索引。可以使用任何合适的可变长度编码方案,包括但不限于哈夫曼码。

[0089] 例如,设 $I_0, \dots, I_7$ 表示上面的八个说明性索引,这些索引分别对应于段“上海” (“Shanghai”)、“浦东” (“Pudong”)、“国际” (“International”)、“机场” (“Airport”)、“陶瓷” (“Ceramic”)、“机厂” (“Factory”)、“西郊” (“Western”)和“百联” (“Brilliance”)。索引 I_0 包括两个条目(条目0和条目2),而索引 I_7 只包括一个条目(条目2)。因此,在一些实施例中,较短的编码可以用于“上海” (“Shanghai”),而较长的编码可以用于“百联” (“Brilliance”)。

[0090] 在一些实施例中,可以使用delta(增量)编码方法来对索引中的一个或多个兴趣点条目标识符进行编码。发明人已经认识和意识到,例如当兴趣点数据库包括大量条目(例如,数百万或数千万个)时, delta编码对于减少用于存储索引的空间量可以是有效的。例如,发明人已经认识和意识到,随着兴趣点数据库的大小增长,每个条目的标识符(例如,自动生成的数据库主键)的长度会相应地增长。因此,在兴趣点数据库中可能频繁出现的像“上海” (“Shanghai”)这样的段的索引可以包括兴趣点条目标识符的长列表,其中,每个标识符可以是一个大数。发明人已经认识和意识到,可以使用delta编码方法来减少为这样的

索引存储的信息量。

[0091] 为了例示说明的目的,假定段(例如,“上海”或“Shanghai”)的索引包括以下兴趣点条目标识符:

[0092] $\dots, 1000000, 1000024, 1000031, \dots$

[0093] 不是存储这些大数中的每个,而是可以存储起始点,比如1000000。对于每个随后的标识符,可以存储该标识符和前一标识符之间的差(或增量(delta))。因此,在该例子中,可以改为存储以下内容:

[0094] $\dots, 1000000, 24, 7, \dots$

[0095] 在解码过程期间,可以通过将24加到1000000来恢复标识符 1000024,可以通过将7加到1000024来恢复标识符1000031,依此类推。发明人已经认识和意识到,可以通过将大数(例如,1000024、1000031等)替换为小数(例如,24、7等)来实现显著的存储节省。

[0096] 即使在解码期间可能需要附加处理时间(例如,以累加几个增量值来恢复标识符),发明人也已经认识和意识到这样的延迟可能不会显著地影响用户体验。例如,在一些实施例中,解码可能是在兴趣点应用加载到存储器中时执行的,所以当启动应用时,用户可能体验到一些延迟。经解码的索引可以保持在存储器中,使得当对用户所说的兴趣点查询进行处理时,可能不需要解码。然而,应意识到,本公开的各方面不限于执行在最前面执行解码,因为在一些实施例中,可以按需执行解码,或者可以采用混合方法(例如,在最前面对更频繁遇到的段的索引进行解码,并且按需对不太频繁遇到的段的索引进行解码)。

[0097] 在动作340,可以存储一个或多个编码的索引,例如,存储在编有索引的兴趣点数据库中。在一些实施例中,可以根据某个合适的排序来对存储的索引进行排序。例如,可以(例如通过使用如上面所讨论的可变长度编码方案)将每个头节点中的兴趣点名称段编码为数字,并且可以对索引进行排序以使得段的编码按照上升次序。发明人已经认识和意识到,以这种方式对索引进行排序可以促进搜索。例如,给定输入段(例如,“浦东”或“Pudong”),可以计算输入段的编码,并且可以使用高效的搜索算法(例如,二分搜索)来快速地标识具有与编码匹配的头节点的索引。然而,应意识到,本公开的各方面不限于存储经排序的索引,因为在一些实施例中,可以在索引被解码并且被加载到存储器中时(例如,当兴趣点识别系统被用户启动时)执行排序。

[0098] 在一些实施例中,除了兴趣点名称的段的索引之外或作为其替代,还可以存储兴趣点条目表。例如,设 $E_0, \dots, E_7$ 分别表示对八个段“上海”(“Shanghai”)、“浦东”(“Pudong”)、“国际”(“International”)、“机场”(“Airport”)、“陶瓷”(“Ceramic”)、“机厂”(“Factory”)、“西郊”(“Western”)和“百联”(“Brilliance”)进行编码的结果。可以生成以下条目并且将这些条目存储在编有索引的兴趣点数据库中。

[0099] 一条目0: $E_0|E_1|E_2|E_3$

[0100] 一条目1: $E_1|E_2|E_4|E_5$

[0101] 一条目2: $E_0|E_6|E_7$

[0102] 因此,在该例子中,段“上海”(“Shanghai”)的每次出现可以替换为对应编码 E_0 ,并且对于其他段是同样的。如果使用可变长度编码方法来生成“上海”(“Shanghai”)的短编码,则段“上海”(“Shanghai”)替换为编码 E_0 的每次替换可以带来一定量的存储减少。因为“上海”(“Shanghai”)在许多兴趣点条目中出现,所以可以通过累积许多少量减少来实现显

著的整体节省。

[0103] 此外,发明人已经认识和意识到,通过将较短的编码分配给更频繁出现的段并且将较长的编码分配给不太频繁出现的段,通过更频繁出现的段实现的存储减少可以比抵消通过不太频繁出现的段引发的存储增加要更多。例如,段“百联”(“Brilliance”)可能只出现在一个或几个兴趣点条目中。即使将段“百联”(“Brilliance”)替换为编码E₇可能带来一定量的存储增加,这样的增加也可能只发生一次或几次。因此,由对不太频繁出现的段(像“百联”(“Brilliance”))使用较长的编码引起的整体增加可以被通过对更频繁出现的段(像“上海”(“Shanghai”))使用较短的编码实现的整体减少抵消。

[0104] 尽管各种例子在上面是结合图3描述的,但是应意识到这样的例子仅仅是出于例示说明的目的提供的。例如,虽然发明人已经认识和意识到将可变长度编码应用于兴趣点名称的段的各种优点,但是本公开的各方面不限于此。在一些实施例中,除了可变长度编码之外或作为其替代,可以使用一种或多种其他类型的编码,或者可以根本不使用编码。此外,本公开的各方面不限于使用十进制数作为兴趣点条目标识符,因为在一些实施例中,可以使用其他值,包括但不限于位串、字符串、十六进制数等。

[0105] 图4示出了根据一些实施例的说明性兴趣点识别系统400。兴趣点识别系统400可以接收输入文本,并且尝试将输入文本与兴趣点数据库420中的一个或多个兴趣点条目匹配。输入文本可以是例如由图2所示的说明性ASR引擎210从用户说话识别的。然而,在一些实施例中,兴趣点识别系统400可以可替代地或附加地用于对来自另一个源的(例如,用户打入的、从手写识别的、通过网络接收的、等等)输入文本进行处理。

[0106] 在一些实施例中,兴趣点数据库420可以包括分段的兴趣点名称。这些段可以例如如上面结合图3描述的那样编索引和/或编码。然而,应意识到,本公开的各方面不限于对兴趣点名称进行分段或者对段编索引或编码。例如,可以在使用未分段的兴趣点数据库情况下应用本文所描述的用于将输入文本与一个或多个兴趣点条目匹配的技术。

[0107] 在一些实施例中,兴趣点识别系统400可以使用如上面结合图1描述的技术中的一种或多种技术来标识在文本上与输入文本匹配的一个或多个兴趣点条目。可替代地或另外地,兴趣点识别系统400可以生成输入文本的音标表示。例如,在图4所示的例子中,兴趣点识别系统400包括文本到读音转换组件430,文本到读音转换组件430可以被编程为对输入文本进行处理并且输出输入文本的音标表示。例如,中文的输入文本可以包括中文字符串(例如,“浦东机厂”)。文本到读音转换组件430可以将每个字符映射到一些适当的系统(比如拼音)中的音标表示(例如,“pu”对“浦”、“dong”对“东”、“ji”对“机”、“chang”对“厂”)。兴趣点识别系统400然后可以在兴趣点数据库420中搜索具有匹配读音的兴趣点名称(例如,在没有声调的情况下,“pu dong ji chang”,或者在有声调的情况下,“pu-3 dong-1 ji-1 chang-3”)。例如,在一些实施例中,兴趣点数据库420可以存储兴趣点名称的段的音标表示,并且可以(例如通过使用32位循环冗余校验)对音标表示进行编码和/或排序以促进搜索(例如,以使得可以使用二分搜索)。

[0108] 在一些实施例中,兴趣点识别系统400可以标识多个候选兴趣点条目。例如,在中文中,字符“场”(如在意指“Airport”的“机场”中)可以具有与字符“厂”(如在意指“Factory”的“机厂”中)相同的读音。因此,下面两个条目都是输入文本(例如,“浦东机厂”)的候选。

[0109] -条目0:上海浦东国际机场(Shanghai Pudong International Airport)

[0110] -条目1:浦东国际陶瓷机厂(Pudong International Ceramic Factory)

[0111] 在图4所示的例子中,兴趣点识别系统400包括兴趣点候选评分组件450,兴趣点候选评分组件450可以被编程为对多个候选兴趣点条目进行评分和/或排名。例如,评分组件450可以将较高得分分配给作为输入文本“浦东机厂”的匹配的条目1,因为条目1在文本和读音上都与输入文本匹配。相比之下,评分组件450可以将较低得分分配给上面的条目0,因为条目0在读音上与输入文本匹配,但是在一个字符中存在错配(即,“场”而不是“厂”)。虽然如此,两个条目都呈现给用户(例如,条目1首先呈现,因为条目1收到了较高得分)。以这种方式,即使用户实际上说“浦东机场”(“Pudong Airport”),但是ASR引擎将语音输入错误地识别为“浦东机厂”(“Pudong Factory”),兴趣点识别系统400也可以能够将预期的兴趣点标识为候选。

[0112] 在一些实施例中,评分组件450可以被编程为使用历史信息来调整分配给候选兴趣点条目的得分。例如,评分组件450可以访问搜索历史数据库460,搜索历史数据库460可以包括与特定用户相关的历史信息和/或与用户群体相关的历史信息。作为一个例子,历史信息可以指示群体中的用户搜索“上海浦东国际机场”(“Shanghai Pudong International Airport”)比“浦东国际陶瓷机厂”(“Pudong International Ceramic factory”)更频繁和/或时间上更近。因此,评分组件450可以将比后者高的得分分配给前者。作为另一个例子,历史信息可以指示发出查询的用户搜索“上海浦东国际机场”(“Shanghai Pudong International Airport”)不如“浦东国际陶瓷机厂”(“Pudong International Ceramic factory”)频繁和/或时间上近。因此,评分组件450可以将比后者低的得分分配给前者。在一些实施例中,评分组件450可以对特定于发出查询的用户的信息给予更大权重。然而,这不是必需的,因为在一些实施例中,评分组件450可以改为对群体信息给予更大权重。

[0113] 在一些实施例中,评分组件450可以被编程为使用上下文信息来调整分配给候选的得分。例如,评分组件450可以被编程为使用上下文信息来对发出兴趣点查询的用户进行分类。分类结果然后可以用于调整分配给候选兴趣点条目的得分。作为一个例子,评分组件450可以被编程为使用上下文信息来确定用户有可能是行人。响应于确定用户有可能是行人,评分组件450可以将较高得分分配给离用户当前地点在步行距离内的兴趣点。

[0114] 作为另一个例子,评分组件450可以被编程为使用上下文信息来确定用户有可能是司机。响应于确定用户有可能是司机,评分组件450 可以将较低得分分配给汽车不太能访问的兴趣点(例如,对私家车封闭的街道,或者众所周知难以找到停车点的地方)。评分组件450可以查阅任何合适的上下文信息源,包括但不限于搜索历史(例如,用户是否频繁地选择步行和/或公共交通作为搜索选项)、地点跟踪(例如,用户的当前移动是否与用户步行和/或使用公共交通一致)、设备标识(例如,接收的查询是否指示与移动电话(而不是集成到车辆中的设备)一致的设备类型、操作系统、用户代理等)等。

[0115] 在一些实施例中,评分组件450可以被编程为使用文本相似性和/或读音相似性来将得分分配给候选兴趣点条目。例如,在图4所示的例子中,说明性兴趣点识别系统400包括文本层模糊匹配组件410,文本层模糊匹配组件410可以被编程为为候选兴趣点条目计算一个或多个文本相似性得分,该一个或多个文本相似性得分指示候选兴趣点条目与输入文本在文本上的相似性程度。另外地或可替代地,说明性兴趣点识别系统400包括读音层模糊

匹配组件440,读音层模糊匹配组件440可以被编程为对候选兴趣点条目计算指示候选兴趣点条目与输入文本在读音上的相似性的一个或多个读音相似性得分。

[0116] 在一些实施例中,评分组件450可以组合文本层模糊匹配组件410 输出的一个或多个文本相似性得分和读音层模糊匹配组件440输出的一个或多个读音相似性得分。例如,评分组件450可以计算文本相似性得分和读音相似性得分的加权和。发明人已经认识和意识到,一些语言(例如,中文)可以具有许多同音异义词,并且因此涉及同音异义词的ASR错误可能是常见的。因此,在一些实施例中,对于具有许多同音异义词的语言(例如,中文),可以给予读音相似性比文本相似性大的权重,以便改进对于识别错误的鲁棒性。

[0117] 在一些实施例中,文本层模糊匹配组件410可以通过在文本上将输入文本与候选兴趣点条目进行比较来生成文本相似性得分。例如,文本层模糊匹配组件410可以被编程为如下基于输入文本和兴趣点名称之间的编辑距离度量来生成文本相似性得分。

[0118] $\text{text_sim}(\text{输入文本}, \text{POI名称})$

[0119] $= 1 - \text{edit_dist}(\text{输入文本}, \text{POI名称}) / \text{max_length}(\text{输入文本}, \text{POI名称})$

[0120] 作为编辑距离度量的一个例子,输入文本“Boston Logan Airport”和候选兴趣点条目“Boston Logan International Airport”之间的 Levenshtein距离可以为1,因为单次编辑(例如,将“International”插入在“Logan”和“Airport”之间)足以将输入文本“Boston Logan Airport”转换为候选兴趣点条目“Boston Logan International Airport”。作为编辑距离度量的另一个例子,输入文本“City Hall Boston”和候选兴趣点条目“Boston City Hall”之间的 Damerau-Levenshtein距离可以为2,因为将输入文本“City Hall Boston”转换为候选兴趣点条目“Boston City Hall”需要至少两次编辑(例如,调换“Boston”,然后调换“Boston”和“City”,或者在末尾删除“Boston”,并且在开头添加“Boston”)。另外地或可替代地,可以使用一个或多个其他的度量(例如,基于删除、插入、替换和/或调换字符而不是单词的度量),因为本公开的各方面不限于使用任何特定的度量。

[0121] 在一些实施例中,在生成文本相似性得分时,文本层模糊匹配组件410可以区分在某个词汇表中出现的文本段(例如,每个均在至少一个兴趣点条目中出现的段)和在该词汇表中没有出现的文本段。例如,可以如下计算输入文本和候选兴趣点条目之间的文本相似性,其中,LCS表示最长公共子序列的程度,M表示每个均出现在至少一个兴趣点条目中的文本段中的字符的数量,N表示文本段中的没有出现在任何兴趣点条目中的字符的数量。

[0122] $(\text{LCS}(\text{输入文本}, \text{POI名称}) - M) / N$

[0123] 例如,文本层模糊匹配组件410可以对输入文本“中国农民银行”(“Chinese Farmer Bank”)进行处理,并且确定段“中国”(“Chinese”)和“银行”(“Bank”)中的每个均出现在一个或多个兴趣点条目中,但是段“农民”(“Farmer”)没有出现在任何兴趣点条目中。因此,可以如下计算输入文本“中国农民银行”(“Chinese Farmer Bank”)和候选兴趣点条目“中国农业银行”(“Chinese Agricultural Bank”)之间的文本相似性。

[0124] $(\text{LCS}(\text{“中国农业银行”}, \text{“中国农民银行”}) - M) / N$

[0125] $= (5 - 4) / 2$

[0126] $= 0.5$

[0127] 尽管本文描述了用于测量文本相似性的各种技术,但是应意识到这样的技术仅仅

是说明性的。本公开的各方面不限于任何特定的测量文本相似性的方式或者使用文本相似性来将输入文本与一个或多个兴趣点条目匹配。可替代地或另外地,可以基于读音相似性来将输入文本与一个或多个兴趣点条目匹配。

[0128] 如上面所讨论的,说明性兴趣点识别系统400的文本到读音转换组件430可以被编程为生成输入文本的音标表示。在一些实施例中,音标表示可以包括音节(syllable)序列,其中,每个音节可以包括音素(phoneme)序列,每个音素可以包括元音或辅音。另外,每个音节可以包括一个或多个标注,比如指示该音节的音调的标注。例如,输入文本“中国龙夜银行”(意指“Chinese Dragon Night Bank”,其有可能包括一个或多个转录错误)可以具有以下音标表示。

[0129] zhong-1 guo-2 long-2 ye-4 yin-2 hang-2

[0130] 另一方面,候选兴趣点条目“中国农业银行”(“Chinese Agricultural Bank”)可以具有以下音标表示。

[0131] zhong-1 guo-2 nong-2 ye-4 yin-2 hang-2

[0132] 在该例子中,输入文本的起始段“中国”(“Chinese”)与候选兴趣点条目的起始段是相同的,输入文本的最后一个段“银行”(“Bank”)与候选兴趣点条目的最后一个段是相同的。输入文本的第四个字符“夜”(“Night”)具有与候选兴趣点条目的第四个字符“业”(“Industry”)相同的读音。输入文本的第三个字符“龙”(“Dragon”)具有与候选兴趣点条目的第三个字符“农”(“Agriculture”)类似的、但不相同的读音——“long-2”对“nong-2”,唯一差别在于辅音“l”对“n”。

[0133] 因此,上面的序列中的六个位置中的五个具有相同的读音。每个这样的位置的相似性得分可以为1。对于第三个位置,“long-2”对“nong-2”,相似性得分可以为0.75。因此,可以如下计算输入文本“中国龙夜银行”和候选兴趣点条目“中国农业银行”之间的模糊最长公共子序列(fLCS)的程度。

[0134] $1+1+0.75+1+1+1=5.75$

[0135] 在一些实施例中,说明性兴趣点识别系统400的读音层模糊匹配组件440可以如下计算读音相似性,其中,fLCS表示模糊最长公共子序列的程度,M表示每个均出现在至少一个兴趣点条目中的文本段中的字符的数量,N表示文本段中的没有出现在任何兴趣点条目中的字符的数量。

[0136] $(\text{fLCS}(\text{输入本文的音标表示}, \text{POI名称的音标表示}) - M) / N$

[0137] 在上面的例子中,段“中国”(“Chinese”)和“银行”(“Bank”)均出现在一个或多个兴趣点条目中,但是段“龙夜”(“Dragon Night”)没有出现在任何兴趣点条目中。因此,可以如下计算读音相似性。

[0138] $(\text{fLCS}(\text{输入本文的音标表示}, \text{POI名称的音标表示}) - M) / N$

[0139] $= (5.75 - 4) / 2$

[0140] $= 0.875$

[0141] 可以使用一种或多种技术的任何合适的组合来计算两个音标表示之间的相似性程度,因为本公开的各方面不限于此。例如,在一些实施例中,可以如下基于两个音节A和B的辅音之间的相似性程度以及A和B的元音之间的相似性程度来计算A和B之间的相似性程度。

[0142] $(\text{sim_con}(\text{A.辅音}, \text{B.辅音}) + \text{sim_vow}(\text{A.元音}, \text{B.元音})) / 2$

[0143] 可以以任何合适的方式定义两个辅音之间的相似性程度,并且对于两个元音之间的相似性程度是同样的。例如,相同辅音之间的相似性程度可以为1,两个可高度混淆的辅音之间的相似性程度(例如,“l”对“n”、“s”对“sh”、“b”对“p”等)可以为0.5,两个可中等混淆的辅音之间的相似性程度(例如,“s”对“z”、“s”对“th”等)可以为0.25,等等。同样地,相同元音之间的相似性程度可以为1,两个可高度混淆的元音之间的相似性程度(例如,“fit”中的“i”对“feet”中的“ee”、“ban”中的“an”对“bang”中的“ang”、“sin”中的“in”对“sing”中的“ing”等)可以为0.5,两个可中等混淆的元音之间的相似性程度(例如,“hot”中的“o”对“hut”中的“u”、“bad”中的“a”对“bed”中的“e”等)可以为0.25,等等。

[0144] 发明人已经认识和意识到,混淆性可能取决于一个或多个因素而变化,这些因素包括但不限于所用的特定的ASR引擎、特定的语言和 /或口音、特定的说话者等。因此,在一些实施例中,辅音和/或元音的分组和/或对于不同组的赋值可以基于测试数据。另外地或可替代地,可以为某些音节对(例如,“wang”对“huang”、“wa”对“hua”、“wu”对“hu”、“wen”对“hun”等)提供一个或多个特殊的规则。

[0145] 在一些实施例中,说明性兴趣点识别系统400所接收并处理的识别的文本可以包括语音识别系统(例如,图2所示的说明性语音识别系统200)输出的n个最佳结果(针于某个合适的n)。n个最佳结果可以包括一个或多个单词的n个序列,其中,每个序列是用户说话的可能匹配。兴趣点识别系统400可以对这n个序列中的一些或全部进行处理以标识可能匹配的兴趣点条目。然而,应意识到,本公开的各方面不限于从语音识别系统接收n个最佳结果,因为在一些实施例中,一个或多个单词的单个序列可以作为输入提供给兴趣点识别系统400。

[0146] 在一些实施例中,对于n个最佳结果中的每个序列,说明性兴趣点识别系统400可以将一个或多个兴趣点候选标识为可能匹配该序列。例如,评分组件450可以被编程为维护具有相应得分的兴趣点候选的列表。给定针对n个最佳结果中的第i个序列的候选,可以如下计算得分,其中,wf是适当的加权函数,sim_score是(例如,按照如上面所讨论的文本相似性和读音相似性的加权和计算得到的)候选和第i个序列之间的相似性。

[0147] $\text{candidate_score}(\text{候选}, n\text{个最佳结果中的第}i\text{个序列})$

[0148] $= \text{sim_score}(\text{候选}, n\text{个最佳结果中的第}i\text{个序列}) * \text{wf}(i) / (\text{wf}(1) + \dots + \text{wf}(n))$

[0149] 如果兴趣点条目是n个最佳结果中的多个序列的候选,则该兴趣点条目的得分可以是 $\text{candidate_score}(\text{兴趣点条目}, n\text{个最佳结果中的第}i\text{个序列})$ 在对于其来说该兴趣点条目是候选的所有i值上的总和。

[0150] 加权函数wf可以以任何合适的方式选择。例如,在一些实施例中,加权函数可以选自一组合适的函数,包括但不限于以下函数。

[0151] $\text{wf}(i) = 1/i$

[0152] $\text{wf}(i) = 1/2^i$

[0153] $\text{wf}(i) = (n-i+1)/n$

[0154] $\text{wf}(i) = 1$

[0155] 例如,可以将这些函数中的每个应用于测试数据,并且可以选择具有最高准确度(例如,最高F得分)的函数。然而,应意识到,本公开的各方面不限于任何特定的用于选择加

权函数的方式或者根本不使用任何加权函数。

[0156] 在一些实施例中,兴趣点识别系统400可以使用评分组件450计算出的得分来对候选兴趣点条目进行排名,并且针对某个合适的n输出n个最佳结果(这个n可以与兴趣点识别系统400作为输入接收的一个或多个单词的序列的数量是相同的或不同的)。这些得分可以但无需连同n个最佳结果一起输出。在一些实施例中,n可以等于1,在这种情况下,兴趣点识别系统400可以输出单个兴趣点候选。

[0157] 在一些实施例中,兴趣点识别系统400可以基于相应得分来向用户(例如在视觉上、听觉上等等)呈现一个或多个候选兴趣点条目(例如,得分按下降的次序以使得最佳匹配首先呈现)。兴趣点识别系统400可以但无需将一次呈现给用户的候选兴趣点条目的数量限制为一个条目、两个条目、三个条目等。这可以减轻可能正在步行或开车的用户上的认知负荷。

[0158] 尽管上面结合图4描述了各种例子,但是应意识到这样的例子仅仅是出于例示说明的目的提供的。

[0159] 图5示出了根据一些实施例的用于将输入文本与一个或多个候选兴趣点条目匹配的说明性过程500。例如,说明性过程500可以由兴趣点识别系统(例如,图1所示的说明性兴趣点识别系统100和/或图4所示的说明性兴趣点识别系统400)执行以对从用户接收的兴趣点查询进行处理。

[0160] 在动作510,可以以某种合适的方式(比如通过使用本文所描述的分段技术中的一种或多种)对输入文本进行分段。例如,可以将输入文本“西郊百联商场”(“Western Brilliance Shopping Mall”)分段为三个段,“西郊”|“百联”|“商场”(“Western”|“Brilliance”|“Shopping Mall”)。

[0161] 在动作520,可以为在动作510标识的至少一个段检索索引。例如,在按编码的形式对兴趣点名称的段进行排序的实施例中,可以对在动作510标识的段进行编码,并且可以使用所得的编码来在经编码的段的列表中搜索匹配。

[0162] 在一些实施例中,为段检索到的索引可以为编码的形式(例如,已经使用delta编码方案进行了编码)。可以对这样的索引进行解码以恢复对应段出现在其中的兴趣点条目的一个或多个标识符。然而,应意识到,本公开的各方面不限于索引的编码和后续解码,因为在一些实施例中,可以在不对索引进行编码的情况下存储索引,以使得可以不执行解码。

[0163] 在一些实施例中,对于标识出的段可能没有找到对应的索引,这可能表明该段没有出现在任何已知的兴趣点条目中。然而,在一些实施例中,如上面结合图4所讨论的,在评估输入文本和候选兴趣点条目之间的相似性(例如,文本相似性和/或读音相似性)时可以考虑这样的段。

[0164] 在一些实施例中,对于其找到索引的段可以被放置在第一列表中,而对于其没有找到索引的段可以被放置在第二列表中。在动作530,可以确定是否存在这样的至少一个兴趣点条目,即,第一列表中的所有段都出现在该兴趣点条目中。例如,可以对第一列表中的每个段标识一个或多个兴趣点条目的集合(例如,包括为该段检索出的索引中的所有兴趣点条目),并且可以取所有这样的集合的交集。

[0165] 如果交集为非空,则可以在动作540将交集中的一个或多个兴趣点条目作为候选输出。否则,在动作535,可以从第一列表移除至少一个段,并且将这些段放置到第二列表

中,并且过程500可以返回到动作530以取与第一列表中的段相对应的所有集合的交集。因为至少一个段已经从第一列表移除,所以交集可能变为非空。如果是这样,则过程500可以继续进行到动作540。否则,过程500可以再次继续进行到动作535以移除至少一个其他的段。可以重复该操作,直到交集变为非空为止。

[0166] 可以使用任何合适的技术来选择将从第一列表移除的一个或多个段。例如,在一些实施例中,可以使用一种或多种统计技术来对兴趣点数据库(例如,图2所示的说明性分段的兴趣点数据库230)进行分析并且基于信息内容来对兴趣点名称的段进行评分。例如,很少出现的段可以被看作具有比频繁出现的段更高的信息量。因此,可以在动作535移除具有最低出现频率的段。

[0167] 可替代地或另外地,类别单词可以被移除(例如,“Hotel”、“Supermarket”等),而名称可以被保留(例如,“Sheraton”、“Carrefour”等)。例如,说出输入文本“西郊百联商场”(“Western Brilliance Shopping Mall”)的用户可能意图搜索“上海西郊百联购物中心”(“Shanghai Western Brilliance Shopping Center”)。输入文本“西郊百联商场”(“Western Brilliance Shopping Mall”)可能一开始导致空的交集,因为可能不存在三个段“西郊”(“Western”)、“百联”(“Brilliance”)和“商场”(Shopping Mall)都出现在其中的条目。通过移除类别单词“商场”(Shopping Mall)并且只取分别与段“西郊”(“Western”)和“百联”(“Brilliance”)相对应的两个候选集合的交集,可以得到非空交集,该非空交集可以包括预期的兴趣点条目“上海西郊百联购物中心”(“Shanghai Western Brilliance Shopping Center”)。

[0168] 尽管上面结合图5描述了各种例子,但是应意识到这样的例子仅仅是出于例示说明的目的提供的。例如,本公开的各方面不限于按编码的形式对兴趣点名称的段进行排序,因为在一些实施例中,可以按解码的形式对兴趣点名称的段进行排序,并且可以使用在动作510标识出的段来在不首先对该段进行编码的情况下标识段列表中的匹配。

[0169] 此外,在一些实施例中,可以将来自不同地理区域(例如,不同国家、省、城市等)的兴趣点条目编译到单独的数据库中。以这种方式,可以在任何给定时间将较少量的信息(例如,只有一个数据库)保持在存储器中。在一些这样的实施例中,如果过程500未能标识任何候选兴趣点条目,则可以将目前加载到存储器中的数据库移到高速缓存中,并且可以加载不同的数据库,并且可以使用新加载的数据库来执行过程500。可以作为将来自第一列表的段移到第二列表以获得可能非空的交集的补充或代替来进行该操作。

[0170] 图6示意性地示出了在其上可以实现本公开的任何方面的说明性计算机100。例如,图1-图2和图4所示的说明性组件中的任何一个或多个(例如,ASR引擎110、兴趣点识别组件120和/或兴趣点数据库130)可以在计算机1000上实现。

[0171] 如本文所使用的,“移动设备”可以是足够小以使得它可以构建到或安装到车辆中或者被用户携带的任何计算设备。移动设备的例子可以包括但不限于集成到车辆中的计算设备、移动电话、寻呼机、便携式媒体播放器、电子书阅读器、手持游戏机、个人数字助理(PDA)和平板电脑。在一些情况下,移动设备的重量可以最多是1磅、1.5磅或2磅,和/或移动设备的最大尺寸可以最多为6英寸、9英寸或1英尺。另外,移动设备可以包括使得用户能够在不同地点使用该设备的特征。例如,移动设备可以包括功率储存(例如,蓄电池),以使得移动设备可以在不插入到电源插座中的情况下使用达到某个持续时间,或者可以依赖

于车辆的蓄电池。作为另一个例子,移动设备可以包括被配置为在不物理地连接到网络连接点的情况下提供网络连接的无线网络接口。

[0172] 在图6所示的实施例中,计算机1000包括处理单元1001和非暂时性计算机可读存储介质1002,处理单元1001具有一个或多个处理器,非暂时性计算机可读存储介质1002可以包括例如易失性和/或非易失性存储器。存储器1002可以存储将处理单元1001编程为执行本文所描述的功能中的任何一个的一个或多个指令。除了存储器1002之外,计算机1000还可以包括其他类型的非暂时性计算机可读介质,比如储存器1005(例如,一个或多个盘驱动器)。储存器1005还可以存储可以加载到1002中的一个或多个应用程序和/或应用程序所用的资源(例如,软件库)。

[0173] 计算机1000可以具有一个或多个输入设备和/或输出设备,比如图6所示的设备1006和1007。除了其他方面,这些设备可以用于呈现用户界面。可以用于提供用户界面的输出设备的例子包括用于呈现视觉输出的打印机或显示屏幕以及用于呈现听觉输出的扬声器或其他声音生成设备。可以用于用户界面的输入设备的例子包括键盘和指点设备,比如鼠标、触控板和数字化平板。作为另一个例子,输入设备 1007可以包括用于捕捉音频信号的麦克风,输出设备1006可以包括用于视觉地渲染识别的文本的显示屏幕和/或用于听觉地渲染识别的文本的扬声器。

[0174] 如图6所示,计算机1000还可以包括一个或多个网络接口(例如,网络接口1010)以使得能够经由各种网络(例如,网络1020)通信。网络的例子包括局域网或广域网,比如企业网或互联网。这样的网络可以基于任何合适的技术,可以根据任何合适的协议操作,并且可以包括有线网络、无线网络或光纤网络。

[0175] 如此描述了至少一个实施例的几个方面,要意识到的是,各种变化、修改和改进是本领域技术人员将容易想到的。这样的变化、修改和改进意图在本公开的精神和范围内。因此,前面的描述和附图仅仅是作为例子。

[0176] 本公开的上述实施例可以以许多方式中的任何一种方式实现。例如,实施例可以使用硬件、软件或它们的组合来实现。当用软件实现时,软件代码可以在任何合适的一个处理器或一批处理器上执行,不管是在单个计算机中提供的,还是分布在多个计算机之间的。

[0177] 此外,本文所概述的各种方法或过程可以编码为软件,该软件可以在利用各种操作系统或平台中的任何一个的一个或多个处理器上执行。另外,这样的软件可以使用若干合适的程序设计语言和/或程序设计或脚本撰写工具中的任何一个来编写,并且还可以编译为可执行机器语言代码或在框架或虚拟机上执行的中间代码。

[0178] 在这方面,本文所公开的构思可以实施为编码有一个或多个程序的一个非暂时性计算机可读介质(或多个计算机可读介质)(例如,计算机存储器、一个或多个软盘、紧凑盘、光学盘、磁带、闪存、现场可编程门阵列或其他半导体器件中的电路配置、或其他非暂时性有形计算机存储介质),所述一个或多个程序当在一个或多个计算机或其他处理器上执行时执行实现上面所讨论的本公开的各种实施例的方法。所述一个或多个计算机可读介质可以是可运输的,以使得存储在其上的一个程序或多个程序可以加载到一个或多个不同的计算机或其他的处理器上以实现如上面所讨论的本公开的各种方面。

[0179] 术语“程序”或“软件”在本文中用于指代可以用于将计算机或其他处理器编程为实现如上面所讨论的本公开的各种方面的任何类型的计算机代码或计算机可执行指令集。

另外,应意识到,根据该实施例的一个方面,当被执行时执行本公开的方法的一个或多个计算机程序无需驻留在单个计算机或处理器上,而是可以以模块化的方式分布在若干不同的计算机或处理器之间来实现本公开的各种方面。

[0180] 计算机可执行指令可以为许多形式,比如一个或多个计算机或其他设备执行的程序模块。一般来说,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等。通常,在各种实施例中,程序模块的功能性可以根据需要组合或分布。

[0181] 此外,数据结构可以以任何合适的形式存储在计算机可读介质中。为简化例示说明,数据结构可能被示为在数据结构中具有通过位置关连的字段。这样的关系可以同样地通过在计算机可读介质中为具有地点的字段分配传达字段之间的关系的关系的存储来实现。然而,可以使用任何合适的机制来建立数据结构的字段中的信息之间的关系,包括通过使用指针、标签或建立数据元素之间的关系的其他机制。

[0182] 本公开的各种特征和方面可以独立使用、按两个或更多个的任何组合使用、或者按前面描述的实施例中沒有具体讨论的各种布置使用,因此在其应用上不限于前面的描述中阐述的或在附图中例示说明的组件的细节和布置。例如,一个实施例中描述的各方面可以按任何方式与其他实施例中描述的各方面组合。

[0183] 此外,本文所公开的构思可以实施为方法,已经提供了该方法的例子。作为该方法的一部分执行的动作可以以任何合适的方式排序。因此,可以构造动作按不同于所示次序的次序执行的实施例,这些实施例可以包括同时执行一些动作,即使这些动作在说明性实施例中所示为顺序的动作。

[0184] 权利要求中用于修饰权利要求元素的比如“第一”、“第二”、“第三”等的序数词的使用独自并不暗示一个权利要求元素优于另一个权利要求元素的任何优先权、优先级或次序或者执行方法的动作的时间次序,而是仅用作区分具有某个名称的一个权利要求元素和具有相同名称(但是使用序数词)的另一个元素的标签来区分权利要求元素。

[0185] 此外,本文所使用的措辞和术语是出于描述的目的,而不应被认为是限制。本文中“包括”、“包含”、“具有”、“涉及”及其变型的使用意在于包含其后列出的各项及其等同物以及附加项。

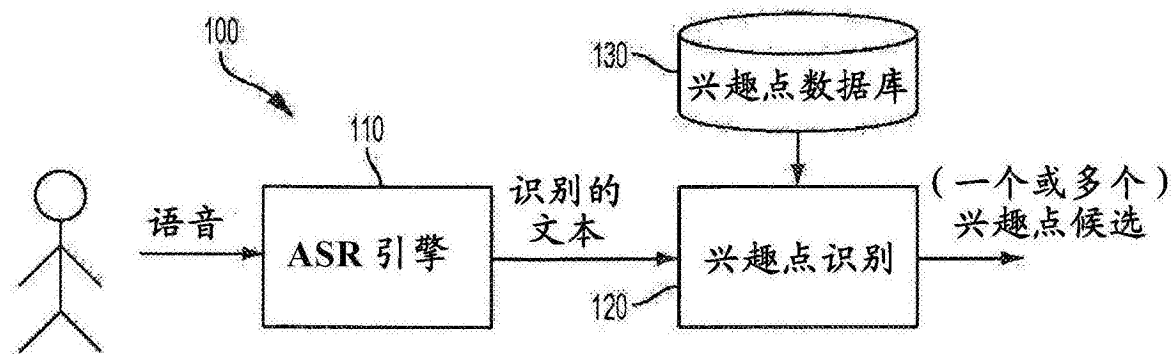


图1

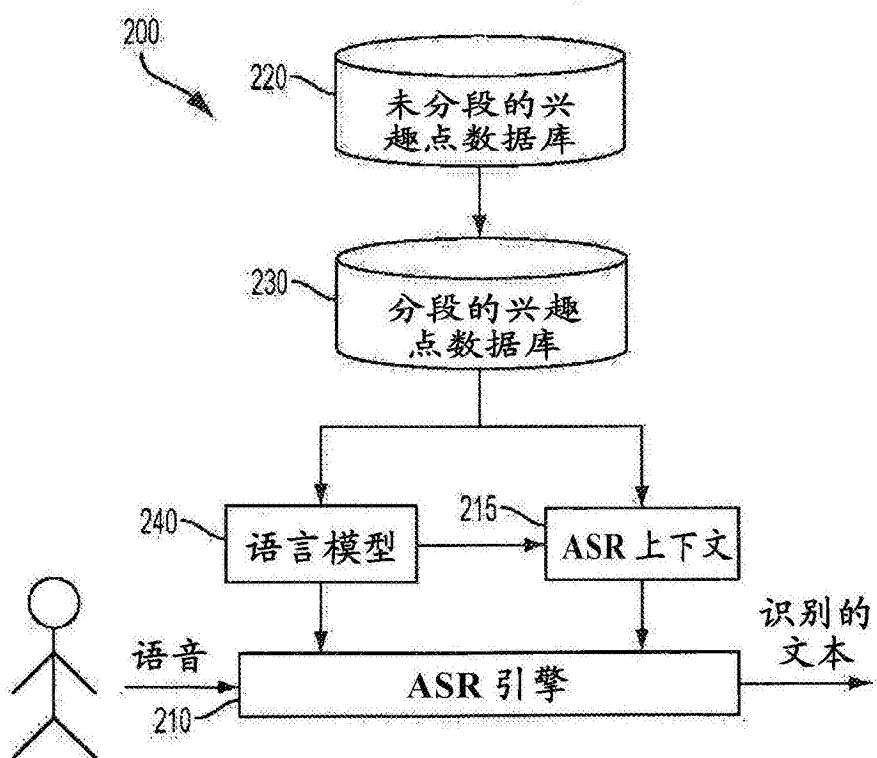


图2

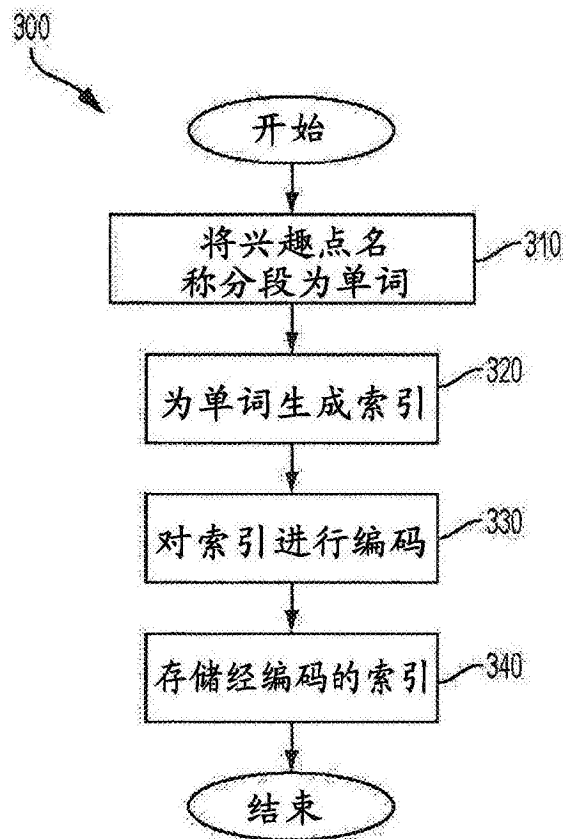


图3

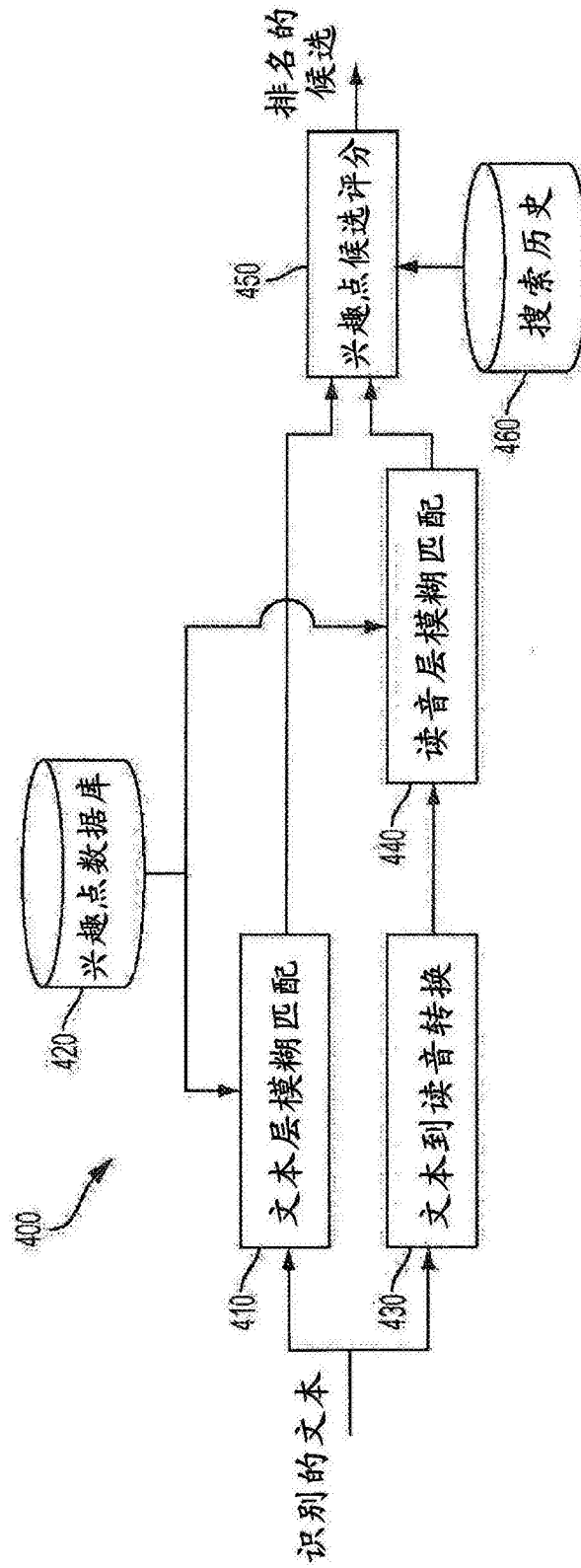


图4

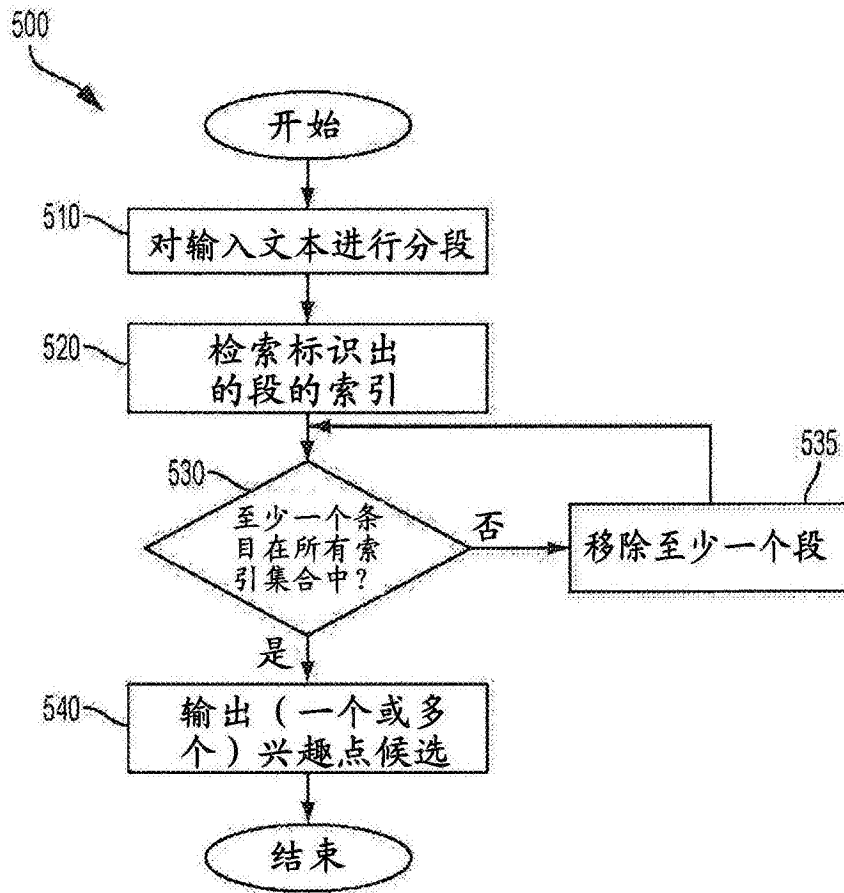


图5

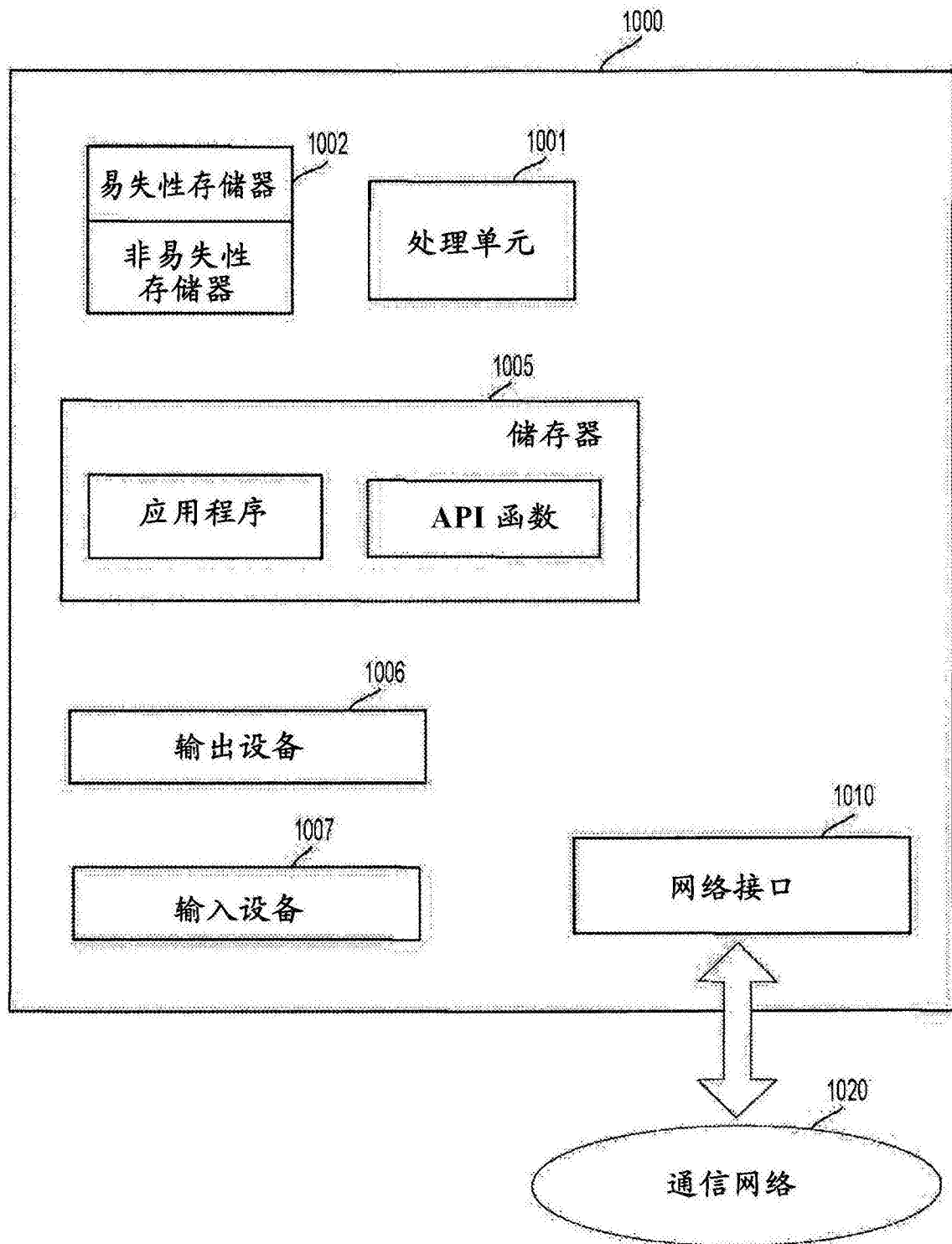


图6