(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 10 月 24 日 (24.10.2024)

WIPO | PCT

(10) 国际公布号
WO 2024/217470 A1

(51) 国际专利分类号:
G10L 13/033 (2013.01)

(21) 国际申请号: PCT/CN2024/088407

(22) 国际申请日: 2024 年 4 月 17 日 (17.04.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202310433734.8 2023年4月20日 (20.04.2023) CN

(71) 申请人:维沃移动通信有限公司(**VIVO MOBILE COMMUNICATION CO., LTD.**) [CN/CN]; 中国广东省东莞市长安镇维沃路 1 号, Guangdong 523863 (CN)。

(72) 发明人:邢梦晗(**XING, Menghan**); 中国广东省东莞市长安镇维沃路1号, Guangdong 523863 (CN)。

(74) 代理人: 北京远志博慧知识产权代理事务所（特殊普通合伙）(**BOHUI INTELLECTUAL PROPERTY**); 中国北京市海淀区交大东路31号东区10号楼等17幢31幢108, Beijing 100044 (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)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

(54) **Title:** VOICE RECORDING METHOD AND APPARATUS, ELECTRONIC DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 语音录制方法、装置、电子设备及存储介质

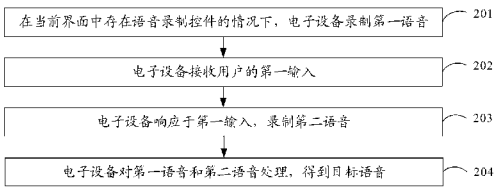


图 1

201 When there is a voice recording control in a current interface, an electronic device records first voice
202 The electronic device receives a first input from a user
203 In response to the first input, the electronic device records second voice
204 The electronic device processes the first voice and the second voice to obtain target voice

(57) **Abstract:** The present application relates to the field of voice technology, and discloses a voice recording method and apparatus, an electronic device, and a storage medium. The method comprises: when there is a voice recording control in a current interface, recording first voice; receiving a first input from a user, the first input being an input to the voice recording control; in response to the first input, recording second voice; and processing the first voice and the second voice to obtain target voice.

(57) 摘要: 本申请公开了一种语音录制方法、装置、电子设备及存储介质，属于语音技术领域。该方法包括：在当前界面中存在语音录制控件的情况下，录制第一语音；接收用户的第一输入，该第一输入为对语音录制控件的输入；响应于第一输入，录制第二语音；对第一语音和第二语音处理，得到目标语音。

WO 2024/217470 A1

语音录制方法、装置、电子设备及存储介质

相关申请的交叉引用

本申请主张在2023年04月20日在中国提交的申请号为202310433734.8的中国专利的优先权，其全部内容通过引用包含于此。

技术领域

本申请属于语音技术领域，具体涉及一种语音录制方法、装置、电子设备及存储介质。

背景技术

随着电子设备的不断发展，电子设备的功能和应用也越来越丰富，用户会经常使用聊天软件进行语音聊天，例如，在“通讯A”应用的某个聊天框中，用户可以先点击该聊天框中的语音按钮，以触发电子设备显示“按住说话”控件，然后用户可以长按“按住说话”控件并开始说话，从而电子设备开始录制语音，并在录制完成后向该聊天框对应的联系人发送录制的语音。

然而，在上述过程中，经常出现用户已经开始说话，但是没有及时按住语音录制控件（即上述的“按住说话”控件）进行录音的情况，如此，造成电子设备录制的语音缺失。

发明内容

本申请实施例的目的是提供一种语音录制方法、装置、电子设备及存储介质，能够保证电子设备录制语音的完整性。

第一方面，本申请实施例提供了一种语音录制方法，该方法包括：在当前界面中存在语音录制控件的情况下，录制第一语音；接收用户的第一输入，该第一输入为对语音录制控件的输入；响应于第一输入，录制第二语音；对第一语音和第二语音处理，得到目标语音。

第二方面，本申请实施例提供了一种语音录制装置，该语音录制装置包括：录制模块、接收模块和处理模块。录制模块，用于在当前界面中存在语音录制控件的情况下，录制第一语音。接收模块，用于接收用户的第一输入，该第一输入为对语音录制控件的输入。录制模块，还用于响应于接收模块接收的第一输入，录制第二语音。处理模块，用于对录制模块得到的第一语音和第二语音处理，得到目标语音。

第三方面，本申请实施例提供了一种电子设备，该电子设备包括处理器和存储器，所述存储器存储可在所述处理器上运行的程序或指令，所述程序或指令被所述处理器执行时实现如第一方面所述的方法的步骤。

第四方面，本申请实施例提供了一种可读存储介质，所述可读存储介质上存储程序或指令，所述程序或指令被处理器执行时实现如第一方面所述的方法的步骤。

第五方面，本申请实施例提供了一种芯片，所述芯片包括处理器和通信接口，所述通信接口和所述处理器耦合，所述处理器用于运行程序或指令，实现如第一方面所述的方法。

第六方面，本申请实施例提供一种计算机程序产品，该程序产品被存储在存储介质中，该程序产品被至少一个处理器执行以实现如第一方面所述的方法。

在本申请实施例中，在当前界面中存在语音录制控件的情况下，电子设备录制第一语音，并在接收到对语音录制控件的第一输入后，录制第二语音，从而对第一语音和第二语音处理，得到目标语音。本方案中，由于电子设备在检测到当前界面中存在语音录制控件的情况下，可以先录制第一语音，并在实际接收到用户对语音录制控件的输入时，停止录制第一语音，并开始录制第二语音，然后电子设备可以将用户输入之前录制的第一语音和用户输入之后录制的第二语音进行处理，以得到目标语音。如此，避免了只有在用户触发对录制控件的输入后电子设备才开始录制语音，容易产生语音缺失的问题，从而保证了电子设备录制语音的完整性。

附图说明

图 1 是本申请实施例提供的一种语音录制方法的流程示意图之一；

图 2（A）是本申请实施例提供的一种对语音按钮进行输入的界面实例示意图；

图 2（B）是本申请实施例提供的一种在会话界面显示“按住说话”控件的界面实例示意图；

图 3 是本申请实施例提供的一种语音录制方法的流程示意图之二；

图 4 是本申请实施例提供的一种在会话界面对“按住说话”控件进行输入的界面实例示意图；

图 5 是本申请实施例提供的一种语音录制方法的流程示意图之三；

图 6 是本申请实施例提供的一种语音录制方法的流程示意图之四；

图 7 是本申请实施例提供的一种语音录制方法的流程示意图之五；

图 8 是本申请实施例提供的一种语音录制方法的流程示意图之六；

图 9 是本申请实施例提供的一种在会话界面显示提示信息的界面实例示意图；

图 10 是本申请实施例提供的一种语音录制装置的结构示意图之一；

图 11 是本申请实施例提供的一种语音录制装置的结构示意图之二；

图 12 是本申请实施例提供的一种电子设备的硬件结构示意图之一；

图 13 是本申请实施例提供的一种电子设备的硬件结构示意图之二。

具体实施方式

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地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员获得的所有其他实施例，都属于本申请保护的范围。

本申请的说明书和权利要求书中的术语“第一”、“第二”等是用于区别类似的对象，而不用于描述特定的顺序或先后次序。应该理解这样使用的术语在适当情况下可以互换，以便本申请的实施例能够以除了在这里图示或描述的那些以外的顺序实施，且“第一”、“第二”等所区分的对象通常为一类，并不限定对象的个数，例如第一对象可以是一个，

也可以是多个。此外，说明书以及权利要求中“和/或”表示所连接对象的至少其中之一，字符“/”，一般表示前后关联对象是一种“或”的关系。

下面结合附图，通过具体的实施例及其应用场景对本申请实施例提供的语音录制方法进行详细地说明。

本申请实施例中的语音录制方法可以应用于发送语音的场景。

目前，在用户使用“通讯A”应用与联系人B聊天时，若用户想要发送语音，则需要先点击聊天框中的语音按钮，以触发电子设备显示“按住说话”控件，然后用户可以长按该“按住说话”控件并开始说话，从而电子设备开始录制语音，并在录制完成后向联系人B发送录制的语音。然而，在上述过程中，经常出现用户已经开始说话，但是没有及时按住语音录音控件（即上述的“按住说话”控件）进行录音的情况，如此，造成电子设备录制的语音缺失。

在本申请实施例提供的方案中，在电子设备检测到当前界面中存在 “按住说话”控件的情况下，电子设备可以先录制第一语音；在用户进行第一输入，例如：用户点击当前界面中的“按住说话”控件时，电子设备停止录制第一语音，并开始录制第二语音；然后电子设备可以将用户输入之前录制的第一语音和用户输入之后录制的第二语音进行处理，以得到目标语音。如此，避免了只有在用户触发对录制控件的输入后电子设备才开始录制语音，容易产生语音缺失的问题，从而保证了电子设备录制语音的完整性。

本申请实施例提供的语音录制方法的执行主体可以为语音录制装置，该语音录制装置可以为电子设备，或电子设备中的功能模块或实体。以下以电子设备为例，对本申请实施例提供的技术方案进行说明。

本申请实施例提供一种语音录制方法，图1示出了本申请实施例提供的一种语音录制方法的流程图，该方法可以应用于电子设备。如图1所示，本申请实施例提供的语音录制方法可以包括下述的步骤201至步骤204。

步骤201、在当前界面中存在语音录制控件的情况下，电子设备录制第一语音。

可选地，本申请实施例中，上述的当前界面可以为电子设备当前显示的某个应用中的界面，该应用可以为具有语音录制功能的应用，该界面可以为具有语音录制功能的应用中的会话界面，该会话界面为电子设备用户与其他联系人的会话界面。

可选地，本申请实施例中，上述的具有语音录制功能的应用可以包括以下至少一项：通讯类应用、购物类应用、短视频类应用等。具体的可以根据实际使用需求确定，本申请实施例不作限制。

可选地，本申请实施例中，上述语音录制控件可以为上述的具有语音录制功能的应用中用于开始录制语音的控件。

可选地，本申请实施例中，电子设备在显示上述应用中的一个会话界面后，该会话界面中包括语音控件/语音按钮，该语音控件用于开启语音输入功能，用户可以对该语音控件进行点击输入，以触发电子设备显示上述的语音录制控件，开始录制第一语音。

举例说明，以电子设备为手机为例进行说明，如图2（A）所示，手机显示与联系人B的会话界面10，用户可以点击该会话界面10中包括的语音按钮11，以触发手机在会话界面10中显示如图2（B）所示的语音录制控件，例如“按住说话”控件12，从而手机开始录制第一语音。

可选地，本申请实施例中，在上述的当前界面中存在语音控件的情况下，可以通过电子设备的本地录音应用（例如：录音机）录制第一语音。

可选地，本申请实施例中，电子设备可以将第一语音保存在本地录音应用中，或者文件管理应用中。

可选地，本申请实施例中，结合图 1，如图 3 所示，上述步骤 201 具体可以通过下述的步骤 201a 实现。

步骤 201a、在当前界面中存在语音录制控件，且检测到用户的目标行为信息的情况下，电子设备录制第一语音。

本申请实施例中，上述用户的目标行为信息包括以下至少一项：用户的声音信息，用户注视屏幕的视线信息，用户的嘴巴张合信息。

可选地，本申请实施例中，上述用户的声音信息可以包括以下至少一项：声音大小信息、音色、音调、声音频率等。

可选地，本申请实施例中，上述用户注视屏幕的视线信息可以包括以下至少一项：视线角度、视线方向、注视点位置等。

可选地，本申请实施例中，上述用户的嘴巴张合信息可以包括以下至少一项：嘴巴的张合状态信息、嘴巴的张合程度、嘴巴张合的口型信息。

可选地，本申请实施例中，电子设备可以通过前置摄像头来检测用户的视线信息，以确定是否存在用户注视屏幕的视线信息。

可选地，本申请实施例中，电子设备可以通过麦克风来检测用户的声音信息，以确定用户是否发出声音。

可选地，本申请实施例中，电子设备可以通过前置摄像头来检测用户的嘴巴张合信息，以确定用户是否有嘴巴张合动作。

可选地，本申请实施例中，在用户的行为满足以下至少之一：用户的眼睛注视屏幕、用户发出声音、用户有嘴巴张合动作时，电子设备可以开始录制第一语音。

可选地，本申请实施例中，电子设备可以通过前置摄像头获取用户的面部特征信息，以根据该面部特征信息判断电子设备前的用户是否为使用该电子设备的用户，以在确定电子设备前的用户为使用该电子设备的用户，且该用户的行为满足以下至少之一：用户的眼睛注视屏幕、用户发出声音、用户有嘴巴张合动作时，开始录制第一语音。

可选地，本申请实施例中，电子设备在检测到声音信息时，可以根据音色判断是否为使用该电子设备的用户的声音信息，以在确定声音信息为使用该电子设备的用户的声音信息时，录制第一语音。

本申请实施例中，通过在上述的当前界面中存在语音录制控件时，进一步检测用户的行为信息，以在检测到用户的目标行为信息时，录制第一语音，使得电子设备可以更精准地判断出用户是否有录制语音的需求，从而确定是否预先录制第一语音。如此，提高了电子设备录制语音的灵活性。

步骤 202、电子设备接收用户的第一输入。

本申请实施例中，上述第一输入为对语音录制控件的输入。

可选地，本申请实施例中，上述第一输入为用于录制语音的输入，以触发电子设备开始录制语音。

可选地，本申请实施例中，上述第一输入包括但不限于：用户通过手指或者手写笔等触控装置对语音录制控件进行触控输入，或者为用户输入的特定手势，或者为点击输入。具体的可以根据实际使用需求确定，本发明实施例不作限定。

可选地，本申请实施例中，上述特定手势可以为单击手势、滑动手势、拖动手势、压力识别手势、长按手势、面积变化手势、双按手势、双击手势中的任意一种。

可选地，本申请实施例中，上述第一输入还可以为其他可行性输入，例如：在上述的当前界面显示语音录制控件的情况下，用户对电子设备的物理按键的输入，或者为用户输入的语音指令，以触发电子设备开始录制第二语音。

步骤 203、电子设备响应于第一输入，录制第二语音。

可以理解，电子设备在录制第一语音的过程中，若接收到第一输入，则停止录制第一语音，并开始录制第二语音。

可选地，本申请实施例中，电子设备可以通过上述的当前界面对应的应用录制第二语音。

可选地，本申请实施例中，电子设备可以将第二语音保存在上述的当前界面对应的应用中，或者文件管理应用中。

可选地，本申请实施例中，在上述第一输入为用户对语音录制控件的点击输入的情况下，用户可以在电子设备录制第二语音的过程中，再次点击语音录制控件，以触发电子设备停止录制第二语音。

可选地，本申请实施例中，在上述第一输入为用户对语音录制控件的长按输入的情况下，用户开始长按输入时，电子设备停止录制第一语音，并开始录制第二语音，在用户停止长按输入时，电子设备停止录制第二语音。

举例说明，结合图 2（B），如图 4 所示，用户可以对会话界面 10 中的“按住说话”控件 12 进行输入，例如：长按输入，以触发手机停止录制第一语音，并开始录制第二语音，在手机录制第二语音的过程中，用户可以停止对“按住说话”控件 12 的长按输入，以触发手机停止录制第二语音。

步骤 204、电子设备对第一语音和第二语音处理，得到目标语音。

可选地，本申请实施例中，电子设备可以直接将第一语音拼接在第二语音之前，以得到目标语音。

示例性地，电子设备录制的第一语音为“你今天有空吗？”，且录制的第二语音为“我们去吃饭吧”，电子设备可以直接将第一语音拼接在第二语音之前，得到目标语音“你今天有空吗？我们去吃饭吧”。

可选地，本申请实施例中，电子设备可以对第一语音和第二语音中的至少一个进行降噪处理，并对降噪处理后的第一语音和第二语音进行拼接，以得到目标语音。此处的降噪处理是指将语音中的噪音或杂音进行删除。

示例性地，电子设备录制的第一语音为“今天###天气不错”，且录制的第二语音为“我们出去玩吧”，需要说明的是，此处的“#”表示噪音或杂音；电子设备可以对第一语音进行降噪处理，并对降噪处理后的第一语音和第二语音进行拼接，以得到目标语音“今天天气不错，我们出去玩吧”。

可选地，本申请实施例中，电子设备可以检测第一语音中是否有静音部分，并在含有

静音部分的情况下，将第一语音中的静音部分进行裁切，以得到处理后的第一语音，并对处理后的第一语音和第二语音进行拼接处理，得到目标语音。需要说明的是，上述静音部分为用户未发出声音的部分。

示例性地，假设第一语音的时长为 6 秒，在检测到第一语音的第 1 秒至第 3 秒为静音的情况下，电子设备可以将第一语音中第 1 秒至第 3 秒的静音部分进行裁切，以得到处理后的时长为 3 秒的第一语音，并对处理后的第一语音和第二语音进行拼接处理，得到目标语音。

可选地，本申请实施例中，电子设备可以将第一语音中预设时长内的语音与第二语音进行拼接处理，以得到目标语音，该预设时长可以为电子设备接收到第一输入之前的预设时长。

例如：将第一输入之前的三秒内的语音与第二语音进行拼接处理，以得到目标语音。

本申请实施例提供一种语音录制方法，在当前界面中存在语音录制控件的情况下，电子设备录制第一语音，并在接收到对语音录制控件的第一输入后，录制第二语音，从而对第一语音和第二语音处理，得到目标语音。本方案中，由于电子设备在检测到当前界面中存在语音录制控件的情况下，可以先录制第一语音，并在实际接收到用户对语音录制控件的输入时，停止录制第一语音，并开始录制第二语音，然后电子设备可以将用户输入之前录制的第一语音和用户输入之后录制的第二语音进行处理，以得到目标语音。如此，避免了只有在用户触发对录制控件的输入后电子设备才开始录制语音，容易产生语音缺失的问题，从而保证了电子设备录制语音的完整性。

可选地，本申请实施例中，结合图 1，如图 5 所示，上述步骤 204 具体可以通过下述的步骤 204a 和步骤 204b 实现，或者，结合图 1，如图 6 所示，上述步骤 204 具体可以通过下述的步骤 204a 和步骤 204c 实现。

步骤 204a、电子设备将第一语音与第二语音进行内容比对。

可选地，本申请实施例中，电子设备可以将第一语音和第二语音的内容均转为文字后，再对比第一语音和第二语音的内容，以获得第一语音和第二语音的相似度。若相似度大于一个阈值，则确认第一语音的内容和第二语音的内容相同；若相似度小于或等于一个阈值，则确认第一语音的内容和第二语音的内容不同。

步骤 204b、在第一语音与第二语音的内容相同的情况下，电子设备删除第一语音，并将第二语音作为目标语音。

示例性地，在第一语音的内容为“吃饭了吗”，第二语音的内容同样为“吃饭了吗”的情况下，电子设备可以删除第一语音，并将第二语音作为目标语音。

示例性地，在第一语音的内容为“我们一会儿去商场吧”，第二语音的内容为“一会儿我们去商场吧”的情况下，电子设备同样可以删除第一语音，将第二语音作为目标语音。

可选地，本申请实施例中，在第一语音与第二语音的内容相同时，电子设备也可以删除第二语音，并将第一语音作为目标语音。

步骤 204c、在第一语音与第二语音的内容不同的情况下，电子设备对第一语音和第二语音拼接处理，得到目标语音。

示例性地，第一语音的内容为“早上好”，第二语音的内容为“吃饭了吗”，即第一语音与第二语音的内容不同时，电子设备可以将第一语音和第二语音进行拼接处理，以得

到目标语音。

需要说明的是，针对此处对第一语音和第二语音拼接处理的具体说明，可以参见上述实施例的步骤 204 中对拼接处理的描述，即与上述步骤 204 中的拼接处理的具体方案相同，此处不再赘述。

本申请实施例中，电子设备在得到第一语音和第二语音之后，通过比对第一语音和第二语音的内容，使得电子设备可以根据比对的结果对第一语音和第二语音进行处理。如此，提高了电子设备处理录制语音的灵活性。

可选地，本申请实施例中，在电子设备对第一语音和第二语音进行拼接处理之后，电子设备可以显示菜单栏，该菜单栏用于用户对拼接后的语音进行处理。

可选地，本申请实施例中，上述菜单栏可以包括以下至少一项：变声、语音加速、语音减速、添加背景音乐等。

可选地，本申请实施例中，结合图 1，如图 7 所示，上述步骤 204 具体可以通过下述的步骤 204d 和步骤 204e 实现，或者，结合图 1，如图 8 所示，上述步骤 204 具体可以通过下述的步骤 204d 和步骤 204f 实现。

步骤 204d、电子设备对第一语音和第二语音拼接处理。

需要说明的是，针对此处对第一语音和第二语音拼接处理的具体说明，可以参见上述实施例的步骤 204 中对拼接处理的描述，即与上述步骤 204 中的拼接处理的具体方案相同，此处不再赘述。

步骤 204e、在拼接后的语音为语义连贯状态的情况下，电子设备将拼接后的语音作为目标语音。

可选地，本申请实施例中，电子设备可以使用自然语言处理（Natural Language Processing，NLP）技术对第一语音和第二语音进行语义比对，以判断第一语音和第二语音中，字、词、以及句子之间是否连贯。

可选地，本申请实施例中，电子设备可以分别获取第一语音和第二语音的语义特征信息，然后根据该语义特征信息，得到第一语音和第二语音的语义相似度，若相似度大于一个阈值，则确认第一语音与第二语音为语义连贯状态；若相似度小于或等于一个阈值，则确认第一语音与第二语音为语义不连贯状态。

示例性地，在第一语音为“你”，第二语音为“起床了吗”，拼接后的语音为“你起床了吗”，即拼接后的语音为语义连贯状态的情况下，电子设备可以将拼接后的语音作为目标语音。

示例性地，在第一语音为“你这会儿忙吗？”，第二语音为“我需要你的帮助”，拼接后的语音为“你这会儿忙吗？我需要你的帮助”，即拼接后的语音为语义连贯状态的情况下，电子设备可以将拼接后的语音作为目标语音。

步骤 204f、在拼接后的语音为语义不连贯状态的情况下，电子设备将第一语音或第二语音作为目标语音。

示例性地，在第一语音为“早上好，你”，第二语音为“饭了吗”，拼接后的语音为“早上好，你饭了吗”，即拼接后的语音为语义不连贯状态的情况下，电子设备可以将第一语音：“早上好，你”或第二语音：“饭了吗”作为目标语音。

可选地，本申请实施例中，电子设备可以对第一语音和第二语音的语音长度、语音质

量、语音内容的重要程度等进行对比，以根据对比结果确定最终作为目标语音的语音。

示例性地，在第一语音的语音长度大于第二语音的语音长度时，电子设备可以将第一语音作为目标语音。

示例性地，在第一语音的语音质量高于第二语音的语音质量，例如第一语音的语音清晰度高于第二语音的语音清晰度时，电子设备可以将第一语音作为目标语音。

示例性地，在第一语音的语音内容的重要程度低于第二语音的语音内容的重要程度，例如第一语音的语音内容为“忙完了没？”，第二语音的语音内容为“下午三点会议室开会”时，电子设备可以将第二语音作为目标语音。

本申请实施例中，电子设备在得到第一语音和第二语音之后，通过对第一语音和第二语音进行预拼接处理，并检测拼接后的语音的语义连贯状态，使得电子设备可以根据拼接后的语音的语义连贯状态对拼接后的语音进行对应的处理。如此，提高了电子设备处理录制语音的灵活性和多样性。

可选地，本申请实施例中，本申请实施例提供的语音录制方法还包括下述的步骤 301 至步骤 303，或者，本申请实施例提供的语音录制方法还包括下述的步骤 301、步骤 302 和步骤 304。

步骤 301、在第一语音和第二语音为语义不连贯状态的情况下，电子设备在当前界面中显示提示信息。

本申请实施例中，上述提示信息用于提示是否将拼接后的语音作为目标语音。

可选地，本申请实施例中，电子设备可以在上述的当前界面中显示一个窗口，该窗口中包括上述的提示信息，以及第一控件和第二控件，该第一控件用于确定将拼接后的语音作为目标语音，第二控件用于确定将第一语音或第二语音作为目标语音。

例如，上述提示信息可以为“是否将拼接后的语音作为发送语音”。

举例说明，如图 9 所示，手机确定拼接后的语音为语义不连贯状态时，可以在会话界面 10 中显示窗口 13，该窗口 13 中包括提示信息“是否将拼接后的语音作为发送语音”，并且包括第一控件，例如：“是”控件 14，以及包括第二控件，例如：“否”控件 15。

步骤 302、电子设备接收用户对提示信息的第二输入。

可选地，本申请实施例中，上述第二输入可以为对第一控件或第二控件的输入。

可选地，本申请实施例中，上述第二输入包括但不限于：用户通过手指或者手写笔等触控装置对上述第一控件或第二控件进行触控输入，或者为用户输入的特定手势，或者为点击输入，或者为其他可行性输入。具体的可以根据实际使用需求确定，本发明实施例不作限定。

可选地，本申请实施例中，上述特定手势可以为单击手势、滑动手势、拖动手势、压力识别手势、长按手势、面积变化手势、双按手势、双击手势中的任意一种。

步骤 303、电子设备响应于第二输入，在第二输入确认是的情况下，将拼接后的语音作为目标语音。

可以理解，在第二输入为对第一控件的输入时，电子设备可以将拼接后的语音作为目标语音。

步骤 304、电子设备响应于第二输入，在第二输入确认否的情况下，将第一语音或第二语音作为目标语音。

可以理解，在第二输入为对第二控件的输入时，电子设备可以将第一语音或第二语音作为目标语音。

举例说明，结合上述图 9，用户可以对“是”控件 14 进行第二输入，以触发手机将拼接后的语音作为目标语音，或者用户可以对“否”控件 15 进行第二输入，以触发手机将第一语音或第二语音作为目标语音。

需要说明的是，针对电子设备将第一语音或第二语音作为目标语音的具体方案，可以参见上述实施例的上述步骤 204f 中的描述，此处不再赘述。

可选地，本申请实施例中，在拼接后的语音为语义不连贯状态的情况下，电子设备也可以删除第一语音和第二语音，并提醒用户重新录制语音。

本申请实施例中，电子设备在确定拼接后的语音为语义不连贯状态时，通过显示提示信息，使得用户可以根据该提示信息灵活地选择最终发送的语音，即目标语音。如此，提高了电子设备发送录制语音的多样性和灵活性。

需要说明的是，本申请实施例中，在执行上述步骤 204d 之后，在第一语音和第二语音为语义不连贯状态的情况下：电子设备可以执行上述步骤 204f；或者，电子设备可以执行上述步骤 301 至步骤 303；或者，电子设备可以执行上述步骤 301、步骤 302 和步骤 304。

可选地，本申请实施例中，本申请实施例提供的语音录制方法还可以包括下述的步骤 401。

步骤 401、在当前界面中存在语音录制控件，且录制第一语音的过程中关闭当前界面的情况下，电子设备删除录制的第一语音。

可选地，本申请实施例中，电子设备可以在检测到上述的当前界面中不存在语音录制控件，例如用户触发电子设备在上述的当前界面显示文字输入键盘的情况下，删除录制的第一语音。

可选地，本申请实施例中，电子设备可以在检测到退出上述的当前界面对应的应用时，删除录制的第一语音。

可选地，本申请实施例中，电子设备可以在检测到显示会话列表界面时，删除录制的第一语音。

可选地，本申请实施例中，电子设备可以在接收到来电信息、视频通话、或语音通话时，删除录制的第一语音。

本申请实施例中，电子设备在录制第一语音的过程中，通过检测上述的当前界面是否关闭，使得在上述的当前界面关闭的情况下，电子设备可以删除录制的第一语音。如此，提高了电子设备录制语音的灵活性。

需要说明的是，本申请实施例提供的语音录制方法，执行主体可以为语音录制装置。本申请实施例中以语音录制装置执行语音录制方法为例，说明本申请实施例提供的语音录制装置。

图 10 示出了本申请实施例中涉及的语音录制装置的一种可能的结构示意图。如图 10 所示，该语音录制装置 70 可以包括：录制模块 71、接收模块 72 和处理模块 73。

其中，录制模块 71，用于在当前界面中存在语音录制控件的情况下，录制第一语音。接收模块 72，用于接收用户的第一输入，该第一输入为对语音录制控件的输入。录制模块 71，还用于响应于接收模块 72 接收的第一输入，录制第二语音。处理模块 73，用于对录

制模块 71 得到的第一语音和第二语音处理，得到目标语音。

本申请实施例提供一种语音录制装置，由于语音录制装置在检测到当前界面中存在语音录制控件的情况下，可以先录制第一语音，并在实际接收到用户对语音录制控件的输入时，停止录制第一语音，并开始录制第二语音，然后语音录制装置可以将用户输入之前录制的第一语音和用户输入之后录制的第二语音进行处理，以得到目标语音。如此，避免了只有在用户触发对录制控件的输入后语音录制装置才开始录制语音，容易产生语音缺失的问题，从而保证了语音录制装置录制语音的完整性。

在一种可能的实现方式中，上述录制模块 71，具体用于在当前界面中存在语音录制控件，且检测到用户的目标行为信息的情况下，录制第一语音，用户的目标行为信息包括以下至少一项：用户的声音信息，用户注视屏幕的视线信息，用户的嘴巴张合信息。

在一种可能的实现方式中，上述处理模块 73，具体用于将第一语音与第二语音进行内容比对；在第一语音与第二语音的内容相同的情况下，删除第一语音，并将第二语音作为目标语音；或者，在第一语音与第二语音的内容不同的情况下，对第一语音和第二语音拼接处理，得到目标语音。

在一种可能的实现方式中，上述处理模块 73，具体用于对第一语音和第二语音拼接处理；在拼接后的语音为语义连贯状态的情况下，将拼接后的语音作为目标语音；或者，在拼接后的语音为语义不连贯状态的情况下，将第一语音或第二语音作为目标语音。

在一种可能的实现方式中，结合图 10，如图 11 所示，本申请实施例提供的语音录制装置 70 还包括：删除模块 74。该删除模块 74，用于在当前界面中存在语音录制控件，且录制第一语音的过程中关闭当前界面的情况下，删除录制的第一语音。

本申请实施例中的语音录制装置可以是电子设备，也可以是电子设备中的部件，例如集成电路或芯片。该电子设备可以是终端，也可以为除终端之外的其他设备。示例性的，电子设备可以为手机、平板电脑、笔记本电脑、掌上电脑、车载电子设备、移动上网装置（Mobile Internet Device，MID）、增强现实(Augmented Reality，AR)/虚拟现实(Virtual Reality，VR)设备、机器人、可穿戴设备、超级移动个人计算机（Ultra-Mobile Personal Computer，UMPC）、上网本或者个人数字助理（personal digital assistant，PDA）等，还可以为服务器、网络附属存储器（Network Attached Storage，NAS）、个人计算机（personal computer，PC）、电视机（television，TV）、柜员机或者自助机等，本申请实施例不作具体限定。

本申请实施例中的语音录制装置可以为具有操作系统的装置。该操作系统可以为安卓（Android）操作系统，可以为 ios 操作系统，还可以为其他可能的操作系统，本申请实施例不作具体限定。

本申请实施例提供的语音录制装置能够实现上述方法实施例实现的各个过程，为避免重复，这里不再赘述。

可选地，如图 12 所示，本申请实施例还提供一种电子设备 900，包括处理器 901 和存储器 902，存储器 902 上存储有可在所述处理器 901 上运行的程序或指令，该程序或指令被处理器 901 执行时实现上述方法实施例的各个步骤，且能达到相同的技术效果，为避免重复，这里不再赘述。

需要说明的是，本申请实施例中的电子设备包括上述所述的移动电子设备和非移动电

子设备。

图 13 为实现本申请实施例的一种电子设备的硬件结构示意图。

该电子设备 100 包括但不限于：射频单元 101、网络模块 102、音频输出单元 103 、输入单元 104、传感器 105、显示单元 106、用户输入单元 107 、接口单元 108、存储器 109、以及处理器 110 等部件。

本领域技术人员可以理解，电子设备 100 还可以包括给各个部件供电的电源（比如电池），电源可以通过电源管理系统与处理器 110 逻辑相连，从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。图 13 中示出的电子设备结构并不构成对电子设备的限定，电子设备可以包括比图示更多或更少的部件，或者组合某些部件，或者不同的部件布置，在此不再赘述。

其中，输入单元 104，用于在当前界面中存在语音录制控件的情况下，录制第一语音。

用户输入单元 107，用于接收用户的第一输入，该第一输入为对语音录制控件的输入。

输入单元 104，还用于响应于用户输入单元 107 接收的第一输入，录制第二语音。

处理器 110，用于对输入单元 104 得到的第一语音和第二语音处理，得到目标语音。

本申请实施例提供一种电子设备，由于电子设备在检测到当前界面中存在语音录制控件的情况下，可以先录制第一语音，并在实际接收到用户对语音录制控件的输入时，停止录制第一语音，并开始录制第二语音，然后电子设备可以将用户输入之前录制的第一语音和用户输入之后录制的第二语音进行处理，以得到目标语音。如此，避免了只有在用户触发对录制控件的输入后电子设备才开始录制语音，容易产生语音缺失的问题，从而保证了电子设备录制语音的完整性。

可选地，输入单元 104，具体用于在当前界面中存在语音录制控件，且检测到用户的目标行为信息的情况下，录制第一语音，用户的目标行为信息包括以下至少一项：用户的声音信息，用户注视屏幕的视线信息，用户的嘴巴张合信息。

可选地，处理器 110，具体用于将第一语音与第二语音进行内容比对；在第一语音与第二语音的内容相同的情况下，删除第一语音，并将第二语音作为目标语音；或者，具体用于将第一语音与第二语音进行内容比对，在第一语音与第二语音的内容不同的情况下，对第一语音和第二语音拼接处理，得到目标语音。

可选地，处理器 110，具体用于对第一语音和第二语音拼接处理；在拼接后的语音为语义连贯状态的情况下，将拼接后的语音作为目标语音；或者，具体用于对第一语音和第二语音拼接处理，在拼接后的语音为语义不连贯状态的情况下，将第一语音或第二语音作为目标语音。

可选地，处理器 110，还用于在当前界面中存在语音录制控件，且录制第一语音的过程中关闭当前界面的情况下，删除录制的第一语音。

本申请实施例提供的电子设备能够实现上述方法实施例实现的各个过程，且能达到相同的技术效果，为避免重复，这里不再赘述。

本实施例中各种实现方式具有的有益效果具体可以参见上述方法实施例中相应实现方式所具有的有益效果，为避免重复，此处不再赘述。

应理解的是，本申请实施例中，输入单元 104 可以包括图形处理器（Graphics Processing Unit，GPU）1041 和麦克风 1042，图形处理器 1041 对在视频捕获模式或图像捕获模式中

由图像捕获装置（如摄像头）获得的静态图片或视频的图像数据进行处理。显示单元 106 可包括显示面板 1061，可以采用液晶显示器、有机发光二极管等形式来配置显示面板 1061。用户输入单元 107 包括触控面板 1071 以及其他输入设备 1072 中的至少一种。触控面板 1071，也称为触摸屏。触控面板 1071 可包括触摸检测装置和触摸控制器两个部分。其他输入设备 1072 可以包括但不限于物理键盘、功能键（比如音量控制按键、开关按键等）、轨迹球、鼠标、操作杆，在此不再赘述。

存储器 109 可用于存储软件程序以及各种数据。存储器 109 可主要包括存储程序或指令的第一存储区和存储数据的第二存储区，其中，第一存储区可存储操作系统、至少一个功能所需的应用程序或指令（比如声音播放功能、图像播放功能等）等。此外，存储器 109 可以包括易失性存储器或非易失性存储器，或者，存储器 109 可以包括易失性和非易失性存储器两者。其中，非易失性存储器可以是只读存储器(Read-Only Memory，ROM)、可编程只读存储器(Programmable ROM，PROM)、可擦除可编程只读存储器(Erasable PROM，EPROM)、电可擦除可编程只读存储器(Electrically EPROM，EEPROM)或闪存。易失性存储器可以是随机存取存储器(Random Access Memory，RAM)，静态随机存取存储器(Static RAM，SRAM)、动态随机存取存储器(Dynamic RAM，DRAM)、同步动态随机存取存储器(Synchronous DRAM，SDRAM)、双倍数据速率同步动态随机存取存储器(Double Data Rate SDRAM，DDRSDRAM)、增强型同步动态随机存取存储器(Enhanced SDRAM，ESDRAM)、同步连接动态随机存取存储器(Synch link DRAM，SLDRAM)和直接内存总线随机存取存储器(Direct Rambus RAM，DRRAM)。本申请实施例中的存储器 109 包括但不限于这些和任意其它适合类型的存储器。

处理器 110 可包括一个或多个处理单元；可选的，处理器 110 集成应用处理器和调制解调处理器，其中，应用处理器主要处理涉及操作系统、用户界面和应用程序等的操作，调制解调处理器主要处理无线通信信号，如基带处理器。可以理解的是，上述调制解调处理器也可以不集成到处理器 110 中。

本申请实施例还提供一种可读存储介质，所述可读存储介质上存储有程序或指令，该程序或指令被处理器执行时实现上述方法实施例的各个过程，且能达到相同的技术效果，为避免重复，这里不再赘述。

其中，所述处理器为上述实施例中所述的电子设备中的处理器。所述可读存储介质，包括计算机可读存储介质，如计算机只读存储器 ROM、随机存取存储器 RAM、磁碟或者光盘等。

本申请实施例另提供了一种芯片，所述芯片包括处理器和通信接口，所述通信接口和所述处理器耦合，所述处理器用于运行程序或指令，实现上述方法实施例的各个过程，且能达到相同的技术效果，为避免重复，这里不再赘述。

应理解，本申请实施例提到的芯片还可以称为系统级芯片、系统芯片、芯片系统或片上系统芯片等。

本申请实施例提供一种计算机程序产品，该程序产品被存储在存储介质中，该程序产品被至少一个处理器执行以实现如上述方法实施例的各个过程，且能达到相同的技术效果，为避免重复，这里不再赘述。

需要说明的是，在本文中，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、物品或者装置不仅包括那些要素，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、物品或者装置所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括该要素的过程、方法、物品或者装置中还存在另外的相同要素。此外，需要指出的是，本申请实施方式中的方法和装置的范围不限按示出或讨论的顺序来执行功能，还可包括根据所涉及的功能按基本同时的方式或按相反的顺序来执行功能，例如，可以按不同于所描述的次序来执行所描述的方法，并且还可以添加、省去、或组合各种步骤。另外，参照某些示例所描述的特征可在其他示例中被组合。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以计算机软件产品的形式体现出来，该计算机软件产品存储在一个存储介质（如ROM/RAM、磁碟、光盘）中，包括若干指令用以使得一台终端（可以是手机，计算机，服务器，或者网络设备等）执行本申请各个实施例所述的方法。

上面结合附图对本申请的实施例进行了描述，但是本申请并不局限于上述的具体实施方式，上述的具体实施方式仅仅是示意性的，而不是限制性的，本领域的普通技术人员在本申请的启示下，在不脱离本申请宗旨和权利要求所保护的范围情况下，还可做出很多形式，均属于本申请的保护之内。

权 利 要 求 书

1. 一种语音录制方法，所述方法包括：

在当前界面中存在语音录制控件的情况下，录制第一语音；

接收用户的第一输入，所述第一输入为对所述语音录制控件的输入；

响应于所述第一输入，录制第二语音；

对所述第一语音和所述第二语音处理，得到目标语音。

2. 根据权利要求 1 所述的方法，其中，所述在当前界面中存在语音录制控件的情况下，录制第一语音，包括：

在所述当前界面中存在所述语音录制控件，且检测到用户的目标行为信息的情况下，录制所述第一语音，所述用户的目标行为信息包括以下至少一项：用户的声音信息，用户注视屏幕的视线信息，用户的嘴巴张合信息。

3. 根据权利要求 1 所述的方法，其中，所述对所述第一语音和所述第二语音处理，得到目标语音，包括：

将所述第一语音与所述第二语音进行内容比对；

在所述第一语音与所述第二语音的内容相同的情况下，删除所述第一语音，并将所述第二语音作为所述目标语音；或者，

在所述第一语音与所述第二语音的内容不同的情况下，对所述第一语音和所述第二语音拼接处理，得到所述目标语音。

4. 根据权利要求 1 或 3 所述的方法，其中，所述对所述第一语音和所述第二语音处理，得到目标语音，包括：

对所述第一语音和所述第二语音拼接处理；

在拼接后的语音为语义连贯状态的情况下，将拼接后的语音作为所述目标语音；或者，

在拼接后的语音为语义不连贯状态的情况下，将所述第一语音或所述第二语音作为所述目标语音。

5. 根据权利要求 1 所述的方法，其中，所述方法还包括：

在所述当前界面中存在所述语音录制控件，且录制所述第一语音的过程中关闭所述当前界面的情况下，删除录制的所述第一语音。

6. 一种语音录制装置，所述语音录制装置包括：录制模块、接收模块和处理模块；

所述录制模块，用于在当前界面中存在语音录制控件的情况下，录制第一语音；

所述接收模块，用于接收用户的第一输入，所述第一输入为对所述语音录制控件的输入；

所述录制模块，还用于响应于所述接收模块接收的所述第一输入，录制第二语音；

所述处理模块，用于对所述录制模块得到的所述第一语音和所述第二语音处理，得到目标语音。

7. 根据权利要求 6 所述的装置，其中，所述录制模块，具体用于在所述当前界面中存在所述语音录制控件，且检测到用户的目标行为信息的情况下，录制第一语音，所述用户的目标行为信息包括以下至少一项：用户的声音信息，用户注视屏幕的视线信息，用户的嘴巴张合信息。

8. 根据权利要求 6 所述的装置，其中，所述处理模块，具体用于：

将所述第一语音与所述第二语音进行内容比对；

在所述第一语音与所述第二语音的内容相同的情况下，删除所述第一语音，并将所述第二语音作为所述目标语音；或者，在所述第一语音与所述第二语音的内容不同的情况下，对所述第一语音和所述第二语音拼接处理，得到所述目标语音。

9. 根据权利要求 6 或 8 所述的装置，其中，所述处理模块，具体用于：

对所述第一语音和所述第二语音拼接处理；

在拼接后的语音为语义连贯状态的情况下，将拼接后的语音作为所述目标语音；或者，在拼接后的语音为语义不连贯状态的情况下，将所述第一语音或所述第二语音作为所述目标语音。

10. 根据权利要求 6 所述的装置，其中，所述语音录制装置还包括：删除模块；所述删除模块，用于在所述当前界面中存在所述语音录制控件，且录制所述第一语音的过程中关闭所述当前界面的情况下，删除录制的所述第一语音。

11. 一种电子设备，包括处理器和存储器，所述存储器存储可在所述处理器上运行的程序或指令，所述程序或指令被所述处理器执行时实现如权利要求 1-5 中任一项所述的语音录制方法的步骤。

12. 一种可读存储介质，所述可读存储介质上存储程序或指令，所述程序或指令被处理器执行时实现如权利要求 1-5 中任一项所述的语音录制方法的步骤。

13. 一种芯片，所述芯片包括处理器和通信接口，所述通信接口和所述处理器耦合，所述处理器用于运行程序或指令，实现如权利要求 1-5 中任一项所述的语音录制方法。

14. 一种计算机程序产品，所述程序产品被至少一个处理器执行以实现如权利要求 1-5 中任一项所述的语音录制方法。

15. 一种电子设备，包括所述电子设备被配置成用于执行如权利要求 1-5 中任一项所述的语音录制方法。

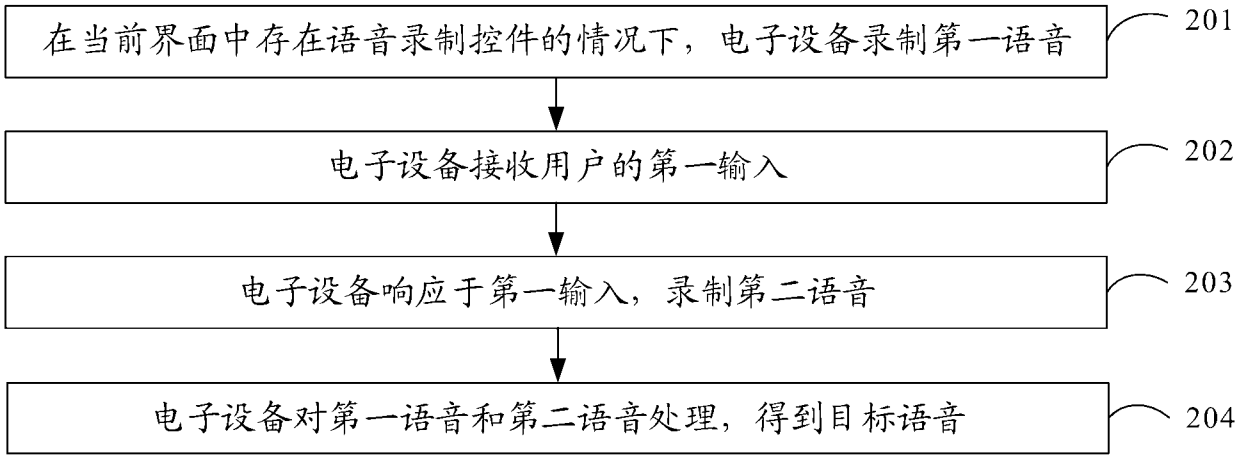

在当前界面中存在语音录制控件的情况下，电子设备录制第一语音
201
电子设备接收用户的第一输入
202
电子设备响应于第一输入，录制第二语音
203
电子设备对第一语音和第二语音处理，得到目标语音
204

图 1

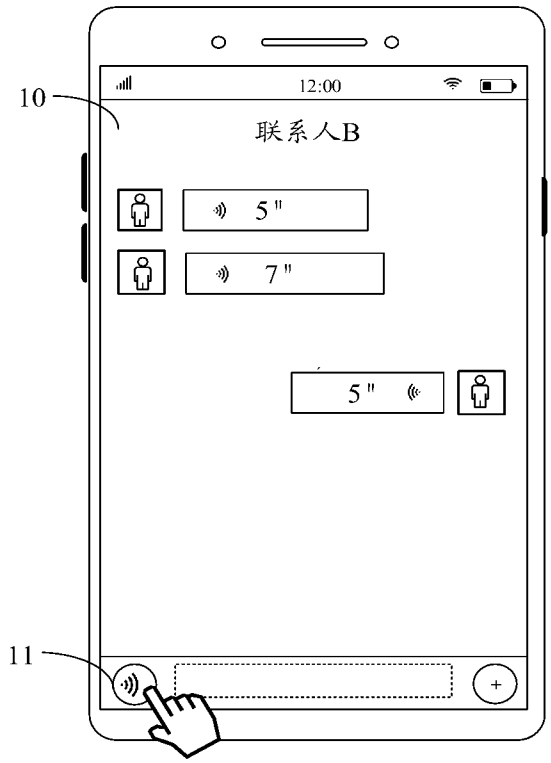

10
12:00
联系人B
5"
7"
5"
11

图 2（A）

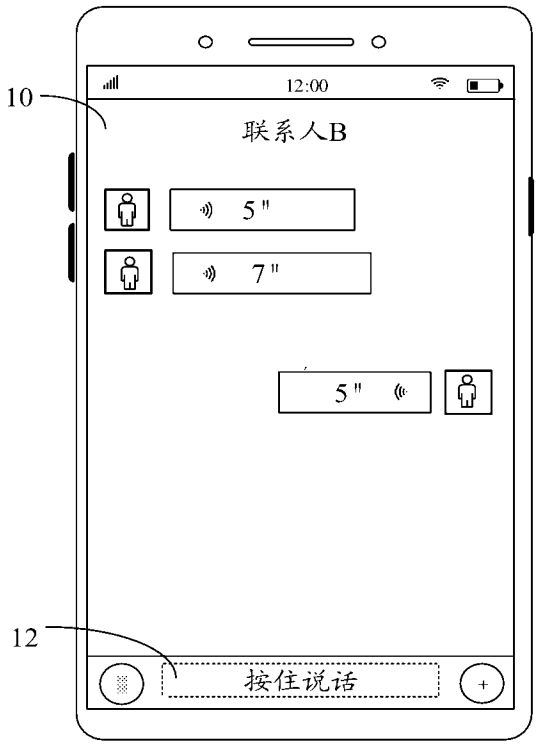

10
12:00
联系人B
5"
7"
5"
12
按住说话

图 2（B）

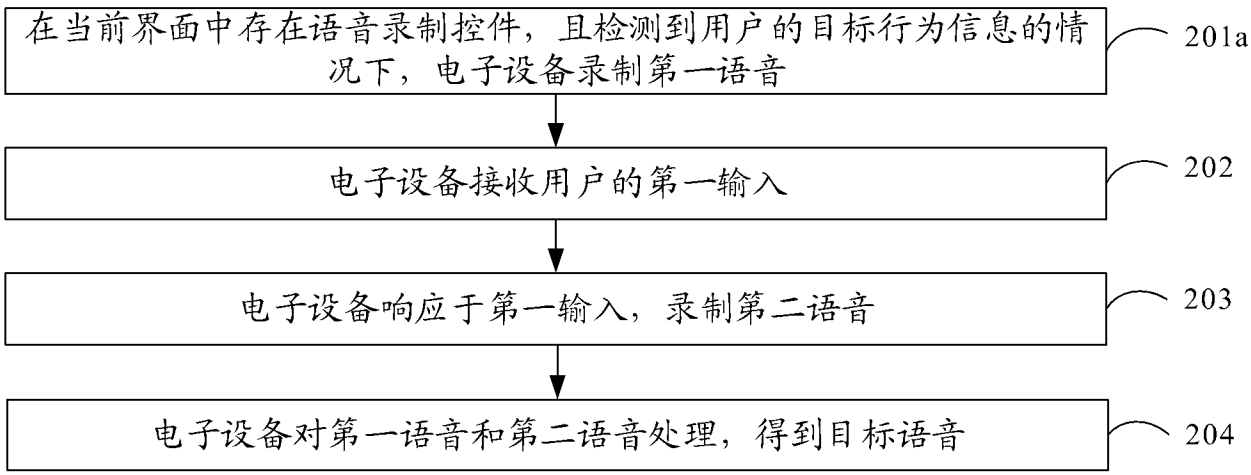

在当前界面中存在语音录制控件，且检测到用户的目标行为信息的情况下，电子设备录制第一语音
201a
电子设备接收用户的第一输入
202
电子设备响应于第一输入，录制第二语音
203
电子设备对第一语音和第二语音处理，得到目标语音
204

图 3

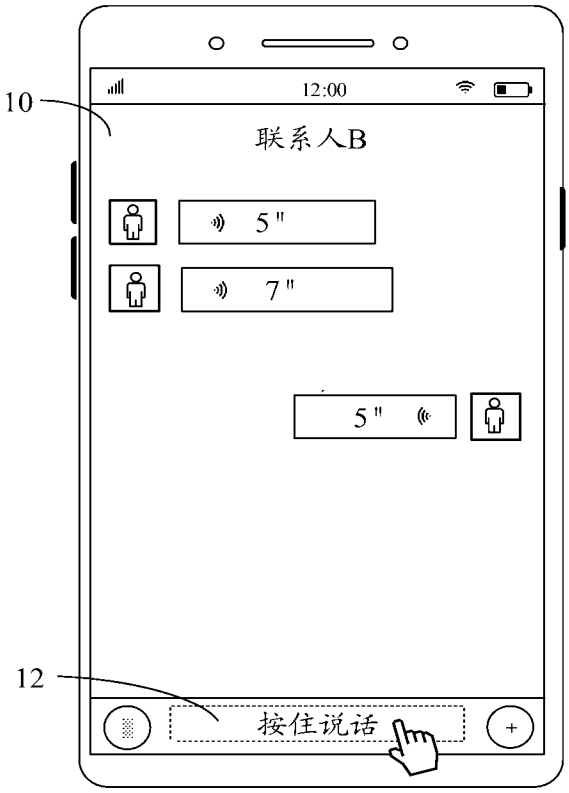

10
12:00
联系人B
5"
7"
5"
12
按住说话

图 4

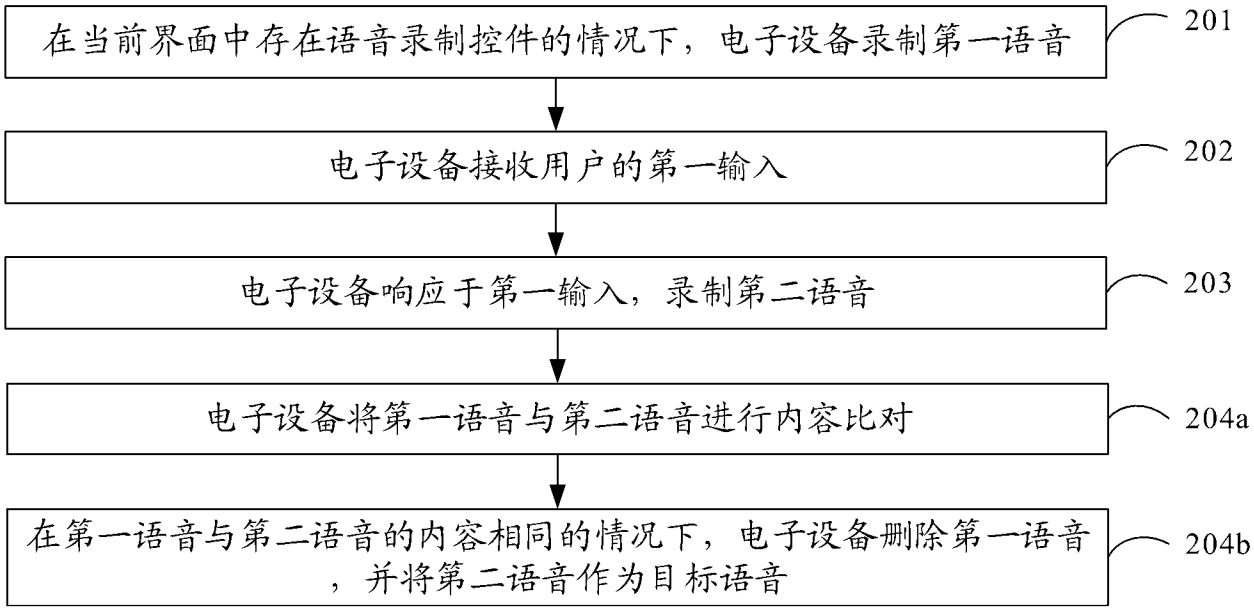

在当前界面中存在语音录制控件的情况下，电子设备录制第一语音
201
电子设备接收用户的第一输入
202
电子设备响应于第一输入，录制第二语音
203
电子设备将第一语音与第二语音进行内容比对
204a
在第一语音与第二语音的内容相同的情况下，电子设备删除第一语音，并将第二语音作为目标语音
204b

图 5

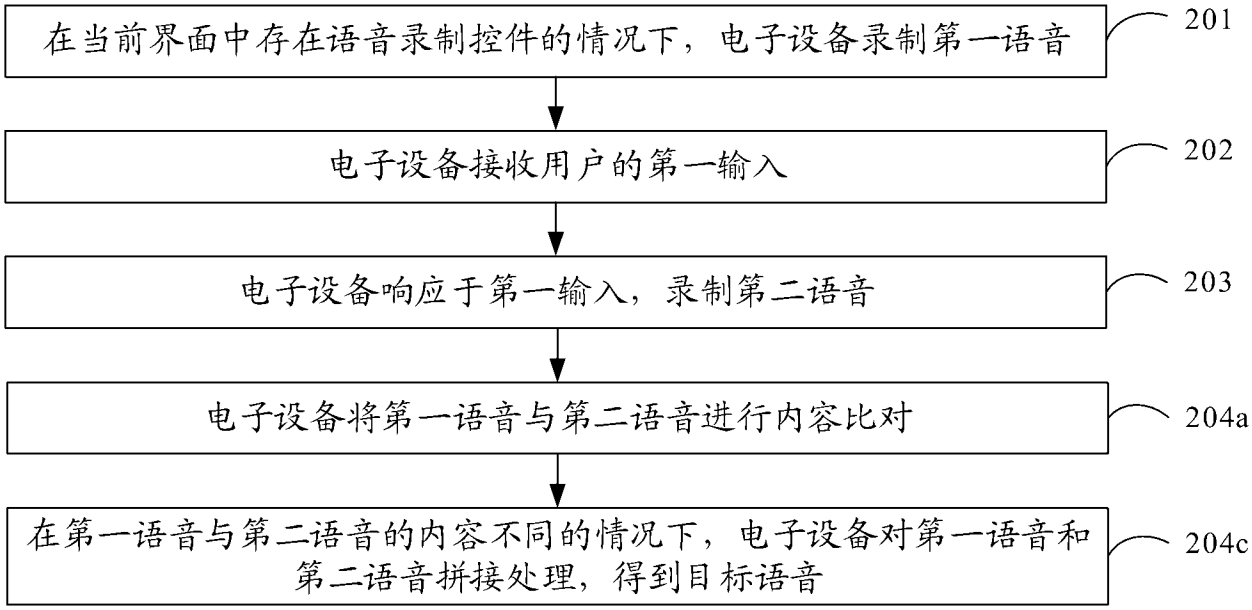

在当前界面中存在语音录制控件的情况下，电子设备录制第一语音
201
电子设备接收用户的第一输入
202
电子设备响应于第一输入，录制第二语音
203
电子设备将第一语音与第二语音进行内容比对
204a
在第一语音与第二语音的内容不同的情况下，电子设备对第一语音和第二语音拼接处理，得到目标语音
204c

图 6

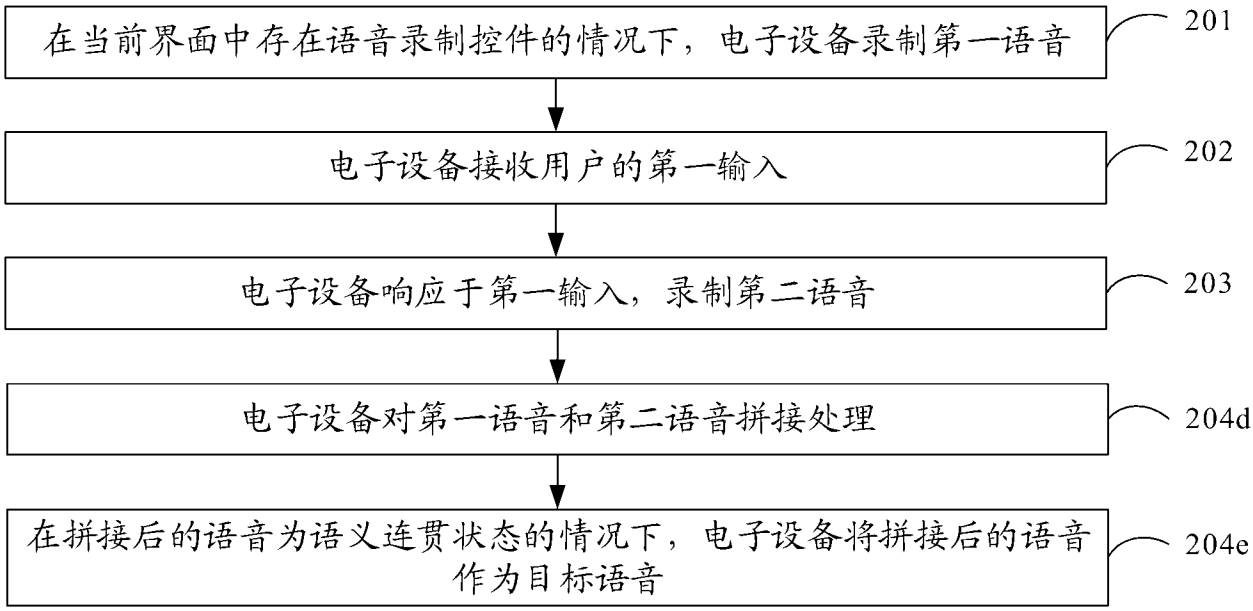

在当前界面中存在语音录制控件的情况下，电子设备录制第一语音
201
电子设备接收用户的第一输入
202
电子设备响应于第一输入，录制第二语音
203
电子设备对第一语音和第二语音拼接处理
204d
在拼接后的语音为语义连贯状态的情况下，电子设备将拼接后的语音作为目标语音
204e

图 7

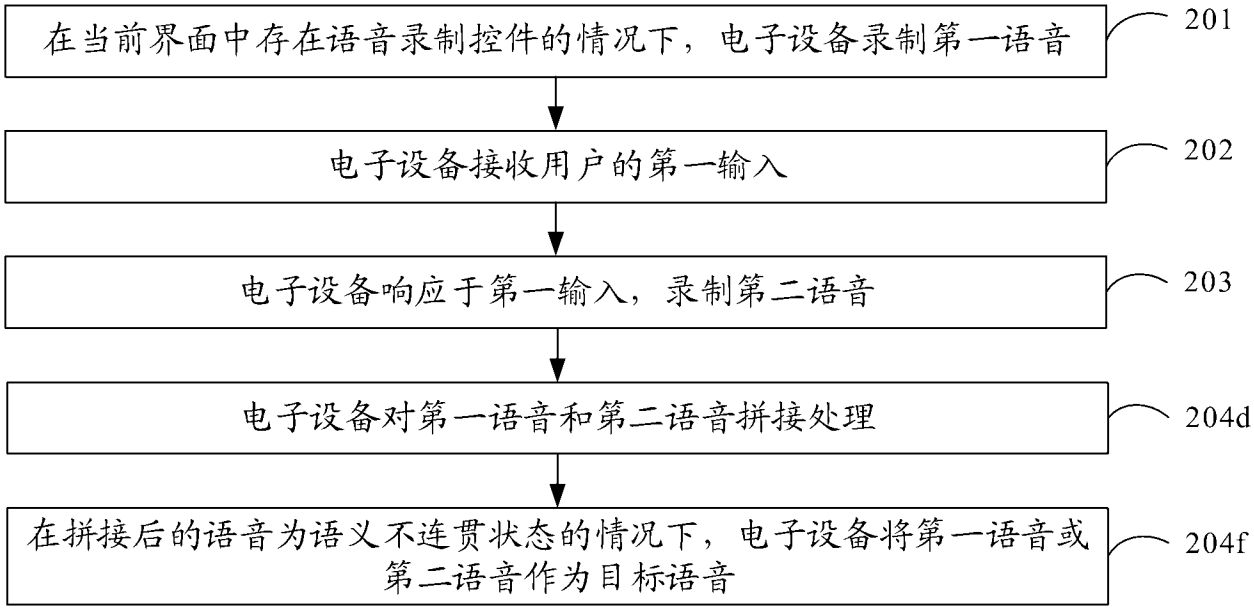

在当前界面中存在语音录制控件的情况下，电子设备录制第一语音
201
电子设备接收用户的第一输入
202
电子设备响应于第一输入，录制第二语音
203
电子设备对第一语音和第二语音拼接处理
204d
在拼接后的语音为语义不连贯状态的情况下，电子设备将第一语音或第二语音作为目标语音
204f

图 8

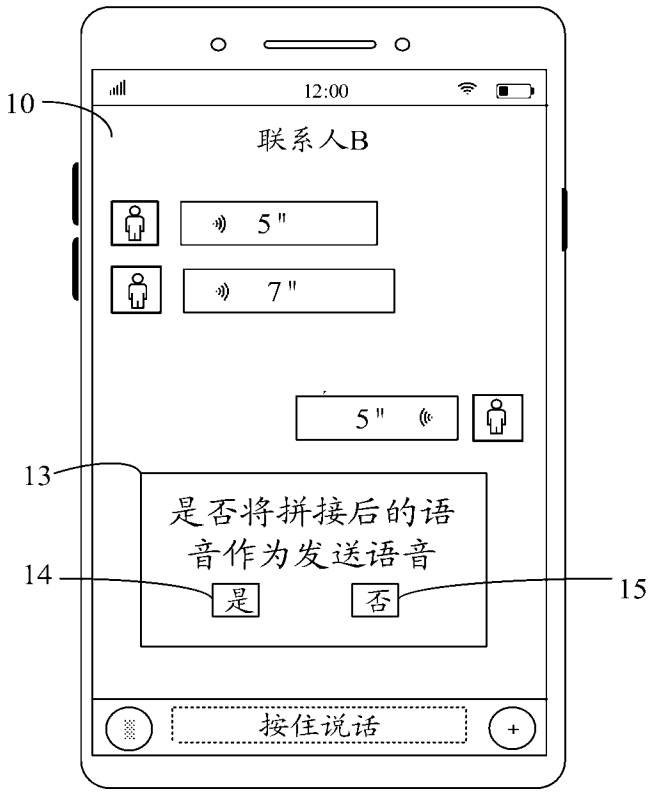

10
12:00
联系人B
5"
7"
5"
13
是否将拼接后的语音作为发送语音
14
是
否
15
按住说话

图 9

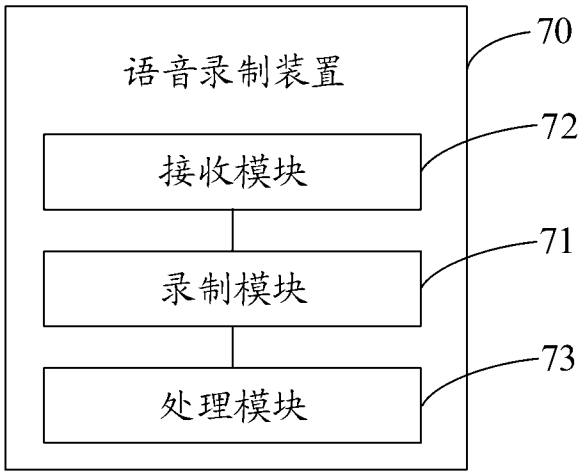

语音录制装置
70
接收模块
72
录制模块
71
处理模块
73

图 10

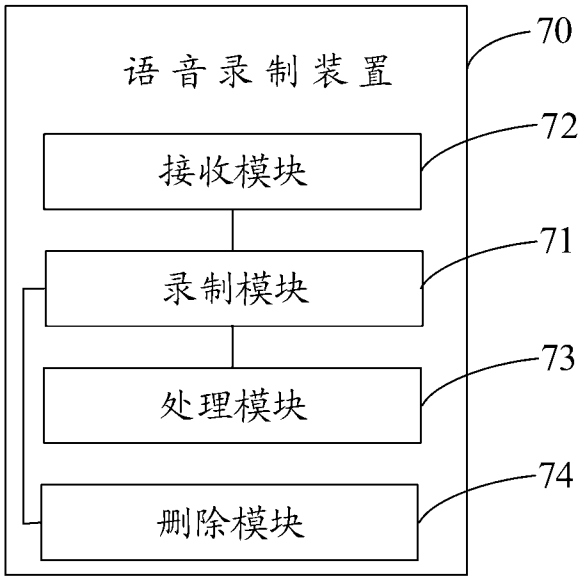

语音录制装置
70
接收模块
72
录制模块
71
处理模块
73
删除模块
74

图 11

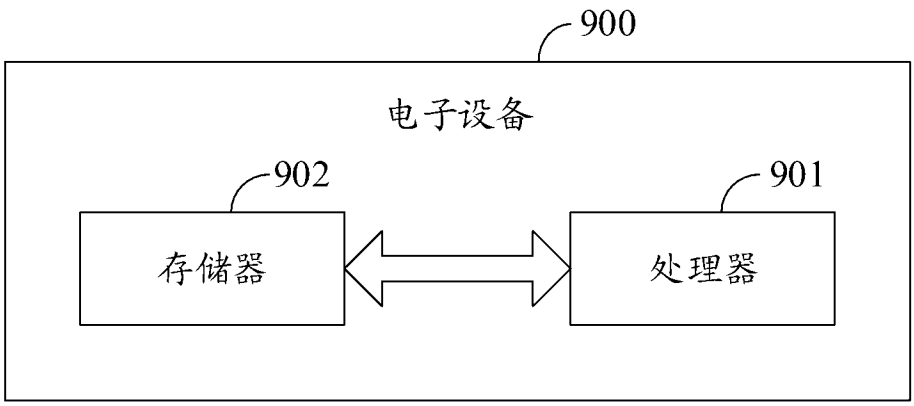

900
电子设备
902
存储器
901
处理器

图 12

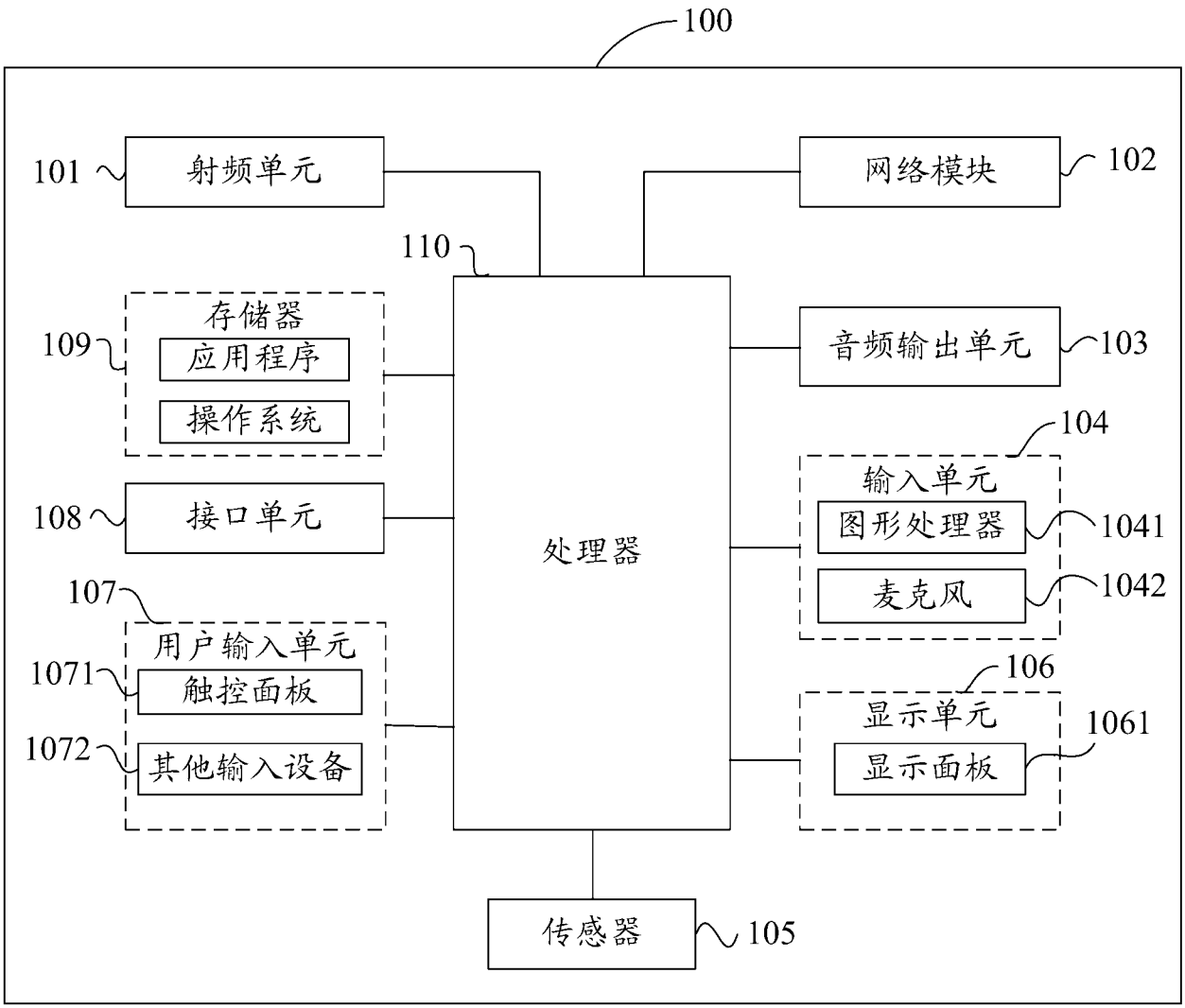

100
101
射频单元
网络模块
102
110
存储器
109
应用程序
操作系统
音频输出单元
103
104
输入单元
108
接口单元
处理器
图形处理器
1041
107
麦克风
1042
用户输入单元
106
1071
触控面板
显示单元
1061
1072
其他输入设备
显示面板
传感器
105

图 13

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2024/088407

A. CLASSIFICATION OF SUBJECT MATTER

G10L13/033(2013.01)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:G10L13/-

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)

CNTXT, ENTXTC, WPABSC, ENTXT, DWPI, WPABS, CJFD: 语音, 声音, 录制, 录音, 按键, 按钮, 按压, 触发, 缺少, 缺失, 延迟, 及时, 时间差, 启动, 开启, 起动, 提前, 预先, 预设, 目标, 行为, 动作, 视线, 注视, 嘴巴, 嘴唇, voice, sound, record+, key, button, press+, trigger, delay, start+, target

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 116543745 A (VIVO COMMUNICATION TECHNOLOGY CO., LTD.) 04 August 2023 (2023-08-04) see description, paragraphs 0028-0168, and figures 1-13	1-15
Y	CN 106531167 A (BEIJING UNISOUND INFORMATION TECHNOLOGY CO., LTD.) 22 March 2017 (2017-03-22) description, paragraphs 0050-0111, and figures 1-6	1, 3-6, 8-15
Y	CN 105760084 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) 13 July 2016 (2016-07-13) description, paragraphs 0021-0097, and figures 1-7	1, 3-6, 8-15
A	CN 106383644 A (GUANGDONG GENIUS TECHNOLOGY CO., LTD.) 08 February 2017 (2017-02-08) entire document	1-15
A	CN 107886975 A (GUANGDONG OPPO MOBILE TELECOMMUNICATIONS CO., LTD.) 06 April 2018 (2018-04-06) entire document	1-15

☑ 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	Date of mailing of the international search report
17 July 2024	**23 July 2024**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 108694947 A (GUANGDONG OPPO MOBILE COMMUNICATIONS CO., LTD.) 23 October 2018 (2018-10-23) entire document	1-15
A	CN 111492426 A (TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)) 04 August 2020 (2020-08-04) entire document	1-15
A	KR 20180084469 A (NAVER CORP.) 25 July 2018 (2018-07-25) entire document	1-15

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/088407

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 116543745 A	04 August 2023	None	
CN 106531167 A	22 March 2017	CN 106531167 B	10 December 2019
CN 105760084 A	13 July 2016	CN 105760084 B	25 January 2019
CN 106383644 A	08 February 2017	None	
CN 107886975 A	06 April 2018	CN 107886975 B	10 March 2020
CN 108694947 A	23 October 2018	WO 2020001165 A1	02 January 2020
		CN 108694947 B	19 June 2020
CN 111492426 A	04 August 2020	WO 2019123425 A1	27 June 2019
		EP 3729421 A1	28 October 2020
		US 2020380977 A1	03 December 2020
		US 11423896 B2	23 August 2022
		CN 111492426 B	02 February 2024
KR 20180084469 A	25 July 2018	None	

国际检索报告

国际申请号
PCT/CN2024/088407

A. 主题的分类

G10L13/033(2013.01)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:G10L13/-

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））

CNTXT，ENTXTC，WPABSC，ENTXT，DWPI，WPABS，CJFD：语音，声音，录制，录音，按键，按钮，按压，触发，缺少，缺失，延迟，及时，时间差，启动，开启，起动，提前，预先，预设，目标，行为，动作，视线，注视，嘴巴，嘴唇， voice, sound, record+, key, button, press+, trigger, delay, start+, target

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 116543745 A (维沃移动通信有限公司) 2023年8月4日 (2023 - 08 - 04) 参见说明书第0028-0168段，附图1-13	1-15
Y	CN 106531167 A (北京云知声信息技术有限公司) 2017年3月22日 (2017 - 03 - 22) 说明书第0050-0111段，附图1-6	1、3-6、8-15
Y	CN 105760084 A (百度在线网络技术（北京）有限公司) 2016年7月13日 (2016 - 07 - 13) 说明书第0021-0097段，附图1-7	1、3-6、8-15
A	CN 106383644 A (广东小天才科技有限公司) 2017年2月8日 (2017 - 02 - 08) 全文	1-15
A	CN 107886975 A (广东欧珀移动通信有限公司) 2018年4月6日 (2018 - 04 - 06) 全文	1-15
A	CN 108694947 A (OPPO广东移动通信有限公司) 2018年10月23日 (2018 - 10 - 23) 全文	1-15
A	CN 111492426 A (瑞典爱立信有限公司) 2020年8月4日 (2020 - 08 - 04) 全文	1-15

☑ 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	国际检索报告邮寄日期
2024年7月17日	2024年7月23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	受权官员 杨晓林 电话号码 (+86) 010-53962460

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
A	KR 20180084469 A (NAVER CORP.) 2018年7月25日 (2018 - 07 - 25) 全文	1-15

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/088407

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	116543745	A	2023年8月4日	无			
CN	106531167	A	2017年3月22日	CN	106531167	B	2019年12月10日
CN	105760084	A	2016年7月13日	CN	105760084	B	2019年1月25日
CN	106383644	A	2017年2月8日	无			
CN	107886975	A	2018年4月6日	CN	107886975	B	2020年3月10日
CN	108694947	A	2018年10月23日	WO	2020001165	A1	2020年1月2日
				CN	108694947	B	2020年6月19日
CN	111492426	A	2020年8月4日	WO	2019123425	A1	2019年6月27日
				EP	3729421	A1	2020年10月28日
				US	2020380977	A1	2020年12月3日
				US	11423896	B2	2022年8月23日
				CN	111492426	B	2024年2月2日
KR	20180084469	A	2018年7月25日	无			