



(21)申请号 201510340528.8

(22)申请日 2015.06.18

(65)同一申请的已公布的文献号
申请公布号 CN 105320268 A

(43)申请公布日 2016.02.10

(30)优先权数据
14305964.0 2014.06.20 EP

(73)专利权人 交互数字CE专利控股公司
地址 法国巴黎

(72)发明人 V.阿利奥梅 C.贝拉德 P.朱厄特

(74)专利代理机构 北京市柳沈律师事务所
11105
代理人 赵碧洋

(51)Int.Cl.

G06F 3/01(2006.01)

(56)对比文件

US 2012089392 A1,2012.04.12,
CN 102473040 A,2012.05.23,
CN 101910960 A,2010.12.08,
US 2014007115 A1,2014.01.02,
CN 102541438 A,2012.07.04,

审查员 张雨微

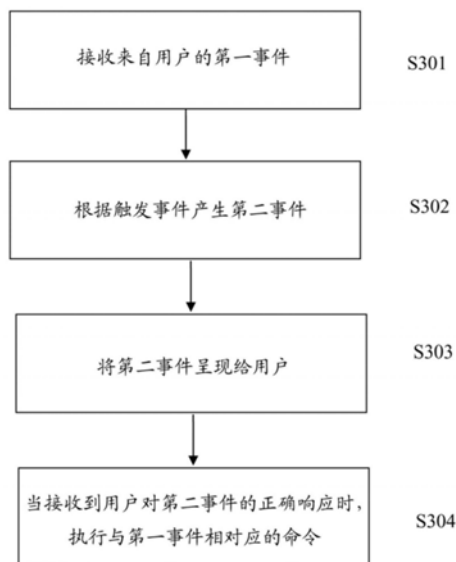
权利要求书1页 说明书6页 附图3页

(54)发明名称

装置以及用户控制装置的方法

(57)摘要

提出一种装置以及用户控制装置的方法。所提出的装置包括:检测单元,用于检测与包括所述装置的多个设备有关的由用户执行的第一事件;控制单元,用于根据由检测单元所检测到的第一事件来产生第二事件;输出单元,用于将第二事件呈现给用户,其中控制单元根据由检测单元所检测到的用户对第二事件的响应来产生用于在装置上执行第一事件的命令。所提出的装置和方法适合于在执行与来自用户的第一事件相对应的命令之前,提供第二事件,以便用户确认第一事件的意图。



1. 一种用于与用户交互的装置(1),包括
检测单元(201),用于检测与包括所述装置的多个设备有关的、由用户所执行的第一事件,其中所述多个设备中的多于一个设备可以同时响应第一事件;
控制单元(202),用于根据由所述检测单元(201)所检测到的第一事件来产生第二事件;以及
输出单元(203),将第二事件呈现给用户,其中
控制单元(202)根据由所述检测单元所检测到的用户对第二事件的响应来产生用于在所述装置上执行第一事件的命令;并且
所述第二事件包含文本消息,所述文本消息呈现给用户以确认在所述多个设备中的所述装置而非其他设备上执行第一事件的意图。
2. 根据权利要求1所述的装置(1),其中检测单元(201)检测手势和/或声音形式的第一事件。
3. 根据权利要求1所述的装置(1),其中第一事件是用于在装置与用户之间发起交互会话的触发事件。
4. 根据权利要求3所述的装置(1),其中第二事件是用户用以验证发起交互会话的意图的验证事件。
5. 根据权利要求4所述的装置(1),其中验证事件包括用户用以口头重复的文本消息。
6. 根据权利要求4所述的装置(1),其中验证事件包括要求用户执行活动的文本、图像、音频或视频形式的询问。
7. 根据权利要求1所述的装置(1),进一步包括
控制单元(202)将命令传送给执行第一事件所需要的组件和相关设备。
8. 根据权利要求1所述的装置(1),其中输出单元(203)以由第二事件的形式所确定的方式将第二事件呈现给用户。
9. 一种用户控制装置的方法,包括,在装置的级别:
接收(S301)与包括所述装置的多个设备有关的、由用户所执行的第一事件,其中所述多个设备中的多于一个设备可以同时响应第一事件;
根据第一事件来产生(S302)第二事件;
将第二事件呈现(S303)给用户;以及
根据用户对第二事件的响应,在所述装置上执行(S304)与第一事件相对应的命令,其中,
所述第二事件包含文本消息,所述文本消息呈现给用户以确认在所述多个设备中的所述装置而非其他设备上执行第一事件的意图。
10. 根据权利要求9所述的方法,其中第一事件是以手势和/或声音的形式。
11. 根据权利要求9所述的方法,其中第二事件包括用户用以口头重复的文本消息。
12. 根据权利要求9所述的方法,其中第二事件包括要求用户执行活动的文本、图像、音频或视频形式的询问。
13. 一种非临时性计算机可读介质,其上存储有计算机程序,所述计算机程序包括处理器可执行的用于实现根据权利要求9至12中的至少一项所述的方法的步骤的程序代码指令。

装置以及用户控制装置的方法

技术领域

[0001] 本公开一般涉及人机接口 (HMI)。具体地,本公开涉及一种装置以及用户控制装置的方法。

背景技术

[0002] 这部分旨在向读者介绍技术的各个方面,其可能与下面所描述和/或要求的本公开的各个方面相关。相信该论述对于向读者提供背景信息以便于更好地理解本公开的各个方面是有帮助的。因此,应当理解这些陈述要从这个角度来阅读,而不是作为对现有技术的承认。

[0003] 在人机接口 (HMI) 的领域中,基于手势识别和/或声音识别处理的系统最近分别在利用家庭控制台的视频游戏应用和智能电话上的导航服务中展示了一些好的用例。例如,新的微软控制台 XBOX One 将从不完全切断,等待通过简单的声音接通命令 (通过“Xbox 开启”句子) 来激活。

[0004] 随着所提及的趋势,人们能够容易地想象到,在不久的将来,在同一位置 (家庭,公共交通等) 处的许多设备可以通过音频或成像传感器接受来自语音或手势的命令。例如,在家庭,具有相同类型并且具有比邻地进行操作的这种能力的产品可能变成通常的情况。再次以 XBOX One 作为示例,在接近的区域中具有两个类似的控制台是可能的选择。

[0005] 在房间中的用户与其设备之一 (例如在同一房间中) 之间的通常的音频/视频交互序列的情况下,可能需要或考虑下面的功能或场景:

[0006] (1) 用户想要通过声音或手势与其接近的设备进行交互;

[0007] (2) 用户对所期望的设备执行命令 (语音或手势),其被假定发起交互 (例如,设备的设置菜单调用、媒体选择等);以及

[0008] (3) 来自用户的触发命令可能在没有消除操作的歧义的情况下发起与不止一个设备的交互,导致冲突和不期望的情况 (例如,两个接近的 XBOX One 可能通过捕捉相同的语音命令而接通)。

[0009] 交互的触发事件 (即检测到特定手势或特定声音命令) 可能是非有意地并且超出设备使用环境地执行。但是,如果不需要,则不应当导致这种交互会话的真正初始化。这里期待一些双重检查过程。

[0010] 简而言之,在基于手势识别和/或声音识别的人机交互中,有时在有限区域中存在要控制的若干设备。因为可能难以确定要控制哪个设备 (若干设备可能同时响应命令),所以存在歧义问题。

[0011] 因此,需要一种允许语音/手势控制的设备比邻地 (诸如在家庭住所中) 进行操作的消歧 (disambiguation) 系统。

[0012] 由于邻近而引起的上述歧义问题已经在使用无线游戏控制手柄 (paddle) 和一个或多个接近的控制台的背景下的家庭娱乐系统中被意识到并且进行了讨论。已知的解决方案尝试通过初始的 (一次使用的) 配对处理来解决该问题,其通过由最终用户明确进行的交

互式的配对处理来定义每个控制手柄与所选择的控制台之间的绑定(认为PS3和Wii控制台正在使用这种方法)。但是,预料到这些解决方案不经常被执行,其通常将要求用户和与配对处理有关的设备的物理交互,并且因此最终无法容易地扩展到在日常的语音/手势多设备控制使用中。而且,在一些情况下,用户可能远离要控制的设备一定距离,其中传统的基于按钮的接口无法被应用于控制。

[0013] 因此,至少对于基于手势识别和/或声音识别的系统,仍然需要改进人机接口的消歧系统。

发明内容

[0014] 本公开描述一种装置以及用户控制装置的方法。所提出的装置和方法适合于在执行与来自用户的第一事件相对应的命令之前,提供第二事件,以使用户确认第一事件的意图。本发明的实施例在具有基于手势识别和/或声音识别的HMI的系统中实现。本发明的实施例改进基于手势识别和/或声音识别的系统的人机接口的消歧系统。

[0015] 根据本公开的第一方面,提出一种装置。该装置包括:检测单元,用于检测与包括所述装置的多个设备有关的由用户执行的第一事件;控制单元,用于根据由检测单元所检测到的第一事件来产生第二事件;输出单元,用于将第二事件呈现给用户,其中控制单元根据由检测单元所检测到的用户对第二事件的响应来产生用于在装置上执行第一事件的命令。

[0016] 在实施例中,检测单元检测手势和/或声音形式的第一事件。

[0017] 在实施例中,第一事件是用于在装置与用户之间发起交互会话的触发事件。

[0018] 在实施例中,第二事件是用户用以验证发起交互会话的意向的验证事件。

[0019] 在实施例中,验证事件包括用户用以口头重复的文本消息。

[0020] 在实施例中,验证事件包括要求用户执行活动的文本、图像、音频或视频形式的询问。

[0021] 在实施例中,控制单元将命令传送给执行第一事件所需要的组件和相关设备。

[0022] 在实施例中,输出单元以由第二事件的形式所确定的方式将第二事件呈现给用户。

[0023] 根据本公开的第二方面,提出一种用户控制装置的方法。该方法包括,在装置的级别:接收与包括所述装置的多个设备有关的由用户执行的第一事件;根据触发事件来产生第二事件;将第二事件呈现给用户;以及根据用户对第二事件的响应,在所述装置上而不是在多个设备中的其他设备上,执行与第一事件相对应的命令。

[0024] 在实施例中,第一事件是以手势和/或声音的形式。

[0025] 在实施例中,第二事件包括用户用以口头重复的文本消息。

[0026] 在实施例中,第二事件包括要求用户执行活动的文本、图像、音频或视频形式的询问。

[0027] 根据本公开的第三方面,提出一种计算机程序产品,其可以从通信网络下载和/或记录在计算机可读取和/或处理器可执行的介质上。该计算机程序包括用于实现根据本公开的第二方面的方法的步骤的程序代码指令。

[0028] 根据本公开的第四方面,提出一种非临时性计算机可读介质,其包括记录于其上

并且能够由处理器运行的计算机程序产品。该非临时性计算机可读介质包括用于实现根据本公开的第二方面的方法的步骤的程序代码指令。

[0029] 要理解的是,在下面的本发明的详细描述中将发现本发明的更多方面和优点。

附图说明

[0030] 包括附图以与用于解释本发明的实施例的原理的描述一起提供对本发明的实施例的进一步理解。本发明并不局限于所述实施例。

[0031] 在附图中:

[0032] 图1例示根据本发明的实施例的可以实现本公开的示例性系统;

[0033] 图2是示出根据本发明的实施例的图1的系统中的装置的结构框图;

[0034] 图3是示出根据本发明的实施例的用户控制装置的方法流程图;以及

[0035] 图4例示在通过声音命令控制的媒体设备的环境中实现的根据本发明的实施例的方法的示例性示例。

具体实施方式

[0036] 现在将结合附图详细地描述本发明的实施例。在下面的描述中,为了简洁,可能省略对已知功能和配置的一些详细描述。

[0037] 图1例示可以实现本公开的示例性系统100。系统100可以是基于手势识别和/或声音识别处理的任何种类的HMI系统,诸如微软控制台XBOXOne、PS3和任天堂Wii控制台等。如图1所示,用户可以通过声音或手势与装置1、2、……和n中的一个交互。对于系统100,将不给出更多的细节。

[0038] 图2是示出根据本发明的实施例的系统100中的装置1的结构框图。

[0039] 如图2所示,装置1包括用于检测关于装置1的由用户执行的第一事件的检测单元201。第一事件可以是用户可以用以与装置1交互的各种形式的手势和/或声音。在一个示例中,第一事件可以是被定义为用于发起用户与装置1之间的交互会话的特定手势和/或语音事件的触发事件。不同的传感器和检测技术可以被用于检测单元201。将不给出更多的细节,因为其对于本领域技术人员是众所周知的,并且超出本公开的范围。

[0040] 装置1进一步包括用于根据由检测单元201检测的第一事件来产生第二事件的控制单元202。检测来自用户的第一事件的在图1中示出的系统100中的每个装置可以根据装置打算验证由用户发起的交互的意图的方式来产生单独的第二事件。例如,在第一事件是发起交互会话的触发事件的情况下,第二事件可以是用户验证他/她发起交互会话的意图的验证事件。

[0041] 第二事件可以是以下面的形式:

[0042] 1) 文本消息,诸如用户用以口头重复的独特的单词、数字代码等;以及

[0043] 2) 以文本、图像、音频或视频的形式的要求用户执行诸如手势这样的活动的精细询问(elaborate query)。

[0044] 对于上面的形式1),例如它可以是用户用以口头重复的简单的独特数字。

[0045] 对于上面的形式2),例如它可以是文本询问,例如“用你的手画圆”,使用户执行诸如手势这样的活动。它也可以是声音询问,诸如用于用户执行的“重复当前显示的数字”。询

问也可以是显示的图像/视频的形式。例如,可以拍摄手势的短视频/动画以使用户来模仿。

[0046] 可以意识到,其他适当的形式也可以被应用于第二事件。将不提供更多的细节。

[0047] 装置1进一步包括用于将第二事件呈现给用户的输出单元203。可以意识到,第二事件可以根据第二事件的形式以适当的方式来呈现。例如,如果第二事件是文本消息,则前面板LCD可以用来呈现第二事件。以精细询问的形式的第二事件可以在连接且活动(linked-and-active)的电视机上呈现给用户。如果第二事件是音频询问,则可能需要扬声器,在这种情况下,优选地使用随机延迟来呈现第二事件,以避免与系统100中的其他装置(如果有)重叠。

[0048] 用于呈现第二事件的设备可以是装置1的一部分,或者是独立于装置1的设备。

[0049] 在一个示例中,输出单元203可以允许用户设置关于呈现第二事件的设备的优选。例如,如果存在LCD前显示器和扬声器输出用于呈现,则输出单元203可以允许用户设置哪一个为优选。

[0050] 用户将根据第二事件的要求对来自他/她真正打算控制的装置的所呈现的第二事件进行正确的响应。

[0051] 如果第二事件是显示在前面板LCD上的要求用户读出的文本消息,则用户可以通过读出所显示的文本单词或代码来进行响应。例如,第二事件可以是文本消息“重复当前显示的数字:0011”。然后,用户可以通过读出数字“0011”来做出正确的响应。

[0052] 如果第二事件是要求用户做出附加手势的精细询问,则用户可以通过执行由精细询问所指示的附加手势来做出响应。例如,第二事件可以是文本询问“用你的手画圆”。然后,用户可以通过执行第二事件所要求的手势来进行正确的响应。

[0053] 如果第二事件是音频形式的给用户的问题,则用户可以通过对该音频询问给出答案来进行响应。

[0054] 在一个示例中,输出单元203可以允许用户根据设备交互能力,关于要做出的响应的类型,针对每个设备设置优选。例如,用户可以设置声音响应在一种情况下是优选的,或者手势响应在另一种情况下是优选的。

[0055] 检测单元201将检测来自用户的响应。当检测到对第二事件的正确的响应时,控制单元202将产生用于执行第一事件的命令,并且将该命令传送给系统100中的必要组件或设备。在一个示例中,第一事件是用于在用户与装置1之间发起交互会话的触发事件。在这种情况下,如果检测单元201接收到对第二事件的正确的响应,则控制单元202将接通装置1的必要组件以发起与用户的交互会话。可以意识到,交互会话所需要的系统100中的其他设备也将被接通和操作。

[0056] 对它们的第二事件在延迟之后没有检测到正确的响应或者检测到无效的响应的其他装置(例如,在图1中所示的装置2至n)将不发起与用户的交互会话。

[0057] 图3是示出根据本发明的实施例的用户控制装置的方法的流程图。

[0058] 如图3所示,该方法包括由装置接收来自用户的第一事件的步骤S301。

[0059] 在步骤S302,装置根据触发事件来产生第二事件。

[0060] 在步骤S303,装置将第二事件呈现给用户。

[0061] 在步骤S304,当接收到用户对第二事件的正确的响应时,装置执行与第一事件相对应的命令。

[0062] 示例

[0063] 接下来,将参考图4,在通过声音命令控制的媒体设备的背景下描述根据本发明的实施例的方法的示例。

[0064] 在该示例中,媒体设备401装配有麦克风,诸如具有LCD面板显示器402并且连接到在图4中没有示出的外部显示器(例如,电视机)的增强型蓝光读取器。在这个意义上,媒体设备401是声音控制的。

[0065] 在图4中的步骤1,用户可以执行用于发起与媒体设备401的交互会话的触发事件。触发事件可以是特定的单词或手势,其被视为是来自蓝光设备的激活事件。例如,触发事件用于读取插入的盘的目的。

[0066] 当接收到触发事件时,媒体设备401将开始消歧过程,以检查用户是否真地想要建立交互会话。

[0067] 消歧过程包括:首先,根据接收的触发事件来产生特定的质疑询问(challenge query),如图4中的步骤2所示。在该示例中,质疑询问是任意文本代码。

[0068] 然后,任意文本代码以适当的方式呈现给用户。具体地,该任意文本代码可以连同设备信息(诸如设备别名、型号等)一起在电视机还没有激活的情况下显示在LCD面板显示器402上,或者在其他情况下显示在电视机上,或者甚至显示在二者上。如图4所示,显示下面的任意文本代码:

[0069] DVD播放器TechniC.

[0070] 会话发起确认

[0071] 请重复代码:

[0072] 1456

[0073] 如下所述,任意文本代码将帮助用户确认他/她想要开始与其进行交互会话的设备。

[0074] 在该示例中,作为对质疑询问的响应,如在图4中示出的步骤3,用户重复所显示的代码文本以确认发起与该媒体设备401的交互会话的意图。

[0075] 当接收到响应时,在图4中的步骤4,媒体设备401执行与第一事件相对应的命令,即发起交互会话。为此,需要涉及媒体设备401或系统中的设备的一些组件。例如,如果电视机在仅使用LCD面板显示器402的消歧过程期间处于睡眠模式,则可能需要由蓝光读取器将其激活。

[0076] 本发明的实施例提供一种计算机程序产品,其可以从通信网络下载和/或记录在计算机可读取和/或处理器可执行的介质上,包括用于实现上述方法的步骤的程序代码指令。

[0077] 本发明的实施例提供一种非临时性计算机可读介质,其包括记录于其上并且能够由处理器运行的计算机程序产品,包括用于实现上述方法的步骤的程序代码指令。

[0078] 要理解的是,本发明可以实现为各种形式的硬件、软件、固件、专用处理器或其组合。而且,软件优选地实现为有形地实施在程序存储设备上的应用程序。应用程序可以被上载到包括任何适当架构的机器并由其执行。优选地,机器在具有诸如一个或多个中央处理单元(CPU)、随机存取存储器(RAM)以及输入/输出(I/O)接口这样的硬件的计算机平台上实现。计算机平台也包括操作系统和微指令代码。这里描述的各种处理和功能可以是经

由操作系统执行的微指令代码的一部分或应用程序的一部分(或者其组合)。另外,诸如另外的数据存储设备和打印设备这样的各种其他外围设备可以连接到计算机平台。

[0079] 还要理解的是,因为在附图中所示的构成系统组件和方法步骤中的一些优选地以软件来实现,所以系统组件(或处理步骤)之间的实际连接可以根据本发明被编程的方式而不同。给出本文中的教导,本领域的普通技术人员将能够想到本发明的这些以及类似的实现方式或配置。

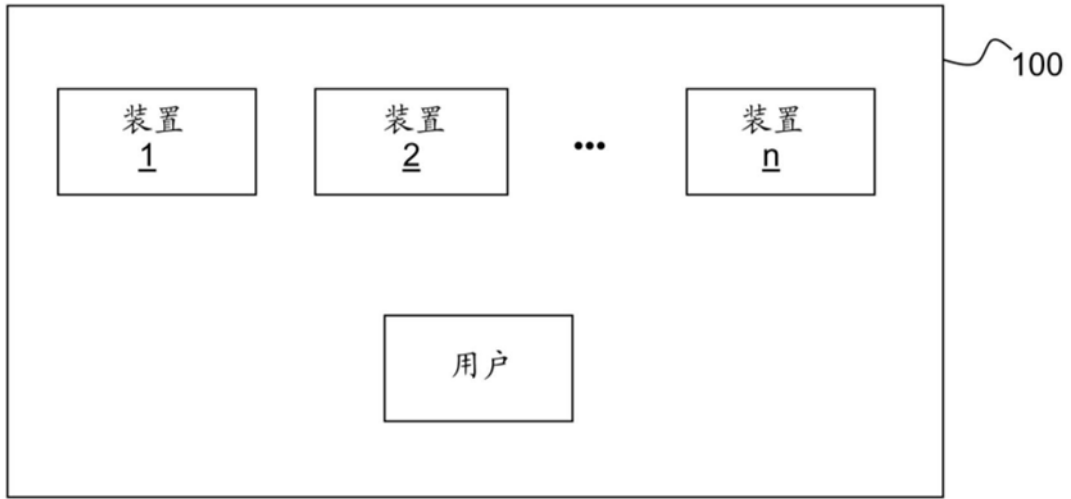


图1

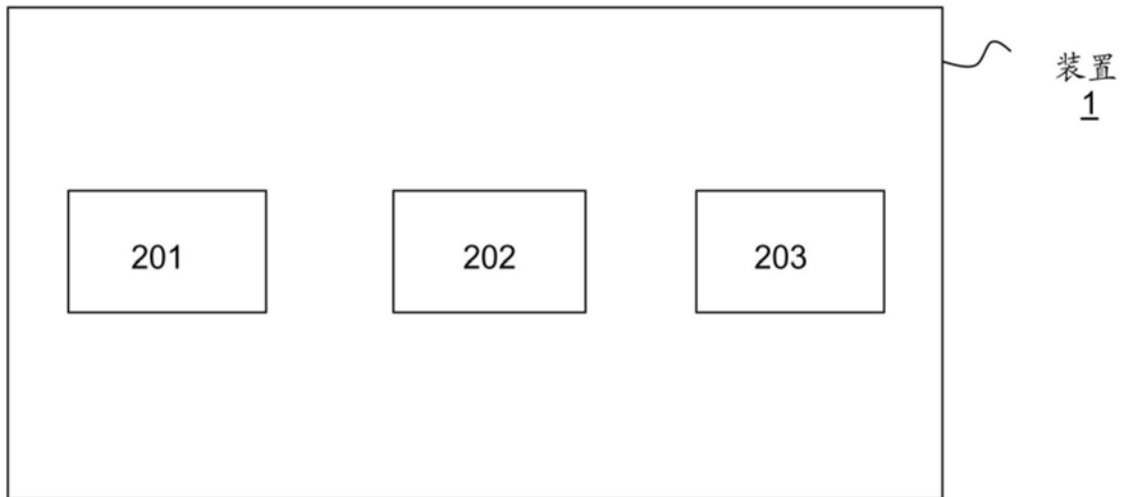


图2

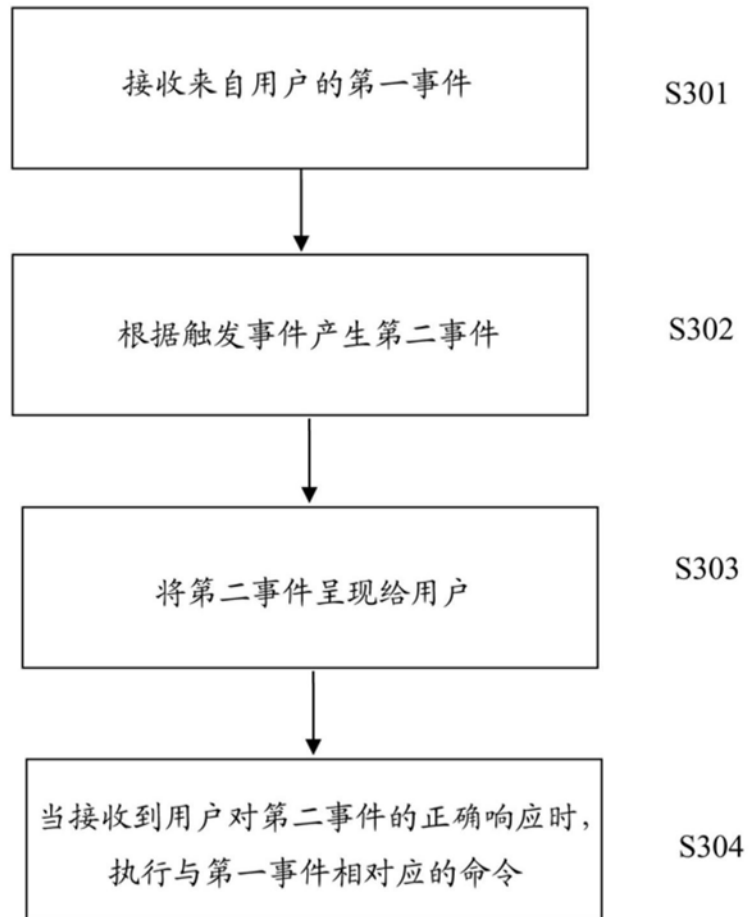


图3

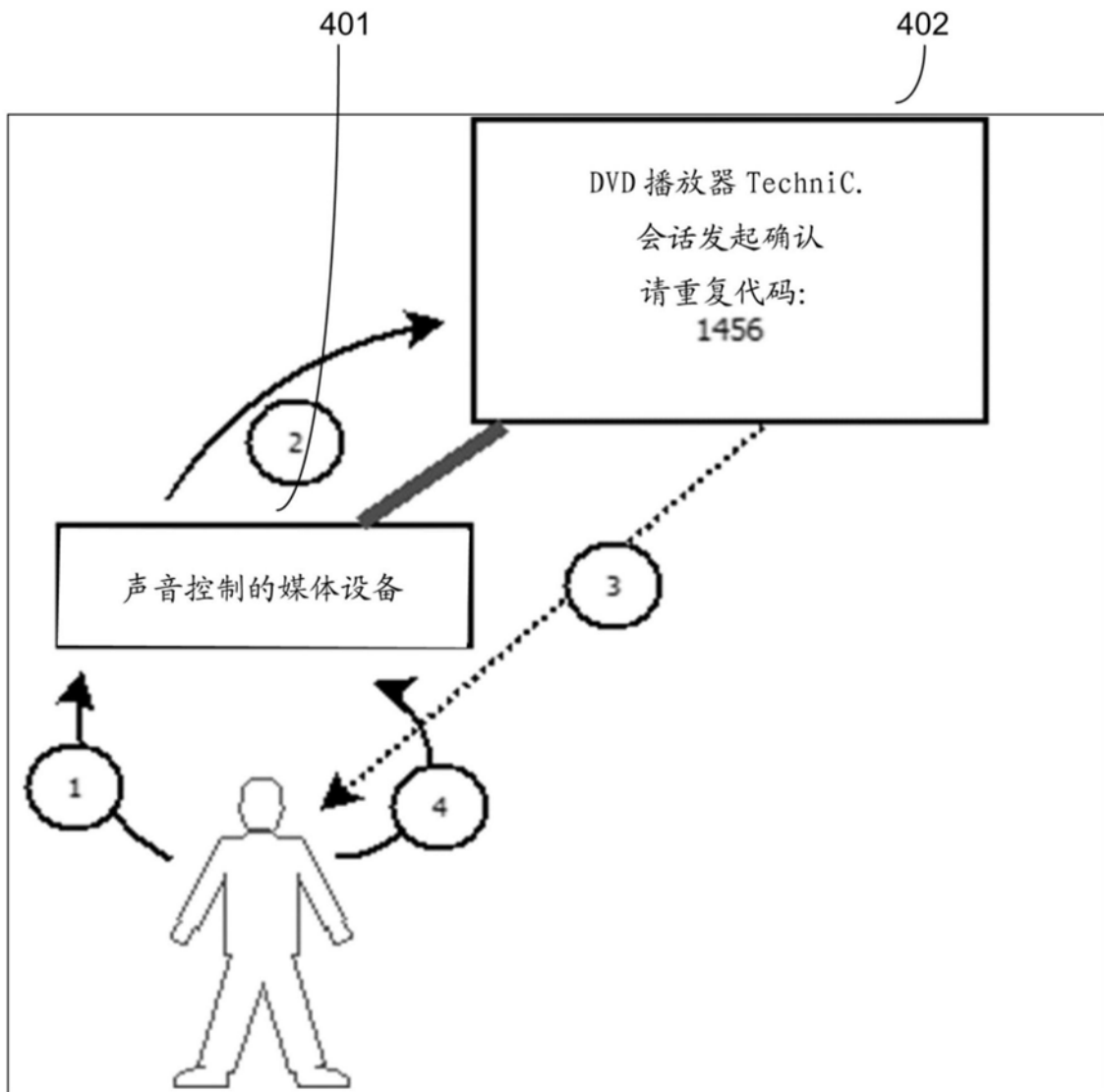


图4