



(12)发明专利申请

(10)申请公布号 CN 106131583 A

(43)申请公布日 2016. 11. 16

(21)申请号 201610513186.X

(22)申请日 2016.06.30

(71)申请人 北京小米移动软件有限公司
地址 100085 北京市海淀区清河中街68号
华润五彩城购物中心二期9层01房间

(72)发明人 韩尚佑 李志刚

(74)专利代理机构 北京博思佳知识产权代理有限公司 11415

代理人 林祥

(51) Int. Cl.

H04N 21/2187(2011.01)

H04N 21/439(2011.01)

H04N 21/6437(2011.01)

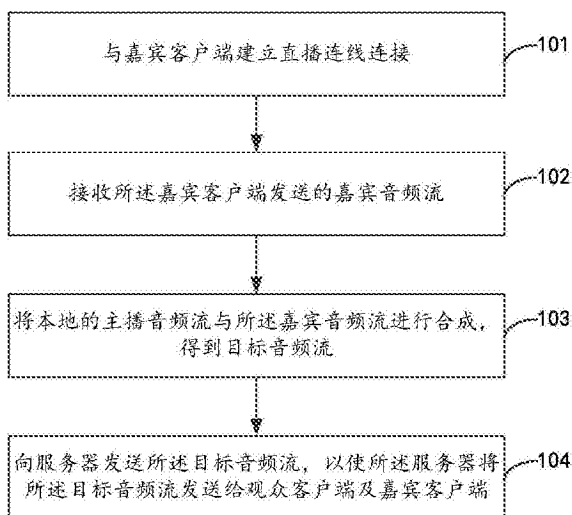
权利要求书3页 说明书11页 附图5页

(54)发明名称

一种直播处理方法、装置、终端设备及系统

(57)摘要

本公开是关于一种直播处理方法、装置、终端设备及系统。该方法包括：与嘉宾客户端建立直播连线连接；接收所述嘉宾客户端发送的嘉宾音频流；将本地的主播音频流与所述嘉宾音频流进行合成，得到目标音频流；向服务器发送所述目标音频流，以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。本公开的方案，能解决多人直播时的通信延迟问题，提高用户体验。



1. 一种直播处理方法,其特征在于,用于主播客户端,包括:
与嘉宾客户端建立直播连线连接;
接收所述嘉宾客户端发送的嘉宾音频流;
将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流;
向服务器发送所述目标音频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。
2. 根据权利要求1所述的方法,其特征在于,所述方法还包括:
接收所述嘉宾客户端发送的嘉宾视频流;
将本地的主播视频流与所述嘉宾视频流进行合成,得到目标视频流;
向服务器发送所述目标视频流,以使所述服务器将所述目标视频流发送给观众客户端及嘉宾客户端。
3. 根据权利要求1所述的方法,其特征在于,所述方法还包括:
当检测到所述主播音频流或所述嘉宾音频流中存在预设关键词时,在直播画面中显示所述预设关键词对应的动画。
4. 根据权利要求2所述的方法,其特征在于,所述方法还包括:
当检测到所述主播视频流或所述嘉宾视频流中存在预设动作或预设表情时,在直播画面中显示所述预设动作或所述预设表情对应的动画。
5. 根据权利要求1所述的方法,其特征在于,所述接收所述嘉宾客户端发送的嘉宾音频流,包括:
基于用户数据报协议UDP的实时传输协议RTP接收所述嘉宾客户端发送的嘉宾音频流。
6. 根据权利要求1所述的方法,其特征在于,所述与嘉宾客户端建立直播连线连接,包括:
根据用户选择指令,向所述嘉宾客户端发送连麦邀请消息,所述用户选择指令中携带有所述嘉宾客户端的标识信息;
当接收到所述嘉宾客户端返回的连麦确认消息时,与所述嘉宾客户端建立直播连线连接。
7. 一种直播处理方法,其特征在于,用于嘉宾客户端,包括:
与主播客户端建立直播连线连接;
向所述主播客户端发送嘉宾音频流,以使所述主播客户端将本地的主播音频流与所述嘉宾音频流进行合成得到目标音频流;
接收所述服务器发送的所述目标音频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。
8. 根据权利要求7所述的方法,其特征在于,所述方法还包括:
向所述主播客户端发送嘉宾视频流,以使所述主播客户端将本地的主播视频流与所述嘉宾视频流进行合成得到目标视频流;
接收所述服务器发送的所述目标视频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。
9. 根据权利要求7所述的方法,其特征在于,所述向所述主播客户端发送嘉宾音频流包括:

基于用户数据报协议UDP的实时传输协议RTP向所述主播客户端发送嘉宾音频流。

10. 根据权利要求7所述的方法,其特征在於,所述与主播客户端建立直播连线连接,包括:

接收所述主播客户端根据用户选择指令发送的连麦邀请消息,所述用户选择指令中携带有所述嘉宾客户端的标识信息;

向所述主播客户端发送连麦确认消息,与所述主播客户端建立直播连线连接。

11. 一种直播处理装置,其特征在於,位于主播客户端,包括:

第一连接模块,用于与嘉宾客户端建立直播连线连接;

第一音频接收模块,用于接收所述嘉宾客户端发送的嘉宾音频流;

音频合成模块,用于将本地的主播音频流与所述第一音频接收模块接收的嘉宾音频流进行合成,得到目标音频流;

音频输出模块,用于向服务器发送所述音频合成模块合成的目标音频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。

12. 根据权利要求11所述的装置,其特征在於,所述装置还包括:

第一视频接收模块,用于接收所述嘉宾客户端发送的嘉宾视频流;

视频合成模块,用于将本地的主播视频流与所述第一视频接收模块接收的嘉宾视频流进行合成,得到目标视频流;

视频输出模块,用于向服务器发送所述视频合成模块合成的目标视频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。

13. 根据权利要求11所述的装置,其特征在於,所述装置还包括:

动画模块,用于当检测到所述主播音频流或所述嘉宾音频流中存在预设关键词时,在直播画面中显示所述预设关键词对应的动画;或者,

当检测到所述主播视频流或所述嘉宾视频流中存在预设动作或预设表情时,在直播画面中显示所述预设动作或所述预设表情对应的动画。

14. 一种直播处理装置,其特征在於,位于嘉宾客户端,包括:

第二连接模块,用于与主播客户端建立直播连线连接;

音频发送模块,用于向所述主播客户端发送嘉宾音频流,以使所述主播客户端将本地的主播音频流与所述嘉宾音频流进行合成得到目标音频流;

第二音频接收模块,用于接收所述服务器发送的所述目标音频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。

15. 根据权利要求14所述的装置,其特征在於,所述装置还包括:

视频发送模块,用于向所述主播客户端发送嘉宾视频流,以使所述主播客户端将本地的主播视频流与所述嘉宾视频流进行合成得到目标视频流;

第二视频接收模块,用于接收所述服务器发送的所述目标视频流,其中所述目标视频流由所述主播客户端合成后发送给所述服务器。

16. 一种终端设备,其特征在於,包括:

处理器和用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为:

与嘉宾客户端建立直播连线连接;

接收所述嘉宾客户端发送的嘉宾音频流；
将本地的主播音频流与所述嘉宾音频流进行合成，得到目标音频流；
向服务器发送所述目标音频流，以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。

17. 一种终端设备，其特征在于，包括：

处理器和用于存储处理器可执行指令的存储器；

其中，所述处理器被配置为：

与主播客户端建立直播连线连接；

向所述主播客户端发送嘉宾音频流，以使所述主播客户端将本地的主播音频流与所述嘉宾音频流进行合成得到目标音频流；

接收所述服务器发送的所述目标音频流，其中所述目标音频流由所述主播客户端合成后发送给所述服务器。

18. 一种直播处理系统，其特征在于，包括：

主播客户端，用于与嘉宾客户端建立直播连线连接，接收所述嘉宾客户端发送的嘉宾音频流，将本地的主播音频流与所述嘉宾音频流进行合成，得到目标音频流，向服务器发送所述目标音频流；

服务器，用于将所述目标音频流发送给观众客户端及所述嘉宾客户端；

嘉宾客户端，用于向所述主播客户端发送嘉宾音频流，接收所述服务器发送的所述目标音频流。

19. 根据权利要求18所述的系统，其特征在于：

所述主播客户端接收所述嘉宾客户端发送的嘉宾视频流，将本地的主播视频流与所述嘉宾视频流进行合成，得到目标视频流，向服务器发送所述目标视频流；

所述服务器将从所述主播客户端接收的目标视频流发送给观众客户端及嘉宾客户端。

一种直播处理方法、装置、终端设备及系统

技术领域

[0001] 本公开涉及移动通信技术领域,尤其涉及一种直播处理方法、装置、终端设备及系统。

背景技术

[0002] 随着移动直播软件的日益发展,移动直播渐渐成为一种深受用户喜爱的直播方式。从早年的PC直播、网页直播、到现在的手机APP(Application,应用程序)直播,移动直播被越来越多的用户所接受和喜爱。在移动直播中,主播的主播客户端(简称主播端)可以通过直播软件向观众进行直播。当前的直播类产品,大多数是一人主播,目前也出现有连麦模式的直播,观众可以同时看到两个主播进行双人直播互动。所谓麦就是麦克风的简称,连麦是指在麦序模式下实现多人同时上麦说话的目的。

[0003] 但是现有的连麦模式的直播处理方法中,存在以下问题:两个主播客户端分属两路视频,这期间的交互会出现视频延迟,导致两个主播之间说话时快时慢,经常会有冲突,观众和主播都会觉得视听体验差,影响使用体验。

发明内容

[0004] 本公开提供了一种直播处理方法、装置、终端设备及系统,能解决多人直播时的通信延迟问题,提高用户体验。

[0005] 根据本公开实施例的第一方面,提供一种直播处理方法,用于主播客户端,包括:

[0006] 与嘉宾客户端建立直播连线连接;

[0007] 接收所述嘉宾客户端发送的嘉宾音频流;

[0008] 将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流;

[0009] 向服务器发送所述目标音频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。

[0010] 可选的,所述方法还包括:

[0011] 接收所述嘉宾客户端发送的嘉宾视频流;

[0012] 将本地的主播视频流与所述嘉宾视频流进行合成,得到目标视频流;

[0013] 向服务器发送所述目标视频流,以使所述服务器将所述目标视频流发送给观众客户端及嘉宾客户端。

[0014] 可选的,所述方法还包括:

[0015] 当检测到所述主播音频流或所述嘉宾音频流中存在预设关键词时,在直播画面中显示所述预设关键词对应的动画。

[0016] 可选的,所述方法还包括:

[0017] 当检测到所述主播视频流或所述嘉宾视频流中存在预设动作或预设表情时,在直播画面中显示所述预设动作或所述预设表情对应的动画。

[0018] 可选的,所述接收所述嘉宾客户端发送的嘉宾音频流,包括:

- [0019] 基于用户数据报协议UDP的实时传输协议RTP接收所述嘉宾客户端发送的嘉宾音频流。
- [0020] 可选的,所述与嘉宾客户端建立直播连线连接,包括:
- [0021] 根据用户选择指令,向所述嘉宾客户端发送连麦邀请消息,所述用户选择指令中携带有所述嘉宾客户端的标识信息;
- [0022] 当接收到所述嘉宾客户端返回的连麦确认消息时,与所述嘉宾客户端建立直播连线连接。
- [0023] 根据本公开实施例的第二方面,提供一种直播处理方法,用于嘉宾客户端,包括:
- [0024] 与主播客户端建立直播连线连接;
- [0025] 向所述主播客户端发送嘉宾音频流,以使所述主播客户端将本地的主播音频流与所述嘉宾音频流进行合成得到目标音频流;
- [0026] 接收所述服务器发送的所述目标音频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。
- [0027] 可选的,所述方法还包括:
- [0028] 向所述主播客户端发送嘉宾视频流,以使所述主播客户端将本地的主播视频流与所述嘉宾视频流进行合成得到目标视频流;
- [0029] 接收所述服务器发送的所述目标视频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。
- [0030] 可选的,所述向所述主播客户端发送嘉宾音频流包括:
- [0031] 基于用户数据报协议UDP的实时传输协议RTP向所述主播客户端发送嘉宾音频流。
- [0032] 可选的,所述与主播客户端建立直播连线连接,包括:
- [0033] 接收所述主播客户端根据用户选择指令发送的连麦邀请消息,所述用户选择指令中携带有所述嘉宾客户端的标识信息;
- [0034] 向所述主播客户端发送连麦确认消息,与所述主播客户端建立直播连线连接。
- [0035] 根据本公开实施例的第三方面,提供一种直播处理装置,位于主播客户端,包括:
- [0036] 第一连接模块,用于与嘉宾客户端建立直播连线连接;
- [0037] 第一音频接收模块,用于接收所述嘉宾客户端发送的嘉宾音频流;
- [0038] 音频合成模块,用于将本地的主播音频流与所述第一音频接收模块接收的嘉宾音频流进行合成,得到目标音频流;
- [0039] 音频输出模块,用于向服务器发送所述音频合成模块合成的目标音频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。
- [0040] 可选的,所述装置还包括:
- [0041] 第一视频接收模块,用于接收所述嘉宾客户端发送的嘉宾视频流;
- [0042] 视频合成模块,用于将本地的主播视频流与所述视频接收模块接收的嘉宾视频流进行合成,得到目标视频流;
- [0043] 视频输出模块,用于向服务器发送所述视频合成模块合成的目标视频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。
- [0044] 可选的,所述装置还包括:
- [0045] 动画模块,用于当检测到所述主播音频流或所述嘉宾音频流中存在预设关键词

时,在直播画面中显示所述预设关键词对应的动画;或者,

[0046] 当检测到所述主播视频流或所述嘉宾视频流中存在预设动作或预设表情时,在直播画面中显示所述预设动作或所述预设表情对应的动画。

[0047] 根据本公开实施例的第四方面,提供一种直播处理装置,位于嘉宾客户端,包括:

[0048] 第二连接模块,用于与主播客户端建立直播连线连接;

[0049] 音频发送模块,用于向所述主播客户端发送嘉宾音频流,以使所述主播客户端将本地的主播音频流与所述嘉宾音频流进行合成得到目标音频流;

[0050] 第二音频接收模块,用于接收所述服务器发送的所述目标音频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。

[0051] 可选的,所述装置还包括:

[0052] 视频发送模块,用于向所述主播客户端发送嘉宾视频流,以使所述主播客户端将本地的主播视频流与所述嘉宾视频流进行合成得到目标视频流;

[0053] 第二视频接收模块,用于接收所述服务器发送的所述目标视频流,其中所述目标视频流由所述主播客户端合成后发送给所述服务器。

[0054] 根据本公开实施例的第五方面,提供一种终端设备,包括:

[0055] 处理器和用于存储处理器可执行指令的存储器;

[0056] 其中,所述处理器被配置为:

[0057] 与嘉宾客户端建立直播连线连接;

[0058] 接收所述嘉宾客户端发送的嘉宾音频流;

[0059] 将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流;

[0060] 向服务器发送所述目标音频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。

[0061] 根据本公开实施例的第六方面,提供一种终端设备,包括:

[0062] 处理器和用于存储处理器可执行指令的存储器;

[0063] 其中,所述处理器被配置为:

[0064] 与主播客户端建立直播连线连接;

[0065] 向所述主播客户端发送嘉宾音频流,以使所述主播客户端将本地的主播音频流与所述嘉宾音频流进行合成得到目标音频流;

[0066] 接收所述服务器发送的所述目标音频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。

[0067] 根据本公开实施例的第七方面,提供一种直播处理系统,包括:

[0068] 主播客户端,用于与嘉宾客户端建立直播连线连接,接收所述嘉宾客户端发送的嘉宾音频流,将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流,向服务器发送所述目标音频流;

[0069] 服务器,用于将所述目标音频流发送给观众客户端及所述嘉宾客户端;

[0070] 嘉宾客户端,用于向所述主播客户端发送嘉宾音频流,接收所述服务器发送的所述目标音频流。

[0071] 可选的,所述主播客户端接收所述嘉宾客户端发送的嘉宾视频流,将本地的主播视频流与所述嘉宾视频流进行合成,得到目标视频流,向服务器发送所述目标视频流;

[0072] 所述服务器将从所述主播客户端接收的目标视频流发送给观众客户端及嘉宾客户端。

[0073] 本公开的实施例提供的技术方案可以包括以下有益效果：

[0074] 本公开在直播过程中，主播客户端接收嘉宾客户端发送的嘉宾音频流后，可以先将本地的主播音频流与所述嘉宾音频流进行合成，得到目标音频流，再向服务器发送所述目标音频流以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端，这样就使得在观众客户端及嘉宾客户端面前播放的是合成处理后的流畅的音频，不会存在之前因为时延导致主播客户端与嘉宾客户端之间的音频对话显得时快时慢或冲突的问题，因此可以解决直播过程中的通信延迟问题，提高直播的音频效果和用户的使用体验。

[0075] 本公开还可以对视频流进行合成，可以将本地的主播视频流与所述嘉宾视频流进行合成得到目标视频流后，再向服务器发送所述目标视频流，以使所述服务器将所述目标视频流发送给观众客户端及嘉宾客户端，这样就使得在观众客户端及嘉宾客户端面前播放的是合成处理后的流畅的视频，因此可以提高直播的视频效果和用户的使用体验。

[0076] 本公开还可以当检测到所述主播音频流或所述嘉宾音频流中存在预设关键词时，在直播画面中显示所述预设关键词对应的动画；或者，当检测到所述主播视频流或所述嘉宾视频流中存在预设动作或预设表情时，在直播画面中显示所述预设动作或所述预设表情对应的动画，这样可以提高直播过程中的用户视觉系效果和使用体验。

[0077] 本公开在直播过程中，主播客户端还可以通过UDP(User Datagram Protocol, 用户数据报协议)的RTP(Real-time Transport Protocol, 实时传输协议)与所述嘉宾客户端进行音频收发，UDP主要就是用于对高速传输和实时性有较高要求的通信或广播通信，网络负荷比一般直播所使用的TCP(Transmission Control Protocol, 传输控制协议)小，传输速度也要比TCP快。另外RTP协议是为数据提供了具有实时特征的端对端传送服务，如在组播或单播网络服务下的交互式视频音频或模拟数据等。因此本公开使用UDP协议进行音频通话也可以极大减小通信延迟。

[0078] 应当理解的是，以上的一般描述和后文的细节描述仅是示例性和解释性的，并不能限制本公开。

附图说明

[0079] 此处的附图被并入说明书中并构成本说明书的一部分，示出了符合本公开的实施例，并与说明书一起用于解释本公开的原理。

[0080] 图1是本公开根据一示例性实施例示出的一种直播处理方法的流程图。

[0081] 图2是本公开根据一示例性实施例示出的一种直播处理方法的另一流程图。

[0082] 图3是本公开根据一示例性实施例示出的一种直播处理方法的另一流程图。

[0083] 图4是本公开根据一示例性实施例示出的一种直播处理装置的框图。

[0084] 图5是本公开根据一示例性实施例示出的一种直播处理装置的另一框图。

[0085] 图6是本公开根据一示例性实施例示出的一种直播处理装置的另一框图。

[0086] 图7是本公开根据一示例性实施例示出的一种直播处理装置的另一框图。

[0087] 图8是本公开根据一示例性实施例示出的一种终端设备的一结构框图。

[0088] 图9是本公开根据一示例性实施例示出的一种终端设备的一结构框图。

[0089] 图10是本公开根据一示例性实施例示出的一种设备的一结构框图。

[0090] 图11是本公开根据一示例性实施例示出的一种直播处理系统的一结构框图。

具体实施方式

[0091] 这里将详细地对示例性实施例进行说明,其示例表示在附图中。下面的描述涉及附图时,除非另有表示,不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本公开相一致的所有实施方式。相反,它们仅是与如所附权利要求书中所详述的、本公开的一些方面相一致的装置和方法的例子。

[0092] 在本公开使用的术语是仅仅出于描述特定实施例的目的,而非旨在限制本公开。在本公开和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式,除非上下文清楚地表示其他含义。还应当理解,本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

[0093] 应当理解,尽管在本公开可能采用术语第一、第二、第三等来描述各种信息,但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如,在不脱离本公开范围的情况下,第一信息也可以被称为第二信息,类似地,第二信息也可以被称为第一信息。取决于语境,如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[0094] 本公开提供了一种直播处理方法,能解决多人直播时的通信延迟问题,提高用户体验。

[0095] 图1是本公开根据一示例性实施例示出的一种直播处理方法的流程图。

[0096] 该方法可以应用于主播客户端中,如图1所示,该方法可以包括以下步骤:

[0097] 在步骤101中,与嘉宾客户端建立直播连线连接。

[0098] 该步骤中,主播客户端可以根据用户选择指令,向所述嘉宾客户端发送连麦邀请消息,所述用户选择指令中携带有所述嘉宾客户端的标识信息;当接收到所述嘉宾客户端返回的连麦确认消息时,与所述嘉宾客户端建立直播连线连接。

[0099] 在步骤102中,接收所述嘉宾客户端发送的嘉宾音频流。

[0100] 该步骤中,主播客户端可以基于用户数据报协议UDP的实时传输协议RTP接收所述嘉宾客户端发送的嘉宾音频流。

[0101] 在步骤103中,将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流。

[0102] 该步骤中,主播客户端可以将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流。

[0103] 在步骤104中,向服务器发送所述目标音频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。

[0104] 其中,上述方法还可以包括:接收所述嘉宾客户端发送的嘉宾视频流;将本地的主播视频流与所述嘉宾视频流进行合成,得到目标视频流;向服务器发送所述目标视频流,以使所述服务器将所述目标视频流发送给观众客户端及嘉宾客户端。

[0105] 其中,上述方法还可以包括:当检测到所述主播音频流或所述嘉宾音频流中存在预设关键词时,在直播画面中显示所述预设关键词对应的动画。

[0106] 其中,上述方法还可以包括:当检测到所述主播视频流或所述嘉宾视频流中存在预设动作或预设表情时,在直播画面中显示所述预设动作或所述预设表情对应的动画,这样可以提高直播过程中的用户视觉系效果和使用体验。

[0107] 由该实施例可见,本公开在直播过程中,主播客户端接收嘉宾客户端发送的嘉宾音频流后,可以先将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流,再向服务器发送所述目标音频流以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端,这样就使得在观众客户端及嘉宾客户端面前播放的是合成处理后的流畅的音频,不会存在之前因为时延导致主播客户端与嘉宾客户端之间的音频对话显得时快时慢或冲突的问题,因此可以解决直播过程中的通信延迟问题,提高直播的音频效果和用户的使用体验。

[0108] 图2是本公开根据一示例性实施例示出的一种直播处理方法的另一流程图。

[0109] 该方法可以应用于嘉宾客户端中,如图2所示,该方法可以包括以下步骤:

[0110] 在步骤201中,与主播客户端建立直播连线连接。

[0111] 该步骤中,嘉宾客户端可以接收所述主播客户端根据用户选择指令发送的连麦邀请消息,所述用户选择指令中携带有所述嘉宾客户端的标识信息;向所述主播客户端发送连麦确认消息,与所述主播客户端建立直播连线连接。

[0112] 在步骤202中,向所述主播客户端发送嘉宾音频流,以使所述主播客户端将本地的主播音频流与所述嘉宾音频流进行合成得到目标音频流。

[0113] 该步骤中,嘉宾客户端可以基于用户数据报协议UDP的实时传输协议RTP向所述主播客户端发送嘉宾音频流。

[0114] 在步骤203中,接收所述服务器发送的所述目标音频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。

[0115] 其中,上述方法还可以包括:向所述主播客户端发送嘉宾视频流,以使所述主播客户端将本地的主播视频流与所述嘉宾视频流进行合成得到目标视频流;接收所述服务器发送的所述目标视频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。

[0116] 从上述实施例可以看出,本公开在直播过程中,嘉宾客户端是将嘉宾音频流发送给主播客户端以使所述主播客户端将本地的主播音频流与所述嘉宾音频流进行合成得到目标音频流,主播客户端再向服务器发送所述目标音频流以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端,这样就使得在观众客户端及嘉宾客户端面前播放的是合成处理后的流畅的音频,不会存在之前因为时延导致主播客户端与嘉宾客户端之间的音频对话显得时快时慢或冲突的问题,因此可以解决直播过程中的通信延迟问题,提高直播的音频效果和用户的使用体验。

[0117] 图3是本公开根据一示例性实施例示出的一种直播处理方法的另一流程图。

[0118] 该方法可以应用于终端设备中。图3相对于图1和图2更详细描述了本发明方案,并描述了主播客户端和嘉宾客户端的交互过程。

[0119] 如图3所示,该方法可以包括以下步骤:

[0120] 在步骤301中,主播客户端与嘉宾客户端建立直播连线连接。

[0121] 该步骤中,主播客户端与嘉宾客户端可以通过无线网络建立直播连线连接,也就是主播和嘉宾可以进行两人直播。

[0122] 在步骤302中,主播客户端基于用户数据报协议UDP的实时传输协议RTP接收所述嘉宾客户端发送的嘉宾音频流。

[0123] 在步骤303中,主播客户端将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流,并发送给服务器。

[0124] 在步骤304中,主播客户端接收所述嘉宾客户端发送的嘉宾视频流。

[0125] 在步骤305中,主播客户端将本地的主播视频流与所述嘉宾视频流进行合成,得到目标视频流,并发送给服务器。

[0126] 需说明的是,步骤302、303与步骤304、305之间没有必然的顺序关系。

[0127] 在步骤306中,服务器将所述目标音频流和所述目标视频流发送给观众客户端及嘉宾客户端。

[0128] 需说明的是,上述方法中,当检测到所述主播音频流或所述嘉宾音频流中存在预设关键词时,在直播画面中显示所述预设关键词对应的动画;或者,当检测到所述主播视频流或所述嘉宾视频流中存在预设动作或预设表情时,在直播画面中显示所述预设动作或所述预设表情对应的动画。

[0129] 还需说明的是,上述方法中可以只有音频流合成而没有视频流合成。

[0130] 也就是说,本公开在A主播客户端设备和B嘉宾客户端设备(也可以是其他具备音视频通话能力的设备)之间,可以由A主播客户端设备先将本地的主播音频流与B嘉宾客户端设备的嘉宾音频流进行合成,得到目标音频流,再向服务器发送所述目标音频流以使所述服务器将所述目标音频流发送给观众客户端及B嘉宾客户端设备,这样就使得在观众客户端及B嘉宾客户端设备面前播放的是合成处理后的流畅的音频,从而提高用户体验。

[0131] 与前述应用功能实现方法实施例相对应,本公开还提供了一种直播处理装置、终端设备、设备、系统及相应的的实施例。

[0132] 图4是本公开根据一示例性实施例示出的一种直播处理装置的框图。

[0133] 该装置可以是位于主播客户端中。如图4所示,在一种直播处理装置中可以包括:第一连接模块41、第一音频接收模块42、音频合成模块43、音频输出模块44。

[0134] 第一连接模块41,用于与嘉宾客户端建立直播连线连接。

[0135] 第一音频接收模块42,用于接收所述嘉宾客户端发送的嘉宾音频流。

[0136] 音频合成模块43,用于将本地的主播音频流与所述第一音频接收模块42接收的嘉宾音频流进行合成,得到目标音频流。

[0137] 音频输出模块44,用于向服务器发送所述音频合成模块43合成的目标音频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。

[0138] 由该实施例可见,本公开在直播过程中,主播客户端接收嘉宾客户端发送的嘉宾音频流后,可以先将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流,再向服务器发送所述目标音频流以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端,这样就使得在观众客户端及嘉宾客户端面前播放的是合成处理后的流畅的音频,不会存在之前因为时延导致主播客户端与嘉宾客户端之间的音频对话显得时快时慢或冲突的问题,因此可以解决直播过程中的通信延迟问题,提高直播的音频效果和用户的使用体验。

[0139] 图5是本公开根据一示例性实施例示出的一种直播处理装置的另一框图。

[0140] 图5中,直播处理装置包括第一连接模块41、第一音频接收模块42、音频合成模块43、音频输出模块44,功能可参见图4描述。

[0141] 其中,直播处理装置还可以包括:第一视频接收模块45、视频合成模块46、视频输出模块47。

[0142] 第一视频接收模块45,用于接收所述嘉宾客户端发送的嘉宾视频流。

[0143] 视频合成模块46,用于将本地的主播视频流与所述第一视频接收模块45接收的嘉宾视频流进行合成,得到目标视频流。

[0144] 视频输出模块47,用于向服务器发送所述视频合成模块46合成的目标视频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。

[0145] 其中,直播处理装置还可以包括:动画模块48。

[0146] 动画模块48,用于当检测到所述主播音频流或所述嘉宾音频流中存在预设关键词时,在直播画面中显示所述预设关键词对应的动画;或者,当检测到所述主播视频流或所述嘉宾视频流中存在预设动作或预设表情时,在直播画面中显示所述预设动作或所述预设表情对应的动画。

[0147] 图6是本公开根据一示例性实施例示出的一种直播处理装置的另一框图。

[0148] 该装置可以是设置在嘉宾客户端中。如图6所示,在一种直播处理装置中可以包括:第二连接模块51、音频发送模块52、第二音频接收模块53。

[0149] 第二连接模块51,用于与主播客户端建立直播连线连接。

[0150] 音频发送模块52,用于向所述主播客户端发送嘉宾音频流,以使所述主播客户端将本地的主播音频流与所述嘉宾音频流进行合成得到目标音频流。

[0151] 第二音频接收模块53,用于接收所述服务器发送的所述目标音频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。

[0152] 图7是本公开根据一示例性实施例示出的一种直播处理装置的另一框图。

[0153] 图7中,直播处理装置包括:第二连接模块51、音频发送模块52、第二音频接收模块53,功能可参见图6所示。

[0154] 其中,直播处理装置还可以包括:视频发送模块54、第二视频接收模块55。

[0155] 视频发送模块54,用于向所述主播客户端发送嘉宾视频流,以使所述主播客户端将本地的主播视频流与所述嘉宾视频流进行合成得到目标视频流。

[0156] 第二视频接收模块55,用于接收所述服务器发送的所述目标视频流,其中所述目标视频流由所述主播客户端合成后发送给所述服务器。

[0157] 上述装置中各个单元的功能和作用的实现过程具体详见上述方法中对应步骤的实现过程,在此不再赘述。

[0158] 对于装置实施例而言,由于其基本对应于方法实施例,所以相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的,其中作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本公开方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0159] 图8是本公开根据一示例性实施例示出的一种终端设备的一结构框图。

[0160] 如图8所示,在终端设备中可以包括:处理器601和用于存储处理器601可执行指令的存储器602;

[0161] 其中,处理器601被配置为:

[0162] 与嘉宾客户端建立直播连线连接;

[0163] 接收所述嘉宾客户端发送的嘉宾音频流;

[0164] 将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流;

[0165] 向服务器发送所述目标音频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。

[0166] 还需说明的是,存储器602存储的其他程序,具体参见前面方法流程中的描述,此处不再赘述,处理器601还用于执行存储器602存储的其他程序。

[0167] 图9是本公开根据一示例性实施例示出的一种终端设备的一结构框图。

[0168] 如图9所示,在终端设备中可以包括:处理器601和用于存储处理器601可执行指令的存储器702;

[0169] 其中,处理器701被配置为:

[0170] 与主播客户端建立直播连线连接;

[0171] 向所述主播客户端发送嘉宾音频流,以使所述主播客户端将本地的主播音频流与所述嘉宾音频流进行合成得到目标音频流;

[0172] 接收所述服务器发送的所述目标音频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。

[0173] 图10是本公开根据一示例性实施例示出的一种设备的一结构框图。

[0174] 例如,设备800可以是移动电话,计算机,数字广播终端,消息收发设备,游戏控制台,平板设备,医疗设备,健身设备,个人数字助理等。

[0175] 参照图8,设备800可以包括以下一个或多个组件:处理组件802,存储器804,电源组件806,多媒体组件808,音频组件810,输入/输出(I/O)的接口812,传感器组件814,以及通信组件816。

[0176] 处理组件802通常控制设备800的整体操作,诸如与显示,电话呼叫,数据通信,相机操作和记录操作相关联的操作。处理组件802可以包括一个或多个处理器820来执行指令,以完成上述的方法的全部或部分步骤。此外,处理组件802可以包括一个或多个模块,便于处理组件802和其他组件之间的交互。例如,处理组件802可以包括多媒体模块,以方便多媒体组件808和处理组件802之间的交互。

[0177] 存储器804被配置为存储各种类型的数据以支持在设备800的操作。这些数据的示例包括用于在设备800上操作的任何应用程序或方法的指令,联系人数据,电话簿数据,消息,图片,视频等。存储器804可以由任何类型的易失性或非易失性存储设备或者它们的组合实现,如静态随机存取存储器(SRAM),电可擦除可编程只读存储器(EEPROM),可擦除可编程只读存储器(EPROM),可编程只读存储器(PROM),只读存储器(ROM),磁存储器,快闪存储器,磁盘或光盘。

[0178] 电源组件806为设备800的各种组件提供电力。电源组件806可以包括电源管理系统,一个或多个电源,及其他与为设备800生成、管理和分配电力相关联的组件。

[0179] 多媒体组件808包括在设备800和用户之间的提供一个输出接口的屏幕。在一些实

施例中,屏幕可以包括液晶显示器(LCD)和触摸面板(TP)。如果屏幕包括触摸面板,屏幕可以被实现为触摸屏,以接收来自用户的输入信号。触摸面板包括一个或多个触摸传感器以感测触摸、滑动和触摸面板上的手势。触摸传感器可以不仅感测触摸或滑动动作的边界,而且还检测与触摸或滑动操作相关的持续时间和压力。在一些实施例中,多媒体组件808包括一个前置摄像头和/或后置摄像头。当设备800处于操作模式,如拍摄模式或视频模式时,前置摄像头和/或后置摄像头可以接收外部的多媒体数据。每个前置摄像头和后置摄像头可以是一个固定的光学透镜系统或具有焦距和光学变焦能力。

[0180] 音频组件810被配置为输出和/或输入音频信号。例如,音频组件810包括一个麦克风(MIC),当设备800处于操作模式,如呼叫模式、记录模式和语音识别模式时,麦克风被配置为接收外部音频信号。所接收的音频信号可以被进一步存储在存储器804或经由通信组件816发送。在一些实施例中,音频组件810还包括一个扬声器,用于输出音频信号。

[0181] I/O接口812为处理组件802和外围接口模块之间提供接口,上述外围接口模块可以是键盘,点击轮,按钮等。这些按钮可包括但不限于:主页按钮、音量按钮、启动按钮和锁定按钮。

[0182] 传感器组件814包括一个或多个传感器,用于为设备800提供各个方面的状态评估。例如,传感器组件814可以检测到设备800的打开/关闭状态,组件的相对定位,例如组件为设备800的显示器和小键盘,传感器组件814还可以检测设备800或设备800一个组件的位置改变,用户与设备800接触的存在或不存在,设备800方位或加速/减速和设备800的温度变化。传感器组件814可以包括接近传感器,被配置用来在没有任何的物理接触时检测附近物体的存在。传感器组件814还可以包括光传感器,如CMOS或CCD图像传感器,用于在成像应用中使用。在一些实施例中,该传感器组件814还可以包括加速度传感器,陀螺仪传感器,磁传感器,压力传感器或温度传感器。

[0183] 通信组件816被配置为便于设备800和其他设备之间有线或无线方式的通信。设备800可以接入基于通信标准的无线网络,如WiFi,2G或3G,或它们的组合。在一个示例性实施例中,通信组件816经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中,通信组件816还包括近场通信(NFC)模块,以促进短程通信。例如,在NFC模块可基于射频识别(RFID)技术,红外数据协会(IrDA)技术,超宽带(UWB)技术,蓝牙(BT)技术和其他技术来实现。

[0184] 在示例性实施例中,设备800可以被一个或多个应用专用集成电路(ASIC)、数字信号处理器(DSP)、数字信号处理设备(DSPD)、可编程逻辑器件(PLD)、现场可编程门阵列(FPGA)、控制器、微控制器、微处理器或其他电子元件实现,用于执行上述方法。

[0185] 在示例性实施例中,还提供了一种包括指令的非临时性计算机可读存储介质,例如包括指令的存储器804,上述指令可由设备800的处理器820执行以完成上述方法。例如,非临时性计算机可读存储介质可以是ROM、随机存取存储器(RAM)、CD-ROM、磁带、软盘和光数据存储设备等。

[0186] 一种非临时性计算机可读存储介质,当存储介质中的指令由终端设备的处理器执行时,使得终端设备能够执行一种直播处理方法,方法包括:

[0187] 与嘉宾客户端建立直播连线连接;

[0188] 接收所述嘉宾客户端发送的嘉宾音频流;

[0189] 将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流;

[0190] 向服务器发送所述目标音频流,以使所述服务器将所述目标音频流发送给观众客户端及嘉宾客户端。

[0191] 一种非临时性计算机可读存储介质,当存储介质中的指令由终端设备的处理器执行时,使得终端设备能够执行一种直播处理方法,方法包括:

[0192] 与主播客户端建立直播连线连接;

[0193] 向所述主播客户端发送嘉宾音频流,以使所述主播客户端将本地的主播音频流与所述嘉宾音频流进行合成得到目标音频流;接收所述服务器发送的所述目标音频流,其中所述目标音频流由所述主播客户端合成后发送给所述服务器。

[0194] 图11是本公开根据一示例性实施例示出的一种直播处理系统的一结构框图。

[0195] 如图11所示,在一种直播处理系统中包括:主播客户端91、嘉宾客户端92、服务器93。

[0196] 主播客户端91,用于与嘉宾客户端92建立直播连线连接,接收所述嘉宾客户端92发送的嘉宾音频流,将本地的主播音频流与所述嘉宾音频流进行合成,得到目标音频流,向服务器93发送所述目标音频流。

[0197] 服务器93,用于将所述目标音频流发送给观众客户端及所述嘉宾客户端92。

[0198] 嘉宾客户端92,用于向所述主播客户端发送嘉宾音频流,接收所述服务器发送的所述目标音频流。

[0199] 其中,所述主播客户端91还可以接收所述嘉宾客户端92发送的嘉宾视频流,将本地的主播视频流与所述嘉宾视频流进行合成,得到目标视频流,向服务器93发送所述目标视频流;所述服务器93将从所述主播客户端91接收的目标视频流发送给观众客户端及嘉宾客户端92。

[0200] 主播客户端91的结构可以参见图4、图5和图8所示,嘉宾客户端92的结构可以参见图6、图7和图9所示,此处不再赘述。

[0201] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本公开的其他实施方案。本公开旨在涵盖本公开的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本公开的一般性原理并包括本公开未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本公开的真正范围和精神由下面的权利要求指出。

[0202] 应当理解的是,本公开并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本公开的范围仅由所附的权利要求来限制。

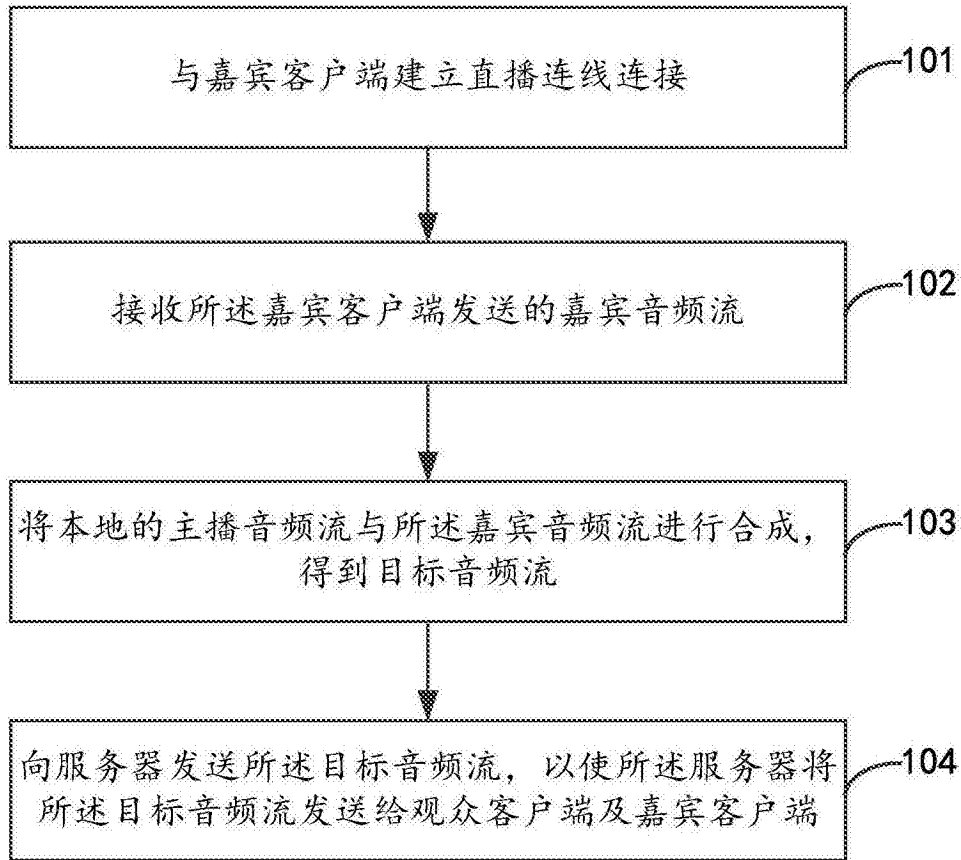


图1

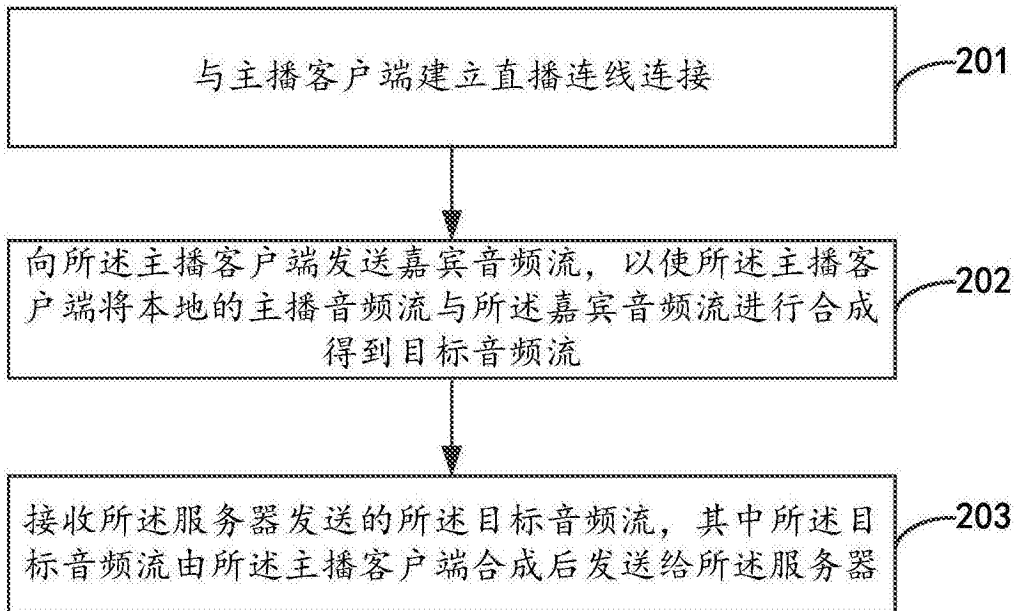


图2

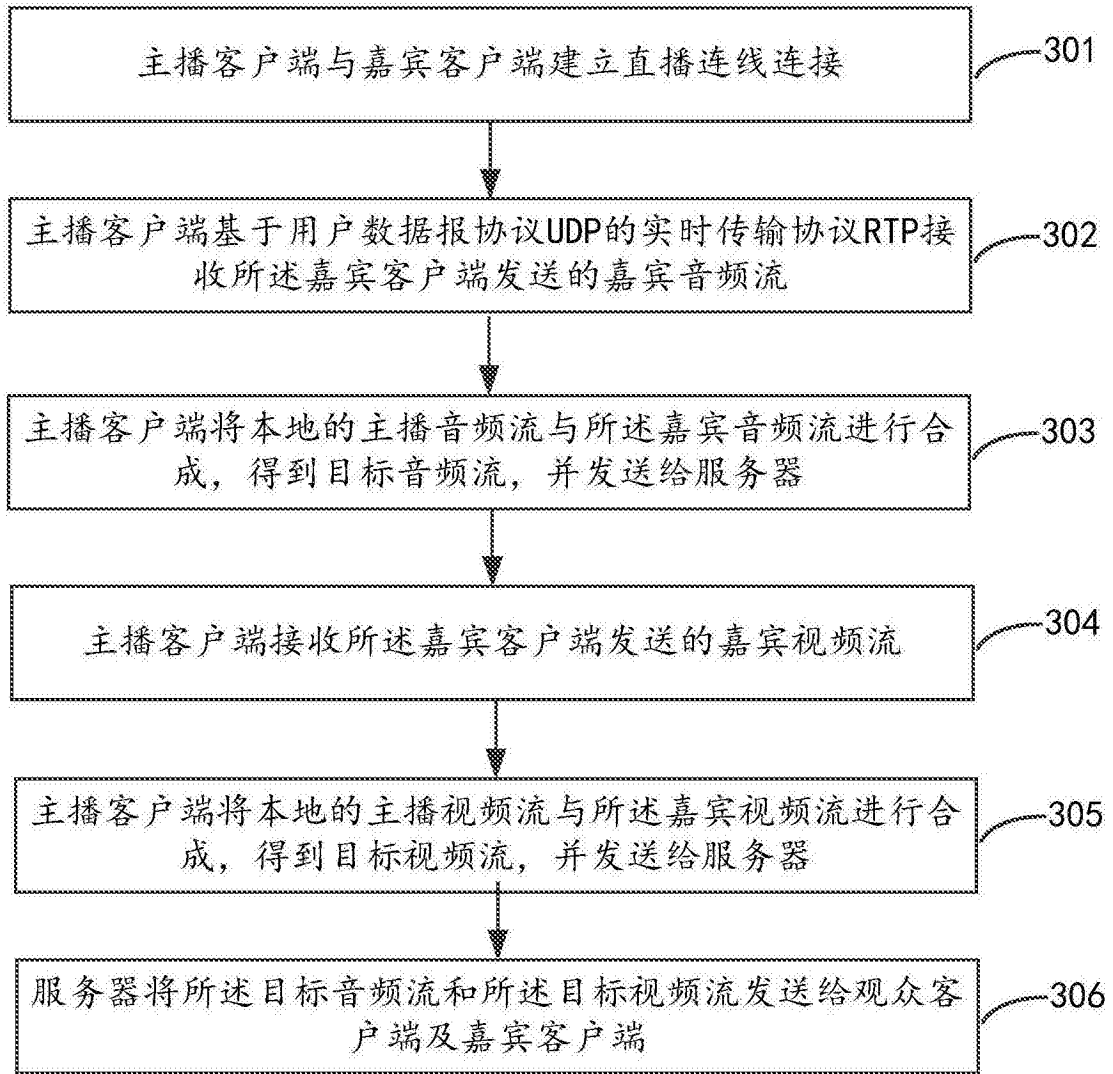


图3

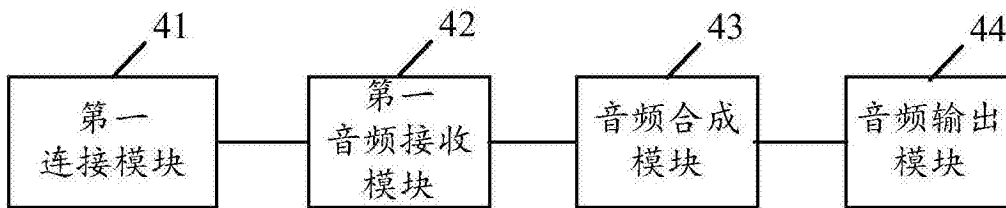


图4

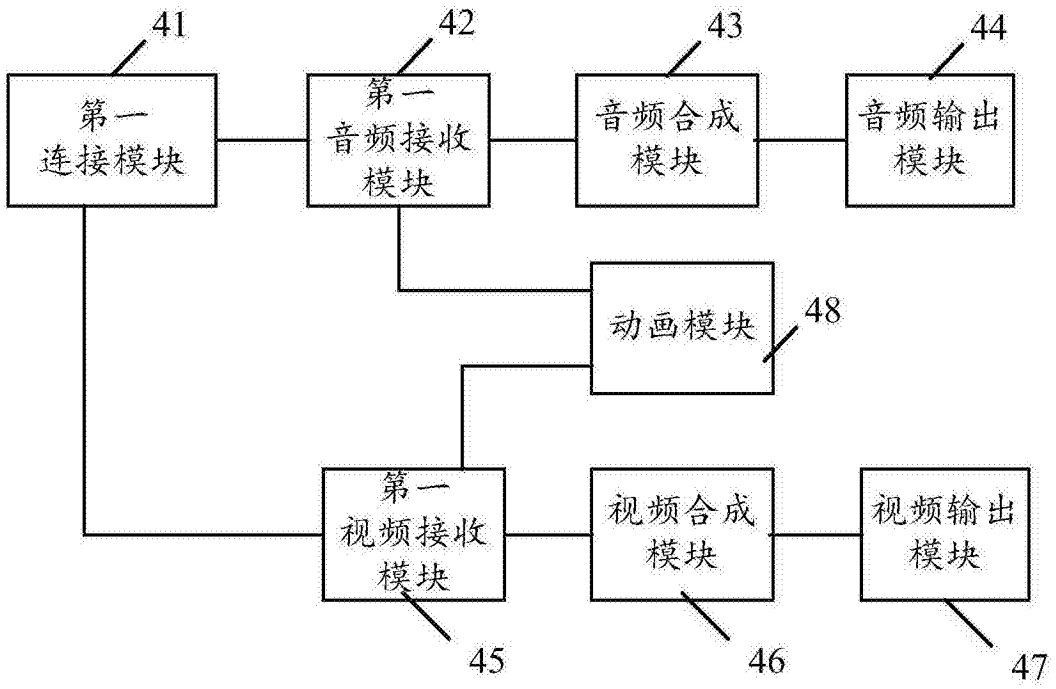


图5

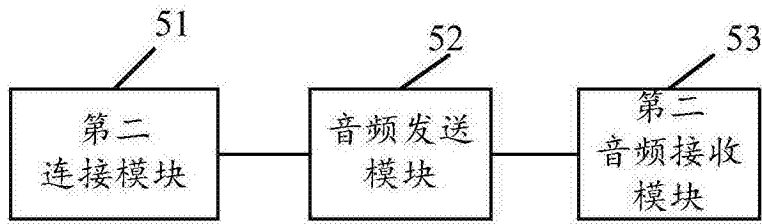


图6

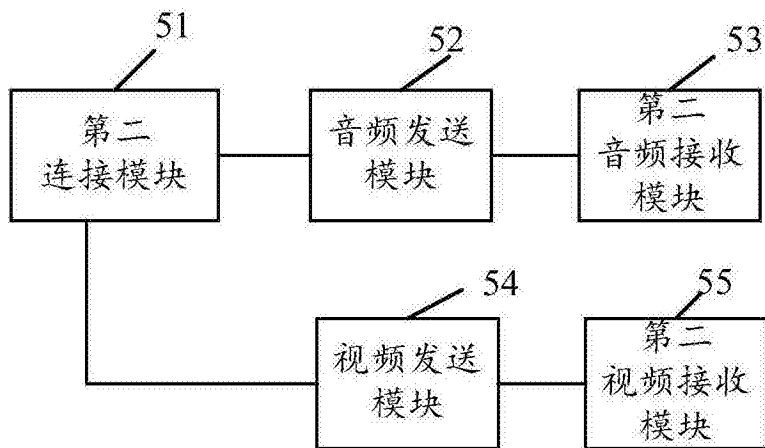


图7

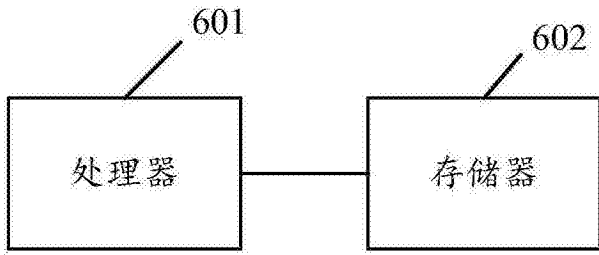


图8

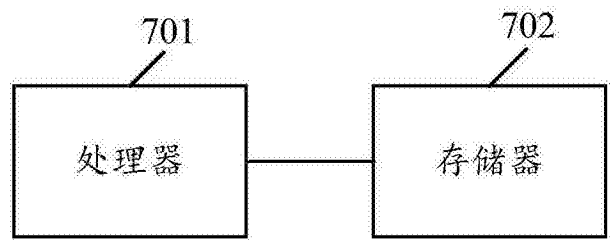


图9

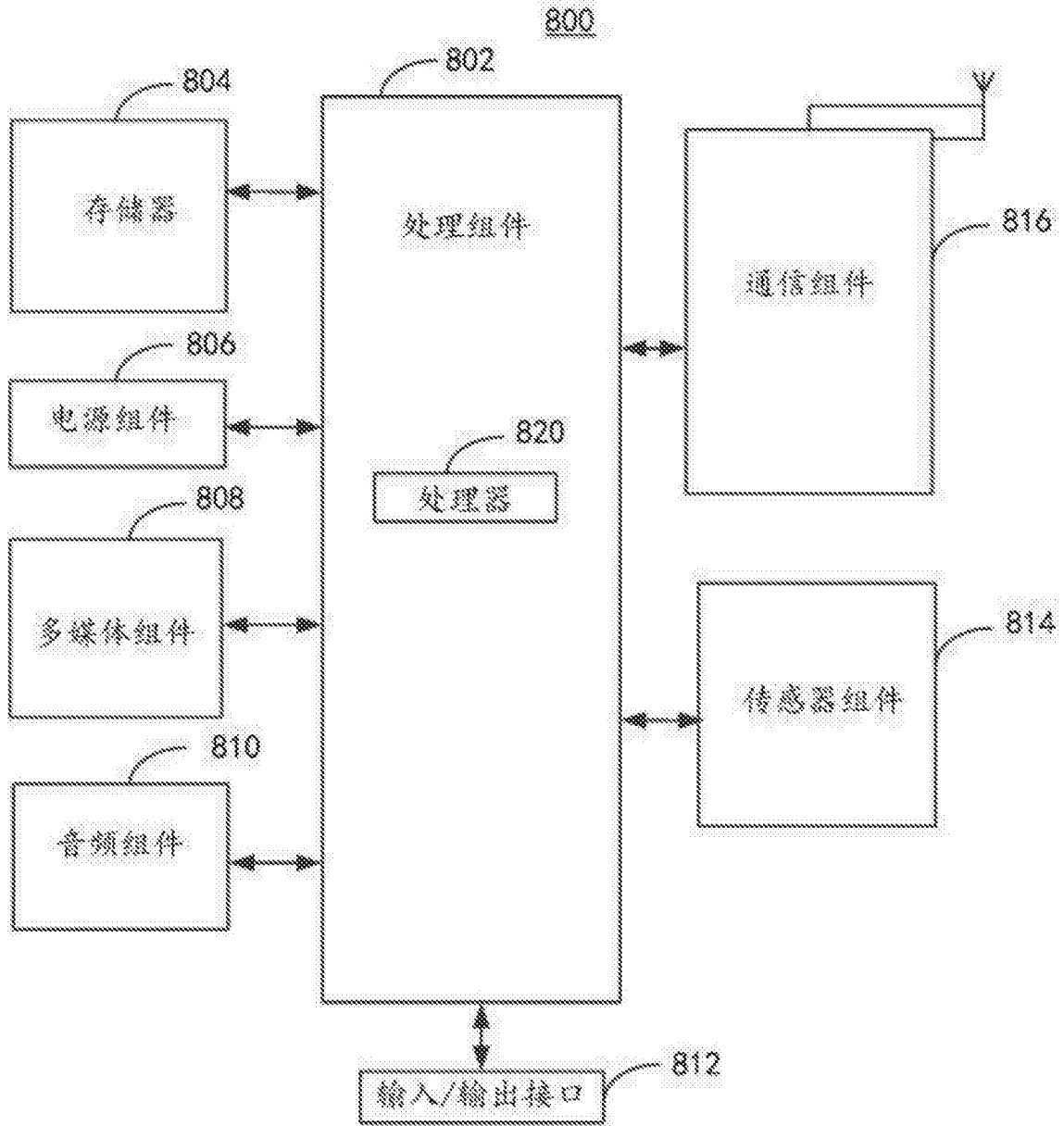


图10

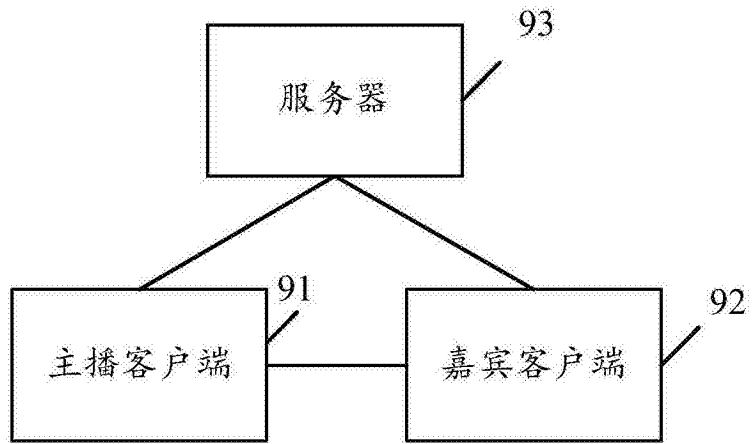


图11