

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2024 年 6 月 20 日 (20.06.2024)



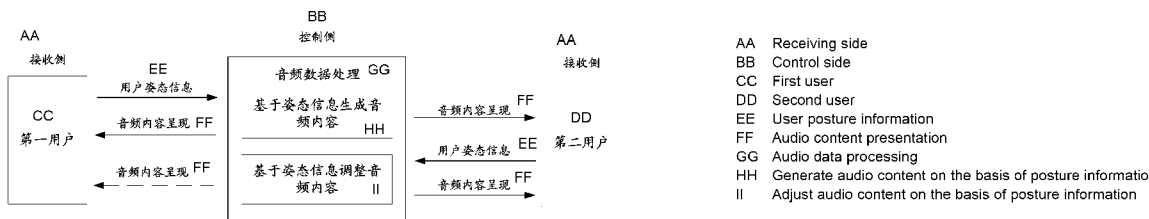
(10) 国际公布号
WO 2024/125478 A1

- (51) 国际专利分类号:
G06F 3/01 (2006.01)
- (21) 国际申请号: PCT/CN2023/138019
- (22) 国际申请日: 2023 年 12 月 12 日 (12.12.2023)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202211599599.6 2022年12月12日 (12.12.2022) CN
- (71) 申请人: 索尼(中国)有限公司(SONY (CHINA) LIMITED) [CN/CN]; 中国北京市朝阳区新源南路 1 号平安国际金融中心商业栋 3 层 301 室, Beijing 100027 (CN)。索尼集团公司(SONY GROUP CORPORATION) [JP/JP]; 日本东京都港区港南 1-7-1, Tokyo 108-0075 (JP)。
- (72) 发明人: 高玮隆(GAO, Weilong); 中国上海市杨浦区四平路 1239 号, Shanghai 200092 (CN)。徐艺晨(XU, Yichen); 中国上海市杨浦区四平路 1239 号, Shanghai 200092 (CN)。金烨鑫(JIN, Yexin); 中国上海市杨浦区四平路 1239 号, Shanghai 200092 (CN)。韩佳(HAN, Jia); 中国上海市浦东新区居里路 361 号, Shanghai 201203 (CN)。

- (74) 代理人: 中国贸促会专利商标事务所有限公司(CCPIT PATENT AND TRADEMARK LAW OFFICE); 中国北京市西城区复兴门内大街 158 号远洋大厦 F10 层, Beijing 100031 (CN)。
- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

(54) Title: AUDIO PRESENTATION METHOD AND DEVICE

(54) 发明名称: 音频呈现方法和设备



(57) Abstract: The present disclosure relates to an audio presentation method and device. Provided is a receiving-side device used for interactive audio presentation, comprising a processing circuit which is configured for receiving, from a control-side device used for interactive audio presentation, information related to audio content to be presented, wherein the audio content to be presented comprises audio content set on the basis of user posture information, and presenting the audio content, wherein presenting the audio content comprises presenting the audio content in a tactile manner.

(57) 摘要: 本公开涉及音频呈现方法和设备。提供了一种用于交互式音频呈现的接收侧设备, 包括处理电路, 其被配置为接收来自用于交互式音频呈现的控制侧设备的待呈现音频内容的相关信息, 其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容, 以及使得呈现音频内容, 其中呈现音频内容包括以触觉方式呈现音频内容。

[见续页]

本国际公布:

— 包括国际检索报告(条约第21条(3))。

音频呈现方法和设备

相关申请的交叉引用

本申请是以申请号为 202211599599.6、申请日为 2022 年 12 月 12 日的中国申请
5 为基础，并主张其优先权，该中国申请的公开内容在此作为整体引入本申请中。

技术领域

本公开涉及音频信号处理，特别涉及音频信号呈现。

10 背景技术

音乐往往在人类生活中能够带给听众快乐和美好。音乐存在各种各样的形式并且可以在各种场合中被呈现给听众。然而，收听音乐对于听力受损的人士是比较困难的，而且随着生活水平进步，越来越多的听力受损人士也希望能够感受到音乐的魅力，享受音乐的乐趣。

15 因此，需要提供改进的音乐呈现方案。

发明内容

提供该发明内容部分以便以简要的形式介绍本公开的构思，这些构思将在后面的具体实施方式部分被详细描述。

20 本公开提供了对音频信号呈现进行优化，特别地优化针对特定用户的音频信号呈现。

本公开还提供了优化的交互式音频信号呈现。

在本公开的一个方面，提供了一种用于交互式音频呈现的接收侧设备，所述设备包括处理电路，被配置为：接收来自用于交互式音频呈现的控制侧设备的待呈现音频
25 内容的相关信息，其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容，以及使得呈现音频内容，其中呈现音频内容包括以触觉方式呈现音频内容。

在本公开的另一方面，提供了一种用于交互式音频呈现的控制侧设备，所述设备包括处理电路，被配置为：获取音频内容呈现指示信息，所述音频内容呈现指示信息包括基于音频要被呈现给的用户姿态信息的指示信息，将待呈现音频内容的相关信息
30 发送给用于音频交互呈现的接收侧设备，其中所述待呈现音频内容包括基于用户的

姿态信息设定的音频内容。

在本公开的另一方面，提供了一种用于交互式音频呈现的接收侧的方法，包括：接收来自用于交互式音频呈现的控制侧设备的待呈现音频内容的相关信息，其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容，以及使得呈现音频内容，其中呈现音频内容包括以触觉方式呈现音频内容。

在本公开的另一方面，提供了一种用于交互式音频呈现的控制侧的方法，包括：获取音频内容呈现指示信息，所述音频内容呈现指示信息包括基于音频要被呈现给的用户的姿态信息的指示信息，将待呈现音频内容的相关信息发送给用于音频交互呈现的接收侧设备，其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容。

在本公开的还另一方面，提供了一种设备，包括至少一个处理器和至少一个存储设备，所述至少一个存储设备其上存储有程序代码和/或指令，该程序代码和/或指令在由所述至少一个处理器执行时可使得所述至少一个处理器执行如本文所述的方法。

在本公开的仍另一方面，提供了一种存储有程序代码和/或指令的存储介质，该程序代码和/或指令在由处理器执行时可以使得执行如本文所述的方法。

在本公开的仍另一方面，提供了一种程序产品，所述程序产品包含程序代码和/或指令，该程序代码和/或指令在由处理器执行时可使得所述处理器执行如本文所述的方法。

在本公开的仍另一方面，提供了一种计算机程序，所述计算机程序包含程序代码和/或指令，该程序代码和/或指令在由处理器执行时可使得所述处理器执行如本文所述的方法。

从参照附图的示例性实施例的以下描述，本公开的其它特征将变得清晰。

附图说明

下面参照附图说明本公开的优选实施例。此处所说明的附图用来提供对本公开的进一步理解，各附图连同下面的具体描述一起包含在本说明书中并形成说明书的一部分，用于解释本公开。应当理解的是，下面描述中的附图仅仅涉及本公开的一些实施例，而非对本公开构成限制。

图 1 示出了根据本公开的实施例的音频呈现的概念图。

图 2A 示出了根据本公开的实施例的交互式音频呈现的概念图。

图 2B 示出了根据本公开的实施例的交互式音频呈现的流程图。

图 3A 示出了根据本公开的实施例的用于交互式音频呈现的接收侧设备的框图。

图 3B 示出了根据本公开的实施例的用于交互式音频呈现的接收侧方法的流程图。

图 4A 示出了根据本公开的实施例的用于交互式音频呈现的控制侧设备的框图。

图 4B 示出了根据本公开的实施例的用于交互式音频呈现的控制侧方法的流程图。

5 图 5 示出了根据本公开的实施例的待呈现音频内容设定的概念性流程图。

图 6A 到 6C 示出了示例性姿态检测的示意图。

图 7A 示出了根据本公开的第一用户（演奏者或表演者）的示例性姿态的示意图。

图 7B 示出了根据本公开的第二用户（听众）的示例性姿态的示意图。

图 8 示出了根据本公开的实施例的示例性音频转换的曲线图。

10 图 9 示出了根据本公开的实施例的接收侧设备的示例性实现。

图 10 示出了根据本公开的实施例的示例性实现场景。

图 11 示出了示出了能够实现本公开的实施例的计算机系统的示例性硬件配置的框图。

应当明白，为了便于描述，附图中所示出的各个部分的尺寸并不一定是按照实际
15 的比例关系绘制的。在各附图中使用了相同或相似的附图标记来表示相同或者相似的部件。因此，一旦某一项在一个附图中被定义，则在随后的附图中可能不再对其进行进一步讨论。

具体实施方式

20 下面将结合本公开的附图，对本公开实施例中的技术方案进行清楚、完整地描述，但是应理解，所描述的实施例仅仅是本公开的一部分实施例，而不是全部的实施例。附图以及下文对实施例的描述实际上也仅仅是说明性的，而不作为对本公开及其应用或使用的任何限制。应当理解的是，本公开可以通过各种形式来实现，而且不应该被解释为仅限于这里阐述的实施例。

25 此外，在下文中结合附图对本公开的示例性实施例进行描述时，为了清楚和简明起见，在说明书中并未描述实施例的所有特征。应当注意，为了避免因不必要的细节而模糊了本公开，在附图中仅仅示出了与根据本公开的方案密切相关的处理步骤和/或设备结构，而省略了与本公开关系不大的其他细节。

应当理解，本公开的方法实施方式中记载的各个步骤可以按照不同的顺序执行，
30 和/或并行执行。此外，方法实施方式可以包括附加的步骤和/或省略执行示出的步骤。

本公开的范围在此方面不受限制。除非另外具体说明，否则在这些实施例中阐述的部件和步骤的相对布置、数字表达式和数值应被解释为仅仅是示例性的，不限制本公开的范围。

在本公开中，术语“第一”、“第二”等仅仅用于区分元件或者步骤，而不是要指示时间顺序、优先选择或者重要性。

音乐是人类共有的精神食粮。人们喜欢音乐，除了音乐好听，能给我们带来情绪上的不同感受以外，特定的音乐对人的身心健康有着积极作用。随着生活水平的日益提高，越来越多的人有机会收听到音乐，也越来越喜欢收听音乐。

但是，应该认识到，当今世界中还存在大量听力受损的人士，尤其在中国大约由 2700 万左右的听力受损人士，其中的大部分人士是轻度到重度听力受损，甚至在年轻人中也有约 1150 万听力受损的人士。这些听力受损人士尤其希望能感受到音乐的美。然而，通常的各种社交活动中的音乐播放主要是面向听力未受损的人士，而没有特别考虑听力受损人士的收听情况。比如年轻人喜欢去各种音乐吧，然而普通的音乐吧中听力受损人士往往由于听力受损而无法如听力无损的人士那样收听音乐。

因此希望提供改进的音频呈现方案，尤其为听力受损人士提供改进的音频呈现。

此外，考虑到在一些音乐播放场景中，例如在各种音乐吧、KTV、演唱会等场景中，听众往往希望能够参与到音乐中，实现希望的音乐互动，因此系统提供改进的互动音频呈现方案，尤其是为听力受损人士提供改进的互动音频呈现。

一方面，本公开提出了一种改进的音频呈现方案，特别地，能够以触觉方式将音频内容呈现给听力受损人士。更特别地，待呈现的音频内容可以经由佩戴在听众身上的触感提供装置作为相应的振动提供给听众。

另一方面，本公开提出了一种改进的交互音频呈现方案，特别地，可以通过检测听众的特定输入来影响音频内容，例如正在播放的音频内容，从而使得音频内容能够更以用户更加希望的方式、节奏等等呈现给听众。更特别地，本公开尤其提出了根据姿态信息来实现对音频内容的影响，从而实现更加方便的交互。

以下结合示例性示例来具体描述根据本公开的实施例。应指出，在本公开的上下文中所提及的音频内容可以为各种适当形式，作为示例可以与音频有关，其可以涵盖任何适当类型的音乐信号，诸如音乐旋律、音轨、音元、音序、音效等等。例如，音频内容可对应于完整的音乐或者其一部分，甚至可为对应于用户特定输入，例如与特定用户姿态相对应的音乐片段。

图 1 示出了根据本公开的实施例的音频呈现的概念图。根据本公开的实施例的音频呈现尤其适合于听力受损人士，可以基于姿态信息来实现音频呈现。

首先，采集与音频呈现相关的数据/信息。特别地，在基于姿态信息实现音频呈现的场景中，采集参与音频呈现的成员的姿态信息。作为示例，参与音频呈现的人员可包括听众，尤其是听力受损的人士。参与音频呈现的人员还可以包括负责音频呈现的特定成员，例如主持人、DJ、演奏者，表演者等等。当然，所采集的与音频呈现相关的数据/信息还可以包括其它数据/信息，诸如成员的参数信息（包括诸如身份 ID 等），启动和/或停止音频呈现的指令、其它与音频呈现控制相关的数据/信息等。

其次，基于所采集的数据/信息进行音频处理。特别地，可以基于所采集的数据/信息，尤其是姿态信息，进行待呈现音频内容的设定，稍后将此进行详细描述。

然后，将所设定的待呈现音频内容呈现给听众。特别地，对于听力受损的人士可以通过触觉方式来实现音频内容呈现。当然，还可以通过其他方式来实现音频呈现。例如，可以通过视频、视觉特效、灯光特效等来呈现给听众，以便更加丰富音频呈现效果。

以下将参照图 2A-2B 描述根据本公开的实施例的交互式音频呈现，其中图 2A 示出了根据本公开的实施例的交互式音频呈现的概念图，图 2B 示出了根据本公开的实施例的交互式音频呈现的流程图。

根据本公开的一些实施例的交互式音频呈现可以适用于各种应用场景，例如在各种音乐吧、聚会地点参与音乐互动的音乐现场场景。在这样的场景中包含第一用户和第二用户。其中，第一用户可以是场景中的负责或主导音频呈现的人士，例如主持人，DJ 等等，其可以启动、暂停、结束、设定、调整要呈现给用户的音频内容。第二用户可以是场景的音频呈现对象，例如音乐吧、聚会地点的顾客、参与者、听众等等。这里，第一和第二用户中的至少一者尤其可以是听力受损人士。但是应理解，第一和第二用户也可以不是参与现场的人士，例如可以是通过网络、云等来参与音乐的人士。

一方面，通过获取第一用户的姿态信息来设定音频内容，然后将所设定的音频内容呈现给第一用户和第二用户中的至少一者。在一些实施例中，可以基于第一用户的姿态信息生成或创建音频内容，例如与第一用户的姿态信息相对应的音频内容，例如由与第一用户的各姿态信息相对应的各音频单元来组合得到的音频内容。在另一些实施例中，第一用户的姿态信息可以仅仅指示音频呈现启动、暂停、停止等，从而在用户姿态信息表示启动时，可以开始呈现特定音频内容，例如播放特定音乐，诸如预先设定的音频/音乐。音频内容可被以各种适当方式呈现给用户，例如触觉、视频、视觉

特效、灯光等等，这里将不再详细描述。

另一方面，通过获取第二用户的姿态信息来对音频内容进行调整，例如对于正在播放的音频/音乐进行调整，然后将调整后的音频内容呈现给第一用户和第二用户中的至少一者。音频内容的调整可被以各种适当方式来实现，例如调整音频播放的音量、旋律等等，这样的调整可以相应地反映到触觉实现中。

根据本公开的实施例，音频数据处理可以被以适当的方式来实现，例如可以由软件、硬件、固件等实现。其可以位于用于音频呈现的系统的控制侧，由用于音频呈现的控制侧设备来实现，例如网络中的服务器、控制设备等。另一方面，音频呈现给用户可对应于系统的接收侧，其可以配置有接收侧设备，接收音频内容以便以适当的方式呈现给用户。例如，接收侧设备可以与各种呈现设备，例如触觉、视觉特效、灯光等呈现设备相配合以呈现给用户。当然，呈现设备可也被包含在接收侧设备中。

用于第一用户和第二用户的接收侧设备可以不同，例如可以对于不同的用户根据用户的需求采用不同的接收侧设备，和/或音频内容呈现装置。用于第一用户和第二用户的接收侧设备也可以相同。例如，这样的接收侧设备和/或音频内容呈现装置可能能够单独设置功能，从而可以针对不同的用户根据用户的需求来设置不同的功能配置，例如对于不同用户相应地开启或关闭某些功能。

以下将参照附图来描述根据本公开的实施例的实现。

图 3A 示出了根据本公开的实施例的用于交互式音频呈现的接收侧设备的框图。该接收侧设备 300 包括处理电路 302，其被配置为接收待呈现音频内容的相关信息，其中所述待呈现音频内容可包括基于用户的姿态信息设定的音频内容，以及使得呈现音频内容，其中使得呈现音频内容可包括使得以触觉方式呈现音频内容。

根据本公开的一些实施例，音频内容可以被适当的方式提供到接收侧设备，从而音频内容相关信息也相应地为各种适当的形式。

在一些实施例中，音频内容可以为各种适当的格式，并且直接作为音频内容相关信息提供到接收侧。例如，音频内容为要播放的音乐，可以为各种适当的音乐格式，例如 mp3、midi，其它适当格式等，并且被直接发送给接收侧。

在另一些实施例中，音频内容相关信息可以是指示音频内容的信息，例如为音频内容的索引。作为示例，音频内容和音频索引可以预先关联设定和存储，并且在应用过程中可以根据音频索引来调用对应的音频内容。

在还另一些实施例中，音频内容相关信息可以是将音频内容转换得到的信息/数

据，特别地，在预先知晓音频内容要以何种方式呈现给用户的情况，可以在控制侧将音频内容转换成适合于该呈现方式的信息/数据，然后作为相关信息传递到接收侧。

根据本公开的实施例，待呈现音频内容可被各种适当的方式设定，包括但不限于生成、创建、调整等。特别地，可以基于用户的姿态信息来设定。根据本公开的实施例，用户的姿态信息包括用户的姿态（包括例如特定部位的姿势、空间位置等）、姿态运动信息中的至少一者，其中，姿态运动信息包括姿态运动轨迹、运动加速度中的至少一者。作为示例，姿态运动信息可以包含特定姿态的移动方向、移动速度、移动加速度、移动频率等等。作为示例，在用户姿态对应于用户手指姿态的情况下，用户姿态可以包括是特定手势、空间位置等，姿态的动作可指的是手势的动作，例如特定手势如何摆动，摆动的速度、摆动的方向、摆动的频率等等。

根据本公开的一些实施例，该接收侧设备的应用场景中可包括各种类型的用户，尤其包括第一用户和第二用户，如上所述。这里，音频内容可包括基于第一用户和第二用户中的至少一者的姿态信息设定的音频内容。

图5示出了根据本公开的实施例的待呈现音频内容设定的概念性流程图。其中，在音频内容呈现场景中，当用户调整姿态时，获取或者检测用户的姿态，生成用户姿态相关的信息和/或数据，由此设定待呈现音频内容以便呈现给用户。应指出，待呈现音频内容设定通常可在系统的控制侧来实现，特别地由控制侧设备来实现。

以下将参照附图来进一步描述音频内容设定和呈现中的具体实现。

用户姿态的获取或检测

根据本公开的实施例，用户姿态的获取或者检测可通过各种适当的方式来执行。

在一些实施例中，可以通过视频采集、图像捕捉等方式来获取用户姿态。例如可以通过相机/摄像头获取用户的动作，然后从所获取的用户动作的图像或视频中进行用户姿态分析，以获取用户姿态相关的信息/数据。在一种示例性实现中，可以通过相机动作捕捉、相机颜色捕捉等来实现。可以在用户的特定部位设置特定颜色或者特定标签，然后通过相机颜色识别来获取相应部位的动作。例如，可以在用户的至少一个手指上贴上特定颜色的贴片，然后通过相机颜色识别来捕获用户手指姿态/动作相关的信息/数据，如图6A所示。

这样可以适用于宽范围的应用场景。例如在各种聚会场景中，可利用预先安装的聚会场所的摄像头，可以捕获聚会现场场景中所关注的用户的姿态；还例如在远程场景中，分别利用各用户自身专用的摄像头来捕获各用户的姿态，然后上传至网络；由

此可以在服务器端或云端进行姿态捕获。

在另一些实施例中，可以通过相机骨骼捕获来获取用户姿态。例如，可以通过相机捕获用户手部特定部位的动作，例如手指整体轮廓、骨骼等的动作来获取相应部位的姿态，如图 6B 所示。作为一个示例，可通过投影仪等设备通过特定算法来检测手指骨骼的运动状态，来获取手指姿态。

在另一些实施例中，用户可以佩戴特定的姿态捕捉装置，诸如动作捕捉传感器、陀螺仪等，然后根据姿态捕捉装置的数据来获取用户姿态相关的信息/数据，如图 6C 所示。在一些实施例中，所要获取的用户姿态信息为用户手部姿态信息，并且姿态捕捉装置可包括能够佩戴在用户的至少一个手指上的动作捕捉器件，并且姿态信息是基于佩戴有动作捕捉器件的各手指的姿态信息和/或它们的组合的。

应指出，在此情况下，用户姿态相关的信息/数据可以看做在接收侧获取并提供给控制侧以供进行音频内容设定。特别地，所述控制侧设备进一步配置为获取经由姿态捕捉装置确定的用户的姿态信息，并且将所获取的用户姿态信息发送给控制侧设备。应指出，接收侧设备还可以提供其它适当的信息，例如该用户的参数信息，诸如用户身份 ID 等等。

以下将描述根据本公开的实施例的姿态捕获和转换的示例性实现。该实现可以是在网络场景下示例性实现。其中用户在摄像头之前挥动手，由此可以通过计算机摄像头来捕获用户手指的运动，例如，可通过比较相邻图片之间的像素差异来确定用户手指的运动状态和轨迹，并且相应地生成手指运动数据，这样的数据可被以各种适当的方式表示和存储，例如可以包括每个手指的数据编号、以及对应的数据，包括但不限于摆动速度、摆动位置，时间点等等。由此可以确定用户的手指姿态等等。这样可以通过本领域已知的各种方式来实现，这里将不再详细描述。然后基于所确定的用户手指的运动数据来确定对应的音频内容，例如执行将其转换成 MIDI 和音乐化处理。

音频转换

根据本公开的实施例，可以基于所获取的用户姿态的相关信息/数据来设定相应的音频内容以供呈现给用户。

在一些实施例中，待呈现音频内容可以包括基于与所述第一用户的姿态信息相对应的音频单元或者特定组合构建的音频内容。特别地，可以基于姿态数据与音频单元之间的关联性或对应性来设定音频内容。音频单元可以是音频内容的组成单元，例如可对应于音元、音序、音频片段等等中的至少一者。由此，可以获取用户的至少一个

姿态，然后使用对应于所述至少一个姿态的音频单元来生成音频内容。在一些实施例中，可以进行音频单元的组合来生成音频内容。特别地，在用户连续作出动作的情况下，可以将用户的各个姿态所对应的音频单元进行组合以得到待呈现的音频内容。在另一些示例中，还可对组合得到的音频内容进行适当的处理，例如滤波、平滑化等等。

5 在一些实施例中，用户姿态与音频单元之间的关联性/对应性可被预先构建，例如可训练各种手势并且为每种手势设定对应的音频单元。在一些实施例中，第一用户可以提供相对精细的姿态信息，例如多个手指的姿态信息，针对音频内容进行相应的控制，例如控制音频的多个音轨，生成更加精确的音频内容，从而更加准确地呈现音频。作为示例，用户姿态可与相对应的音频单元相关联地存储在数据库中。用户姿态与音频单元可以被各种适当的方式来存储。例如，每一用户姿态和与其对应的音频单元可被以列表方式存储，以映射方式存储等等。作为示例，数据库中可包括用户姿态、相应的音频单元、用户姿态变化方式、相应的音频单元变化方式等等，但并不局限于此。只要能够从数据库所存储的数据，基于所获取的用户姿态来生成音频内容和/或改变音频内容即可。

15 在一些实施例中，音频内容设定可采用适当的方式来执行。作为示例，可以采用机器学习或深度学习算法来基于姿态数据来设定音频内容，从而使得根据姿态数据转化的音频 **MIDI** 信号得到更优的滤波和平滑，以增强其音乐性。机器学习或深度学习算法可以包含本领域中已知的各种算法，这里将不再详细描述。在一些实施例中，机器学习或深度学习算法也可预先基于训练数据来训练，训练可采用各种适当的方式来执行，这里将不再详细描述。训练好的 **AI** 模型输入为姿态数据，输出为 **MIDI** 信号以用于呈现音频。进一步地，训练的 **AI** 模型输入为多个用户的姿态数据和表演侧的初始音频内容，输出为经调整的 **MIDI** 信号以用于呈现音频，以实现观众对音乐的共同创作。

25 根据本公开的一些实施例，待呈现的音频内容可以包括由用户姿态信息所指定的特定音频内容。特别地，特定用户姿态可对应于特定的音频内容，由此在检测到特定用户姿态时可以直接将完整的音频内容发送给接收侧设备以供呈现。

30 在一些实施例中，用户的姿态信息还可以对应于音频内容呈现指示信息，其例如可指示音频呈现的特定操作，诸如启动、暂停、停止等，由此可以在检测到该姿态信息时，可以对音频内容呈现执行相应操作。这里的音频内容可以是预先设定的，或者与用户姿态相关联的。

应指出上述的音频内容创建和/或生成可尤其对应于基于本公开的第一用户的姿态信息来设定音频内容的情况，如图 7A 示出了根据本公开的第一用户（演奏者）的示例性姿态的示意图，其中不同的姿态可对应于不同的音乐呈现操作，例如连续移动可对应于表演，握拳可对应于击鼓，对于音乐开始录制、暂停、结束以及其他的操作也都可对应于其他的姿态。例如，可对应于在音频内容呈现场景中，由诸如演奏者、表演者、主持人等第一用户来主导音频内容的生成和/或创建。

音频呈现

根据本公开的实施例，所创建或生成的音频内容可被以各种适当方式呈现给用户。特别地，将音频内容的相关信息转换为适用于音频呈现装置的数据；并且将转换得到的数据提供给音频呈现装置。作为示例，转换得到的数据可以是音频呈现装置的驱动数据或输入数据，以便音频呈现装置能够以特定方式将音频内容呈现给用户。数据转换可通过各种适当的方式来实现。例如，根据本公开的实施例，可以通过各种适当的方式将音频数据改变为触觉数据，例如采用模拟信号方式、FFT(Fast Fourier Transform)过滤方式等等。当然还可采用本领域已知的其它方式，这里将不再详细描述。

图 8 示出了数据转换的示意图，其中对于不同类型的音乐数据将转换得到各自的波形数据并用以驱动音频呈现数据。由于所得到的波形数据往往能够体现不同类型的音乐数据的特性，因此音频呈现装置也能够将音乐的特性、旋律等准确地呈现给用户。

在本公开的一些实施例中，接收侧设备可以与音频呈现装置分离，在另一些实施例中，接收侧设备可以与音频呈现设备集成在一起，特别地，接收侧设备可以包括音频呈现装置。

在本公开的一些实施例中，音频内容可被以各种适当的时序方式提供给接收侧。在一些实施例中，一旦基于用户的姿态信息可得到可播放/呈现的音频单元/片段，就将其之发送到接收侧。在另一些实施例中，可以每次将预定数量的音频单元/片段，甚至是整个音频内容，才发送给接收侧。

应指出，音频内容也可以是被以其它方式设定的待呈现的音频内容，例如接收到特定播放/呈现指令而开始播放/呈现的音频内容，按照预先设定的顺序/指令而开始播放/呈现的音频内容，例如演奏厅、现场等中预定的音频内容，这里将不再详细描述。

根据本公开的实施例，音频呈现装置为触感提供装置，使得经由所述触感提供装

置将音频内容以触觉方式提供给用户。这样，尤其针对听力受损人士，音频内容可以触觉方式呈现给用户。在本公开的一些实施例中，触感提供装置可以包括佩戴在用户手、手腕、手臂等中至少一者上的触觉反馈装置，例如为手套、腕带、臂带等的样式，可以为用户的至少一个手指、手背、手腕、手臂等提供触觉反馈。

5 在一些实施例中，在接收侧接收到的音频内容相关信息为已经由音频内容针对触觉呈现方式转换得到的信息/数据的情况下，可以将所接收到的信息/数据直接转发给触觉装置。在另一些示例中，在所接收到的音频内容相关信息为音频内容本身的情况下，接收侧设备可以将音频内容转换为适合于触觉呈现方式的信息/数据，然后将之提供转发给触觉装置。这样，接收侧装置可以包括转换单元，其被配置为针对音频呈现
10 方式转换得到的信息/数据，以便进行音频呈现。

根据本公开的实施例，触感提供装置可被以各种适当的方式来实现。作为示例，触感提供装置可以包括振动器，其能够将与音频内容的特性，如旋律等，相对应的振动提供给用户，从而让听力受损人士感受到音乐旋律。例如，触感提供装置可以通过惯性执行机构(**Inertial actuator**)，压电半导体换能器(**Piezoelectric actuator**)，电活性
15 高分子驱动器(**Electro-active polymer actuator, EAP**)等等实现，这里将不再详细描述。

在一些实施例中，在所采集的用户姿态为用户手指姿态/动作的情况下，每个用户手指可以对应于特定的音轨，来设定（例如，生成或影响）音频内容的不同的音色。

根据本公开的实施例，所述触感提供装置包括至少一个触觉单元，其中每个触觉单元可对应于待呈现音频内容中的特定音轨。特别地，在一些实施例中，触感提供装
20 置可以包括手套或指套样式的触觉反馈装置，可以为至少一个手指部件设置有振动马达以便提供振动反馈。特别地，在姿态采集阶段每个手指对应不同音轨的情况下，在触觉反馈阶段，手套或指套样式的触觉反馈设备的每个手指部件的震动马达可根据对应音轨的声音强度、节奏进行震动反馈。

在一些实施例中，进一步地，为了提升听力受损人士的用户体验并在应用成本上
25 取得平衡，触感提供装置可以被设定为仅针对听力受损人士难以听到的音轨进行触觉反馈，例如特定频率音频（诸如高频音轨）等进行触觉反馈。作为示例，在手套或指套样式的触觉反馈装置的情况下，可以预先规定多个手指与多个音轨的对应关系，由此被采集每个手指的姿态都可以进行相应音轨的控制，但仅在特定音轨、例如高频音轨对应的手指上设置有触觉反馈单元，从而仅将该特定音轨的对应音频内容触觉反馈
30 给用户。

在一些实施例中，还可以主要提供用户敏感的音频频率、节奏等进行反馈。作为示例，考虑到一般听众对鼓点等节奏韵律较为敏感，反馈设备可以被设定为仅针对鼓点所在的音轨进行触觉反馈。这样能够增强用户感受音乐的用户体验。例如用户在收听音乐时在特定的节奏旋律感受到反馈，进一步提高用户体验。

5 在一些实施例中，在触感提供装置包括手套或指套样式的触觉反馈装置的情况下，其还可被适当地设置以便于识别、控制以及用户操作的简易性。作为示例，可以预先规定特定的一个或多个手指专用于姿态控制，而另一个或多个手指专用于触觉反馈。

在存在演奏人员和听众的情况下，演奏人员的手指佩戴的呈现设备可与听众是对应的。例如，演奏人员的手指佩戴呈现设备可与听众是相同的，其中演奏人员的手指
10 与听众的手指是相对应的，例如，同一手指对应于同一音轨。还例如，演奏人员的手指佩戴设备可以与听众不同，但是对应关系是预先设定的。

根据本公开的实施例，接收侧设备还可以使得音频内容以其它适当的方式呈现给用户。作为示例，可以采取声音、视频展示、灯光、视觉特效等方式来呈现给用户。在此情况下，接收侧设备可以将音频内容转换为适合于其它呈现方式的信息/数据，然
15 后将之提供转发给相对应的呈现装置。当然，应指出，在多种呈现装置可以由相同格式的数据驱动的情况下，可将音频内容转换之后公共地用于各种呈现装置。

作为一些示例，音频内容可被以音频方式提供给用户。作为示例，可以通过扬声器等将音频内容播放给用户。特别地，在播放之前可以在对音频内容进行进一步的处理，例如将音频内容转换为适合听力受损人士的低频内容。这里，还可包含必要的音
20 频播放软件、音频播放设备等等，这里将不再详细描述。这样的扬声器可以是便携式设备的扬声器，影院场景中的扬声器、KTV、酒吧、聚会场所等设置的扬声器等，或者是其它适当类型的扬声器等等。

作为另一些示例，音频内容可被以视频方式提供给用户。例如，通过视频呈现设备来提供给用户。作为示例，可以通过各种类型的屏幕，例如投影仪、计算机屏幕、
25 便携式设备的屏幕等等，以各种适当的视频呈现给用户。这样的视频可以是与音频内容对应的视频轨迹、特效、图片、短视频等等，可以被以预先设定和存储。

作为另一示例，可通过灯光效果来呈现音频，特别地，呈现装置的灯光可以根据音频内容的节奏相应地闪烁。这样的呈现装置可以是固定地设置的，例如固定屏幕、闪烁器件等，也可以是便携式的，例如便携式设备的屏幕、闪烁器件，诸如腕带、摆
30 件等。作为示例，灯光效果可以通过电子腕带上的 LED 来实现。

交互性音频呈现

本公开进一步提出了优化的交互式音频呈现。

根据本公开的实施例，可以基于用户姿态来实现音频互动。特别地，可以获取用户的姿态来对于音频内容进行调整。具体而言，在包括演奏者和听众的场景中，演奏者可以如上所述地向听众呈现音频内容，听众在获取音频内容之后，可以通过其动作来进行反馈，例如通过动作来表达用户的情绪，根据用户的动作来适应性地调整音频内容等等。由此，可以实现用户互动。

根据本公开的实施例，待呈现的音频内容包括基于所述用户的姿态信息对音频内容进行调整而获得的音频内容。在一些实施例中，调整包括以下中的至少一者：增大或减小音频内容的音量；调整音频内容的节奏；增强音频内容的效果；为音频内容增加附加效果。

特别地，作为示例，用户姿态可以对应于特定音频单元、音频片段等，并且特定用户姿态可以对应于针对该特定音频单元的修改、例如增强或减弱该特定音频单元的强度、变换该音频单元的节奏等等。在一些实施例中，在用户正对于特定音频片段作出特定动作时，可以相应地调整该音频片段的呈现效果。例如，特定动作可以指示要增大该音频片段的呈现效果，例如增大音量、增加触觉效果等；减小该音频片段的呈现效果，例如减小音量，减小触觉效果等等。这里的调整可以如前文所述地音频内容修改那样执行。作为示例，以用户手势相对于身体上下左右做动作为例，整个手在左边时指示低音，在右边时指示高音，向上时指示特定音符的高八度，向下时指示特定音符的低八度。

在一些示例中，用户可以通过动作来表达其喜欢该特定音频内容的情绪。例如，通过特定的挥手动作等。这样，可以在聚会场景中通过视频呈现该情绪。

应指出，交互式音频呈现的应用场景尤其适合于基于第二用户（例如听众）的姿态来实现对于所呈现的音频内容的影响。在一些实施例中，第二用户可以提供相对粗略的姿态信息，例如仅一个手指的姿态信息，针对音频内容进行相应的控制，例如控制鼓点、音量，从而简化用户的操作。图 7B 示出了根据本公开的第二用户（听众）的示例性姿态的示意图，例如可通过连续移动来进行音频内容控制。

应指出，第一用户也可以参与到交互式音频呈现的场景中，例如也可以基于第一用户的姿态来对音乐内容进行控制或调整。在此情况下，第一用户也可以被当做特定的第二用户，基于两者的用户姿态来调整音频内容。

在一些实施例中，基于用户姿态调整音频内容可以遵照各种适当的准则来进行。特别地，在存在至少一个第二用户、或者所获取的至少一个第二用户的姿态的情况下，可以基于至少一个第二用户的姿态信息的统计值来调整音频内容的呈现。这样，可以更加全面的考虑第二用户的需求来实现对于音频内容的影响。

5 根据本公开的实施例，至少一个第二用户的姿态信息的统计值包括关于用户姿态信息的优先级的统计值，并且根据至少一个第二用户的姿态信息中的最高优先级姿态信息来调整音频内容的呈现。在一些实施例中，关于用户姿态信息的优先级的统计值如下地被确定：对至少一个音频用户的姿态信息进行加权处理，其中加权处理是基于各姿态信息的数量、各姿态信息的优先级、各姿态信息对应的用户的优先级中的至少
10 一者来进行的。

特别地，在存在多个第二用户的情况下尤其基于多个第二用户的姿态信息的统计值来调整音频内容的呈现，这样能考虑到多个第二用户的群体感受来进行音频设定，实现了群体共同创作的音频呈现反馈，从而能够提升听众的临场感。

在一些示例中，可以对于用户动作姿态设定优先级，然后按照用户姿态的优先级
15 来提供反馈。例如，将各个用户的动作进行汇总，并且按照优先级进行排序，然后根据优先级最高的动作来进行相应的音频内容调整。

在一些示例中，可以根据用户动作的数量来提供反馈。例如，将各个用户的动作进行汇总，统计相同或者相似动作的数量，并且按照数量最大的动作来进行相应的音频内容调整。

20 在一些示例中，还可以进一步考虑用户的优先级。特别地，可以为用户设定优先级，并且根据优先级最高的用户的动作来进行相应的音频内容调整。

在一些实施例中，还可以进一步基于用户动作的优先级、用户的优先级、用户动作的数量等中的至少两者来提供反馈。例如，可以将各个用户的动作进行汇总，并且进行数学统计来获得反馈结果，从而根据反馈结果来进行相应的音频内容调整。

25 在一些示例中，可以设定每个用户的优先级值，设定每个用户的动作的优先级值，然后，将所获取的各用户动作进行统计，以获得每种用户动作的统计值，例如通过将用户优先级值或动作优先级值乘以该动作的数量，由此而得到该用户动作的统计值。然后，根据统计值最高的动作来进行相应的音频内容调整。

通过本公开，可以实现方便的互动。例如在线上、线下的任何场景中可以方便地
30 获取用户的反馈，实现互动，尤其是在线上、线下两者都存在的情况下，也可获得用

户的反馈。而且能够及时根据用户的反馈来进行音频内容的调整，满足用户的修改。还可以采用适当方式进行音频内容调整，这样能够更加适当地调整音频内容，获得更好的呈现效果。

在上述装置的结构示例中，处理电路 302 可以是通用处理器的形式，也可以是专用处理电路，例如 ASIC。例如，处理电路 202 能够由电路（硬件）或中央处理设备（诸如，中央处理单元（CPU））构造。此外，处理电路 302 上可以承载用于使电路（硬件）或中央处理设备工作的程序（软件）。该程序能够存储在存储器（诸如，布置在存储器中）或从外面连接的外部存储介质中，以及经由网络（诸如，互联网）下载。

根据本公开的实施例，处理电路 302 可以包括用于实现上述功能的各个单元，例如接收单元 304，被配置为接收来自用于交互式音频呈现的控制侧设备的待呈现音频内容的相关信息，其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容，以及控制单元 306，被配置为使得呈现音频内容，其中呈现音频内容包括以触觉方式呈现音频内容。控制单元 306 可控制发送单元 308 将音频内容或者其相关信息提供给音频呈现设备，以供音频内容呈现。在一些实施例中，音频呈现装置可以被包含在接收侧设备中，特别地，被包含在控制单元中，从而可以直接由控制单元控制音频呈现设备来呈现音频内容。

在一些实施例中，处理电路 302 还可以包括获取单元 310，被配置为可以获取获取经由姿态捕捉装置确定的用户的姿态信息，并且经由发送单元 308 将所获取的用户姿态信息发送给控制侧设备。特别地，获取单元 310 可以与姿态捕捉装置分离，并且从姿态捕捉装置获取用户姿态信息。在另一种实现中，获取单元 310 可包含姿态捕捉装置。

根据本公开的实施例，处理单路 302 还可以包括转换单元 312，被配置为将音频内容的相关信息转换为适于音频呈现装置的数据；以及经由发送单元 308 将转换得到的数据提供给音频呈现装置。

应注意，尽管图 3 中将各个单元示为分立的单元，但是这些单元中的一个或多个也可以合并为一个单元，或者拆分为多个单元。此外，一些单元也可并不被包含在处理电路甚至接收侧设备中，因此可用虚线示出。作为示例，获取单元 310、转换单元 312 甚至可以在处理电路 302 之外，因此其也可被用虚线示出。

应注意，上述各个单元仅是根据其所实现的具体功能划分的逻辑模块，而不是用

于限制具体的实现方式，例如可以以软件、硬件或者软硬件结合的方式来实现。在实际实现时，上述各个单元可被实现为独立的物理实体，或者也可由单个实体（例如，处理器（CPU 或 DSP 等）、集成电路等）来实现。此外，上述各个单元在附图中用虚线示出指示这些单元可以并不实际存在，而它们所实现的操作/功能可由处理电路本身来实现。

应理解，图 3A 仅仅是用于音频呈现的接收侧设备的概略性结构配置，设备 300 还可以包括其他可能的部件，诸如存储器、网络接口、控制器、通信单元等，为了清楚起见这些部件并未示出。特别地，处理电路可以与存储器相关联。例如，处理电路可以直接或间接（例如，中间可能连接有其它部件）连接到存储器，以进行图像处理相关数据的存取。存储器可以存储由处理电路 302 产生的各种数据和/或信息。存储器还可以位于优化设备内但在处理电路之外，或者甚至位于优化设备之外。存储器可以是易失性存储器和/或非易失性存储器。例如，存储器可以包括但不限于随机存储存储器（RAM）、动态随机存储存储器（DRAM）、静态随机存取存储器（SRAM）、只读存储器（ROM）、闪存存储器。

以下将参照图 3B 来描述根据本公开的实施例的用于失真图像增强的模型训练方法的流程图。在方法 310 中，在步骤 S311（接收步骤），接收来自用于交互式音频呈现的控制侧设备的待呈现音频内容的相关信息，其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容，以及在步骤 S313（控制步骤），使得呈现音频内容，其中呈现音频内容包括以触觉方式呈现音频内容。在一些实施例中，可选地，方法 310 还可包括步骤 S312（转换步骤），将音频内容的相关信息转换为适于音频呈现装置的数据；从而转换得到的数据可被提供给音频呈现装置。

应指出，这些步骤可以由任何适当的设备或设备元件来执行，例如前述的接收侧设备，接收侧设备中的处理电路、处理电路中的相应元件等等。应指出，根据本公开的实施例的音频呈现方法还可包含其他步骤，例如前文所述的各种进一步的处理。而且这些进一步的处理也可通过适当的设备或者设备元件来执行，这里将不再详细描述。

以下将参照图 4A 来描述根据本公开的实施例的用于音频呈现的控制侧设备的框图。该控制侧设备 400 包括处理电路 402，被配置为：获取音频内容呈现指示信息，所述音频内容呈现指示信息包括基于音频要被呈现给的用户姿态信息的指示信息，将待呈现音频内容的相关信息发送给用于音频交互呈现的接收侧设备，其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容。

根据本公开的实施例，所述处理电路 **402** 可进一步配置为：获取用户的姿态信息，以及基于所获取的用户的姿态信息来设定待呈现音频内容。

根据本公开的实施例，所述处理电路 **402** 可进一步配置为：确定至少一个第二用户的姿态信息的统计值，所述统计值包括关于用户姿态信息的优先级的统计值，并且
5 根据至少一个第二用户的姿态信息中的最高优先级姿态信息来设定待呈现音频内容。

根据本公开的实施例，处理电路 **402** 可以采用各种适当的方式来实现，如上文所述的处理电路 **302** 那样，这里将不再详细描述。特别地，根据本公开的实施例，处理电路 **402** 可以包括用于实现上述功能的各个单元，例如获取单元 **404**，被配置为获取音频内容呈现指示信息，所述音频内容呈现指示信息包括基于音频要被呈现给的用户
10 的姿态信息的指示信息，发送单元 **406**，被配置为将待呈现音频内容的相关信息发送给用于音频交互呈现的接收侧设备，其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容。

根据本公开的实施例，处理电路 **402** 可包括设定单元 **408**，被配置为基于所获取的用户的姿态信息来设定待呈现音频内容。

15 根据本公开的实施例，处理电路可包括确定单元 **410**，被配置为确定至少一个第二用户的姿态信息的统计值，所述统计值包括关于用户姿态信息的优先级的统计值，由此设定单元 **408** 可根据至少一个第二用户的姿态信息中的最高优先级姿态信息来设定待呈现音频内容。

根据本公开的实施例，处理单路 **402** 还可以包括转换单元 **412**，被配置为将音频
20 内容转换成适合于接收侧设备接收的信息，甚至可以转换为适于音频呈现装置的数据。

以下将参照图 **4B** 来描述根据本公开的实施例的用于音频呈现的控制方法的流程图。在方法 **410** 中，在步骤 **S411**（获取步骤），获取音频内容呈现指示信息，所述音频内容呈现指示信息包括基于音频要被呈现给的用户姿态信息的指示信息，以及在步骤 **S413**（发送步骤），将待呈现音频内容的相关信息发送给用于音频交互呈现的接收侧设备，其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容。在
25 一些实施例中，可选的，在该方法中，在步骤 **S412**（设定步骤），基于所获取的用户的姿态信息来设定待呈现音频内容。

应指出，这些步骤可以由任何适当的设备或设备元件来执行，例如前述的控制侧设备，控制侧设备中的处理电路、处理电路中的相应元件等等。应指出，根据本公开
30 的实施例的音频呈现方法还可包含其他步骤，例如前文所述的各种进一步的处理。而

且这些进一步的处理也可通过适当的设备或者设备元件来执行，这里将不再详细描述。

根据本公开的实施例，还提供了一种用于交互式音频呈现的系统，该系统可包括如上所述的控制侧设备和接收侧设备，其中接收侧设备可以与至少一个用户相关联，例如与包括第一用户和第二用户的多个用户相关联，其中每个用户佩戴相应的或者相

5 关联的接收侧设备。

在一些实施例中，控制侧设备接收音频呈现用户的姿态信息，并基于音频呈现用户的姿态信息设定待呈现音频内容。而接收侧设备接收待呈现音频内容的相关信息，并且使得呈现音频内容，其中特别地，呈现音频内容包括以触觉方式将音频内容呈现给用户。

10 根据本公开的实施例，还提供了一种用于交互式音频呈现的方法，该方法基于如上所述的控制侧方法和接收侧方法。

作为示例，在操作中，可以获取第一用户（例如聚会、酒吧等场景中的主持者、演唱会的演奏者、各种活动的表演者等等）的姿态（例如，手指姿态等）以生成、创建或者启动音乐，从而通过第二用户（例如听众、观众等）的接收设备，例如佩戴在

15 手上的触觉反馈设备来触觉呈现给第二用户。另一方面，在音乐呈现给第二用户期间，还可以获取第二用户的姿态（例如，手指姿态等），这反映了用户的体验、反馈、需求等等，并且可根据第二用户的姿态来对音乐进行调整以更适应于用户，进一步提高用户的体验。

20 【实现示例】

以下将以手指或手部姿态实现音频内容的呈现和/或反馈为例来描述本公开的实施例的实现。

应指出，同样可应用于脚趾或脚步姿态，进一步可应用于身体其他部分的姿态，还可应用于特定器件的姿态。特定器件可被称为互动器件，该器件的姿态也可应用于

25 生成/调整音频内容。作为一个示例，可以是特定器件，例如拟人玩偶、手持器件、佩戴在身体上的各种器件等，可以采集这些器件的特定姿态，来实现音频内容呈现和/或反馈。作为示例，在舞台、聚会等场景中，入场观众可以发布特定的手持器件，例如荧光棒等，这样可以根据入场观众的手持器件的姿态/动作来作为观众反馈，以相应地调整音频内容。

30 图9示出了根据本公开的接收侧设备的实现方式。其可以实现为可以佩戴在用户

手指上的指套/手套的形式。

作为示例，其可以包含各种适当的元件/装置。其中，**901** 指示接收侧设备的数据接收、发送单元，例如可以接收音频内容相关信息，并且提供特定数据以驱动触觉提供装置 **903** 和灯光特效呈现装置 **904**。附加地，**901** 还可以实现音频内容数据进行数据转换。可选地，触觉提供装置 **903** 和灯光特效呈现装置 **904** 也可被包括在接收侧设备中。

可选地，接收侧设备中还可以包括姿态获取装置 **902**，其可以获取手指运动数据，并经由 **901** 提供到控制侧设备。尽管图 9 中示出了接收侧设备中仅包括单个姿态获取装置 **902**，并且姿态获取装置 **902** 仅佩戴在一个手指上，但是这仅仅示例性的，而且是可被佩戴在其他手指上，或者佩戴在更多的手指上。

应指出，佩戴在用户上的手套可以相同或者不同。

作为示例，佩戴在演奏者的手上的接收侧设备中，两个或更多个手指都可被配置有姿态捕获装置，从而能够更加准确地检测演奏者的手势，以便更加准确地设定、创建或合成音频内容。作为对比，佩戴在听众手上的接收侧设备，姿态捕获装置可以仅佩戴在一个手指上，而触觉提供装置也可以佩戴在另一个手指上，这样可以简化听众的操作，便于听众使用

应指出，在本公开的实施例中，接收侧设备还可包括集成天线。

附加地，接收侧设备还可以包括电池、数据传递收发器件，例如天线，以及可选的数据处理单元，例如相位器、滤波器等，这里将不再详细描述。

图 10 示出了根据本公开的实施例的演奏者和听众进行交互式音频呈现的示例性实现的示意图。

如图 10 中所示，演奏者进行相对精细的姿态操作，然后基于所获取的姿态信息来设定音乐，例如将姿态转换成音乐，从而以适当方式呈现给演奏者。并且这样的音乐可以呈现给听众。例如在现场场景中。另一方面，听众也可进行姿态操作，特别地，为了可以进行相对简单的姿态操作，然后基于所获取的姿态信息来影响音乐，例如将调整音乐的节奏、旋律等等，从而以适当方式呈现给听众，例如通过声音、触觉、视觉反馈等方式来呈现给听众。当然，这样调整后的音乐也可被呈现演奏者。这样实现了交互式音频呈现。

这样的实现可以体现各种适当的应用场景中。例如在一种音乐吧，尤其是适合于或者可以接待听力受损人士的音乐吧中。在音乐吧运营期间，听众在希望参与时可以

在入口处领取到适当的接收设备，如上所述的手套式设备，然后在音乐吧中的活动中，通过适当地肢体运动，尤其是佩戴有手套式装置的手指摆动，来基于手指摆动所对应的姿态来设定观众收听的音乐。具体如上所述，这里将不再详细描述。

应指出，本公开的技术方案可以应用于各种适当的任务，包括但不限于听力受损环境。在一些实施例中，该任务包括生成音乐并且将音乐以视频方式、音频方式等提供给用户的任务。

在另一些实施例中，本公开的技术方案还可应用于提供其它内容，以其它方式。例如提供音频内容，或者将视频中的音频内容，如电影、电视剧等中的对白，提供给听力受损的用户。在另一些实施例中，本公开的技术方案同样可以用于听力正常用户。

本公开的技术可被用于许多应用。例如，本公开的技术可被用于现场音频呈现应用场合。还可以用于远程音乐会、演奏会等等，同样可以捕获用户的姿态，在云端设定或调整音频内容，然后将音频内容通过网络提供给用户。

在一些实施例中，本公开的方案可以通过软件算法来实现，从而可以方便地集成在包含呈现设备的各种类型的设备中，例如包含各种呈现装置的设备，诸如指套。特别地，本公开的方法可作为计算机程序、指令等由便携设备的处理器来执行，以便进行音频呈现进行增强处理。

另外，应当理解，上述系列处理和设备也可以通过软件和/或固件实现。在通过软件和/或固件实现的情况下，从存储介质或网络向具有专用硬件结构的计算机，例如图 11 所示的通用个人计算机 1100 安装构成该软件的程序，该计算机在安装各种程序时，能够执行各种功能等等。图 11 是示出根据本公开的实施例的中可采用的优化设备的个人计算机的示例结构的框图。在一个例子中，该个人计算机可以对应于根据本公开的上述示例性优化设备。

在图 11 中，中央处理单元（CPU）1101 根据只读存储器（ROM）1102 中存储的程序或从存储部分 1108 加载到随机存取存储器（RAM）1103 的程序执行各种处理。在 RAM 1103 中，也根据需要存储当 CPU 1101 执行各种处理等时所需的数据。

CPU 1101、ROM 1102 和 RAM 1103 经由总线 1104 彼此连接。输入/输出接口 1105 也连接到总线 1104。

下述部件连接到输入/输出接口 1105：输入部分 1106，包括键盘、鼠标等；输出部分 1107，包括显示器，比如阴极射线管（CRT）、液晶显示器（LCD）等，和扬声器等；存储部分 1108，包括硬盘等；和通信部分 1109，包括网络接口卡比如 LAN 卡、

调制解调器等。通信部分 1109 经由网络比如因特网执行通信处理。

根据需要，驱动器 1110 也连接到输入/输出接口 1105。可拆卸介质 1111 比如磁盘、光盘、磁光盘、半导体存储器等等根据需要被安装在驱动器 1110 上，使得从中读出的计算机程序根据需要被安装到存储部分 1108 中。

5 在通过软件实现上述系列处理的情况下，从网络比如因特网或存储介质比如可拆卸介质 1111 安装构成软件的程序。

本领域技术人员应当理解，这种存储介质不局限于图 11 所示的其中存储有程序、与设备相分离地分发以向用户提供程序的可拆卸介质 1111。可拆卸介质 1111 的例子包含磁盘（包含软盘（注册商标））、光盘（包含光盘只读存储器（CD-ROM）和数字通用盘（DVD））、磁光盘（包含迷你盘（MD）（注册商标））和半导体存储器。
10 或者，存储介质可以是 ROM 1102、存储部分 1108 中包含的硬盘等等，其中存有程序，并且与包含它们的设备一起被分发给用户。

应指出，文中所述的方法和设备可被实现为软件、固件、硬件或它们的任何组合。有些组件可例如被实现为在数字信号处理器或者微处理器上运行的软件。其他组件可
15 例如实现为硬件和/或专用集成电路。

另外，可采用多种方式来实行本公开的方法和系统。例如，可通过软件、硬件、固件或它们的任何组合来实行本公开的方法和系统。上文所述的该方法的步骤的顺序仅是说明性的，并且除非另外具体说明，否则本公开的方法的步骤不限于上文具体描述的顺序。此外，在一些实施例，本公开还可具体化为记录介质中记录的程序，包
20 括用于实施根据本公开的方法的机器可读指令。因此，本公开还涵盖了存储用于实施根据本公开的方法的程序的记录介质。这样的存储介质可以包括但不限于软盘、光盘、磁光盘、存储卡、存储棒等等。

本领域技术人员应当意识到，在上述操作之间的边界仅仅是说明性的。多个操作可以结合成单个操作，单个操作可以分布于附加的操作中，并且操作可以在时间上至少部分重叠地执行。而且，另选的实施例可以包括特定操作的多个实例，并且在其他各种实施例中可以改变操作顺序。但是，其它的修改、变化和替换同样是可能的。因此，本说明书和附图应当被看作是说明性的，而非限制性的。

另外，本公开的实施方式还可以包括以下示例性实施例实现（EEE）。

EEE 1、一种用于交互式音频呈现的接收侧设备，所述设备包括处理电路，被配置为：接收来自用于交互式音频呈现的控制侧设备的待呈现音频内容的相关信息，其
30

中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容，以及使得呈现音频内容，其中呈现音频内容包括以触觉方式呈现音频内容。

5 EEE 2、根据 EEE 1 所述的接收侧设备，其中，用户的姿态信息包括所述用户的姿态、姿态运动信息中的至少一者，其中，姿态运动信息包括姿态运动方向、轨迹、运动加速度中的至少一者。

EEE 3、根据 EEE 1 所述的接收侧设备，其中，所述处理电路进一步配置为：获取经由姿态捕捉装置确定的用户的姿态信息，并且将所获取的用户姿态信息发送给控制侧设备。

10 EEE 4、根据 EEE 3 所述的接收侧设备，其中，所述姿态捕获装置包括能够佩戴在用户的至少一个手指上的动作捕捉器件，并且姿态信息是基于佩戴有动作捕捉器件的各手指的姿态信息和/或它们的组合的。

15 EEE 5、根据 EEE 1 所述的接收侧设备，其中，用户的姿态信息包括第一用户的姿态信息，并且其中，待呈现的音频内容包括由第一用户姿态信息所指定的特定音频内容、基于与所述第一用户的姿态信息相对应的音频单元或者特定组合构建的音频内容中的至少一者。

EEE 6、根据 EEE 1 所述的接收侧设备，其中，用户的姿态信息包括第二用户的姿态信息，并且其中，待呈现的音频内容包括基于所述第二用户的姿态信息对音频内容进行调整而获得的音频内容。

20 EEE 7、根据 EEE 6 所述的接收侧设备，其中，基于所述第二用户的姿态信息对音频内容进行调整包括以下中的至少一者：

增大或减小音频内容的音量；

调整音频内容的节奏；

增强音频内容的效果；

为音频内容增加附加效果。

25 EEE 8、根据 EEE 7 所述的接收侧设备，其中，基于所述第二用户的姿态信息对音频内容进行调整包括：

基于多个第二用户的姿态信息的统计值来调整音频内容的呈现。

30 EEE 9、根据 EEE 8 所述的接收侧设备，其中，多个第二用户的姿态信息的统计值包括关于用户姿态信息的优先级的统计值，并且根据多个第二用户的姿态信息中的最高优先级姿态信息来调整音频内容的呈现。

EEE 10、根据 EEE 9 所述的接收侧设备，其中，关于用户姿态信息的优先级的统计值如下地被确定：

对多个音频用户的姿态信息进行加权处理，其中加权处理是基于各姿态信息的数量、各姿态信息的优先级、各姿态信息对应的用户的优先级中的至少一者来进行的。

5 EEE 11、根据 EEE 1 所述的接收侧设备，其中，使得呈现音频内容进一步包括：将音频内容的相关信息转换为适于适用于音频呈现装置的数据；以及将转换得到的数据提供给音频呈现装置。

EEE 12、根据 EEE 1-11 中任一项所述的接收侧设备，其中，所述音频呈现装置为触感提供装置，使得经由所述触感提供装置将音频内容以触觉方式提供给用户。

10 EEE 13、根据 EEE 12 所述的接收侧设备，其中，所述触感提供装置包括至少一个触觉单元，其中每个触觉单元对应于待呈现音频内容中的特定音轨。

EEE 14、一种用于交互式音频呈现的控制侧设备，所述设备包括处理电路，被配置为：获取音频内容呈现指示信息，所述音频内容呈现指示信息包括基于音频要被呈现给的用户的姿态信息的指示信息，以及将待呈现音频内容的相关信息发送给用于音频交互呈现的接收侧设备，其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容。

EEE 15、根据 EEE 14 所述的控制侧设备，其中，所述处理电路进一步配置为：获取用户的姿态信息，以及基于所获取的用户的姿态信息来设定待呈现音频内容。

20 EEE 16、根据 EEE 15 所述的控制侧设备，其中，用户的姿态信息包括第一用户的姿态信息，并且其中，待呈现的音频内容包括由第一用户姿态信息所指定的特定音频内容、基于与所述第一用户的姿态信息相对应的音频单元或者特定组合构建的音频内容中的至少一者，和/或

其中，用户的姿态信息包括第二用户的姿态信息，并且其中，待呈现的音频内容包括基于所述第二用户的姿态信息对音频内容进行调整而获得的音频内容。

25 EEE 17、根据 EEE 16 所述的控制侧设备，其中，所述处理电路进一步配置为：确定多个第二用户的姿态信息的统计值，所述统计值包括关于用户姿态信息的优先级的统计值，并且

根据多个第二用户的姿态信息中的最高优先级姿态信息来设定待呈现音频内容。

30 EEE 18、一种用于交互式音频呈现的接收侧的方法，包括：接收来自用于交互式音频呈现的控制侧设备的待呈现音频内容的相关信息，其中所述待呈现音频内容包括

基于用户的姿态信息设定的音频内容，以及使得呈现音频内容，其中呈现音频内容包括以触觉方式呈现音频内容。

EEE 19、根据 EEE 18 所述的方法，还包括：获取经由姿态捕捉装置确定的用户的姿态信息，并且将所获取的用户姿态信息发送给控制侧设备。

5 EEE 20、根据 EEE 18 所述的方法，其中，使得呈现音频内容进一步包括：将音频内容的相关信息转换为适于适用于音频呈现装置的数据；以及将转换得到的数据提供给音频呈现装置。

EEE 21、一种用于交互式音频呈现的控制侧的方法，包括：获取音频内容呈现指示信息，所述音频内容呈现指示信息包括基于音频要被呈现给的用户姿态信息的指示信息，并且将待呈现音频内容的相关信息发送给用于音频交互呈现的接收侧设备，
10 其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容。

EEE 22、根据 EEE 21 所述的方法，还包括：获取用户的姿态信息，以及基于所获取的用户的姿态信息来设定待呈现音频内容。

EEE 23、根据 EEE 21 所述的方法，还包括：确定多个第二用户的姿态信息的统计值，所述统计值包括关于用户姿态信息的优先级的统计值，并且根据多个第二用户的姿态信息中的最高优先级姿态信息来设定待呈现音频内容。
15

EEE 24、一种交互式音频呈现系统，包括：用于交互式音频呈现的控制侧设备，被配置为接收音频呈现用户的姿态信息，并基于音频呈现用户的姿态信息设定待呈现音频内容；以及用于交互式音频呈现的接收侧设备，被配置为接收待呈现音频内容的相关信息，并且使得呈现音频内容，其中呈现音频内容包括以触觉方式将音频内容呈现给用户。
20

EEE 25、根据 EEE 24 所述的系统，其中，所述接收侧设备进一步配置为：获取用户的姿态信息，并且将所述姿态信息发送至所述控制侧设备。

EEE 26、根据 EEE 24 所述的系统，其中，所述控制侧设备进一步配置为：获取
25 多个用户的姿态信息，并且基于多个用户的姿态信息的统计值设定待呈现音频内容。

EEE 27、一种交互式音频呈现方法，包括：接收音频呈现用户的姿态信息，基于音频呈现用户的姿态信息设定待呈现音频内容；以及使得呈现音频内容，其中呈现音频内容包括以触觉方式将音频内容呈现给用户。

EEE 28、一种设备，包括至少一个处理器；和至少一个存储设备，所述至少一个
30 存储设备在其上存储指令，该指令在由所述至少一个处理器执行时，使所述至少一个

处理器执行根据 **EEE 18-23** 和 **27** 中任一项所述的方法。

EEE 29、一种存储指令的存储介质，该指令在由处理器执行时能使得执行处理器根据 **EEE 18-23** 和 **27** 中任一项所述的方法。

EEE 30、一种计算机程序产品，所述计算机程序产品包含指令，该指令在由处理器执行时能使得处理器执行根据 **EEE 18-23** 和 **27** 中任一项所述的方法。

EEE 31、一种计算机程序，所述计算机程序包含指令，该指令在由处理器执行时能使得处理器执行根据 **EEE 18-23** 和 **27** 中任一项所述的方法。

虽然已经详细说明了本公开及其优点，但是应当理解在不脱离由所附的权利要求所限定的本公开的精神和范围的情况下可以进行各种改变、替代和变换。而且，本公开实施例的术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

虽然已详细描述了本公开的一些具体实施例，但是本领域技术人员应当理解，上述实施例仅是说明性的而不限制本公开的范围。本领域技术人员应该理解，上述实施例可以被组合、修改或替换而不脱离本公开的范围和实质。本公开的范围是通过所附的权利要求限定的。

权 利 要 求

- 1、一种用于交互式音频呈现的接收侧设备，所述设备包括处理电路，被配置为：
接收来自用于交互式音频呈现的控制侧设备的待呈现音频内容的相关信息，其中
所述待呈现音频内容包括基于用户的姿态信息设定的音频内容，以及
使得呈现音频内容，其中呈现音频内容包括以触觉方式呈现音频内容。
- 2、根据权利要求 1 所述的接收侧设备，其中，用户的姿态信息包括所述用户的姿态、姿态运动信息中的至少一者，
其中，姿态运动信息包括姿态运动方向、轨迹、运动加速度中的至少一者。
- 3、根据权利要求 1 或 2 所述的接收侧设备，其中，所述处理电路进一步配置为：
获取经由姿态捕捉装置确定的用户的姿态信息，并且
将所获取的用户姿态信息发送给控制侧设备。
- 4、根据权利要求 3 所述的接收侧设备，其中，所述姿态捕获装置包括能够佩戴在用户的至少一个手指上的动作捕捉器件，并且姿态信息是基于佩戴有动作捕捉器件的各手指的姿态信息和/或它们的组合的。
- 5、根据权利要求 1-4 中任一项所述的接收侧设备，其中，用户的姿态信息包括第一用户的姿态信息，并且
其中，待呈现的音频内容包括由第一用户姿态信息所指定的特定音频内容、基于与所述第一用户的姿态信息相对应的音频单元或者特定组合构建的音频内容中的至少一者。
- 6、根据权利要求 1-5 中任一项所述的接收侧设备，其中，用户的姿态信息包括第二用户的姿态信息，并且
其中，待呈现的音频内容包括基于所述第二用户的姿态信息对音频内容进行调整而获得的音频内容。

7、根据权利要求 6 所述的接收侧设备，其中，基于所述第二用户的姿态信息对音频内容进行调整包括以下中的至少一者：

增大或减小音频内容的音量；
调整音频内容的节奏；
增强音频内容的效果；
为音频内容增加附加效果。

8、根据权利要求 6 或 7 所述的接收侧设备，其中，基于所述第二用户的姿态信息对音频内容进行调整包括：

基于多个第二用户的姿态信息的统计值来调整音频内容的呈现。

9、根据权利要求 8 所述的接收侧设备，其中，多个第二用户的姿态信息的统计值包括关于用户姿态信息的优先级的统计值，并且根据多个第二用户的姿态信息中的最高优先级姿态信息来调整音频内容的呈现。

10、根据权利要求 9 所述的接收侧设备，其中，关于用户姿态信息的优先级的统计值如下地被确定：

对多个音频用户的姿态信息进行加权处理，其中加权处理是基于各姿态信息的数量、各姿态信息的优先级、各姿态信息对应的用户的优先级中的至少一者来进行的。

11、根据权利要求 1-10 中任一项所述的接收侧设备，其中，使得呈现音频内容进一步包括：

将音频内容的相关信息转换为适于适用于音频呈现装置的数据；以及
将转换得到的数据提供给所述音频呈现装置。

12、根据权利要求 11 所述的接收侧设备，其中，所述音频呈现装置为触感提供装置，使得经由所述触感提供装置将音频内容以触觉方式提供给用户。

13、根据权利要求 12 所述的接收侧设备，其中，所述触感提供装置包括至少一个触觉单元，其中每个触觉单元对应于待呈现音频内容中的特定音轨。

14、一种用于交互式音频呈现的控制侧设备，所述设备包括处理电路，被配置为：
获取音频内容呈现指示信息，所述音频内容呈现指示信息包括基于音频要被呈现给的用户的姿态信息的指示信息，以及

将待呈现音频内容的相关信息发送给用于音频交互呈现的接收侧设备，其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容。

15、根据权利要求 14 所述的控制侧设备，其中，所述处理电路进一步配置为：
获取用户的姿态信息，以及
基于所获取的用户的姿态信息来设定待呈现音频内容。

16、根据权利要求 14 或 15 所述的控制侧设备，
其中，用户的姿态信息包括第一用户的姿态信息，并且其中，待呈现的音频内容包括由第一用户姿态信息所指定的特定音频内容、基于与所述第一用户的姿态信息相对应的音频单元或者特定组合构建的音频内容中的至少一者，和/或

其中，用户的姿态信息包括第二用户的姿态信息，并且其中，待呈现的音频内容包括基于所述第二用户的姿态信息对音频内容进行调整而获得的音频内容。

17、根据权利要求 16 所述的控制侧设备，其中，所述处理电路进一步配置为：
确定多个第二用户的姿态信息的统计值，所述统计值包括关于用户姿态信息的优先级的统计值，并且

根据多个第二用户的姿态信息中的最高优先级姿态信息来设定待呈现音频内容。

18、一种用于交互式音频呈现的接收侧的方法，包括：
接收来自用于交互式音频呈现的控制侧设备的待呈现音频内容的相关信息，其中所述待呈现音频内容包括基于用户的姿态信息设定的音频内容，以及
使得呈现音频内容，其中呈现音频内容包括以触觉方式呈现音频内容。

19、根据权利要求 18 所述的方法，还包括：
获取经由姿态捕捉装置确定的用户的姿态信息，并且

将所获取的用户姿态信息发送给控制侧设备。

20、根据权利要求 18 或 19 所述的方法，其中，使得呈现音频内容进一步包括：
将音频内容的相关信息转换为适于适用于音频呈现装置的数据；以及
将转换得到的数据提供给音频呈现装置。

21、一种用于交互式音频呈现的控制侧的方法，包括：
获取音频内容呈现指示信息，所述音频内容呈现指示信息包括基于音频要被呈现
给的用户的姿态信息的指示信息，并且
将待呈现音频内容的相关信息发送给用于音频交互呈现的接收侧设备，其中所述
待呈现音频内容包括基于用户的姿态信息设定的音频内容。

22、根据权利要求 21 所述的方法，还包括：
获取用户的姿态信息，以及
基于所获取的用户的姿态信息来设定待呈现音频内容。

23、根据权利要求 21 或 22 所述的方法，还包括：
确定多个第二用户的姿态信息的统计值，所述统计值包括关于用户姿态信息的优
先级的统计值，并且
根据多个第二用户的姿态信息中的最高优先级姿态信息来设定待呈现音频内容。

24、一种交互式音频呈现系统，包括：
用于交互式音频呈现的控制侧设备，被配置为接收音频呈现用户的姿态信息，并
基于音频呈现用户的姿态信息设定待呈现音频内容；以及
用于交互式音频呈现的接收侧设备，被配置为接收待呈现音频内容的相关信息，
并且使得呈现音频内容，其中呈现音频内容包括以触觉方式将音频内容呈现给用户。

25、根据权利要求 24 所述的系统，其中，所述接收侧设备进一步配置为：获取
用户的姿态信息，并且将所述姿态信息发送至所述控制侧设备。

26、根据权利要求 24 或 25 所述的系统，其中，所述控制侧设备进一步配置为：获取多个用户的姿态信息，并且基于多个用户的姿态信息的统计值设定待呈现音频内容。

27、一种交互式音频呈现方法，包括：接收音频呈现用户的姿态信息，基于音频呈现用户的姿态信息设定待呈现音频内容；以及使得呈现音频内容，其中呈现音频内容包括以触觉方式将音频内容呈现给用户。

28、一种设备，包括：

至少一个处理器；和

至少一个存储设备，所述至少一个存储设备在其上存储指令，该指令在由所述至少一个处理器执行时，使所述至少一个处理器执行根据权利要求 18-23 和 27 中任一项所述的方法。

29、一种存储指令的存储介质，该指令在由处理器执行时能使得执行处理器根据权利要求 18-23 和 27 中任一项所述的方法。

30、一种计算机程序产品，所述计算机程序产品包含指令，该指令在由处理器执行时能使得处理器执行根据权利要求 18-23 和 27 中任一项所述的方法。

31、一种计算机程序，所述计算机程序包含指令，该指令在由处理器执行时能使得处理器执行根据权利要求 18-23 和 27 中任一项所述的方法。

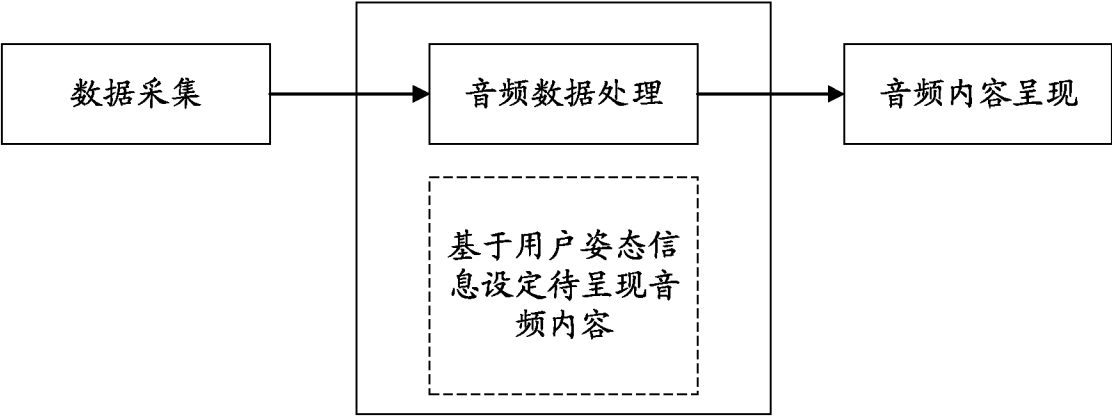


图 1

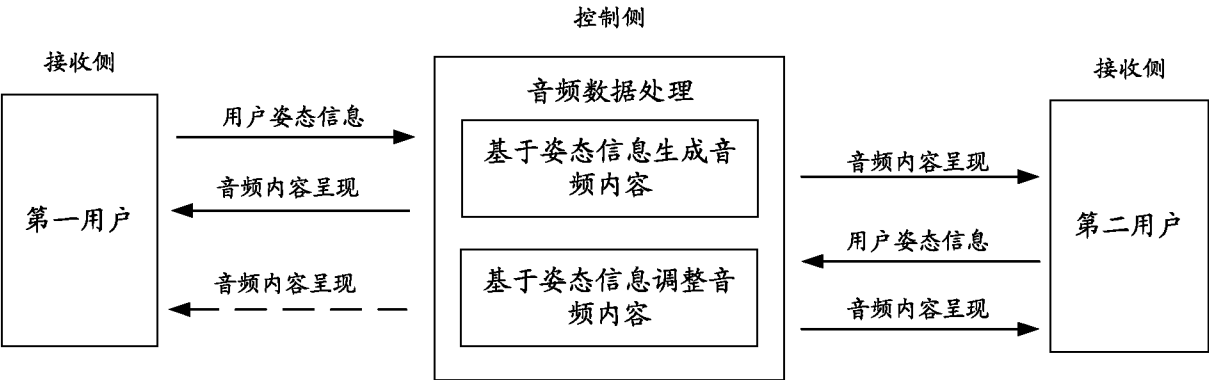


图 2A

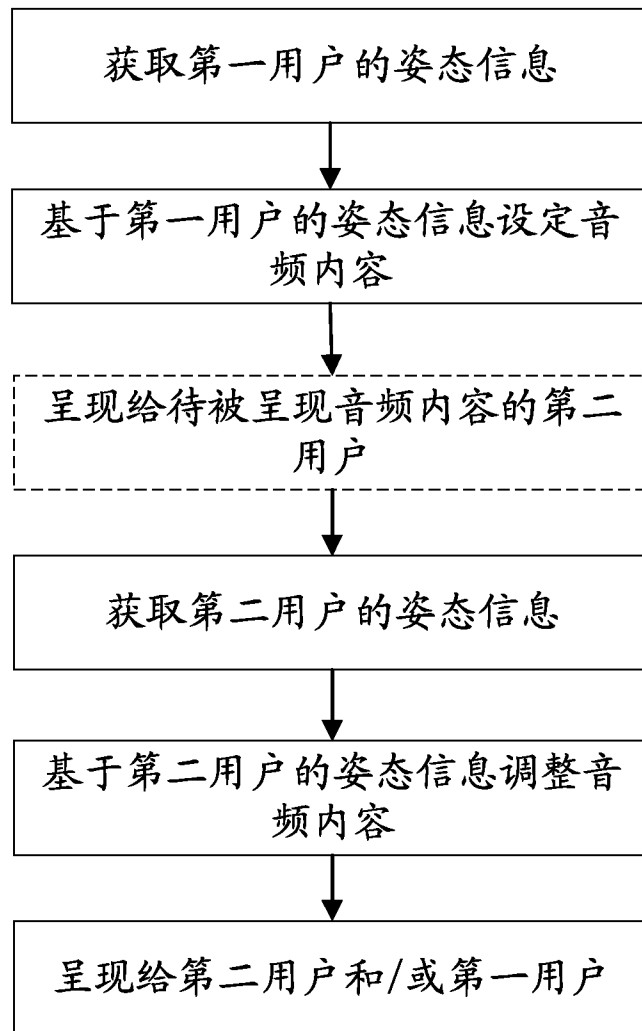


图 2B

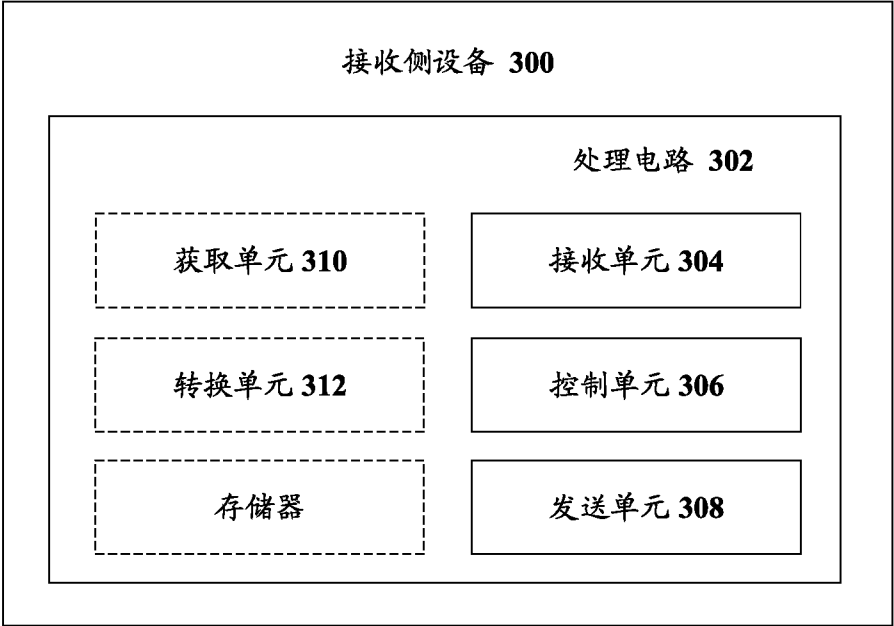


图 3A

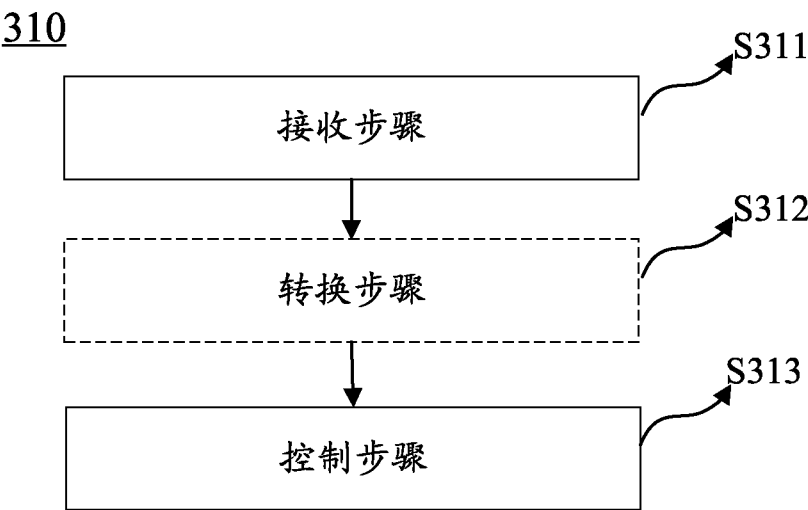


图 3B

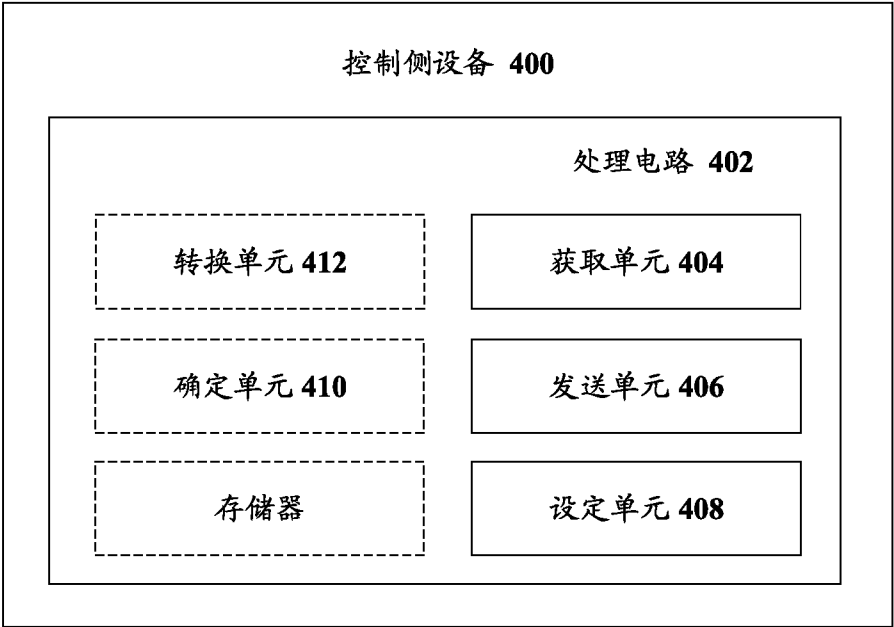


图 4A

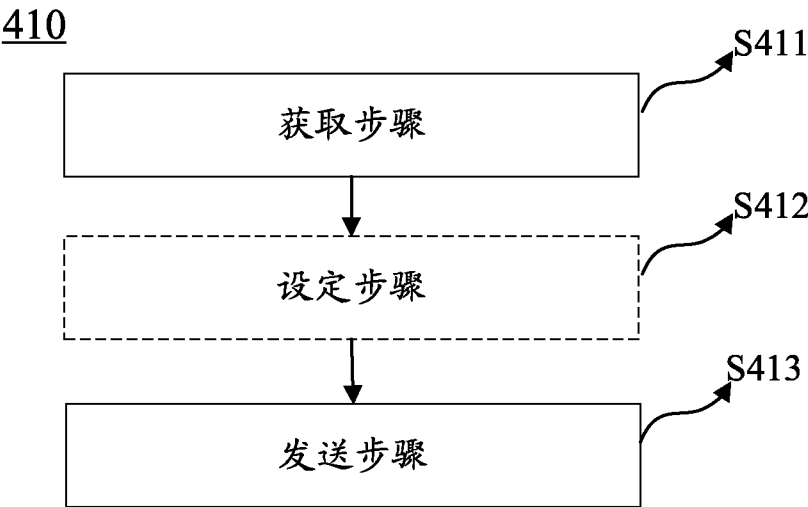


图 4B

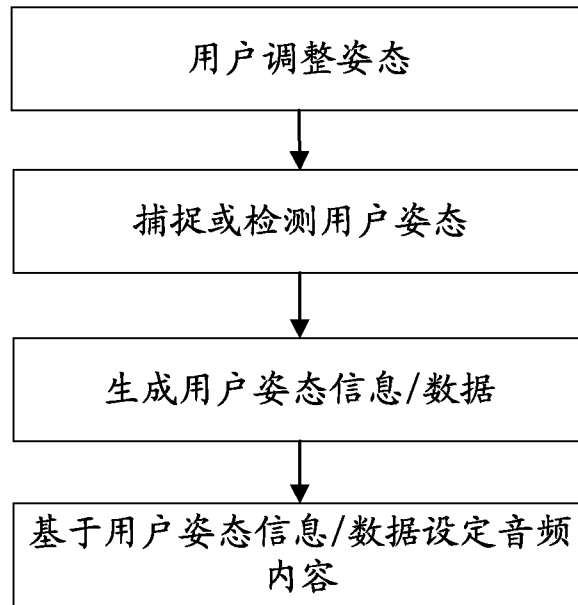


图 5

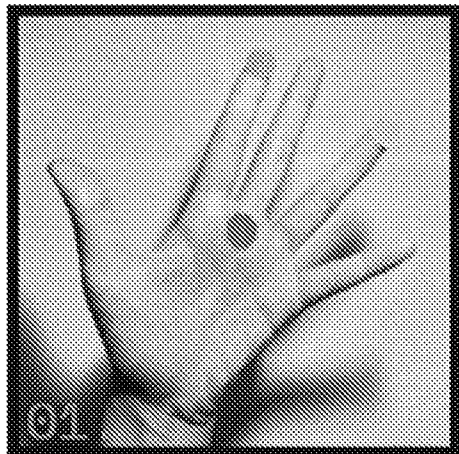


图 6A

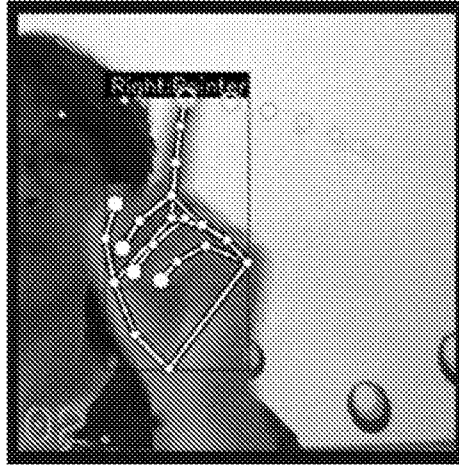


图 6B

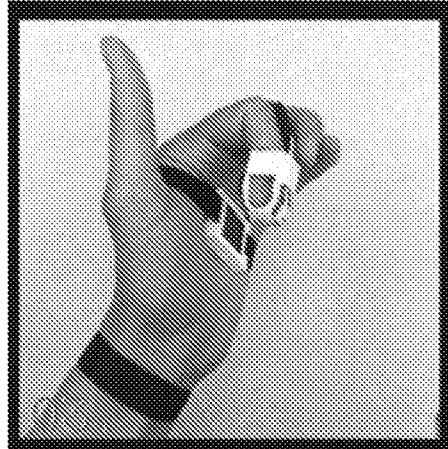


图 6C

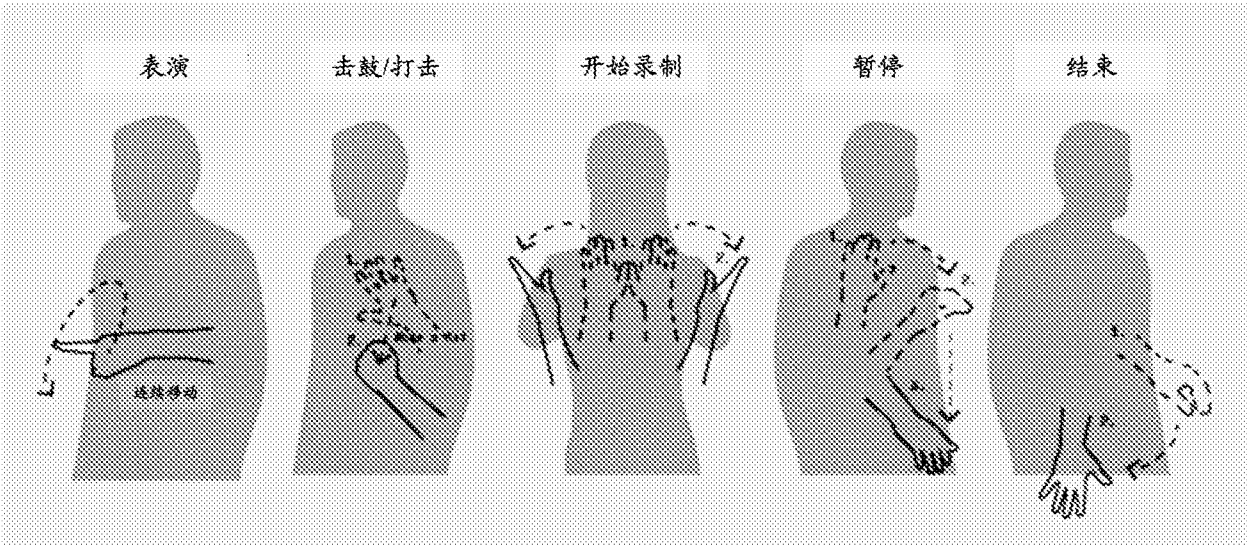


图 7A

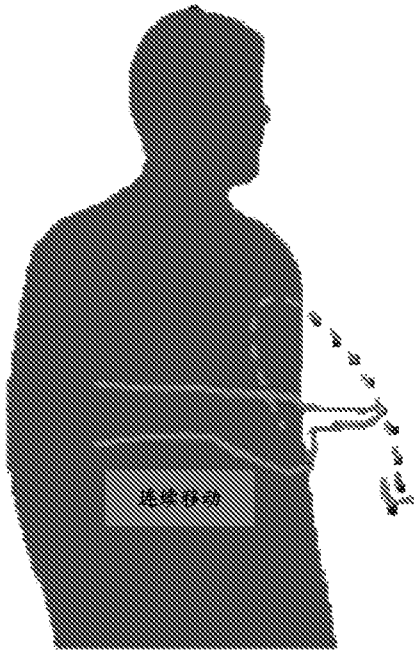


图 7B

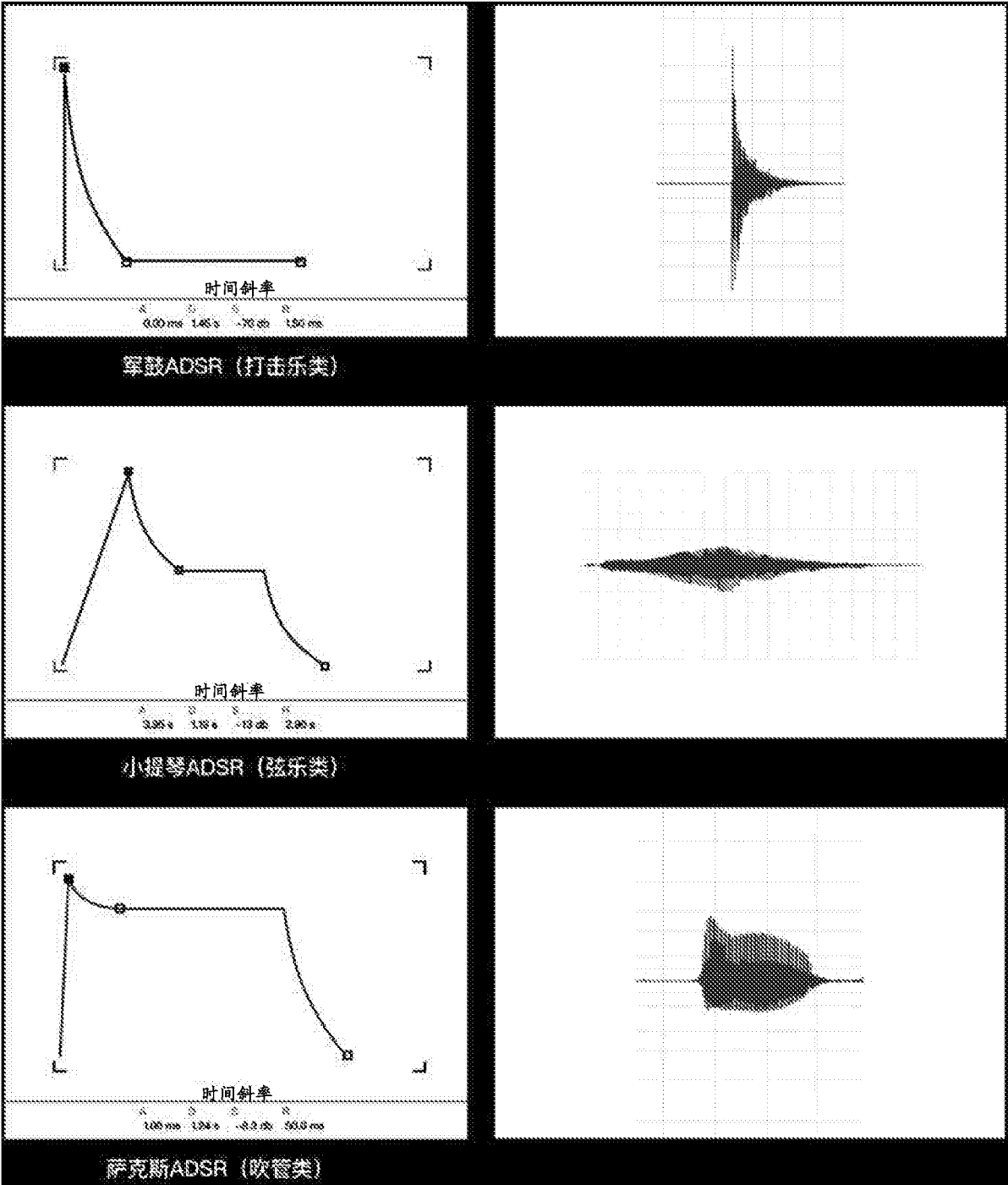


图 8

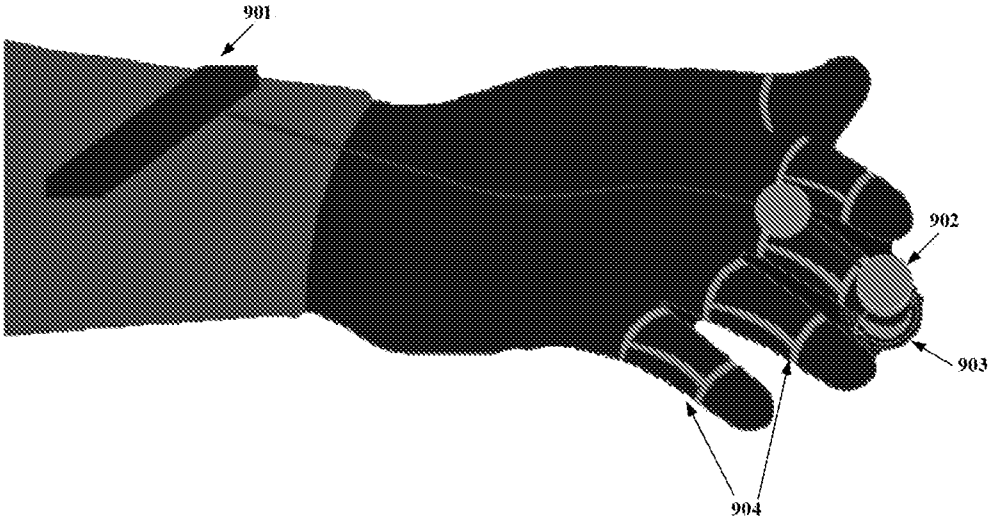


图 9

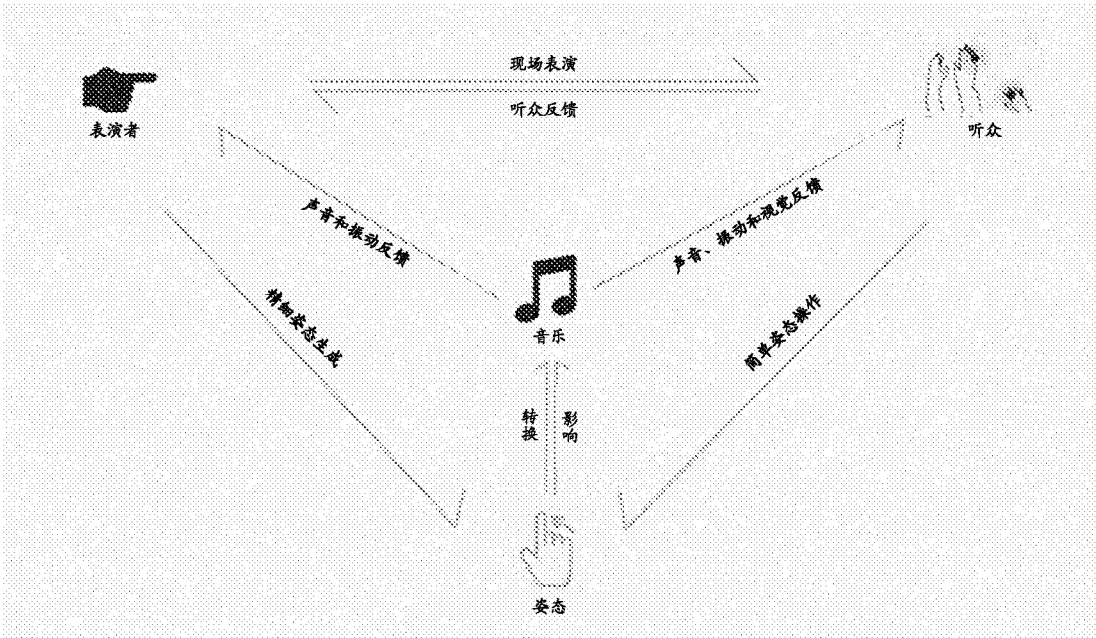


图 10

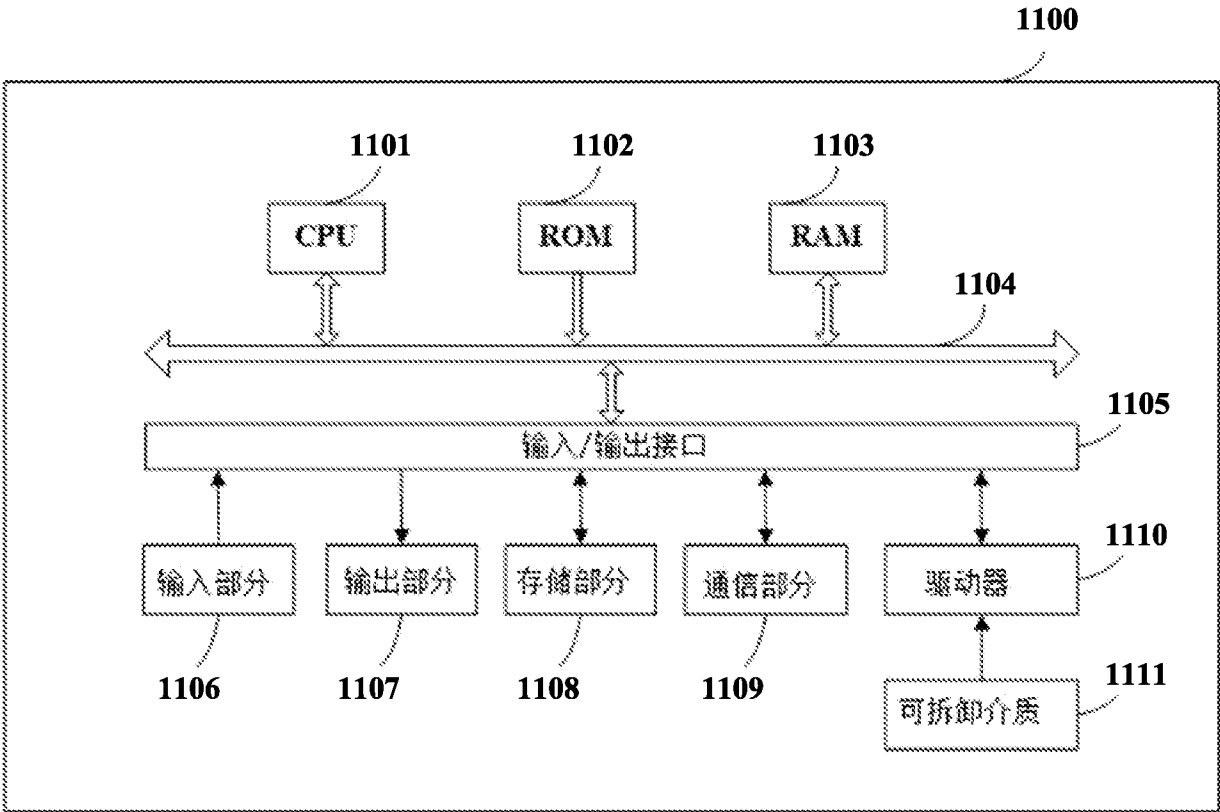


图 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/138019

A. CLASSIFICATION OF SUBJECT MATTER

G06F 3/01(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC:G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

VEN, CNABS, CNTXT, WOTXT, EPTXT, USTXT, CNKI, IEEE: 播放, 触感, 触觉, 动作, 耳蜗, 感官, 感知, 歌曲, 观众, 交互, 接触觉, 接近觉, 力觉, 手势, 听众, 停止, 统计, 音乐, 音量, 音频, 音视频, 音效, 影音, 暂停, 震动, 振动, 知觉, 姿态, play, haptics, touch, music, audience, listener, statistics, gesture, attitude

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2011228962 A1 (NATIONAL UNIVERSITY OF SINGAPORE) 22 September 2011 (2011-09-22) claims 1-21, and description, paragraphs [0031]-[0115]	1-7, 11-16, 18-22, 24-25, 27-31
A	CN 112817557 A (HISENSE VISUAL TECHNOLOGY CO., LTD.) 18 May 2021 (2021-05-18) entire document	1-31
A	US 2015103154 A1 (SONY CORP.) 16 April 2015 (2015-04-16) entire document	1-31
A	US 2019373355 A1 (BOSE CORP.) 05 December 2019 (2019-12-05) entire document	1-31



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
 “D” document cited by the applicant in the international application
 “E” earlier application or patent but published on or after the international filing date
 “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 “O” document referring to an oral disclosure, use, exhibition or other means
 “P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

21 February 2024

Date of mailing of the international search report

26 February 2024

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
China No. 6, Xitucheng Road, Jimenqiao, Haidian District,
Beijing 100088

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/138019

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)		Publication date (day/month/year)	
US	2011228962	A1	22 September 2011	US	8638966	B2	28 January 2014
				WO	2010033086	A1	25 March 2010
				SG	194367	A1	29 November 2013
				CA	2738746	A1	25 March 2010
				CA	2738746	C	10 November 2015

CN	112817557	A	18 May 2021	None			

US	2015103154	A1	16 April 2015	None			

US	2019373355	A1	05 December 2019	EP	3803537	A1	14 April 2021
				US	10820086	B2	27 October 2020
				WO	2019231752	A1	05 December 2019

A. 主题的分类 G06F 3/01(2006.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类																	
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) IPC:G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) VEN, CNABS, CNTXT, WOTXT, EPTXT, USTXT, CNKI, IEEE:播放,触感,触觉,动作,耳蜗,感官,感知,歌曲,观众,交互,接触觉,接近觉,力觉,手势,听众,停止,统计,音乐,音量,音频,音视频,音效,影音,暂停,震动,振动,知觉,姿态,play, haptics,touch,music,audience,listener,statistics,gesture,attitude																	
C. 相关文件 <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>US 2011228962 A1 (NATIONAL UNIVERSITY OF SINGAPORE) 2011年9月22日 (2011 - 09 - 22) 权利要求1-21, 说明书第[0031]段-[0115]段</td> <td>1-7、11-16、18-22、24-25、27-31</td> </tr> <tr> <td>A</td> <td>CN 112817557 A (海信视像科技股份有限公司) 2021年5月18日 (2021 - 05 - 18) 全文</td> <td>1-31</td> </tr> <tr> <td>A</td> <td>US 2015103154 A1 (SONY CORPORATION) 2015年4月16日 (2015 - 04 - 16) 全文</td> <td>1-31</td> </tr> <tr> <td>A</td> <td>US 2019373355 A1 (BOSE CORPORATION) 2019年12月5日 (2019 - 12 - 05) 全文</td> <td>1-31</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	US 2011228962 A1 (NATIONAL UNIVERSITY OF SINGAPORE) 2011年9月22日 (2011 - 09 - 22) 权利要求1-21, 说明书第[0031]段-[0115]段	1-7、11-16、18-22、24-25、27-31	A	CN 112817557 A (海信视像科技股份有限公司) 2021年5月18日 (2021 - 05 - 18) 全文	1-31	A	US 2015103154 A1 (SONY CORPORATION) 2015年4月16日 (2015 - 04 - 16) 全文	1-31	A	US 2019373355 A1 (BOSE CORPORATION) 2019年12月5日 (2019 - 12 - 05) 全文	1-31
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
X	US 2011228962 A1 (NATIONAL UNIVERSITY OF SINGAPORE) 2011年9月22日 (2011 - 09 - 22) 权利要求1-21, 说明书第[0031]段-[0115]段	1-7、11-16、18-22、24-25、27-31															
A	CN 112817557 A (海信视像科技股份有限公司) 2021年5月18日 (2021 - 05 - 18) 全文	1-31															
A	US 2015103154 A1 (SONY CORPORATION) 2015年4月16日 (2015 - 04 - 16) 全文	1-31															
A	US 2019373355 A1 (BOSE CORPORATION) 2019年12月5日 (2019 - 12 - 05) 全文	1-31															
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。 ★ 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “D” 申请人在国际申请中引证的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件																	
国际检索实际完成的日期 2024年2月21日		国际检索报告邮寄日期 2024年2月26日															
ISA/CN的名称和邮寄地址 中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088		授权官员 李文浩 电话号码 (+86) 010-53961341															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2023/138019

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
US	2011228962	A1	2011年9月22日	US 8638966 B2	2014年1月28日
				WO 2010033086 A1	2010年3月25日
				SG 194367 A1	2013年11月29日
				CA 2738746 A1	2010年3月25日
				CA 2738746 C	2015年11月10日
CN	112817557	A	2021年5月18日	无	
US	2015103154	A1	2015年4月16日	无	
US	2019373355	A1	2019年12月5日	EP 3803537 A1	2021年4月14日
				US 10820086 B2	2020年10月27日
				WO 2019231752 A1	2019年12月5日