



(12) 发明专利申请

(10) 申请公布号 CN 112669884 A

(43) 申请公布日 2021.04.16

(21) 申请号 202011632424.1

(22) 申请日 2020.12.31

(71) 申请人 广州酷狗计算机科技有限公司
地址 510660 广东省广州市天河区黄埔大道中315号自编1-17

(72) 发明人 邓俊锋

(74) 专利代理机构 北京三高永信知识产权代理有限公司 11138

代理人 郭翱杰

(51) Int.Cl.
G11B 20/10 (2006.01)

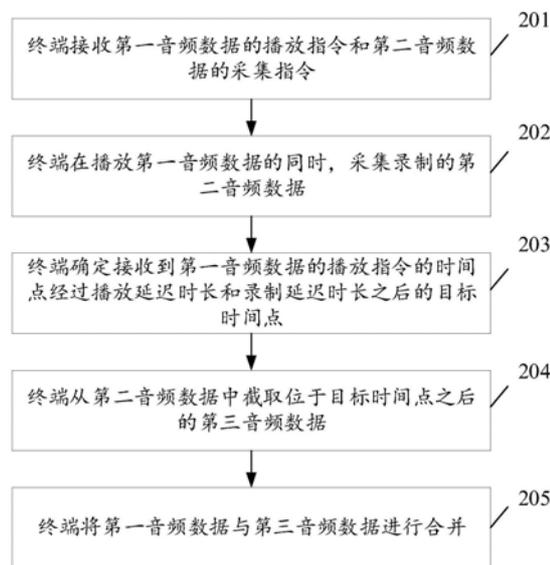
权利要求书2页 说明书17页 附图11页

(54) 发明名称

音频数据处理方法、装置、设备及存储介质

(57) 摘要

本申请实施例公开了一种音频数据处理方法、装置、设备及存储介质,属于音频处理技术领域。该方法包括:接收第一音频数据的播放指令和第二音频数据的采集指令;在播放第一音频数据的同时,采集录制的第二音频数据;确定接收到第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点,从第二音频数据中截取位于目标时间点之后的第三音频数据;将第一音频数据与第三音频数据进行合并,本申请实施例将第二音频数据中播放延迟时长和录制延迟时长对应的空白数据排除,以使播放第一音频数据中包含的声音与在播放第一音频数据的同时终端所处环境中产生的声音实现了对齐,进而提高了录制音频数据的效果。



1. 一种音频数据处理方法,其特征在于,所述方法包括:
接收第一音频数据的播放指令和第二音频数据的采集指令;
在播放所述第一音频数据的同时,采集录制的所述第二音频数据;
确定接收到所述第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点,所述播放延迟时长包括播放所述第一音频数据的延迟时长,所述录制延迟时长包括对录制的任一音频数据进行传输所产生的延迟时长;
从所述第二音频数据中截取位于所述目标时间点之后的第三音频数据;
将所述第一音频数据与所述第三音频数据进行合并。
2. 根据权利要求1所述的方法,其特征在于,所述在播放第一音频数据的同时,采集录制的第二音频数据,包括:
基于第一线程播放所述第一音频数据的同时,基于第二线程进行音频录制,并采集录制的所述第二音频数据。
3. 根据权利要求2所述的方法,其特征在于,所述基于第一线程播放所述第一音频数据的同时,基于第二线程进行音频录制,并采集录制的所述第二音频数据之前,所述方法还包括:
启动所述第一线程和所述第二线程;
在经过所述第一线程和所述第二线程的启动时间点之后的第一预设时长后,执行所述基于第一线程播放所述第一音频数据的同时,基于第二线程进行音频录制,并采集录制的所述第二音频数据的步骤。
4. 根据权利要求3所述的方法,其特征在于,所述在经过所述第一线程和所述第二线程的启动时间点之后的第一预设时长后,执行所述基于第一线程播放所述第一音频数据的同时,基于第二线程进行音频录制,并采集录制的所述第二音频数据的步骤之前,所述方法还包括:
在所述第一线程和所述第二线程的启动时间点之后的所述第一预设时长内,基于所述第一线程播放第四音频数据,所述第四音频数据为静音数据,基于所述第二线程进行音频录制,采集录制的第五音频数据,并将已采集的所述第五音频数据删除。
5. 根据权利要求1所述的方法,其特征在于,所述确定接收到所述第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点之前,所述方法还包括:
根据第一缓存延迟时长和第一传输延迟时长,确定所述播放延迟时长;
其中,所述第一缓存延迟时长为将所述第一音频数据存储至缓存区所产生的延迟时长,所述第一传输延迟时长为将所述缓存区中的所述第一音频数据传输给播放组件所产生的延迟时长,所述播放组件用于播放所述第一音频数据。
6. 根据权利要求5所述的方法,其特征在于,所述缓存区位于音频接口,所述根据第一缓存延迟时长和第一传输延迟时长,确定所述播放延迟时长之前,所述方法还包括:
调用所述音频接口,获取将所述第一音频数据存储至所述缓存区内所产生的第一缓存延迟时长。
7. 根据权利要求5所述的方法,其特征在于,音频接口用于将所述缓存区中的所述第一音频数据传输给所述播放组件,所述根据第一缓存延迟时长和第一传输延迟时长,确定所

述播放延迟时长之前,所述方法还包括:

调用所述音频接口,获取将所述第一音频数据传输至所述播放组件所产生的第一传输延迟时长。

8. 根据权利要求1所述的方法,其特征在于,所述确定接收到所述第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点之前,所述方法还包括:

根据第二缓存延迟时长和第二传输延迟时长,确定所述录制延迟时长;

其中,所述第二缓存延迟时长为将任一音频数据存储至缓存区所产生的延迟时长,所述第二传输延迟时长为将所述缓存区中的所述任一音频数据传输给目标应用所产生的延迟时长,所述目标应用用于播放所述第一音频数据以及录制所述第二音频数据。

9. 根据权利要求8所述的方法,其特征在于,所述缓存区位于音频接口,所述根据第二缓存延迟时长和第二传输延迟时长,确定所述录制延迟时长之前,所述方法还包括:

调用所述音频接口,获取将所述任一音频数据存储至所述缓存区内所产生的第二缓存延迟时长。

10. 根据权利要求8所述的方法,其特征在于,音频接口用于将所述缓存区中的所述任一音频数据传输给所述目标应用,所述根据第二缓存延迟时长和第二传输延迟时长,确定所述录制延迟时长之前,所述方法还包括:

调用所述音频接口,获取将所述任一音频数据传输至所述目标应用所产生的第二传输延迟时长。

11. 根据权利要求8所述的方法,其特征在于,所述根据第二缓存延迟时长和第二传输延迟时长,确定所述录制延迟时长,包括:

将所述第二缓存延迟时长、所述第二传输延迟时长与第二预设时长的总和,确定为所述录制延迟时长,所述第二预设时长用于指示用户的反应时长。

12. 一种音频数据处理装置,其特征在于,所述装置包括:

接收模块,用于接收第一音频数据的播放指令和第二音频数据的采集指令;

操作模块,用于在播放所述第一音频数据的同时,采集录制的所述第二音频数据;

时间点确定模块,用于确定接收到所述第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点,所述播放延迟时长包括播放所述第一音频数据的延迟时长,所述录制延迟时长包括对录制的任一音频数据进行传输所产生的延迟时长;

截取模块,用于从所述第二音频数据中截取位于所述目标时间点之后的第三音频数据;

合并模块,用于将所述第一音频数据与所述第三音频数据进行合并。

13. 一种计算机设备,其特征在于,所述计算机设备包括处理器和存储器,所述存储器中存储有至少一条程序代码,所述至少一条程序代码由所述处理器加载并执行,以实现如权利要求1至11任一权利要求所述的音频数据处理方法。

14. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质中存储有至少一条程序代码,所述至少一条程序代码由处理器加载并执行,以实现如权利要求1至11任一权利要求所述的音频数据处理方法。

音频数据处理方法、装置、设备及存储介质

技术领域

[0001] 本申请实施例涉及音频处理技术领域,特别涉及一种音频数据处理方法、装置、设备及存储介质。

背景技术

[0002] 相关技术能够在播放第一音频数据的过程中,由用户跟随着第一音频数据发出声音,从而录制用户发出的第二音频数据,将第一音频数据与第二音频数据进行合并。

[0003] 但是,终端播放第一音频数据和录制第二音频数据的过程均存在延迟时长,导致第一音频数据和第二音频数据之间的延迟时长大,两者不匹配,进而影响音频数据的录制效果。

发明内容

[0004] 本申请实施例提供了一种音频数据处理方法、装置、设备及存储介质,将第二音频数据中播放延迟时长和录制延迟时长对应的空白数据排除,以使播放第一音频数据中包含的声音与在播放第一音频数据的同时终端所处环境中产生的声音实现了对齐,进而提高了录制音频数据的效果。所述技术方案如下:

[0005] 一方面,提供了一种音频数据处理方法,所述方法包括:

[0006] 接收第一音频数据的播放指令和第二音频数据的采集指令;

[0007] 在播放所述第一音频数据的同时,采集录制的所述第二音频数据;

[0008] 确定接收到所述第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点,所述播放延迟时长包括播放所述第一音频数据的延迟时长,所述录制延迟时长包括对录制的任一音频数据进行传输所产生的延迟时长;

[0009] 从所述第二音频数据中截取位于所述目标时间点之后的第三音频数据;

[0010] 将所述第一音频数据与所述第三音频数据进行合并。

[0011] 可选地,所述在播放第一音频数据的同时,采集录制的第二音频数据,包括:

[0012] 基于第一线程播放所述第一音频数据的同时,基于第二线程进行音频录制,并采集录制的所述第二音频数据。

[0013] 可选地,所述基于第一线程播放所述第一音频数据的同时,基于第二线程进行音频录制,并采集录制的所述第二音频数据之前,所述方法还包括:

[0014] 启动所述第一线程和所述第二线程;

[0015] 在经过所述第一线程和所述第二线程的启动时间点之后的第一预设时长后,执行所述基于第一线程播放所述第一音频数据的同时,基于第二线程进行音频录制,并采集录制的所述第二音频数据的步骤。

[0016] 可选地,所述在经过所述第一线程和所述第二线程的启动时间点之后的第一预设时长后,执行所述基于第一线程播放所述第一音频数据的同时,基于第二线程进行音频录制,并采集录制的所述第二音频数据的步骤之前,所述方法还包括:

[0017] 在所述第一线程和所述第二线程的启动时间点之后的所述第一预设时长内,基于所述第一线程播放第四音频数据,所述第四音频数据为静音数据,基于所述第二线程进行音频录制,采集录制的第五音频数据,并将已采集的所述第五音频数据删除。

[0018] 可选地,所述确定接收到所述第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点之前,所述方法还包括:

[0019] 根据第一缓存延迟时长和第一传输延迟时长,确定所述播放延迟时长;

[0020] 其中,所述第一缓存延迟时长为将所述第一音频数据存储至缓存区所产生的延迟时长,所述第一传输延迟时长为将所述缓存区中的所述第一音频数据传输给播放组件所产生的延迟时长,所述播放组件用于播放所述第一音频数据。

[0021] 可选地,所述缓存区位于音频接口,所述根据第一缓存延迟时长和第一传输延迟时长,确定所述播放延迟时长之前,所述方法还包括:

[0022] 调用所述音频接口,获取将所述第一音频数据存储至所述缓存区内所产生的第一缓存延迟时长。

[0023] 可选地,音频接口用于将所述缓存区中的所述第一音频数据传输给所述播放组件,所述根据第一缓存延迟时长和第一传输延迟时长,确定所述播放延迟时长之前,所述方法还包括:

[0024] 调用所述音频接口,获取将所述第一音频数据传输至所述播放组件所产生的第一传输延迟时长。

[0025] 可选地,所述确定接收到所述第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点之前,所述方法还包括:

[0026] 根据第二缓存延迟时长和第二传输延迟时长,确定所述录制延迟时长;

[0027] 其中,所述第二缓存延迟时长为将所述任一音频数据存储至缓存区所产生的延迟时长,所述第二传输延迟时长为将所述缓存区中的所述任一音频数据传输给目标应用所产生的延迟时长,所述目标应用用于播放所述第一音频数据以及录制所述第二音频数据。

[0028] 可选地,所述缓存区位于音频接口,所述根据第二缓存延迟时长和第二传输延迟时长,确定所述录制延迟时长之前,所述方法还包括:

[0029] 调用所述音频接口,获取将所述任一音频数据存储至所述缓存区内所产生的第二缓存延迟时长。

[0030] 可选地,音频接口用于将所述缓存区中的所述任一音频数据传输给所述目标应用,所述根据第二缓存延迟时长和第二传输延迟时长,确定所述录制延迟时长之前,所述方法还包括:

[0031] 调用所述音频接口,获取将所述任一音频数据传输至所述目标应用所产生的第二传输延迟时长。

[0032] 可选地,所述根据第二缓存延迟时长和第二传输延迟时长,确定所述录制延迟时长,包括:

[0033] 将所述第二缓存延迟时长、所述第二传输延迟时长与第二预设时长的总和,确定为所述录制延迟时长,所述第二预设时长用于指示用户的反应时长。

[0034] 另一方面,提供了一种音频数据处理装置,所述装置包括:

[0035] 接收模块,用于接收第一音频数据的播放指令和第二音频数据的采集指令;

- [0036] 操作模块,用于在播放所述第一音频数据的同时,采集录制的所述第二音频数据;
- [0037] 时间点确定模块,用于确定接收到所述第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点,所述播放延迟时长包括播放所述第一音频数据的延迟时长,所述录制延迟时长包括对录制的任一音频数据进行传输所产生的延迟时长;
- [0038] 截取模块,用于从所述第二音频数据中截取位于所述目标时间点之后的第三音频数据;
- [0039] 合并模块,用于将所述第一音频数据与所述第三音频数据进行合并。
- [0040] 可选地,所述操作模块,用于基于第一线程播放所述第一音频数据的同时,基于第二线程进行音频录制,并采集录制的所述第二音频数据。
- [0041] 可选地,所述装置还包括:
- [0042] 启动模块,用于启动所述第一线程和所述第二线程;
- [0043] 所述操作模块,用于在经过所述第一线程和所述第二线程的启动时间点之后的第一预设时长后,执行所述基于第一线程播放所述第一音频数据的同时,基于第二线程进行音频录制,并采集录制的所述第二音频数据的步骤。
- [0044] 可选地,所述操作模块,用于在所述第一线程和所述第二线程的启动时间点之后的所述第一预设时长内,基于所述第一线程播放第四音频数据,所述第四音频数据为静音数据,基于所述第二线程进行音频录制,采集录制的第五音频数据,并将已采集的所述第五音频数据删除。
- [0045] 可选地,所述装置还包括:
- [0046] 时长确定模块,用于根据第一缓存延迟时长和第一传输延迟时长,确定所述播放延迟时长;
- [0047] 其中,所述第一缓存延迟时长为将所述第一音频数据存储至缓存区所产生的延迟时长,所述第一传输延迟时长为将所述缓存区中的所述第一音频数据传输给播放组件所产生的延迟时长,所述播放组件用于播放所述第一音频数据。
- [0048] 可选地,所述缓存区位于音频接口,所述装置还包括:
- [0049] 时长获取模块,用于调用所述音频接口,获取将所述第一音频数据存储至所述缓存区内所产生的第一缓存延迟时长。
- [0050] 可选地,音频接口用于将所述缓存区中的所述第一音频数据传输给所述播放组件,所述装置还包括:
- [0051] 时长获取模块,用于调用所述音频接口,获取将所述第一音频数据传输至所述播放组件所产生的第一传输延迟时长。
- [0052] 可选地,所述装置还包括:
- [0053] 时长确定模块,用于根据第二缓存延迟时长和第二传输延迟时长,确定所述录制延迟时长;
- [0054] 其中,所述第二缓存延迟时长为将任一音频数据存储至缓存区所产生的延迟时长,所述第二传输延迟时长为将所述缓存区中的所述任一音频数据传输给目标应用所产生的延迟时长,所述目标应用用于播放所述第一音频数据以及录制所述第二音频数据。
- [0055] 可选地,所述缓存区位于音频接口,所述装置还包括:

[0056] 时长获取模块,用于调用所述音频接口,获取将所述任一音频数据存储至所述缓存区内所产生的第二缓存延迟时长。

[0057] 可选地,所述音频接口用于将所述缓存区中的所述第二音频数据传输给所述目标应用,所述装置还包括:

[0058] 时长获取模块,用于调用所述音频接口,获取将所述任一音频数据传输至所述目标应用所产生的第二传输延迟时长。

[0059] 可选地,所述时长确定模块,用于将所述第二缓存延迟时长、所述第二传输延迟时长与第二预设时长的总和,确定为所述录制延迟时长,所述第二预设时长用于指示用户的反应时长

[0060] 另一方面,提供了一种计算机设备,所述计算机设备包括处理器和存储器,所述存储器中存储有至少一条程序代码,所述至少一条程序代码由所述处理器加载并执行,以实现如上述方面所述的音频数据处理方法。

[0061] 另一方面,提供了一种计算机可读存储介质,所述计算机可读存储介质中存储有至少一条程序代码,所述至少一条程序代码由处理器加载并执行,以实现如上述方面所述的音频数据处理方法。

[0062] 再一方面,提供了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机程序代码,该计算机程序代码存储在计算机可读存储介质中,计算机设备的处理器从计算机可读存储介质读取该计算机程序代码,处理器执行该计算机程序代码,使得该计算机设备实现如上述方面所述的音频数据处理方法。

[0063] 本申请实施例提供的音频数据处理方法、装置、设备及介质,在播放第一音频数据的同时,采集录制的第二音频数据,由于在播放第一音频数据的过程中会产生播放延迟时长,在对录制的第二音频数据进行传输的过程中会产生录制延迟时长,从采集录制的第二音频数据中截取第三音频数据,将第二音频数据中播放延迟时长和录制延迟时长对应的空白数据排除,以使播放第一音频数据中包含的声音与在播放第一音频数据的同时终端所处环境中产生的声音实现了对齐,进而提高了录制音频数据的效果。

附图说明

[0064] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请实施例的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0065] 图1是本申请实施例提供的一种实施环境的结构示意图。

[0066] 图2是本申请实施例提供的一种音频数据处理方法的流程图。

[0067] 图3是本申请实施例提供的一种播放延迟时长和录制延迟时长的示意图。

[0068] 图4是本申请实施例提供的一种第一音频数据与第三音频数据对齐的示意图。

[0069] 图5是本申请实施例提供的一种音频数据处理方法的流程图。

[0070] 图6是本申请实施例提供的一种第一音频数据的传输过程的示意图。

[0071] 图7是本申请实施例提供的一种第二音频数据的传输过程的示意图。

[0072] 图8是本申请实施例提供的一种音频数据处理方法的流程图。

- [0073] 图9是本申请实施例提供的一种音频数据处理方法的流程图。
- [0074] 图10是本申请实施例提供的一种音频数据处理装置的结构示意图。
- [0075] 图11是本申请实施例提供的另一种音频数据处理装置的结构示意图。
- [0076] 图12是本申请实施例提供的一种终端的结构示意图。
- [0077] 图13是本申请实施例提供的一种服务器的结构示意图。

具体实施方式

[0078] 为使本申请实施例的目的、技术方案和优点更加清楚,下面将结合附图对本申请实施方式作进一步地详细描述。

[0079] 可以理解,本申请所使用的术语“第一”、“第二”、“第三”、“第四”、“第五”、“第六”等可在本文中用于描述各种概念,但除非特别说明,这些概念不受这些术语限制。这些术语仅用于将一个概念与另一个概念区分。举例来说,在不脱离本申请的范围的情况下,可以将第一排列顺序称为第二排列顺序,将第二排列顺序称为第一排列顺序。

[0080] 本申请所使用的术语“每个”、“多个”、“至少一个”、“任一”等,至少一个包括一个、两个或两个以上,多个包括两个或两个以上,而每个是指对应的多个中的每一个,任一是指多个中的任意一个。举例来说,多个元素包括3个元素,而每个是指这3个元素中的每一个元素,任一是指这3个元素中的任意一个,可以是第一个,可以是第二个,也可以是第三个。

[0081] 图1是本申请实施例提供的一种实施环境的结构示意图。参见图1,该实施环境包括终端101和服务器102,终端101和服务器102之间通过无线或者有线网络连接,终端101和服务器102之间通过建立的连接进行数据传输。

[0082] 其中,终端101具有音频播放功能和音频录制功能,能够播放音频数据,也能够进行音频录制,采集录制的音频数据。终端101的音频播放功能和音频录制功能相互独立,或者也可以相互配合。例如,终端101播放音频数据时还可以录制音频数据。

[0083] 可选地,终端101播放的音频数据为终端101中存储的音频数据,或者为从服务器102中获取的音频数据。终端101录制的音频数据能够发送给服务器102,由服务器102存储。

[0084] 服务器102用于为终端101服务,该服务器102中存储有音频数据,若接收到终端101发送的播放请求,根据该播放请求将对应的音频数据发送给终端101,由终端101播放。该服务器102还能够接收终端101发送的已录制得到的音频数据,存储该音频数据。

[0085] 可选地,本申请实施例中的终端101从服务器102中获取第一音频数据,播放该第一音频数据的同时,还能够录制第二音频数据,进而采用本申请实施例提供的音频数据处理方法进行音频合并,将合并生成的音频数据存储存储在服务器102中。

[0086] 其中,本申请实施例中的终端为手机、平板电脑、计算机等多种类型的终端,服务器为一台服务器,或者由若干服务器组成的服务器集群,或者是一个云计算服务中心。

[0087] 本申请实施例提供的方法,应用于音频数据录制场景中。

[0088] 例如,该音频数据录制场景为歌曲录制场景,终端播放歌曲伴奏,用户随着播放的歌曲伴奏进行歌唱,终端采用本申请实施例提供的方法,将录制得到的歌唱音频中由于播放延迟时长和录制延迟时长而产生的空白音频删除,再将删除后的歌唱音频与歌曲伴奏合并,保证歌曲伴奏与录制的歌唱音频的同步,提高了歌曲音频的录制效果。

[0089] 或者,该音频数据录制场景为有声小说录制场景,终端播放背景音频,用户随着背

景音频的节奏进行朗诵,终端采用本申请实施例提供的方法,将录制得到的用户的朗诵音频中由于播放延迟时长和录制延迟时长而产生的空白音频删除,进而将删除后的朗诵音频与背景音频合并,得到合并后的音频数据。

[0090] 图2是本申请实施例提供的一种音频数据处理方法的流程图。参见图2,该方法包括:

[0091] 201、终端接收第一音频数据的播放指令和第二音频数据的采集指令。

[0092] 其中,播放指令用于指示终端播放该第一音频数据,采集指令用于指示终端采集录制的第二音频数据。

[0093] 202、终端在播放第一音频数据的同时,采集录制的第二音频数据。

[0094] 终端根据播放指令和采集指令,确定在播放第一音频数据的同时,采集录制的第二音频数据。

[0095] 其中,该第一音频数据为歌曲音频、伴奏音频、有声小说音频等等,本申请实施例并不做限定。该第二音频数据为终端录制得到的音频数据。例如,该第二音频数据包含用户发出的声音、环境声音,或者为其他声音。

[0096] 在本申请实施例中,终端播放第一音频数据,则用户能够听到该第一音频数据,该用户还能够基于听到的第一音频数据发出声音,则终端能够在播放第一音频数据的同时,采集录制得到的第二音频数据,则第二音频数据中会包含着用户发出的声音。

[0097] 203、终端确定接收到第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点。

[0098] 其中,播放延迟时长包括播放第一音频数据的延迟时长。并且,终端在接收到第一音频数据的播放指令后,开始加载第一音频数据并对该第一音频数据进行处理,处理完毕后才能播放该第一音频数据的音频,在加载并对第一音频数据进行处理的过程中所耗费的时长为播放延迟时长。

[0099] 录制延迟时长包括对任一音频数据进行处理所产生的延迟时长。终端在录制音频数据的过程中,由于需要对音频数据进行传输,导致对音频数据的传输过程需要耗费时长,在此过程中所耗费的时长为录制延迟时长。

[0100] 终端在播放第一音频数据的过程中,由于需要先对第一音频数据进行处理才能进行播放,处理的过程需要耗费时长,即播放延迟时长。终端在采集录制得到的第二音频数据的过程中,由于在播放第一音频数据的过程中,已经采集得到第二音频数据,但是采集的第二音频数据中包括的用户的音频数据是需要进行传输的,进行处理耗费的时长即录制延迟时长。

[0101] 在本申请实施例中,由于终端同时播放音频数据和录制音频数据,不仅需要在播放过程中对第一音频数据进行处理,而且还需要对录制的第二音频数据进行传输,因此得到的第二音频数据中,在第二音频数据的起始时间点之后的播放延迟时长和录制延迟时长内存在空白数据。

[0102] 例如,如图3所示,终端同时播放第一音频数据和录制第二音频数据,但是由于开始播放第一音频数据,经过播放延迟时长后第一音频数据才由终端播放出来,并且此时终端采集录制得到的第二音频数据,再对第二音频数据进行传输耗费录制延迟时长,以第二音频数据的起始时间点开始,经过该播放延迟时长和录制延迟时长后,第二音频数据包含

的声音才与第一音频数据包含的声音对齐。

[0103] 在本申请实施例中,为了保证第二音频数据中不包括由于播放延迟时长和录制延迟时长导致的空白数据,终端需要从第二音频数据中将空白数据删除,截取不包括空白数据的音频数据,截取得到的音频数据为附和第一音频数据而产生的音频数据。

[0104] 终端采集录制得到的第二音频数据的时间点与播放第一音频数据的时间点相同,而由于从接收播放第一音频数据的播放指令的时间点开始,位于播放延迟时长和录制延迟时长内的第二音频数据为空白数据,因此从接收到第一音频数据的播放指令的时间点开始,确定经过播放延迟时长和录制延迟时长后的目标时间点,该目标时间点之前的第二音频数据均为空白数据,而该目标时间点之后的第二音频数据才包括与第一音频数据匹配的第三音频数据。

[0105] 204、终端从第二音频数据中截取位于目标时间点之后的第三音频数据。

[0106] 在本申请实施例中,终端确定目标时间点后,则能够从第二音频数据中截取位于该目标时间点之后的第三音频数据,该第三音频数据中包括的声音与第一音频数据中包括的声音同步。

[0107] 205、终端将第一音频数据与第三音频数据进行合并。

[0108] 终端截取得到第三音频数据后,则能够将第一音频数据和第三音频数据进行合并,终端播放的第一音频数据中包含的声音,与在播放第一音频数据的同时终端所处环境中产生的声音实现了对齐。例如在播放第一音频数据中包含的歌曲伴奏的同时,用户跟随该歌曲伴奏进行歌唱,则能实现歌曲伴奏与用户歌唱的歌曲之间对齐。

[0109] 例如,在图3的基础上,图4是本申请实施例提供的一种第一音频数据与第三音频数据对齐的示意图,参见图4,通过对第二音频数据的截取,将第二音频数据中对应播放延迟时长和录制延迟时长的音频数据排除,得到的第三音频数据的起始时间点与第一音频数据的起始时间点对齐,因此第三音频数据中包括的音频与第一音频数据中包括的音频也匹配。

[0110] 例如,若第一音频数据包含的声音的第一个字为“我”,在第二音频数据中播放延迟时长和录制延迟时长对应的音频为空白音频,而在第三音频数据的起始时间点处包括的声音中的第一个字“我”,因此将第三音频数据和第一音频数据进行合并,保证了第一音频数据中的声音与第三音频数据中的声音对齐。

[0111] 本申请实施例提供的方法,在播放第一音频数据的同时,采集录制的第二音频数据,由于在播放第一音频数据的过程中会产生播放延迟时长,在对录制的第二音频数据进行传输的过程中会产生录制延迟时长,从采集录制的第二音频数据中截取第三音频数据,将第二音频数据中播放延迟时长和录制延迟时长对应的空白数据排除,以使播放第一音频数据中包含的声音与在播放第一音频数据的同时终端所处环境中产生的声音实现了对齐,进而提高了录制音频数据的效果。

[0112] 图5是本申请实施例提供的一种音频数据处理方法的流程图,参见图3,该方法包括:

[0113] 501、终端接收第一音频数据的播放指令和第二音频数据的采集指令。

[0114] 502、终端在播放第一音频数据的同时,采集录制的第二音频数据。

[0115] 在本申请实施例中,若终端播放第一音频数据的同时,还会录制第二音频数据,而

终端播放第一音频数据完成时,也会结束第二音频数据的录制。

[0116] 可选地,该第一音频数据为终端中已存储的音频数据,或者为终端从服务器中获取的音频数据,或者为采用其他方式获取的音频数据。

[0117] 其中,终端中安装有目标应用,该目标应用具有音频数据录制功能,终端采用该音频数据录制功能,基于该目标应用播放第一音频数据,同时基于该目标应用录制第二音频数据。

[0118] 其中,该目标应用为终端中安装的应用,例如该目标应用为歌曲应用、小说应用、聊天应用、直播应用或者其他类型的应用。

[0119] 例如,该目标应用为歌曲应用,则终端播放的第一音频数据为歌曲音频,录制的第二音频数据包括用户的歌唱音频。或者,该目标应用为小说应用,则终端播放的第一音频数据为背景音频,录制的第二音频数据包括用户的朗诵音频。该歌曲应用可以为卡拉OK应用。

[0120] 在一些实施例中,终端能够基于用户标识登录该目标应用,再从目标应用中选取第一音频数据,进而控制终端在播放第一音频数据的同时,录制第二音频数据。例如,该用户标识为用户昵称、用户的手机号、用户账号或者其他标识,本申请实施例并不做限定。

[0121] 可选地,终端基于该目标应用显示至少一个第一音频数据的标识,响应于对任一一个第一音频数据的标识的触发操作,播放该第一音频数据,同时录制第二音频数据。

[0122] 其中,该第一音频数据的标识为音频数据的名称、音频数据的歌手,或者其他类型的标识。该触发操作为单击操作、双击操作或者其他类型的操作。

[0123] 需要说明的是,本申请实施例中的终端能够通过扬声器播放第一音频数据,或者通过与终端连接的耳机播放第一音频数据,或者采用其他方式播放第一音频数据。另外,终端能够通过麦克风采集第二音频数据。

[0124] 503、终端根据第一缓存延迟时长和第一传输延迟时长,确定播放延迟时长。

[0125] 其中,第一缓存延迟时长为将第一音频数据存储至缓存区的延迟时长,第一传输延迟时长为将缓存区中的第一音频数据传输给播放组件的延迟时长,播放组件用于播放第一音频数据。

[0126] 在本申请实施例中,终端在播放第一音频数据的过程中,将播放的第一音频数据存储至缓存区中,在存储至缓存区中的过程中会产生存储的延迟时长,该延迟为第一缓存延迟时长。

[0127] 可选地,该第一缓存延迟时长与缓存区的容量成正比,也即是缓存区的容量越大,则第一缓存延迟时长越大,而缓存区的容量越小,则第一缓存延迟时长越小。

[0128] 另外,终端将第一音频数据存储至缓存区中后,还需要将缓存区中存储的第一音频数据传输给播放组件,在传输过程中仍会耗费时长,该延时即为第一传输延迟时长。该播放组件用于播放第一音频数据。例如,该播放组件为扬声器、耳机或者其他组件。

[0129] 在一些实施例中,终端将第一缓存延迟时长和第一传输延迟时长的总和确定为播放延迟时长。

[0130] 例如,终端确定的第一缓存延迟时长为20毫秒,而第一传输延迟时长为30毫秒,则确定的时长总和50毫秒为播放延迟时长。

[0131] 在本申请实施例中,仅对终端如何根据第一缓存延迟时长和第一传输延迟时长确定播放延迟时长进行了说明。下面将对终端如何获取第一缓存延迟时长和第一传输延迟时

长进行说明。

[0132] 首先,先对本申请实施例所涉及到的音频接口进行说明。例如,图6示出了终端调用音频接口传输第一音频数据的过程。该终端中包括扬声器、音频接口和目标应用,终端中的目标应用将第一音频数据发送给音频接口,该音频接口会先将接收的第一音频数据存储于缓存区中,再将缓存区中的第一音频数据发送给播放组件,以完成音频数据的播放。

[0133] 在一些实施例中,终端调用音频接口,获取将第一音频数据存储至缓存区内所产生的第一缓存延迟时长。

[0134] 其中,该缓存区位于音频接口,则该音频接口能够确定缓存区的容量,进而能够获取将第一音频数据存储至缓存区内所产生的第一缓存延迟时长。

[0135] 在另一些实施例中,调用音频接口,获取将第一音频数据传输至播放组件所产生的第一传输延迟时长。

[0136] 其中,音频接口用于将缓存区中的第一音频数据传输给播放组件。则音频接口将获取的第一音频数据存储至缓存区中后,能够将缓存区中的第一音频数据传输至播放组件,在传输过程中音频接口会记录发送的时间点以及播放组件返回的接收成功的时间点,进而能够确定第一音频数据传输至播放组件所产生的第一传输延迟时长。

[0137] 可选地,该第二传输延迟时长为音频接口的返回值。该终端的系统为Android(一种操作系统),或者其他操作系统,本申请实施例不做限定。该音频接口为AAudio(一种音频接口),或者其他类型的接口。

[0138] 504、终端根据第二缓存延迟时长和第二传输延迟时长,确定录制延迟时长。

[0139] 其中,第二缓存延迟时长为将录制的音频数据存储至缓存区的延迟时长,第二传输延迟时长为将缓存区中的音频数据传输给目标应用的延迟时长,该录制延迟时长包括对录制的任一音频数据进行传输所产生的延迟时长。

[0140] 在本申请实施例中,终端在录制音频数据的过程中,将录制的音频数据存储于缓存区中,在存储至缓存区中的过程中会产生存储的延迟时长,该延迟为第二缓存延迟时长。

[0141] 可选地,该第二缓存延迟时长与缓存区的容量成正比,也即是缓存区的容量越大,则第二缓存延迟时长越大,而缓存区的容量越小,则第二缓存延迟时长越小。

[0142] 另外,终端将音频数据存储至缓存区中后,还需要将缓存区中存储的音频数据传输给目标应用,在传输过程中仍会产生延时,该延时即为第二传输延迟时长。该目标应用用于播放第一音频数据以及录制第二音频数据。

[0143] 在一些实施例中,终端将第二缓存延迟时长和第二传输延迟时长的时长总和确定为录制延迟时长。

[0144] 例如,终端确定的第二缓存延迟时长为30毫秒,而第二传输延迟时长为50毫秒,则确定的时长总和80毫秒为录制延迟时长。

[0145] 在本申请实施例中,仅对终端如何根据第二缓存延迟时长和第二传输延迟时长确定录制延迟时长进行了说明。下面将对终端如何获取第二缓存延迟时长和第二传输延迟时长进行说明。

[0146] 首先,先对本申请实施例所涉及到的音频接口进行说明。例如,图7示出了终端调用音频接口传输任一音频数据的过程。该终端中包括麦克风、音频接口和目标应用,终端采用该麦克风采集任一音频数据,再将该任一音频数据发送给音频接口,该音频接口会先将

接收的任一音频数据存储至缓存区中,再将缓存区中的任一音频数据发送给目标应用,以完成音频数据的录制。

[0147] 在一些实施例中,终端调用音频接口,获取将任一音频数据存储至缓存区内所产生的第二缓存延迟时长。

[0148] 其中,该缓存区位于音频接口,则该音频接口能够确定缓存区的容量,进而能够获取将任一音频数据存储至缓存区内所产生的第二缓存延迟时长。

[0149] 在另一些实施例中,调用音频接口,获取将任一音频数据传输至目标应用所产生的第二传输延迟时长。

[0150] 其中,音频接口用于将缓存区中的任一音频数据传输给目标应用。则音频接口将获取的任一音频数据存储至缓存区中后,能够将缓存区中的任一音频数据传输至目标应用,在传输过程中音频接口会记录发送的时间点以及目标应用返回的接收成功的时间点,进而能够确定任一音频数据传输至目标应用所产生的第二传输延迟时长。

[0151] 可选地,该第二传输延迟时长为延时接口返回值。该音频接口为AAudio(一种音频接口),或者其他类型的接口。

[0152] 本申请实施例提供的方法,利用现有的音频接口即可直接获取第二传输延迟时长和第二缓存延迟时长,进而能够确定录制延迟时长,提高了获取录制延迟时长的效率,进而提高了截取第三音频数据的准确率,提高录制音频数据的效果。

[0153] 需要说明的第一点是,本申请实施例仅是以根据第二缓存延迟时长和第二传输延迟时长确定录制延迟时长为例进行说明。在另一实施例中,终端还能够获取用户的反应时长,额外考虑用户的反应时长以确定录制延迟时长。

[0154] 终端在录制第二音频数据的过程中,若终端开始播放第一音频数据后,用户听到该第一音频数据不会立即跟随第一音频数据发出声音,而是会反应一段时长后才会跟随第一音频数据发出声音,因此终端在确定录制延迟时长时,需要考虑用户的反应时长。

[0155] 可选地,终端将第二缓存延迟时长、第二传输延迟时长与第二预设时长的总和,确定为录制延迟时长。其中,第二预设时长用于指示用户的反应时长。

[0156] 该第二缓存延迟时长和第二传输延迟时长的获取过程与上述过程同理,而该的第二预设时长由终端设置,或者由操作人员设置,或者采用其他方式设置,本申请实施例并不对第二预设时长的设置方式进行限定。

[0157] 另外,需要说明的第二点是,本申请实施例中的步骤502-503为可选步骤,在另一实施例中还可以不执行502-503获取播放延迟时长和录制延迟时长,终端已提前获取到播放延迟时长和录制延迟时长,后续执行步骤504即可。

[0158] 505、终端确定接收到第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点。

[0159] 在本申请实施例中,终端播放第一音频数据和录制第二音频数据,由于在播放第一音频数据的过程中,会存在播放延迟时长,在录制第二音频数据的过程中,会存在录制延迟时长,在第二音频数据对应的播放延迟时长和录制延迟时长内会存在没有音频数据的情况,因此终端需要确定第二音频数据中包括用户的音频数据的时间点。

[0160] 终端通过上述步骤502-503确定了播放第一音频数据的播放延迟时长和录制延迟时长,并且还能够确定接收到第一音频数据的播放指令的时间点,因此能够确定以接收到

第一音频数据的播放指令的时间点为起点,经过播放延迟时长和录制延迟时长后的时间点,将该时间点确定为目标时间点。终端确定的该目标时间点为终端录制得到用户的声音的时间点,后续能够以该目标时间点为起点,从第二音频数据中截取音频数据。

[0161] 506、终端从第二音频数据中截取位于目标时间点之后的第三音频数据。

[0162] 终端通过步骤504确定目标时间点后,也即是确定了第二音频数据中包括与第一音频数据对应的声音的时间点,从该第二音频数据中截取位于该目标时间点之后的第三音频数据,截取的第三音频数据能够认为与第一音频数据相匹配。

[0163] 507、终端将第一音频数据与第三音频数据进行合并。

[0164] 在本申请实施例中,终端播放的第一音频数据的声音与截取的第三音频数据的声音对齐。

[0165] 本申请实施例提供的方法,在播放第一音频数据的同时,录制第二音频数据,由于在播放第一音频数据的过程中会产生播放延迟时长,在对录制的第二音频数据进行传输也会产生录制延迟时长,从采集录制的第二音频数据中截取第三音频数据,将第二音频数据中播放延迟时长和录制延迟时长对应的空白数据排除,以使终端播放第一音频数据中包含的声音与在播放第一音频数据的同时终端所处环境中产生的声音实现了对齐,进而提高了录制音频数据的效果。

[0166] 并且,利用终端设置的音频接口即可直接获取传输延迟时长和缓存延迟时长,进而能够确定录制延迟时长,提高了获取录制延迟时长的效率,进而提高了截取第三音频数据的准确率,提高录制音频数据的效果。

[0167] 上述图5实施例仅是对如何根据播放延迟时长和录制延迟时长,从第二音频数据中截取音频数据进而与第一音频数据进行合并为例进行说明。而终端是基于第一线程播放第一音频数据,基于第二线程录制第二音频数据,为了保证第一线程和第二线程同时开始工作,采用图8所示的实施例进行说明。图8是本申请实施例提供的一种音频数据处理方法的流程图,参见图8,该方法包括:

[0168] 801、终端启动第一线程和第二线程。

[0169] 在本申请实施例中,终端包括多个线程,不同的线程的功能不同。例如,通过线程能够播放音频,或者,通过线程录制音频,或者,通过线程拍摄视频等。而本申请实施例中的第一线程用于播放音频数据,第二线程用于录制音频数据。

[0170] 若终端需要根据播放的第一音频数据录制第二音频数据,则需要基于第一线程播放第一音频数据,基于第二线程录制第二音频数据,而在执行上述步骤之前,需要先启动第一线程和第二线程,再基于第一线程和第二线程执行后续过程。

[0171] 802、终端在经过第一线程和第二线程的启动时间点之后的第一预设时长后,基于第一线程播放第一音频数据的同时,基于第二线程进行音频录制,并采集录制的第二音频数据。

[0172] 在本申请实施例中,终端在控制第一线程和第二线程启动过程中,第一线程和第二线程的启动时间会出现差别,若此时直接执行播放第一音频数据,录制第二音频数据的过程,也会造成第一音频数据和第二音频数据之间存在延迟时长,因此为了保证第一线程和第二线程同时工作,需要先让第一线程和第二线程空转第一预设时长,再由第一线程和第二线程同时执行播放第一音频数据和录制第二音频数据的过程。

[0173] 终端启动第一线程和第二线程后,此时第一线程和第二线程处于工作状态,为了实现第一线程播放第一音频数据的同时,录制第二音频数据,先由第一线程和第二线程运行,但是并未执行播放第一音频数据和录制第二音频数据的过程,在经过启动时间点之后的第一预设时长后,再基于第一线程播放第一音频数据的同时,基于第二线程录制第二音频数据。

[0174] 可选地,终端确定到达第一线程和第二线程的启动时间点之后的第一预设时长的时间点后,同时向第一线程发送播放第一音频数据的播放指令,以及向第二线程发送采集录制的第二音频数据的采集指令,则第一线程和第二线程同时接收到指令,进而同时播放第一音频数据,以及采集录制的第二音频数据。

[0175] 需要说明的是,本申请实施例仅是以第一线程和第二线程启动后经过第一预设时长执行步骤802为例进行说明。在另一实施例中,终端启动第一线程和第二线程后的第一预设时长内,第一线程和第二线程也会工作。

[0176] 可选地,在第一线程和第二线程的启动时间点之后的第一预设时长内,基于第一线程播放第四音频数据,第四音频数据为静音数据,基于第二线程录制第五音频数据,并将已录制的第五音频数据删除。

[0177] 终端基于第一线程播放静音数据,此时用户无法收听到声音,但是第一线程在正常工作,第二线程也会录制第五音频数据,但是第二线程会实时将录制的第五音频数据删除,不会保留录制的第五音频数据,若终端确定开始播放第一音频数据,则立即暂停基于第一线程播放第三音频数据以及基于第二线程录制第五音频数据,同时开始基于第一线程播放第一音频数据,并基于第二线程录制第二音频数据。

[0178] 本申请实施例提供的方法,启动第一线程和第二线程后,由第一线程播放静音数据,第二线程录制第五音频数据并删除,能够保证第一线程和第二线程都处于工作状态,再同时基于第一线程播放第一音频数据,基于第二线程录制第二音频数据,排除了第一线程和第二线程在同时启动时产生的延迟时长,进而排除了第一线程播放第一音频数据和第二线程录制第二音频数据的延迟时长,确保第一音频数据和第二音频数据相匹配,提高了录制音频数据的效果。

[0179] 803、终端确定接收到第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点。

[0180] 804、终端从第二音频数据中截取位于目标时间点之后的第三音频数据。

[0181] 805、终端将第一音频数据与第三音频数据进行合并。

[0182] 其中,步骤803-805与上述步骤503-505同理,在此不再赘述。

[0183] 例如,如图9所示,第一线程为播放线程,第二线程为录音线程,播放线程播放静音数据,以实现空转状态,录音线程录制音频数据,但是同时也会将录制的音频数据删除,从而也能够实现空转状态,经过预设时长后,播放线程播放伴奏,而录音线程确定录制延迟时长,将从播放线程开始播放伴奏的时间点经过录制延迟时长之后的时间点,对录制的音频数据进行截取,终端再将播放线程播放的伴奏与截取后的音频数据进行合并,得到合并后的音频数据。

[0184] 本申请实施例提供的方法,启动第一线程和第二线程并经过预设时长后,才同时基于第一线程播放第一音频数据,基于第二线程进行音频录制,并采集录制的第二音频数

据,防止出现启动第一线程和第二线程的时间不同步,导致开始播放第一音频数据和开始录制第二音频数据的时间点不同步的情况,进而防止出现录制的第二音频数据与第一音频数据出现延迟时长的情况,进而提高了录制音频数据的效果。

[0185] 图10是本申请实施例提供的一种音频数据处理装置的结构示意图。参见图10,该装置包括:

[0186] 接收模块1001,用于接收第一音频数据的播放指令和第二音频数据的采集指令;

[0187] 操作模块1002,用于在播放第一音频数据的同时,采集录制的第二音频数据;

[0188] 时间点确定模块1003,用于确定接收到第一音频数据的播放指令的时间点经过播放延迟时长和录制延迟时长之后的目标时间点,播放延迟时长包括播放第一音频数据的延迟时长,录制延迟时长包括对录制的任一音频数据进行传输所产生的延迟时长;

[0189] 截取模块1004,用于从第二音频数据中截取位于目标时间点之后的第三音频数据;

[0190] 合并模块1005,用于将第一音频数据与第三音频数据进行合并。

[0191] 可选地,操作模块1002,用于基于第一线程播放第一音频数据的同时,基于第二线程进行音频录制,并采集录制的第二音频数据。

[0192] 可选地,装置还包括:

[0193] 启动模块1006,用于启动第一线程和第二线程;

[0194] 操作模块1002,用于在经过第一线程和第二线程的启动时间点之后的第一预设时长后,执行基于第一线程播放第一音频数据的同时,基于第二线程进行音频录制,并采集录制的第二音频数据的步骤。

[0195] 可选地,操作模块1002,用于在第一线程和第二线程的启动时间点之后的第一预设时长内,基于第一线程播放第四音频数据,第四音频数据为静音数据,基于第二线程进行音频录制,采集录制的第五音频数据,并将已采集的第五音频数据删除。

[0196] 可选地,装置还包括:

[0197] 时长确定模块1007,用于根据第一缓存延迟时长和第一传输延迟时长,确定播放延迟时长;

[0198] 其中,第一缓存延迟时长为将第一音频数据存储至缓存区所产生的延迟时长,第一传输延迟时长为将缓存区中的第一音频数据传输给播放组件所产生的延迟时长,播放组件用于播放第一音频数据。

[0199] 可选地,缓存区位于音频接口,装置还包括:

[0200] 时长获取模块1008,用于调用音频接口,获取将第一音频数据存储至缓存区内所产生的第一缓存延迟时长。

[0201] 可选地,音频接口用于将缓存区中的第一音频数据传输给播放组件,装置还包括:

[0202] 时长获取模块1008,用于调用音频接口,获取将第一音频数据传输至播放组件所产生的第一传输延迟时长。

[0203] 可选地,装置还包括:

[0204] 时长确定模块1007,用于根据第二缓存延迟时长和第二传输延迟时长,确定录制延迟时长;

[0205] 其中,第二缓存延迟时长为将任一音频数据存储至缓存区所产生的延迟时长,第

二传输延迟时长为将缓存区中的任一音频数据传输给目标应用所产生的延迟时长,目标应用用于播放第一音频数据以及录制第二音频数据。

[0206] 可选地,缓存区位于音频接口,装置还包括:

[0207] 时长获取模块1008,用于调用音频接口,获取将任一音频数据存储至缓存区内所产生的第二缓存延迟时长。

[0208] 可选地,音频接口用于将缓存区中的第二音频数据传输给目标应用,装置还包括:

[0209] 时长获取模块1009,用于调用音频接口,获取将任一音频数据传输至目标应用所产生的第二传输延迟时长。

[0210] 可选地,时长确定模块1007,用于将第二缓存延迟时长、第二传输延迟时长与第二预设时长的总和,确定为录制延迟时长,第二预设时长用于指示用户的反应时长。

[0211] 需要说明的是:上述实施例提供的音频数据处理装置在录制音频数据时,仅以上述各功能模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能模块完成,即将终端的内部结构划分成不同的功能模块,以完成以上描述的全部或者部分功能。另外,上述实施例提供的音频数据处理装置的实施例与音频数据处理方法实施例属于同一构思,其具体实现过程详见方法实施例,这里不再赘述。

[0212] 本公开实施例还提供了一种计算机设备,该计算机设备包括处理器和存储器,存储器中存储有至少一条程序代码,至少一条程序代码由处理器加载并执行,以实现如上述实施例中的音频数据处理方法。

[0213] 可选地,该计算机设备提供为终端。图12是本申请实施例提供的一种终端的结构示意图。该终端1200可以是便携式移动终端,比如:智能手机、平板电脑、MP3播放器(Moving Picture Experts Group Audio Layer III,动态影像专家压缩标准音频层面3)、MP4(Moving Picture Experts Group Audio Layer IV,动态影像专家压缩标准音频层面4)播放器、笔记本电脑或台式电脑。终端1200还可能被称为用户设备、便携式终端、膝上型终端、台式终端等其他名称。

[0214] 终端1200包括有:处理器1201和存储器1202。

[0215] 处理器1201可以包括一个或多个处理核心,比如4核心处理器、8核心处理器等。处理器1201可以采用DSP(Digital Signal Processing,数字信号处理)、FPGA(Field-Programmable Gate Array,现场可编程门阵列)、PLA(Programmable Logic Array,可编程逻辑阵列)中的至少一种硬件形式来实现。处理器1201也可以包括主处理器和协处理器,主处理器是用于对在唤醒状态下的数据进行处理的处理器,也称CPU(Central Processing Unit,中央处理器);协处理器是用于对在待机状态下的数据进行处理的低功耗处理器。在一些实施例中,处理器1201可以集成有GPU(Graphics Processing Unit,图像处理器),GPU用于负责显示屏所需要显示的内容的渲染和绘制。一些实施例中,处理器1201还可以包括AI(Artificial Intelligence,人工智能)处理器,该AI处理器用于处理有关机器学习的计算操作。

[0216] 存储器1202可以包括一个或多个计算机可读存储介质,该计算机可读存储介质可以是非暂态的。存储器1202还可包括高速随机存取存储器,以及非易失性存储器,比如一个或多个磁盘存储设备、闪存存储设备。在一些实施例中,存储器1202中的非暂态的计算机可读存储介质用于存储至少一个程序代码,该至少一个程序代码用于被处理器1201所执行以

实现本申请中方法实施例提供的音频数据处理方法。

[0217] 在一些实施例中,终端1200还可选包括有:外围设备接口1203和至少一个外围设备。处理器1201、存储器1202和外围设备接口1203之间可以通过总线或信号线相连。各个外围设备可以通过总线、信号线或电路板与外围设备接口1203相连。具体地,外围设备包括:射频电路1204、显示屏1205、摄像头组件1206、音频电路1207、定位组件1208和电源1209中的至少一种。

[0218] 外围设备接口1203可被用于将I/O (Input/Output, 输入/输出) 相关的至少一个外围设备连接到处理器1201和存储器1202。在一些实施例中,处理器1201、存储器1202和外围设备接口1203被集成在同一芯片或电路板上;在一些其他实施例中,处理器1201、存储器1202和外围设备接口1203中的任意一个或两个可以在单独的芯片或电路板上实现,本实施例对此不加以限定。

[0219] 射频电路1204用于接收和发射RF (Radio Frequency, 射频) 信号,也称电磁信号。射频电路1204通过电磁信号与通信网络以及其他通信设备进行通信。射频电路1204将电信号转换为电磁信号进行发送,或者,将接收到的电磁信号转换为电信号。可选地,射频电路1204包括:天线系统、RF收发器、一个或多个放大器、调谐器、振荡器、数字信号处理器、编解码芯片组、用户身份模块卡等等。射频电路1204可以通过至少一种无线通信协议来与其它终端进行通信。该无线通信协议包括但不限于:万维网、城域网、内联网、各代移动通信网络(2G、3G、4G及5G)、无线局域网和/或WiFi (Wireless Fidelity, 无线保真) 网络。在一些实施例中,射频电路1204还可以包括NFC (Near Field Communication, 近距离无线通信) 有关的电路,本申请对此不加以限定。

[0220] 显示屏1205用于显示UI (User Interface, 用户界面)。该UI可以包括图形、文本、图标、视频及其它们的任意组合。当显示屏1205是触摸显示屏时,显示屏1205还具有采集在显示屏1205的表面或表面上方的触摸信号的能力。该触摸信号可以作为控制信号输入至处理器1201进行处理。此时,显示屏1205还可以用于提供虚拟按钮和/或虚拟键盘,也称软按钮和/或软键盘。在一些实施例中,显示屏1205可以为一个,设置在终端1200的前面板;在另一些实施例中,显示屏1205可以为至少两个,分别设置在终端1200的不同表面或呈折叠设计;在另一些实施例中,显示屏1205可以是柔性显示屏,设置在终端1200的弯曲表面上或折叠面上。甚至,显示屏1205还可以设置成非矩形的不规则图形,也即异形屏。显示屏1205可以采用LCD (Liquid Crystal Display, 液晶显示屏)、OLED (Organic Light-Emitting Diode, 有机发光二极管) 等材质制备。

[0221] 摄像头组件1206用于采集图像或视频。可选地,摄像头组件1206包括前置摄像头和后置摄像头。前置摄像头设置在终端的前面板,后置摄像头设置在终端的背面。在一些实施例中,后置摄像头为至少两个,分别为主摄像头、景深摄像头、广角摄像头、长焦摄像头中的任意一种,以实现主摄像头和景深摄像头融合实现背景虚化功能、主摄像头和广角摄像头融合实现全景拍摄以及VR (Virtual Reality, 虚拟现实) 拍摄功能或者其它融合拍摄功能。在一些实施例中,摄像头组件1206还可以包括闪光灯。闪光灯可以是单色温闪光灯,也可以是双色温闪光灯。双色温闪光灯是指暖光闪光灯和冷光闪光灯的组合,可以用于不同色温下的光线补偿。

[0222] 音频电路1207可以包括麦克风和扬声器。麦克风用于采集用户及环境的声波,并

将声波转换为电信号输入至处理器1201进行处理,或者输入至射频电路1204以实现语音通信。出于立体声采集或降噪的目的,麦克风可以为多个,分别设置在终端1200的不同部位。麦克风还可以是阵列麦克风或全向采集型麦克风。扬声器则用于将来自处理器1201或射频电路1204的电信号转换为声波。扬声器可以是传统的薄膜扬声器,也可以是压电陶瓷扬声器。当扬声器是压电陶瓷扬声器时,不仅可以将电信号转换为人类可听见的声波,也可以将电信号转换为人类听不见的声波以进行测距等用途。在一些实施例中,音频电路1207还可以包括耳机插孔。

[0223] 定位组件1208用于定位终端1200的当前地理位置,以实现导航或LBS (Location Based Service,基于位置的服务)。定位组件1208可以是基于美国的GPS (Global Positioning System,全球定位系统)、中国的北斗系统或俄罗斯的伽利略系统的定位组件。

[0224] 电源1209用于为终端1200中的各个组件进行供电。电源1209可以是交流电、直流电、一次性电池或可充电电池。当电源1209包括可充电电池时,该可充电电池可以是有线充电电池或无线充电电池。有线充电电池是通过有线线路充电的电池,无线充电电池是通过无线线圈充电的电池。该可充电电池还可以用于支持快充技术。

[0225] 在一些实施例中,终端1200还包括有一个或多个传感器1210。该一个或多个传感器1210包括但不限于:加速度传感器1211、陀螺仪传感器1212、压力传感器1213、指纹传感器1214、光学传感器1215以及接近传感器1216。

[0226] 加速度传感器1211可以检测以终端140建立的坐标系的三个坐标轴上的加速度大小。比如,加速度传感器1211可以用于检测重力加速度在三个坐标轴上的分量。处理器1201可以根据加速度传感器1211采集的重力加速度信号,控制显示屏1205以横向视图或纵向视图进行用户界面的显示。加速度传感器1211还可以用于游戏或者用户的运动数据的采集。

[0227] 陀螺仪传感器1212可以检测终端1200的机体方向及转动角度,陀螺仪传感器1212可以与加速度传感器1211协同采集用户对终端1200的3D动作。处理器1201根据陀螺仪传感器1212采集的数据,可以实现如下功能:动作感应(比如根据用户的倾斜操作来改变UI)、拍摄时的图像稳定、游戏控制以及惯性导航。

[0228] 压力传感器1213可以设置在终端1200的侧边框和/或显示屏1205的下层。当压力传感器1213设置在终端1200的侧边框时,可以检测用户对终端1200的握持信号,由处理器1201根据压力传感器1213采集的握持信号进行左右手识别或快捷操作。当压力传感器1213设置在显示屏1205的下层时,由处理器1201根据用户对显示屏1205的压力操作,实现对UI界面上的可操作性控件进行控制。可操作性控件包括按钮控件、滚动条控件、图标控件、菜单控件中的至少一种。

[0229] 指纹传感器1214用于采集用户的指纹,由处理器1201根据指纹传感器1214采集到的指纹识别用户的身份,或者,由指纹传感器1214根据采集到的指纹识别用户的身份。在识别出用户的身份为可信身份时,由处理器1201授权该用户执行相关的敏感操作,该敏感操作包括解锁屏幕、查看加密信息、下载软件、支付及更改设置等。指纹传感器1214可以被设置在终端1200的正面、背面或侧面。当终端1200上设置有物理按键或厂商Logo时,指纹传感器1214可以与物理按键或厂商Logo集成在一起。

[0230] 光学传感器1215用于采集环境光强度。在一个实施例中,处理器1201可以根据光

学传感器1215采集的环境光强度,控制显示屏1205的显示亮度。具体地,当环境光强度较高时,调高显示屏1205的显示亮度;当环境光强度较低时,调低显示屏1205的显示亮度。在另一个实施例中,处理器1201还可以根据光学传感器1215采集的环境光强度,动态调整摄像头组件1206的拍摄参数。

[0231] 接近传感器1216,也称距离传感器,设置在终端1200的前面板。接近传感器1216用于采集用户与终端1200的正面之间的距离。在一个实施例中,当接近传感器1216检测到用户与终端1200的正面之间的距离逐渐变小时,由处理器1201控制显示屏1205从亮屏状态切换为息屏状态;当接近传感器1216检测到用户与终端1200的正面之间的距离逐渐变大时,由处理器1201控制显示屏1205从息屏状态切换为亮屏状态。

[0232] 本领域技术人员可以理解,图12中示出的结构并不构成对终端1200的限定,可以包括比图示更多或更少的组件,或者组合某些组件,或者采用不同的组件布置。

[0233] 可选地,该计算机设备提供为服务器。图13是根据一示例性实施例示出的一种服务器的结构示意图,该服务器1300可因配置或性能不同而产生比较大的差异,可以包括一个或一个以上处理器(Central Processing Units,CPU)1301和一个或一个以上的存储器1302,其中,存储器1302中存储有至少一条程序代码,至少一条程序代码由处理器1301加载并执行以实现上述各个方法实施例提供的方法。当然,该服务器还可以具有有线或无线网络接口、键盘以及输入输出接口等部件,以便进行输入输出,该服务器还可以包括其他用于实现设备功能的部件,在此不做赘述。

[0234] 本申请实施例还提供了一种计算机可读存储介质,该计算机可读存储介质中存储有至少一条程序代码,该至少一条程序代码由处理器加载并执行,以实现上述实施例的音频数据处理方法。

[0235] 本申请实施例还提供了一种计算机程序产品或计算机程序,该计算机程序产品或计算机程序包括计算机程序代码,该计算机程序代码存储在计算机可读存储介质中,计算机设备的处理器从计算机可读存储介质读取该计算机程序代码,处理器执行该计算机程序代码,使得该计算机设备实现如上述实施例所述的音频数据处理方法。

[0236] 本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成,也可以通过程序来指令相关的硬件完成,该程序可以存储于一种计算机可读存储介质中,上述提到的存储介质可以是只读存储器,磁盘或光盘等。

[0237] 以上仅为本申请实施例的可选实施例,并不用以限制本申请实施例,凡在本申请实施例的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

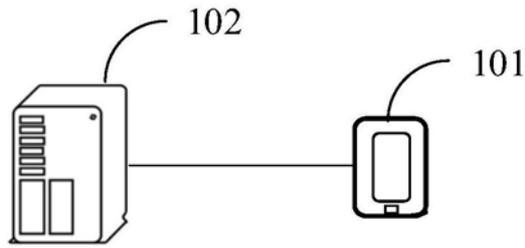


图1

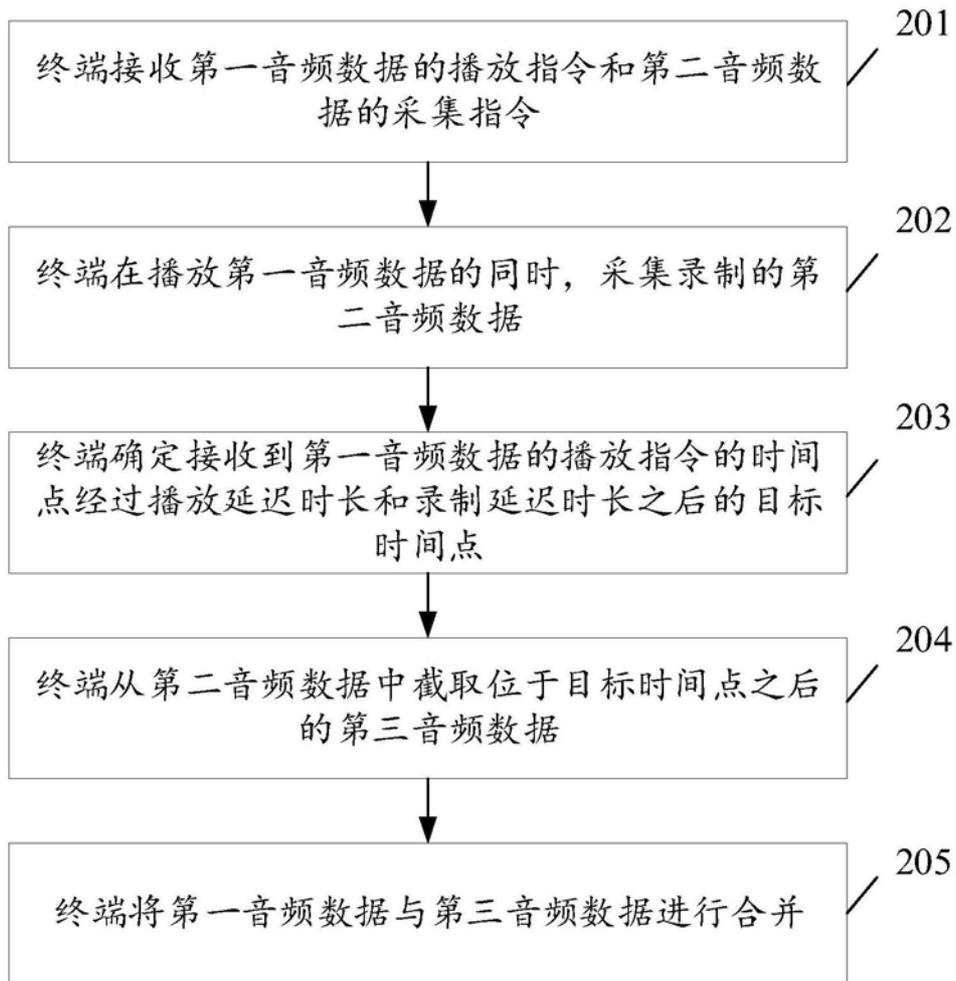


图2

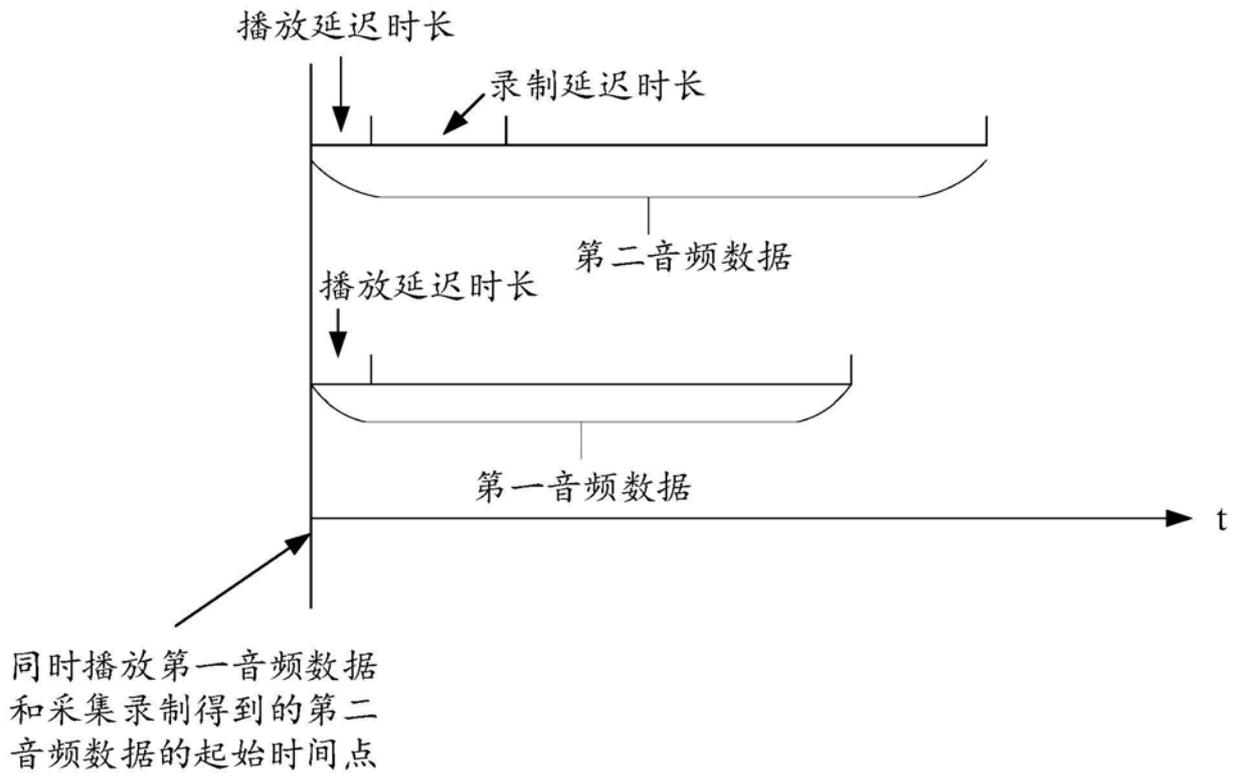


图3

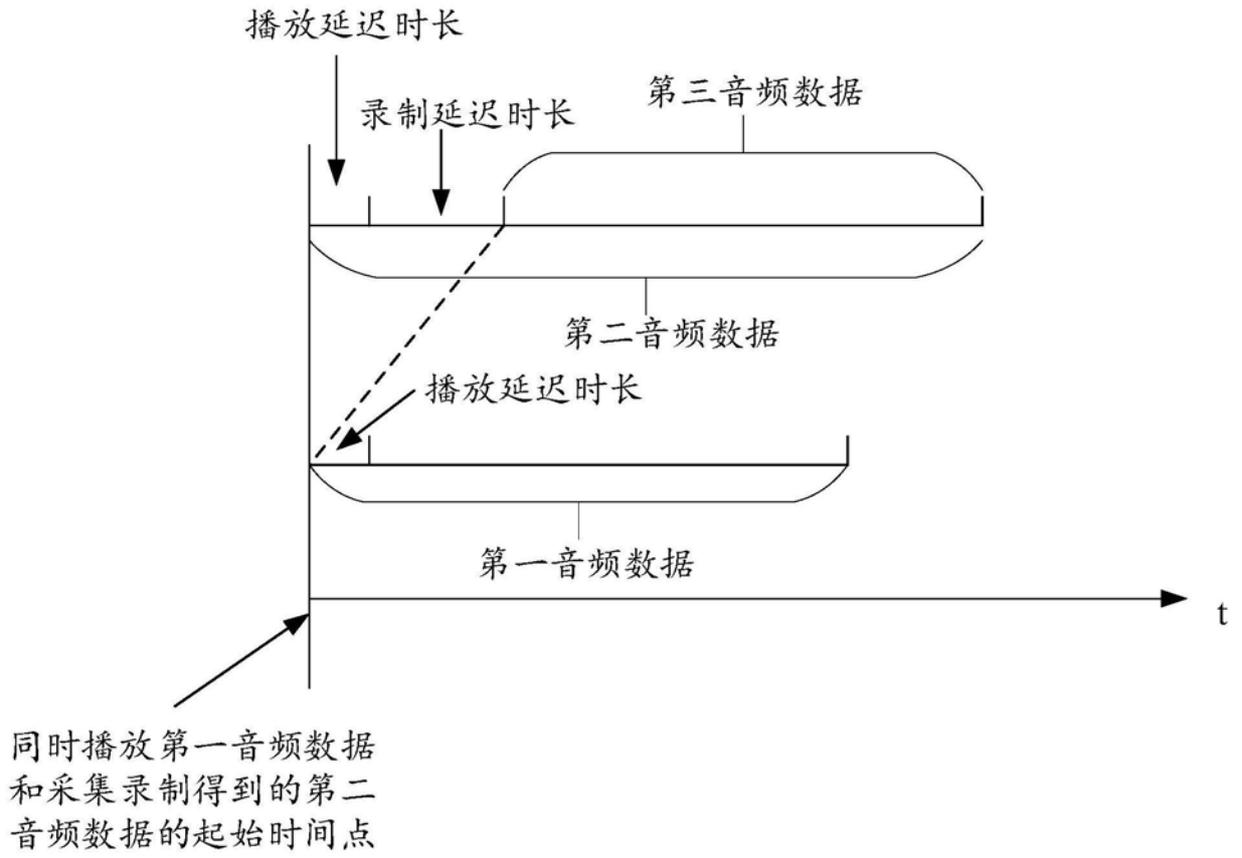


图4

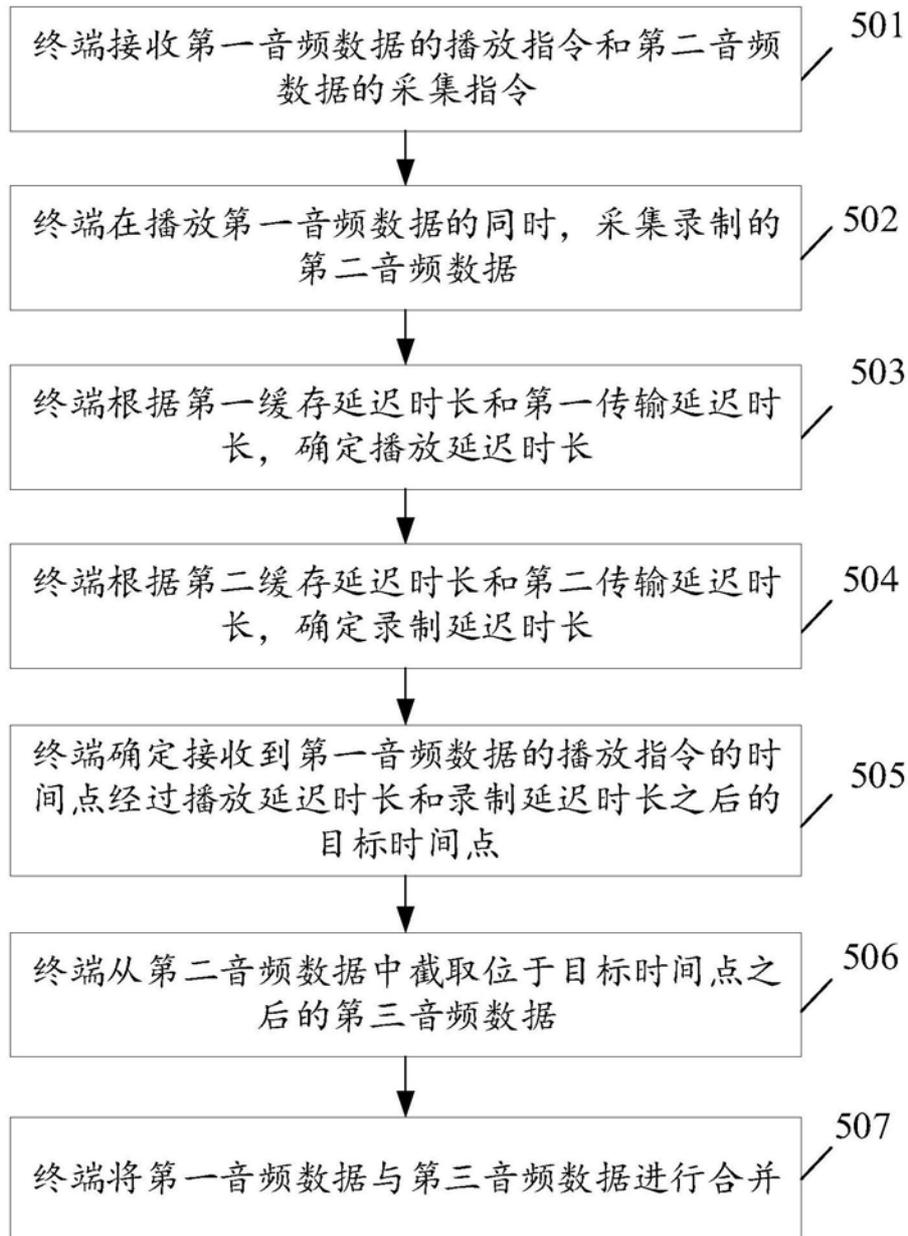


图5

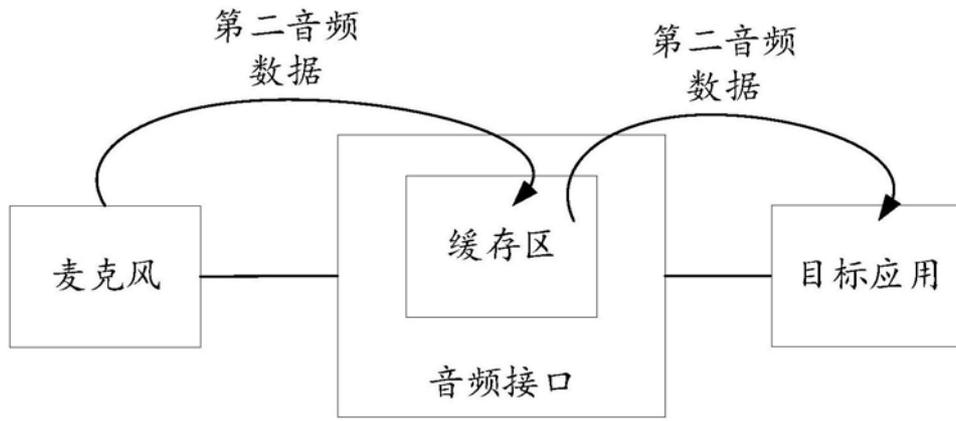


图6

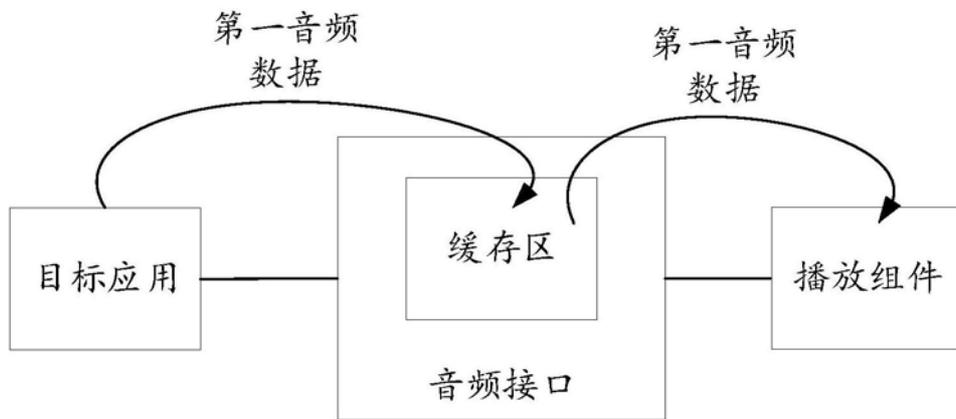


图7

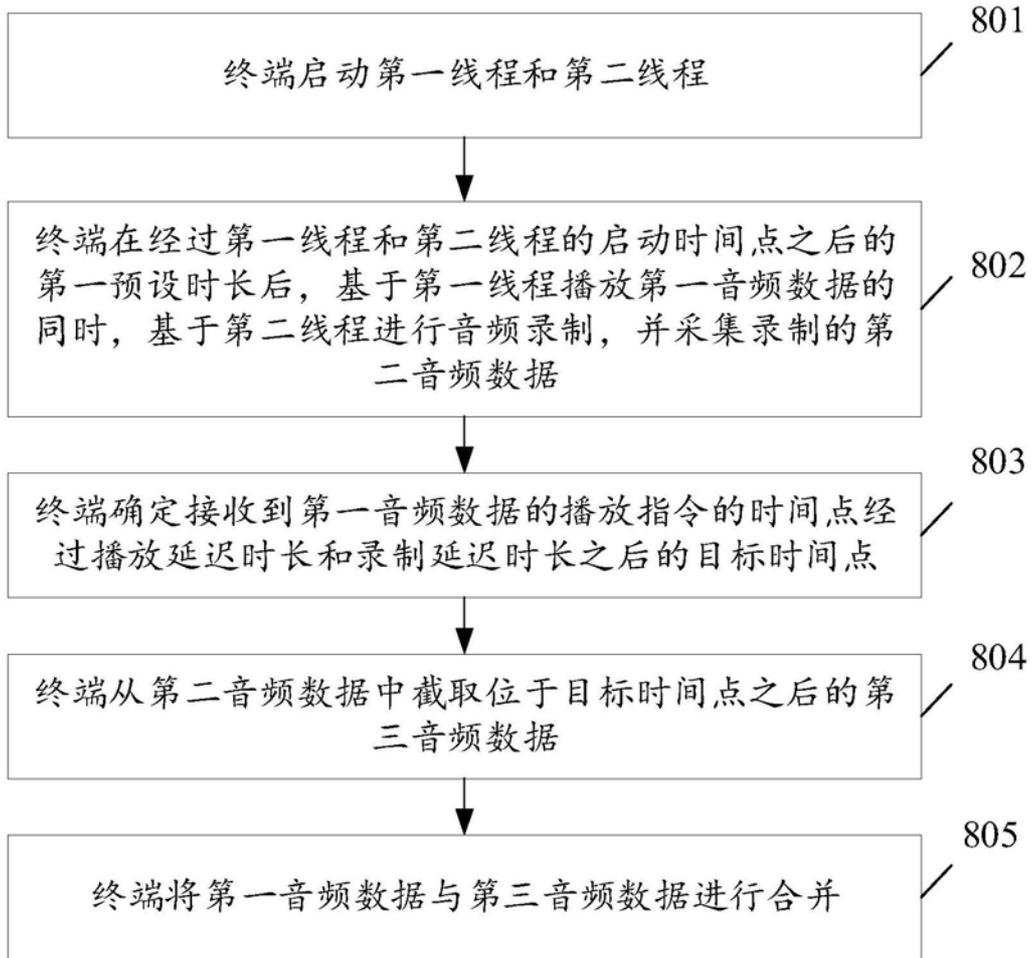


图8

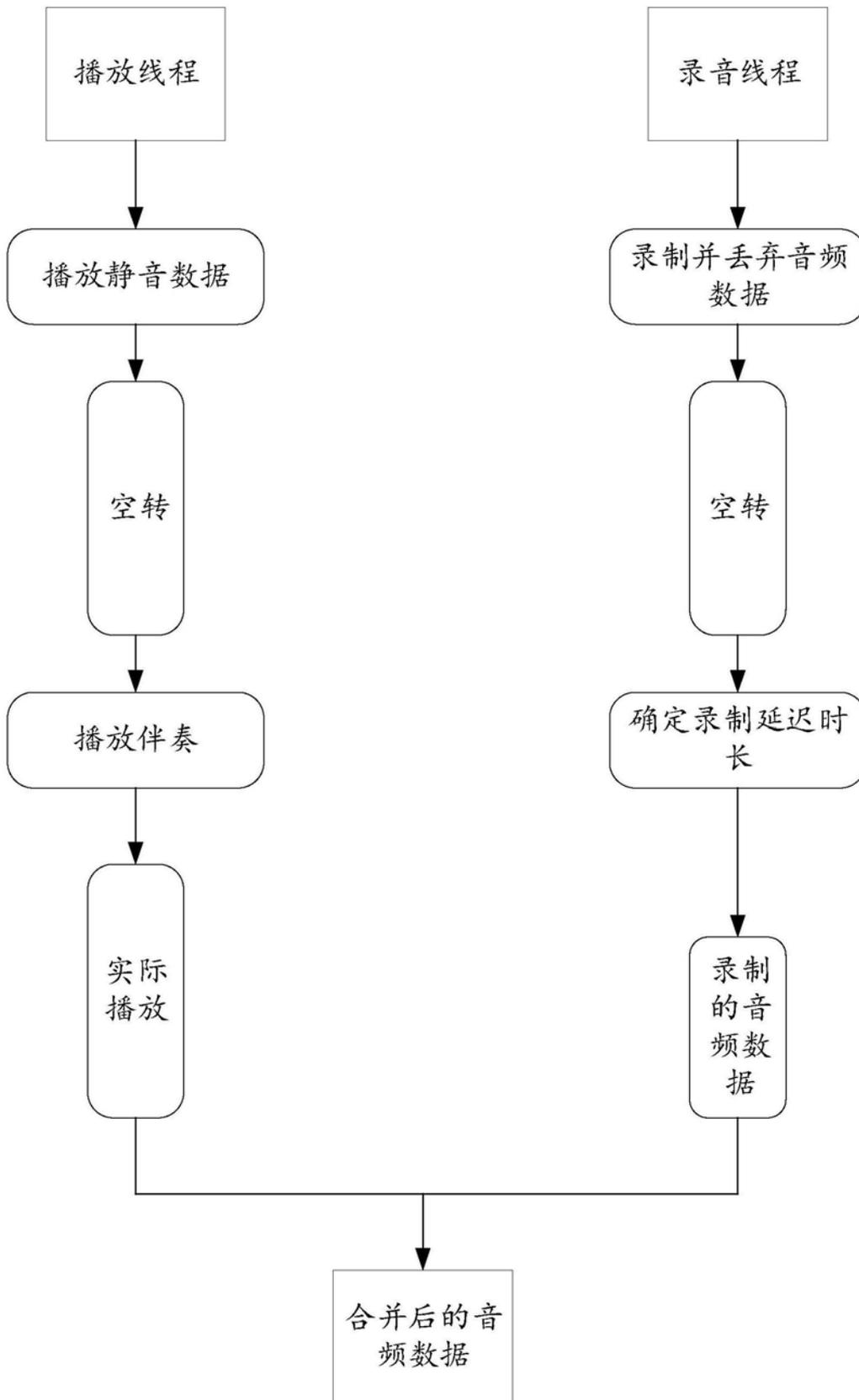


图9

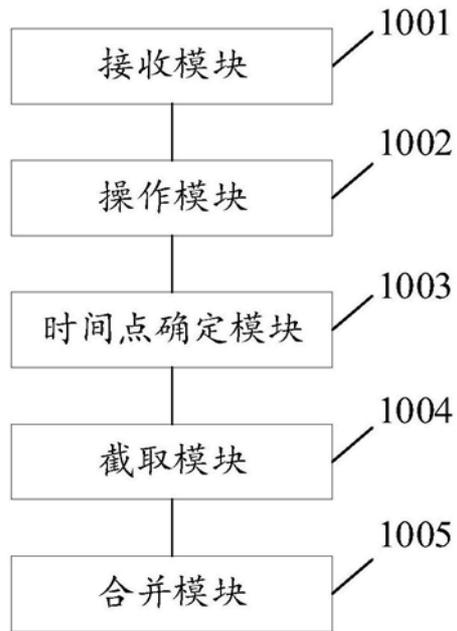


图10

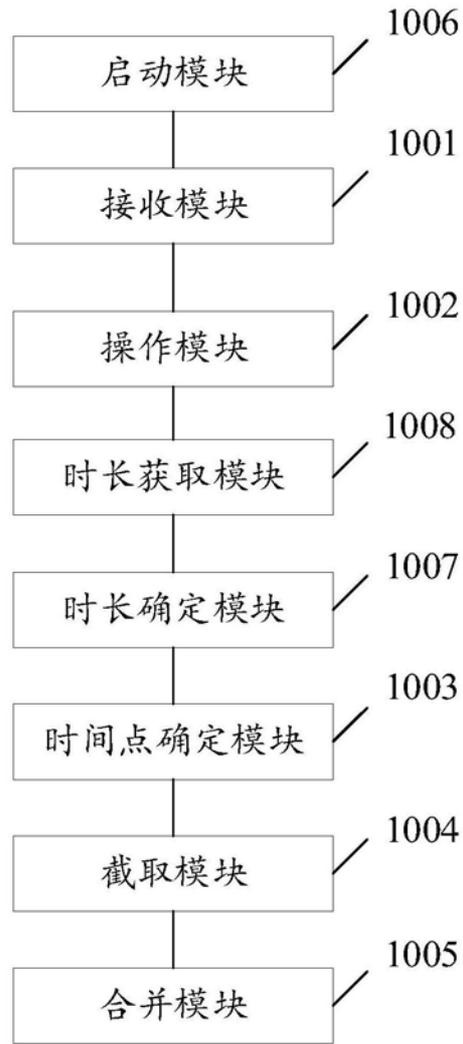


图11

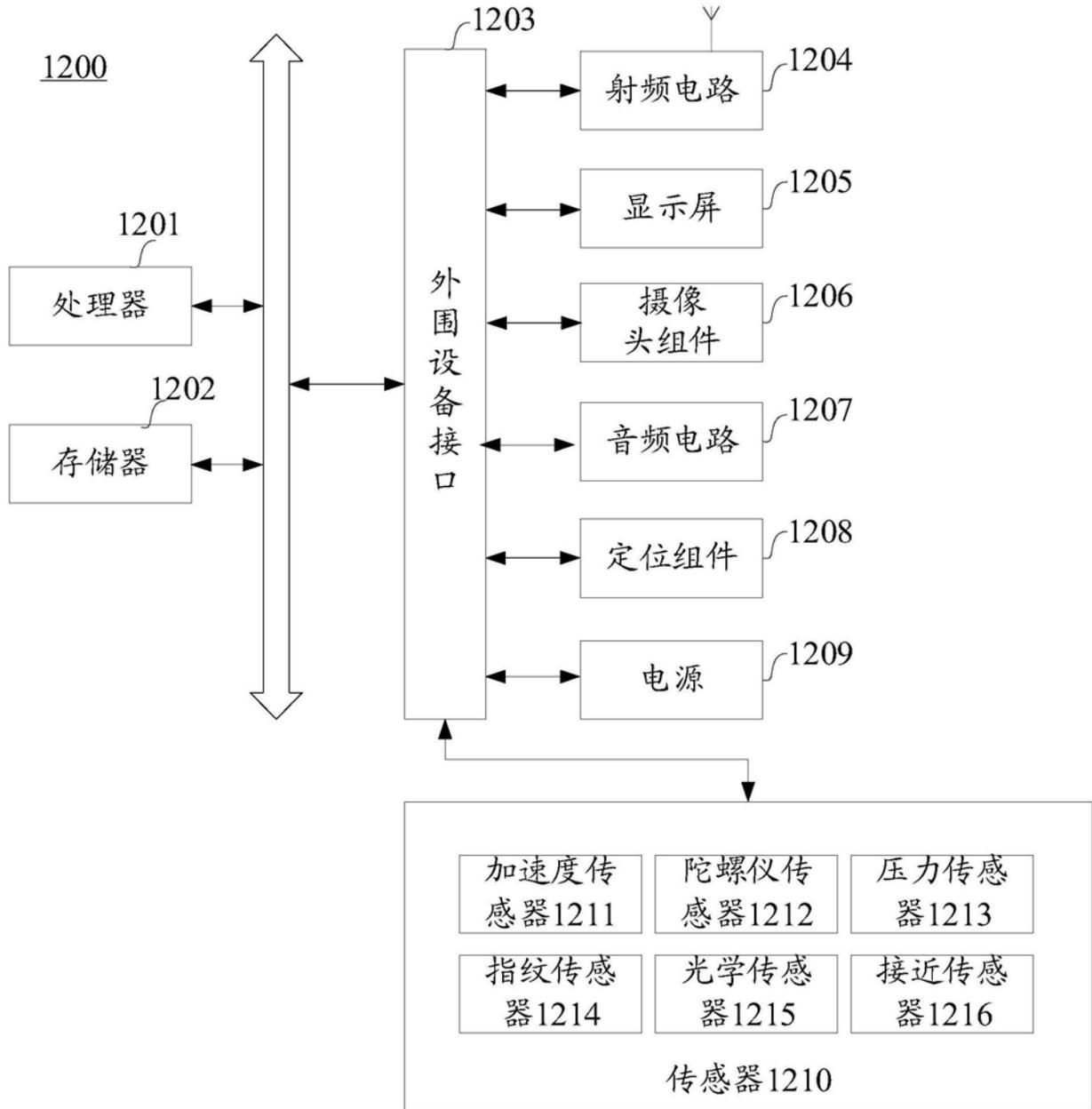


图12

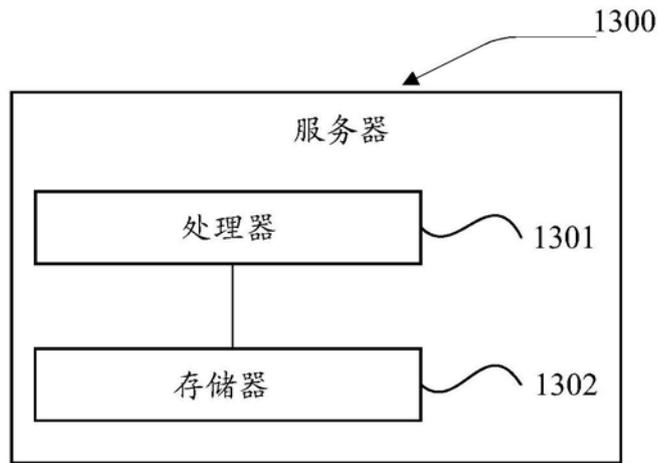


图13