



(12) 发明专利

(10) 授权公告号 CN 110741433 B

(45) 授权公告日 2024. 02. 02

(21) 申请号 201880038644.6

(22) 申请日 2018.09.11

(65) 同一申请的已公布的文献号
申请公布号 CN 110741433 A

(43) 申请公布日 2020.01.31

(30) 优先权数据
15/702,164 2017.09.12 US

(85) PCT国际申请进入国家阶段日
2019.12.11

(86) PCT国际申请的申请数据
PCT/US2018/050344 2018.09.11

(87) PCT国际申请的公布数据
W02019/055372 EN 2019.03.21

(73) 专利权人 谷歌有限责任公司
地址 美国加利福尼亚州

(72) 发明人 桑德罗·弗兹
塞巴斯蒂安·米柳斯
扬·奥尔索斯

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

(51) Int.Cl.
G10L 25/51 (2013.01)
H04L 67/141 (2022.01)
H04L 67/52 (2022.01)
H04W 4/12 (2009.01)
G06F 40/20 (2020.01)
G06F 3/16 (2006.01)
G06N 3/08 (2023.01)
G10L 15/26 (2006.01)

(56) 对比文件
US 2017245125 A1, 2017.08.24
US 2014156833 A1, 2014.06.05
US 2016336009 A1, 2016.11.17
CN 101233559 A, 2008.07.30
US 2014172953 A1, 2014.06.19

审查员 千雪

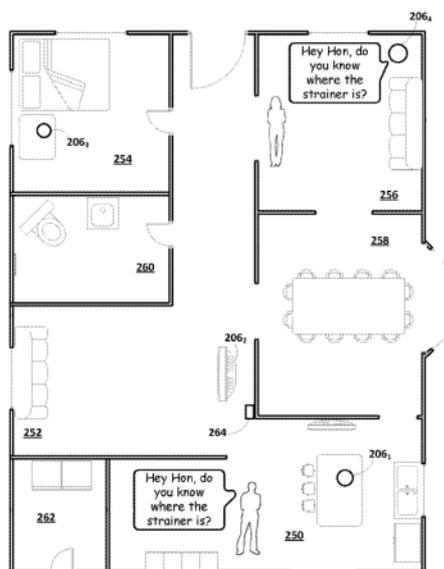
权利要求书2页 说明书19页 附图6页

(54) 发明名称

使用多个计算设备的对讲式通信

(57) 摘要

描述了与使用关于环境分布的多个计算设备的改进的对讲式通信相关的技术。在各种实施方式中,可以例如在多个计算设备的第一计算设备的麦克风处从第一用户接收语音输入。可以分析语音输入,并且基于该分析,可以确定第一用户旨在将消息传达给第二用户。可以确定第二用户相对于多个计算设备的位置,从而基于第二用户的位置,可以从多个计算设备中选择能够提供第二用户可感知的音频或视觉输出的第二计算设备。然后,第二计算设备可以被操作以提供将消息传达给第二用户的音频或视觉输出。



1. 一种用于使用多个计算设备的对讲式通信的方法,包括:

访问已训练的机器学习模型,其中,所述机器学习模型是使用标记的语音输入的语料库来训练的,以预测语音输入是指示应当被忽略的背景对话还是指示将消息传达给一个或多个其他用户的用户意图;

在多个计算设备中的第一计算设备的麦克风处,从第一用户接收语音输入;

分析所述语音输入,其中,所述分析包括将指示所述语音输入的音频记录的数据应用为已训练的机器学习模型上的输入以生成输出,所述输出预测所述语音输入是应当被忽略的背景对话还是所述语音输入指示将消息传达给一个或多个其他用户的用户意图;

基于所述分析,确定所述第一用户旨在将所述消息传达给所述一个或多个其他用户;以及

使得所述多个计算设备中的一个或多个其他计算设备提供将所述消息传达给所述一个或多个其他用户的音频或视觉输出。

2. 根据权利要求1所述的方法,进一步包括:

在所述第一计算设备的所述麦克风处接收附加语音输入;

分析所述附加语音输入,其中,所述分析包括将指示所述语音输入的音频记录的数据应用为已训练的机器学习模型上的输入以生成附加输出,所述附加输出预测所述附加语音输入指示应当被忽略的背景噪音;

响应于所述附加输出指示所述附加语音输入指示应当被忽略的背景噪音,忽略所述附加语音输入。

3. 根据权利要求1所述的方法,其中,所述机器学习模型是使用标记的语音输入的语料库来训练的,以及其中,应用于所述语音输入的标记包括:

第一标记,所述第一标记指示将消息传达给一个或多个其他用户的用户意图;以及

第二标记,所述第二标记指示多个用户之间的背景对话。

4. 根据权利要求3所述的方法,其中,应用于所述语音输入的所述标记进一步包括第三标记,所述第三标记指示加入与自动化助理的人机对话的用户意图。

5. 根据权利要求1所述的方法,进一步包括:

确定所述一个或多个用户中的第二用户相对于所述多个计算设备的位置;以及

基于所述第二用户的位置,从所述多个计算设备中选择能够提供所述第二用户能够感知的音频或视觉输出的第二计算设备;

其中,所述使得包括使得所述第二计算设备提供将所述消息传达给所述第二用户的音频或视觉输出。

6. 根据权利要求1所述的方法,其中,所述使得包括使用所有所述多个计算设备,向所述一个或多个其他用户广播所述消息。

7. 根据权利要求1所述的方法,其中,所述分析包括对所述语音输入的音频记录执行语音到文本处理,以生成文本输入作为指示所述音频记录的数据,其中,所述文本输入被应用为已训练的机器学习的输入。

8. 一种用于使用多个计算设备的对讲式通信的系统,包括一个或多个处理器以及与所述一个或多个处理器可操作地耦合的存储器,其中,所述存储器存储指令,响应于由一个或多个处理器执行指令,所述指令使得所述一个或多个处理器执行下述操作:

访问已训练的机器学习模型,其中,所述机器学习模型是使用标记的语音输入的语料库来训练的,以预测语音输入是指示应当被忽略的背景对话还是指示将消息传达给一个或多个其他用户的用户意图;

在多个计算设备中的第一计算设备的麦克风处,从第一用户接收语音输入;

分析所述语音输入,其中,所述分析包括将指示所述语音输入的音频记录的数据应用为已训练的机器学习模型上的输入以生成输出,所述输出预测所述语音输入是应当被忽略的背景对话还是所述语音输入指示将消息传达给一个或多个其他用户的用户意图;

基于所述分析,确定所述第一用户旨在将所述消息传达给所述一个或多个其他用户;以及

使得所述多个计算设备中的一个或多个其他计算设备提供将所述消息传达给所述一个或多个其他用户的音频或视觉输出。

9. 根据权利要求8所述的系统,进一步包括:

在所述第一计算设备的所述麦克风处接收附加语音输入;

分析所述附加语音输入,其中,所述分析包括将指示所述语音输入的音频记录的数据应用为已训练的机器学习模型上的输入以生成附加输出,所述附加输出预测所述附加语音输入指示应当被忽略的背景噪音;

响应于所述附加输出指示所述附加语音输入指示应当被忽略的背景噪音,忽略所述附加语音输入。

10. 根据权利要求8所述的系统,其中,所述机器学习模型是使用标记的语音输入的语料库来训练的,以及其中,应用于所述语音输入的标记包括:

第一标记,所述第一标记指示将消息传达给一个或多个其他用户的用户意图;以及

第二标记,所述第二标记指示多个用户之间的背景对话。

11. 根据权利要求10所述的系统,其中,应用于所述语音输入的所述标记进一步包括第三标记,所述第三标记指示加入与自动化助理的人机对话的用户意图。

12. 根据权利要求8所述的系统,进一步包括:

确定所述一个或多个用户中的第二用户相对于所述多个计算设备的位置;以及

基于所述第二用户的位置,从所述多个计算设备中选择能够提供所述第二用户能够感知的音频或视觉输出的第二计算设备;

其中,所述使得包括使得所述第二计算设备提供将所述消息传达给所述第二用户的音频或视觉输出。

13. 根据权利要求8所述的系统,其中,所述使得包括使用所有所述多个计算设备,向所述一个或多个其他用户广播所述消息。

14. 根据权利要求8所述的系统,其中,所述分析包括对所述语音输入的音频记录执行语音到文本处理,以生成文本输入作为指示所述音频记录的数据,其中,所述文本输入被应用为已训练的机器学习的输入。

使用多个计算设备的对讲式通信

技术领域

[0001] 本申请涉及用于使用多个计算设备的对讲式通信的方法和系统。

背景技术

[0002] 人们可以使用在本文中称为“自动化助理”(也称为“聊天机器人”、“交互式个人助理”、“智能个人助理”、“个人语音助理”、“会话代理”等)的交互式软件应用加入人机对话。例如,人们(当它们与自动化助理互动时,可以被称为“用户”)可以使用在一些情况下可以被转换为文本然后进行处理的语音自然语言输入(即话语)和/或通过提供文本(例如,键入的)自然语言输入来提供命令、查询和/或请求。

[0003] 在一些情况下,自动化助理可以包括本地安装在客户端设备上并由用户直接加入的自动化助理“客户端”,以及利用云的虚拟的无限资源来帮助自动化助理客户端响应用户的查询的基于云的对方。例如,自动化助理客户端可以向基于云的对方提供用户查询的音频记录(或其文本转换)以及指示用户身份的数据(例如证书)。基于云的对方可以对查询执行各种处理,以将各种结果返回给自动化助理客户端,然后该自动化助理客户端可以向用户提供相应的输出。为简洁和简化起见,术语“自动化助理”在本文中描述为“服务”特定用户时,可以指安装在特定用户的客户端设备上的自动化助理客户端,以及与自动化助理客户端交互以响应用户的查询的任何基于云的对方。

[0004] 许多用户可以使用多个设备来使自动化助理加入。例如,某些用户可能拥有计算设备的协调“生态系统”,其包括一个或多个智能电话、一个或多个平板电脑、一个或多个车辆计算系统、一个或多个可穿戴计算设备、一个或多个智能电视和/或一个或多个独立的交互式扬声器,以及其他更传统的计算设备。用户可以使用这些设备中的任何加入与自动化助理的人机对话(假设已安装自动化助理客户端)。在一些情况下,这些设备可能散布在用户的家中或工作场所中。例如,诸如智能电话、平板电脑、智能手表等的移动计算设备可以在用户的身上和/或用户最后将它们放置的任何地方(例如,在充电站处)。诸如传统的台式计算机、智能电视和独立的交互式扬声器的其他计算设备可能会更固定,但仍然可以位于用户家或工作场所内的各个位置(例如,房间)处。

[0005] 存在使多个用户(例如,家庭、同事、同居者等)能够利用多个计算设备的分布式性质来促进多个用户之间的对讲式语音通信的技术。然而,这些技术限于用户发出明确的命令以将消息传达到明确限定的计算设备。例如,希望将消息传达到能够听到的距离之外的另一位置(例如,在另一房间中)的第二用户的第一用户必须首先确定第二用户的位置。只有这样,第一用户才能明确调用到在第二用户的位置处或附近的计算设备的对讲通信信道,以便第一用户可以将消息传达给第二用户位置处的第二用户。如果第一用户不知道第二用户的位置,则第一用户可能被迫简单地使得消息在可用于对讲式通信的所有计算设备上广播。此外,如果第一用户不知道第二用户不在能够听到的距离内(例如,第一用户正在做饭并且没有注意到第二用户离开厨房),则第一用户可能没有意识到对讲式通信是必要的,并且可能会在空房间说出该消息。

发明内容

[0006] 本文描述了用于使用围绕诸如房屋、公寓、营业场所等的环境分布的多个计算设备的改进的对讲式通信的技术。例如,本文描述了用于使得能够确定环境中内的多个用户的位置,使得 (i) 可以自动地确定语音消息的预期接收者是否在能听到说话者的范围内,以及 (ii) 可以识别预期接收者附近的合适的计算设备,并且将其用于输出消息,以便预期接收者接收它的技术。另外,本文描述了用于自动地确定用户话语是否构成以下内容的技术: (a) 调用自动化助理以正常使用的命令; (b) 将语音消息传达给可能潜在地要求本文所述的对讲式通信的另一用户的尝试;和/或 (c) 不要求动作的其他背景噪音/对话。另外,本文描述了用于允许使用所公开的技术接收到的对讲式消息的接收者发出至少部分地基于说话者传达的原始消息来(例如,使用自然语言处理)处理的请求(例如,对自动化助理搜索查询或其他命令,诸如订购比萨、播放歌曲等)的技术。

[0007] 在各种实施方式中,可以通过使用各种技术,配置有本公开的所选方面的计算设备来确定环境或区域内的用户的位置。例如,一个或多个计算设备可以配备有各种类型的存在传感器,诸如可以确定用户是否在附近的无源红外(“PIR”)传感器、照相机、麦克风、超声波传感器等。这些计算设备可以以诸如智能电话、独立交互式扬声器、智能电视、其他智能电器(例如,智能恒温器、智能冰箱等)、网络摄像机等等的各种形式出现。附加地或替代地,其他类型的信号,诸如用户携带的移动计算设备(例如,智能电话、智能手表)发出的信号可以被其他计算设备检测到,并用于(例如,使用飞行时间、三角测量等)确定用户的位置。确定用户在环境内的位置以用在本文所述的各种技术中可以视用户对此类确定明确提供的授权而定。在各种实施方式中,可以响应于确定用户话语构成将语音消息传达给可能需要对讲式通信的另一用户的尝试来“按需”确定用户的位置。在各种其他实施方式中,可以周期性地和/或以其他间隔来确定用户的位置,以及可以是用在确定语音消息的预期接收者是否在能听到语音消息的说话者的范围内中和/或识别语音消息的预期接收者附近的合适的计算设备中的最近确定的位置。

[0008] 作为一个示例,各种独立的交互式扬声器和/或智能电视可以分布在家中的各个位置。这些设备中的每一个可以包括能够检测附近的人存在的一个或多个传感器(例如,麦克风、照相机、PIR传感器等)。在一些实施例中,这些设备可以简单地检测是否存在人。在其他实施例中,这些设备不仅能够检测存在,而且能够将所检测的人与例如家庭的其他已知成员区分开。由这些独立的交互式扬声器和/或智能电视生成的存在信号可以被收集并用来确定/跟踪人们在特定时间点所处的位置。然后,根据本文所述的技术,这些检测到的位置可以用于各种目的,诸如确定说话者提供的话语是否可能被预期的接收者听到(例如,说话者和预期的接收者是在不同的房间还是在相同的房间),和/或选择应当使用多个扬声器和/或电视中的哪个来将话语输出到预期的接收者。

[0009] 在另一方面中,本文描述了用于自动地确定用户话语是否构成以下内容的技术: (a) 调用自动化助理以正常使用的命令; (b) 将语音消息传达给可能需要本文所述的对讲式通信的另一用户的尝试;和/或 (c) 不需要任何动作的其他背景噪音/会话。在一些实施方式中,可以使用包括录制的话语(和/或录制的话语的特征)的训练示例来训练机器学习分类器(例如,神经网络),录制的话语被分类(标记)为例如使用对讲式通信链路将消息传达另一用户的命令、与自动化助理进行常规人机对话的命令或不直接对自动化助理的会话(例

如背景会话和/或噪音)。

[0010] 在一些实施例中,语音到文本(“STT”)可能不会在每个话语上自动地执行。相反,可以训练机器学习分类器以识别语音输入的音频记录中的音素,并且特别是使用上述标记中的一个对集体音素进行分类。例如,通常使用一个或多个调用短语来调用常规的自动化助理。在一些情况下,训练了简单的调用机器学习模型(例如,分类器),以将这些调用短语与用来确定用户何时调用自动化助理的任何其他短语(例如,识别与“Hey,Assistant(嘿,助理)”相关联的音素)区分开。使用本文描述的技术,可以(进一步)训练相同的调用机器学习模型或不同的机器学习模型,以将话语分类为旨在将消息传达给另一用户,这可能也需要可能不需要使用本文所述的对讲式通信。在一些实施方式中,例如,可以与调用机器学习模型并行地,或者在调用机器学习模型确定用户没有调用自动化助理之后,使用该机器学习模型来确定用户是否可以从使用对讲式通信以使得远程计算设备将消息传达给另一用户中受益。

[0011] 在一些实施方式中,可以训练或者“订做”机器学习模型,以便可以识别用户说出的姓名,并将这些姓名附加到其他个人上。例如,自动化助理可以检测第一话语,诸如“Jan, can you pass me the salt?(Jan,把盐递给我好吗?)”。自动化助理可以检测第二话语,大概来自Jan,诸如“Sure,here you go(好,给你)”。从这些话语和相关的音素中,自动化助理可以学习到当用户向Jan发出请求时,应当利用Jan的语音来定位该人。假设稍后,Jan在分开的房间里打电话。当用户说出如“Jan,where are my shoes(Jan,我的鞋子在哪里)”的内容时,自动化助理可以从该话语(特别是“Jan,...”)确定该话语包含针对个人Jan的消息。自动化助理还可以确定Jan可能不在能够听到的范围内,因此该消息应当作为对讲消息传达给Jan。通过检测附近的客户端设备上的Jan的语音,自动化助理可以定位Jan并且选择附近的客户端设备以输出说话者的消息。

[0012] 在其他实施方式中,用户可以使用传统的调用短语来调用自动化助理,然后明确地命令自动化助理以使得一些其他计算设备输出要传达给接收者的消息。如上所述,可以基于接收者的检测到的位置来自动地选择另一计算设备,或者由说话的用户明确地指定另一计算设备。

[0013] 在另一方面中,本文描述了用于允许使用所公开的技术接收的对讲式消息的接收者利用所接收的对讲消息中提供的场境(context)来执行其他动作,诸如向自动化助理发出搜索查询或命令的技术。例如,在感知所传达的对讲消息之后,接收者可以例如在她接收到所传达的对讲消息的计算设备或另一计算设备处发出搜索查询。然后例如可以通过服务于第二用户的自动化助理来获得响应于搜索查询的搜索结果。在一些实施方式中,可以至少部分地基于原始传达的对讲消息的内容来偏向或排名搜索结果。附加地或替代地,在一些实施方式中,可以至少部分地基于原始传达的对讲消息的内容来消除接收者的搜索查询的歧义。

[0014] 在原始话语被接收者用户用来向下游请求提供场境以保护隐私的一些实施方式中,仅在确定接收者做出下游请求时才可以转录(STT)原始说话者的话语。如果接收者简单地收听消息而没有做进一步动作,则不会执行SST。在其他实施方式中,可以总是使用SST来处理原始说话者的话语(例如,在确定要通过对讲式通信来传达话语时),但是所得的转录可以仅本地存储和/或存储有限的时间量(例如足够长的时间以给接收者用户有足够的时

间来做出一些下游请求)。

[0015] 在一些实施方案中,一个或多个计算装置可以等待直到预期用户能够感知(例如在能够听到的范围内的)消息,直到他们使用本文所述的技术传达消息为止。例如,假设第一用户将消息传达给了预期的接收者,但是预期的接收者已经暂时离开。在一些实施方式中,检测接收者的第一计算设备在接收者返回时可以输出原始消息。

[0016] 在一些实施方式中,提供了一种由一个或多个处理器执行的方法,该方法包括:在多个计算设备中的第一计算设备的麦克风处从第一用户接收语音输入;分析语音输入;基于该分析,确定第一用户旨在向第二用户传达消息;确定第二用户相对于多个计算设备的位置;基于第二用户的位置,从多个计算设备中选择能够提供第二用户可感知的音频或视觉输出的第二计算设备;以及使得第二计算设备排他地提供将消息传达给第二用户的音频或视觉输出(例如,只有第二计算设备提供输出,其他计算设备除外)。

[0017] 本文公开的技术的这些和其他实施方式可以可选地包括以下特征中的一个或多个。

[0018] 在各种实施方式中,分析可以包括将语音输入的音频记录应用为已训练的机器学习模型上的输入以生成输出,其中,输出指示第一用户旨在将消息传达给第二用户。在各种实施方式中,可以使用标记的话语的语料库来训练机器学习模型,以及其中,应用到话语的标记包括指示向另一用户传达消息的命令的第一标记和指示加入与自动化助理的人机对话的第二标记的命令。在各种实施方式中,应用于话语的标记可以进一步包括指示背景会话的第三标记。

[0019] 在各种实施方式中,可以响应于基于第二用户的位置确定第二用户不在能够听到第一用户的范围内来执行选择。在各种实施方式中,可以至少部分地基于由第二用户操作的移动计算设备生成的一个或多个信号来确定第二用户的位置。本领域技术人员通过阅读说明书将理解到,本文所述的概念和主题可以确保以对用于传达和接收消息的技术设备有效的方式,将消息传达给预期的人并由其接收。这可以包括消息被预期的人在适当的时间传达和传递,以使预期的人可以充分地理解该消息,并且不需要消息被为此目的的技术设备重新传达/接收。技术设备可以包括上面提到的多个计算设备,以及可以在设备之间传达消息的网络。传达消息的方式和时间的效率可以至少使得在计算设备之间更有效地使用网络,以及更有效地使用计算设备内、被采用来传达和接收消息的计算资源。

[0020] 在各种实施方式中,可以至少部分基于由除第一计算设备以外的多个计算设备中的一个或多个生成的一个或多个信号来确定第二用户的位置。在各种实施方式中,一个或多个信号可以包括指示由除第一计算设备之外的多个计算设备中的一个或多个使用无源红外或超声检测到第二用户的第二信号。在各种实施方式中,一个或多个信号可以包括指示由除第一计算设备之外的多个计算设备中的一个或多个使用相机或麦克风检测到第二用户的第二信号。

[0021] 在各种实施方式中,分析可以包括确定语音输入包括经由多个计算设备中的一个或多个,将消息作为对讲消息传达给第二用户的显式命令。在各种实施方式中,分析可以包括对语音输入执行语音到文本处理以生成文本输入,以及对文本输入执行自然语言处理以确定用户旨在将消息传达给第二用户。

[0022] 在各种实施方式中,该方法可以进一步包括:在第二计算设备提供音频或视觉输

出之后,识别由第二用户发出的搜索查询;获得响应于该搜索查询的搜索结果,其中,该获得至少部分地基于来自第一用户的语音输入;以及使得多个计算设备中的一个或多个来提供指示至少一些搜索结果的输出。

[0023] 在另一方面中,一种方法可以包括:访问已训练的机器学习模型,其中,使用标记的语音输入的语料库训练机器学习模型,以预测语音输入是指示应当被忽略的背景对话还是指示向一个或多个其他用户传达消息的用户意图;

[0024] 在多个计算设备中的第一计算设备的麦克风处,从第一用户接收语音输入;分析语音输入,其中,分析包括将指示语音输入的音频记录的数据应用为已训练的机器学习模型上的输入以生成输出,其中,输出指示第一用户旨在将消息传达给一个或多个其他用户;基于分析,确定第一用户旨在将消息传达给一个或多个其他用户;以及使得多个计算设备中的一个或多个其他计算设备提供将消息传达给一个或多个其他用户的音频或视觉输出。

[0025] 在各种实施方式中,该方法可以进一步包括:在第一计算设备的麦克风处接收附加语音输入;分析附加语音输入,其中,分析包括将指示语音输入的音频记录的数据应用为已训练的机器学习模型上的输入以生成附加输出,其中,附加输出指示附加语音输入指示应当被忽略的背景噪音;响应于附加输出指示附加语音输入指示应当被忽略的背景噪音,忽略附加语音输入。

[0026] 在各种实施方式中,使用标记的语音输入的语料库来训练机器学习模型,以及其中,应用于语音输入的标记包括:第一标记,第一标记指示将消息传达到一个或多个其他用户的用户意图;以及第二标记,第二标记指示多个用户之间的背景对话。在各种实施方式中,应用于语音输入的标记进一步包括第三标记,第三标记指示加入与自动化助理的人机对话的用户意图。

[0027] 在各种实施方式中,该方法可以进一步包括:确定一个或多个用户中的第二用户相对于多个计算设备的位置;以及基于第二用户的位置,从多个计算设备中选择能够提供第二用户可感知的音频或视觉输出的第二计算设备。在各种实施方式中,使得包括使得第二计算设备提供将消息传达给第二用户的音频或视觉输出。

[0028] 在各种实施方式中,使得包括使用所有多个计算设备,向一个或多个其他用户广播消息。在各种实施方式中,分析包括对语音输入的音频记录执行语音到文本处理,以生成作为为指示音频记录的数据的文本输入,其中,文本输入被应用为已训练的机器学习的输入。

[0029] 在另一方面中,一种方法可以包括:在多个计算设备中的第一计算设备的麦克风处,从第一用户接收语音输入;分析语音输入;

[0030] 基于分析,确定第一用户旨在向第二用户传达消息;使得多个计算设备中的一个或多个其他计算设备提供将消息传达给第二用户的音频或视觉输出;在多个计算设备中的一个或多个其他计算设备提供音频或视觉输出之后,识别第二用户在多个计算设备的第二计算设备处发出的搜索查询;至少部分地基于来自第一用户的语音输入,消除由第二用户发出的搜索查询的至少一部分的歧义;获得响应于已消除歧义的搜索查询的搜索结果;以及使得其他计算设备中的一个或多个提供指示至少一些搜索结果的输出。

[0031] 在各种实施方式中,使得多个计算设备中的一个或多个其他计算设备提供将消息传达给第二用户的音频或视觉输出包括:确定第二用户相对于多个计算设备的位置;基于

第二用户的位置,从多个计算设备中选择能够提供第二用户可感知的音频或视觉输出的第二计算设备;以及使得第二计算设备排他地提供将消息传达给第二用户的音频或视觉输出。在各种实施方式中,至少部分地基于由除第一计算设备外的多个计算设备中的一个或多个生成的一个或多个信号来确定第二用户的位置。在各种实施方式中,一个或多个信号包括指示第二用户被除第一计算设备以外的多个计算设备中的一个或多个使用无源红外或超声检测到的信号。在各种实施方式中,一个或多个信号包括指示第二用户被除第一计算设备以外的多个计算装置中的一个或多个使用相机或麦克风检测到的信号。

[0032] 在各种实施方式中,分析包括确定语音输入包括经由多个计算设备中的一个或多个,将消息作为对讲消息传达给第二用户的显式命令。

[0033] 此外,一些实施方式包括一个或多个计算设备的一个或多个处理器,其中,一个或多个处理器可操作以执行在相关联的存储器中存储的指令,以及其中,配置该指令以使得执行上述方法的任何一个。一些实施方式还包括一个或多个非暂时性计算机可读存储介质,其存储被一个或多个处理器可执行来执行上述方法的任何一个的计算机指令。

[0034] 应当理解到,本文更详细所述的前述概念和附加概念的所有组合被认为是本文公开的主题的一部分。例如,出现在本公开的结尾处的所要求保护的的主题的所有组合被认为是本文公开的主题的一部分。

附图说明

[0035] 图1A是可以实施本文公开的实施方式的示例性环境的框图。

[0036] 图1B示意性地描绘了根据各种实现方式,如何应用已训练的分类器来基于用户话语和/或位置生成输出的一个示例。

[0037] 图2、3和4根据各种实施方式,描绘了各种用户和自动化助理之间的示例性对话,包括对讲式通信。

[0038] 图5描绘了示出根据本文公开的实施方式的示例性方法的流程图。

[0039] 图6示出了计算设备的示例性架构。

具体实施方式

[0040] 现在转到图1A,示出了可以实施本文公开的技术的示例性环境。示例性环境包括多个客户端计算设备 106_{1-N} 。每个客户端设备106可以执行自动化助理客户端118的相应实例。可以在经由通常以110所示的一个或多个局域网和/或广域网(例如互联网)通信地耦合到客户端设备 106_{1-N} 的一个或多个计算系统(统称为“云”计算系统)上实施一个或多个基于云的自动化助理组件119,诸如自然语言处理器122。而且,在一些实施例中,多个客户端设备 106_{1-N} 可以经由一个或多个局域网(“LAN”,包括Wi-Fi LAN、网状网络等)彼此通信地耦合。

[0041] 在一些实施方式中,多个客户端计算设备 106_{1-N} (在本文中也简称为“客户端设备”)可以以各种方式彼此关联,以便于执行本文所述的技术。例如,在一些实施方式中,多个客户端计算设备 106_{1-N} 可以借助于经由一个或多个LAN通信地耦合而彼此关联。例如,在多个客户端计算设备 106_{1-N} 跨诸如家、建筑物、校园等的特定区域或环境部署的情况下可能就是这种情况。附加地或可替代地,在一些实施方式中,多个客户端计算设备 106_{1-N} 可以借

助于它们是由一个或多个用户(例如,个人、家庭、组织的员工、其他预定义的小组等)操作的客户端设备106的协调生态系统的成员而彼此关联。

[0042] 如背景技术中所提及的,通过与一个或多个基于云的自动化助理组件119的交互,自动化助理客户端118的实例可以形成从用户的角度看似乎是自动化助理120的逻辑实例,用户可以加入与该自动化助理120的人机对话。在图1A中描绘了这种自动化助理120的两个实例。虚线包围的第一自动化助理120A服务于操作第一客户端设备106₁的第一用户(未示出),并且包括自动化助理客户端118₁和一个或多个基于云的自动化助理组件119。点划线包围的第二自动化助理120B服务于操作另一客户端设备106_N的第二用户(未示出),并且包括自动化助理客户端118_N和一个或多个基于云的自动化助理组件119。因此,应当理解到,与在客户端设备106上执行的自动化助理客户端118进行交互的每个用户实际上可以与她或他自己的自动化助理120的逻辑实例进行交互。为了简洁和简化起见,在本文中用作“服务于”特定用户的术语“自动化助理”将指在由用户操作的客户端设备106上执行的自动化助理客户端118和一个或多个基于云的自动化助理组件119(可以在多个自动化助理客户端118之间共享)的组合。还应当理解到,在一些实施方式中,自动化助理120可以响应于来自任何用户的请求,而不管该用户是否实际上由自动化助理120的该特定实例“服务”。

[0043] 客户端设备106_{1-N}可以包括例如下述中的一个或多个:台式计算设备、手提计算设备、平板计算设备、移动电话计算设备、用户的车辆的计算设备(例如,车载通信系统、车载娱乐系统、车载导航系统)、独立的交互式扬声器、诸如智能电视的智能电器和/或包括计算设备的用户的可穿戴装置(例如,具有计算设备的用户的手表、具有计算设备的用户的眼镜、虚拟或者增强现实计算设备)。可以提供附加和/或替选的客户端计算设备。

[0044] 在各种实施方式中,一个或多个客户端计算设备106_{1-N}可以包括一个或多个存在传感器105_{1-N},其被配置为提供指示所检测到的存在,特别是人类存在的信号。存在传感器105_{1-N}可以具有各种形式。一些客户端设备106可以配备有一个或多个数字相机,其被配置为捕获并提供指示在其视场中检测到的运动的信号。附加地或替代地,一些客户端设备106可以配备有其他类型的基于光的存在传感器105,诸如测量从其视场内的物体发出的红外(“IR”)光的无源红外(“PIR”)传感器。附加地或替代地,一些客户端设备106可以配备有检测声学(或压力)波的存在传感器105,诸如一个或多个麦克风。

[0045] 附加地或替代地,在一些实施方式中,存在传感器105可以被配置为检测与人类存在相关联的其他现象。例如,在一些实施例中,客户端设备106可以配备有存在传感器105,该存在传感器105检测由例如特定用户携带/操作的移动客户端设备106所发出的各种类型的波(例如,无线电波、超声波、电磁波等)。例如,一些客户端设备106可以被配置为发出人类不可感知的波,诸如超声波或红外波,其可以被其他客户端设备106检测到(例如,经由诸如具有超声波能力的麦克风的超声/红外接收器)。

[0046] 附加地或替代地,各种客户端设备106可以发出其他类型的人类无法感知的波,诸如可以由一个或多个其他客户端设备106检测到并用于确定操作用户的特定位置的无线电波(例如,Wi-Fi、蓝牙、蜂窝等)。在一些实施方式中,Wi-Fi三角测量可用于例如基于到/来自客户端设备106的Wi-Fi信号来检测人的位置。在其他实施方式中,诸如飞行时间、信号强度等的其他无线信号特性可以由各种客户端设备106单独或共同使用,以基于由他们携带的客户端设备106发出的信号来确定特定人的位置。

[0047] 附加地或替代地,在一些实施方式中,一个或多个客户端设备106可以执行语音识别以从其语音中识别个人。例如,出于提供/限制对各种资源的访问的目的,一些自动化助理120可以被配置为使语音与用户的简档匹配。在一些实施方式中,然后例如可以通过可以被并入例如灯、灯开关、智能恒温器、安全摄像机等中的一个或多个其他存在传感器来跟踪说话者的运动。在一些实施方式中,基于这些检测到的运动,可以预测个体的位置,并且当其他个体(即,说话者)向第一个体用消息提供话语时,可以假设该位置是个体的位置。在一些实施方式中,可以简单地假设个体处于他或她与自动化助理120交互的最后位置,尤其是如果自从上一次交互以来没有经过太多时间。

[0048] 每个客户端计算设备 106_{1-N} 可以操作各种不同的应用,诸如,多个消息交换客户端 107_{1-N} 中的相应一个。消息交换客户端 107_{1-N} 可以采取各种形式,并且形式可能在客户端计算设备 106_{1-N} 之间而不同和/或可以在客户端计算设备 106_{1-N} 中的单个客户端计算设备 106_{1-N} 上操作多种形式。在一些实施方式中,一个或多个消息交换客户端 107_{1-N} 可以采取短消息服务(“SMS”)和/或多媒体消息服务(“MMS”)客户端、在线聊天客户端(例如,即时通讯软件、互联网中继聊天或者“IRC”等)、与社交网络相关联的消息传递应用、专用于与自动化助理120会话的个人助理消息传递服务等。在一些实施方式中,一个或多个消息交换客户端 107_{1-N} 可以经由网页或者由客户端计算设备106的web浏览器(未描绘)或者其它应用渲染的其它资源来实施。

[0049] 如在本文中更详细地所述,自动化助理120经由一个或多个客户端设备 106_{1-N} 的用户接口输入和输出设备来加入和一个或多个用户的人机对话会话。在一些实施方式中,自动化助理120可以响应于由用户经由客户端设备 106_{1-N} 中的一个客户端设备的一个或多个用户接口输入设备提供的用户接口输入来加入和用户的人机对话会话。在这些实施方式的一些中,用户接口输入明确地指向自动化助理120。例如,消息交换客户端 107_{1-N} 中的一个可以是专用于与自动化助理120会话的个人助理消息传递服务,以及可以将经由该个人助理消息传递服务提供的用户接口输入自动地提供给自动化助理120。同样,例如,基于指示要调用自动化助理120的特定用户接口输入,在一个或多个消息交换客户端 107_{1-N} 中,可以将用户接口输入明确地指向自动化助理120。例如,特定用户接口输入可以是一个或多个键入的字符(例如,@AutomatedAssistant)、与硬件按钮和/或虚拟按钮的用户交互(例如,敲击、长击)、口头命令(例如,“Hey Automated Assistant(你好,自动化助理)”)和/或其它特定用户接口输入。

[0050] 在一些实施方式中,自动化助理120可以响应于用户接口输入来加入对话会话,即使在未将用户接口输入明确地指向自动化助理120时。例如,自动化助理120可以检查用户接口输入的内容并且响应于存在于用户接口输入中的某些术语并且/或基于其它线索来加入对话会话。在许多实施方式中,自动化助理120可以加入交互式语音响应(“IVR”),从而用户可以发出命令、搜索等,并且自动化助理可以利用自然语言处理和/或一个或多个语法来将话语转换为文本,并且相应地对该文本做出响应。在一些实施方式中,自动化助理120可以附加地或替代地对话语做出响应而无需将话语转换为文本。例如,自动化助理120可以将语音输入转换成(指示存在于语音输入中的实体的)实体表示和/或其他“非文本”表示中的嵌入,并且对这种非文本表示进行操作。因此,本文描述为基于从语音输入转换的文本进行操作的实施方式可以附加地和/或替代地直接对语音输入和/或语音输入的其他非文本表

示进行操作。

[0051] 客户端计算设备106_{1-N}和操作基于云的自动化助理组件119的计算设备中的每一个可以包括用于存储数据和软件应用的一个或多个存储器、用于访问数据并且执行应用的一个或多个处理器、以及便于通过网络进行通信的其它组件。可以由一个或多个客户端计算设备106_{1-N}和/或由自动化助理120执行的操作可以被分布在多个计算机系统中。例如,自动化助理120可以被实施为在通过网络彼此耦合的一个或多个位置中的一个或多个计算机上运行的计算机程序。

[0052] 如上所述,在各种实施方式中,每个客户端计算设备106_{1-N}可以操作自动化助理客户端118。在各种实施例中,每个自动化助理客户端118可以包括相应的语音捕获/文本到语音转换(“TTS”)/STT模块114。在其他实施方式中,语音捕获/TTS/STT模块114的一个或多个方面可以与自动化助理客户端118分开地实施。

[0053] 每个语音捕获/TTS/STT模块114可以被配置为执行一个或多个功能:例如经由麦克风(在一些情况下可以包括存在传感器105)来捕获用户的语音;将捕获的音频转换为文本(和/或其他表示或嵌入);和/或将文本转换为语音。例如,在一些实施方式中,因为客户端设备106可能在计算资源(例如,处理器周期、存储器、电池等)方面相对受限,每个客户端设备106本地的语音捕获/TTS/STT模块114可以被配置为将有限数量的不同语音短语——尤其是调用自动化助理120和/或对讲式通信的短语——转换为文本(或其他形式,诸如低维度的嵌入)。其他语音输入可以被发送到基于云的自动化助理组件119,其可以包括基于云的TTS模块116和/或基于云的STT模块117。

[0054] 在一些实施方式中,可以有意地在例如借助于在同一LAN上彼此关联的一个或多个客户端设备106上排他地操作对如本文所述的对讲式通信的实施方式有用的组件。在一些这样的实施方式中,可以在例如,在互联网防火墙之后,在一个或多个客户端设备106上训练和/或存储本文中其他地方描述的任何机器学习模型,以便可以保护由机器学习生成或与之相关联的训练数据和其他信息的隐私。并且在一些这样的实施方式中,在调用对讲式通信中,可以不涉及基于云的STT模块117、基于云的TTS模块116和/或自然语言处理器122的基于云的方面。

[0055] 基于云的STT模块117可以被配置为利用云的虚拟上无限的资源来将由语音捕获/TTS/STT模块114捕获的音频数据转换成文本(然后可以将其提供给自然语言处理器122)。基于云的TTS模块116可以被配置为利用云的虚拟上无限的资源来将文本数据(例如,由自动化助理120制定的自然语言响应)转换为计算机生成的语音输出。在一些实施方式中,TTS模块116可以将计算机生成的语音输出提供给客户端设备106以例如使用一个或多个扬声器来直接输出。在其他实施方式中,可以将由自动化助理120生成的文本数据(例如,自然语言响应)提供给语音捕获/TTS/STT模块114,然后,可以将文本数据转换成本地输出的计算机生成的语音。

[0056] 自动化助理120(并且特别是基于云的自动化助理组件119)可以包括自然语言处理器122、上述TTS模块116、上述STT模块117以及其他组件,在下文中,将更详细地描述其中一些。在一些实施方式中,可以在与自动化助理120分离的组件中省略、组合和/或实施自动化助理120的一个或多个引擎和/或模块。并且如上所述,在一些实施方式中,为了保护隐私,可以至少部分在客户端设备106上(例如,排除云)实施自动化助理120的一个或多个组

件,诸如自然语言处理器122、语音捕获/TTS/STT模块114等。在一些这样的实施方式中,语音捕获/TTS/STT模块114可以被充分地配置为执行本公开的所选方面以实现对讲式通信,而在一些情况下,当适合时,将其他非对讲相关的自然语言处理方面留给基于云的组件。

[0057] 在一些实施方式中,自动化助理120响应于在与自动化助理120的人机对话会话期间,由客户端设备106_{1-N}中的一个的用户生成的各种输入来生成响应内容。自动化助理120可以提供响应内容(例如,当与用户的客户端设备分开时通过一个或多个网络),以作为对话会话的一部分呈现给用户。例如,自动化助理装置120可以响应于经由客户端设备106_{1-N}中的一个提供的自由形式的自然语言输入而生成响应内容。如本文所使用的,自由形式的输入是由用户制定的并且不限于呈现给用户以选择的一组选项的输入。

[0058] 如本文所使用的,“对话会话”可以包括在用户与自动化助理120(以及在一些情况下,其他人类参与者)之间进行一个或多个消息的逻辑独立(logically-self-contained)交换。自动化助理120可以基于诸如,会话之间的时间推移、会话之间的用户场境(例如,位置、在安排会议之前/期间/之后等)的变化、检测到除了用户与自动化助理之间的对话之外的用户与客户端设备之间的一个或多个中间交互(例如,用户暂时切换应用、用户离开,然后回到独立的语音激活的产品)、会话之间的客户端设备的锁定/休眠、用于与自动化助理120的一个或多个实例接口的客户端设备的变化等的各种信号来区分与用户的多个对话会话。

[0059] 自动化助理120的自然语言处理器122处理由经由客户端设备106_{1-N},由用户生成的自然语言输入,并且可以生成供自动化助理120的一个或多个其它组件使用的注释输出。例如,自然语言处理器122可以处理由用户经由客户端设备106₁的一个或多个用户接口输入设备生成的自然语言自由形式输入。所生成的注释输出包括自然语言输入的一个或多个注释,并且可选地包括自然语言输入的一个或多个(例如,所有)术语。

[0060] 在一些实施方式中,自然语言处理器122被配置为识别和注释自然语言输入中的各种语法信息。例如,自然语言处理器122可以包括词性标记器的部分,该词性标记器被配置为用其语法角色来注释术语。例如,词性标记器的该部分可以用诸如,“名词”、“动词”、“形容词”、“代词”等的其词性来标记每个术语。同样,例如,在一些实施方式中,自然语言处理器122可以附加和/或替选地包括相关性分析器(未示出),该相关性分析器配置为确定自然语言输入中的术语之间的句法关系。例如,相关性分析器可以确定哪些术语修饰句子的其它术语、主语和动词(例如,解析树),并且可以做出这种相关性的注释。

[0061] 在一些实施方式中,自然语言处理器122可以附加地和/或替选地包括实体标记器(未示出),该实体标记器被配置为注释一个或多个段中的实体引用,诸如,对人(例如,包括文学人物、名人、社会名人)、组织、位置(真实与虚构)等的引用。在一些实施方式中,关于实体的数据可以存储在一个或多个数据库中,诸如存储在知识图(未示出)中。在一些实施方式中,知识图可以包括表示已知实体(在一些情况下,还表示实体属性)的节点,以及连接节点并表示实体之间关系的边。例如,“banana(香蕉)”节点可以(例如作为子节点)被连接到“fruit(水果)”节点,“fruit”节点继而可以(例如作为子节点)被连接到“produce(产品)”和/或“food(食物)”节点。作为另一示例,称为“Hypothetical Café(Hypothetical咖啡厅)”的餐馆可以由还包括诸如其地址、所供应的食物的类型、营业时间、联系信息等属性的节点表示。在一些实施方式中,“Hypothetical Café”节点可以通过边(例如,表示父子关

系)连接到一个或多个其他节点,诸如“restaurant(餐馆)”节点、“business(企业)”节点、表示餐厅所在的城市和/或州的节点等。

[0062] 自然语言处理器122的实体标记器可以以高粒度级(例如使得能够识别对诸如人的实体类的所有引用)和/或以低粒度级(例如,使得能够识别对诸如特定人的特定实体的所有引用),标记对实体的引用。实体标记器可以依赖自然语言输入的内容以消解特定实体和/或可以可选地与知识图或者其它实体数据库通信以解析特定实体。

[0063] 在一些实施方式中,自然语言处理器122可以附加和/或替选地包括共指消解器(未示出),该共指消解器被配置为基于一个或多个语境线索来分组或者“聚类”对相同的实体的引用。例如,可以利用共指消解器来将自然语言输入“I liked Hypothetical Café last time we ate there(我喜欢我们上次去过的Hypothetical Café)”中的术语“there”解析为“Hypothetical Café”。

[0064] 在一些实施方式中,自然语言处理器122的一个或多个组件可以依赖来自自然语言处理器122的一个或多个其它组件的注释。例如,在一些实施方式中,指定的实体标记器在注释所有提及的特定实体时,可以依赖来自共指消解器和/或相关性分析器的注释。同样,例如,在一些实施方式中,共指消解器可以在聚类对相同实体的引用时,依赖来自相关性分析器的注释。在一些实施方式中,在处理特定自然语言输入时,自然语言处理器122的一个或多个组件可以使用除了特定自然语言输入之外的相关在前输入和/或其它相关数据来确定一个或多个注释。

[0065] 在各种实施方式中,基于云的自动化助理组件119可以包括对讲通信分析服务(“ICAS”)138和/或对讲通信位置服务(“ICLS”)140。在其他实施方式中,例如,可以在一个或多个客户端设备106和/或另一计算机系统中(例如,在所谓的“云”中),与基于云的自动化助理组件119分开地实施服务138和/或140。

[0066] 在各种实施方式中,ICAS138可以被配置为基于各种信号和/或数据点来确定如何和/或何时便于使用多个客户端设备106的多个用户之间的对讲式通信。例如,在各种实施方式中,ICAS138可以被配置为分析由第一用户在多个相关联的客户端设备106_{1-N}的客户端设备106的麦克风处提供的语音输入。在各种实施方式中,ICAS138可以分析第一用户的语音输入并基于该分析确定该语音输入包含旨在用于第二用户的消息。

[0067] 可以采用各种技术作为分析的一部分,以确定第一用户是否旨在向第二用户传达消息。在一些实施方式中,第一用户的语音输入的音频记录可以被应用为已训练的机器学习分类器上的输入以生成输出。输出可以指示第一用户的语音输入包含旨在用于第二用户的消息。可以训练各种类型的机器学习分类器(或更一般地,“模型”)以提供这种输出,包括但不限于各种类型的神经网络(例如,前馈、卷积等)。

[0068] 在一些实施方式中,可以使用用户话语的标记音素来训练诸如神经网络的机器学习模型,以学习将话语嵌入到较低维度的表示中。然后可以使用可以包括原始音素的较低维度的表示的这些嵌入(例如,作为训练模型的输入)来识别用户旨在何时使用本文所述的对讲式通信,和/或用户的话语何时包含旨在用于另一人的消息。例如,标记的话语可以被嵌入到降低维度的空间中,例如,使得它们被聚类为与对讲式通信和非对讲式通信相关联的组。然后,可以嵌入新的、未标记的话语,并且可以基于其嵌入(例如,在欧几里得空间中)最接近哪一集群进行分类。

[0069] 在一些实施方式中,可以使用以标记的用户话语的语料库的形式的训练数据来训练神经网络(或其他分类器)(在这种情况下,训练是“受监督的”)。应用于话语的语料库的标记可以包括例如,指示包含旨在用于另一用户的话语的第一标记、指示加入与自动化助理120的人机对话的命令的第二标记,和/或指示背景噪音的第三标记(可以被忽略)。标记的训练示例可以被用作未训练的神经网络的输入。可以通过反向传播、随机梯度下降、目标函数优化等技术来确定和使用未训练(或未完全训练)的神经网络的输出与标记之间的差异(即误差)来调整神经网络的一个或多个隐藏层的各种权重来减少误差。

[0070] 如背景技术中所述,可以已经训练了诸如神经网络的机器学习分类器来识别(例如,分类)旨在调用自动化助理120的语音的音素或其他音频特性。在一些实施方案中,可以进一步训练同一分类器来识别(例如,分类)自动化助理120的明确调用,并且确定话语是否包含旨在用于第二用户的消息。在其他实施方式中,可以将单独的机器学习分类器用于这两个任务中的每一个,例如,一个接一个或并行地。

[0071] 除了确定捕获的(记录的)话语包含旨在用于另一用户的消息外,还可以例如基于说话者和预期接收者的相应位置来确定是否保证对讲式通信。在各种实施方式中,ICLS140可以例如使用与一个或多个客户端设备 106_{1-N} 相关联的存在传感器105来确定预期接收者相对于多个客户端设备 106_{1-N} 的位置。例如,ICLS140可以确定哪个客户端设备106最接近预期接收者,和/或预期接收者在哪一房间(在一些情况下,可以与部署在该房间中的客户端设备相关联)。在各种实施方式中,基于由ICLS140确定的预期接收者的位置,ICAS138可以从多个客户端设备 106_{1-N} 中选择能够提供预期接收者可以感知的音频或视觉输出的第二客户端设备106。例如,如果最后检测到预期的接收者走进特定区域,则可以选择最接近该区域的客户端设备106。

[0072] 在一些实施方式中,ICLS140可以被提供为例如基于云的自动化助理组件119的一部分和/或与其分开提供。在其他实施方式中,ICAS 138和ICLS140可以一起在单个模型或引擎中实现。在各种实施方式中,ICLS140可以被配置为基于例如由与分布在整个区域中的多个客户端设备 106_{1-N} 集成的存在传感器105提供的信号,跟踪诸如在家、工作场所、校园等之内的感兴趣区域内的人的位置。基于这些跟踪的位置,ICLS140和/或ICAS138可以被配置为便于使用如本文所述的多个客户端设备 106_{1-N} 的该区域中的人之间的对讲式通信。

[0073] 在一些实施方式中,ICLS140可以创建和/或维护位于特定区域中的人的列表或数据库,和/或相对于在该区域中部署的多个客户端设备 106_{1-N} 的他们的最后已知位置。在一些实施方式中,当人因为已经移动到不同的位置而被不同的客户端设备检测到时,可以例如实时地更新该列表/数据库。例如,如果例如在一些预定时间间隔(例如一小时)内,在整个区域中未检测到该人和/或如果最后检测到该人通过入口或出口区域(例如前门、后门等),则ICLS140可以从列表/数据库中删除该人。在其他实施方式中,ICLS140可以例如每几分钟、几小时等周期性地更新列表/数据库。

[0074] 在一些实施方式中,ICAS138和/或ICLS140(并且更一般地,自动化助理120)可以被配置为使用来自存在传感器105的信号来区分不同的人,而不是简单地检测普通人的存在。例如,假设客户端设备106包括作为存在传感器105的麦克风。自动化助理120可以被配置为使用各种说话者识别和/或语音识别技术来不仅确定有人在附近,而且确定谁在场。这些说话者识别和/或语音识别技术可以包括但不限于隐马尔可夫模型、高斯混合模型、频率

估计、已训练的分类器、深度学习、模式匹配算法、矩阵表示、矢量量化、决策树等。

[0075] 如果配备有麦克风的客户端设备106附近的人没有说话,则可以采用其他技术来识别该人。假设客户端设备106包括照相机和/或PIR传感器作为存在传感器105。在一些实施方式中,可以使用由这种存在传感器105捕获的标记训练数据来训练机器学习视觉识别分类器以可视地识别人。在一些实施方式中,用户可以通过在一个或多个配备有照相机/PIR传感器的客户端设备106处调用训练例程来使视觉识别分类器被训练。例如,用户可以站在存在传感器105的视场中,并且通过诸如“Hey Assistant, I am Jan and this is what I look like (嘿助理,我是Jan,这就是我的样子)”的短语来调用自动化助理120。在一些实施方式中,自动化助理120可以提供提示用户在存在传感器105的视场内的各个位置周围移动的听觉或视觉输出,同时存在传感器105捕获用户的一个或多个快照。然后,可以将这些快照加标记(例如,标记有“Jan”),并且用作用于视觉识别分类器的监督训练的标记训练示例。在其他实施方式中,可以自动地生成用于视觉识别的标记训练示例,例如无需用户知道。例如,当用户在存在传感器105的视场中时,例如可以由自动化助理120分析由用户携带的移动客户端设备106发出的信号(例如,无线电波、超声波),以对存在传感器105捕获的快照确定用户的身份(因此确定标记)。

[0076] 且在又一实施方式中,可以采用除了音频和/或视觉提示之外的其他类型的提示来区分用户。例如,可以由例如自动化助理120分析由用户携带的客户端设备106发出的无线电、超声波和/或其他类型的无线信号(例如,红外、调制光等),以辨别附近用户的身份。在一些实施方式中,用户的移动客户端设备106可以包括可以用来识别用户的网络标识符,诸如“Jan’s Smartphone (Jan的智能手机)”。

[0077] 现在参考图1B,示意性地描绘了示例性数据流以说明一种可能的方式,在该方式中,可以应用已训练的机器学习分类器来分析用户话语并且确定除了其他以外是否采用对讲式通信。在图1B中,可以训练音素分类器142(可以是自动化助理120的组件),使得可以将一种或多种话语和一个或多个人的位置用作音素分类器142的输入。然后,音素分类器142可以将话语的分类生成输出。在图1B中,这些分类包括“调用助理(invoke assistant)”、“传达消息(convey message)”和“背景噪音(background noise)”,但是附加的和/或替代的标记是可能的。

[0078] 传统的音素分类器已经存在,它们可以检测诸如“Hey, Assistant”、“OK Assistant (好的,助理)”等的显式调用短语。在一些实施方式中,音素分类器142可以包括相同的功能,使得当输入话语包括这样的调用短语时,音素分类器142的输出是“调用助理”。一旦调用了自动化助理120,如本领域所公知的,用户就可以加入与自动化助理120的人机对话。

[0079] 然而,在一些实施方式中,可以进一步训练音素分类器142以识别发信号告知向另一用户传达消息的用户意图的其他音素。例如,用户可能经常使用诸如“Hey, <name> (嘿,<名称>)”的短语来引起人的注意。更一般地,音素分类器142可以操作以匹配定制短语、单词等。附加地或可替代地,为了引起另一个人的注意,通常是先说出另一个人的姓名,有时以稍高音量和/或带有特定的语调,或者使用其他类型的语调。在各种实施方式中,可以训练音素分类器142来识别这样的音素并生成诸如“传达消息”的输出,以信号告知可能潜在地保证对讲式通信的情形。在各种实施方式中,可以可选地单独训练单独的语调模型,以识别

寻求与另一人进行通信的话语(例如,将这类话语与随意话语区分开)并生成指示这类话语的存在的输出(例如,这类话语存在的可能性)。在确定是否保证对讲式通信时,可以综合考虑对于给定的用户话语,来自音素分类器和语调模型的输出。

[0080] 在一些实施方式中,可以例如由ICLS140提供一个或多个人的位置,作为音素分类器142的输入。除了或代替话语,可以使用这些人的位置来确定是否保证对讲式通信。例如,如果接收者位置离说话者的位置足够近(例如,在能听到说话者的位置的范围内),这可能影响音素分类器142以产生诸如“背景噪音”的输出,即使话语包含旨在用于另一人的消息。另一方面,假设预期接收者的位置不在能听到说话者位置的范围之内。这可能影响音素分类器142以产生诸如“传达消息”的输出,这可能增加采用对讲式通信的可能性。附加地或替代地,可以实施两步法,其中,首先确定说话者的话语是否包含旨在用于另一用户的消息,然后确定另一用户是否在能听到说话者的范围内。如果对这两个问题的回答都为“是”,则可以实施对讲式通信,以将消息传达给预期接收者。

[0081] 现在参考图2,示出了包括多个房间250-262的家庭平面图。在至少一些房间中部署了多个客户端设备206₁₋₄。每个客户端设备206可以实施配置有本公开的所选方面的自动化助理客户端118的实例,并且可以包括能够捕获附近人说出的话语的一个或多个输入设备,诸如麦克风。例如,在房间250(在该示例中是厨房)中部署采取独立交互式扬声器形式的第一客户端设备206₁。在房间252(在该示例中,为小房间)中部署采取所谓的“智能”电视形式的第二客户端设备206₂(例如,具有实现自动化助理客户端118的实例的一个或多个处理器的联网电视)。在房间254(在该示例中为卧室)中部署采用交互式独立扬声器形式的第三客户端设备206₃。在房间256(在该示例中为起居室)中部署采用另一个交互式独立扬声器形式的第四客户端设备206₄。

[0082] 尽管在图2中未示出,但是多个客户端设备106₁₋₄可以经由一个或多个有线或无线LAN(例如,图1A中的110₂)彼此和/或与其他资源(例如,因特网)通信地耦合。另外,也可以存在,例如由家中的一个或多个个人携带其他客户端设备,特别是诸如智能电话、平板电脑、笔记本电脑、可穿戴设备等移动设备,并且可以连接也可以不连接到同一局域网。应当理解到,图2和其他附图中所示的客户端设备的配置仅是一个示例;可以在任意多个房间和/或除家以外的区域部署更多或更少的客户端设备106。

[0083] 在图2的示例中,第一用户Jack说出问题“Hey Hon, do you know where the strainer is? (嘿,亲爱的,你知道过滤器在哪里吗?)”时,他正在厨房250中。也许Jack不知道的是,他的妻子Jan不在厨房250中,而是在起居室256中,因此很可能没有听到Jack的问题。如上所述,被配置有本公开的所选方面的第一客户端设备206₁可以检测Jack的话语。可以使用上述技术来分析话语的记录,以确定Jack的话语包含旨在用于Jan的消息。第一客户端设备206₁还可以例如基于在所有多个客户端设备206₁₋₄之间共享的信息来确定Jan在起居室256中(或至少离第四客户端设备206₄最近)。例如,客户端设备206₄可以已经例如使用一个或多个集成存在传感器(例如,图1A中的105)检测到Jan在起居室256中。

[0084] 基于Jan的检测到的位置和/或Jack的话语的属性(在一些实施方式中,可以使用如上所述的已训练的机器学习模型对其进行分类),第一客户端设备206₁可以确定Jack旨在其消息用于Jan,并且Jan不在能听到Jack的范围内。因此,第一客户端设备206₁可以将Jack的话语记录(或者在一些情况下,Jack的话语的转录文本)(在一个或多个上述LAN上)

推送到最接近Jan的客户端设备,在该示例中,为第四客户端设备206₄。在接收到该数据时,第四客户端设备206₄可以通过例如至少部分地在第四客户端设备206₄上执行的自动化助理120,如图2所示,将Jack的消息可听地输出到Jan,从而实现Jack和Jan之间的对讲式通信。

[0085] 在图2的示例中(以及在本文其他地方所述的类似示例中),Jack的问题使用如上所述,为独立交互式扬声器的第四客户端设备206₄可听地输出到Jan。但是,这并不意味着是限制性的。在各种实施方式中,可以使用其他输出方式将Jack的消息传达给Jan。例如,在Jan携带的移动客户端设备(未示出)连接到Wi-Fi LAN的一些实施方式中,该移动设备可以例如使用诸如在Jan的移动客户端设备上执行的诸如消息交换客户端107的应用,输出Jack的消息,作为可听录音或作为可视地传达给Jan的文本消息。

[0086] 在各种实施方式中,在客户端设备106之间交换以促进对讲通信的话语的记录和/或STT转录可以用于多种附加目的。在一些实施例中,它们可以被用来对用户和自动化助理120之间的下游人机对话提供场境。例如,在一些情况下,记录的话语和/或其STT转录可以用于消除提供给自动化助理120的实例的请求的歧义,无论该请求是来自最初提供话语的用户、该话语的预期接收者,甚至在涉及多个客户端设备106的对讲式通信之后与自动化助理120交互的另一用户。

[0087] 图3描绘了与图2中描绘的相同的家和客户端设备206₁₋₄的分布。在图3中,Jan(仍在起居室256中)说出话语“Hey Jack,you should leave soon to pick up Bob from the airport(嘿,Jack,你应该快点去机场接Bob)”。例如,ICLS140可以确定Jack在不能听到Jan的范围内的另一房间里。例如,ICLS140可以例如基于由“智能”恒温器264的板载摄像机和/或PIR传感器提供的信号,确定Jack位于小房间252中。基于该确定和/或Jan的话语(例如,使用上述机器学习模型中的一个)已经被分类为旨在用于Jack的消息的确定,可以识别Jack检测到的位置附近的客户端设备,诸如客户端设备206₂以输出Jan的话语。在一些实施方式中,可以将Jan的记录的话语从诸如客户端设备206₄的记录了它的、Jan附近的另一计算设备推送到Jack附近识别的客户端设备206₂,并且可听地输出(或因为客户端设备206₂是具有显示能力的智能电视,可视地输出)。

[0088] 图4展示了图3所示的示例性后续方案。在经由客户端设备206₂接收到Jan传达的消息后,Jack说“OK Assistant—when is the next tram leaving?(好的,助理,下一辆电车什么时候离开?)”没有其他信息,该请求或搜索查询可能太含糊,无法回答,因此可能需要自动化助理120向Jack请求消除歧义的信息。然而,使用本文所述的技术,自动化助理120可以基于Jan的原始话语来消除Jack的请求的歧义,以确定到机场的电车是Jack感兴趣的电车。附加地或可替代地,自动化助理120可以简单地获取所有附近电车的正常结果,然后基于Jack的话语对那些结果进行排名,例如,使得到机场的电车排名最高。无论哪种情况,在图4中,自动化助理120都会在客户端设备206₂上提供音频输出,即“Next tram to the airport leaves in 10minutes(到机场的下一趟电车在10分钟内离开)”。

[0089] 图5是示出根据本文公开的实施方式的示例性方法500的流程图。为了方便起见,参照执行该操作的系统来描述流程图的操作。该系统可以包括各种计算机系统的各种组件,诸如实现自动化助理120的计算系统的一个或多个组件。此外,尽管以特定顺序示出了方法500的操作,但这并不意味着是限制性的。可以重新排序、省略或添加一个或多个操作。

[0090] 在框502处,系统可以在多个计算设备中的第一计算设备的输入设备处从第一用

户接收自由形式的自然语言输入。在许多实施方式中,该自由形式的自然语言输入可以语音输入,即来自第一用户的话语的形式出现,尽管这不是必需的。应当理解到,该语音输入不一定要由第一用户在自动化助理120处导向,而是可以包括由第一用户提供的、由配置有本公开的所选方面的客户端设备捕获和/或记录的任何语音。

[0091] 在框504处,系统可以分析语音输入。可以分析语音输入的各个方面(例如,音素),包括但不限于语调、音量、识别的短语等。

[0092] 在一些实施方式中,系统可以分析除语音输入之外的其他信号。这些其他信号可以包括例如房屋的环境中的多个人。例如,如果只有一个人在场,则无法使用对讲功能。如果仅存在两个人,则可以将另一个人的位置自动地确定为应该提供对讲输出的位置。如果存在两个以上的人,则系统可以尝试各种技术(例如,语音识别、面部识别、无线信号识别等)以尝试将人们彼此区分开。

[0093] 在框506处,系统可以基于该分析来确定第一用户旨在向第二用户传达消息(例如,语音输入包含旨在用于第二用户的消息)。如前所述,自动化助理120可以采用各种技术,诸如以记录的话语形式,在标记的训练示例上训练的分类器,以分析语音输入和/或确定第一用户的语音输入是为调用自动化助理120以加入进一步的人机对话的命令(例如,“Hey,Assistant”)、旨在将消息传达给第二个用户(或多个其他用户)的话语还是其他背景噪音。在一些实施方式中,除了已训练的机器学习模型之外或代替已训练的机器学习模型,可以实施基于规则的方法。例如,可以例如使用被设计成匹配旨在用户之间传达消息的话语、诸如语音可扩展标记语言或“VXML”的技术来定义一种或多种简单的IVR语法。

[0094] 在框508(其可以在框502-506之后或持续的基础上),系统可以确定第二用户相对于多个计算设备的位置。在一些实施例中,ICLS140可以维护诸如家庭或工作场所等区域的人及其它们的最后已知(即,最后检测到)的位置的列表或数据库。在一些这样的实施方式中,系统可以在该列表或数据库简单地查询第二用户的位置。在其他实施方式中,系统可以例如根据需要(例如,当确定第一用户的话语包含旨在要传达给第二用户的消息时)主动轮询环境中的多个客户端设备以找出第二用户。这可能使得客户端设备激活存在传感器(图1A中的105)以便他们可以检测附近是否有人(例如第二用户)。

[0095] 在框510处,系统可以基于第二用户的位置,从多个计算设备中选择能够提供第二用户可感知的音频或视觉输出的第二计算设备。在一些实施方式中,第二计算设备可以是部署在环境的特定区域中的固定客户端设备(例如,独立的交互式扬声器、智能电视、台式计算机等)。在其他实施方式中,第二计算设备可以是第二用户携带的移动客户端设备。在一些这样的实施方式中,由于作为同一协调生态系统的一部分和/或加入同一无线LAN(或简单地处于预定距离之内),移动客户端设备可以成为系统所考虑的多个计算设备的一部分。

[0096] 在框512处,系统可以使得在框510处识别的第二计算设备提供将消息传达给第二用户的音频或视觉输出。例如,在ICAS138和/或ICLS140为基于云的一些实施方式中,一个或另一个可以使得第一用户话语的记录被转发(例如,流式传输)到在框510处选择的第二计算设备。可以配置有本公开的所选方面的第二计算设备可以通过输出转发的记录来进行响应。

[0097] 尽管本文描述的示例包括第一用户尝试将消息传达给单个其他用户,但这并不意

味着限制。在各种实施方式中,用户的话语可能旨在用于多个其他用户,诸如说话者家庭的多个成员、该地区或环境中的所有人等。在一些这样的实施方式中,可以(进一步)训练一个或多个上述机器学习分类器以确定话语是否包含旨在用于单个接收者或多个接收者的消息。如果回答为是,则系统可以以各种方式将消息传达到多个位置处的多个预期接收者。在一些简单的实施方式中,系统可以简单地使得消息被推送到该区域中的所有客户端设备(例如,协调生态系统的所有客户端设备和/或连接到Wi-Fi LAN的所有客户端设备),从而有效地广播该信息。在其他实施方式中,系统(例如,ICLS 140)可以在个人的基础上,确定所有预期接收者的位置,并且仅在每个预期接收者附近的那些客户端设备上输出消息。

[0098] 在一些实施方式中,自动化助理120可以等待,直到消息的预期接收者能够感知消息(例如,在能听到的范围内),直到使用本文所述的技术传达消息为止。例如,假设第一用户将消息传达给预期的接收者,但是预期的接收者已经瞬间离开。在一些实施方式中,消息可以被暂时延迟,直到一个或多个计算设备检测到预期的接收者。检测到接收者的第一计算设备在他们返回时可以输出原始消息。在一些实施方式中,各种信号,诸如从他们携带的移动设备获得的预期接收者的位置坐标(例如,全球定位系统或“GPS”),可以被用来确定使用(至少不是与LAN上的任何设备)对讲式通信无法达到预期接收者。在一些实施方式中,该消息(例如,说话者的话语的记录)可以被转发到接收者的移动设备。在其他实施方式中,自动化助理120可以确定预期的接收者是不可达到的,并且可以例如在最接近说话者的客户端设备(例如,捕获说话者话语的设备)处提供输出,以通知说话者此时接收者无法达到。在一些这样的实施方式中,自动化助理120可以提示用户允许将消息转发到接收者的移动设备,例如,通过输出如“I can't reach Jan directly right now.Would you like me to send a message to their phone?(我现在不能直接联系Jan。您想向他们的手机发送消息吗?)”之类的内容。

[0099] 如果在框512输出第一用户的消息之后的一些时间,第二用户发出自由形式的自然语言输入,则可选框514-518可以或可以不发生。在框514,系统可以识别在框512处由第二计算设备提供音频或视觉输出之后,由第二用户发出的自由形式的自然语言输入,诸如语音输入。第二用户的语音输入可以包括例如命令和/或搜索查询。在一些实施例中,命令和/或搜索查询本身可能太含糊,以至于不能正确地如图4中Jack的话语的情况解释。

[0100] 在框516,系统可以至少部分地基于在框502从第一用户接收的自由形式自然语言输入来分析在框514处识别的第二用户的自由形式自然语言输入。换句话说,第一用户的话语可以被转录并且用于向第二用户的后续请求提供场境。在框518处,系统可以基于由第一用户的原始自由形式自然语言输入提供的场境来制定对第二用户的自然语言输入的响应。例如,如果第二用户的自由形式自然语言输入包括搜索查询(诸如图4中的Jack的查询),则系统可以至少部分地基于来自第一用户的语音输入来获得响应于搜索查询的搜索结果。例如,可以基于第一用户的原始话语来消除第二用户的搜索查询的歧义,和/或可以基于第一用户的原始话语来对一个或多个响应搜索结果进行排名。在框518处,系统然后可以使多个计算设备中的一个或多个提供指示至少一些搜索结果的输出,如图4所示。

[0101] 在各种实施方式中,用户可以在其家、工作场所或另一环境中预先配置(例如,委托)客户端计算设备,以可用于加入本文所述的对讲式通信。例如,在一些实施方式中,用户可以例如使用图形用户界面和/或通过加入与自动化助理120的人机对话会话,来为每个固

定客户端计算设备分配“位置”，诸如“kitchen(厨房)”、“dining room(饭厅)”等。因此，在一些这样的实施方式中，用户可以明确地调用自动化助理120以促进与特定位置的对讲式通信。例如，用户可以提供以下语音输入，以将消息传达给另一位用户：“Hey Assistant, tell Oliver in the kitchen that we need more butter(嗨，助理，告诉在厨房的Oliver，我们需要更多黄油)”。

[0102] 更一般地，在一些实施方式中，用户在调用对讲式通信时可以显式指定消息的接收者。如果用户也未指定接收者的位置，则例如与ICLS 140相关联的本文所述的技术可以被自动地用于确定接收者的位置，并且选择将使用哪一计算设备来将消息输出到接收者。然而，如上所述，用户不需要显示地调用对讲式通信。而是，如上所述，可以考虑各种信号和/或数据点(例如，机器学习分类器的输出、预期接收者的位置等)而无需来自用户的显示指示，确定应当使用对讲式通信自动地传达用户的消息。

[0103] 图6是可以可选地用于执行本文所述的技术的一个或多个方面的示例性计算设备610的框图。在一些实施方式中，客户端计算设备、用户控制的资源引擎130和/或其它组件中的一个或多个可以包括示例性计算设备610的一个或多个组件。

[0104] 计算设备610通常包括经由总线子系统612与多个外围设备通信的至少一个处理器614。这些外围设备可以包括存储子系统624(包括例如存储器子系统625和文件存储子系统626)、用户接口输出设备620、用户接口输入设备622和网络接口子系统616。输入设备和输出设备允许用户与计算设备610交互。网络接口子系统616提供到外部网络的接口并且耦合至其它计算设备中的相应接口设备。

[0105] 用户接口输入设备622可以包括键盘、诸如，鼠标、轨迹球、触摸板或者图形输入板的指向设备、扫描仪、并入显示器中的触摸屏、诸如，语音识别系统、麦克风和/或其它类型的输入设备的音频输入设备。一般而言，使用术语“输入设备”旨在包括将信息输入到计算设备610中或者到通信网络上的所有可能类型的设备和方法。

[0106] 用户接口输出设备620可以包括显示子系统、打印机、传真机、或者非可视显示器，诸如，音频输出设备。显示子系统可以包括阴极射线管(CRT)、诸如，液晶显示器(LCD)的平板设备、投影设备、或者用于创建可见的图像的一些其它机制。显示子系统还可以提供非可视显示器，诸如，经由音频输出设备。一般而言，使用术语“输出设备”旨在包括将信息从计算设备610输出至用户或者至另一机器或者计算设备的所有可能类型的设备和方法。

[0107] 存储子系统624存储提供本文所述的模块中的一些或者所有的功能的编程和数据结构。例如，存储子系统624可以包括执行图5的方法的所选方面以及实现图1A中图示的各个组件的逻辑。

[0108] 这些软件模块通常由处理器614单独或者与其它处理器组合执行。用在存储子系统624中的存储器625可以包括多个存储器，该多个存储器包括用于在程序执行期间存储指令和数据的主随机存取存储器(RAM)630以及存储固定指令的只读存储器(ROM)632。文件存储子系统626可以为程序和数据文件提供永久存储，并且可以包括硬盘驱动器、软盘驱动器连同相关联的可移动介质、CD-ROM驱动器、光盘驱动器、或者可移动介质盒。实施某些实施方式的功能的模块可以由文件存储子系统626存储在存储子系统624中或者可以存储在可由处理器614访问的其它机器中。

[0109] 总线子系统612提供用于使计算设备610的各种组件和子系统按照预期彼此通信

的机制。虽然总线子系统612被示意性地示出为单个总线,但是总线子系统的备选实施方式可以使用多个总线。

[0110] 计算设备610可以为各种类型,包括工作站、服务器、计算集群、刀片式服务器、服务器群、或者任何其它数据处理系统或者计算设备。由于计算机和网络的性质多变,所以图6中图示的计算设备610的描述仅仅旨在作为为了说明一些实施方式的目的的特定示例。计算设备610的许多其它配置可能具有比图6中图示的计算设备更多或者更少的组件。

[0111] 在本文所讨论的某些实施方式可以收集或者使用关于用户的个人信息(例如,从其它电子通信提取到的用户数据、关于用户的社交网络的信息、用户的位置、用户的时间、用户的生物特征信息,以及用户的活动和人口统计信息、用户之间的关系等)的情况下,为用户提供控制是否收集信息、是否存储个人信息、是否使用个人信息、以及如何收集、存储和使用关于用户的信息的一个或多个机会。即,本文所讨论的系统和方法仅在从相关用户接收到可以这样做的明确授权时,才收集、存储和/或使用用户个人信息。

[0112] 例如,为用户对是否收集关于该特定用户或者与程序或者功能特征相关的其它用户的用户信息提供控制程序或者功能部件。为将收集个人信息的每个用户呈现一个或多个选项以允许控制与该用户相关的信息收集、提供关于是否收集信息和关于收集信息的哪些部分的许可或者授权。例如,可以通过通信网络,为用户提供一个或多个这样的控制选项。另外,某些数据在被存储或者使用之前,可以以一种或者多种方式来处理,从而移除个人身份信息。作为一个示例,可以处理用户的身份,从而无法确定个人身份信息。作为另一示例,可以将用户的地理位置泛化到更大的区域,从而无法确定用户的具体位置。

[0113] 虽然在本文中已经描述和示出了多个实施方式,但是可以利用用于执行功能和/或获得结果和/或本文所述的一个或多个优点的各种其它装置和/或结构,并且认为这类变形和/或修改中的每一个在本文所述的实施方式的范围内。更一般地,本文所述的所有参数、尺寸、材料和配置意指示例性的,并且实际参数、尺寸、材料和/或配置将取决于使用了本教导的具体应用。本领域的技术人员仅仅使用常规实验即将认识或者能够确定本文所述的具体实施方式的许多等同物。因此,将理解到,前述实施方式仅仅以举例的方式呈现,并且在所附权利要求书和其等同物的范围内,可以不同于所具体描述和所要求保护的方式来实践实施方式。本公开的实施方式是针对本文所述的每一单个特征、系统、制品、材料、套件和/或方法。另外,如果这些特征、系统、制品、材料、套件和/或方法并非相互矛盾,则两个或以上这些特征、系统、制品、材料、套件和/或方法的任意组合被包括在本公开的范围。

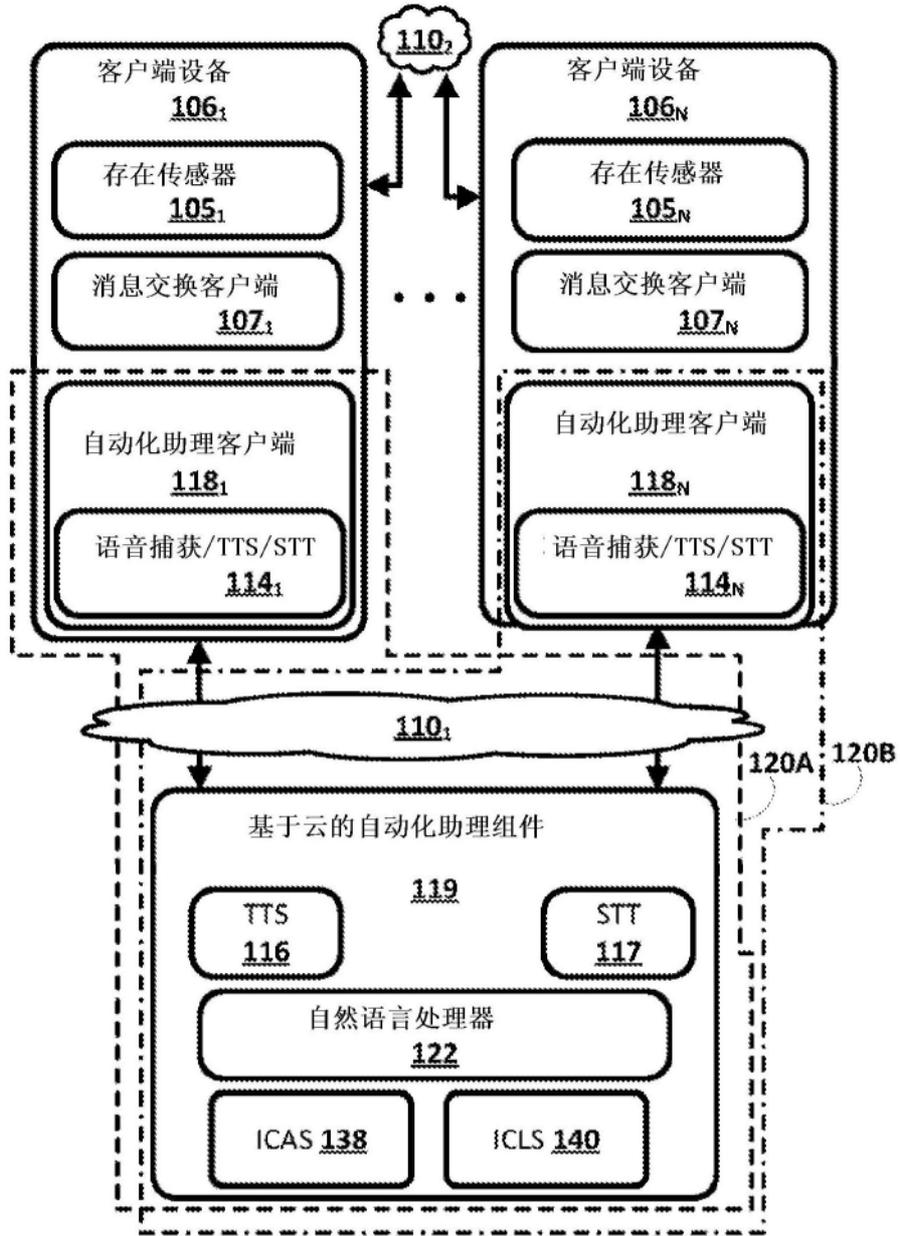


图1A

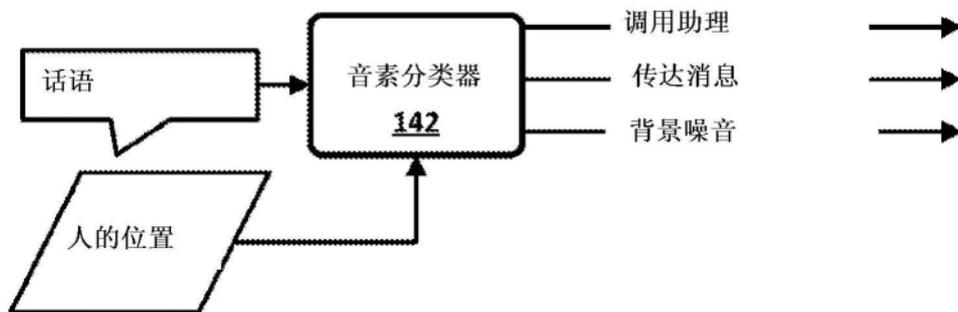


图1B

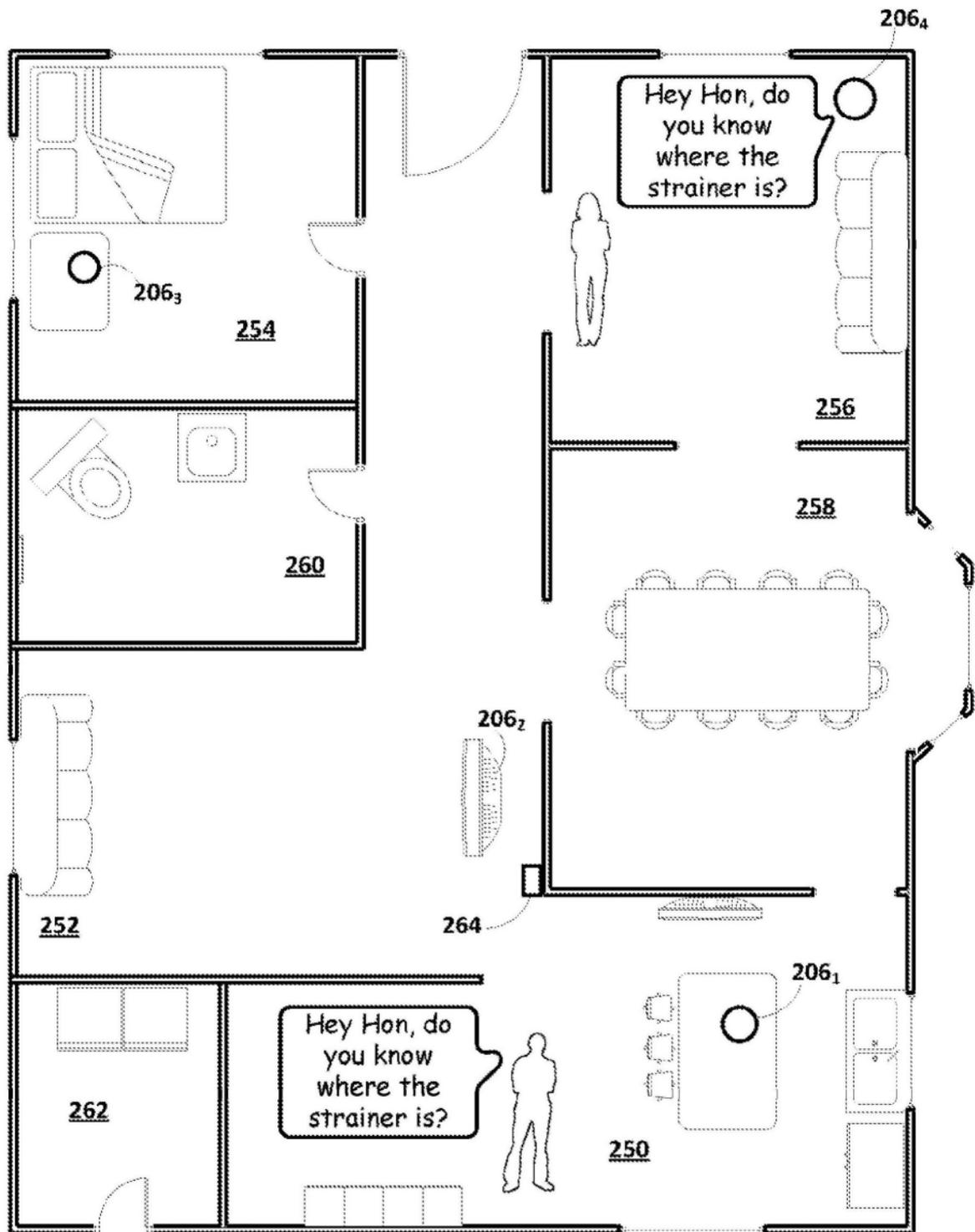


图2

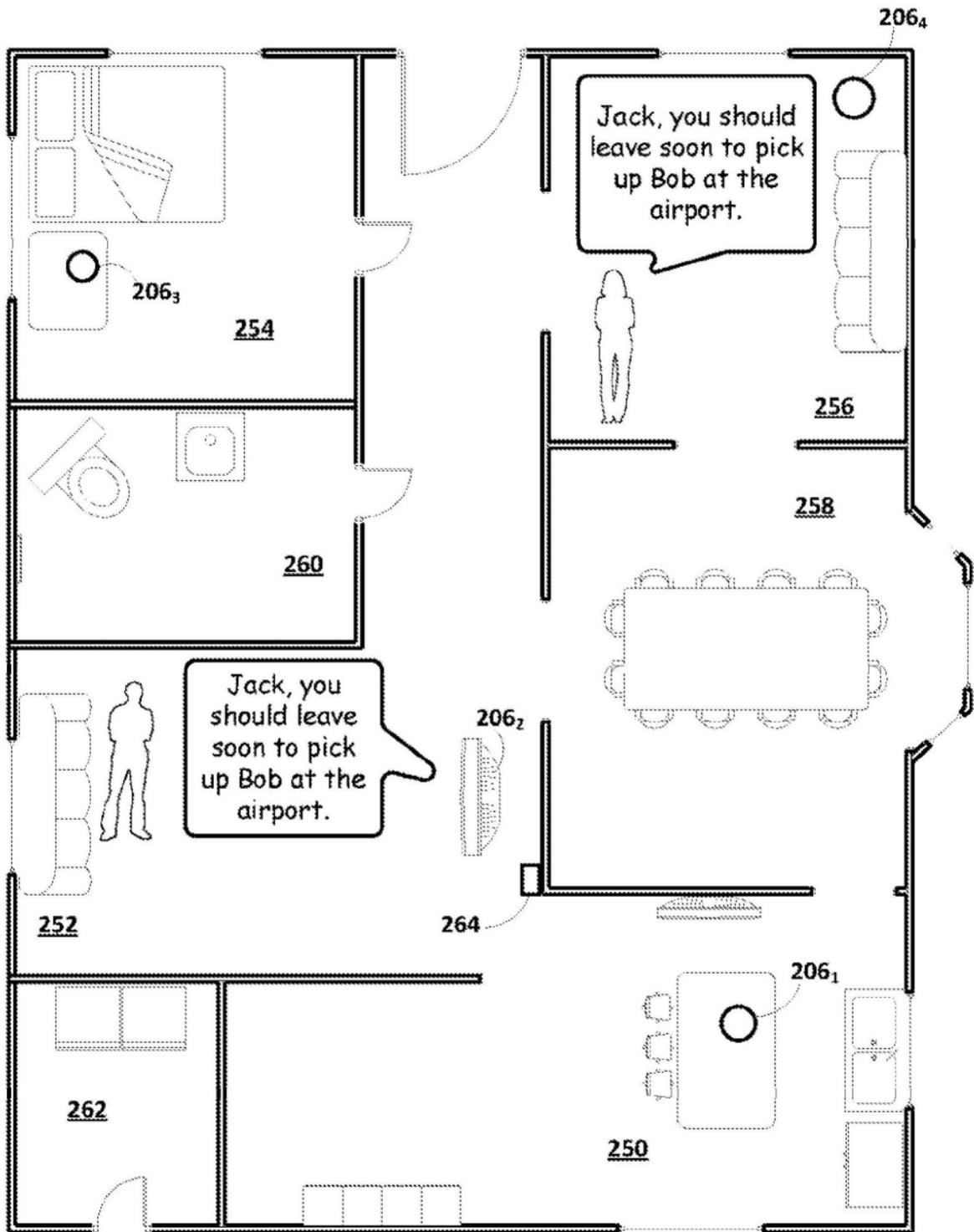


图3

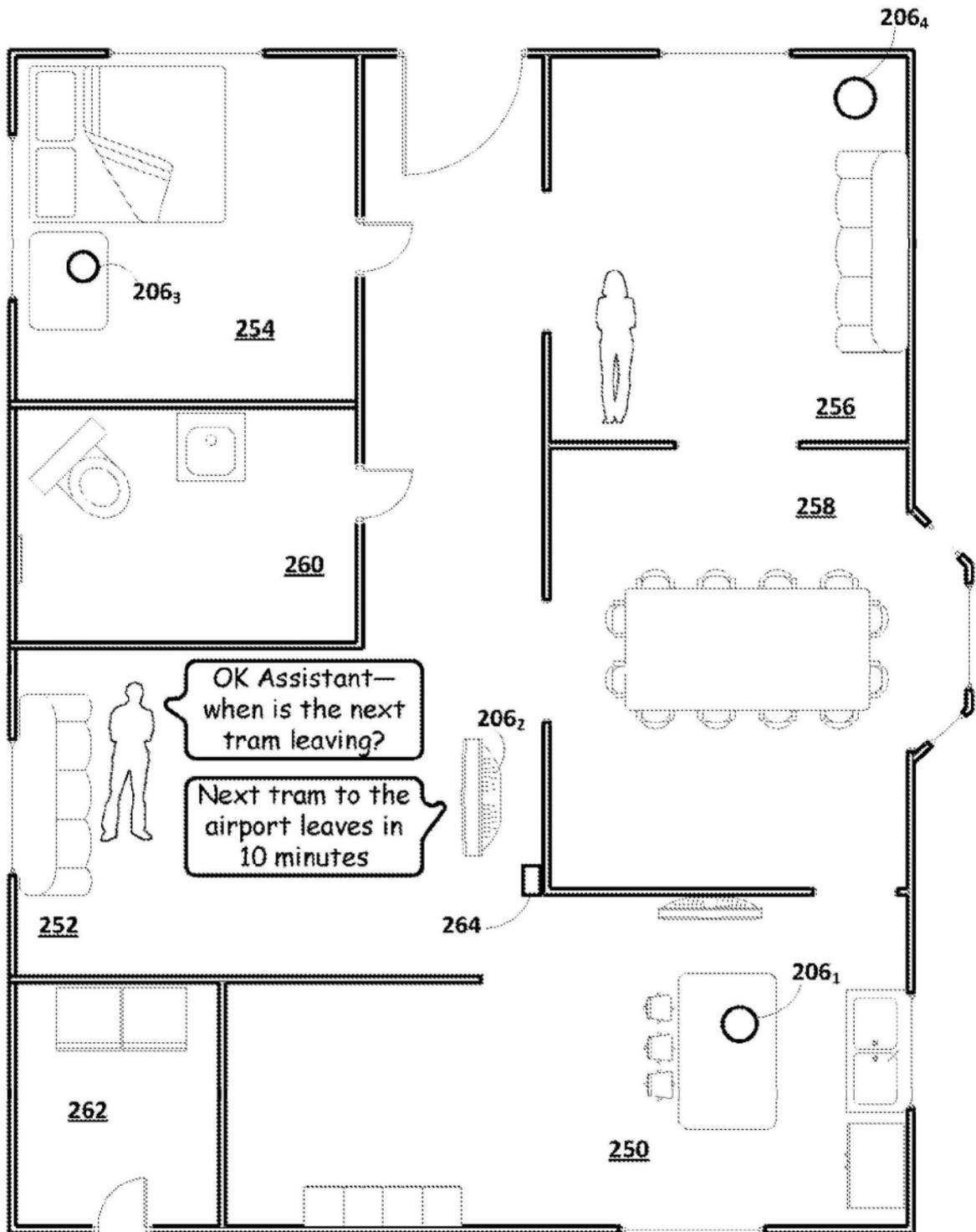


图4

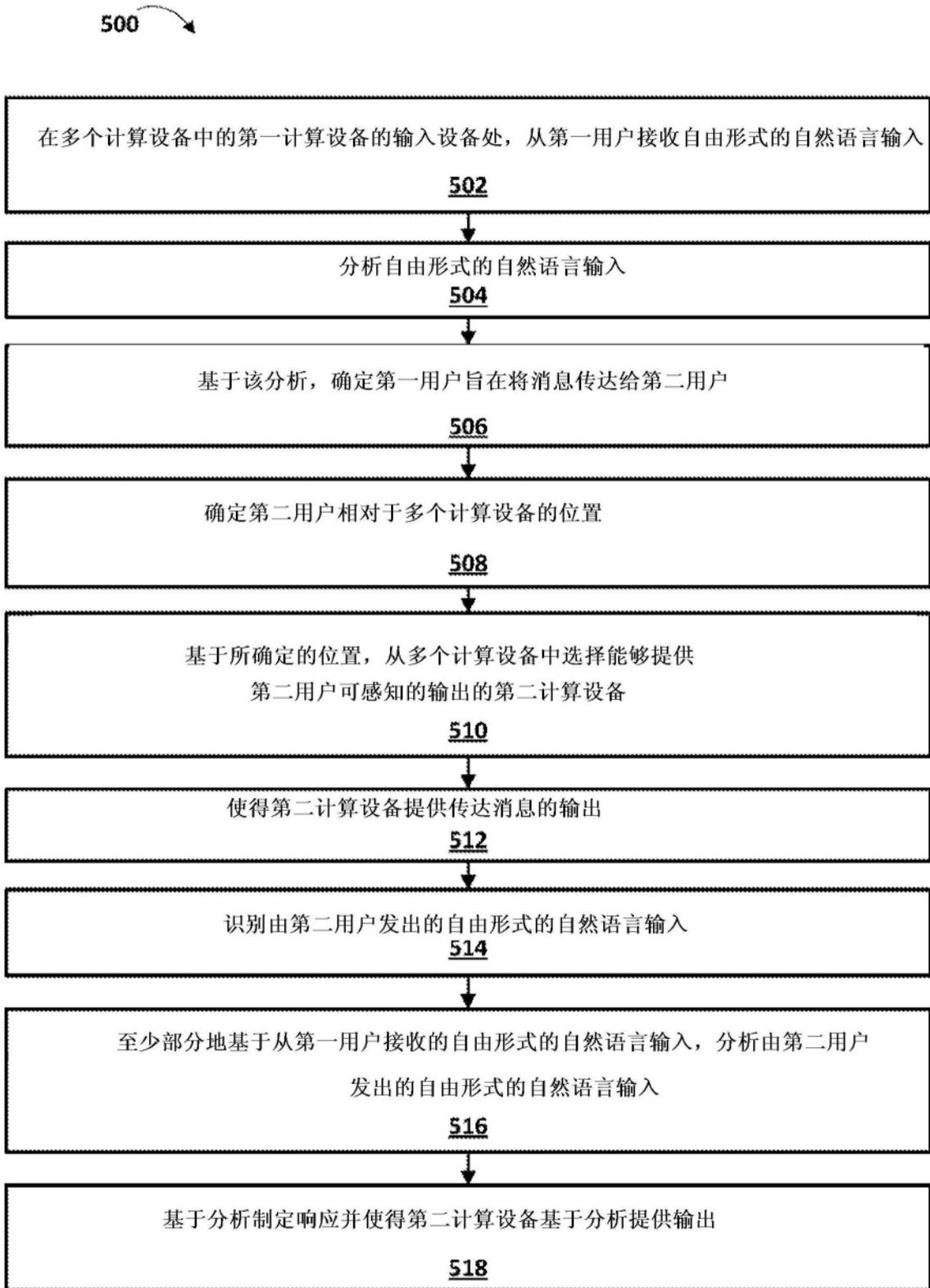


图5

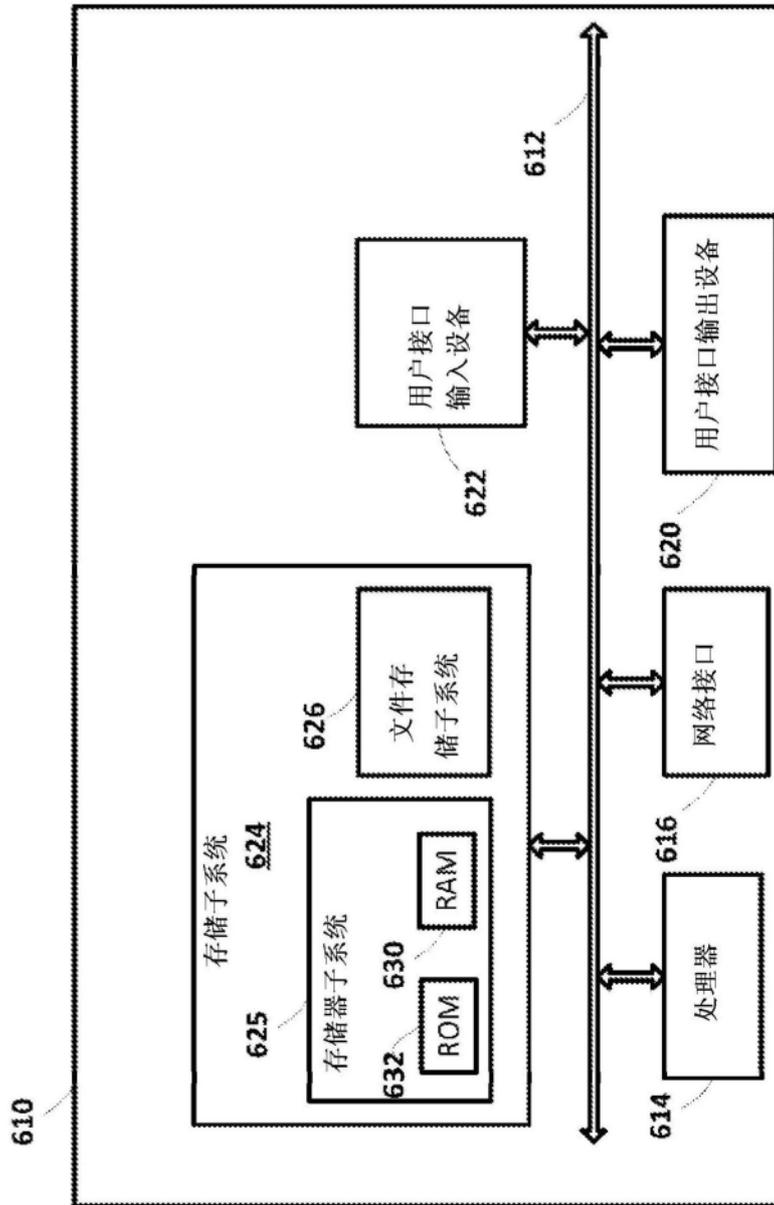


图6