



## (12) 发明专利

(10) 授权公告号 CN 103049090 B

(45) 授权公告日 2016. 05. 18

(21) 申请号 201210555317. 2

(22) 申请日 2012. 12. 19

(30) 优先权数据

13/332, 348 2011. 12. 20 US

(73) 专利权人 微软技术许可有限责任公司

地址 美国华盛顿州

(72) 发明人 陈瑞龙 什韦塔克·帕特尔

丹尼尔·莫里斯 西德汉特·古普塔

(74) 专利代理机构 北京集佳知识产权代理有限公司

公司 11227

代理人 郎晓虹 陈炜

(51) Int. Cl.

G06F 3/02(2006. 01)

(56) 对比文件

US 2011/0103448 A1, 2011. 05. 05,

US 2007/0121097 A1, 2007. 05. 31,

TW 201112049 A1, 2011. 04. 01,

审查员 邢丽超

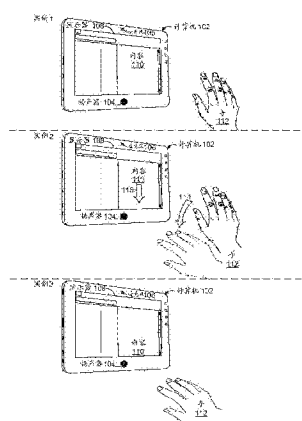
权利要求书2页 说明书11页 附图6页

(54) 发明名称

用户控制姿势检测

(57) 摘要

本发明公开了用户控制姿势检测。一个示例允许扬声器和麦克风执行第一功能。所述示例同时利用扬声器和麦克风来执行第二功能。第二功能包括使用麦克风捕捉来自扬声器的声音信号并且检测在声音信号中的多普勒频移。将在计算机附近执行的用户控制姿势与多普勒频移相关联，并且将用户控制姿势映射到控制功能。



1. 一种计算机 (802), 包括 :

扬声器 (804) 和麦克风 (806) ; 以及

姿势识别组件 (808), 所述姿势识别组件被配置为使扬声器发射给定频率的声音, 并且使麦克风捕捉所述声音, 所述姿势识别组件还被配置为在给定频率周围评估被捕捉声音的频谱, 以检测由用户姿势引起的多普勒效应频移, 所述姿势识别组件还被配置为将多普勒效应频移映射到个体控制功能, 并且使所述计算机执行所述控制功能, 并且其中, 所述姿势识别组件还被配置为使所述扬声器和所述麦克风在发射和捕捉期间执行与姿势识别不同的其他功能。

2. 如权利要求 1 所述的计算机, 还包括用户输入设备, 并且其中, 所述姿势识别组件还被配置成 : 当用户使用所述用户输入设备时, 不评估所述被捕捉声音。

3. 如权利要求 2 所述的计算机, 其中, 所述用户输入设备包括物理键盘、虚拟键盘、鼠标、触控板或触摸屏。

4. 如权利要求 1 所述的计算机, 其中, 所述扬声器包括单个扬声器或多个扬声器, 并且所述麦克风包括单个麦克风或多个麦克风。

5. 如权利要求 1 所述的计算机, 其中, 所述姿势识别组件还被配置为允许用户定义自计算机起在三维空间中放射状延伸的范围, 所述控制功能将在所述范围内执行而在所述范围外不执行。

6. 如权利要求 5 所述的计算机, 其中, 所述姿势识别组件还被配置为评估多普勒效应频移的振幅, 以确定在所述范围内是否出现用户姿势。

7. 如权利要求 1 所述的计算机, 其中, 所述扬声器的其他功能是播放音乐, 并且其中, 给定频率的声音包括音乐的一部分。

8. 如权利要求 1 所述的计算机, 其中, 所述姿势识别组件被配置为测试频率范围和在所测试的频率范围中评估被捕捉的声音, 以及当与所述给定频率相关的干扰比所述频率范围内的其他频率相对较小时, 选择所述给定频率。

9. 如权利要求 8 所述的计算机, 其中, 所述干扰包括由所述扬声器产生的其他声音或背景噪声。

10. 如权利要求 8 所述的计算机, 其中, 所述姿势识别组件被配置为根据一个或更多个参数来重新测试所述频率范围。

11. 如权利要求 10 所述的计算机, 其中, 所述一个或更多个参数涉及预定义的时间段或涉及在超过预定义阈值的给定频率处检测到的干扰。

12. 一种识别用户控制姿势的方法, 包括 :

从计算机发射 (902) 声音信号 ;

在所述计算机处捕捉 (904) 声音 ;

将所发射的声音信号的频率范围与所捕捉的声音的同一频率范围相比较 (906), 以检测由用户控制姿势所产生的多普勒频移 ;

将所述用户控制姿势映射 (908) 到控制功能 ; 以及

在所述计算设备上执行 (910) 所述控制功能,

其中, 所述发射包括从单个扬声器组同时发射在恒定频率处的声音信号和其他声音信号。

13. 如权利要求 12 所述的识别用户控制姿势的方法,其中,多个麦克风执行所述捕捉,并且所述捕捉还包括比较所检测的、所述多个麦克风的多普勒频移的差异,以进一步识别关于所述用户控制姿势的信息。

## 用户控制姿势检测

### 技术领域

[0001] 本公开涉及用户控制姿势检测。

### 背景技术

[0002] 在日常生活中计算机普遍存在。用户往往拥有许多不同的计算机或与许多不同的计算机进行交互,而不是只拥有“一台计算机”。例如,一般而言,用户可能在工作的地方拥有台式计算机,在家里拥有膝上型计算机、智能电话、平板式计算机、视频游戏控制台和/或电子书,等等。此外,用户往往与集成到其他设备中的其他计算机进行交互。例如,用户的车可能具有用户与之交互以获取方向、收听音乐等的计算机。已经为所述不同类型的计算机开发了各种不同的用户控制器。例如,鼠标和物理键盘往往是控制台式计算机和与台式计算机交互的最常见的方式。对于笔记本式计算机,鼠标不是非常方便,其通常使用触控板式设备来代替鼠标。鼠标或触控板都不适用于智能手机和平板式计算机,智能手机和笔记本计算机往往具有用户可以接触以控制设备的触摸屏(例如,虚拟键盘)。

[0003] 这些控制技术中的每一种都具有以下缺点:例如,通过接触触摸屏来控制设备常常弄污整个屏幕,和/或如果用户佩戴手套则不能通过接触触摸屏来控制设备。可能更重要的是,所述控制技术不能适用于所有的设备类型。因此用户必须学习控制他/她的不同设备的不同技术。例如,假设在一天中用户在他/她的每一个计算设备上收听音乐并且希望关闭该音乐。当在台式计算机上收听音乐时,用户可使用鼠标来使音乐播放应用从“播放”变为“停止”。用户可能以类似的方式使用笔记本式计算机的触控板。在车中,用户可能使用特定的硬件开关来停止音乐。在平板式计算机或智能电话上,用户可触摸在触摸屏上的“停止”图标。这种连贯性的缺乏带来了很多不便。此外,当用户尝试控制用户不熟悉的计算机时,会更加不方便。本构思涉及可被应用来控制计算机和可被容易地应用到任何(或所有)计算机(诸如以上列出的计算机等)的技术。

### 发明内容

[0004] 所描述的实施例涉及用户控制姿势。一个示例允许扬声器和麦克风执行第一功能。所述示例同时利用扬声器和麦克风来执行第二功能。第二功能包括使用麦克风捕捉来自扬声器的声音信号并且检测声音信号中的多普勒频移。将用户在计算机附近做出的控制姿势和多普勒频移相关联并且将用户的控制姿势映射到控制功能。

### 附图说明

[0005] 附图示出了在本申请中所表示的构思的实施例。通过参考结合附图的以下说明,可更容易地理解图示的实施例的特征。在各个附图中相同的附图标记可在任何情况下用于指示相同的元件。此外,每一个附图标记的最左端的数字表示第一次引入该附图标记的附图和相关描述。

[0006] 图1-图7示出了根据本构思的一些实施例的基于多普勒的用户控制姿势检测方

案的示例。

[0007] 图 8 示出了根据本构思的一些实施例的可使用基于多普勒的用户控制姿势检测技术的示例系统。

[0008] 图 9 示出了根据本构思的一些实施例的基于多普勒的用户控制姿势检测方法或技术的示例的流程图。

## 具体实施方式

### [0009] 综述

[0010] 本申请涉及利用用户控制姿势(以下,“用户姿势”)来控制计算机。更具体地,可通过计算机发射音频信号和检测所发射的信号,来检测用户姿势。可基于多普勒效应或多普勒频移来分析在发射信号和检测信号之间的任何差异以检测用户姿势。从被分析的音频信号数据中可以识别出各个用户姿势。各个用户姿势可被映射到计算机的特定控制功能。可以实现所述特定控制功能以控制计算机。

### [0011] 使用情形示例

[0012] 为了进行说明,首先来看图 1,图 1 示出了表示为平板式计算机的计算机 102。计算机 102 包括扬声器 104、麦克风 106 和显示器 108。在所述示例中,在显示器 108 的一部分上显示有内容 110(诸如文本)。

[0013] 图 1 分成实例 1 到实例 3,以帮助示出用户可通过作出用户姿势来控制计算机的方式。在实例 1-3 中,计算机 102 可以通过扬声器 104 生成音频信号。麦克风 106 可捕捉所述音频信号。可将所捕捉的信号与发射信号相比较,以得出差异。这些差异可由于用户姿势对信号的影响(例如,其多普勒效应)而产生。不同的用户姿势产生不同的多普勒效应,并且从而可区分和识别各个用户姿势。

[0014] 在图示的示例中,用户姿势是用户的右手 112 作出的手势。(在一些实施例中,所述姿势可以替代地由左手执行,这里为了避免在图页上造成混乱而没有示出)。在所述示例中,如实例 1 所示,用户从手 112 大体上与显示器 108 平行的位置开始。然后,如实例 2 中所示,用户通过向下倾斜(箭头 114)手而作出用户姿势。用户姿势 114 作为接收信号与发射信号之间的差异通过计算机 102 来检测。换一种说法,用户姿势可以导致被捕捉信号的频率特征(frequency profile)相对于没有用户姿势的情况下所期望的频率特征而变化。用户的手可引起所述差异或变化,进而在接收或捕捉到的信号上产生多普勒频移。在所述示例中,计算机将所检测到的姿势映射到“向下滚动命令”。然后,计算机向下滚动内容 110(箭头 116)。

[0015] 实例 3 示出了在姿势结束时用户的手 112,此时大体上垂直于显示器 108。此外,当与实例 1 相比时,内容 110 已经向下滚动。此外还注意到,在实例 1-3 中用户不需要实际接触计算机来完成所述控制功能。

[0016] 如上所述,计算机 102 包括扬声器 104 和麦克风 106。扬声器和麦克风可以是只用于用户姿势检测功能的专用硬件设备。然而,并不需如此。事实上,大多数计算机已经包括至少一个麦克风和至少一个扬声器。可以利用这些现有的硬件设备减少涉及增加另外的硬件设备的复杂性和成本。此外,现有的扬声器和麦克风可以实现本姿势检测功能,而不会中断或削弱所述设备的现有用途。例如,假设在实例 1 中,用户正在进行电话呼叫(仅音频)

或网络聊天(音频和视频(为了简明而未示出,计算机 102 可包括摄像头))。因此,在电话呼叫期间,扬声器 104 正在播放来自另外一方的音频信号,而麦克风 106 正在捕捉音频信号(诸如用户讲话)。同时,扬声器可以发送一个或更多个音频信号以在姿势检测功能中使用。在一些实施例中,所述音频信号可以是在用户不能听到的频率上发射的稳定的音调。例如,大多数的扬声器可以发射高达 22k-24k 赫兹的信号,并且大多数的麦克风可以捕捉这些信号。然而,人类通常不能听到任何高于大约 18k-20k 赫兹的声音。因此,扬声器 104 和麦克风 106 可以同时用于电话呼叫和用户姿势检测,而用户只会听到正常的电话呼叫。当用户没有作出姿势时,发射的姿势检测信号和捕捉到的姿势检测信号通常相同。例如,稳定的 20k 赫兹信号被发射和接收。

[0017] 当用户进行电话呼叫时,用户可能正在观看内容 110。用户可能希望观看内容 110 上面的内容并且能够做出用户姿势 114 来使内容向下滚动。该运动改变音频信号,使得被捕捉的信号与发射的姿势检测信号稍有不同。可以分析所述差异以确定用户的姿势(例如,在不同姿势之间进行区分)。一旦完成姿势,则在发射的姿势检测信号和检测到的姿势检测信号之间的差异减小。不管由扬声器和麦克风执行的双重功能如何,用户的电话呼叫在正常模式下进行(例如,用户甚至不需要知道到同时影响扬声器 104 和麦克风 106 的潜在姿势检测功能)。在其他的示例中,可利用用户姿势来影响第一功能和/或与第一功能协同操作。例如,可使用用户姿势来减弱电话呼叫或控制电话呼叫的音量等。

[0018] 此外,该用户姿势检测实施例不依赖于用户手部状态。例如,用户可以佩戴手套或用户的手可以是脏的或油腻的。所述状态可能妨碍触摸屏检测用户触摸的能力。然而,本实施例不需要用户接触并且可不受手套、污垢等的影响。

[0019] 此外,该用户姿势识别实施例可以容易地应用到每一台用户计算机,使得用户可以使用相同的用户姿势来控制他的/她的各种不同的计算机。例如,用户可以使用图示的姿势,以与图 1 中用户控制平板式计算机 102 相同的方式来控制他的/她的台式计算机、内置式汽车计算机、智能手机和笔记本式计算机(等等)。

[0020] 图 2-图 7 共同地示出了涉及计算机 202 的另一个基于多普勒的用户姿势检测实施例。在该情形中,计算机 202 被表示为笔记本式计算机。计算机 202 包括两个扬声器 204 (1)和 204 (2)、两个麦克风 206 (1)和 206 (2)、显示器 208、以及表示为键盘/触控板 210 的用户输入设备。在所述示例中,在所述计算机上运行音乐播放器应用 212。如图 2 中 214 所示,音乐播放器处于“停止”状态(例如,未播放)。假设在所述计算机上运行基于多普勒的用户姿势检测特征。例如,基于多普勒的用户姿势检测特征可以使扬声器 204 (1)和 204 (2)中的一个或两个发射一个或更多个信号,并且麦克风 206 (1)和 206 (2)中的一个或两个捕捉所述信号。如以上与图 1 有关的说明,可以分析发射信号和捕捉到的信号之间的差异并且所述差异可以被识别为用于控制计算机的用户姿势。

[0021] 图 3 示出了用户分别使用他的/她的左手和右手 302 和 304 来使用键盘/触控板 210。在该实施例中,使用键盘/触控板 210 阻止基于多普勒的用户姿势检测特征以避免将打字运动认作是用户姿势。例如,基于多普勒的用户姿势检测特征可停止使得信号从扬声器发出,或当用户使用键盘/触控板时,其可简单地不分析捕捉到的信号。

[0022] 在其他示例中,基于多普勒的用户姿势检测特征可与其他输入机制共同工作以提供在其他情形下用户不可使用的控制功能,或使所述控制功能对用户更友好。例如,在图

中未示出的示例中,在三维(3-D)应用中,用户可使用触控板来选择对象。接着用户能够使用触控板来“拖拽”所选择的对象。此外,通过执行与图 1 有关的、上面阐述和说明的边到边(side-to-side)的控制姿势,用户能够“翻转”或“旋转”所选择的对象。当然这只是一个示例,而且用户可以利用传统输入机制,结合基于多普勒的用户姿势检测来设计许多其他“组合”控制选择,从而根据需要控制计算机。

[0023] 继续所说明的示例,图 4 示出了用户离开键盘 / 触控板 210 的后续时间(例如,物理上不再进行键输入或触控板输入)。如此,基于多普勒的用户姿势检测特征再一次被激活。图 5 示出了用户如箭头 502 所示向前(朝向显示器 208)移动他的 / 她的左手 302,同时如箭头 504 所示向后(离开显示器 208)移动他的 / 她的右手 304。所述用户运动影响被扬声器 204 (1)和 / 或 204 (2)捕捉的声音。所述效果可被基于多普勒的用户姿势检测特征检测并且识别为用户姿势。在所述示例中,用户姿势被映射至音乐播放器或音乐播放器应用 212 的状态改变。如此,基于多普勒的用户姿势检测特征使音乐播放器从图 4 的停止状态 214 变为图 5 的“播放”状态 506。因此,如 508 处所示,扬声器发出音乐声音。

[0024] 图 6- 图 7 示出了即使用户不在计算机的前方并且不在屏幕的视线内,用户也可以使用用户姿势来控制计算机 202 的方案。在图 6 中,当用户接近计算机 202 时,如 508 处所示,计算机仍在输出音乐声音。通过重复关于图 4- 图 5 所讨论的用户姿势(即在使一只手朝向计算机移动的同时,使另外一只手移离计算机),用户可以改变计算机的音乐播放状态。在图 4- 图 5 中,用户姿势使计算机的音乐播放应用从停止状态切换到播放状态。在图 6 中,计算机仍处于播放状态中,并且重复用户姿势将计算机切换回到停止状态。这表现为图 7 中计算机没有发出音符。用户姿势可以被计算机识别并且用来控制计算机,而用户不需要在计算机的前方或在计算机的诸如摄像设备等某些设备的视线内。事实上,在一些实施例中,只要用户在自计算机起放射状地延伸的所定义的范围,用户就可以从任何方向或维度来控制计算机。

[0025] 总结关于图 1- 图 7 的讨论,基于多普勒的用户姿势检测实施例检测在计算机周围的移动并且使用检测到的移动的特性(诸如速度、方向和 / 或振幅)来识别一组丰富的用户姿势。例如,通过追踪方向和速度,向上或向下移动手的用户姿势可以实时滚动网页。基于多普勒的用户姿势检测实施例还可以检测向相反方向移动的两只手,以及其他可检测的用户姿势。可在具有麦克风和扬声器的任何类型的计算机上实现所述构思。麦克风和扬声器不需要专用于姿势检测,并且可以同时执行其他功能。所阐述的实施例是鲁棒的,并且在繁忙和喧闹的环境以及安静的环境中均可良好地执行。所述构思可以在多个不同类型的设备上实现,以使得用户可以使用相同的用户姿势来控制任何设备。

[0026] 系统示例

[0027] 图 8 示出了可以实现上面引入的基于多普勒的用户姿势检测构思的系统 800 的示例。系统 800 包括计算机 802,在所述示例中所述计算机 802 被表示为智能手机。在所述示例中,计算机 802 包括扬声器 804、麦克风 806、基于多普勒的用户姿势识别组件 808、操作系统 810、处理器 812 和存储器 814。基于多普勒的用户姿势识别组件 808 可以是应用程序、应用部件或操作系统 810 的一部分。

[0028] 计算机 802 可以配置成通过网络 818 与其他的计算机和 / 或计算资源(诸如位于云 816 中的计算资源)进行通信。

[0029] 在这里使用的术语“计算机”或“计算设备”可指具有一定处理能力和 / 或存储能力的任何类型的设备。计算机的示例可以包括常规计算设备(诸如个人计算机、移动手机、智能手机、个人数字助理、照相机、摄像机),或各种不断发展的或还在开发的计算设备类型。计算机还可以是更大设备的组件。例如,在许多新兴的车中包含仪表盘计算机。

[0030] 处理能力可以由一个或多个处理器(诸如处理器 812 等)来提供,所述处理器可以运行具有计算机可读指令形式的数据以提供功能。诸如计算机可读指令等数据可以存储在存储器(诸如存储器 814)上。存储器可以位于所述计算设备的内部和 / 或外部。例如,存储卡 820 可被认为是由计算机以可拆卸的方式容纳的存储器。存储器可以包括易失性或非易失性的存储器、硬盘、闪存设备和 / 或光学存储设备(例如,CD、DVD 等)等中的任何一个或更多个。本文使用的术语“计算机可读介质”可以包括瞬时或非瞬时指令。相比之下,术语“计算机可读存储介质”不包括瞬时实例。计算机可读存储介质可以包括“计算机可读存储设备”。计算机可读存储设备的示例包括易失性存储介质(诸如 RAM)和非易失性存储介质(诸如,硬盘、光盘和闪存)。

[0031] 基于多普勒的用户姿势识别组件 808 可以包括和 / 或访问信号控制模块 822、信号分析模块 824、用户交互模块 826 和数据表(datatable) 828。

[0032] 简言之,信号控制模块 822 可以控制由基于多普勒的用户姿势识别组件 808 发射何种信号来供使用。信号控制模块 822 还可以控制信号捕捉。信号分析模块 824 可以分析所发射的和捕捉到的信号,以识别引起多普勒频移的用户姿势。用户交互模块 826 可以提供接口(诸如向用户解释如何正确地作出预安装姿势的 GUI)和 / 或可以允许用户定义各个姿势并且允许用户将所述姿势映射到各个计算机控制命令。这些映射关系可以存储在数据表 828 中。

[0033] 回顾上文,基于多普勒的用户姿势识别组件 808 可以使用公知的多普勒频移现象来感知移动。该移动可被检测为响应于移动对象的声波频率偏移,即被称为多普勒效应的效应。所述频率偏移与对象移动的速度和源频率成比例。在本实施例中,初始源(扬声器)和收听器(麦克风)是静止的,因此,在没有任何移动的情况下,没有频率变化。然而当用户移动他的 / 她的手时,其反射声波,引起频率的偏移。所述频率(fr)由麦克风测量并且可通过以下可被信号分析模块 824 使用的等式来描述。

$$[0034] \quad f_r = f_i \cdot \left( \frac{c+v}{c-v} \right)$$

[0035] 其中,  $f_r$  = 麦克风感知的频率;

[0036]  $f_i$  = 来自扬声器的初始频率;

[0037]  $c$  = 声音在空气中的速度;

[0038]  $v$  = 目标 / 手的速度。

[0039] 信号控制模块 822 可以产生连续的导频音(或多个导频音),所述导频音通过计算机的扬声器 804 播放。通常使用高频(日用的音频系统通常在 18-22kHz 的范围内),以使用户听不到导频音。然而,可以使用诸如低至 6kHz 以下等的低频。此外,对于给定的速度,频率越高,偏移越大,这使得在以给定的分辨率评估移动时计算起来更为容易。上限在很大程度上根据仅能够产生高达 22kHz 音频的大多数膝上型电脑和电话扬声器系统而变化。还注意到,在一些实施例中,信号控制模块可以被配置成使用与另一功能相关联的信号而不是



产生专用的导频音。例如,如果扬声器正在为用户播放音乐,则可以利用音乐的频率成分作为导频音。然后,发射的音乐频率可以与捕捉到的频率相比较以识别由用户姿势引起的差异。

[0040] 由于硬件的变化和在扬声器和麦克风系统中的滤波作用,在一些实施例中,信号控制模块 822 可以执行初始校准以发现一个或多个高性能的(并且可能是最佳的)音频(不需要用户的干涉)。在一种实施例中,信号控制模块 822 执行 500 毫秒的频率扫描,并且保持对峰值振幅测量以及所检测到的候选移动事件的数目(即,潜在的误报(false positives))进行跟踪。信号控制模块 822 可以选择所检测到的错误事件最少并且峰值极其孤立(即,振幅比扫描范围内的下一个最高峰值大至少 3dB)的最高频率。注意,可以周期地重复频率扫描,和/或可在出现一些其他事件时或根据各种参数来重复频率扫描。例如,如果在一些点处在所捕捉的信号中检测到过量的噪声(例如,干扰),则可以重复频率扫描以搜寻“更干净的频率(cleaner frequency)”。例如,噪声可以与预定义的阈值相比较。如果噪声级超过预定义的阈值,则重复所述测试。潜在导频音的测试和重复测试可以使基于多普勒的用户姿势检测技术在不同的和/或变化的环境中良好地执行。

[0041] 此外,信号控制模块 822 可以同时地发射分别具有不同但稳定的频率的多个导频音。可完全分析相应的被捕捉信号以确定冗余度或获取更准确的结果。替代地,导频音中的一个可以表现出“更好”或“最好”(例如,“更干净”)的特性,并且其可被选择来用于分析。

[0042] 在发出高频导频音的情形下,在一个实施例中,在计算机 802 附近(根据速度为大约一米的范围)的任何移动将导致麦克风拾取多普勒偏移反射,以 44,100Hz 对麦克风进行连续采样。

[0043] 在一些实施例中,信号分析模块 824 可以缓冲从麦克风 806 引入的时域信号,并且使用 2048 点汉明窗矢量(Hamming window vectors)计算快速傅里叶变换(FFT)(Fast Fourier Transform)。这产生均匀分布在 22,050Hz 的频谱宽度上的 1024 点大小的矢量。在计算出每一 FFT 矢量之后,信号分析模块 824 对每一 FFT 矢量进一步进行处理。例如,在检测移动之后,信号分析模块可以执行信号调节、带宽提取,最后提取特征。

[0044] 在一些情形中,信号调节可被认为是利用与用户移动手所能达到的快速程度有关的测试数据。一些实施例可以在用户在计算机前方移动手时所能达到的最快速度为大约 3.9m/sec 的假设下进行操作。因此,所述实施例可以将感兴趣的信号适当地限定在 6m/sec。给定所述的取样速率和 FFT 采样数,这在发射的峰值的两侧上产生大约 33 个频率仓(frequency bin)。(一些实施例可以在以下假定下操作:用户使用作出用户姿势的某种类型的点击设备来产生更快速的运动。例如,用户可以持有和移动比用户移动他的/她的手时更快的标尺或铅笔。所述快速的运动可以被映射到单个控制命令)。

[0045] 信号分析模块 824 可以将频率偏移解释为用户姿势。例如,信号分析模块可以检测用户姿势,所述用户姿势有效地增加(或减少)导频音的带宽(即,窗口平均和频谱泄露使峰值的运动模糊)。为检测该用户姿势,信号分析模块 824 可以通过独立地扫描两侧上的频率仓直到振幅下降导频音峰值的 10%,来计算导频音的带宽。换一种说法,本技术不需要分析整个频谱,替代地,重点分析导频音周围的频谱(例如,在每一侧上的 200Hz 内)。此外,使用相对的振幅下降可以允许系统动态响应(诸如当用户改变扬声器的音量时)。

[0046] 在大多数的情况下,所述仓分析足以推断移动。然而,如果偏移足够大,则反射信号可与导频音的峰值分离,而非模糊所述峰值。为了解决所述问题,信号分析模块 824 可以执行第二次扫描,搜寻第一次扫描的停止点以外的范围。如果发现具有原始音调能量的至少 30% 的能量的第二峰值,则可以重复第一次扫描,以寻找由第二峰值计算得出的振幅下降。

[0047] 移动检测和特征提取:在一些实施例中,频率矢量具有 21.5Hz 的每仓分辨率。对于 20kHz 的导频音,这相当于检测慢至 18.5cm/sec 的运动。信号分析模块 824 自身可以考虑导频音的带宽,例如,在没有移动的情形下导频音的带宽为约~80Hz。根据日用硬件的声音系统的质量,导频音的带宽可以在 60Hz ~ 120Hz 的范围内变化(在音调的每一侧上为 1-3 个仓)。因此,当例如存在 4 个或更多个仓的频率偏移带宽时,信号分析模块 824 可能认为出现了“移动事件”。所述阈值允许在忽略由带宽变化而导致的误报的同时,检测手的十分缓慢的运动。

[0048] 可测量特性

[0049] 除了基本的频率偏移,信号分析模块 824 还可以计算用于推断姿势的其他有用特征。

[0050] 速度:测量到的频率变化与对象的绝对速度成比例。信号分析模块 824 可以测量在原始频率(例如,导频音)和反射频率之间的差异以区分慢速的、中速的和快速的姿势。

[0051] 方向:可以根据频率偏移的符号来确定手是朝向计算机移动还是移离计算机。正偏移表示朝向计算机移动,而负偏移表示移离计算机。在多个麦克风的情形中,信号分析模块 824 可以比较多个麦克风的被检测的多普勒偏移的差异,从而进一步识别与用户控制姿势有关的信息。

[0052] 目标的邻近和大小:随着目标移近计算机,所观察到的信号振幅增加,并且所观察到的信号振幅还随目标的大小和反射率增加。例如,更大的手或张开的手掌表现出比更小的或紧攥的手更大的振幅。

[0053] 时间变化:测量上述三个特性随时间的变化,允许信号分析模块 824 观察变化速率并且使用其来过滤杂散信号。例如,可以可靠地滤除持续时间段非常短的任何移动,而持续更长时间的移动事件可以用于识别类似走向计算机或离开计算机等活动。

[0054] 姿势和使用情况

[0055] 可以合并上述特征以形成复杂的姿势。例如,如以上关于图 1 的说明,滚动可以被有效地检测并且用来控制网络浏览器或文字处理应用,等等。一些实施例可以使用控制机制(clutching mechanism)以便当手回到特定位置时,减少和/或阻止无意中作出的滚动。例如,可以使用速度特征和仅当满足某一速度准则时才进行滚动来实现。

[0056] 双击或点击姿势:通过观察方向随时间的变化,可计算方向改变的频率。所述方向频率的值可以用于检测点击姿势,其还可用于将较慢的点击与快速点击(非常类似于鼠标双击)区分开。

[0057] 如图 2- 图 7 示出了双手来回移动的姿势,其要求双手同时在相反的方向上移动来激活该姿势。通过检测在同一 FFT 矢量中是否存在上下偏移的频率成分来检测该姿势。

[0058] 邻近和轻弹:当在相同的方向上检测到至少 N 个连续移动事件时,激活所述姿势。大的 N 可以表示人是否在行走(与移动手时相反,此时具有相对小的 N 个连续移动)。当用

户离开或走向计算机时,行走姿势可以用于使计算机自动休眠或自动唤醒计算机。以上列出的可检测姿势是出于解释目的而提供的,而许多其他的用户姿势是可以检测的。

[0059] 用户交互模块 826 可以在计算机上设置用户姿势。例如,一些实施例可以具有一种或更多种预定义的用户姿势。在所述情况下,用户交互模块可生成界面(诸如图形用户界面(graphical user interface(GUI) 等)),所述界面解释和 / 或显示用户如何做出各个用户姿势。用户交互模块还可以允许用户定义计算机控制映射至用户姿势的方式。例如,GUI 可以显示如何作出图 1 中示出的滚动姿势。

[0060] 然后,GUI 可以允许用户将用户姿势映射到他 / 她选择的计算机控制功能。在数据表 828 的水平行 832 中显示了所述的一个示例。所述数据表包括姿势名称列 834、姿势特征列 836 和控制功能列 838。在行 832 中,姿势名称是“滚动”并且姿势特征是“A”。“A”表示针对该姿势所记录的图解频率特征,为了简明而未示出该图解频率特征。在本情形中,滚动姿势被映射到如上述的、用户可调节的“向下滚动”控制功能。例如,用户可以将相应的控制功能改变为在关于图 1 说明的音乐播放器应用中的“降低音量”。

[0061] 在另一示例中,假设用户希望向用户‘教授’新的姿势。GUI 可以提供该特征。例如,GUI 可以指示用户执行一次或更多次用户姿势。在直到获得满意的姿势特征的用户姿势执行期间,用户交互模块 826 和 / 或信号分析模块 824 可以分析所述信号。例如,假设用户执行关于上述图 2- 图 7 示出的两手来回移动姿势。如数据表中的行 840 所示,信号分析模块 824 可以将该姿势特征识别为“B”。用户交互模块 826 可以允许用户定义相应的控制功能(诸如通过为用户提供控制功能列表)。在本情形中,假设用户选择了音乐播放器的‘改变播放状态’。

[0062] 在随后的时间点,用户可能希望控制计算机 802,并且可执行用户姿势。信号分析模块 824 可以捕捉和分析用户姿势,以确定所执行的姿势的被捕捉姿势特征是否与在数据表 828 中的任何姿势特征相匹配(或如相似度阈值所定义的近似匹配)。如果存在近似匹配,可以执行相应的控制功能(例如,来自匹配姿势特征的水平行)。

[0063] 在一些实施例中,如果作出了过多的错误近似匹配,则用户可以将相似度阈值调高。在相反情况下(即姿势没有被识别为匹配),用户可以通过用户交互模块 826 将所述相似度阈值调低。

[0064] 此外,用户交互模块 826 可以允许用户设置和 / 或调整检测的范围。例如,用户可以设置振幅阈值,使得仅仅频率特征中超过该振幅阈值的事件映射至数据表 828。替代地,用户可调整扬声器的音量,扬声器进而控制所发射的导频音的响度(例如,振幅)。音量控制可以用来调节有效检测范围,这在用户明确不希望拾取周围运动的拥挤情况下是有用的。

[0065] 在一些实施例中,基于多普勒的用户姿势识别组件 808 可以由制备计算机和 / 或向最终用户(例如,消费者)销售该计算机的制造商或中间商预先安装在计算机 802 上。在其他的实施例中,基于多普勒的用户姿势识别组件 808 可以被用户安装。例如,用户可以在网络 818 上下载基于多普勒的用户姿势识别组件作为计算机可读指令集,接下来将该计算机可读指令集以有形的形式存储在计算机上(诸如存储在存储器 814 上)。在另一个实施例中,用户可以在诸如存储卡 820 等存储设备上获取基于多普勒的用户姿势识别组件,以在他的 / 她的计算机上安装基于多普勒的用户姿势识别组件。

[0066] 此外,在所说明的实施例中,计算机 802 被配置为具有通用处理器 812 和存储器

814。在一些替代配置中,计算机可以包括芯片上系统(System on a chip, SOC)式设计。在所述情况下,计算机的各个组件提供的功能可被集成到单个 SOC 或多个连接的 SOC 上。在这样的一个示例中,所述计算机可以包括共享资源和专用资源。接口可有助于在共享资源和专用资源之间的通信。顾名思义,专用资源可以认为是包括专用于实现特定功能的各个部分。

[0067] 共享资源可以是可被多个功能所使用的存储器,处理单元等。在所述示例中,共享资源可以包括处理器。

[0068] 方法示例

[0069] 图 9 示出了与本构思的至少一些实施例相应的技术或方法 900 的流程图。

[0070] 在块 902 中,所述方法可以从计算机发射声音信号。声音信号可以从在计算机上的扬声器中发射。在一些实施例中,这些扬声器可以是可用来在计算机上实现其他功能的现有多用途扬声器。

[0071] 在块 904 中,所述方法可以捕捉计算机的声音信号。可以使用麦克风来捕捉所述声音信号。在一些实施例中,所述麦克风可以是可用来在计算机上实现其他功能的现有多用途麦克风。

[0072] 在块 906 中,所述方法可以对所发射的声音信号的频率范围和所捕捉的声音信号的同一频率范围进行比较,以检测用户控制姿势所产生的多普勒频移。

[0073] 在块 908 中,所述方法可以将用户控制姿势映射到控制功能。

[0074] 在块 910 中,所述方法可以在计算设备上执行所述控制功能。

[0075] 上述方法的说明次序不想被解释作限制,并且任何数量的所述块可以用任何次序来组合,以实现所述方法或替代方法。此外,可以在任何适合的硬件、软件、固件或其组合上实现所述方法,以使计算设备可以执行所述方法。在一种情况下,所述方法作为指令集而存储在计算机可读存储介质上,以便通过计算机运行所述指令集来使计算机执行所述方法。

[0076] 根据本公开的实施例,至少提供了如下技术方案:

[0077] 1. 一种计算机,包括:

[0078] 扬声器和麦克风;以及

[0079] 姿势识别组件,所述姿势识别组件被配置为使扬声器发射给定频率的声音,并且使麦克风捕捉所述声音,所述姿势识别组件还被配置为评估在给定频率周围的被捕捉声音,以检测由用户姿势引起的多普勒效应频移,所述姿势识别组件还被配置为将多普勒效应频移映射到个体控制功能,并且使所述计算机执行所述控制功能,并且其中,所述姿势识别组件还被配置为使所述扬声器和所述麦克风在发射和捕捉期间执行其他的功能。

[0080] 2. 如技术方案 1 所述的计算机,还包括用户输入设备,并且其中,所述姿势识别组件还被配置成:,当用户使用所述用户输入设备时,不评估所述被捕捉声音。

[0081] 3. 如技术方案 2 所述的计算机,其中,所述用户输入设备包括物理键盘、虚拟键盘、鼠标、触控板或触摸屏。

[0082] 4. 如技术方案 1 所述的计算机,其中,所述扬声器包括单个扬声器或多个扬声器,并且所述麦克风包括单个麦克风或多个麦克风。

[0083] 5. 如技术方案 1 所述的计算机,其中,所述姿势识别组件还被配置为允许用户定义自计算机起的范围,所述控制功能将在所述范围内执行而在所述范围外不执行。

[0084] 6. 如技术方案 5 所述的计算机, 其中, 所述姿势识别组件还被配置为评估多普勒效应频移的振幅, 以确定在所述范围内是否出现用户运动。

[0085] 7. 如技术方案 5 所述的计算机, 其中, 所述范围自计算机起在三维空间中放射状延伸。

[0086] 8. 如技术方案 1 所述的计算机, 其中, 所述扬声器的其他功能是播放音乐, 并且其中, 给定频率的声音包括音乐的一部分。

[0087] 9. 如技术方案 1 所述的计算机, 其中, 所述姿势识别组件被配置为测试频率范围和评估相应的被捕捉声音, 以及当与所述给定频率相关的干扰比所述频率范围内的其他频率相对较小时, 选择所述给定频率。

[0088] 10. 如技术方案 9 所述的计算机, 其中, 所述干扰包括由所述扬声器产生的其他声音或背景噪声。

[0089] 11. 如技术方案 9 所述的计算机, 其中, 所述姿势识别组件被配置为根据一个或更多个参数来重新测试所述频率范围。

[0090] 12. 如技术方案 11 所述的计算机, 其中, 所述一个或更多个参数涉及预定义的时间段或涉及在超过预定义阈值的给定频率处检测到的干扰。

[0091] 13. 一种计算机可读存储介质, 其上存储有指令, 当所述指令由计算设备运行时使所述计算设备执行包括以下的动作:

[0092] 从计算机发射声音信号;

[0093] 在所述计算机处捕捉声音;

[0094] 将所发射的声音信号的频率范围与所捕捉的声音的同一频率范围相比较, 以检测由用户控制姿势所产生的多普勒频移;

[0095] 将所述用户控制姿势映射到控制功能; 以及

[0096] 在所述计算设备上执行所述控制功能。

[0097] 14. 如技术方案 13 所述的计算机可读存储介质, 其中, 所述发射包括从单个扬声器组同时发射在恒定频率处的声音信号和其他声音信号。

[0098] 15. 如技术方案 13 所述的计算机可读存储介质, 其中, 多个麦克风执行所述捕捉, 并且所述捕捉还包括比较所检测的、所述多个麦克风的多普勒频移的差异, 以进一步识别关于所述用户控制姿势的信息。

[0099] 16. 如技术方案 15 所述的计算机可读存储介质, 其中, 所述比较多个麦克风的多普勒频移的差异以进一步识别关于所述用户控制姿势的信息包括: 确定所述用户控制姿势是单手姿势还是双手姿势。

[0100] 17. 如技术方案 13 所述的计算机可读存储介质, 其中, 所述比较包括利用多普勒频移来区分所述用户控制姿势与产生其他多普勒频移的其他用户控制姿势。

[0101] 18. 一种计算机可读存储介质, 其上存储有指令, 当所述指令由计算设备运行时使所述计算设备执行包括以下的动作:

[0102] 使扬声器和麦克风执行第一功能;

[0103] 同时利用所述扬声器和所述麦克风执行第二功能, 所述第二功能包括:

[0104] 使用所述麦克风捕捉来自所述扬声器的声音信号;

[0105] 检测所述声音信号中的多普勒频移;

[0106] 将在计算机附近执行的用户控制姿势与多普勒频移相关联 ; 以及

[0107] 将所述用户控制姿势映射到控制功能。

[0108] 19. 如技术方案 18 所述的计算机可读存储介质, 其中, 所述第一功能包括电话功能。

[0109] 20. 如技术方案 18 所述的计算机可读存储介质, 其中, 所述控制功能与所述第一功能相关。

[0110] 结束语

[0111] 虽然使用特定于结构特征和 / 或方法步骤的语言描述了与基于多普勒的用户姿势检测有关的技术、方法、设备、系统等, 但应该理解, 所附权利要求书中定义的主题不一定限于所述的特定特征或步骤。更确切地, 公开了如实现所要求保护的方法、设备、系统等的示例性形式的特定功能和步骤。此外, 在本文中使用的术语“a”和“an”意在引入一种元件, 其可以是单数或多数。因此, “a”和“an”可以用来涵盖单数意义(例如, 一个)和多数意义(例如, 多个)。

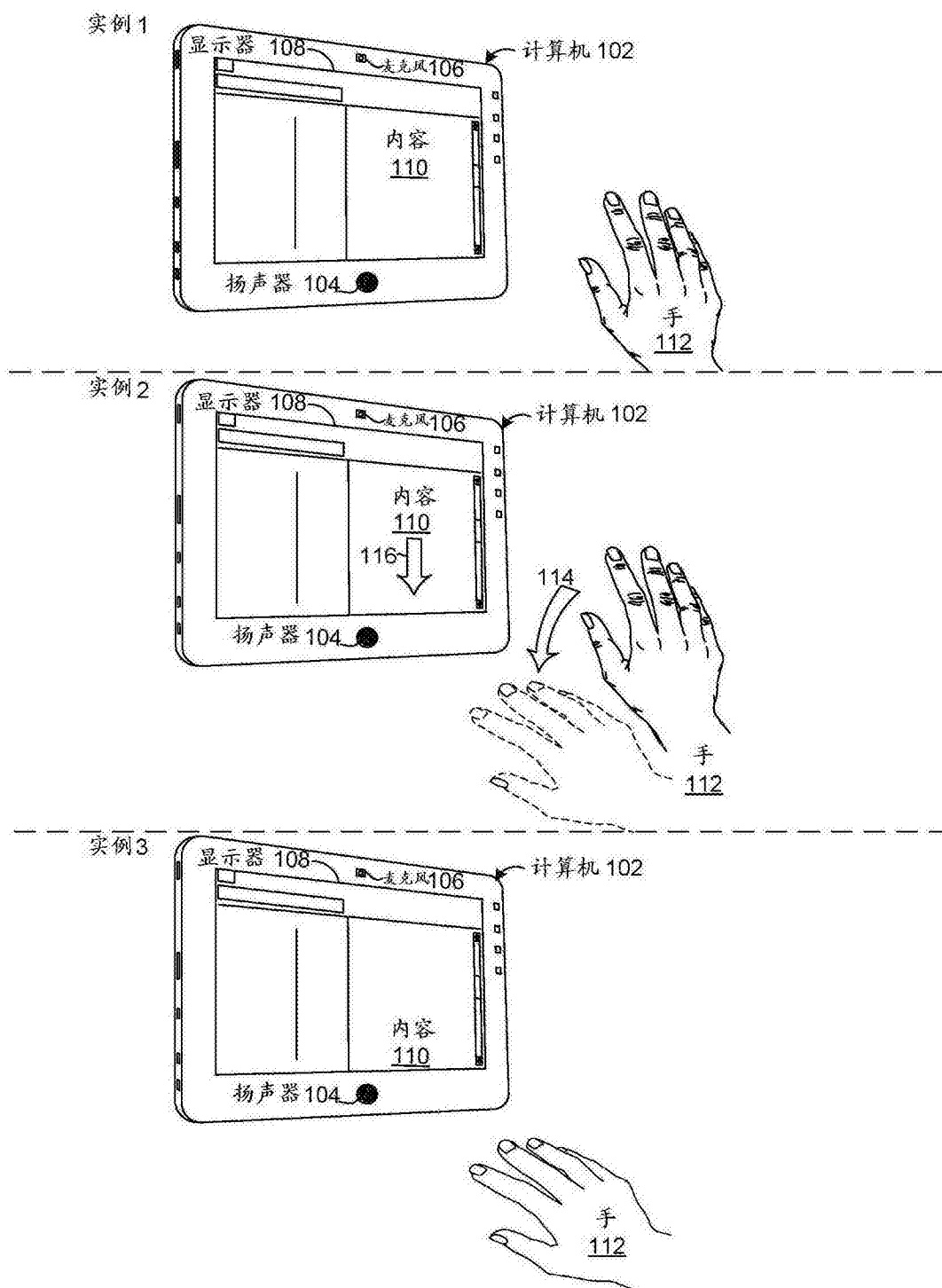


图 1

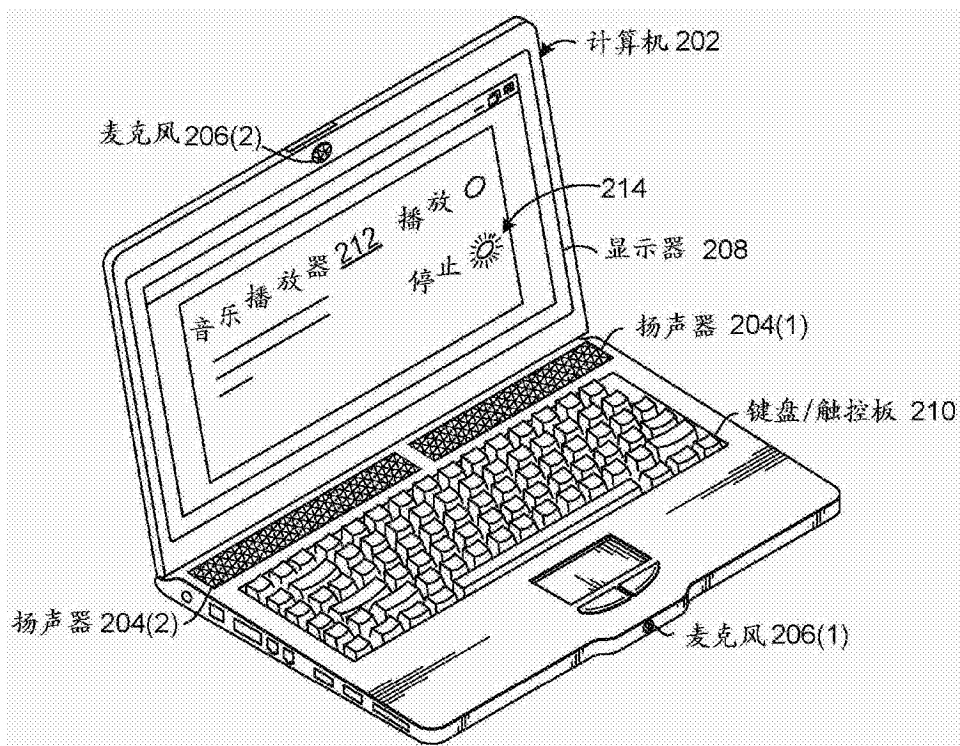


图 2

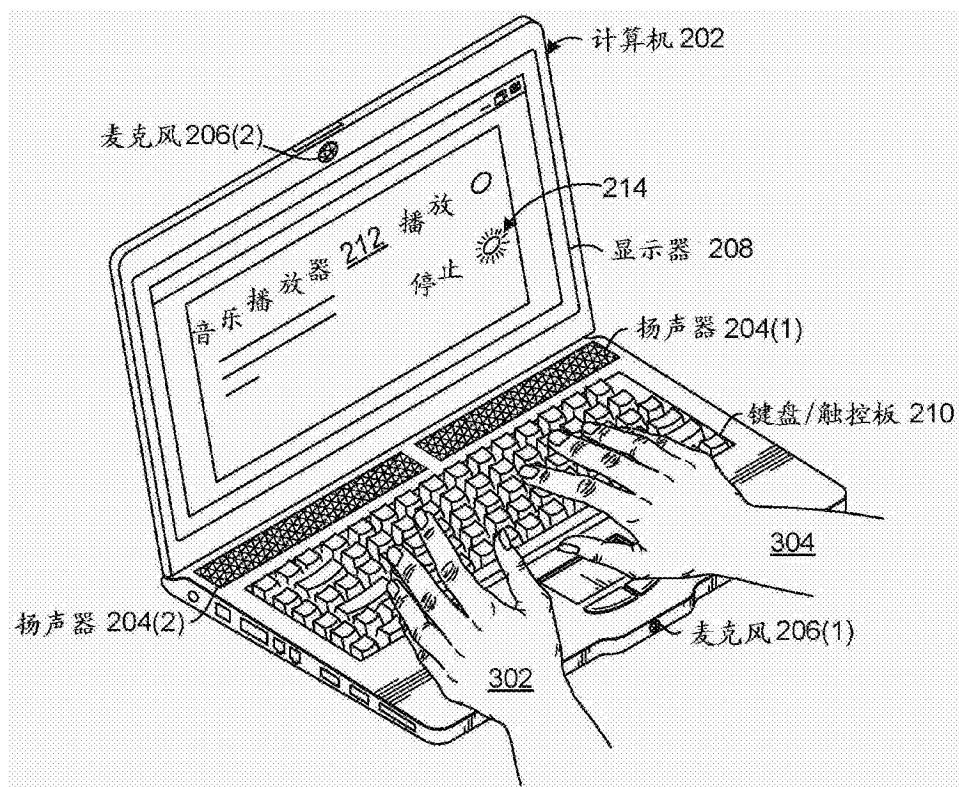


图 3



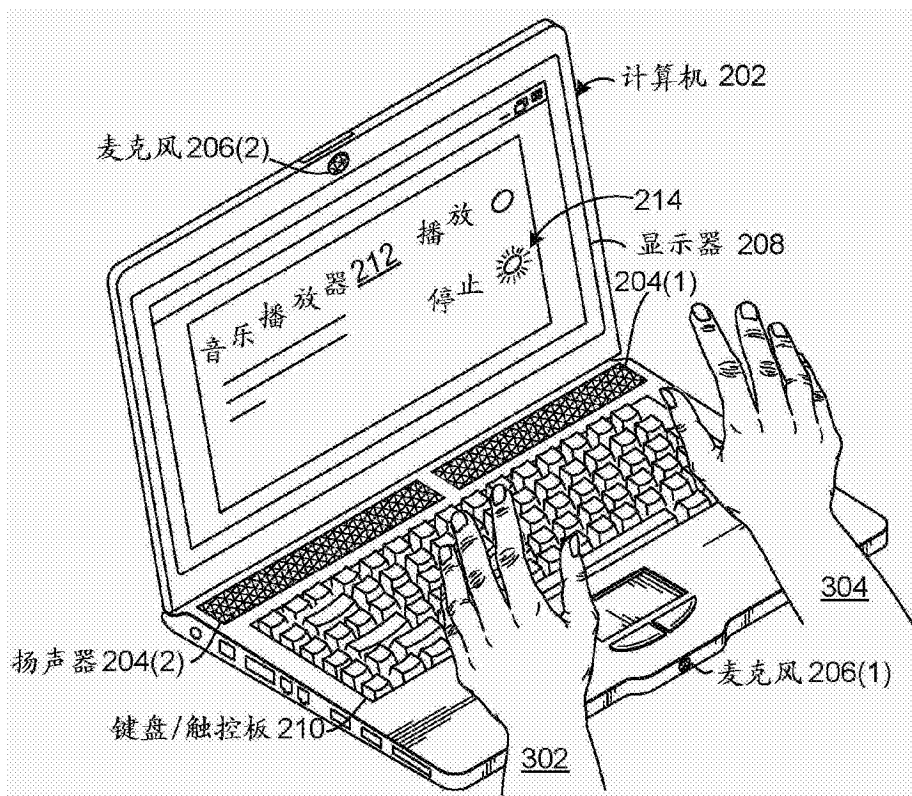


图 4

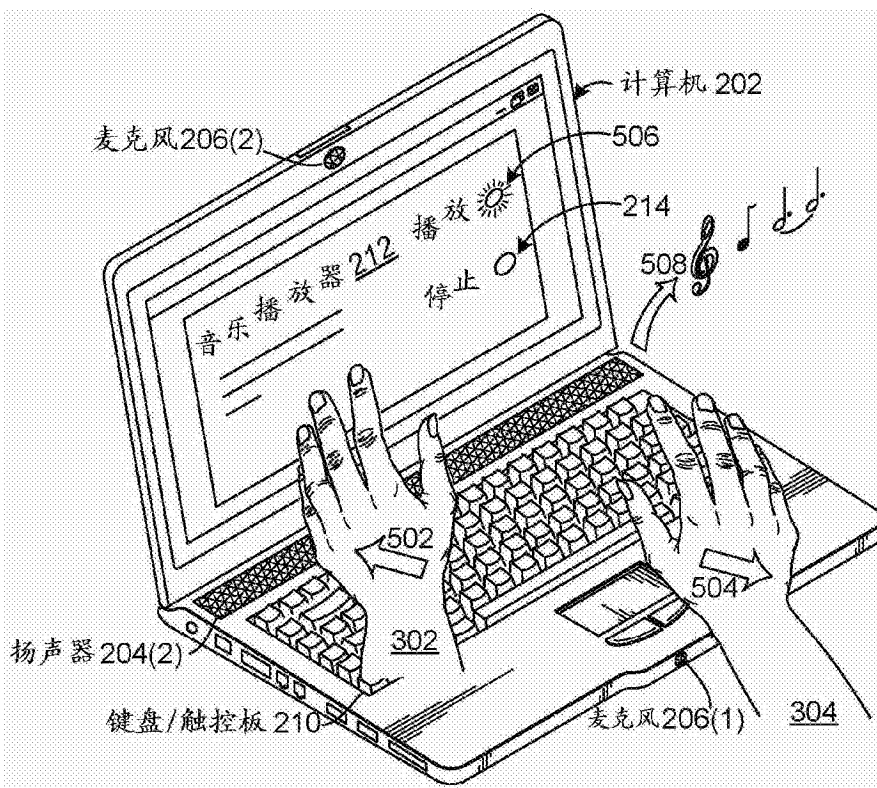


图 5

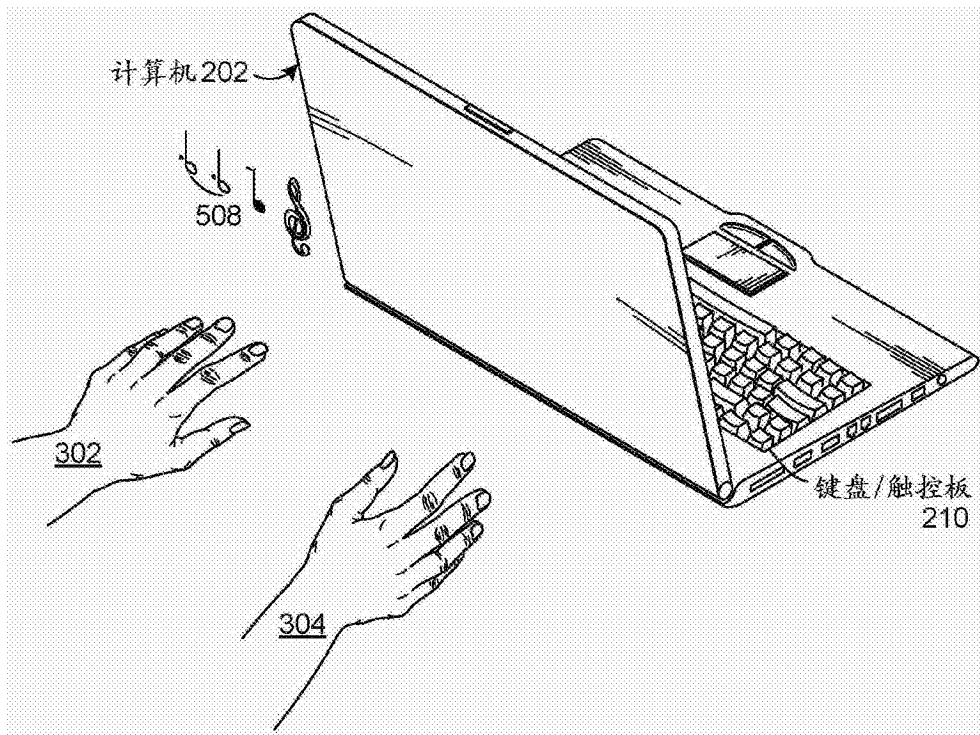


图 6

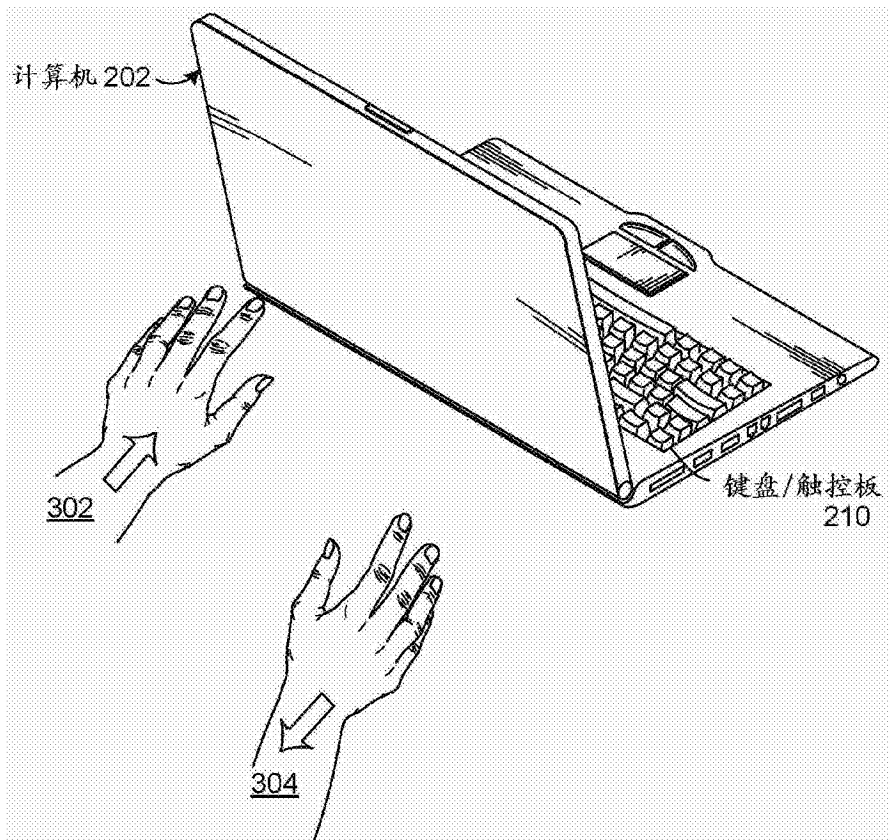


图 7

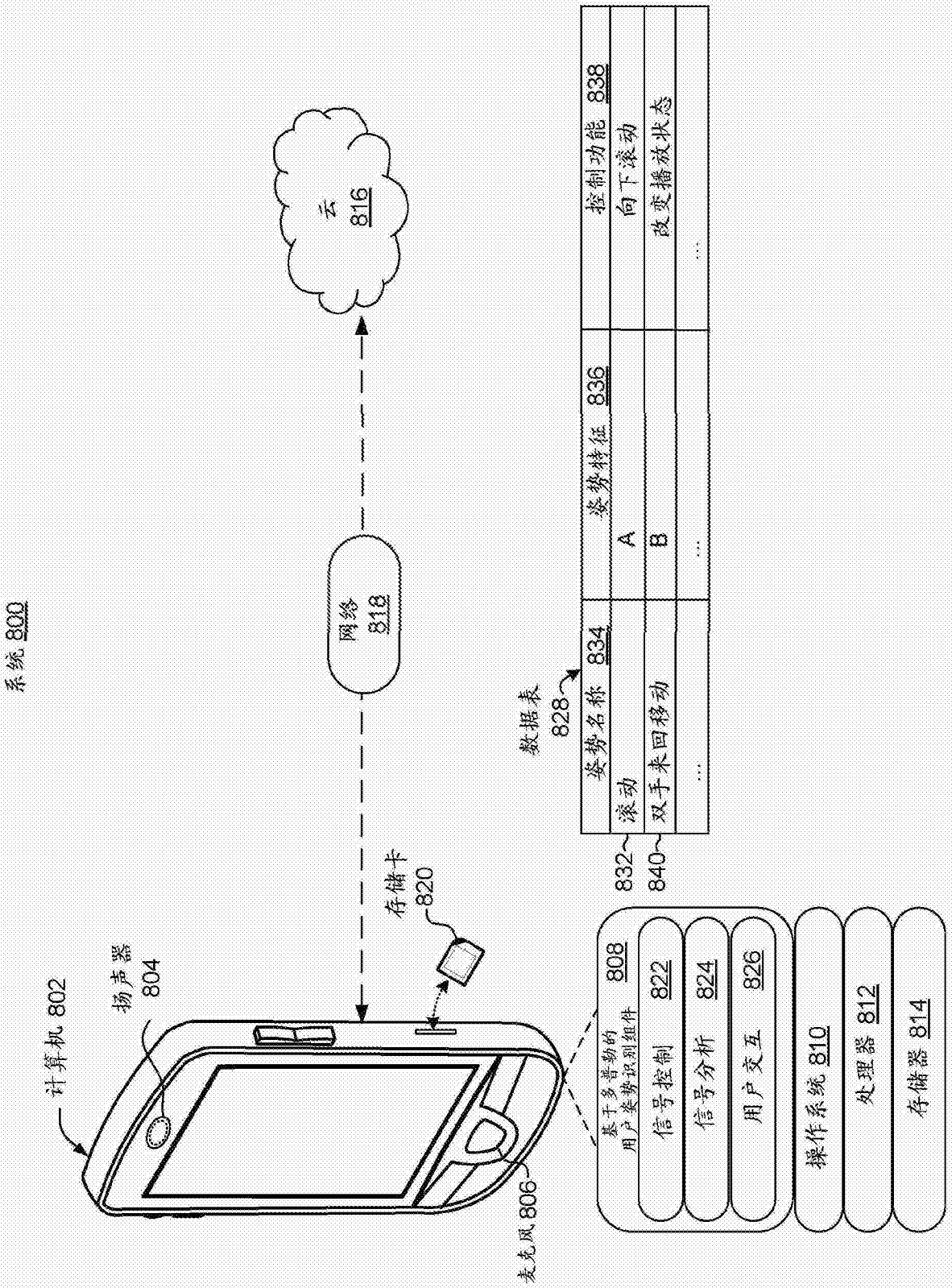


图 8

## 方法 900

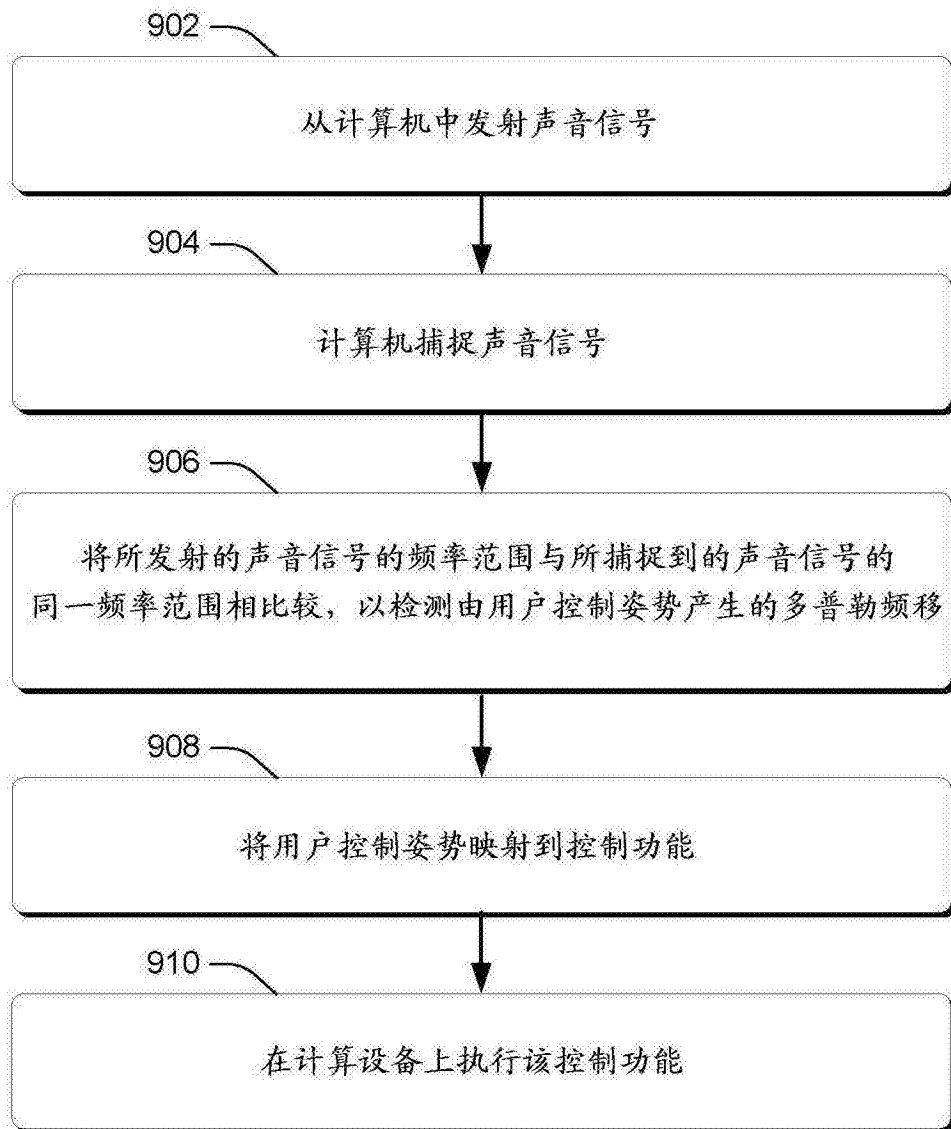


图 9