



(12) 发明专利申请

(10) 申请公布号 CN 104469491 A

(43) 申请公布日 2015. 03. 25

(21) 申请号 201410454355. 8

(22) 申请日 2014. 09. 09

(30) 优先权数据

14/026, 154 2013. 09. 13 US

(71) 申请人 索尼公司

地址 日本东京

(72) 发明人 P·辛塔尼 F·祖斯塔克

(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 郑宗玉

(51) Int. Cl.

H04N 21/442(2011. 01)

H04N 21/439(2011. 01)

H04N 21/466(2011. 01)

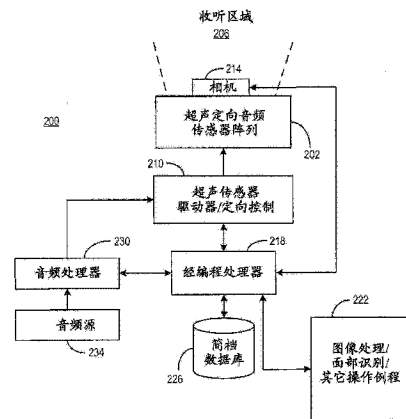
权利要求书2页 说明书13页 附图6页

(54) 发明名称

音频传递方法和系统

(57) 摘要

一种音频传递方法和系统。拍摄收听区域的图像并且处理图像以定位听众在房间中的位置。检索与听众关联的存储的听众简档并且基于听众的简档创建音频特性。将定向音频束导向听众的耳朵并且调节定向束以追踪听众的移动。此摘要将不被认为是限制性的,因为其它实施例可以偏离该摘要中描述的特征。



1. 一种音频传递方法,包括:
使用图像拍摄装置拍摄收听区域的图像;
在一个或多个经编程处理器处:
处理所述图像以定位听众在所述收听区域中的位置,
处理所述图像以识别所述收听区域中的所述听众的面部,
处理所述图像以定位所述听众的耳朵的位置,
检索与识别出的面部关联的存储的听众简档,
基于所述听众简档调节一个或多个音频特性,以及
控制定向音频束,以将所述定向音频束导向所述听众的耳朵;
使用所述图像拍摄装置拍摄所述听众的后续一系列图像;以及
在所述一个或多个经编程处理器处:
通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动,以及
按照所述听众在所述收听区域内的移动来调节所述定向音频束。
2. 根据权利要求1所述的方法,其中,所述定向音频束包括多频道音频节目的缩混,所述多频道音频节目包括多个频道。
3. 根据权利要求2所述的方法,其中,调节所述定向音频束的步骤包括改变所述多频道音频节目的混音。
4. 根据权利要求3所述的方法,其中,所述多频道音频节目包括中心频道节目,其中,改变所述多频道音频节目的混音包括增加对于移动到与所述定向音频束的源最接近的位置的所述听众的耳朵而言的中心频道节目的幅度。
5. 根据权利要求1所述的方法,其中,所述定向音频束包括超声音频束。
6. 根据权利要求1所述的方法,其中,控制的步骤包括控制伺服电机,所述伺服电机定位用万向架安装的超声传感器阵列。
7. 一种音频传递方法,包括:
使用图像拍摄装置拍摄收听区域的图像;
在一个或多个经编程处理器处:
处理所述图像以定位听众在所述收听区域中的位置,
处理所述图像以识别所述收听区域中的所述听众的面部,
处理所述图像以定位所述听众的左耳和右耳的位置,
检索与识别出的面部关联的存储的听众简档,
基于所述听众简档调节一个或多个音频特性,以及
控制左频道定向音频束和右频道定向音频束,以将所述左定向音频束和所述右定向音频束分别导向所述听众的左耳和右耳;
使用所述图像拍摄装置拍摄所述听众的后续一系列图像;以及
在所述一个或多个经编程处理器处:
通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动,以及
按照所述听众的左耳和右耳在所述收听区域内的移动,调节由所述左定向音频束和所

述右定向音频束携带的音频的混音。

8. 一种音频传递系统,包括:

图像拍摄装置,被配置为拍摄收听区域的图像;

处理所述图像以定位听众在所述收听区域中的位置的模块;

处理所述图像以识别所述收听区域中的所述听众的面部的模块;

处理所述图像以定位所述听众的耳朵的位置的模块;

检索与识别出的面部关联的存储的听众简档的模块;

基于所述听众简档调节一个或多个音频特性的模块;

控制定向音频束,以将所述定向音频束导向所述听众的耳朵的模块;

所述图像拍摄装置还被配置为拍摄所述听众的后续一系列图像;

通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动的模块;以及

按照所述听众在所述收听区域内的移动来调节所述定向音频束的模块。

9. 一种音频传递系统,包括:

图像拍摄装置,被配置为拍摄收听区域的图像;

处理所述图像以定位听众在所述收听区域中的位置的模块;

处理所述图像以识别所述收听区域中的所述听众的面部的模块;

处理所述图像以定位所述听众的左耳和右耳的位置的模块;

检索与识别出的面部关联的存储的听众简档的模块;

基于所述听众简档调节一个或多个音频特性的模块;

控制左定向音频束和右定向音频束,以将所述左定向音频束和所述右定向音频束分别导向所述听众的左耳和右耳的模块;

所述图像拍摄装置还被配置为拍摄所述听众的后续一系列图像;通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动的模块;以及

按照所述听众的左耳和右耳在所述收听区域内的移动来调节由所述左定向音频束和所述右定向音频束携带的音频的混音的模块。

10. 一种音频传递方法,包括:

在经编程处理器处检索并读取存储的听众简档,以确定与听众关联的音频特性设置;以及

在音频混音器处,所述经编程处理器基于存储的所述听众简档将多频道音频节目的频道的混音调节成数量相等或减少的频道。

音频传递方法和系统

[0001] 版权和商标声明

[0002] 本专利文献的一部分公开内容包含受到版权保护的内容。当专利文献或专利公开内容出现在专利和商标局专利文件或记录中时,版权拥有者对专利文献或专利公开内容的副本复制不提出异议,但以其它方式保留所有任何版权的权利。商标是它们各自拥有者的财产。

技术领域

[0003] 本发明涉及音频传递方法和系统,尤其涉及音频可达性。

背景技术

[0004] 美国的高级通信服务法案要求应对各种残疾,其中一种是听力残疾。该法案要求电视设备供应方采取措施来尝试改进向具有听力残疾的人呈现音频。

发明内容

[0005] 根据本发明的一个实施例,提供了一种音频传递方法,包括:使用图像拍摄装置拍摄收听区域的图像;在一个或多个经编程处理器处:处理所述图像以定位听众在所述收听区域中的位置,处理所述图像以识别所述收听区域中的所述听众的面部,处理所述图像以定位所述听众的耳朵的位置,检索与识别出的面部关联的存储的听众简档,基于所述听众简档调节一个或多个音频特性,以及控制定向音频束,以将所述定向音频束导向所述听众的耳朵;使用所述图像拍摄装置拍摄所述听众的后续一系列图像;以及在所述一个或多个经编程处理器处:通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动,以及按照所述听众在所述收听区域内的移动来调节所述定向音频束。

[0006] 根据本发明的另一个实施例,提供了一种音频传递方法,包括:使用图像拍摄装置拍摄收听区域的图像;在一个或多个经编程处理器处:处理所述图像以定位听众在所述收听区域中的位置,处理所述图像以识别所述收听区域中的所述听众的面部,处理所述图像以定位所述听众的左耳和右耳的位置,检索与识别出的面部关联的存储的听众简档,基于所述听众简档调节一个或多个音频特性,以及控制左频道定向音频束和右频道定向音频束,以将所述左定向音频束和所述右定向音频束分别导向所述听众的左耳和右耳;使用所述图像拍摄装置拍摄所述听众的后续一系列图像;以及在所述一个或多个经编程处理器处:通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动,以及按照所述听众的左耳和右耳在所述收听区域内的移动,调节由所述左定向音频束和所述右定向音频束携带的音频的混音。

[0007] 根据本发明的另一个实施例,提供了一种音频传递系统,包括:图像拍摄装置,被配置为拍摄收听区域的图像;处理所述图像以定位听众在所述收听区域中的位置的模块;处理所述图像以识别所述收听区域中的所述听众的面部的模块;处理所述图像以定位所述听众的耳朵的位置的模块;检索与识别出的面部关联的存储的听众简档的模块;基于所述

听众简档调节一个或多个音频特性的模块；控制定向音频束，以将所述定向音频束导向所述听众的耳朵的模块；所述图像拍摄装置还被配置为拍摄所述听众的后续一系列图像，通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动的模块，以及按照所述听众在所述收听区域内的移动来调节所述定向音频束的模块。

[0008] 根据本发明的另一个实施例，提供了一种音频传递系统，包括：图像拍摄装置，被配置为拍摄收听区域的图像；处理所述图像以定位听众在所述收听区域中的位置的模块；处理所述图像以识别所述收听区域中的所述听众的面部的模块；处理所述图像以定位所述听众的左耳和右耳的位置的模块；检索与识别出的面部关联的存储的听众简档的模块；基于所述听众简档调节一个或多个音频特性的模块；控制左定向音频束和右定向音频束，以将所述左定向音频束和所述右定向音频束分别导向所述听众的左耳和右耳的模块；所述图像拍摄装置还被配置为拍摄所述听众的后续一系列图像；通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动的模块；以及按照所述听众的左耳和右耳在所述收听区域内的移动来调节由所述左定向音频束和所述右定向音频束携带的音频的混音的模块。

[0009] 根据本发明的另一个实施例，提供了一种音频传递方法，包括：在经编程处理器处检索并读取存储的听众简档，以确定与听众关联的音频特性设置；以及在音频混音器处，所述经编程处理器基于存储的所述听众简档将多频道音频节目的频道的混音调节成数量相等或减少的频道。

附图说明

[0010] 通过参照随后结合附图的详细描述，可以最佳地理解例示操作的构造和方法的某些示例性实施例以及目的和优点，在附图中：

[0011] 图 1 是符合本发明的某些实施例的电视音频系统的示例。

[0012] 图 2 是符合本发明的某些实施例的听众简档的示例实现方式。

[0013] 图 3 由图 3A、图 3B 和图 3C 组成，示出符合本发明的某些实施例的定向音频系统中听众的头部转动带来的影响的示例。

[0014] 图 4 是示出符合本发明的某些实施例的操作的方法的流程图示例。

[0015] 图 5 是以符合本发明的某些实施例的方式调节音频的方法的流程图示例。

[0016] 图 6 是符合本发明的某些实施例的定向音频系统的框图表示的示例。

[0017] 图 7 是以符合本发明的某些实施例的方式将超声音频阵列导向某个位置的布置的示例。

具体实施方式

[0018] 虽然本发明容许有许多不同形式的实施例，但是在附图中示出并且将在本文中详细描述具体实施例，要理解的是，这些实施例的当前公开内容将被视为原理的示例，而不是旨在将本发明限于所示出和描述的具体实施例。在下面的描述中，使用相同的参考标号描述附图的几个视图中的相同、类似或对应的部件。

[0019] 如本文中所使用的，术语“a”或“an”被定义为一个或一个以上。如本文中所使用的，术语“多个”被定义为两个或两个以上。如本文中所使用的，术语“其它”被定义为至少

两个或更多个。如本文中所使用的,术语“包括”和/或“具有”被定义为包括(即,开放式语言)。如本文中所使用的,术语“耦合”被定义为连接,尽管不一定是直接连接,并且不一定是机械连接。如本文中所使用的,术语“程序”或“计算机程序”或类似术语被定义为被设计为在计算机系统上执行的指令序列。“程序”或“计算机程序”可以包括子程序、函数、过程、应用、对象方法、对象实现、可执行应用、小应用程序、小服务程序、源代码、对象代码、脚本、程序模块、共享程序库/动态加载库和/或被设计为在计算机系统上执行的其它指令序列。如本文中所使用的,术语“电视接收器装置”或类似术语旨在涵盖任何电视接收器,包括电视机、机顶盒(STB)或被构造用于接收电视节目的其它装置。“显示器”或类似术语可以形成电视装置或计算机系统上的能够接收包括音频的内容的部分。符合本文中的教导的装置可以被例示为STB、独立式条形音箱、或外部附加音频装置、或具有音频能力但没有调谐器的监视器以及其它实现形式。

[0020] 在整个该文献中,对“一个实施例”、“某些实施例”、“实施例”、“实现方式”、“示例”或类似术语的引用意指结合该实施例描述的特定特征、结构或特性被包括在本发明的至少一个实施例中。因此,在整个该说明书中的各种位置出现这类短语不一定都是指同一实施例。此外,特定特征、结构或特性可以按任何合适的方式组合在一个或多个实施例中而没有限制。

[0021] 如本文中所使用的,术语“或”将被解释为包括式的,或者意指任一个或任意组合。因此,“A、B或C”意指“下面中的任一种:A;B;C;A和B;A和C;B和C;A、B和C”。只有当元件、功能、步骤或动作的组合以某种方式在本质上是互斥的时候,才会出现该定义的例外。

[0022] 术语“音频特性”将被解释为意指电子音频信号中可以调节的属性,包括但不限于音量、均衡化、压缩、房间模拟、频道混合等。

[0023] 如之前所述的,美国的高级通信服务法案要求应对各种残疾,其中一种是听力残疾。该法案要求电视设备供应方采取措施来尝试改进向具有听力残疾的人呈现音频。

[0024] 注意,对于不同人而言,听力残疾的差异很大,并且听力残疾通常是不对称的。听力丧失可能限于一只耳朵,或者一只耳朵中的听力丧失或多或少比另一只耳朵中的严重。另外,对于不同人而言,甚至对于同一个人的不同耳朵而言,受影响的频率存在差异。当具有不同听力能力的多个人在同一电视观看区域中时,这种听力残疾会造成困难。这可导致电视音频主要被调节成应对听力最差的人的听力,而这对于其它听众而言会是不舒服的大声。

[0025] 可以使用超声技术使音频信号具有高度方向性,在超声技术中,使用小超声传感器的阵列发送极具方向性的超声束。该高度方向性主要是由于使传感器接近所发送的超声信号的波长造成的。通过向着听众的耳朵发送两个超声信号,音频可以被编码成这两个信号之间的频率差。由于空气和耳朵中的非线性,导致出现这两个超声信号的混合,从而导致和信号和差信号。差信号表示原始编码音频并且可以被听众听到。通过将这样的两组束导向听众的左耳和右耳,可以实现立体音频编程。

[0026] 可以有利地利用该机制对听力受损的人听到音频起到改善作用。一般地,例如,当观看电视(TV)时,对于听力受损的人,需要高音量水平才能够欣赏电视节目。遗憾的是,这会以牺牲听力没有受损并且更喜欢较低音量水平的其它听众为代价。

[0027] 因此,向听众传递音频可以是按照个体的听力特性并且结合超声传递来量身定制

的,个体化的音频可以被导向个体。此外,可以使用图像识别通过相机来识别个体,然后将量身定制的声音导向到识别出的个体。可以按若干种方式进行声音的瞄准。可以使用相控阵传感器,但是利用这个方法存在一些限制,例如方向性的粒度(角度),还有可以同时被作为服务对象的听众的数量。

[0028] 优选的方法是使用如上所述的个体化声音的超声传递。声音被频率转变到超声范围,例如大约 40 kHz。然后,超声声音与其它超声声音撞击,从而导致和、差和基谐波。听众只听到差信号。由于超声声音的波长是传感器的尺寸的可预见部分,因此这导致声音的传递极具方向性。这使得能够将声音导向个别的接收方。

[0029] 为了瞄准声音,一种技术是具有若干可调区域,若干可调区域可以是固定的或预设的。听众通常坐在离散的固定位置,这些位置通常由房间中的椅子或沙发的相对固定位置决定。因此,一旦就坐,仅将必须识别听众并且需要确定该听众在预设位置之中的位置。如果用户手动地标识他自己,则对听众的识别将被简化,或者当使用诸如可识别的 RFID、Bluetooth(蓝牙)、拥有遥控器或许多遥控器之一、拥有蜂窝电话等技术时,对听众的识别会更复杂。在优选的实现方式中,使用相机或其它图像拍摄装置利用面部识别和存储的听众简档来定位和识别听众,并且在空间上表征每个听众。

[0030] 现在转到图 1,考虑符合某些实施例的非限制性的示例电视系统实现方式,在该实现方式中,使用超声音频在若干听众之间隔离音频。在这个例示中,显示器 20 或诸如电视接收器装置的其它装置(STB、外部音频处理装置等)具有对收听区域 28 进行成像的集成相机 24。与显示器 20 关联或者与显示器 20 集成一体的音频系统采用超声传感器的阵列 32,可以采用该阵列 32 使用以上讨论的超声技术将目标明确的定向音频束导向诸如听众 36 和听众 40 的一个或多个听众。在某些实现方式中,这些听众 36 和 40 可以是电视机的经常的观众,因此常常出现在收听区域 28 中。

[0031] 为了定制每个听众的音频体验,可以为每个听众创建简档(profile),并且可以为未辨认出的听众提供默认简档或客人简档。通过对收听区域进行成像,相机 24 可以用于提供图像,在分析图像时,可以确定 1) 每个听众的位置,2) 每个听众的头部和耳朵的位置,3) 辨认每个注册的和有简档的听众,或者将听众归类于客人,4) 追踪听众的移动,5) 记录对于听众的收听体验有意义的移动,6) 针对听众简档中阐明的听众喜好或听力能力,量身定制音频节目。以此方式,如果听众 36 具有正常的听力而听众 40 的听力能力受损,则可以根据听众的需要和喜好单独地对待每个听众,使得对其他听众的影响最小。在另一个实施例中,简档中可以包括优选的语言,因此可以提供多种语言。可以使用各种音频语言子频道来适应优选与主音频频道中提供的语言不同的语言的听众、或者优选在安装期间指示的默认语言的听众。在另一个实施例中,单词取代机可以选择性为那些被识别出的并且与父母控制限制或制约关联的特定听众替换有异议的单词或短语。

[0032] 举例来说,而非限制地,考虑如图 2 中所示的电视系统中的实现方式和名字叫“乔治”的听众的简档画面 50。在这个示例简档画面(例如,从电视的菜单系统中调取的)中,听众可以提供用于参考的图像 52 并且可以在 56 选择优选的可以使用的语言,该语言是选自可用的音频语言子频道,如果可能的话。还注意的是,听众简档可以是包括未明确示出的其它偏好、特性和/约束的更大用户简档的一部分。当在拍摄收听区域的图像时,电视机的相机 24 可以使用这个图像作为用于面部识别的参考,以从简档 50 中检索乔治的音频特

性。在这个示例中,乔治的右耳的听力比左耳的差,并且这个被反映在音量设置 60 中,在音量设置 60 中,右耳音量是满的而左耳音量是大约一半。另外,在 64 处表明,与如 68 中所示的难以听见较高频率的右耳相比,左耳具有听见低频、中频和高频的平衡能力。在这个示例中,可假定具有正常听力的人在音量处于较低水平(例如,大约 25%)的情况下具有接近平坦的频率均衡。

[0033] 使用这个简档作为模板,音频系统可以向乔治播送专用音频信号,在该信号中,右频道音量十分高并且左音量高于正常音量。另外,将调节右频道中的音频,使得与低频相比,在中频和高频下提供更大的音量。可以在音频系统的辅助下或者基于听众的偏好,用实验方法建立此简档。在一个实施例中,音频设置将通过测试听众的听力并且按照听众对音频设置协议的响应修改音频特性来引导用户设置人物的简档。在这种实现方式的示例中,可以产生测试音调,并且用户可以进行响应以确定特定用户在什么水平可以听见特定频率范围。因此,用户可以手动地调节均衡化以提高他或她的听见的的能力,或者音频系统可以推导出用在简档中的合适的均衡化。

[0034] 在另一个示例实现方式中,单词或短语可以在以听得见的形式播放(例如,每个频道播放一次)的同时显示在显示器上,并且询问用户关于理解所显示的口述的单词或短语的能力。例如,由于大多数听力问题开始于听见高频分量的能力的下降,因此,可以播放具有明显高频内容的诸如“spoon(勺子)”、“ship(船)”、“ticket(票)”等单词,并且用户可以指示特定 Q、均衡化、滤波和导致能被最佳理解的平衡、和 / 或右侧和左侧的听力相等。该系统可以让每个用户经历培训过程,在该培训过程中,系统地改变滤波特性并且每个用户可以协助将以能被最佳理解的方式听见说话的能力最佳化。一旦创建了简档的数据,就可以使用按钮 74 保存简档,或者作为退出简档并进行保存的自动化设置过程的一部分保存简档,或者听众可以通过使用按钮 78 退出而不进行保存,该按钮 78 将简档恢复回到先前的设置,或者如果之前没有创建简档则恢复回到无简档。

[0035] 在这个示例中,假定将用立体声向听众播送音频节目,但是这不应被视为是限制,因为音频可以按单声道形式平等地播送,对音频束的方向性和准确性的要求较低。此外,尽管可以向左耳和右耳播送音频,但并没有要求超声音频束中没有重叠。

[0036] 注意,当在传统立体声音频系统中用立体声传递环境声音时,立体混音通常是在多频道音频节目中从较大数量的频道中推导出的混音。例如,5.1 频道音频系统具有中心频道、左前频道、右前频道、左后频道、右后频道和超低音频道。在这种多频道音频混音中,一般地,中心频道携带的是正在观看的电视节目或电影中的对话(话语)的整体。类似地,在超低音频道等中处理低频。当这被混音成立体声时,中心频道对话一般被分到左频道和右频道中。由于只有一个或两个频道最常用于电视和其它音频再现,因此可以调节从多频道音频到较少数量频道的音频信号的缩混,以实现对于听力受损的那些人更期望的收听体验。

[0037] 例如,如果听众在存在其它声音时辨认话语的能力受损,则可能有利的是,基于听众的简档,向该听众提供更高水平的中心频道混音。因此,符合某些实施例的音频传递方法采用经编程处理器来检索和读取存储的听众简档,以确定与听众关联的音频特性设置;并且在混音器处,经编程处理器可以基于存储的听众简档,调节多频道音频节目的频道到数量减少的频道的混音,以改善听众的收听体验。

[0038] 现在参照由图 3A、图 3B 和图 3C 组成的图 3,当考虑本教导时,应该理解,当用定向束向听众发送音频时,会出现其它问题。在图 3A 中,当听众 90 被设置成使得左音频束和右音频束(被示出为 L 和 R)容易直接以双耳为目标时,听众将以所期望方式听到立体声音频。但是,当听众 90 如图 3B 中所示转动他的头部时,针对左耳的音频节目将变得比针对右耳的音频节目更突出。进一步以此为例,考虑图 3C,在图 3C 中,右耳完全被头部挡住(如代表右耳束的虚线所指示的),而左耳束容易以左耳为目标。在这种情形下,束的方向性和左音频和右音频的立体声分离可针对听众 90 的缺点进行工作。在这种情况下,当音频因听众头部的运动而丢失或减弱时,通常最佳的是,电视节目或电影对话没有丢失。因此,以与符合本文中的教导的方式,当目标听众移动时(尤其是移动他的头部时),相机 24 通过连续拍摄听众的图像来跟踪这些移动。当系统检测到移动将中断听众的收听体验时,可以改变原始多频道节目内容的缩混,或者可以调节立体声音频的混音。

[0039] 举例来说而非限制地,当检测到头部位置从图 3A 中示出的位置移动到图 3C 的位置时,可以在经编程处理器的控制下自动地操纵混音,以将右频道音频转移到左频道。在另一个实施例中,当头部进行相同移动时,可以在经编程处理器的控制下自动地操纵混音,以将中心频道混音转移到左频道,使得听众最有可能不丢失对话。在各情况下,当处理器调节音频混音时,参考听众的听力简档,使得在以上示例中,如果听众 90 是乔治,如果右频道信息被转移到左频道,则根据左耳和右耳之间的整体听力的差异,可以减小音量,并且类似地调节正常将在右耳中的被发送到左耳的音频的频率均衡,例如,以减小高频含量。在其它实施例中,可以操纵各种频道的混音,以增强听众的体验。例如,如果人的听力使得左耳的话语可懂度差而右耳的话语可懂度好,则可以基于简档信息主要把话语混音到右耳。可以通过改变从较大数量的频道开始的缩混、或者通过仅仅转变左右之间的混音以实现立体声分离(接近或变成单耳的)的减小、或者通过期望的任何其它方式来操纵混音。本领域的技术人员在考虑到本教导的情况下会想到许多其它变形形式。

[0040] 还要注意的,当人具有听力困难时,听力受损的听众经常有旋转他的头部使得最佳的耳朵面对音频源的几乎自动的动作。因此,混音或其它音频特性的当前改变符合利用这种一般人反应的改进。

[0041] 现在参照图 4,示出在 104 开始的一个实现方式示例的流程图 100。在 108,音频系统确定系统是否已经被构造成使用与听众简档关联的定向音频的播送。如果不是这样,则在 112 系统可以恢复到带有传统扬声器的更传统的音频系统。如果是这样,则在 116 拍摄收听区域的一个或多个图像,并且在 120 使用图像分析程序分析图像以试图识别听众及其位置。在图像分析中,人被识别,然后启用面部识别算法,试图识别已经存储了带有听众的音频特性的简档的人。对于辨认出的听众,在简档数据库中检索他们的简档,而对于未辨认出的听众,在 124 检索默认简档或客人简档。然后,在 128,基于听众的简档及其位置,调节音频特性。如之前所讨论的,可以根据听众的耳朵布置调节混音和其它音频特性。

[0042] 一旦加载了音频简档,则在 132 音频被定向播送到在收听区域内的听众物理位置处的辨认出的听众。类似地,在 136,未辨认出的听众使用默认简档或客人简档同时接收在收听区域内的听众物理位置处的定向音频束。为了保持对听众物理位置的连续追踪并且还为了监控他们的头部位置(如果用以上讨论的方式利用),在 140,通过开始重复处理来连续地更新处理,此后,处理行进回到 108。虽然在这个示例处理 100 中没有明确示出,但如果

没有新的听众进入收听区域,则可以跳过框 124。

[0043] 可以用包括如图 5 的 128 所示的示例处理的各种方式来实现处理 100 的功能 128。在这个示例处理实现方式中,在 150,接收多频道音频(例如,立体声、5.1 环绕声、7.1 环绕声等)。在 154,设置每个听众的左耳位置和右耳位置。如果在 158 像图 3 中一样容易以左耳和右耳为目标(使左耳和右耳平衡),则在 162,服从特定听众简档的频道的正常混音被呈现为被分派到听众的音频频道束。但是,如果听众的头部被设置成使得系统确定对于一只耳朵或另一只耳朵的播送将劣化,则在 166 系统确定哪只耳朵最接近定向声音源。然后,在 170 对音频进行重新混音。在这个示例中,该重新混音为最接近定向音频源的耳朵赋予包含对话的频道(例如,中心频道)的更重权重。在其它实现方式中,如果双耳都可以仍然至少部分地接收声音束,则可以增加到达离定向声音源最远的耳朵的音频的音量,从而提供连续立体声体验,直到该系统认为无法依赖到达离定向声音源最远的耳朵的播送以正确地接收声音束为止。在这种情况下,混音可以被转换成单耳的,或者说,对话频道转移到最接近定向声音源的耳朵,或者可以实现其它合适的混音和重新均衡。在任何情况下,在 162 和 170 二者,对于每个听众,处理在 174 返回以完成处理 128。本领域的技术人员在考虑到本教导的情况下会想到许多其它的变形形式。

[0044] 符合某些实现方式的示例系统被示出为图 6 的系统 200。诸如超声传感器 202 的定向音频传感器阵列一般被导向收听区域 206 并且被传感器驱动器和定向控制器 210 驱动。框 210 用于以按照以上讨论的方式向着听众产生定向音频束的方式来驱动超声传感器阵列 202。在经编程处理器 218 的控制下通过使用相机 214 来定位和识别听众,处理器 128 被编程,以在来自存储在非暂态存储介质中并且被示出为 222 的程序指令的程序控制下,执行如之前讨论的用于位置识别和用于面部识别的图像处理。

[0045] 如之前讨论地处理拍摄图像,以识别和定位收听区域 206 中的人。然后,执行 222 的面部识别算法,以将找到的面部与简档数据库 226 中的面部进行比较。当在简档数据库 226 中识别出听众时,经编程处理器(或多个处理器)218 使用简档数据执行音频处理器 230 内的混音和均衡功能,使得来自音频源 234 的音频被调节以按照听众的简档补偿听众的听力。

[0046] 连续更新这个处理,以按照以上讨论的方式识别各种听众的移动并且保持到达每个听众的合适的音频束或多个音频束。

[0047] 可以按照任何操作方式执行音频束的定向。例如,如图 7 中所示,多个超声传感器阵列可以被安装成允许进行至少水平旋转的万向架安装布置,但是优选地允许进行水平方向和垂直方向旋转二者上的二维运动,以便允许超声传感器阵列 250 以收听区域 206 内的大范围的位置为目标。在运行伺服控制算法的经编程处理器 218 的控制下,调节万向架安装,以通过使用伺服控制器 254 驱动用万向架安装的超声传感器阵列 250 来合适地以听众为目标。提供多个这种布置,以便能够在收听区域 206 内在任何给定时间以多个听众为目标。本领域的技术人员应该理解,在考虑到本教导的情况下,为了将定向音频束定向到观众,还可以提供其它布置。

[0048] 因此,按照某些实现方式,一种音频传递方法包括:使用图像拍摄装置拍摄收听区域的图像;在一个或多个经编程处理器处:处理所述图像以定位听众在所述收听区域中的位置,处理所述图像以识别所述收听区域中的所述听众的面部,处理所述图像以定位所述

听众的耳朵的位置,检索与识别出的面部关联的存储的听众简档,基于所述听众简档调节一个或多个音频特性,以及控制定向音频束,以将所述定向音频束导向所述听众的耳朵;使用所述图像拍摄装置拍摄所述听众的后续一系列图像;以及在所述一个或多个经编程处理器处:通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动,以及按照所述听众在所述收听区域内的移动来调节所述定向音频束。

[0049] 在某些实现方式中,所述定向音频束包括多频道音频节目的缩混,所述多频道音频节目包括多个频道。在某些实现方式中,调节所述定向音频束的步骤包括改变所述多频道音频节目的混音。在某些实现方式中,所述多频道音频节目包括中心频道,并且其中,多频道的混音包括增加对于移动到与所述定向音频束的源最接近的位置的所述听众的耳朵而言的中心频道节目的幅度。在某些实现方式中,所述定向音频束包括超声音频束。在某些实现方式中,所述图像拍摄装置包括集成到电视接收器装置中的相机。在某些实现方式中,所述图像拍摄装置包括集成到电子显示装置中的相机。在某些实现方式中,控制的步骤包括控制伺服电机,所述伺服电机定位用万向架安装的超声传感器阵列。

[0050] 另一种音频传递方法包括:使用图像拍摄装置拍摄收听区域的图像。在一个或多个经编程处理器处:通过以下步骤进行处理:处理所述图像以定位听众在所述收听区域中的位置,处理所述图像以识别所述收听区域中的所述听众的面部,处理所述图像以定位所述听众的左耳和右耳的位置,检索与识别出的面部关联的存储的听众简档,基于所述听众简档调节一个或多个音频特性,以及控制左频道定向音频束和右频道定向音频束,以将所述左定向音频束和所述右定向音频束分别导向所述听众的左耳和右耳;使用所述图像拍摄装置拍摄所述听众的后续一系列图像。在所述一个或多个经编程处理器处,所述处理还包括:通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动,以及按照所述听众的左耳和右耳在所述收听区域内的移动,调节由所述左定向音频束和所述右定向音频束携带的音频的混音。

[0051] 在某些实现方式中,所述左定向音频束和所述右定向音频束包括多频道音频节目的立体声缩混,所述多频道音频节目包括中心频道。在某些实现方式中,调节频道的混音包括增加对于所述听众的右耳或左耳中的任一个而言的中心频道节目的幅度,以便增加对于移动到与所述定向音频束的源最接近的位置的所述听众的右耳或左耳中的一个而言的所述中心频道节目的幅度。在某些实现方式中,所述定向音频束包括超声音频束。在某些实现方式中,所述图像拍摄装置包括集成到电视接收器装置中的相机。在某些实现方式中,所述图像拍摄装置包括集成到电子显示装置中的相机。在某些实现方式中,控制的步骤包括控制伺服电机,所述伺服电机定位用万向架安装的超声传感器阵列。

[0052] 音频传递系统的另一个示例具有图像拍摄装置,该图像拍摄装置被配置为拍摄收听区域的图像。一个或多个经编程处理器被编程为:处理所述图像以定位听众在所述收听区域中的位置,处理所述图像以识别所述收听区域中的所述听众的面部,处理所述图像以定位所述听众的耳朵的位置,检索与识别出的面部关联的存储的听众简档,基于所述听众简档调节一个或多个音频特性,以及控制定向音频束,以将所述定向音频束导向所述听众的耳朵。所述图像拍摄装置还被配置为拍摄所述听众的后续一系列图像;以及所述一个或多个经编程处理器还被编程为:通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动,以及按照所述听众在所述收听区域内的移动来调节所述定向

音频束。

[0053] 在某些实现方式中,所述定向音频束包括多频道音频节目的缩混,所述多频道音频节目包括多个频道。在某些实现方式中,调节所述定向音频束的步骤包括改变所述多频道音频节目的混音。在某些实现方式中,所述多频道音频节目包括中心频道,并且其中,多频道的混音包括增加对于移动到与所述定向音频束的源最接近的位置的所述听众的耳朵而言的中心频道节目的幅度。在某些实现方式中,所述定向音频束包括超声音频束。在某些实现方式中,所述图像拍摄装置包括集成到电视接收器装置中的相机。在某些实现方式中,所述图像拍摄装置包括集成到电子显示装置中的相机。在某些实现方式中,还包括至少一个用万向架安装的超声传感器阵列,并且其中,控制和调节所述定向音频束包括控制伺服电机,所述伺服电机定位所述用万向架安装的超声传感器阵列。

[0054] 另一种音频传递系统具有图像拍摄装置,该图像拍摄装置被配置为拍摄收听区域的图像。一个或多个经编程处理器被编程为:处理所述图像以定位听众在所述收听区域中的位置,处理所述图像以识别所述收听区域中的所述听众的面部,处理所述图像以定位所述听众的左耳和右耳的位置,检索与识别出的面部关联的存储的听众简档,基于所述听众简档调节一个或多个音频特性,以及控制左频道定向音频束和右频道定向音频束,以将所述左定向音频束和所述右定向音频束分别导向所述听众的左耳和右耳。所述图像拍摄装置还被配置为拍摄所述听众的后续一系列图像;以及所述一个或多个经编程处理器还被编程为:通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动,以及按照所述听众的左耳和右耳在所述收听区域内的移动来调节由所述左定向音频束和所述右定向音频束携带的音频的混音。

[0055] 在某些实现方式中,所述左定向音频束和所述右定向音频束包括多频道音频节目的立体声缩混,所述多频道音频节目包括中心频道。在某些实现方式中,调节音频的混音包括增加对于听众的右耳或左耳中的任一个而言的所述中心频道节目的幅度,以增加对于移动到与所述定向音频束的源最接近的位置的所述听众的右耳或左耳中的一个而言的所述中心频道节目的幅度。在某些实现方式中,所述定向音频束包括超声音频束。在某些实现方式中,所述定向音频束包括超声音频束。在某些实现方式中,所述图像拍摄装置包括集成到电视接收器装置中的相机。在某些实现方式中,所述图像拍摄装置包括集成到电子显示装置中的相机。在某些实现方式中,还包括至少一对用万向架安装的超声传感器阵列,并且其中,控制和调节所述定向音频束包括控制伺服电机,所述伺服电机定位所述用万向架安装的超声传感器阵列。

[0056] 符合某些实现方式的一种音频传递方法包括:在经编程处理器处检索并读取存储的听众简档,以确定与听众关联的音频特性设置;以及在音频混音器处,所述经编程处理器基于存储的所述听众简档将多频道音频节目的频道的混音调节成数量相等或减少的频道。

[0057] 在某些实现方式中,所述方法还包括向听众播放数量相等或减少的频道。在某些实现方式中,所述经编程处理器还基于所述听众的位置调节频道的混音。

[0058] 在音频传递方法中,拍摄收听区域的图像并且处理图像以定位听众在房间中的位置。检索与听众关联的存储的听众简档并且基于听众的简档创建音频特性。将定向音频束导向听众的耳朵并且调节定向束以追踪听众的移动。

[0059] 本领域的技术人员应该认识到,在考虑到以上教导的情况下,以上某些示例性实

施例是基于一个或多个经编程处理器的使用。然而,本发明不限于这些示例性实施例,因为可以使用诸如专用硬件和 / 或专用处理器的硬件组件等同物实现其它实施例。类似的,可以使用通用计算机、基于微处理器的计算机、微控制器、光学计算机、模拟计算机、专用处理器、专用电路和 / 或专用硬连线逻辑器件来构造替代的等同实施例。

[0060] 使用诸如处理器 218 的经编程处理器实现或可以实现本文中描述的某些示例实施例,所述经编程处理器执行以上用流程图广义地描述的编程指令,所述编程指令可以被存储在任何合适的非暂态电子或计算机可读存储介质上,其中,如本文中所使用的,术语“非暂态”仅仅旨在排除传播波,并不是诸如当掉电时丢失信息的随机存取存储器或可覆写存储器的装置。然而,本领域的技术人员应该理解,在考虑到本教导的情况下,在不脱离本发明的实施例的情况下,可以用任何数量的变形形式并且用许多合适的编程语言实现上述的那些处理。例如,执行某些操作的次序可以经常变化,可以在不脱离本发明的某些实施例的情况下添加额外操作或者删除一些操作。在不脱离本发明的某些实施例的情况下,可以添加错误捕获、超时等,和 / 或可以在用户界面和信息呈现中进行增强和变形。这种变形被预料到,并且被视为是等同的。

[0061] 虽然已经描述了某些示例性实施例,但明显的是,依据以上描述,对于本领域的技术人员而言,许多替代形式、修改形式、移位形式和变形形式将变得清楚。

[0062] 本发明也可以被如下配置:

[0063] (1) 一种音频传递方法,包括:

[0064] 使用图像拍摄装置拍摄收听区域的图像;

[0065] 在一个或多个经编程处理器处:

[0066] 处理所述图像以定位听众在所述收听区域中的位置,

[0067] 处理所述图像以识别所述收听区域中的所述听众的面部,

[0068] 处理所述图像以定位所述听众的耳朵的位置,

[0069] 检索与识别出的面部关联的存储的听众简档,

[0070] 基于所述听众简档调节一个或多个音频特性,以及

[0071] 控制定向音频束,以将所述定向音频束导向所述听众的耳朵;

[0072] 使用所述图像拍摄装置拍摄所述听众的后续一系列图像;以及

[0073] 在所述一个或多个经编程处理器处:

[0074] 通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动,以及

[0075] 按照所述听众在所述收听区域内的移动来调节所述定向音频束。

[0076] (2) 根据 (1) 所述的方法,其中,所述定向音频束包括多频道音频节目的缩混,所述多频道音频节目包括多个频道。

[0077] (3) 根据 (2) 所述的方法,其中,调节所述定向音频束的步骤包括改变所述多频道音频节目的混音。

[0078] (4) 根据 (3) 所述的方法,其中,所述多频道音频节目包括中心频道节目,其中,改变所述多频道音频节目的混音包括增加对于移动到与所述定向音频束的源最近的位置的所述听众的耳朵而言的中心频道节目的幅度。

[0079] (5) 根据 (1) 所述的方法,其中,所述定向音频束包括超声音频束。

[0080] (6) 根据 (1) 所述的方法, 其中, 所述图像拍摄装置包括集成到电视接收器装置中的相机。

[0081] (7) 根据 (1) 所述的方法, 其中, 所述图像拍摄装置包括集成到电子显示装置中的相机。

[0082] (8) 根据权利要求 1 所述的方法, 其中, 控制的步骤包括控制伺服电机, 所述伺服电机定位用万向架安装的超声传感器阵列。

[0083] (9) 一种音频传递方法, 包括:

[0084] 使用图像拍摄装置拍摄收听区域的图像;

[0085] 在一个或多个经编程处理器处:

[0086] 处理所述图像以定位听众在所述收听区域中的位置,

[0087] 处理所述图像以识别所述收听区域中的所述听众的面部,

[0088] 处理所述图像以定位所述听众的左耳和右耳的位置,

[0089] 检索与识别出的面部关联的存储的听众简档,

[0090] 基于所述听众简档调节一个或多个音频特性, 以及

[0091] 控制左频道定向音频束和右频道定向音频束, 以将所述左定向音频束和所述右定向音频束分别导向所述听众的左耳和右耳;

[0092] 使用所述图像拍摄装置拍摄所述听众的后续一系列图像; 以及

[0093] 在所述一个或多个经编程处理器处:

[0094] 通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动, 以及

[0095] 按照所述听众的左耳和右耳在所述收听区域内的移动, 调节由所述左定向音频束和所述右定向音频束携带的音频的混音。

[0096] (10) 根据 (9) 所述的方法, 其中, 所述左定向音频束和所述右定向音频束包括多频道音频节目的立体声缩混, 所述多频道音频节目包括中心频道节目。

[0097] (11) 根据 (10) 所述的方法, 其中, 调节频道的混音包括增加对于所述听众的右耳或左耳中的任一个而言的中心频道节目的幅度, 以便增加对于移动到与所述定向音频束的源最接近的位置的所述听众的右耳或左耳中的一个而言的所述中心频道节目的幅度。

[0098] (12) 根据 (9) 所述的方法, 其中, 所述定向音频束包括超声音频束。

[0099] (13) 根据 (9) 所述的方法, 其中, 所述图像拍摄装置包括集成到电视接收器装置中的相机。

[0100] (14) 根据 (9) 所述的方法, 其中, 所述图像拍摄装置包括集成到电子显示装置中的相机。

[0101] (15) 根据 (9) 所述的方法, 其中, 控制的步骤包括控制伺服电机, 所述伺服电机定位用万向架安装的超声传感器阵列。

[0102] (16) 一种音频传递系统, 包括:

[0103] 图像拍摄装置, 被配置为拍摄收听区域的图像;

[0104] 处理所述图像以定位听众在所述收听区域中的位置的模块;

[0105] 处理所述图像以识别所述收听区域中的所述听众的面部的模块;

[0106] 处理所述图像以定位所述听众的耳朵的位置的模块;

- [0107] 检索与识别出的面部关联的存储的听众简档的模块；
- [0108] 基于所述听众简档调节一个或多个音频特性的模块；
- [0109] 控制定向音频束,以将所述定向音频束导向所述听众的耳朵的模块；
- [0110] 所述图像拍摄装置还被配置为拍摄所述听众的后续一系列图像；
- [0111] 通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动的模块；以及
- [0112] 按照所述听众在所述收听区域内的移动来调节所述定向音频束的模块。
- [0113] (17) 根据 (16) 所述的系统,其中,所述定向音频束包括多频道音频节目的缩混,所述多频道音频节目包括多个频道。
- [0114] (18) 根据 (17) 所述的系统,其中,调节所述定向音频束的模块包括改变所述多频道音频节目的混音的模块。
- [0115] (19) 根据 (18) 所述的系统,其中,所述多频道音频节目包括中心频道节目,其中,改变所述多频道音频节目的混音的模块包括增加对于移动到与所述定向音频束的源最接近的位置的所述听众的耳朵而言的中心频道节目的幅度的模块。
- [0116] (20) 根据 (16) 所述的系统,其中,所述定向音频束包括超声音频束。
- [0117] (21) 根据 (16) 所述的系统,其中,所述图像拍摄装置包括集成到电视接收器装置中的相机。
- [0118] (22) 根据 (16) 所述的系统,其中,所述图像拍摄装置包括集成到电子显示装置中的相机。
- [0119] (23) 根据 (16) 所述的系统,还包括至少一个用万向架安装的超声传感器阵列,其中,控制所述定向音频束的模块包括控制伺服电机的模块,所述伺服电机定位所述用万向架安装的超声传感器阵列。
- [0120] (24) 一种音频传递系统,包括：
- [0121] 图像拍摄装置,被配置为拍摄收听区域的图像；
- [0122] 处理所述图像以定位听众在所述收听区域中的位置的模块；
- [0123] 处理所述图像以识别所述收听区域中的所述听众的面部的模块；
- [0124] 处理所述图像以定位所述听众的左耳和右耳的位置的模块；
- [0125] 检索与识别出的面部关联的存储的听众简档的模块；
- [0126] 基于所述听众简档调节一个或多个音频特性的模块；
- [0127] 控制左定向音频束和右定向音频束,以将所述左定向音频束和所述右定向音频束分别导向所述听众的左耳和右耳的模块；
- [0128] 所述图像拍摄装置还被配置为拍摄所述听众的后续一系列图像；通过分析所述后续一系列图像来监控所述听众的耳朵在所述收听区域中的位置的移动的模块；以及
- [0129] 按照所述听众的左耳和右耳在所述收听区域内的移动来调节由所述左定向音频束和所述右定向音频束携带的音频的混音的模块。
- [0130] (25) 根据 (24) 所述的系统,其中,所述左定向音频束和所述右定向音频束包括多频道音频节目的立体声缩混,所述多频道音频节目包括中心频道节目。
- [0131] (26) 根据 (25) 所述的系统,其中,调节音频的混音的模块包括增加对于听众的右耳或左耳中的任一个而言的所述中心频道节目的幅度,以增加对于移动到与所述定向音频

束的源最接近的位置的所述听众的右耳或左耳中的一个而言的所述中心频道节目的幅度的模块。

[0132] (27) 根据 (24) 所述的系统,其中,所述定向音频束包括超声音频束。

[0133] (28) 根据 (24) 所述的系统,其中,所述图像拍摄装置包括集成到电视接收器装置中的相机。

[0134] (29) 根据 (24) 所述的系统,其中,所述图像拍摄装置包括集成到电子显示装置中的相机。

[0135] (30) 根据 (24) 所述的系统,还包括至少一对用万向架安装的超声传感器阵列,其中,控制所述定向音频束的模块包括控制伺服电机的模块,所述伺服电机定位所述用万向架安装的超声传感器阵列。

[0136] (31) 一种音频传递方法,包括:

[0137] 在经编程处理器处检索并读取存储的听众简档,以确定与听众关联的音频特性设置;以及

[0138] 在音频混音器处,所述经编程处理器基于存储的所述听众简档将多频道音频节目的频道的混音调节成数量相等或减少的频道。

[0139] (32) 根据 (31) 所述的方法,还包括向听众播放数量相等或减少的频道。

[0140] (33) 根据 (32) 所述的方法,其中,所述经编程处理器还基于所述听众的位置调节频道的混音。

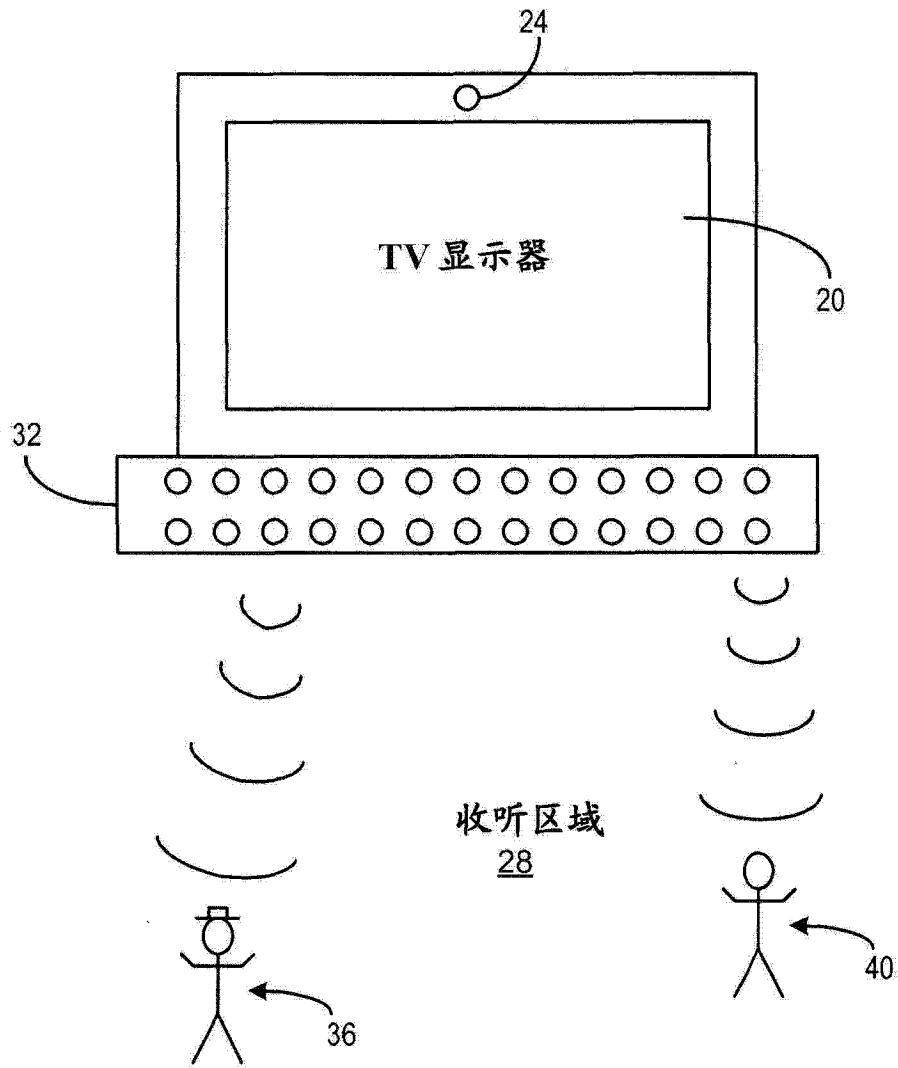


图 1

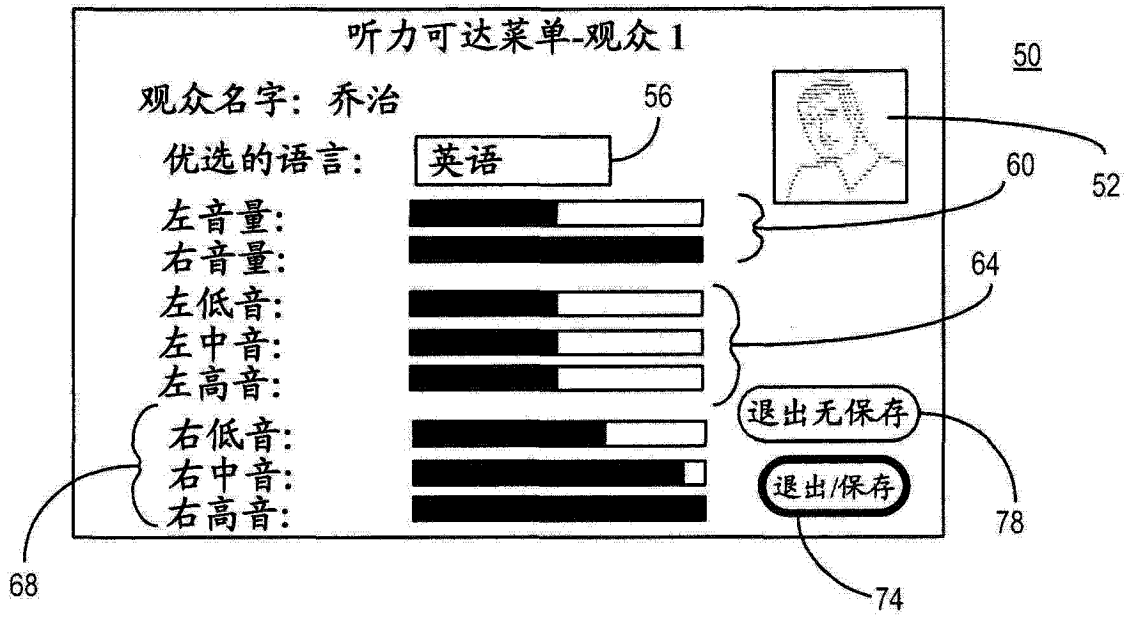


图 2

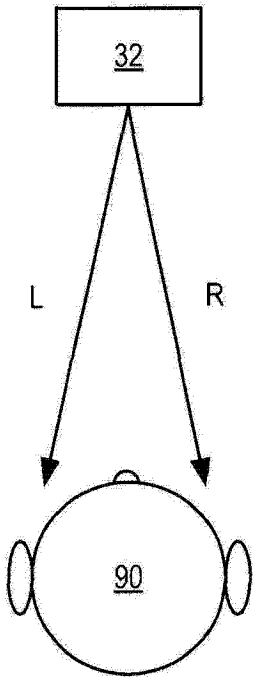


图 3A

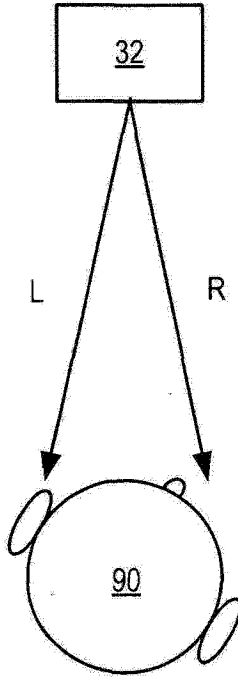


图 3B

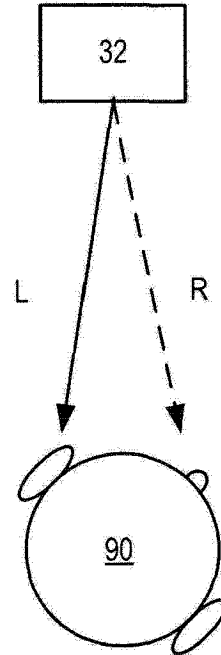


图 3C

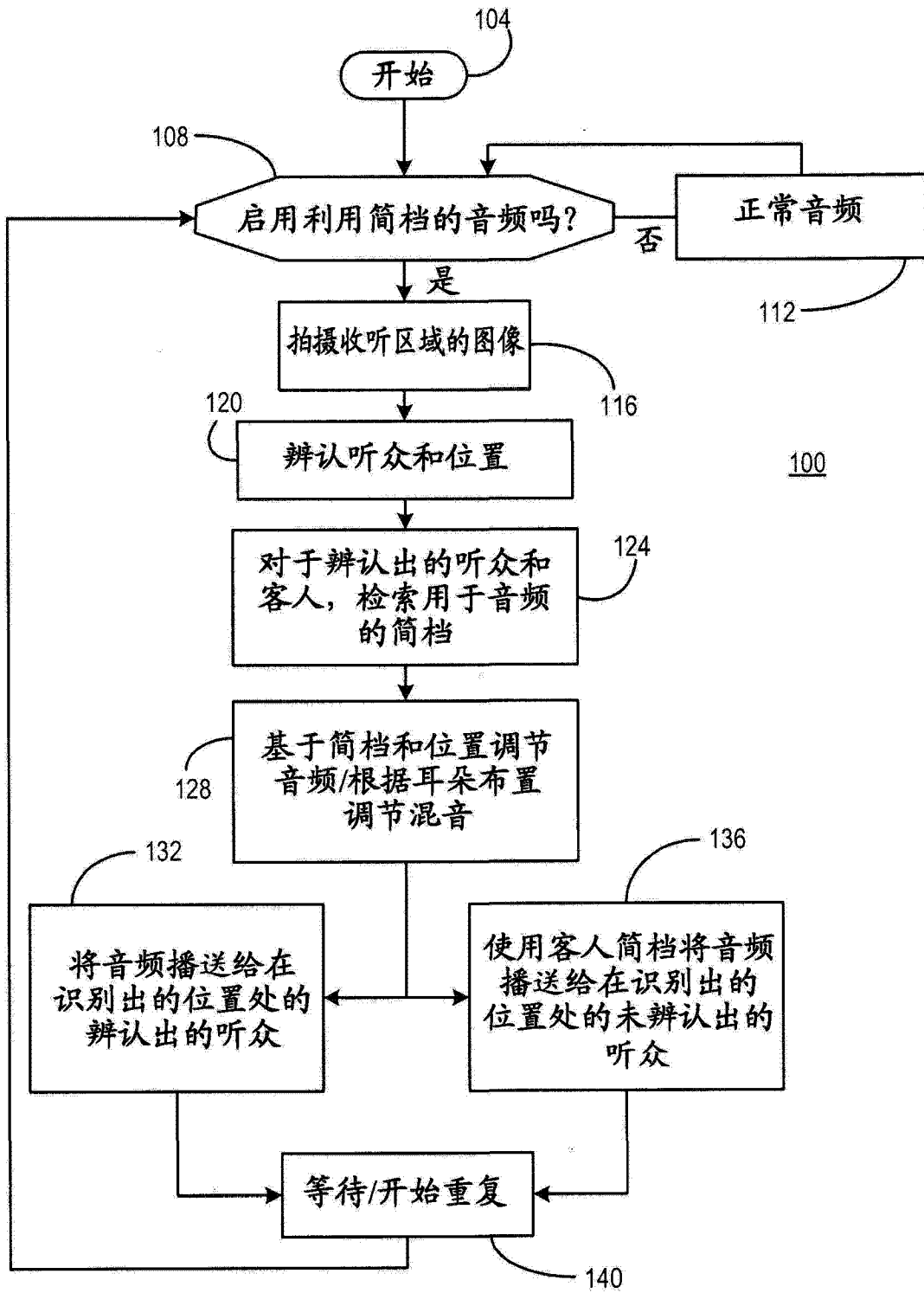


图 4

128

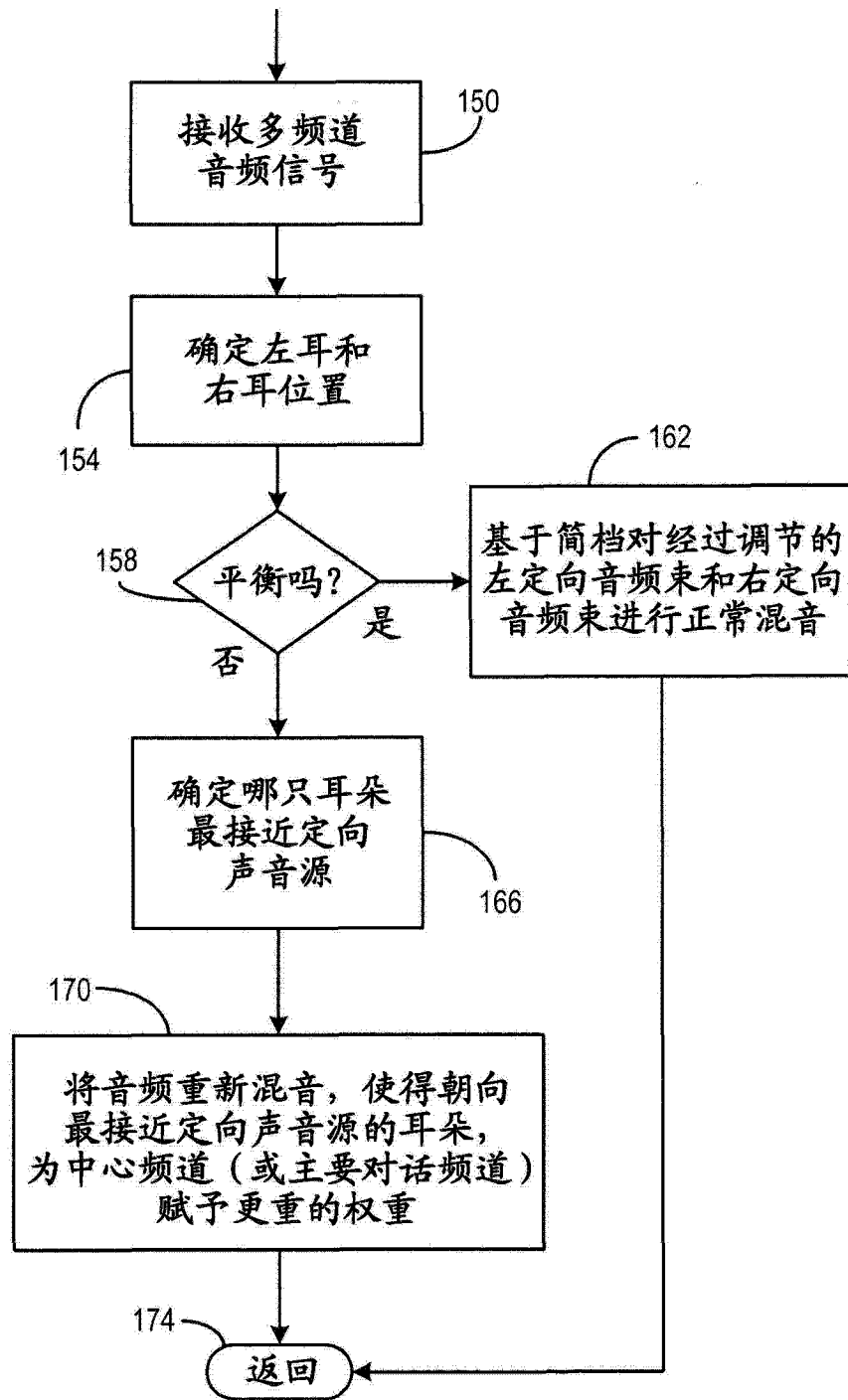


图 5

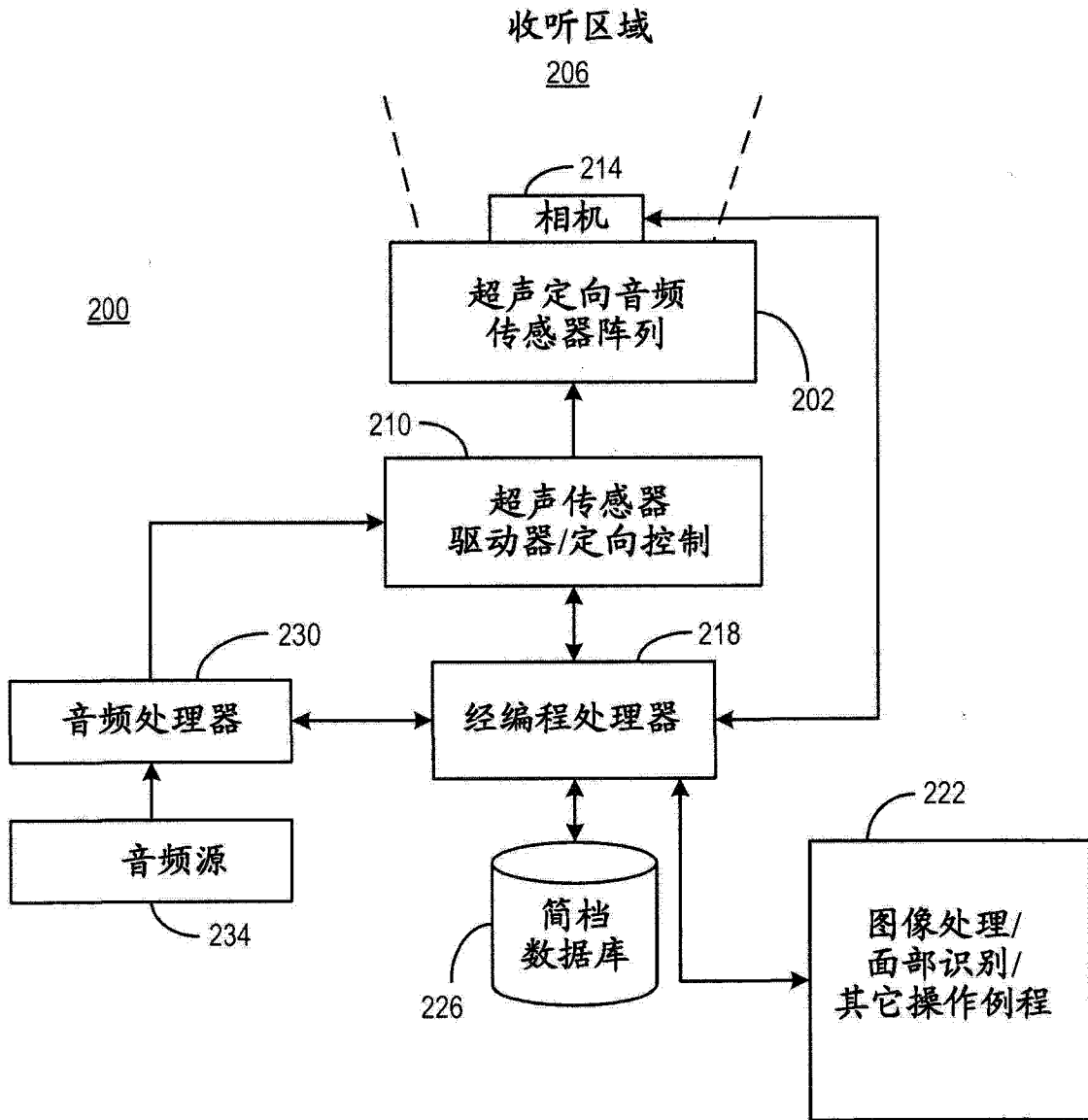


图 6

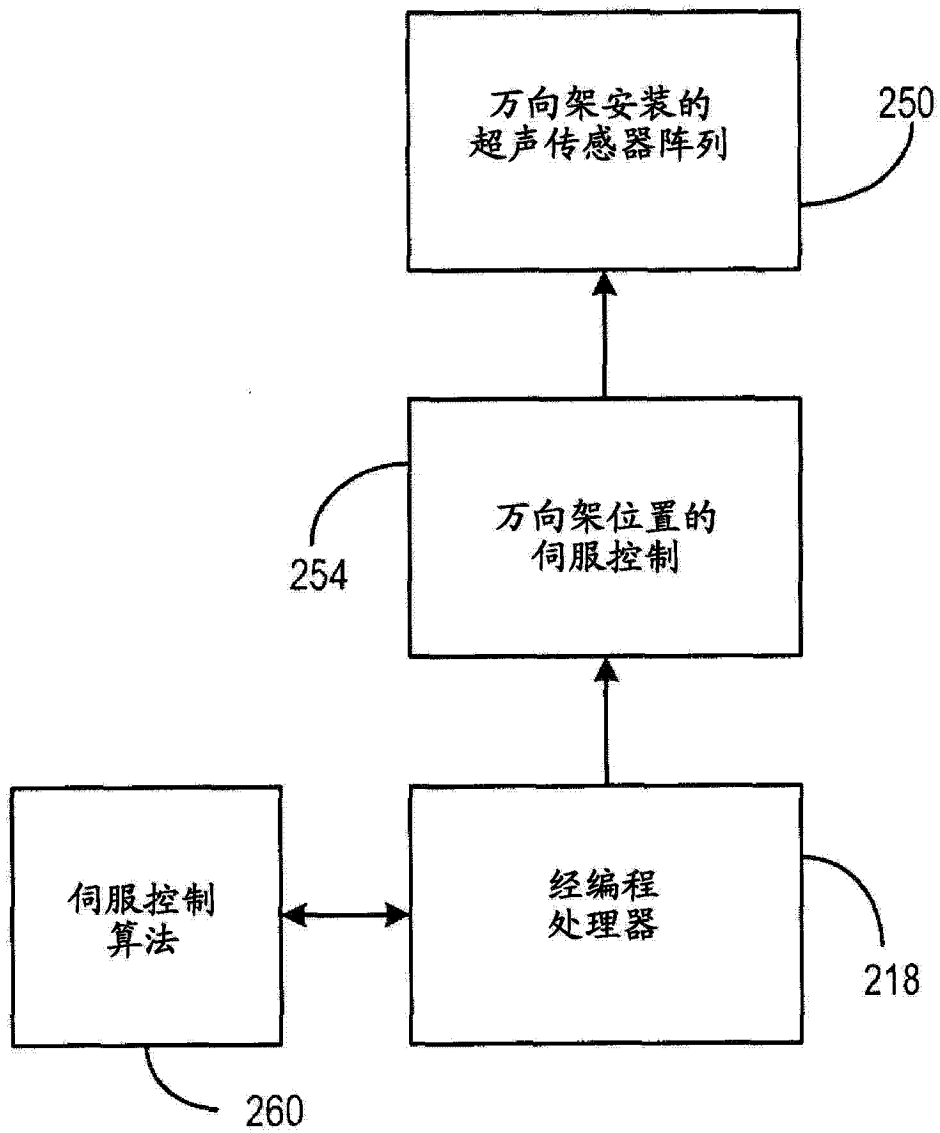


图 7