

(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 7 月 18 日 (18.07.2024)



(10) 国际公布号
WO 2024/149352 A1

- (51) 国际专利分类号:
G10L 15/22 (2006.01)
- (21) 国际申请号: PCT/CN2024/071940
- (22) 国际申请日: 2024 年 1 月 12 日 (12.01.2024)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202310096785.6 2023年1月12日 (12.01.2023) CN
- (71) 申请人: 杭州灵伴科技有限公司
(**HANGZHOU LINGBAN TECHNOLOGY CO., LTD.**)
[CN/CN]; 中国浙江省杭州市余杭区仓前街道良睦路1288号8幢101, Zhejiang 310000 (CN)。
- (72) 发明人: 付济强 (**FU, Jiqiang**); 中国浙江省杭州市余杭区仓前街道良睦路1288号8幢101, Zhejiang

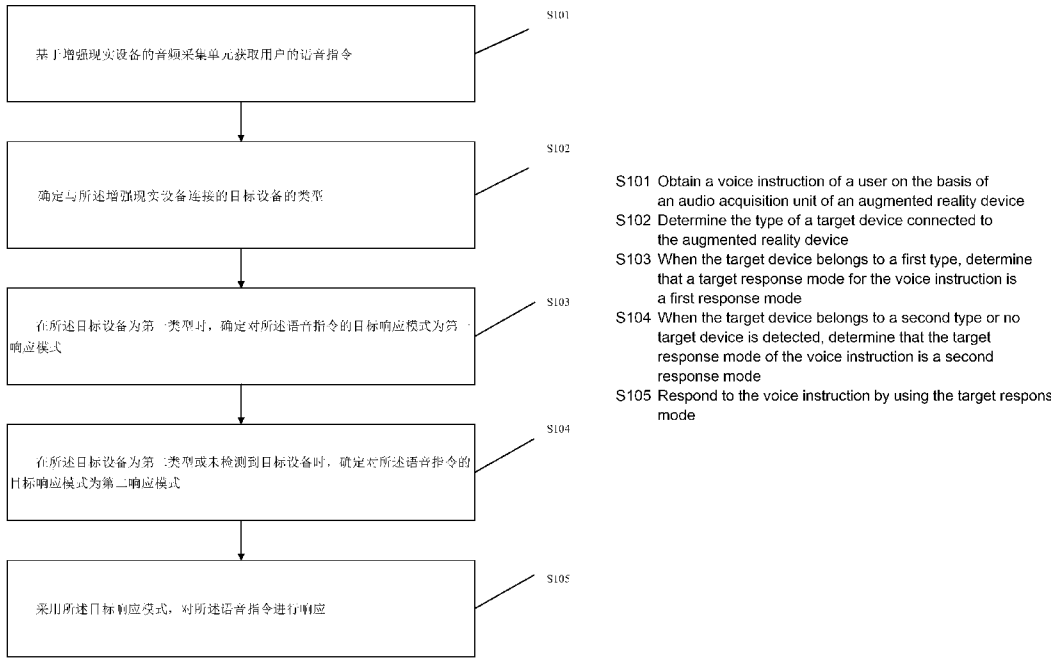
310000 (CN)。刘生华 (**LIU, Shenghua**); 中国浙江省杭州市余杭区仓前街道良睦路1288号8幢101, Zhejiang 310000 (CN)。王文兵 (**WANG, Wenbing**); 中国浙江省杭州市余杭区仓前街道良睦路1288号8幢101, Zhejiang 310000 (CN)。

(74) 代理人: 北京乐知新创知识产权代理事务所 (普通合伙) (**BEIJING LEPATENT INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)**); 中国北京市昌平区北七家镇七北路42号院TBD云集中心4号楼C座6层602室, Beijing 102200 (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,

(54) **Title:** VOICE INTERACTION METHOD AND APPARATUS, AND RELATED DEVICE

(54) 发明名称: 语音交互方法、装置及相关设备



(57) **Abstract:** The present application provides a voice interaction method and apparatus, and a related device. The method comprises: obtaining a voice instruction of a user on the basis of an audio acquisition unit of an augmented reality device; determining the type of a target device connected to the augmented reality device; when the target device belongs to a first type, determining that a target response mode for the voice instruction is a first response mode; when the target device belongs to a second type or no target device is detected, determining that the target response mode of the voice instruction is a second response mode; and responding to the voice instruction by using the target response mode. Technical support can be provided for realizing normal voice interaction under the scenario in which an augmented reality device is not connected to a target device or is connected to different target devices.



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)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本申请提供了一种语音交互方法、装置及相关设备, 其中所述方法包括: 基于增强现实设备的音频采集单元获取用户的语音指令; 确定与所述增强现实设备连接的目标设备的类型; 在所述目标设备为第一类型时, 确定对所述语音指令的目标响应模式为第一响应模式; 在所述目标设备为第二类型或未检测到目标设备时, 确定对所述语音指令的目标响应模式为第二响应模式; 采用所述目标响应模式, 对所述语音指令进行响应。为实现增强现实设备在不连接目标设备, 或连接不同目标设备的场景下都能正常进行语音交互提供了技术支持。

语音交互方法、装置及相关设备

5 相关申请的交叉引用

本申请要求于 2023 年 1 月 12 日提交中国专利局,申请号为 202310096785.6 的中国专利申请的优先权,其全部内容通过引用结合在本申请中。

技术领域

10 本申请涉及语音交互领域,尤其涉及一种语音交互方法、装置及相关设备。

背景技术

通常,增强现实设备用做显示或采集传感器数据,多数情况下需能连接到与增强现实设备适配的终端设备才能使用。

15

发明内容

本申请提供了一种语音交互方法、装置及相关设备,以至少解决现有技术中存在的以上技术问题。

根据本申请的第一方面,提供了一种语音交互方法,所述方法包括:

20

基于增强现实设备的音频采集单元获取用户的语音指令;

确定与所述增强现实设备连接的目标设备的类型;

在所述目标设备为第一类型时,确定对所述语音指令的目标响应模式为第一响应模式;

25

在所述目标设备为第二类型或未检测到目标设备时,确定对所述语音指令的目标响应模式为第二响应模式;

采用所述目标响应模式,对所述语音指令进行响应。

上述方案中,所述采用所述目标响应模式,对所述语音指令进行响应,包括:

30

在所述目标响应模式为第一响应模式时,将所述语音指令交由所述目标设备响应;

在所述目标响应模式为第二响应模式时,基于所述语音指令的类型,确定

由所述增强现实设备和所述目标设备中的其中之一进行响应。

上述方案中，所述基于所述语音指令的类型，确定由所述增强现实设备和所述目标设备中的其中之一进行响应，包括：

5 在所述语音指令为第一类型指令时，由所述增强现实设备对所述语音指令进行响应；

在所述语音指令为第二类型指令时，将所述语音指令交由所述目标设备响应。

上述方案中，所述在所述语音指令为第二类型指令时，将所述语音指令交由所述目标设备响应，包括：

10 在所述语音指令为第二类型指令时，将所述语音指令转换成目标类型指令，将所述目标类型指令交由所述目标设备响应。

上述方案中，所述音频采集单元用于对目标语音数据进行采集，所述目标语音数据包括语音指令；

所述方法还包括：

15 通过确定所述目标语音数据中是否存在语音指令而获取所述语音指令；
和/或，将所述目标语音数据交由所述目标设备。

上述方案中，所述通过确定所述目标语音数据中是否存在语音指令而获取所述语音指令；和/或，将所述目标语音数据交由所述目标设备，包括：

将所述目标语音数据进行降噪处理，得到降噪后的目标语音数据；

20 通过确定所述降噪后的目标语音数据中是否存在语音指令而获取所述语音指令；

和/或，将所述降噪后的目标语音数据交由所述目标设备。

上述方案中，所述通过确定所述降噪后的目标语音数据中是否存在语音指令而获取所述语音指令；和/或，将所述降噪后的目标语音数据交由所述目标设备，包括：

25 基于降噪后的目标语音数据，得到第一目标语音数据和第二目标语音数据；
通过确定所述第一目标语音数据中是否存在语音指令而获取所述语音指令；
和/或，将所述第二目标语音数据交由所述目标设备。

根据本申请的第二方面，提供了一种语音交互装置，所述装置包括：

30 获取单元，用于基于增强现实设备的音频采集单元获取用户的语音指令；

第一确定单元，用于确定与所述增强现实设备连接的目标设备的类型；

第二确定单元，用于在所述目标设备为第一类型时，确定对所述语音指令的目标响应模式为第一响应模式；

第三确定单元，用于在所述目标设备为第二类型或未检测到目标设备时，确定对所述语音指令的目标响应模式为第二响应模式；

响应单元，用于采用所述目标响应模式，对所述语音指令进行响应。

根据本申请的第三方面，提供了一种增强现实设备，所述增强现实设备至少包括本申请所述的语音交互装置。

根据本申请的第四方面，提供了一种电子设备，包括：

至少一个处理器；以及

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行本申请所述的方法。

本申请中，基于增强现实设备的音频采集单元获取用户的语音指令，确定与增强现实设备连接的目标设备的类型，在目标设备为第一类型时，确定对语音指令的目标响应模式为第一响应模式；在目标设备为第二类型或未检测到目标设备时，确定对语音指令的目标响应模式为第二响应模式，采用目标响应模式，对语音指令进行响应。为实现增强现实设备在不连接目标设备，或连接不同目标设备的场景下都能正常进行语音交互提供了技术支持。

应当理解，本部分所描述的内容并非旨在标识本申请的实施例的关键或重要特征，也不用于限制本申请的范围。本申请的其它特征将通过以下的说明书而变得容易理解。

附图说明

通过参考附图阅读下文的详细描述，本申请示例性实施方式的上述以及其他目的、特征和优点将变得易于理解。在附图中，以示例性而非限制性的方式示出了本申请的若干实施方式，其中：

在附图中，相同或对应的标号表示相同或对应的部分。

图 1 示出了本申请实施例语音交互方法的实现流程示意图。

图 2 示出了本申请实施例不同目标响应模式的实现流程示意图一。

图 3 示出了本申请实施例不同目标响应模式的实现流程示意图二。

图 4 示出了本申请实施例增强现实设备端数据流示意图。

图 5 示出了本申请实施例中语音交互装置的组成结构示意图。

图 6 示出了本申请实施例一种电子设备的组成结构示意图。

具体实施方式

为使本申请的目的、特征、优点能够更加的明显和易懂，下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本申请一部分实施例，而非全部实施例。基于本申请中的实施例，本领域技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

相关技术中，增强现实设备仅能和与其适配的终端设备进行通信，无法体现增强现实设备的灵活性和多功能性。

可以理解，增强现实（AR，Augmented Reality）设备是现在主流的可穿戴式设备，智能语音交互作为增强现实设备的主流交互方式，可以做到解放双手，轻松快速的完成对增强现实设备的输入或操控。考虑到外观、佩戴、功耗、发热等因素影响，增强现实设备通常不作为复杂计算单元，仅用作显示（投影）及传感器数据采集功能（如图像、音频、惯性测量单元等）。基于此，增强现实设备只能通过有线或无线的方式连接到与之适配的终端设备，将采集到的传感器数据交给终端设备做算法计算。如果能够实现增强现实设备在不连接终端设备或连接其他终端设备时也可以正常进行语音交互，势必会扩大增强现实设备的功能性。如此，可以为增强现实设备在日常生活中的广泛应用奠定基础。

本申请实施例的技术方案涉及到语音交互的方案，增强现实设备可与不同类型的目标设备连接进行语音交互，也可以在不连接目标设备的情况下进行语音交互，体现了增强现实设备的多功能性和灵活性。基于获取到的语音指令和与增强现实设备连接的目标设备类型，即可采用不同的目标响应模式对语音指令进行响应。为实现增强现实设备在不连接目标设备，或连接不同目标设备的场景下都能正常进行语音交互提供了技术支持，扩大了增强现实设备的使用场景。

下面对本申请实施例的语音交互方法做详细说明。

本申请实施例提供一种语音交互方法，如图 1 所示，所述方法包括：

S101：基于增强现实设备的音频采集单元获取用户的语音指令。

本步骤中，增强现实设备为可以进行 AR 交互的电子设备。增强现实设备可以为智能穿戴设备，包括但不限于智能眼镜、智能手表。本申请中，以增强现实设备为分体式 AR 眼镜为例进行说明。

增强现实设备包括音频采集单元，如麦克风。本步骤中，通过麦克风对用户向增强现实设备发出的语音指令进行采集的方式，而获取用户的语音指令。

可以理解，麦克风包含有麦克风阵列传感器（MicArray），用于采集用户向增强现实设备发出的语音指令。

在实际应用中，用户可以在增强现实设备不被使用时，向增强现实设备发出语音指令。如，在增强现实设备处于黑屏时，向其发出诸如“请亮屏”的语音指令或“请关机”的语音指令。

用户还可以在增强现实设备被使用时，如利用增强现实设备投放电影或播放歌曲等音频时，向增强现实设备发出语音指令。即，在利用增强现实设备输出多媒体数据的情况下，基于增强现实设备的音频采集单元获取用户的语音指令。

可以理解，增强现实设备通常会在进行 AR 交互的过程中输出一些多媒体数据，如图像、音频等。增强现实设备在与目标设备连接的情况下，可以将目标设备中的视频投影至增强现实设备进行输出，如将目标设备中的电影投影至增强现实设备进行输出。在这种情况下，多媒体数据可指目标设备中可被 AR 设备进行投影的视频。还可以，将目标设备中的音频通过增强现实设备进行输出，如通过增强现实设备接听通话语音或语音电话。这种情况下，多媒体数据可指目标设备中可被 AR 设备进行输出的音频。

即，在与目标设备进行连接的情况下，本申请中的增强现实设备作为代替目标设备进行音视频输出的设备。这种代替方案主要考虑到，在某些应用场景中，采用目标设备作为音视频输出设备远没有采用增强现实设备作为音视频输出设备方便、输出效果好。如，在投影的应用场景中，通过增强现实设备对目标设备中的可投影视频进行投影，可令增强现实设备的佩戴者感受到身临其境的效果。还如，在一些接听环境拥挤的场景中，如地铁的早晚高峰情形中，由于地铁中的拥挤不方便接听目标设备的来电，可通过佩戴的分体式 AR 眼镜进行来电的接听。分体式 AR 眼镜可内置在佩戴者佩戴的普通眼镜上，通过分体式 AR 眼镜进行来电接听，可避免在拥挤环境中无法从口袋中掏出手机接听来电情况的发生。

用户可在需要的情形下，向增强现实设备发出语音指令。增强现实设备，具体是麦克风采集语音指令。

示例性地，在增强现实设备为分体式 AR 眼镜的情况下，假定当前场景为用户使用分体式 AR 眼镜进行音视频播放，此时分体式 AR 眼镜输出的多媒体数据为用户播放的音视频信息，当用户发出语音指令“调高音量”，通过麦克风采集用户发出的这一语音指令即可获得到针对音视频播放的语音指令，以通过对语音指令的响应，提高当前音视频播放的音量。

S102：确定与所述增强现实设备连接的目标设备的类型。

本步骤中，目标设备可以是任何与增强现实设备进行语音交互的设备。如智能手机、电脑、个人数字助理等设备。本申请中的增强现实设备连接的目标设备可以是不同类型的终端。如目标设备是自研终端，即，与增强现实设备相适配的终端。可以理解，增强现实设备在被厂商生产出来之后，通常都会有厂家生产的与该增强现实设备相适配的自研终端。该自研终端可以理解为一个没有显示屏幕或显示屏幕较小，但具有计算能力的正常大小的终端。通过将增强现实设备与适配的自研终端连接来实现增强现实设备的智能语音交互功能。

目标设备也可以是其他类型终端，如第三方手机、或者第三方电脑。本申请中，通过将增强现实设备与其他类型终端连接来实现增强现实设备的智能语音交互功能。示例性地，当增强现实设备为分体式 AR 眼镜时，分体式 AR 眼镜可以和与其适配的终端进行连接，还可以和第三方终端进行连接。本申请中，无论是和与其适配的终端进行连接，还是和第三方终端进行连接，分体式 AR 眼镜均可和与分体式 AR 眼镜连接的设备进行智能语音交互。

在实际应用中，语音交互作为增强现实设备主流的交互方式，通常受限在自研终端上。即，用户要想使用增强现实设备，就得购买与之适配的自研终端才可以正常进行语音交互。在这种情况下，一方面，用户在已具备移动手机情况下，出于成本和便携性考虑，不会再额外购买适配的终端。另一方面，增强现实设备在接入个人电脑（PC，Personal Computer）场景下，PC 作为计算单元，也不会再接入终端。在前述两个场景下增强现实设备的语音交互功能就会变的不可用，大大限制了语音交互在 AR 眼镜下的使用场景。

本申请中，增强现实设备可接入不同类型的终端，并通过确定与增强现实设备连接的目标设备的类型，实现增强现实设备的语音交互功能。增强现实设备也可以不接入任何类型的终端，通过自身语音指令控制应用实现一些基础的语音交互功能，如调节音量、调节亮度等。

在实际应用中，增强现实设备可接入不同类型的终端，也可以不接入任何终端，使得用户可以选择只购买增强现实设备，不用再去购买与之适配的自研终端，直接与自己的手机或者电脑进行连接也可以实现增强现实设备的语音交互功能。

S103：在所述目标设备为第一类型时，确定对所述语音指令的目标响应模式为第一响应模式。

本申请中，增强现实设备可接入或连接不同类型的终端，增强现实设备连接的目标设备类型不同，对语音指令的响应模式也会有所不同。

本申请中目标设备的类型包括第一类型和第二类型，第一类型为内部包含语音关键词检测技术服务的终端，如前述的自研终端。第二类型为内部不包含语音关键词检测技术服务的终端，如前述的第三方手机端、电脑端等终端。

示例性地，假定增强现实设备连接的目标设备为前述的自研终端，那与之对应的对语音指令的目标响应模式可以为模式 A（第一响应模式）。假定增强现实设备连接的目标设备为前述的其它类型终端（第三方手机端、电脑端等），那与之对应的对语音指令的目标响应模式可以为模式 B（第二响应模式）。

在实施时，增强现实设备对自身是否连接有目标设备进行识别。如果没有连接设备，则确定对所述语音指令的目标响应模式为第二响应模式。如果连接有设备，获得所连接的设备的标识，基于所连接设备的标识，确定其所连接的设备为第一类型还是为第二类型。如果所连接的设备的标识为标识 A，且标识 A 是表征第一类型的设备的标识，则确定其所连接的设备为第一类型。如果所连接的设备的标识为标识 B，且标识 B 是表征第二类型的设备的标识时，确定其所连接的设备为第二类型。

本申请中，基于与增强现实设备连接的终端是与增强现实设备适配的终端，还是不与增强现实设备适配的第三方终端，预先设置有两种响应模式。其中一种是在与增强现实设备连接的终端是与增强现实设备适配的终端时使用的模式。另一种是在与增强现实设备连接的终端是第三方终端时使用的模式。预先将两种不同类型终端和各类型终端下需要使用的模式设置为一种对应关系。在实施时，基于与增强现实设备连接的终端的类型，在对应关系中查找该类型终端对应的模式，作为对语音指令进行响应的目标响应模式。

S104：在所述目标设备为第二类型或未检测到目标设备时，确定对所述语音指令的目标响应模式为第二响应模式。

由于第二类型的目标设备是不包含语音关键词检测技术服务的终端，因此为保证增强现实设备在接入第二类型的目标设备时仍能进行语音交互，本申请增强现实设备中需运用语音关键词检测技术服务，来实现增强现实设备在连接第二类型目标设备时的正常语音交互。同时，由于本申请增强现实设备中包括语音关键词检测技术服务，因此当增强现实设备没有与目标设备进行连接，即未检测到目标设备时，增强现实设备也可以采用对应的目标响应模式进行基础的语音交互，如调节亮度、调节音量等。

本申请中，在增强现实设备中部署一种低功耗的语音关键词检测技术服务，在不明显增加增强现实设备的功耗或发热的情况下，借助低功耗语音关键词检测技术服务，完成对语音指令的识别。

通过对不同类型的终端实施不同的目标响应模式，满足了增强现实设备在接入不同类型终端下均能保持语音交互的能力。

S105：采用所述目标响应模式，对所述语音指令进行响应。

通过连接的目标设备类型的不同，采用不同的目标响应模式对语音指令进行响应，即采用不同的目标响应模式，分析语音指令是何种指令并对该种指令进行处理。示例性地，分析语音指令是“提高音量”的语音指令，调高增强现实设备的输出音量。分析语音指令是“降低屏幕亮度”的指令，将增强现实设备的屏幕亮度降低。以实现语音指令对应的操作。

在 S101~S105 所示的方案中，增强现实设备不仅可与适配的终端以及第三方终端进行通信，还可以在不接入任何终端的情况下进行语音交互。体现了增强现实设备的灵活性和多功能性。

且本申请中，可基于与增强现实设备连接的目标设备的类型，确定目标响应模式，并利用目标响应模式，实现对语音指令的响应。基于与增强现实设备连接的目标设备类型，即可采用不同的目标响应模式对语音指令进行响应。实现了增强现实设备在接入不同类型终端的情况下保持语音交互能力。为实现增强现实设备在连接不同目标设备的场景下都能正常进行语音交互提供了技术支持。

在一个可选的方案中，所述采用所述目标响应模式，对所述语音指令进行响应，包括：

在所述目标响应模式为第一响应模式时，将所述语音指令交由所述目标设备响应；

在所述目标响应模式为第二响应模式时，基于所述语音指令的类型，确定由所述增强现实设备和所述目标设备中的其中之一进行响应。

其中，语音指令的类型表征该语音指令是由增强现实设备进行响应的指令类型、或由目标设备进行响应的指令类型。

如图 2 所示，本申请中，涉及到的响应主体主要有两类：目标设备和增强现实设备。增强现实设备主要包含语音关键词检测技术服务和语音指令控制应用。当目标响应模式为第一响应模式，即基于增强现实设备连接的目标设备类型为第一类型如自研终端而确定目标响应模式为第一响应模式时，语音指令交由目标设备的操作系统进行全语音控制。当目标响应模式为第二响应模式，即基于增强现实设备连接的目标设备类型为第二类型如第三方手机或电脑、或者未检测到目标设备而确定目标响应模式为第二响应模式时，将语音指令转化为 Event Id 通知给增强现实设备中的语音指令控制应用，语音指令控制应用判断指

令是否为增强现实设备自身控制指令。如果指令为增强现实设备自身控制指令，则由语音指令控制应用对语音指令进行响应，如音量调节、亮度调节等。如果语音指令为目标设备控制的指令，则将 Event Id 转化为通用键盘(USB KeyBoard, UniversalSerialBus KeyBoard) 协议 Id，将 USB KeyBoard Id 交由目标设备，由目标设备对语音指令进行响应。

具体的，如图 3 所示，增强现实设备的语音关键词检测技术服务处于运行状态，由语音关键词检测技术服务判断与增强现实设备连接的目标设备为第一类型还是为第二类型。通过判断接入的目标设备类型，采取不同的目标响应模式。当目标响应模式为第一响应模式，即基于增强现实设备连接的目标设备类型为第一类型如自研终端而确定目标响应模式为第一响应模式时，增强现实设备中的语音关键词检测技术服务进入休眠状态，语音指令交由目标设备的操作系统进行全语音控制。示例性地，假设增强现实设备连接的目标设备为自研终端，由于自研终端里面包含有语音关键词检测技术服务，通常在自研终端上有预定义好的语音指令集，如调高音量指令是为了对多媒体数据的声音进行调节、增加亮度指令是为了对屏幕的显示亮度进行调节、模式切换指令是为了对多媒体数据的显示模式进行调节，如从普通模式调节到 3D 模式，或从 3D 模式调节到普通模式等。

在用户发出对应的语音指令后，增强现实设备判断接入的目标设备为自研终端，可将增强现实设备采集的语音指令交由目标设备的操作系统进行全语音控制。自研终端上运行的语音关键词检测技术服务将语音指令交给自研终端。自研终端、具体是应用处理单元响应语音指令，执行对应于语言指令的操控动作，如声音调节、亮度调节、模式切换等。

在增强现实设备获取到的语音指令为对增强现实设备输出的多媒体数据的声音进行调节的（声音调节）指令，如果与增强现实设备连接的目标设备是与增强现实设备适配的自研终端，则将声音调节指令交由自研终端处理，以通过自研终端实现对增强现实设备输出的多媒体数据的声音的调节。

在增强现实设备获取到的语音指令为对增强现实设备输出的多媒体数据的显示亮度进行调节的指令，如果与增强现实设备连接的目标设备是与增强现实设备适配的自研终端，则将显示亮度调节指令交由自研终端处理，以通过自研终端实现对增强现实设备屏幕的显示亮度的调节。

在增强现实设备获取到的语音指令为对增强现实设备输出的多媒体数据的显示模式进行调节的指令，如果与增强现实设备连接的目标设备是与增强现实设备适配的自研终端，则将显示模式调节指令交由自研终端处理，以通过自研

终端实现对增强现实设备输出的多媒体数据的显示模式的调节。

当目标响应模式为第二响应模式，即基于增强现实设备连接的目标设备类型为第二类型如第三方手机或电脑、或未检测到目标设备而确定目标响应模式为第二响应模式时，增强现实设备的语音关键词检测技术服务保持运行。

5 在一个可选的方案中，所述在所述目标响应模式为第二响应模式时，基于所述语音指令的类型，确定由所述增强现实设备和所述目标设备中的其中之一进行响应，包括：

在所述语音指令为第一类型指令时，由所述增强现实设备对所述语音指令进行响应；

10 在所述语音指令为第二类型指令时，将所述语音指令交由所述目标设备响应。

本申请中，在目标响应模式为第二响应模式的情况下，语音指令也包括两种类型，第一类型指令为增强现实设备可以响应的指令，如前述的声音调节指令、显示亮度调节指令及显示模式调节指令等。第二类型指令为增强现实设备
15 无法响应、但目标设备可以响应的指令，如“返回上一步”指令、“确定”指令及“主菜单”指令等等。

本申请中，在目标响应模式为第二响应模式的情况下，如果语音指令是诸如声音调节指令、显示亮度调节指令及显示模式调节指令等增强现实设备可以响应的指令，则由增强现实设备对语音指令进行响应。如果语音指令是诸如“返回上一步”指令、“确定”指令及“主菜单”指令等，则将语音指令交由目标设备
20 进行响应。

基于此，本申请中，在第二响应模式下，根据语音指令类型的不同，由不同的响应主体对语音指令进行响应的方案，可避免语音指令全部为响应主体中的其中之一进行响应而导致的负担过重的问题。本申请中，不同类型的语音指令
25 由不同的响应主体进行响应，在工程上易于实施，具有很高的可行性。

在一个可选的方案中，所述在所述语音指令为第二类型指令时，将所述语音指令交由所述目标设备响应，包括：

在所述语音指令为第二类型指令时，将所述语音指令转换成目标类型指令，将所述目标类型指令交由所述目标设备响应。

30 本申请中，如果语音指令为增强现实设备无法响应、但目标设备可以响应的指令，则需要将语音指令转化为目标设备可以识别或者响应的类型，进而实现对语音指令的响应。

示例性地，假设增强现实设备连接的目标设备为第三方手机或电脑端，当

增强现实设备中的语音关键词检测技术服务识别到语音指令后，将语音指令转化为 Event Id 通知给增强现实设备中的语音指令控制应用，语音指令控制应用在收到指令 Event Id 后，将指令 Event Id 与预先定义的指令功能对应的编号进行对应，根据对应的编号代表的预定义的指令功能做出对应的反馈。其中，预定义的指令功能为预先定义的指令集中不同编号代表的指令功能，如编号 1 代表声音增大功能，编号 2 代表亮度增大功能，编号 3 代表模式切换功能等。

当指令 Event Id 对应的指令为第一类型指令，即当指令 Event Id 与为增强现实设备预定义的指令功能对应的编号存在对应关系，也就是说指令 Event Id 对应的指令为增强现实设备可以响应的自身控制相关指令，如声音调节、亮度调节、显示切换、模式切换等与增强现实设备硬件相关的指令时，语音指令控制应用直接通过系统应用程序编程接口（API，Application Programming Interface）完成对应的控制操作。当指令 Event Id 对应的指令为第二类型指令，即当指令 Event Id 与为目标设备预定义的指令功能对应的编号存在对应关系，也就是说当指令 Event Id 对应的指令为需要目标设备进行响应的指令时，通过将指令 Event Id 转化为 USB KeyBoard 协议 Id 发送给接入的目标设备。目标设备对 USB KeyBoard 协议 Id 进行响应，执行对增强现实设备输出的多媒体数据的操作。示例性地，将指令“返回上一步”定义为目标设备的键盘“F1”键，当用户点击“F1”键时，增强现实设备输出的多媒体数据就会返回上一步操作。或者将“主菜单”定义为目标设备的键盘“F2”键，当用户点击“F2”键时，增强现实设备就会弹出主菜单界面。或者还可以将“确认”定义为目标设备的键盘“enter”键，当用户点击“enter”键时就可以进行确认操作。

本申请中的增强现实设备的前述执行过程可由设置在增强现实设备中的高性能专用芯片来实现，高性能专用芯片的设置可提升计算效率，实现对语音指令的快速响应。此外，本申请中，采用低功耗语音关键词检测技术服务，具有所需算力低、资源消耗少等显著优势。在不明显增加功耗的情况下，赋予了增强现实设备 AR 能力，极大程度提升了增强现实设备自身控制能力。再有，语音指令可定制、可扩展，弥补了通过按键控制的不足之处。借助通用的 USB KeyBoard 协议，增强现实设备将识别到的语音指令转化为目标设备可识别或响应的目标类型，将目标设备对语音指令的响应看做是用户操作（如前述的目标设备的键盘“F1”键、目标设备的键盘“F2”键以及目标设备的键盘“enter”键等）按键而对语音指令执行的响应。这种方案相当于将增强现实设备作为目标设备的控制外设，在低成本接入的情况下赋予目标设备语音交互能力。在实际应用中增加了增强现实设备的产品竞争力。

在一个可选的方案中，所述音频采集单元用于对目标语音数据进行采集，所述目标语音数据包括语音指令；

所述方法还包括：

通过确定所述目标语音数据中是否存在语音指令而获取所述语音指令；

5 和/或，将所述目标语音数据交由所述目标设备。

本方案中，增强现实设备还包括关键词检测单元及控制应用单元。关键词检测单元包含有语音关键词检测技术服务，用于判断接入的目标设备类型，以及识别语音指令。控制应用单元包含有语音指令控制应用，用于判断指令类型，将不同类型的指令交由不同的响应主体进行响应。目标语音数据包括语音指令
10 和/或在增强现实设备进行多媒体数据输出时产生的语音数据。

可以理解，本申请中的语音指令指的是通过语音形式输入至增强现实设备的指令，是一种指令数据。麦克风阵列传感器采集到的目标语音数据可能是指令数据。除此之外，麦克风阵列传感器采集到的语音数据可能是进行多媒体数据输出时产生的语音数据，如在通过增强现实设备接听来电时，麦克风阵列传感器采集到的语音数据可能是通话内容。或者在通过增强现实设备进行电影的
15 投影时，麦克风阵列传感器采集到的语音数据可能是电影中的音频内容。

增强现实设备通过麦克风阵列传感器采集到目标语音数据，识别目标语音数据是否包含语音指令、和/或是否包含增强现实设备进行多媒体数据输出时产生的语音数据。如果经识别发现目标语音数据既包含语音指令，又包含增强现实设备进行多媒体数据输出时产生的语音数据，则可将这两种语音数据进行分离，以对两种语音数据执行不同的处理。
20

当目标语音数据中包含语音指令时，增强现实设备通过关键词检测单元判断不同的目标设备类型，从而采用不同的响应模式。具体的，当目标设备为第一类型时，由目标设备对语音指令进行响应。当目标设备为第二类型或未检测到目标设备时，通过关键词检测单元识别语音指令，控制应用单元判断不同的指令类型，从而选用不同的响应主体对语音指令进行响应。
25

当目标语音数据中包含增强现实设备进行多媒体数据输出时产生的语音数据，可将该语音数据交由目标设备，由目标设备对该语音数据进行录制、语义识别等处理。

30 前述方案中，通过将采集到的目标语音数据传送到增强现实设备端和/或目标设备端，在工程上易于实施，保证了增强现实设备端和/或目标设备对目标语音数据的正常使用。

在一个可选的方案中，所述通过确定所述目标语音数据中是否存在语音指

令而获取所述语音指令；和/或，将所述目标语音数据交由所述目标设备，包括：
将所述目标语音数据进行降噪处理，得到降噪后的目标语音数据；
通过确定所述降噪后的目标语音数据中是否存在语音指令而获取所述语音指令；

5 和/或，将所述降噪后的目标语音数据交由所述目标设备。

本方案中，增强现实设备还包括降噪单元，用于对目标语音数据进行降噪处理，得到降噪后的目标语音数据。增强现实设备通过麦克风阵列传感器采集到目标语音数据，将采集到的目标语音数据通过降噪单元对目标语音数据进行降噪处理，得到降噪后的目标语音数据。将降噪后的目标语音数据传送给增强现实设备端和/或目标设备端。当目标语音数据中包含语音指令时，增强现实设备端通过关键词检测单元判断不同的目标设备类型，从而采用不同的响应模式。当目标设备类型为第一类型时，由目标设备对语音指令进行响应。当目标设备类型为第二类型或未检测到目标设备时，通过关键词检测单元识别语音指令，控制应用单元判断不同的指令类型，从而选用不同的响应主体对语音指令进行响应。

当目标语音数据中包含增强现实设备进行多媒体数据输出时产生的语音数据，可将降噪后的语音数据交由目标设备，由目标设备对该降噪后的语音数据进行录制、语义识别等处理。

通过将采集到的目标语音数据进行降噪，使得经过降噪后的目标语音数据是不包含噪声或包含噪声较少的语音数据，可以提高传送给增强现实设备端和/或目标设备端的语音数据的质量，从而实现对目标语音数据的准确响应。

在一个可选的方案中，所述通过确定所述降噪后的目标语音数据中是否存在语音指令而获取所述语音指令；和/或，将所述降噪后的目标语音数据交由所述目标设备，包括：

25 基于降噪后的目标语音数据，得到第一目标语音数据和第二目标语音数据；
通过确定所述第一目标语音数据中是否存在语音指令而获取所述语音指令；
和/或，将所述第二目标语音数据交由所述目标设备。

本方案中，增强现实设备还包括复制分流单元，用于将降噪后的目标语音数据进行复制，得到两份降噪后的目标语音数据，将两份降噪后的目标语音数据进行分流，得到第一目标语音数据和第二目标语音数据。如图 4 所示，增强现实设备将读取到的麦克风音频数据 Audio In 送给特定的降噪单元进行定向降噪，得到增强后的不含噪声或含噪声较少的音频数据 Clean Audio，不含噪声或含噪声较少的音频数据通过复制分流单元得到第一目标语音数据和第二目标语

音数据。可以理解，第一目标语音数据和第二目标语音数据均为与经过降噪后的不含噪声或含噪声较少的 Clean Audio 相同的语音数据。第二目标语音数据直接通过硬件有线或无线方式传输给目标设备 Audio Out，第一目标语音数据由增强现实设备端处理。进一步的，当目标语音数据中存在语音指令时，增强现实设备端获取语音指令，通过关键词检测单元判断不同的目标设备类型，从而采用不同的响应模式。具体的，当目标设备为第一类型时，由目标设备对语音指令进行响应。当目标设备为第二类型或未检测到目标设备时，通过关键词检测单元识别语音指令，控制应用单元判断不同的指令类型，从而选用不同的响应主体对语音指令进行响应。

通过将降噪后的目标语音数据进行复制并分流，得到第一目标语音数据和第二目标语音数据，第一目标语音数据由增强现实设备端进行处理，是为了进行语音指令识别以及由不同响应主体进行响应的方案。第二目标语音数据传送到目标设备端是为了让采集到的目标语音数据被目标设备的其他应用所用。示例性地，当用户进行通话时，采集到的目标语音数据作为通话内容被目标设备的通话应用所用。对降噪后的目标语音数据进行复制并分流，能够保证增强现实设备端和目标设备端的应用都能互不干扰的获取到各自需要的音频数据。

本申请实施例提供一种语音交互装置，如图 5 所示，所述装置包括：

获取单元 501，用于基于增强现实设备的音频采集单元获取用户的语音指令；

第一确定单元 502，用于确定与所述增强现实设备连接的目标设备的类型；

第二确定单元 503，用于在所述目标设备为第一类型时，确定对所述语音指令的目标响应模式为第一响应模式；

第三确定单元 504，用于在所述目标设备为第二类型或未检测到目标设备时，确定对所述语音指令的目标响应模式为第二响应模式；

响应单元 505，用于采用所述目标响应模式，对所述语音指令进行响应。

在一个可选的方案中，所述响应单元 505，用于在所述目标响应模式为第一响应模式时，将所述语音指令交由所述目标设备响应；在所述目标响应模式为第二响应模式时，基于所述语音指令的类型，确定由所述增强现实设备和所述目标设备中的其中之一进行响应。

在一个可选的方案中，所述响应单元 505，用于在所述语音指令为第一类型指令时，由所述增强现实设备对所述语音指令进行响应；在所述语音指令为第二类型指令时，将所述语音指令交由所述目标设备响应。

在一个可选的方案中，所述响应单元 505，用于在所述语音指令为第二类型指令时，将所述语音指令转换成目标类型指令，将所述目标类型指令交由所述

目标设备响应。

在一个可选的方案中，所述音频采集单元用于对目标语音数据进行采集，所述目标语音数据包括语音指令；

所述装置还包括：

5 语音数据单元，用于通过确定所述目标语音数据中是否存在语音指令而获取所述语音指令；和/或，将所述目标语音数据交由所述目标设备。

10 在一个可选的方案中，所述装置还包括：降噪单元；所述降噪单元用于将所述目标语音数据进行降噪处理，得到降噪后的目标语音数据；相应地，所述语音数据单元，还用于通过确定所述降噪后的目标语音数据中是否存在语音指令而获取所述语音指令；和/或，将所述降噪后的目标语音数据交由所述目标设备。

15 在一个可选的方案中，所述装置还包括：复制分流单元；所述复制分流单元用于基于降噪后的目标语音数据，得到第一目标语音数据和第二目标语音数据；相应地，所述语音数据单元，还用于通过确定所述第一目标语音数据中是否存在语音指令而获取所述语音指令；和/或，将所述第二目标语音数据交由所述目标设备。

20 需要说明的是，本申请实施例的语音交互装置，由于该装置解决问题的原理与前述的语音交互方法相似，因此，该装置的实施过程、实施原理及有益效果均可以参见前述方法的实施过程、实施原理及有益效果的描述，重复之处不再赘述。

本申请实施例提供一种增强现实设备，所述增强现实设备至少包括本申请所述的语音交互装置。

根据本申请的实施例，本申请还提供了一种电子设备。

25 图 6 示出了可以用来实施本申请的实施例的示例电子设备 600 的示意性框图。电子设备旨在表示各种形式的数字计算机，诸如，膝上型计算机、台式计算机、工作台、个人数字助理、服务器、刀片式服务器、大型计算机、和其它适合的计算机。电子设备还可以表示各种形式的移动装置，诸如，个人数字处理、蜂窝电话、智能电话、可穿戴设备和其它类似的计算装置。本文所示的部件、它们的连接和关系、以及它们的功能仅仅作为示例，并且不意在限制本文中描述的 30 的和/或者要求的本申请的实现。

如图 6 所示，电子设备 600 包括计算单元 601，其可以根据存储在只读存储器（ROM）602 中的计算机程序或者从存储单元 608 加载到随机访问存储器（RAM）603 中的计算机程序，来执行各种适当的动作和处理。在 RAM 603 中，

还可存储电子设备 600 操作所需的各种程序和数据。计算单元 601、ROM 602 以及 RAM 603 通过总线 604 彼此相连。输入/输出 (I/O) 接口 605 也连接至总线 604。

电子设备 600 中的多个部件连接至 I/O 接口 605, 包括: 输入单元 606, 例如键盘、鼠标等; 输出单元 607, 例如各种类型的显示器、扬声器等; 存储单元 608, 例如磁盘、光盘等; 以及通信单元 609, 例如网卡、调制解调器、无线通信收发机等。通信单元 609 允许电子设备 600 通过诸如因特网的计算机网络和/或各种电信网络与其他设备交换信息/数据。

计算单元 601 可以是各种具有处理和计算能力的通用和/或专用处理组件。计算单元 601 的一些示例包括但不限于中央处理单元 (CPU)、图形处理单元 (GPU)、各种专用的人工智能 (AI) 计算芯片、各种运行机器学习模型算法的计算单元、数字信号处理器 (DSP)、以及任何适当的处理器、控制器、微控制器等。计算单元 601 执行上文所描述的各个方法和处理, 例如语音交互方法。例如, 在一些实施例中, 语音交互方法可被实现为计算机软件程序, 其被有形地包含于机器可读介质, 例如存储单元 608。在一些实施例中, 计算机程序的部分或者全部可以经由 ROM 602 和/或通信单元 609 而被载入和/或安装到电子设备 600 上。当计算机程序加载到 RAM 603 并由计算单元 601 执行时, 可以执行上文描述的语音交互方法的一个或多个步骤。备选地, 在其他实施例中, 计算单元 601 可以通过其他任何适当的方式 (例如, 借助于固件) 而被配置为执行语音交互方法。

本文中以上描述的系统和技术和各种实施方式可以在数字电子电路系统、集成电路系统、场可编程门阵列 (FPGA)、专用集成电路 (ASIC)、专用标准产品 (ASSP)、片上系统 (SOC)、复杂可编程逻辑设备 (CPLD)、计算机硬件、固件、软件、和/或它们的组合中实现。这些各种实施方式可以包括: 实施在一个或者多个计算机程序中, 该一个或者多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释, 该可编程处理器可以是专用或者通用可编程处理器, 可以从存储系统、至少一个输入装置、和至少一个输出装置接收数据和指令, 并且将数据和指令传输至该存储系统、该至少一个输入装置、和该至少一个输出装置。

用于实施本申请的方法的程序代码可以采用一个或多个编程语言的任何组合来编写。这些程序代码可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器或控制器, 使得程序代码当由处理器或控制器执行时使流程图和/或框图所示的功能/操作被实施。程序代码可以完全在机器上执行、

部分地在机器上执行，作为独立软件包部分地在机器上执行且部分地在远程机器上执行或完全在远程机器或服务器上执行。

应该理解，可以使用上面所示的各种形式的流程，重新排序、增加或删除步骤。例如，本发申请中记载的各步骤可以并行地执行也可以顺序地执行也可以不同的次序执行，只要能够实现本申请公开的技术方案所期望的结果，本文在此不进行限制。

此外，术语“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或隐含地包括至少一个该特征。在本申请的描述中，“多个”
10 的含义是两个或两个以上，除非另有明确具体的限定。

以上所述，仅为本申请的具体实施方式，但本申请的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本申请揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本申请的保护范围之内。因此，本申请的保护范围应以所述权利要求的保护范围为准。

15

权 利 要 求 书

1、一种语音交互方法，其中，所述方法包括：

基于增强现实设备的音频采集单元获取用户的语音指令；

5 确定与所述增强现实设备连接的目标设备的类型；

在所述目标设备为第一类型时，确定对所述语音指令的目标响应模式为第一响应模式；

在所述目标设备为第二类型或未检测到目标设备时，确定对所述语音指令的目标响应模式为第二响应模式；

10 采用所述目标响应模式，对所述语音指令进行响应。

2、根据权利要求1所述的方法，其中，所述采用所述目标响应模式，对所述语音指令进行响应，包括：

在所述目标响应模式为第一响应模式时，将所述语音指令交由所述目标设备响应；

15 在所述目标响应模式为第二响应模式时，基于所述语音指令的类型，确定由所述增强现实设备和所述目标设备中的其中之一进行响应。

3、根据权利要求2所述的方法，其中，所述基于所述语音指令的类型，确定由所述增强现实设备和所述目标设备中的其中之一进行响应，包括：

20 在所述语音指令为第一类型指令时，由所述增强现实设备对所述语音指令进行响应；

在所述语音指令为第二类型指令时，将所述语音指令交由所述目标设备响应。

4、根据权利要求3所述的方法，其中，所述在所述语音指令为第二类型指令时，将所述语音指令交由所述目标设备响应，包括：

25 在所述语音指令为第二类型指令时，将所述语音指令转换成目标类型指令，将所述目标类型指令交由所述目标设备响应。

5、根据权利要求1所述的方法，其中，所述音频采集单元用于对目标语音数据进行采集，所述目标语音数据包括语音指令；

所述方法还包括：

30 通过确定所述目标语音数据中是否存在语音指令而获取所述语音指令；

和/或，将所述目标语音数据交由所述目标设备。

6、根据权利要求5所述的方法，其中，所述通过确定所述目标语音数据中

是否存在语音指令而获取所述语音指令；和/或，将所述目标语音数据交由所述目标设备，包括：

将所述目标语音数据进行降噪处理，得到降噪后的目标语音数据；

5 通过确定所述降噪后的目标语音数据中是否存在语音指令而获取所述语音指令；

和/或，将所述降噪后的目标语音数据交由所述目标设备。

7、根据权利要求6所述的方法，其中，所述通过确定所述降噪后的目标语音数据中是否存在语音指令而获取所述语音指令；和/或，将所述降噪后的目标语音数据交由所述目标设备，包括：

10 基于降噪后的目标语音数据，得到第一目标语音数据和第二目标语音数据；通过确定所述第一目标语音数据中是否存在语音指令而获取所述语音指令；和/或，将所述第二目标语音数据交由所述目标设备。

8、一种语音交互装置，其中，所述装置包括：

获取单元，用于基于增强现实设备的音频采集单元获取用户的语音指令；

15 第一确定单元，用于确定与所述增强现实设备连接的目标设备的类型；

第二确定单元，用于在所述目标设备为第一类型时，确定对所述语音指令的目标响应模式为第一响应模式；

第三确定单元，用于在所述目标设备为第二类型或未检测到目标设备时，确定对所述语音指令的目标响应模式为第二响应模式；

20 响应单元，用于采用所述目标响应模式，对所述语音指令进行响应。

9、一种增强现实设备，其中，至少包括权利要求8所述的语音交互装置。

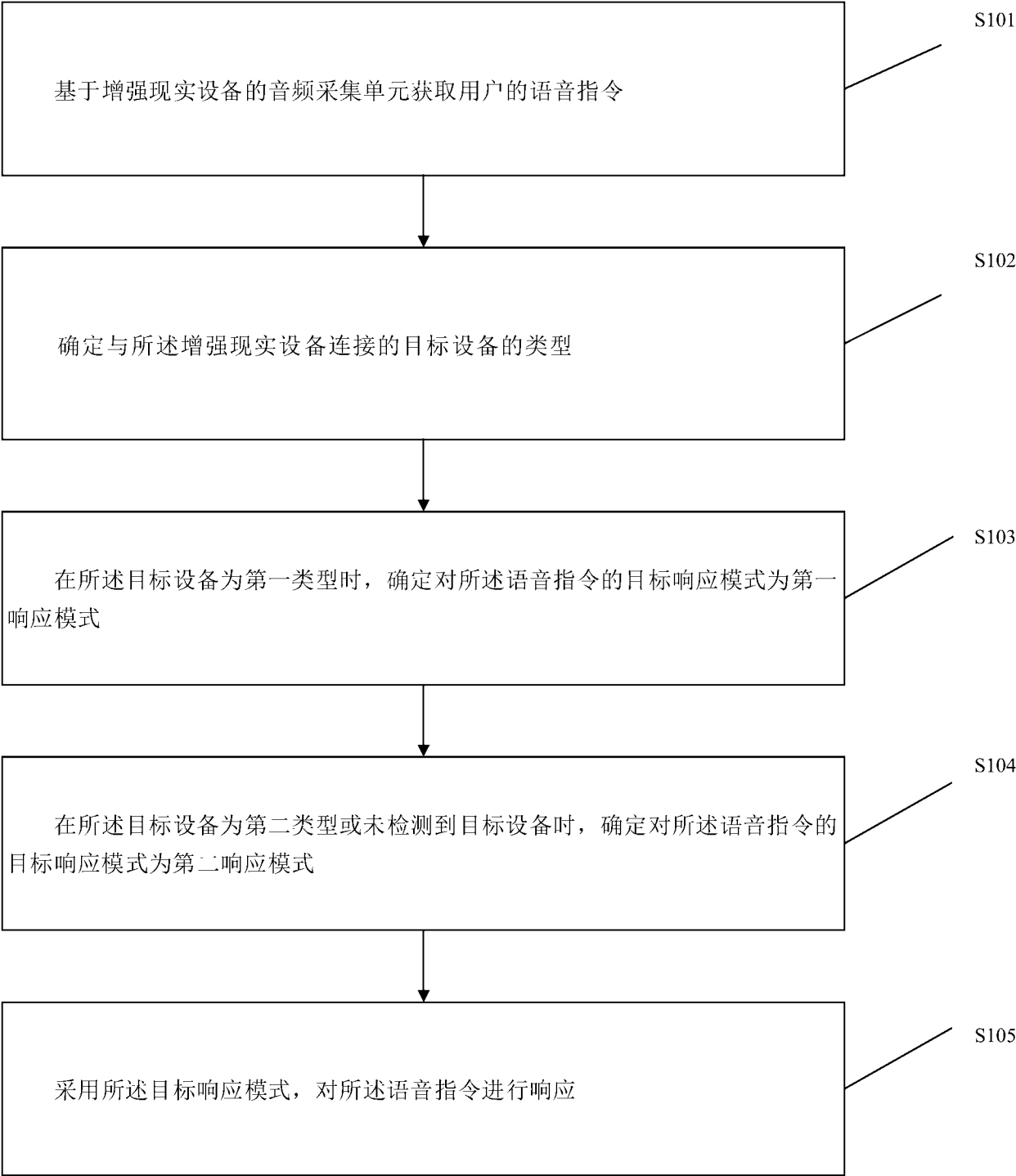
10、一种电子设备，其中，包括：

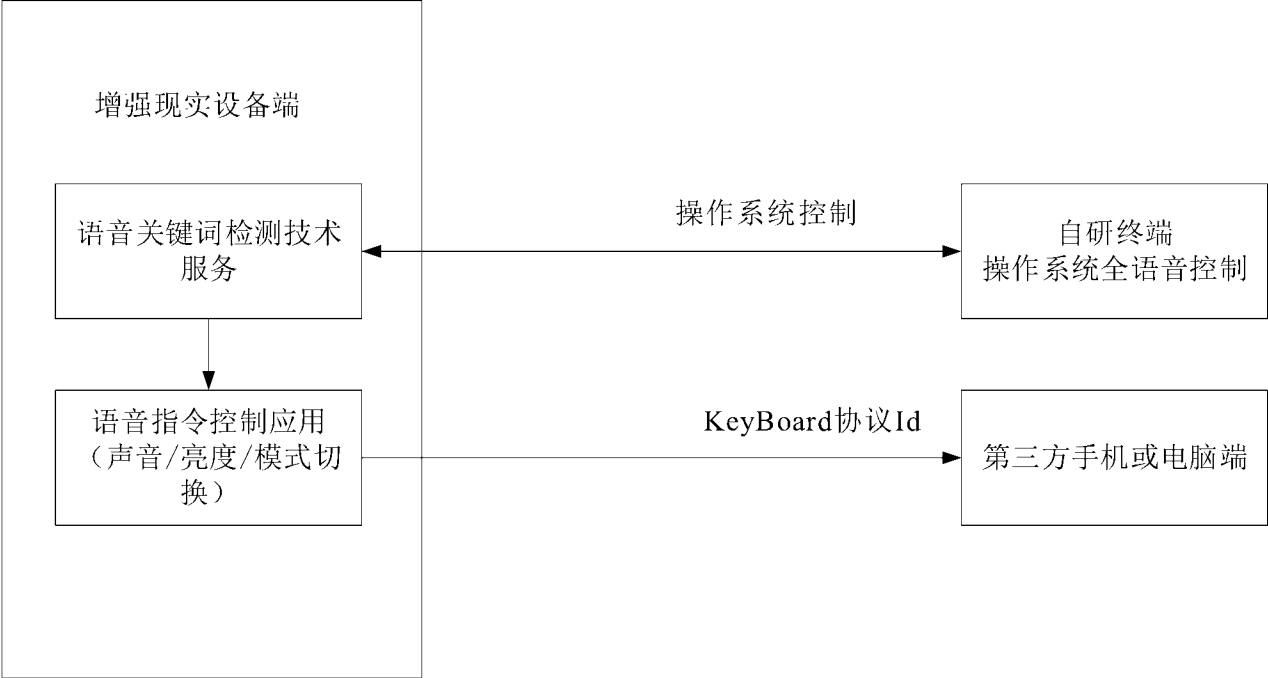
至少一个处理器；以及

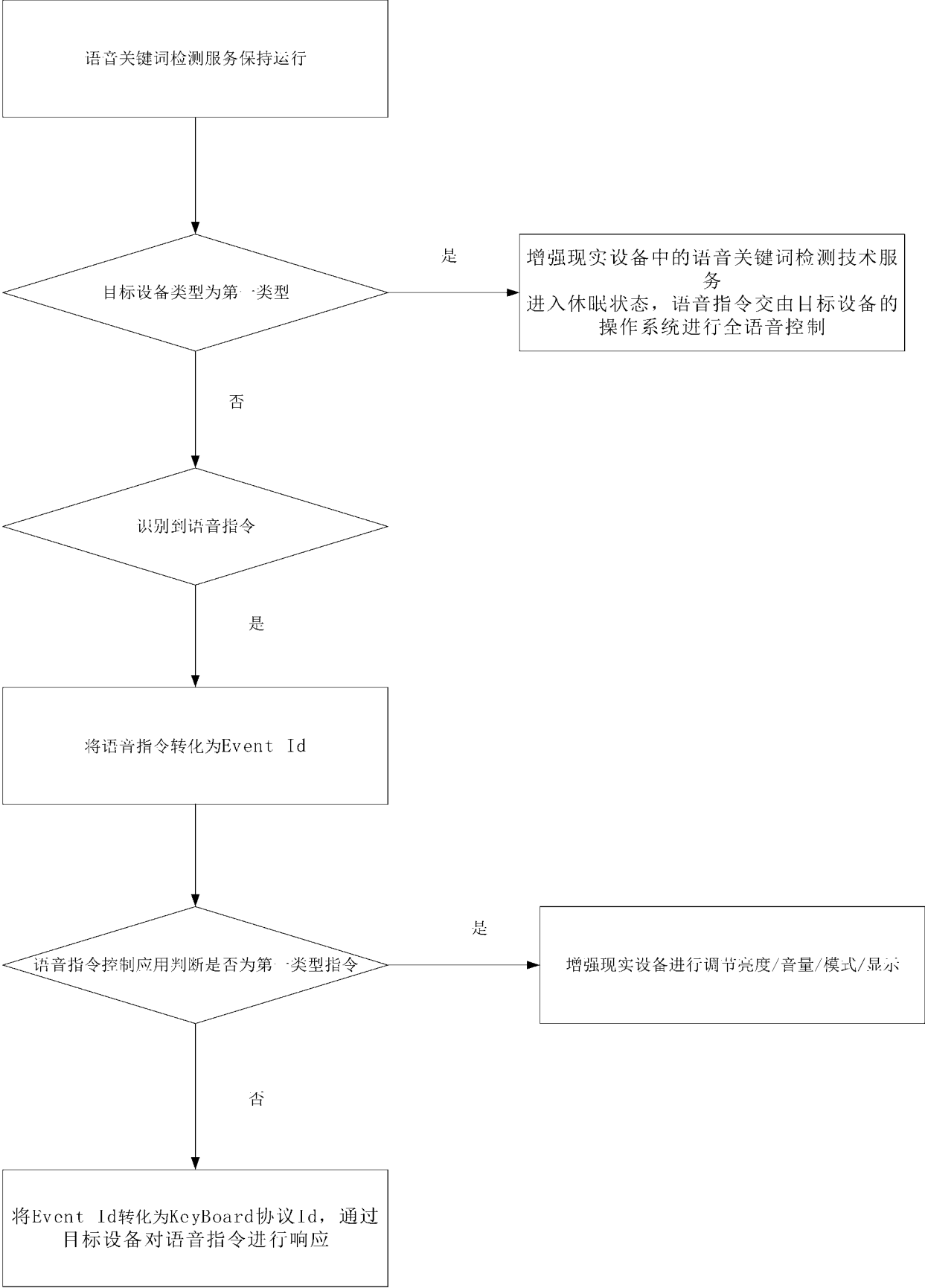
与所述至少一个处理器通信连接的存储器；其中，

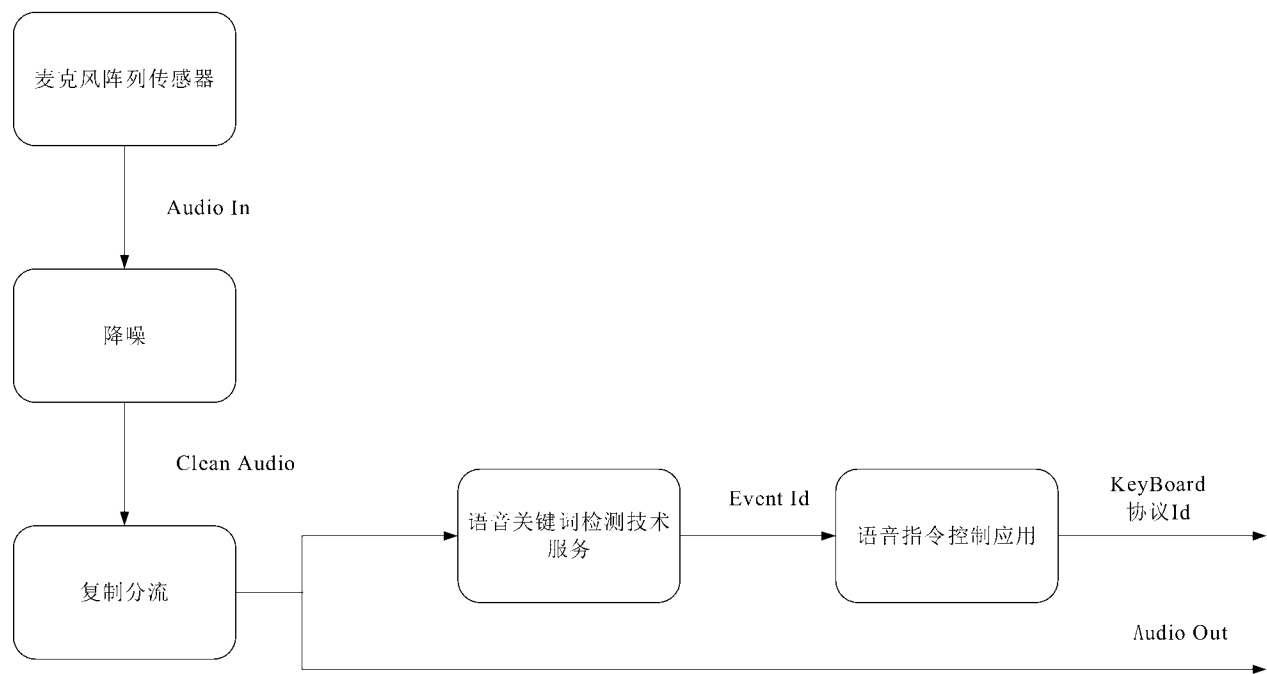
25 所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行权利要求1-7中任一项所述的方法。

1/6

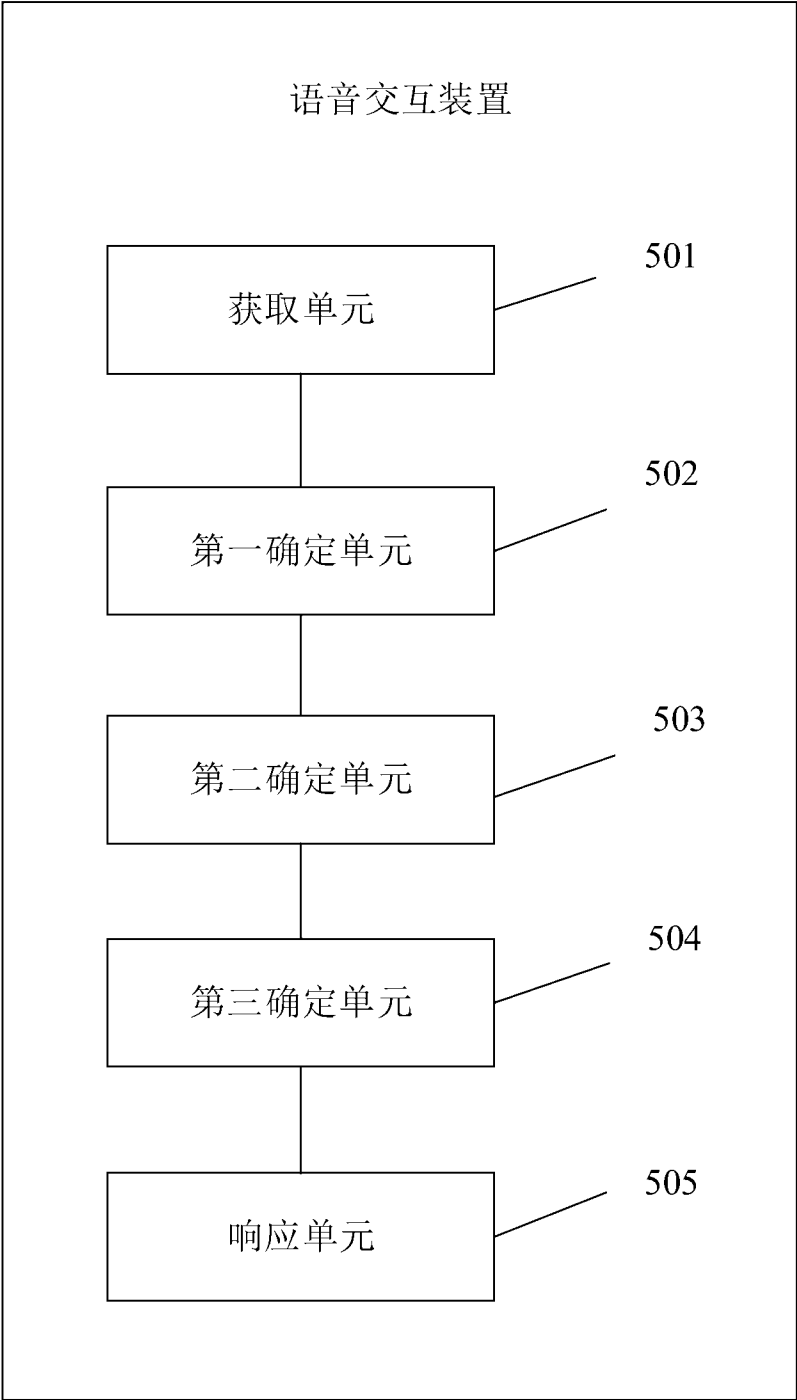


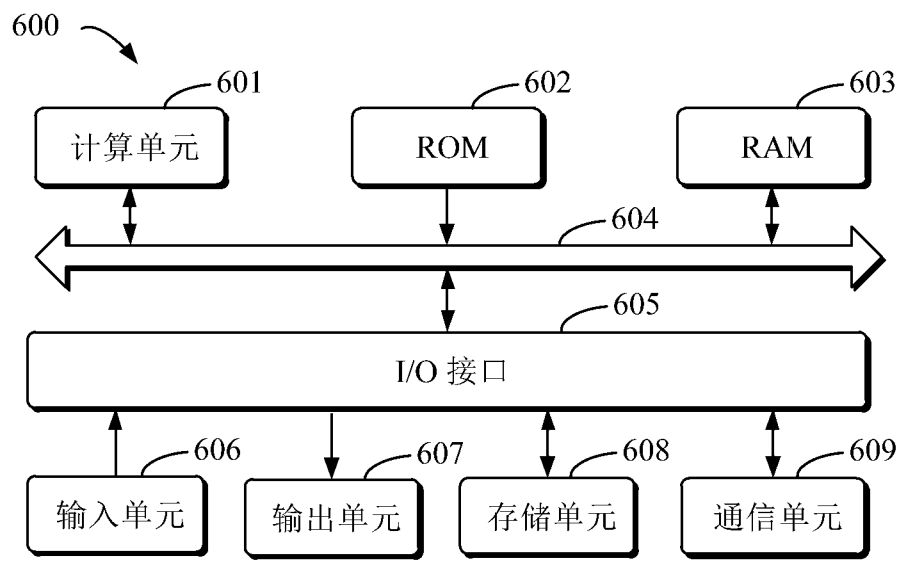






□ 4





INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2024/071940

A. CLASSIFICATION OF SUBJECT MATTER G10L15/22(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:G10L 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) CNABS, CNTXT, CNKI: 增强现实, 语音, 识别, 交互, 终端, 电脑, 兼容, 可穿戴, 类型, 模式; VEN, USTXT, EPTXT, ELSEVIER: AR, voice, recognition, interaction, terminal, computer, compatible, wearable, type, mode		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 116030810 A (HANGZHOU LINGBAN TECHNOLOGY CO., LTD.) 28 April 2023 (2023-04-28) claims 1-10	1-10
X	CN 109658932 A (SHENZHEN SKYWORTH-RGB ELECTRONICS CO., LTD.) 19 April 2019 (2019-04-19) description, paragraphs 31-117, and figures 1-6	1, 5-10
A	US 2015293738 A1 (SAMSUNG DISPLAY CO., LTD.) 15 October 2015 (2015-10-15) entire document	1-10
A	US 2018261224 A1 (JETVOX ACOUSTIC CORP.) 13 September 2018 (2018-09-13) entire document	1-10
A	US 2021366472 A1 (LG ELECTRONICS INC.) 25 November 2021 (2021-11-25) entire document	1-10
A	CN 111966321 A (GUANGDONG OPPO MOBILE COMMUNICATIONS CO., LTD.) 20 November 2020 (2020-11-20) entire document	1-10
☐ 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 March 2024		Date of mailing of the international search report 26 March 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/CN2024/071940

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	116030810	A	28 April 2023	None			
CN	109658932	A	19 April 2019	WO	2020133946	A1	02 July 2020
				CN	109658932	B	17 November 2020
US	2015293738	A1	15 October 2015	KR	20150119573	A	26 October 2015
				KR	102278507	B1	16 July 2021
				US	9921791	B2	20 March 2018
				CN	105022476	A	04 November 2015
				CN	105022476	B	31 March 2020
US	2018261224	A1	13 September 2018	TW	201834437	A	16 September 2018
US	2021366472	A1	25 November 2021	WO	2020213758	A1	22 October 2020
				US	11393465	B2	19 July 2022
				KR	20200122981	A	28 October 2020
				KR	102321855	B1	05 November 2021
CN	111966321	A	20 November 2020	None			

A. 主题的分类

G10L15/22(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:G10L

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))

CNABS,CNXTXT,CNKI:增强现实,语音,识别,交互,终端,电脑,兼容,可穿戴,类型,模式 VEN,USTXT,EPTXT,ELSEVIER: AR,voice,recognition, interaction,terminal,computer, compatible,wearable,type,mode

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 116030810 A (杭州灵伴科技有限公司) 2023年4月28日 (2023 - 04 - 28) 权利要求1-10	1-10
X	CN 109658932 A (深圳创维-RGB电子有限公司) 2019年4月19日 (2019 - 04 - 19) 说明书第31-117段、图1-6	1、5-10
A	US 2015293738 A1 (SAMSUNG DISPLAY CO LTD) 2015年10月15日 (2015 - 10 - 15) 全文	1-10
A	US 2018261224 A1 (JETVOX ACOUSTIC CORP) 2018年9月13日 (2018 - 09 - 13) 全文	1-10
A	US 2021366472 A1 (LG ELECTRONICS INC) 2021年11月25日 (2021 - 11 - 25) 全文	1-10
A	CN 111966321 A (OPPO广东移动通信有限公司) 2020年11月20日 (2020 - 11 - 20) 全文	1-10

☐ 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3月21日

国际检索报告邮寄日期

2024年3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

李琴

电话号码 (+86) 020-28950533

PCT/ISA/210 表(第2页) (2022年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/071940

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)
CN	116030810	A	2023年4月28日	无			
CN	109658932	A	2019年4月19日	WO	2020133946	A1	2020年7月2日
				CN	109658932	B	2020年11月17日
US	2015293738	A1	2015年10月15日	KR	20150119573	A	2015年10月26日
				KR	102278507	B1	2021年7月16日
				US	9921791	B2	2018年3月20日
				CN	105022476	A	2015年11月4日
				CN	105022476	B	2020年3月31日
US	2018261224	A1	2018年9月13日	TW	201834437	A	2018年9月16日
US	2021366472	A1	2021年11月25日	WO	2020213758	A1	2020年10月22日
				US	11393465	B2	2022年7月19日
				KR	20200122981	A	2020年10月28日
				KR	102321855	B1	2021年11月5日
CN	111966321	A	2020年11月20日	无			