(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 113506567 A
(43)申请公布日 2021.10.15

(21)申请号 202110625294.7

(22)申请日 2009.03.06

(30)优先权数据
12/044,310 2008.03.07 US

(62)分案原申请数据
200980115721.4 2009.03.06

(71)申请人 谷歌有限责任公司
地址 美国加利福尼亚州

(72)发明人 大卫·P·辛格尔顿
德巴基特·高什

(74)专利代理机构 中原信达知识产权代理有限责任公司 11219
代理人 周亚荣 邓聪惠

(51)Int.Cl.
G10L 15/183(2013.01)
G01C 21/36(2006.01)
G10L 15/08(2006.01)
G10L 15/19(2013.01)
G10L 15/22(2006.01)
G10L 15/30(2013.01)
H04W 4/02(2018.01)

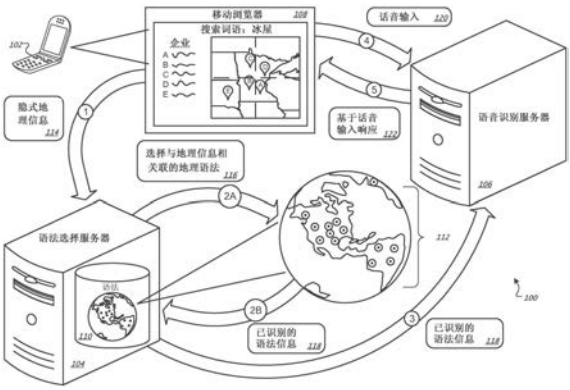
权利要求书9页 说明书10页 附图5页

(54)发明名称

基于场境的语音识别语法选择方法和系统

(57)摘要

本公开涉及基于场境的语音识别语法选择方法和系统。除其它外，本说明书的主题可以实现在包括以下的方法中：接收从与第一计算设备相关联的非语言用户行为取得的地理信息。非语言用户行为暗示用户对地理位置感兴趣。该方法还包括使用取得的地理信息来识别与该地理位置相关联的语法，以及输出语法指示符以供在选择所识别的语法对来自用户的话音输入进行语音识别处理中使用。



CN 113506567 A

1.一种计算机实现的方法,包括:

由计算设备确定所述计算设备的场境;

由所述计算设备接收来自用户的第一输入;

基于来自所述用户的第一输入和所述计算设备的所述场境,由所述计算设备选择来自词语组的词语子集;

由所述计算设备接收来自所述用户的第二输入;以及

由所述计算设备基于来自所述用户的第二输入和来自所述词语组的所述词语子集来提供输出。

2.如权利要求1所述的方法,其中:

确定所述计算设备的所述场境包括:

确定所述计算设备的位置,以及

来自所述词语组的所述词语子集包括在所述计算设备附近的位置名称。

3.如权利要求1所述的方法,其中,来自所述词语组的所述词语子集是由话音识别器使用的语法。

4.如权利要求1所述的方法,其中:

接收来自所述用户的所述第二输入包括:

接收对所述词语子集中的词语的选择,以及

基于来自所述用户的第二输入和来自所述词语组的所述词语子集来提供所述输出包括:

基于对所述词语子集中的所述词语的选择来提供所述输出。

5.如权利要求1所述的方法,其中,接收来自所述用户的所述第一输入包括接收指示所述用户将提供作为话语的所述第二输入的所述第一输入。

6.如权利要求1所述的方法,其中,基于来自所述用户的第二输入和来自所述词语组的所述词语子集来提供所述输出包括:

提供地图,所述地图包括与所述第二输入和所述词语子集相关联的地理位置的视觉指示符。

7.如权利要求1所述的方法,包括:

通过接收由所述用户说出的话语的音频数据来接收来自所述用户的所述第二输入;

由所述计算设备通过使用所述词语子集作为语法对所述音频数据执行自动话音识别来生成所述话语的转录;以及

通过以下来提供所述输出:基于所述话语的所述转录来提供所述输出。

8.一种系统,包括:

一个或多个计算机;以及

一个或多个计算机和一个或多个存储指令的存储设备,所述指令在由所述一个或多个计算机执行时,能够操作以使所述一个或多个计算机执行操作,所述操作包括:

由计算设备确定所述计算设备的场境;

由所述计算设备接收来自用户的第一输入;

基于来自所述用户的第一输入和所述计算设备的所述场境,由所述计算设备选择来自词语组的词语子集;

由所述计算设备接收来自所述用户的第二输入;以及

由所述计算设备基于来自所述用户的第二输入和来自所述词语组的所述词语子集来提供输出。

9.如权利要求8所述的系统,其中:

确定所述计算设备的所述场境包括:

确定所述计算设备的位置,以及

来自所述词语组的所述词语子集包括在所述计算设备附近的位置名称。

10.如权利要求8所述的系统,其中,来自所述词语组的所述词语子集是由话音识别器使用的语法。

11.如权利要求8所述的系统,其中:

接收来自所述用户的所述第二输入包括:

接收对所述词语子集中的词语的选择,以及

基于来自所述用户的第二输入和来自所述词语组的所述词语子集来提供所述输出包括:

基于对所述词语子集中的所述词语的选择来提供所述输出。

12.如权利要求8所述的系统,其中,接收来自所述用户的所述第一输入包括接收指示所述用户将提供作为话语的所述第二输入的所述第一输入。

13.如权利要求8所述的系统,其中,基于来自所述用户的第二输入和来自所述词语组的所述词语子集来提供所述输出包括:

提供地图,所述地图包括与所述第二输入和所述词语子集相关联的地理位置的视觉指示符。

14.如权利要求8所述的系统,其中,所述操作包括:

通过接收由所述用户说出的话语的音频数据来接收来自所述用户的所述第二输入;

由所述计算设备通过使用所述词语子集作为语法对所述音频数据执行自动话音识别来生成所述话语的转录;以及

通过以下来提供所述输出:基于所述话语的所述转录来提供所述输出。

15.一种存储软件的非暂时性计算机可读介质,所述软件包括能够由一个或多个计算机执行的指令,所述指令在这样的执行时,使所述一个或多个计算机执行操作,所述操作包括:

由计算设备确定所述计算设备的场境;

由所述计算设备接收来自用户的第一输入;

基于来自所述用户的第一输入和所述计算设备的所述场境,由所述计算设备选择来自词语组的词语子集;

由所述计算设备接收来自所述用户的第二输入;以及

由所述计算设备基于来自所述用户的第二输入和来自所述词语组的所述词语子集来提供输出。

16.如权利要求15所述的介质,其中:

确定所述计算设备的所述场境包括:

确定所述计算设备的位置,以及

来自所述词语组的所述词语子集包括在所述计算设备附近的位置名称。

17.如权利要求15所述的介质,其中,来自所述词语组的所述词语子集是由话音识别器使用的语法。

18.如权利要求15所述的介质,其中:

接收来自所述用户的所述第二输入包括:

接收对所述词语子集中的词语的选择,以及

基于来自所述用户的第二输入和来自所述词语组的所述词语子集来提供所述输出包括:

基于对所述词语子集中的所述词语的选择来提供所述输出。

19.如权利要求15所述的介质,其中,基于来自所述用户的第二输入和来自所述词语组的所述词语子集来提供所述输出包括:

提供地图,所述地图包括与所述第二输入和所述词语子集相关联的地理位置的视觉指示符。

20.如权利要求15所述的介质,其中,所述操作包括:

通过接收由所述用户说出的话语的音频数据来接收来自所述用户的所述第二输入;

由所述计算设备通过使用所述词语子集作为语法对所述音频数据执行自动话音识别来生成所述话语的转录;以及

通过以下来提供所述输出:基于所述话语的所述转录来提供所述输出。

21.一种计算机实现的方法,包括:

提供地理区域的地图以供在计算设备上显示;

基于在所述计算设备上显示的所述地理区域的所述地图来选择用于话音识别的语法;

在所述计算设备显示所述地理区域的所述地图时,接收与话语相对应的语音数据以发起搜索;

(i)通过使用基于在所述计算设备上显示的所述地理区域的所述地图所选择的所述语法来对所述话语执行话音识别并且(ii)基于从先前提供用于显示的内容得出的地理信息,生成所述话语的转录;以及

更新所述地图以提供识别与所述话语的所述转录相关联的位置的视觉指示符。

22.如权利要求21所述的方法,其中,所述转录是进一步基于从所述地理区域的所述地图得出的地理信息来生成的。

23.如权利要求21所述的方法,其中,所述话语识别位置。

24.如权利要求21所述的方法,其中,所述转录是进一步基于识别一个或多个搜索查询的搜索历史来生成的,所述一个或多个搜索查询由与所述计算设备的用户相关联的一个或多个计算设备提交至一个或多个搜索服务。

25.如权利要求21所述的方法,其中,所述语法是基于地理位置来选择的,所述地理位置是基于与先前所提交的一个或多个搜索查询相关联的搜索结果来识别的。

26.一种系统,包括:

用于存储数据的存储器;以及

硬件处理器,所述硬件处理器能够操作以执行操作,所述操作包括:

提供地理区域的地图以供在计算设备上显示;

基于在所述计算设备上显示的所述地理区域的所述地图来选择用于话音识别的语法；

在所述计算设备显示所述地理区域的所述地图时，接收与话语相对应的语音数据以发起搜索；

(i)通过使用基于在所述计算设备上显示的所述地理区域的所述地图所选择的所述语法来对所述话语执行话音识别并且(ii)基于识别一个或多个搜索查询的搜索历史来生成所述话语的转录，所述一个或多个搜索查询由与所述计算设备的用户相关联的一个或多个计算设备先前提交至一个或多个搜索服务；以及

更新所述地图以提供识别与所述话语的所述转录相关联的位置的视觉指示符。

27.如权利要求26所述的系统，其中，所述转录是进一步基于从所述地理区域的所述地图得出的地理信息来生成的。

28.如权利要求26所述的系统，其中，所述转录是进一步基于从先前提供以供显示的内容得出的地理信息来生成的。

29.如权利要求26所述的系统，其中，所述话语识别位置。

30.如权利要求26所述的系统，其中，所述语法是进一步基于地理位置来选择的，所述地理位置是基于与先前所提交的一个或多个搜索查询相关联的搜索结果来识别的。

31.一种存储指令的非暂时性计算机可读介质，所述指令在被执行时，能够操作以使一个或多个处理器执行操作，所述操作包括：

提供地理区域的地图以供在计算设备上显示；

基于(i)在所述计算设备上显示的所述地理区域的所述地图以及(ii)地理位置来选择用于话音识别的语法，所述地理位置是基于与先前所提交的一个或多个搜索查询相关联的搜索结果来识别的；

在所述计算设备显示所述地理区域的所述地图时，接收与话语相对应的语音数据以发起搜索；

通过使用基于在所述计算设备上显示的所述地理区域的所述地图所选择的所述语法来对所述话语执行话音识别，生成所述话语的转录；以及

更新所述地图以提供识别与所述话语的所述转录相关联的位置的视觉指示符。

32.如权利要求31所述的计算机可读介质，其中，所述转录是进一步基于从所述地理区域的所述地图得出的地理信息来生成的。

33.如权利要求31所述的计算机可读介质，其中，所述转录是进一步基于从先前提供以供显示的内容得出的地理信息来生成的。

34.如权利要求31所述的计算机可读介质，其中，所述操作进一步包括：

通过接收与搜索查询的话语相对应的语音数据来接收与发起搜索的话语相对应的所述语音数据；

在所述计算设备显示所述地理区域的所述地图时，向搜索引擎传送所述搜索查询；

在所述计算设备显示所述地理区域的所述地图时响应于所述搜索查询，从所述搜索引擎接收搜索结果；以及

通过基于所述搜索查询来更新所述地图以提供所述视觉指示符来更新所述地图以提供识别与所述话语的转录相关联的位置的视觉指示符。

35.一种计算机实现的方法，包括：

提供地理区域的地图以供在计算设备上显示；

基于在所述计算设备上显示的所述地理区域的所述地图来选择用于话音识别的语法；

在所述计算设备显示所述地理区域的所述地图时：

接收与搜索查询的话语相对应的语音数据；

通过使用基于所述地理区域的所述地图所选择的所述语法来对所述话语执行话音识别，生成所述搜索查询的所述话语的转录；

向搜索引擎传送所述搜索查询的所述转录；以及

从所述搜索引擎接收响应于所述搜索查询的所述转录的搜索结果；以及

更新所述地图以提供识别与所述搜索结果相关联的位置的视觉指示符。

36.一种计算机实现的方法，包括：

由计算机系统访问搜索历史，所述搜索历史识别与用户相关联的移动计算设备先前提交至一个或多个搜索服务的一个或多个搜索查询；

访问识别先前在所述移动计算设备上显示的一个或多个网页的信息；

至少部分地基于以下来识别来自所述用户的未来语音输入很可能涉及的地理位置：i)所述移动计算设备的所述搜索历史，ii)与所述搜索历史中的所述一个或多个搜索查询相关联的搜索结果，以及iii)先前在所述移动计算设备上显示的所述一个或多个网页中所包括的内容；

由所述计算机系统至少部分地基于所识别的地理位置来从多个语法当中选择第一语法，其中，所述第一语法包括词汇表，所述词汇表与所识别的地理位置相关；以及

由所述计算机系统输出识别所述第一语法的信息，其中，所输出的信息使得用来分析来自移动计算设备的语音输入的语法被改变为所述第一语法。

37.如权利要求36所述的计算机实现的方法，其中，所述搜索查询的所述搜索结果是跨提供给所述一个或多个搜索服务的多个用户的搜索结果集合而聚合的。

38.如权利要求36所述的计算机实现的方法，进一步包括：

至少部分地基于所述移动计算设备的所述搜索历史，识别来自所述用户的未来语音输入很可能涉及的一个或多个话题；

至少部分地基于所识别的一个或多个话题，从所述多个语法当中选择第二语法，其中，所述第二语法包括与所述一个或多个话题相关的词汇表；以及

其中，所输出的信息进一步识别所述第二语法并且使得用来分析来自所述移动计算设备的所述语音输入的所述语法被改变为包括所述第一语法和所述第二语法的多个语法。

39.如权利要求36所述的计算机实现的方法，其中，所述地理位置是独立于所述移动计算设备的当前地理位置来识别的。

40.如权利要求36所述的计算机实现的方法，其中，所述移动计算设备包括智能电话。

41.一种计算机实现的方法，包括：

由计算机系统访问由与用户相关联的移动计算设备先前请求并在所述移动计算设备上显示的内容的信息，其中，所述内容包括先前在所述移动计算设备上显示的一个或多个网页；

至少部分地基于先前被请求并被显示在所述移动计算设备上的所述内容，识别来自所述用户的未来语音输入很可能涉及的地理位置；

由所述计算机系统至少部分地基于所识别的地理位置来从多个语法当中选择第一语法,其中,所述第一语法包括词汇表,所述词汇表与所述地理位置相关;以及

由所述计算机系统输出识别所述第一语法的信息,其中,所输出的信息使得用来分析来自移动计算设备的语音输入的语法被改变为所述第一语法。

42.如权利要求41所述的计算机实现的方法,进一步包括:

识别所述内容的、所述用户表达了至少阈值水平的兴趣的特定部分;以及

其中,所述地理位置是基于所述内容的所述特定部分来识别的。

43.如权利要求42所述的计算机实现的方法,其中,当用户控制的指示器在所述内容的所述特定部分上方悬停了至少阈值时间段时,所述用户被确定为已经对所述特定部分表达了至少所述阈值水平的兴趣。

44.如权利要求41所述的计算机实现的方法,进一步包括:

由所述计算机系统将所述内容分类为一个或多个关键字;

至少部分地基于所述一个或多个关键字,从所述多个语法当中选择第二语法,其中,所述第二语法包括与所述一个或多个关键字相关的词汇表;以及

其中,所输出的信息进一步识别所述第二语法并且使得用来分析来自所述移动计算设备的所述语音输入的所述语法被改变为包括所述第一语法和所述第二语法的多个语法。

45.如权利要求41所述的计算机实现的方法,其中,所述地理位置是基于所述地理位置在所述内容中出现了至少阈值次数而从所述内容中被识别的。

46.如权利要求41所述的计算机实现的方法,其中,所述内容包括在所述移动计算设备上显示的文本内容。

47.一种计算机系统,包括:

一个或多个计算设备;

所述一个或多个计算设备的接口,所述接口被编程来访问识别以下中的一个或多个的信息:i)识别先前由用户输入到移动计算设备的一个或多个搜索查询的搜索历史以及与所述一个或多个搜索查询相关联的搜索结果,以及ii)先前被请求并在所述移动计算设备上向所述用户显示的内容,其中,所述内容包括先前在所述移动计算设备上显示的一个或多个网页;

反向地理编码器,所述反向地理编码器安装在所述一个或多个计算设备上并且被编程为至少部分地基于以下中的一个或多个来识别来自所述用户的未来语音输入很可能涉及的地理位置:i)识别先前由所述用户输入的所述一个或多个搜索查询的所述搜索历史以及关联的搜索结果,以及ii)先前被请求并被显示在所述移动计算设备上的所述内容;以及

语法索引,所述语法索引能够由所述一个或多个计算设备访问并且包括将地理位置与多个语法关联的索引;

语法选择器,所述语法选择器被安装在所述一个或多个计算设备上并且被编程为至少部分地基于由所述反向地理编码器识别的所述地理位置来从所述多个语法当中选择第一语法,其中,所述第一语法包括与所述地理位置相关的词汇表;

其中,所述接口进一步被编程为输出识别所述第一语法的信息,其中,所输出的信息使得用来分析来自所述移动计算设备的语音输入的语法被改变为所述第一语法。

48.如权利要求47所述的计算机系统,其中:

所述一个或多个计算设备被进一步编程为至少部分地基于以下中的一个或多个来识别来自所述用户的未来语音输入很可能涉及的一个或多个话题;i)识别先前由所述用户输入到所述移动计算设备的所述一个或多个搜索查询的所述搜索历史以及关联的搜索结果,以及ii)先前被请求并被显示在所述移动计算设备上的所述内容;

所述语法选择器被进一步编程为至少部分地基于所识别的一个或多个话题来从所述多个语法当中选择第二语法,其中,所述第二语法包括与所述一个或多个话题相关的词汇表;以及

所述接口进一步被编程为输出进一步识别所述第二语法的信息,并且其中,输出所述信息进一步使得用来分析来自所述移动计算设备的所述语音输入的所述语法被改变为包括所述第一语法和所述第二语法的多个语法。

49.如权利要求47所述的计算机系统,其中,所述地理位置是由所述反向地理编码器独立于所述移动计算设备的当前地理位置来被识别的。

50.如权利要求47所述的计算机系统,其中,所述地理位置是由所述反向地理编码器基于所述地理位置在以下中的一个或多个中出现了至少阈值次数而被识别的:i)识别先前由所述用户输入到所述移动计算设备的所述一个或多个搜索查询的所述搜索历史以及关联的搜索结果,以及ii)先前被请求并被显示在所述移动计算设备上的所述内容。

51.如权利要求47所述的计算机系统,其中,所述搜索查询的所述搜索结果是跨提供给所述一个或多个搜索服务的多个用户的搜索结果集合而聚合的。

52.如权利要求47所述的计算机系统,其中,所述反向地理编码器进一步被编程为基于所述内容的、所述用户表达了至少阈值水平的兴趣的特定部分来识别所述地理位置。

53.一种计算机实现的方法,包括:

由计算机系统访问搜索历史,所述搜索历史识别先前由与用户相关联的一个或多个计算设备提交至一个或多个搜索服务的一个或多个搜索查询;

至少部分地基于在所述搜索历史中包括的所述搜索查询,识别来自所述用户的未来语音输入很可能涉及的地理位置,所述地理位置是独立于所述一个或多个计算设备的当前地理位置来被识别的;

在对来自所述一个或多个计算设备的语音输入执行自动化话音识别之前,由所述计算机系统至少部分地基于所述地理位置来从多个语法当中选择第一语法,所述地理位置是至少部分地基于所述搜索历史中所包括的所述搜索查询来识别的,所述第一语法包括词汇表,所述词汇表与至少部分地基于所述搜索历史中所包括的所述搜索查询来识别的所述地理位置相关;以及

由所述计算机系统输出识别所述第一语法的信息,其中,所输出的信息使得用来分析来自所述一个或多个计算设备的随后接收的语音输入的语法被改变为所述第一语法。

54.如权利要求53所述的计算机实现的方法,其中,所述地理位置是进一步基于与所述搜索历史中的所述一个或多个搜索查询相关联的搜索结果来识别的。

55.如权利要求54所述的计算机实现的方法,其中,所述搜索查询的所述搜索结果是跨由所述一个或多个搜索服务提供给多个用户的搜索结果集合而聚合的。

56.如权利要求53所述的计算机实现的方法,进一步包括:

至少部分地基于所述搜索历史,识别来自所述用户的未来语音输入很可能涉及的一个

或多个话题；

至少部分地基于所识别的一个或多个话题，从所述多个语法当中选择第二语法，其中，所述第二语法包括与所述一个或多个话题相关的词汇表；以及

其中，所输出的信息进一步识别所述第二语法并且使得用来分析来自所述移动计算设备的所述语音输入的所述语法被改变为包括所述第一语法和所述第二语法的多个语法。

57.如权利要求53所述的计算机实现的方法，其中，所述一个或多个计算设备包括智能电话。

58.如权利要求53所述的计算机实现的方法，其中，所述地理位置是进一步基于先前被呈现在所述一个或多个计算设备上的内容来识别的。

59.如权利要求58所述的计算机实现的方法，其中，所述内容包括一个或多个网页。

60.如权利要求58所述的计算机实现的方法，其中，所述内容包括一个或多个地图。

61.一种计算机系统，包括：

一个或多个计算设备；

所述一个或多个计算设备的接口，所述接口被编程来访问识别一个或多个搜索查询的搜索历史，所述一个或多个搜索查询先前由与用户相关联的一个或多个客户端计算设备提交至一个或多个搜索服务；

反向地理编码器，所述反向地理编码器安装在所述一个或多个计算设备上并且被编程为至少部分地基于所述搜索历史中包括的所述搜索查询来识别来自所述用户的未来语音输入很可能涉及的地理位置，其中，所述地理位置是独立于所述一个或多个客户端计算设备的当前地理位置来被识别的；以及

语法索引，所述语法索引能够由所述一个或多个计算设备访问并且包括将地理位置与多个语法关联的索引；

语法选择器，所述语法选择器被安装在所述一个或多个计算设备上，并且被编程为在对来自所述一个或多个客户端计算设备的语音输入执行自动化话音识别之前，至少部分地基于所述地理位置来从所述多个语法当中选择第一语法，所述地理位置是由所述反向地理编码器至少部分地基于所述搜索历史中所包括的所述搜索查询来识别的，所述第一语法包括词汇表，所述词汇表与至少部分地基于所述搜索历史中所包括的所述搜索查询来识别的所述地理位置相关；以及

其中，所述接口进一步被编程为输出识别所述第一语法的信息，其中，所输出的信息使得用来分析来自所述一个或多个客户端计算设备的随后接收的语音输入的语法被改变为所述第一语法。

62.如权利要求61所述的计算机系统，其中，所述地理位置是进一步基于与所述搜索历史中的所述一个或多个搜索查询相关联的搜索结果来识别的。

63.如权利要求62所述的计算机系统，其中，所述搜索查询的所述搜索结果是跨由所述一个或多个搜索服务提供给多个用户的搜索结果集合而聚合的。

64.如权利要求61所述的计算机系统，其中：

所述一个或多个计算设备被进一步编程为至少部分地基于所述搜索历史和关联的搜索结果，识别来自所述用户的未来语音输入很可能涉及的一个或多个话题，所述搜索历史识别由所述用户先前输入至所述一个或多个客户端计算设备的一个或多个搜索查询；

所述语法选择器被进一步编程为至少部分地基于所识别的一个或多个话题来从所述多个语法当中选择第二语法，其中，所述第二语法包括与所述一个或多个话题相关的词汇表；以及

所述接口进一步被编程为输出进一步识别所述第二语法的信息，并且其中，输出所述信息进一步使得用来分析来自所述一个或多个客户端计算设备的所述语音输入的所述语法被改变为包括所述第一语法和所述第二语法的多个语法。

65.如权利要求61所述的计算机系统，其中，所述一个或多个客户端计算设备包括智能电话。

66.如权利要求61所述的计算机系统，其中，所述地理位置是进一步基于先前被呈现在所述一个或多个计算设备上的内容来识别的。

67.如权利要求66所述的计算机系统，其中，所述内容包括一个或多个网页。

68.如权利要求66所述的计算机系统，其中，所述内容包括一个或多个地图。

基于场境的语音识别语法选择方法和系统

[0001] 分案说明

[0002] 本申请属于申请日为2009年3月6日的中国发明专利申请No.201710202874.9的分案申请。

技术领域

[0003] 本说明书涉及语音识别。

背景技术

[0004] 多模式应用可以接受多种类型或模式的输入。例如,多模式应用可以从用户接受诸如键入命令和鼠标点击的输入;然而,多模式应用还可以接受诸如语音输入的其它形式的输入。语音识别系统可以使用包括特定词汇表的语法来解释语音输入。

[0005] 某些多模式应用实现表格填充模型,其中应用中的不同输入字段与不同语法相关联。例如,多模式应用可以将包括名称的词汇表的“名称”语法与接受人的名称的“名称”字段相关联。在某人用他或她的鼠标选择“名称”字段时,多模式应用可以选择与该字段相关联的“名称”语法。

发明内容

[0006] 总的来说,本文档描述了选择用于在语音识别中使用的语法,其中该语法基于隐式场境(context)信息来选择,所述隐式场境信息基于一个或多个用户行为。

[0007] 在第一总括方面,描述了一种计算机实现的方法。该方法包括接收从与第一计算设备相关联的非语言用户行为取得的地理信息。该非语言用户行为暗示用户对地理位置感兴趣。该方法还包括使用取得的地理信息来识别与该地理位置相关联的语法,以及输出语法指示符以供在选择所识别的语法对来自用户的话音输入进行语音识别处理中使用。

[0008] 在第二总括方面,描述了一种计算机实现的方法,其包括基于与计算设备相关联的用户行为接收场境信息,其中语音识别处理器不提示(unprompted)该用户行为。该方法包括基于该场境信息取得地理位置,识别与该地理位置相关联的语法,以及输出语法标识符以供在选择所述语法用来处理来自所述用户的话音输入中使用

[0009] 在又一个总括方面,描述了一个系统。该系统包括基于与第一计算设备相关联的非语言用户行为来接收场境信息的接口。该非语言用户行为暗示用户对地理感兴趣。该系统还包括用于识别与从所接收的场境信息取得的地理位置相关联的语法的装置,以及配置为使用所识别的语法来解释从用户接收的话音输入的语音识别服务器。

[0010] 在此描述的系统和技术可以提供以下优势的一个或多个。首先,系统可以通过识别语法的适当子集而不是使用更大的普通语法来提高识别话音输入的速度。额外地,选择语法的子集对于在与该子集相关联的特定场境内给定的话音命令可以提高语音识别的准确性。系统还可以减少在语音识别过程中的用户交互量。此外,可以以对用户透明的方式(例如,基于与软件应用的用户交互而不是基于对语音识别系统的提示的明确用户回答)选

择额外或新的语法。

[0011] 在下面的附图和描述中阐述了一个或多个实施例的细节。其它特征和优势从描述和附图、以及从权利要求将是显而易见的。

附图说明

[0012] 图1是用于选择在计算机实现的语音识别中使用的语法的示例性系统的图。

[0013] 图2是在选择用于语音识别的语法中使用的示例性客户端和示例性音频处理系统的图。

[0014] 图3是用于基于场境信息来选择语法的示例性方法的流程图。

[0015] 图4示出了用于客户端的与音频处理系统交互的用户界面的示例性截屏。

[0016] 图5是可以用于实现所描述的系统和方法的计算设备的框图。

[0017] 在各个附图中相同的参考符号指示相同的元素。

具体实施方式

[0018] 本文档描述了用于选择在话音识别中使用的语法的系统和技术。更具体地，应用可以基于用户与运行该应用的设备如何交互来生成场境信息。所述设备可以将所述场境信息传送给语音识别系统。语音识别系统可以识别与所述场境信息相关联的特定语法，并且可以随后在对从所述设备接收的话音输入(例如，来自所述设备的用户的语音命令)进行音频识别中使用所识别的语法。

[0019] 在一些实施方式中，场境信息包括地理信息。如果用户使用蜂窝电话上的移动浏览器来查看地图，则该蜂窝电话可以将与所述地图显示的地理位置有关的信息传送给语音识别系统。语音识别系统可以基于所接收的信息来识别所述地理位置，以及可以选择与所述地理位置相关联的语法，所述语法包括词的词汇表、句法等。语音识别系统可以使用所选择的语法来解释随后从所述蜂窝电话接收的话音输入。

[0020] 用户可以选择允许与该用户相关联的设备(例如，通过分配给该用户的唯一标识符—诸如cookie)与语音识别系统共享场境信息。如果用户不选择共享场境信息，则语音识别系统可以使用默认语法或为用于在选择在语音识别处理中使用的特定语法时使用的地理或其它场境信息明确提示用户。

[0021] 图1是用于选择在计算机实现的语音识别中使用的语法的示例性系统100的图。在一些实施方式中，示例性系统100基于作为用户行为的结果生成的隐式场境信息来选择语法。系统100包括客户端设备102、语法选择服务器104和语音识别服务器106。

[0022] 在图1的实施方式中，客户端102将场境信息传送给语法选择服务器104，其使用所述场境信息来选择语法。语法选择服务器将与所选择的语法有关的信息发送给语音识别服务器106，其使用所选择的语法来解释从客户端102接收的音频输入。

[0023] 例如，客户端102可以是正在运行移动浏览器108的蜂窝电话。用户可以将搜索输入到移动浏览器以识别出售“冰屋”的企业。浏览器可以显示示出了位于加拿大的南、北达科他州和明尼苏达州的相关企业的地图。用户可能先前已输入了由浏览器使用来识别用于在地图上示出的位置的位置标识符，诸如邮编。用户可以在先前会话中输入位置标识符，以及浏览器可以存储标识符以供在随后会话中使用(例如，位置标识符可以作为cookie被存

储在客户端上)。

[0024] 在其它实施方式中,蜂窝电话运行专用应用108,而不是移动浏览器108。例如,专用应用108可以不能够浏览网页,但是可以被配置为与诸如在线地图应用的特定远程应用对接。

[0025] 在蜂窝电话102上运行的移动浏览器或另一个应用如标记“1”的箭头所指示可以将隐式地理信息114传送给语法选择服务器。在一些实施方式中,场境信息包括基于移动浏览器108显示的地图的隐式地理信息114。例如,隐式地理信息114可以包括识别所显示的地图的中心的坐标。在该示例中,地图的中心与在明尼苏达州的中间的位置一致。

[0026] 语法选择服务器104可以基于所接收的场境信息来选择116语法。在一些实施方式中,如果场境信息包括地理信息114,则语法选择服务器可以访问识别与地理位置相关联的语法112的数据存储110。特定语法可以与特定地理位置相关联。

[0027] 语法112中的每一个可以包括和与相应的地理位置相关联的词汇表相对应的词汇表。例如,与明尼苏达州明尼阿波利斯相关联的语法可以包括描述位于明尼阿波利斯或与明尼阿波利斯关联地发生的企业、兴趣点、事件、新闻等的词或短语。

[0028] 在图1中,隐式地理信息114包括位于明尼苏达州的中间的位置的坐标。语法选择服务器识别锚定在最接近于明尼苏达州的中间的位置处或具有在该位置处的中心的语法。语法选择服务器如标记“2A”和“2B”的箭头所指示可以生成识别语法的信息118。

[0029] 在一些实施方式中,然后,语法选择服务器104如标记“3”的箭头所指示将所识别的语法信息118传送给语音识别服务器106。

[0030] 用户可以对蜂窝电话102说话,蜂窝电话102如箭头“4”所指示将话音输入120传送给语音识别服务器106。语音识别服务器106可以使用语法选择服务器104已识别的语法来解释话音输入120。

[0031] 识别服务器106可以基于话音输入来执行一个或多个动作。在一些实施方式中,语音识别服务器106如标记“5”的箭头所指示可以基于话音输入将响应传送回蜂窝电话102。例如,蜂窝电话102的用户可以可听见地请求对“Paul Bunyan”的新搜索。蜂窝电话102可以将话音搜索请求传送给语音识别服务器106。由于用户当前在移动浏览器108上查看(或先前查看过)的地图以明尼苏达州为中心,所以语音识别服务器106使用锚定在明尼苏达州的中心附近的位置处或以该位置为中心的语法。语音识别服务器106使用该语法来搜索与话音输入“Paul Bunyan”相对应的声音、词、短语。在一些实施方式中,由于词语“Paul Bunyan”相对于世界的其它部分同与明尼苏达州相关联的信息更频繁地关联,所以锚定在明尼苏达州附近或之内的语法可以包括用于解释名称“Paul Bunyan”的信息。

[0032] 语音识别服务器106可以传送文本“Paul Bunyan”,其对应于来自蜂窝电话102的话音输入。蜂窝电话102可以在移动浏览器108上显示所接收的文本。如果用户认可语音识别服务器106执行的翻译,则浏览器可以通过将文本“Paul Bunyan”作为搜索词语提交给搜索引擎来发起新的搜索。在其它实施方式中,语音识别服务器106可以在没有来自蜂窝电话102的用户的认可的情况下使用词语“Paul Bunyan”来发起搜索。语音识别服务器106可以在先前没有传送从话音输入识别的文本的情况下将来自搜索的结果传送给蜂窝电话102。

[0033] 图1的标记箭头指示在系统100中发生的事件的示例性序列。然而,事件的发生不限于所示的序列。例如,序列中的一个或多个步骤可以并行发生。

[0034] 图2是在选择用于语音识别的语法中使用的示例性客户端200和示例性音频处理系统202的图。客户端200和音频处理系统202可以使用在一些实施方式中可以包括因特网和蜂窝网络的网络204来通信。客户端200可以包括进而包括应用环境208的蜂窝电话或其它移动设备206。应用环境208可以包括因特网浏览器210、麦克风接口212以及GPS收发器接口214。音频处理系统202可以包括提供音频处理系统202与客户端200的接口的多模式服务器216、语法选择服务器218以及语音识别服务器220。

[0035] 应用环境208内的应用可以生成或识别地理联系信息222，并且将该信息传送给多模式服务器216。例如，GPS收发器接口214可以基于蜂窝电话206的位置从GPS收发器接收GPS坐标。GPS收发器接口214可以将GPS坐标信息传送给多模式服务器216。

[0036] 在一些实施方式中，GPS坐标信息可以附加为包括在浏览器210提交给多模式服务器216的超文本传输协议(HTTP)POST命令中的统一资源定位符(URI)的部分。在使用不同于浏览器的应用的其它实施方式中，应用可以生成HTTP GET命令，其中命令中的URI包括GPS坐标信息(或其它场境信息)。在另一个实施方式中，GPS坐标或其它场境信息不附加在URI中，而是作为二进制信息包括在HTTP请求(例如，GET或POST)的主体中。

[0037] 在另一个示例中，浏览器210可以传送与浏览器210显示的项有关的地理场境信息。例如，如果用户查看包括百慕大群岛的多次提及的网页，则浏览器210可以传送指定百慕大群岛的地理场境信息。

[0038] 多模式服务器216可以接收地理场境信息222，并且可以将该信息转送给语法选择服务器218。语法选择服务器218可以包括反向地理编码器224，其使用地理场境信息222来识别位置。例如，如果地理场境信息222包括GPS坐标，则反向地理编码器224可以使用在坐标和地理位置之间的已存储的映射来确定对应于GPS坐标的位置。

[0039] 在一些实施方式中，语法选择服务器包括将特定位置与特定语法相关联的语法索引226。例如，语法索引226将位置“百慕大群岛”与包括与该位置相关联的词汇表、句法等的百慕大语法相关联。

[0040] 语法选择服务器218通过识别与反向地理编码器224识别的位置相关联的语法，使用语法索引226来选择语法。语法索引可以在使用语法ID的情况下识别语法中的每一个。

[0041] 语法选择服务器218可以将所选择的语法ID 228传送给多模式服务器216，其进而可以将语法ID 228转送给语音识别系统。在图2中未示出的其它实施方式中，反向地理编码器224可以识别并且返回与最靠近所识别的位置的位置相关联的语法，然后将用于该语法的所选择的语法ID 228传送给多模式服务器216。

[0042] 语音识别系统可以使用语法ID来加载所识别的语法以供在随后的音频处理中使用。例如，语音识别服务器可以将对语法的请求232传送给数据存储230，其中请求232包括语法ID 228。数据存储可以返回语法ID 232指定的语法234。

[0043] 语音识别服务器可以使用语法234来解释随后从蜂窝电话206接收的音频。例如，用户可以说出搜索词语，其由蜂窝电话206内的麦克风接收。麦克风接口212可以将音频236从麦克风传送到多模式服务器216。

[0044] 多模式服务器216可以将音频236传送给语音识别服务器220，其使用音频解码器238来解释音频236。例如，音频解码器238可以加载语法234来将音频236处理成文本表示。例如，语音识别服务器220可以使用文本表示来发起搜索引擎(未示出)的搜索。在另一个示

例中，可以将已解释的音频作为文本240传送给多模式服务器216。多模式服务器216可以将文本240传送回蜂窝电话206。蜂窝电话106可以使用应用环境208中的浏览器210或另一个应用来显示文本。

[0045] 在一些实施方式中，客户端200基于新的用户交互来提交新的地理场境信息。例如，如果用户改变位置，则蜂窝电话206内的GPS收发器可以将新的GPS坐标传送给多模式服务器216。在另一个示例中，用户可以查看与不同的位置相关联的地图。浏览器210可以将新的地图位置传送给多模式服务器216。音频处理系统可以基于新的地理场境信息来选择新的语法，并且基于该新的语法来解释所接收的音频。

[0046] 尽管将多模式服务器216、语法选择服务器218和语音识别服务器220图示为独立设备，但是可以将服务器组合成单个设备或可以使用多个设备来实现单个服务器。

[0047] 图3是用于基于场境信息来选择语法的示例性方法300的流程图。在方法300的示例实施方式中使用系统200和202。然而，包括系统100的其它系统可以实现方法300。

[0048] 在步骤302，在客户端200和音频处理系统202之间创建会话。例如，蜂窝电话206可以(例如，基于HTTP协议)建立与多模式服务器216的通信会话。例如，可以在浏览器访问用于搜索引擎的web接口(例如，搜索网页、交互式地图、允许用户搜索托管在站点上的简档的社交联网站点等)时建立会话。在另一个实施方式中，在蜂窝电话206上起动特定应用时，建立会话。例如，可以在蜂窝电话206上起动专用地图程序时发起会话。

[0049] 在可选步骤304，接收用户ID。例如，蜂窝电话206可以包括将cookie存储在蜂窝电话的存储器内的移动浏览器。Cookie可以包括识别蜂窝电话的用户的标识符。音频处理系统202可能响应于移动浏览器与音频处理系统202或音频处理系统202可以访问的另一个服务器的早先交互，先前已将用户ID传送给浏览器。例如，用户可以访问包括用于搜索引擎的接口的网页。搜索引擎可以向用户发布唯一标识符。音频处理系统202可以访问搜索引擎存储的标识符的列表。

[0050] 在步骤306，接收场境信息。例如，多模式服务器216接收地理场境信息，诸如指定移动设备206的当前位置的GPS坐标。

[0051] 在一些实施方式中，多模式服务器216可以接收其它场境信息，诸如特定于应用的场境信息。客户端202可以传送指定用户访问哪个应用的信息。例如，信息可以指定用户正与浏览器应用交互。此外，信息可以包括应用的用户先前执行的过去导航或其它行为的历史。例如，场境信息可以指定用户已通过指定邮编来请求过地图、在给定的地图上缩小、在地图上向西导航了大约200英里、请求了对地图的卫星查看、请求了在地图上显示兴趣点等。

[0052] 在另一个实施方式中，多模式服务器216可以接收与运行于客户端200上的应用显示的项有关的场境信息，所述客户端200可以是台式机。例如，用户可以查看包括诸如财经新闻内容、娱乐新闻内容、技术新闻内容等的数种类型的内容的门户网页。如果用户的光标悬停于财经新闻内容上，则计算设备可以从环绕光标的区域提取信息(例如，可以提取光标的中心的一定半径内的文本)。可以将所提取的信息的部分或全部包括在传送给多模式服务器的场境信息中。

[0053] 在步骤308，基于所接收的场境信息来选择语法。例如，语法选择服务器218可以选择包括在所接收的场境信息中的财经词汇表的语法，所述场境信息指示用户的鼠标正悬停

于网页上的描述财经信息的内容上。更具体的，语法选择服务器218内的分类模块(未示出)可以对提取的内容分类。语法选择服务器218可以将产生于对提取的内容的分类的一个或多个关键词与通过语法索引226与语法相关联的关键词相匹配。

[0054] 在另一个示例中，语法选择服务器218可以选择与特定地理相关联的语法，其中特定地理对应于指示蜂窝电话206的当前位置的GPS坐标。

[0055] 在又一个示例中，如果所接收的场境信息指定用户正与特定应用交互，则语法选择服务器218可以选择包括特定于应用的词汇表的语法。例如，如果用户正与(例如，驻存于客户端200上或托管并经由浏览器210访问的)日历应用交互，则语法选择服务器218可以选择包括特定于日历的词汇表和特定于日历的语法规则的语法。

[0056] 语法选择服务器218还可以使用所接收的用户ID来选择语法。在一些实施方式中，可以基于用户的过去Web搜索历史来构建语法。例如，如果用户频繁执行与考古学相关联的过去web搜索查询，则语法创建器(未示出)可以构建用于该用户的包括与考古学相关联的词汇表、句法等的个性化语法。

[0057] 在一些实施方式中，可以使用一种或多种类型的场境信息来选择多于一个语法。例如，从用户查看的项取得的场境信息可以与两个或更多个语法关联。在这种情况下，语法选择服务器可以选择多个语法以供语音识别时使用。

[0058] 在步骤310，可以接收音频。例如，设备206的用户可以对着设备206的麦克风说话。麦克风接口212可以将麦克风捕捉到的话音传送给语音识别服务器220。

[0059] 在步骤312，可以使用先前选择的语法来解释所接收的音频。例如，语音识别服务器220可以访问存储语法的数据结构来选择语法选择服务器218识别的语法。音频解码器238可以使用所选择的语法来解释所接收的音频。

[0060] 在步骤314，确定会话是否超时。例如，在步骤302在客户端200和音频处理系统220之间建立的会话可以具有时限。如果超过时限，则方法300可以结束。在另一个实施方式中，如果会话超时，则音频处理系统202提示客户端200建立新的会话。在一些实施方式中，限制会话时间可以阻止客户端独占音频处理系统，尤其在客户端长时间处于非活动状态时。

[0061] 在步骤316，确定场境是否已改变。例如，用户可以改变位置。如果用户移动到新的位置，则GPS收发器可以响应于反映新的位置而更新GPS坐标。可以如先前与步骤306和随后步骤相关联地描述的来接收并处理新的场境信息。在另一个示例中，用户可以访问不同的应用，或使用同一应用来查看不同的数据。应用的改变或对应用的使用可以发起对新的场境信息的传送。

[0062] 如果场境信息没有改变，则方法300可以重复以步骤310开始的步骤。例如，音频处理系统202可以继续使用先前选择的语法来解释任何接收的音频。

[0063] 图4示出了用于客户端的与基于场境信息来选择语法的音频处理系统交互的用户界面的示例性截屏。截屏400包括地图402。用户可以启用客户端的浏览器，并且登录到诸如GOOGLE地图或YAHOO!地图的在线交互式地图服务。用户可以通过输入邮编、区号、城市和州或其它位置标识符来指定到地图的位置。例如，用户可以输入邮编95661。地图服务然后可以传送加州罗斯维尔的相应地图以在浏览器上显示。

[0064] 浏览器(或浏览器显示的网页)可以呈现搜索选项窗口404。用户可以与搜索选项窗口404交互来发起对企业、兴趣点、位置等的搜索，并且可以在地图402上显示结果。搜索

选项窗口404如选项406所指示可以接受“输入的”搜索。例如，用户可以使用小键盘来选择选项406。替选地，用户可以说出选择“输入新的搜索”。用户然后可以经由小键盘输入搜索。

[0065] 搜索选项窗口如选项408所指示还可以接受口语搜索。例如，用户可以使用小键盘或通过说出选择来选择选项408。

[0066] 截屏410示出了在用户选择了指示用户期望说出新的搜索的选项408后显示的示例性界面。在该示例中，客户端可视地提示用户说出企业的类型或企业名称。图412图示了用户说出搜索词语“Fry’s”。

[0067] 在其它实施方式中，客户端可以提示用户说出诸如兴趣点、地理位置等的其它词语。

[0068] 在又其它实施方式中，没有可视地提示用户说出搜索，但是用户可以自行发起搜索。例如，浏览器可以显示显示技术新闻的网页。用户可以说“搜索AJAX”。在又其它实施方式中，客户端可以可听见地提示用户输入或说出搜索或其它话音输入。

[0069] 在另一个实施方式中，用户可以按下设备上的键—诸如蜂窝电话上的“接听呼叫”键—来指示用户意欲发起语音搜索。用户可以在说话时按住该键，或通过按住该键预定时长来发起搜索。在后者实施方式中，语音搜索可以在没有检测到语音信号的情况下经过预定时间量后终止。

[0070] 尽管在先前截屏中没有指示，但是客户端可以传送与客户端显示的地图402有关的地理信息。音频处理系统可以选择与该地理信息相关联的语法以供解释响应于截屏410中显示的提示所接收的音频时使用。

[0071] 截屏416示出了基于口语搜索词语“Fry’s”的已解释的音频。用户可以使用小键盘或通过说出例如与解释的每一个相关联的数字来选择可能的解释中的一个。截屏418示出了加州罗斯维尔的地图402以及识别Fry’s电子商店的位置的图钉图标420。

[0072] 图5是作为客户端或作为服务器或多个服务器的可以用于实现本文档中描述的系统和方法的计算设备500、550的框图。计算设备500意在表示各种形式的数字计算机，诸如膝上型机、台式机、工作站、个人数字助理、服务器、刀片服务器、大型机以及其它适当的计算机。计算设备550意在表示各种形式的移动设备，诸如个人数字助理、蜂窝电话、智能手机以及其它相似的计算设备。额外地，计算设备500或550可以包括通用串行总线(USB)闪存驱动。USB闪存驱动可以存储操作系统和其它应用。USB闪存驱动可以包括输入/输出组件，诸如无线发送器或可以被插入另一个计算设备的USB端口的USB连接器。在此示出的组件、其连接和关系以及其功能意在只是示例性的，并且并不意在限制在本文档中描述和/或主张的本发明的实施方式。

[0073] 计算设备500包括处理器502、存储器504、存储设备506、连接到存储器504和高速扩展端口510的高速接口508以及连接到低速总线514和存储设备506的低速接口512。组件502、504、506、508、510和512中的每一个使用各种总线互连，并且可以视情况被安装在公共主板上或以其它方式安装。处理器502可以处理用于在计算设备500内执行的指令，包括存储在存储器504中或存储设备506上用来在诸如耦接到高速接口508的显示516的外部输入/输出设备上显示GUI的图形信息的指令。在其它实施方式中，可以视情况使用多个处理器和/或多个总线以及多个存储器和存储器类型。并且，可以连接多个计算设备500，其中每一个设备提供部分必需操作(例如，作为服务器簇、刀片服务器组或多处理器系统)。

[0074] 存储器504存储在计算设备500内的信息。在一个实施方式中，存储器504是易失性存储器单元。在另一个实施方式中，存储器504是非易失性存储器单元。存储器504还可以是另一种形式的计算机可读介质，诸如磁或光盘。

[0075] 存储设备506能够为计算设备500提供海量存储。在一个实施方式中，存储设备506可以是或包含计算机可读介质，诸如软盘设备、硬盘设备、光盘设备或带设备、闪存或其它相似的固态存储器设备、或设备阵列，包括在存储区域网络或其它配置中的设备。计算机程序产品可以被有形地包含在信息载体中。计算机程序产品还可以包含在被执行时执行诸如上述那些方法的一个或多个方法的指令。信息载体是计算机或机器可读介质，诸如存储器504、存储设备506、在处理器502上的存储器或传播信号。

[0076] 高速控制器508管理关于计算设备500的带宽密集型操作，而低速控制器512管理较低带宽密集型操作。职责的这样的分配只是示例性的。在一个实施方式中，高速控制器508(例如通过图形处理器或加速器)耦接到存储器504、显示516，以及耦接到可以接受各种扩展卡(未示出)的高速扩展端口510。在该实施方式中，低速控制器512耦接到存储设备506和低速扩展端口514。可以包括各种通信端口(例如USB、蓝牙、以太网、无线以太网)的低速扩展端口可以耦接到诸如键盘、指示设备、扫描仪的一个或多个输入/输出设备，或例如通过网络适配器耦接到诸如交换机或路由器的联网设备。

[0077] 计算设备500可以如图中所示以多种不同的形式实现。例如，计算设备500可以被实现为标准服务器520或更多时间被实现在一组这样的服务器中。计算设备500还可以被实现为机架服务器系统524的一部分。另外，计算设备500可以在诸如膝上型计算机522的个人计算机中实现。替选地，来自计算设备500的组件可以与诸如设备550的移动设备(未示出)中的其它组件相结合。这样的设备中的每一个可以包含计算设备500、550中的一个或多个，以及整个系统可以由相互通信的多个计算设备500、550组成。

[0078] 除其它组件外，计算设备550包括处理器552、存储器564、诸如显示554的输入/输出设备、通信接口566以及收发器568。设备550还可以提供有诸如微驱动或其它设备的存储设备以提供额外的存储。组件550、552、564、554、566和568中的每一个使用各种总线互连，以及组件中的数个可以视情况被安装在公共主板上或以其它方式安装。

[0079] 处理器552可以执行在计算设备550内的指令，包括存储在存储器564中的指令。处理器可以被实现为包括独立和多个模拟和数字处理器的芯片的芯片集。额外地，处理器可以使用多个体系结构中的任何来实现。例如，处理器410可以是CISC(复杂指令集计算机)处理器、RISC(精简指令集计算机)处理器或MISC(最小指令集计算机)处理器。处理器可以提供例如对设备550的其它组件的协调，诸如对用户接口、由设备550运行的应用以及设备550的无线通信的控制。

[0080] 处理器552可以通过耦接到显示554的控制接口558和显示接口556与用户通信。显示554可以是例如TFT(薄膜晶体管液晶显示)显示或OLED(有机发光二极管)显示或其它适当的显示技术。显示接口556可以包括用于驱动显示554向用户呈现图形和其它信息的适当的电路。控制接口558可以从用户接收命令并且将命令进行转换以提交给处理器552。另外，可以提供与处理器552通信的外部接口562，以使得设备550能够与其它设备进行近区域通信。外部接口562可以例如在一些实施方式中提供有线通信，或在其它实施方式中提供无线通信，以及还可以使用多个接口。

[0081] 存储器564存储在计算设备550内的信息。存储器564可以被实现为计算机可读介质、易失性存储器单元或非易失性存储器单元中的一个或多个。还可以提供扩展存储器574并且通过可以包括例如SIMM(单列直插内存模块)卡接口的扩展接口572将其连接到设备550。这样的扩展存储器574可以为设备550提供额外的存储空间，或还可以存储用于设备550的应用或其它信息。具体地，扩展存储器574可以包括指令来实现或补充上述过程，以及还可以包括安全信息。因此，扩展存储器574可以例如作为用于设备550的安全模块提供，以及可以被编程有允许安全使用设备550的指令。另外，安全应用以及额外的信息可以经由SIMM卡提供，诸如以不可非法侵入的方式将标识信息放置在SIMM卡上。

[0082] 存储器可以如下所述包括例如闪存和/或NVRAM存储器。在一个实施方式中，将计算机程序产品有形地包含在信息载体中。计算机程序产品包含在被执行时执行诸如上述那些方法的一个或多个方法的指令。信息载体是计算机或机器可读介质，诸如存储器564、扩展存储器574、在处理器552上的存储器或可以例如通过收发器568或外部接口562接收的传播信号。

[0083] 设备550可以在需要的情况下通过可以包括数字信号处理电路的通信接口566无线地通信。通信接口566可以提供在各种模式或协议下的通信，诸如GSM语音通话、SMS、EMS或MMS消息、CDMA、TDMA、PDC、WCDMA、CDMA2000或GPRS等。可以例如通过射频收发器568来发生这样的通信。另外，诸如使用蓝牙、WiFi或其它这样的收发器(未示出)，可以发生短程通信。另外，GPS(全球定位系统)接收器模块570可以向设备550提供可以由运行在设备550上的应用视情况使用的额外的导航和位置相关的无线数据。

[0084] 设备550还可以使用音频编解码器560可听见地通信，音频编解码器560可以从用户接收口语信息并且将其转换为可使用的数字信息。同样，音频编解码器560可以诸如通过例如在设备550的送受话器中的扬声器，为用户生成可听见的声音。这样的声音可以包括来自语音电话呼叫的声音，可以包括已记录的声音(例如，语音消息、音乐文件等)以及还可以包括由在设备550上操作的应用生成的声音。

[0085] 计算设备550可以如图中所示以多种不同的形式实现。例如，计算设备550可以被实现为蜂窝电话580。计算设备550还可以被实现为智能电话582、个人数字助理或其它相似的移动设备的部分。

[0086] 可以在数字电子电路、集成电路、专门设计的ASIC(专用集成电路)、计算机硬件、固件、软件和/或以上的组合中实现在此描述的系统和技术的各种实施方式。这些各种实施方式可以包括在可编程系统上可执行和/或可解释的一个或多个计算机程序中的实施方式，所述可编程系统包括至少一个可编程处理器，其可以是专用或通用的，被耦接以从存储系统、至少一个输入设备和至少一个输出设备接收数据和指令以及向存储系统、至少一个输入设备和至少一个输出设备传送数据和指令。

[0087] 这些计算机程序(也被称为程序、软件、软件应用或代码)包括用于可编程处理器的机器指令，并且可以以高级程序和/或面向对象的编程语言、和/或汇编/机器语言来实现。如在此所使用的，术语“机器可读介质”、“计算机可读介质”是指用于向可编程处理器提供机器指令和/或数据的任何计算机程序产品、装置和/或设备(例如，磁盘、光盘、存储器、可编程逻辑设备(PLD))，包括接收作为机器可读信号的机器指令的机器可读介质。术语“机器可读信号”是指用于向可编程处理器提供机器指令和/或数据的任何信号。

[0088] 为了提供与用户的交互,在此描述的系统和技术可以在具有下述的计算机上实现:用于向用户显示信息的显示设备(例如,CRT(阴极射线管)或LCD(液晶显示)监视器),以及用户通过其可以向计算机提供输入的键盘和指示设备(例如,鼠标或跟踪球)。也可以使用其它类型的设备来提供与用户的交互;例如,提供给用户的反馈可以是任何形式的感知反馈(例如,视觉反馈、听觉反馈或触觉反馈);以及可以以任何形式,包括声学的、话音的或触觉的输入,接收来自用户的输入。

[0089] 在此描述的系统和技术可以在包括后端组件(例如作为数据服务器)、或包括中间件组件(例如应用服务器)、或包括前端组件(例如具有用户通过其可以与在此描述的系统和技术的实施方式交互的图形用户界面或Web浏览器的客户端计算机)、或者这样的后端、中间件或前端组件的任何组合的计算系统中实现。系统的组件可以通过任何形式或介质的数字数据通信(例如通信网络)互连。通信网络的示例包括局域网("LAN")、广域网("WAN")、对等网络(具有自组或静态成员)、网格计算基础设施以及因特网。

[0090] 计算系统可以包括客户端和服务器。客户端和服务器通常彼此远离并且典型地通过通信网络交互。客户端和服务器的关系依靠在各个计算机上运行并且彼此具有客户端-服务器关系的计算机程序产生。

[0091] 尽管在上面详细描述了一些实施方式,但是其它修改是可能的。例如,可以基于用户查看的文本来取得用于选择地理位置的地理信息。例如,用户可以访问描述伟大俄克拉荷马州的网站。Web浏览器(或其它应用)可以传送指示用户对俄克拉荷马感兴趣的地理信息。在一些实施方式中,如果出现阈值数量的与俄克拉荷马有关的单词,则web浏览器(或其它应用)仅传送地理信息。

[0092] 在另一个实施方式中,用于选择地理位置的地理信息基于用户输入的文本。例如,用户可以将诸如“野火”和“加利福尼亚”的搜索词语输入到搜索引擎界面中。语法选择器还可以将这些与词语关联以识别加利福尼亚内的特定位置。例如,来自搜索引擎用户的聚组的最近搜索结果可以指示结果通常包括加利福尼亚圣地亚哥的提及。语法选择器可以基于最近搜索结果来识别锚定在圣地亚哥附近的语法。

[0093] 额外地,尽管关于蜂窝电话或其它便携式设备内的GPS能力描述了定位设备的当前位置,但是可以使用其它基于位置的检测系统。例如,移动设备的位置还可以经由起源蜂窝小区(COO)移动定位技术、到达时间差(TDOA)信号检测技术、到达时间(TOA)技术、到达角(AoA)测量技术、增强型观察时间差(EOTD)技术等来确定。

[0094] 在其它实施方式中,语法创建器可以基于用户发送或接收的电子邮件来编辑用于用户的个性化语法。语法选择服务器可以基于从客户端接收的用户ID选择要使用的个性化语法。

[0095] 另外,图中所描绘的逻辑流不需要所示的特定顺序或相继顺序才能获得期望的结果。另外,可以提供其它步骤,或可以从所描述的流程除去步骤,以及可以将其它组件添加到所描述的系统或从所描述的系统移除组件。因此,其它实施方式在权利要求的范围内。

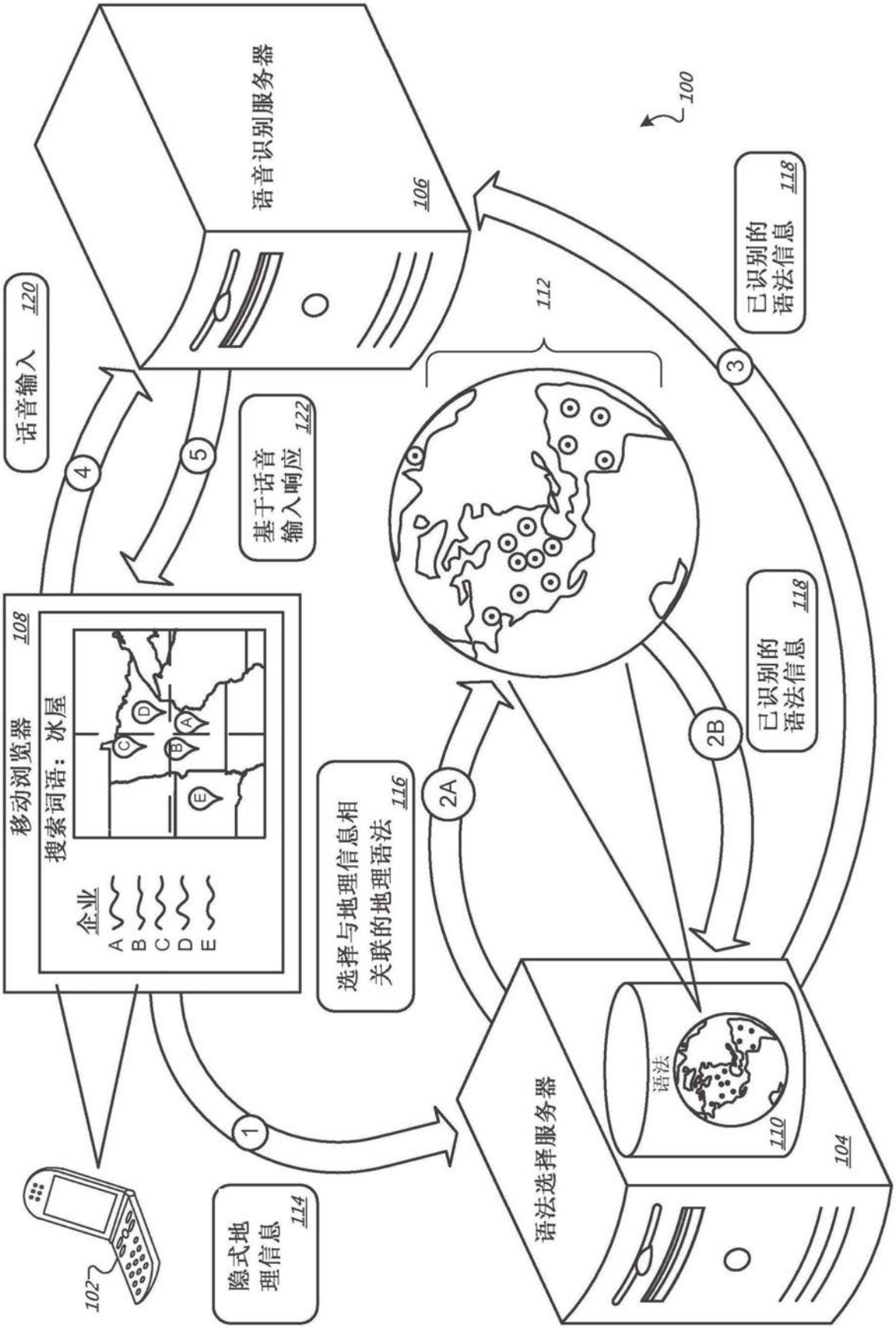

语音识别服务器
106
100
话音输入 120
112
已识别的
语法信息 118
移动浏览器 108
搜索词语：冰屋
企业
A
B
C
D
E
基于话音
输入响应 122
选择与地理信息相
关联的地理语法 116
已识别的
语法信息 118
隐式地
理信息 114
语法选择服务器
语法
110
104
102
1
2A
2B
3
4
5

图1

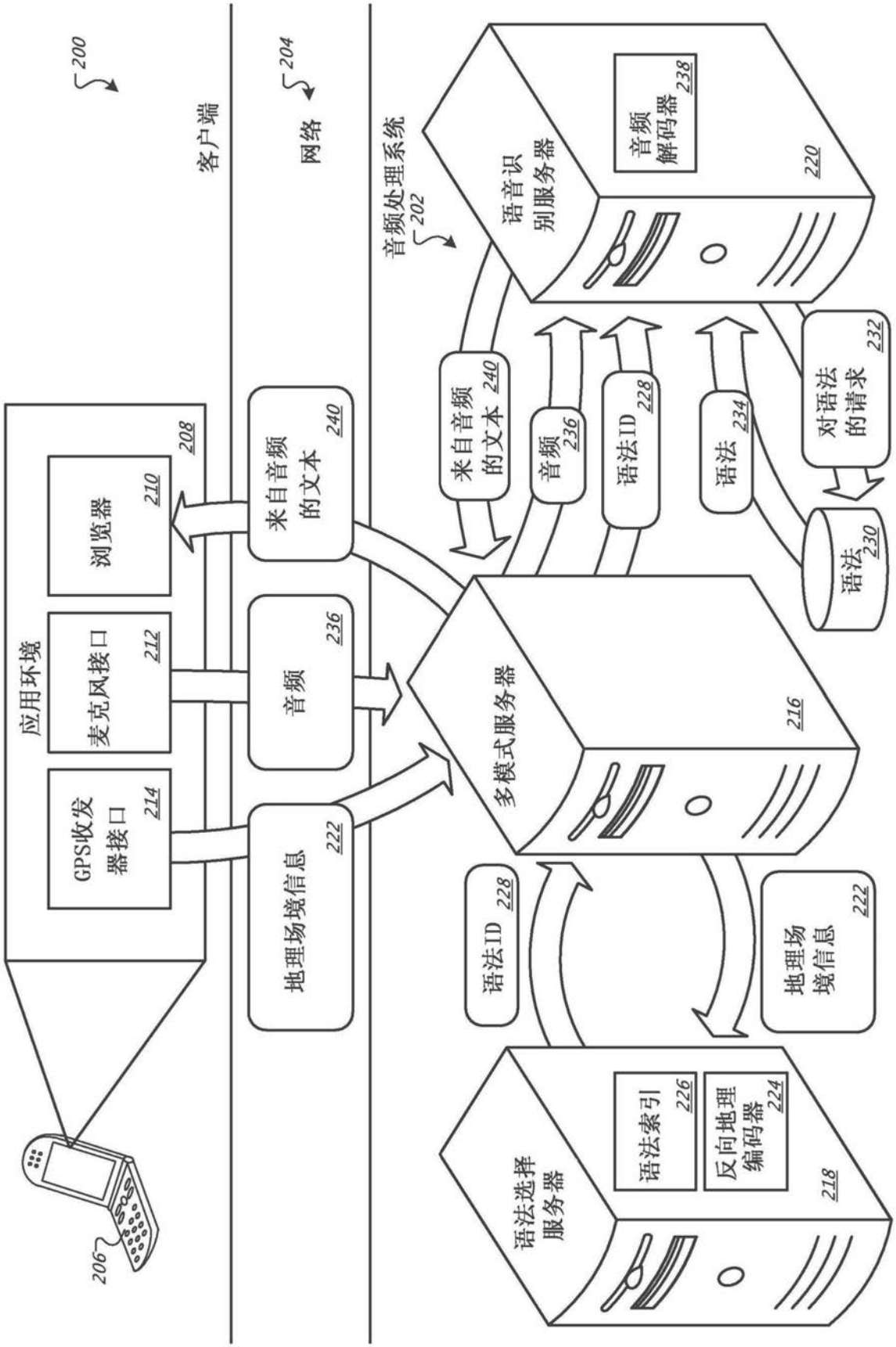

200
客户端
204
网络
音频处理系统
202
应用环境
浏览器
210
208
麦克风接口
212
GPS收发器接口
214
来自音频的文本
240
音频
236
地理场境信息
222
206
多模式服务器
216
语音识别服务器
220
音频解码器
238
来自音频的文本
240
音频
236
语法ID
228
语法
234
对语法的请求
232
语法
230
语法ID
228
地理场境信息
222
语法选择服务器
218
语法索引
226
反向地理编码器
224

图2

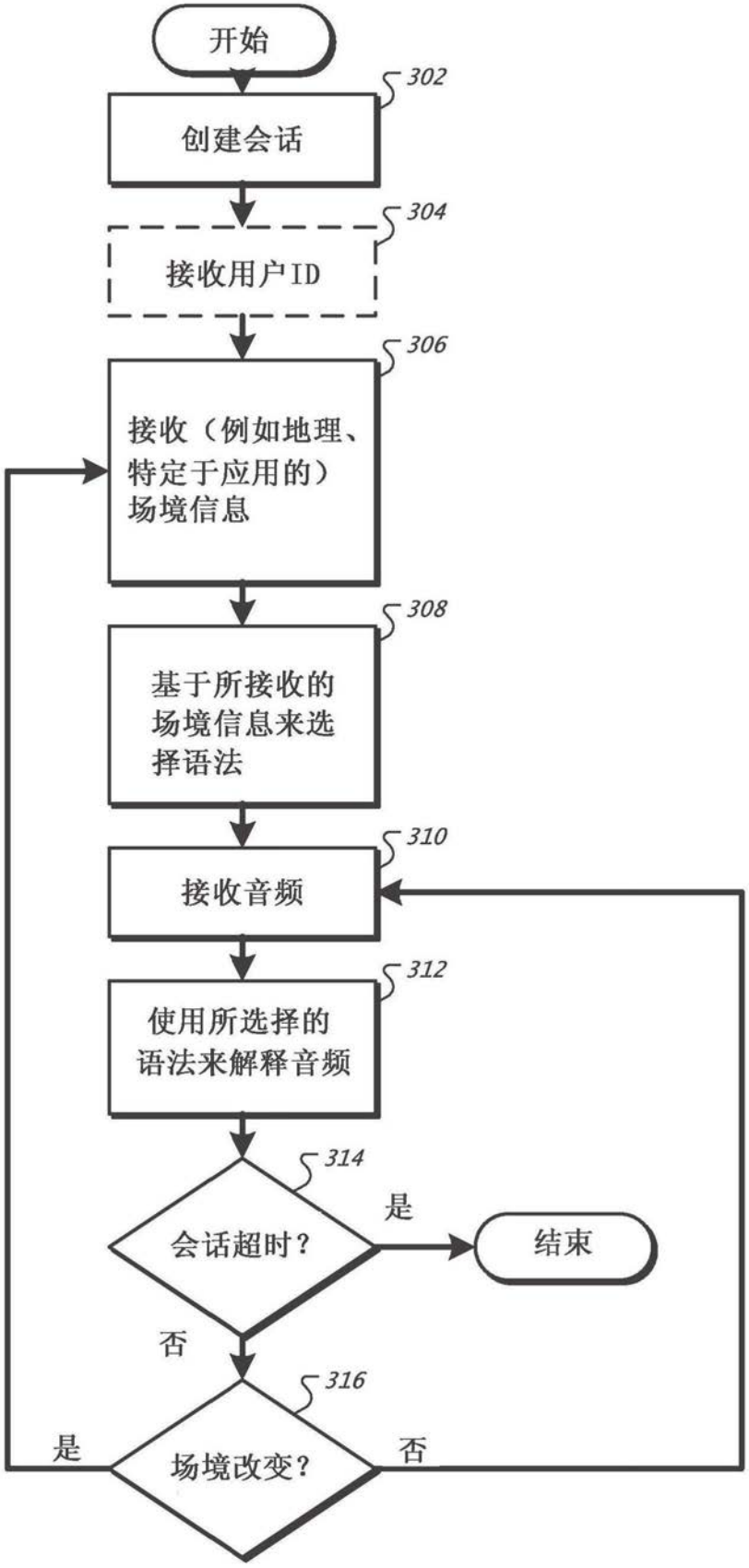

开始
302
创建会话
304
接收用户ID
306
接收（例如地理、特定于应用的）场境信息
308
基于所接收的场境信息来选择语法
310
接收音频
312
使用所选择的语法来解释音频
314
会话超时？
是
结束
否
316
场境改变？
是
否

图3

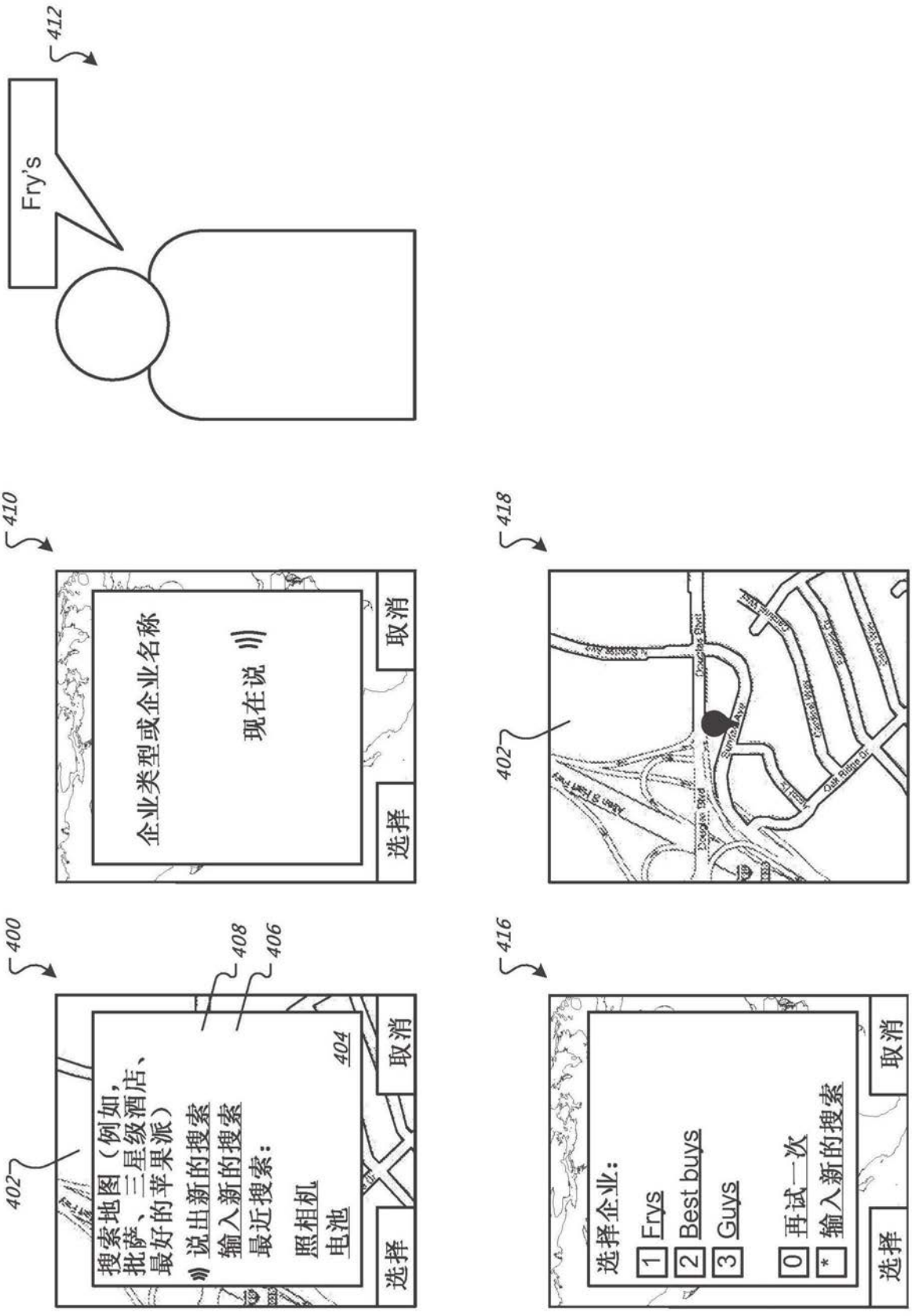

412
Fry's
410
企业类型或企业名称
现在说
选择
取消
418
402
400
搜索地图（例如，批萨、三星级酒店、最好的苹果派）
说出新的搜索
输入新的搜索
最近搜索：
照相机
电池
404
408
406
选择
取消
416
选择企业：
1 Frys
2 Best buys
3 Guys
0 再试一次
* 输入新的搜索
选择
取消

图4

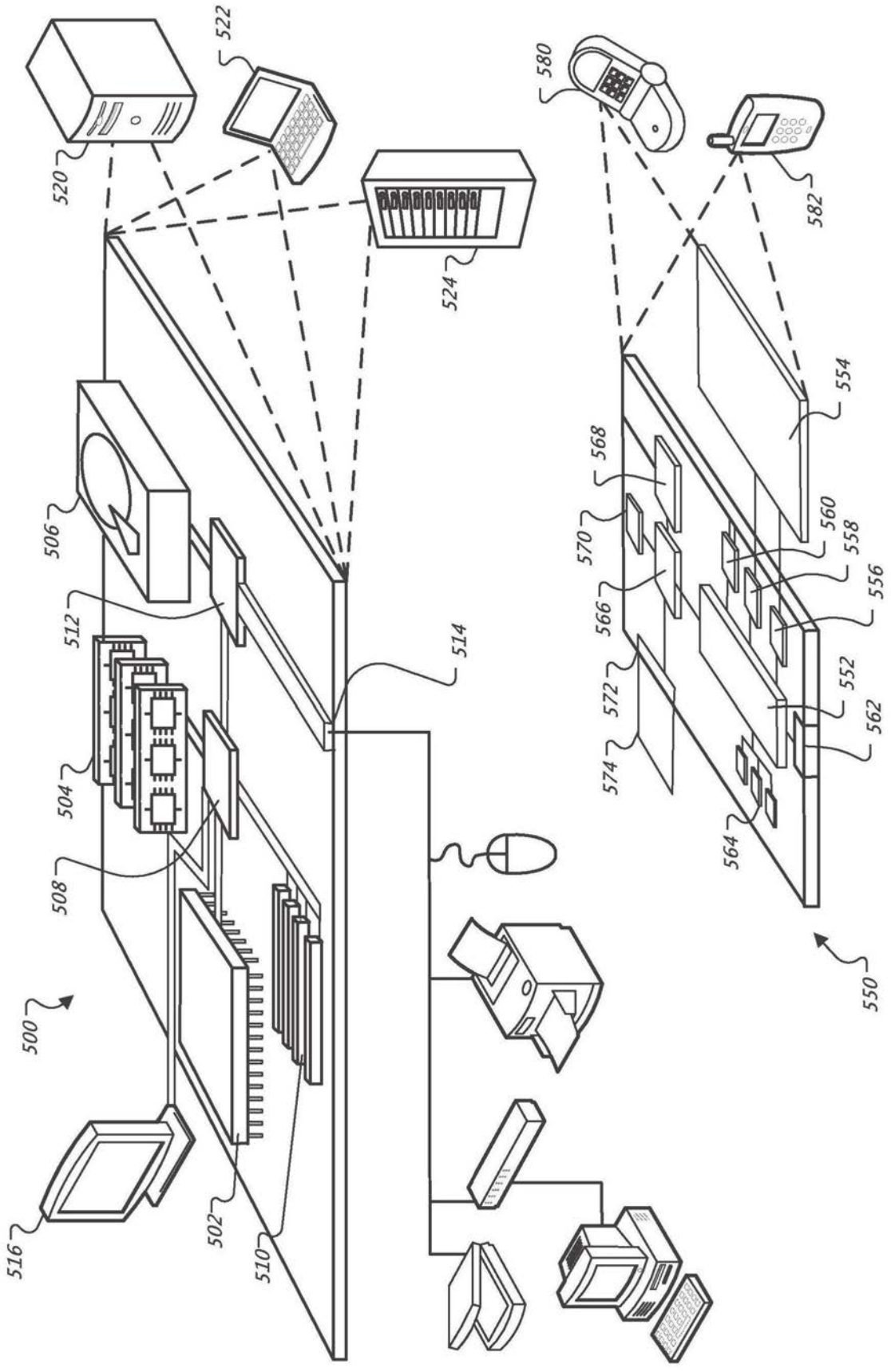

522
520
524
580
582
554
568
560
570
558
566
556
572
552
562
574
564
550
506
512
514
504
508
500
516
502
510

图5