



(12) 发明专利申请

(10) 申请公布号 CN 113473162 A

(43) 申请公布日 2021. 10. 01

(21) 申请号 202110368479.4

H04N 21/43 (2011.01)

(22) 申请日 2021.04.06

H04N 21/439 (2011.01)

(71) 申请人 北京沃东天骏信息技术有限公司

H04N 21/44 (2011.01)

地址 100176 北京市大兴区经济技术开发区

H04N 21/4402 (2011.01)

区科创十一街18号院2号楼4层A402室

H04N 21/4788 (2011.01)

申请人 北京京东世纪贸易有限公司

H04N 21/8547 (2011.01)

(72) 发明人 郑红阳

(74) 专利代理机构 北京派特恩知识产权代理有

限公司 11270

代理人 李昂 张颖玲

(51) Int. Cl.

H04N 21/2187 (2011.01)

H04N 21/233 (2011.01)

H04N 21/234 (2011.01)

H04N 21/2343 (2011.01)

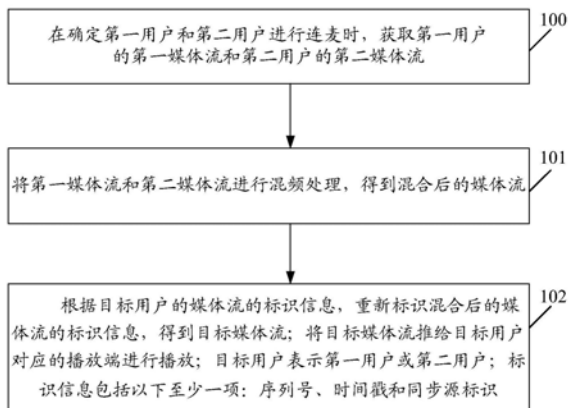
权利要求书2页 说明书11页 附图3页

(54) 发明名称

一种媒体流的播放方法、装置、设备和计算机存储介质

(57) 摘要

本申请实施例提出了一种媒体流的播放方法、装置、电子设备和计算机存储介质,该方法包括:在确定第一用户和第二用户进行连麦时,获取所述第一用户的第一媒体流和所述第二用户的第二媒体流;将所述第一媒体流和所述第二媒体流进行混频处理,得到混合后的媒体流;根据目标用户的媒体流的标识信息,重新标识所述混合后的媒体流的标识信息,得到目标媒体流;将所述目标媒体流推给所述目标用户对应的播放端进行播放;所述目标用户表示所述第一用户或所述第二用户;所述标识信息包括以下至少一项:序列号、时间戳和同步源标识。



1. 一种媒体流的播放方法,其特征在于,所述方法包括:

在确定第一用户和第二用户进行连麦时,获取所述第一用户的第一媒体流和所述第二用户的第二媒体流;

将所述第一媒体流和所述第二媒体流进行混频处理,得到混合后的媒体流;

根据目标用户的媒体流的标识信息,重新标识所述混合后的媒体流的标识信息,得到目标媒体流;将所述目标媒体流推给所述目标用户对应的播放端进行播放;所述目标用户表示所述第一用户或所述第二用户;所述标识信息包括以下至少一项:序列号、时间戳和同步源标识。

2. 根据权利要求1所述的方法,其特征在于,在重新标识所述混合后的媒体流的标识信息之前,所述方法还包括:

创建音频包队列、视频包队列和切换音频包队列;所述视频包队列用于放入混合后的媒体流中的视频包;所述切换音频包队列用于放入混合后的媒体流中的音频包;

在对齐所述视频包队列和所述切换音频包队列后,将所述切换音频包队列中的音频包转移到所述音频包队列。

3. 根据权利要求2所述的方法,其特征在于,所述对齐所述视频包队列和所述切换音频包队列,包括:

在所述视频包队列放入多个视频包、且所述切换音频包队列放入多个音频包后,根据所述多个视频包的时间戳和所述多个音频包的时间戳,确定所述多个视频包的网络时间协议NTP时间和多个音频包的NTP时间;

根据所述多个视频包的NTP时间和多个音频包的NTP时间,确定首次处于相同时刻的基准视频包的时间戳和基准音频包的时间戳;所述基准视频包为多个视频包中的其中一个视频包;所述基准音频包为多个音频包中的其中一个音频包;

基于所述基准视频包的时间戳和基准音频包的时间戳,对齐所述视频包队列和所述切换音频包队列。

4. 根据权利要求3所述的方法,其特征在于,所述方法还包括:

当定时时刻到来时,确定所述视频包队列和所述音频包队列在所述定时时刻对应的设定时间间隔内需要发送的视频包和音频包。

5. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

在确定所述第一用户和所述第二用户进行连麦时,将所述第一用户和所述第二用户之间的媒体流传输过程与所述目标媒体流的推送过程进行分离。

6. 一种媒体流的播放装置,其特征在于,所述装置包括:

获取模块,用于在确定第一用户和第二用户进行连麦时,获取所述第一用户的第一媒体流和所述第二用户的第二媒体流;

混频模块,用于将所述第一媒体流和所述第二媒体流进行混频处理,得到混合后的媒体流;

播放模块,用于根据目标用户的媒体流的标识信息,重新标识所述混合后的媒体流的标识信息,得到目标媒体流;将所述目标媒体流推给所述目标用户对应的播放端进行播放;所述目标用户表示所述第一用户或所述第二用户;所述标识信息包括以下至少一项:序列号、时间戳和同步源标识。

7. 根据权利要求6所述的装置,其特征在于,所述装置还包括同步模块,所述同步模块,在重新标识所述混合后的媒体流的标识信息之前,用于:

创建音频包队列、视频包队列和切换音频包队列;所述视频包队列用于放入混合后的媒体流中的视频包;所述切换音频包队列用于放入混合后的媒体流中的音频包;

在对齐所述视频包队列和所述切换音频包队列后,将所述切换音频包队列中的音频包转移到所述音频包队列。

8. 根据权利要求7所述的装置,其特征在于,所述同步模块,用于对齐所述视频包队列和所述切换音频包队列,包括:

在所述视频包队列放入多个视频包、且所述切换音频包队列放入多个音频包后,根据所述多个视频包的时间戳和所述多个音频包的时间戳,确定所述多个视频包的NTP时间和多个音频包的NTP时间;

根据所述多个视频包的NTP时间和多个音频包的NTP时间,确定首次处于相同时刻的基准视频包的时间戳和基准音频包的时间戳;所述基准视频包为多个视频包中的其中一个视频包;所述基准音频包为多个音频包中的其中一个音频包;

基于所述基准视频包的时间戳和基准音频包的时间戳,对齐所述视频包队列和所述切换音频包队列。

9. 一种电子设备,其特征在于,所述设备包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现权利要求1至5任一项所述的方法。

10. 一种计算机存储介质,其上存储有计算机程序,其特征在于,该计算机程序被处理器执行时实现权利要求1至5任一项所述的方法。

## 一种媒体流的播放方法、装置、设备和计算机存储介质

### 技术领域

[0001] 本申请涉及互联网技术领域,尤其涉及一种媒体流的播放方法、装置、电子设备和计算机存储介质。

### 背景技术

[0002] 近年来,各类用于视频直播的直播平台、直播软件层出不穷,视频直播可以给观看用户带来更实时的社交体验。在直播节目中,两个主播往往通过连麦增加节目效果。在连麦场景中,若两个主播之间进行连麦,往往需要从单个主播切换到两个主播合成的画面,这样两个主播的观看用户可以同时看到两个主播合成的画面;由于单主播的媒体流和双主播的媒体流的数据源不一样,因而,在连麦过程时,往往需要进行媒体流切换。

[0003] 相关技术中,往往采用将原始的单主播的媒体流直接切换到双主播的媒体流的方式进行连麦;这种方式虽然简单易实现,但是由于单主播的媒体流和双主播的媒体流一般不一致,因而,当播放端接收到断层的媒体流后,往往会出现画面中断、音画不同步等各方面的的问题。

### 发明内容

[0004] 本申请提供一种媒体流的播放方法、装置、电子设备和计算机存储介质。

[0005] 本申请的技术方案是这样实现的:

[0006] 本申请实施例提供了一种媒体流的播放方法,所述方法包括:

[0007] 在确定第一用户和第二用户进行连麦时,获取所述第一用户的第一媒体流和所述第二用户的第二媒体流;

[0008] 将所述第一媒体流和所述第二媒体流进行混频处理,得到混合后的媒体流;

[0009] 根据目标用户的媒体流的标识信息,重新标识所述混合后的媒体流的标识信息,得到目标媒体流;将所述目标媒体流推给所述目标用户对应的播放端进行播放;所述目标用户表示所述第一用户或所述第二用户;所述标识信息包括以下至少一项:序列号、时间戳和同步源标识。

[0010] 在一些实施例中,在重新标识所述混合后的媒体流的标识信息之前,所述方法还包括:

[0011] 创建音频包队列、视频包队列和切换音频包队列;所述视频包队列用于放入混合后的媒体流中的视频包;所述切换音频包队列用于放入混合后的媒体流中的音频包;

[0012] 在对齐所述视频包队列和所述切换音频包队列后,将所述切换音频包队列中的音频包转移到所述音频包队列。

[0013] 在一些实施例中,所述对齐所述视频包队列和所述切换音频包队列,包括:

[0014] 在所述视频包队列放入多个视频包、且所述切换音频包队列放入多个音频包后,根据所述多个视频包的时间戳和所述多个音频包的时间戳,确定所述多个视频包的网络时间协议(Network Time Protocol,NTP)时间和多个音频包的NTP时间;

[0015] 根据所述多个视频包的NTP时间和多个音频包的NTP时间,确定首次处于相同时刻的基准视频包的时间戳和基准音频包的时间戳;所述基准视频包为多个视频包中的其中一个视频包;所述基准音频包为多个音频包中的其中一个音频包;

[0016] 基于所述基准视频包的时间戳和基准音频包的时间戳,对齐所述视频包队列和所述切换音频包队列。

[0017] 在一些实施例中,所述方法还包括:

[0018] 当定时时刻到来时,确定所述视频包队列和所述音频包队列在所述定时时刻对应的设定时间间隔内需要发送的视频包和音频包。

[0019] 在一些实施例中,所述方法还包括:

[0020] 在确定所述第一用户和所述第二用户进行连麦时,将所述第一用户和所述第二用户之间的媒体流传输过程与所述目标媒体流的推送过程进行分离。

[0021] 本申请实施例还提出了一种媒体流的播放装置,所述装置包括获取模块、混频模块和播放模块,其中,

[0022] 获取模块,用于在确定第一用户和第二用户进行连麦时,获取所述第一用户的第一媒体流和所述第二用户的第二媒体流;

[0023] 混频模块,用于将所述第一媒体流和所述第二媒体流进行混频处理,得到混合后的媒体流;

[0024] 播放模块,用于根据目标用户的媒体流的标识信息,重新标识所述混合后的媒体流的标识信息,得到目标媒体流;将所述目标媒体流推给所述目标用户对应的播放端进行播放;所述目标用户表示所述第一用户或所述第二用户;所述标识信息包括以下至少一项:序列号、时间戳和同步源标识。

[0025] 本申请实施例提供一种电子设备,所述设备包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现前述一个或多个技术方案提供的媒体流的播放方法。

[0026] 本申请实施例提供一种计算机存储介质,所述计算机存储介质存储有计算机程序;所述计算机程序被执行后能够实现前述一个或多个技术方案提供的媒体流的播放方法。

[0027] 本申请实施例提出了一种媒体流的播放方法、装置、电子设备和计算机存储介质,该方法包括:在确定第一用户和第二用户进行连麦时,获取所述第一用户的第一媒体流和所述第二用户的第二媒体流;将所述第一媒体流和所述第二媒体流进行混频处理,得到混合后的媒体流;根据目标用户的媒体流的标识信息,重新标识所述混合后的媒体流的标识信息,得到目标媒体流;将所述目标媒体流推给所述目标用户对应的播放端进行播放;所述目标用户表示所述第一用户或所述第二用户;所述标识信息包括以下至少一项:序列号、时间戳和同步源标识;如此,本申请实施例在确定两个用户进行连麦时,基于单个用户的媒体流的序列号和时间戳,重新标识两个用户混合后的媒体流的序列号和时间戳,能够解决因单个用户的媒体流与混合后的媒体流的序列号和时间戳不连续而造成的播放端音画不同步的问题;基于单个用户的媒体流的同步源标识,重新标识两个用户混合后的媒体流的同步源标识,能够解决因为单个用户的媒体流与混合后的媒体流的同步源标识不一致而造成的播放端画面中断的问题。

## 附图说明

- [0028] 图1a是本申请实施例中的一种媒体流的播放方法的流程示意图；
- [0029] 图1b是本申请实施例中的一种传输媒体流的结构示意图；
- [0030] 图1c是本申请实施例通过三个队列实现连麦的结构示意图；
- [0031] 图1d是本申请实施例中对视频包和音频包进行同步的结构示意图；
- [0032] 图2是本申请实施例中的另一种媒体流的播放方法的结构示意图；
- [0033] 图3为本申请实施例的媒体流的播放装置的组成结构示意图；
- [0034] 图4为本申请实施例提供的电子设备的结构示意图。

## 具体实施方式

[0035] 以下结合附图及实施例,对本申请进行进一步详细说明。应当理解,此处所提供的实施例仅仅用以解释本申请,并不用于限定本申请。另外,以下所提供的实施例是用于实施本申请的部分实施例,而非提供实施本申请的全部实施例,在不冲突的情况下,本申请实施例记载的技术方案可以任意组合的方式实施。

[0036] 需要说明的是,在本申请实施例中,术语“包括”、“包含”或者其任何其它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方法或者装置不仅包括所明确记载的要素,而且还包括没有明确列出的其它要素,或者是还包括为实施方法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方法或者装置中还存在另外的相关要素(例如方法中的步骤或者装置中的单元,例如的单元可以是部分电路、部分处理器、部分程序或软件等等)。

[0037] 本文中术语“和/或”,仅仅是一种描述关联对象的关联关系,表示可以存在三种关系,例如,I和/或J,可以表示:单独存在I,同时存在I和J,单独存在J这三种情况。另外,本文中术语“至少一种”表示多种中的任意一种或多种中的至少两种的任意组合,例如,包括I、J、R中的至少一种,可以表示包括从I、J和R构成的集合中选择的任意一个或多个元素。

[0038] 例如,本申请实施例提供的媒体流的播放方法包含了一系列的步骤,但是本申请实施例提供的媒体流的播放方法不限于所记载的步骤,同样地,本申请实施例提供的媒体流的播放装置包括了一系列模块,但是本申请实施例提供的媒体流的播放装置不限于包括所明确记载的模块,还可以包括为获取相关时序数据、或基于时序数据进行处理时所设置设置的模块。

[0039] 本申请实施例可以应用于终端设备和服务器集群组成的计算机系统中,服务器集群包括至少一个服务器,服务器与终端设备之间可以进行交互,并可以与众多其它通用或专用计算系统环境或配置一起操作。这里,终端设备可以是瘦客户机、厚客户机、手持或膝上设备、基于微处理器的系统、机顶盒、可编程消费电子产品、网络个人电脑、小型计算机系统,等等,服务器可以是服务器计算机系统小型计算机系统、大型计算机系统和包括上述任何系统的分布式云计算技术环境,等等。

[0040] 终端设备、服务器等电子设备可以在由计算机系统执行的计算机系统可执行指令(诸如程序模块)的一般语境下描述。通常,程序模块可以包括例程、程序、目标程序、组件、逻辑、数据结构等等,它们执行特定的任务或者实现特定的抽象数据类型。计算机系统/服务器可以在分布式云计算环境中实施,分布式云计算环境中,任务是由通过通信网络连接

的远程处理设备执行的。在分布式云计算环境中,程序模块可以位于包括存储设备的本地或远程计算系统存储介质上。

[0041] 在本申请的一些实施例中,媒体流的播放方法可以利用服务器集群中的处理器实现,上述处理器可以为特定用途集成电路(Application Specific Integrated Circuit,ASIC)、数字信号处理器(Digital Signal Processor,DSP)、数字信号处理装置(Digital Signal Processing Device,DSPD)、可编程逻辑装置(Programmable Logic Device,PLD)、现场可编程逻辑门阵列(Field Programmable Gate Array,FPGA)、中央处理器(Central Processing Unit,CPU)、控制器、微控制器、微处理器中的至少一种。

[0042] 图1a是本申请实施例中的一种媒体流的播放方法的流程示意图,如图1a所示,该方法包括如下步骤:

[0043] 步骤100:在确定第一用户和第二用户进行连麦时,获取第一用户的第一媒体流和第二用户的第二媒体流。

[0044] 示例性地,第一用户、第二用户可以表示主播或其它具有连麦需求的用户,其中,主播表示通过直播平台或直播软件进行直播的用户。这里,对于主播的类型本申请实施例不作限制;例如,可以为秀场主播、游戏主播等。下面以第一用户和第二用户为主播为例进行说明。

[0045] 示例性地,第一媒体流可以是第一用户在直播过程中产生的媒体流;第二媒体流可以是第二用户在直播过程中产生的媒体流;其中,第一媒体流和第二媒体流可以表示既包含音频包,又包含视频包的媒体流,也可以表示仅包含音频包的媒体流,还可以表示仅包含视频包的媒体流。

[0046] 示例性地,在第一媒体流和第二媒体流既包含音频包,又包含视频包的情况下,第一用户和第二用户进行连麦可以实现音频数据和视频数据的交互;在第一媒体流和第二媒体流仅包含音频包的情况下,第一用户和第二用户进行连麦可以实现音频数据的交互;在第一媒体流和第二媒体流仅包含视频包的情况下,第一用户和第二用户进行连麦可以实现视频数据的交互。

[0047] 在一种实施方式中,可以通过摄像头或其它带有视频采集功能的设备采集主播在直播过程中产生的视频包;可以通过麦克风或其它带有语音采集功能的设备采集主播在直播过程中产生的音频包;这样,主播在直播过程中产生的视频包和音频包可以形成对应的媒体流。

[0048] 示例性地,当第一用户向第二用户发送连麦请求,且第二用户同意该连麦请求时,确定第一用户和第二用户进行连麦;反之,当第一用户向第二用户发送连麦请求,且第二用户拒绝该连麦请求时,确定第一用户和第二用户未进行连麦。这里,也可以是第二用户向第一用户发送连麦请求,根据第一用户同意或拒绝该连麦请求的结果,确定第一用户和第二用户是否进行连麦;在确定第一用户和第二用户进行连麦时,可以通过服务器集群中的媒体服务器获取第一用户的第一媒体流和第二用户的第二媒体流。

[0049] 步骤101:将第一媒体流和第二媒体流进行混频处理,得到混合后的媒体流。

[0050] 示例性地,在确定第一用户和第二用户进行连麦时,媒体服务器通过会话初始协议(session initiation protocol,sip)将获取到的第一媒体流和第二媒体流发送给服务器集群中的混频服务器;进而,通过混频服务器对第一媒体流和第二媒体流进行混频处理,

可以将第一媒体流和第二媒体流中的音频包和视频包进行合成,得到混合后的媒体流。这里,sip是一种基于文本的应用层控制协议,用于创建、修改和释放一个或多个参与者的会话,为多种即时通信业务提供完整的会话创建和会话更改服务。

[0051] 在一种实施方式中,混频服务器的混频处理可以包括:画面合成、回声消除、降噪和混音等;示例性地,混频服务器可以为多点控制单元(Multi Control Unit,MCU)混频服务器,也可以为其它具有混频功能的服务器,本申请实施例对此不作限制。

[0052] 示例性地,MCU混频服务器将第一媒体流和第二媒体流中的信息流经过同步分离后,抽取出音频、视频、数据、信令等各种信息,并将相同信息送入相应信息处理模块,完成相应信息的处理,比如:音频包的混合、视频包的混合、信令控制等。

[0053] 步骤102:根据目标用户的媒体流的标识信息,重新标识混合后的媒体流的标识信息,得到目标媒体流;将目标媒体流推给目标用户对应的播放端进行播放;目标用户表示第一用户或第二用户;标识信息包括以下至少一项:序列号、时间戳和同步源标识。

[0054] 示例性地,在得到混合后的媒体流后,混频服务器通过sip信令将混合后的媒体流推给服务器集群中的推流服务端;其中,sip信令表示描述开始播放、停止播放和快进播放等命令的信号。推流服务端根据目标用户的媒体流的标识信息,重新标识混合后的媒体流的标识信息,得到目标媒体流;将目标媒体流发送给多媒体视频处理工具ffmpeg(fast forward mpeg),通过ffmpeg对目标媒体流进行编码、封装后,转推给内容分发网络(Content Delivery Network,CDN)再到目标用户对应的播放端接收,解码和播放。

[0055] 这里,播放端可以是位于直播页面中的Flash播放器或直播插件等,用于接收该目标媒体流,并在解码后播放该目标媒体流,以使观看用户能够看到目标用户的直播。

[0056] 在一种实施方式中,标识信息可以包括媒体流中音频包对应的实时传输协议(realtime transport protocol,rtp)包的序列号、时间戳和同步源标识(synchronization source,ssrc)中至少一项;还可以包括媒体流中视频包对应的rtp包的序列号、时间戳、ssrc中至少一项。

[0057] 这里,序列号在rtp包占7位,用于标识发送者所发送的rtp报文的序列号,每发送一个报文,序列号增1;序列号的初始值是随机的,且音频包和视频包对应的序列号是分开计数的。时间戳在rtp包占32位,其反映了rtp报文的第一个八位组的采样时刻;接收者使用时间戳可以计算延迟和延迟抖动,并进行同步控制。ssrc在rtp包占32位,用于标识同步信源,该标识符是随机选择的,同一用户的媒体流具有相同的ssrc。

[0058] 图1b是本申请实施例中的一种传输媒体流的结构示意图,如图1b所示,Soup-worker和nodejs组成媒体服务器;PushService表示推流服务端;这里,nodejs负责sip信令的解析发送;Soup-worker负责视频包和音频包对应的rtp包的发送。

[0059] 由图1b可以看出,在第一用户和第二用户未进行连麦的情况下,第一用户或第二用户的媒体流直接通过媒体服务器,转推给推流服务端PushService,推流服务端PushService发送给ffmpeg,进而转推给CDN。在第一用户和第二用户进行连麦的情况下,第一用户和第二用户的媒体流通过媒体服务器分流后,推给MCU混频服务器进行混频处理,得到混合后的媒体流;然后将混合后的媒体流发给推流服务端PushService,在重新标识混合后的媒体流的标识信息后,发送给ffmpeg,进而转推给CDN。

[0060] 示例性地,由于第一媒体流和第二媒体流的数据源分别来自于第一用户和第二用

户;而混合后的媒体流的数据源来自于混频服务器;可见,第一媒体流、第二媒体流和混合后的媒体流的ssrc均不一致;因而,基于单个用户的媒体流的同步源标识,重新标识两个用户混合后的媒体流的同步源标识,能够解决因为单个用户的媒体流与混合后的媒体流的同步源标识不一致而造成的播放端画面中断的问题;进一步地,基于单个用户的媒体流的序列号和时间戳,重新标识两个用户混合后的媒体流的序列号和时间戳,能够解决因单个用户的媒体流与混合后的媒体流的序列号和时间戳不连续而造成的播放端音画不同步的问题。

[0061] 在一些实施例中,在重新标识混合后的媒体流的标识信息之前,该方法还可以包括:创建音频包队列、视频包队列和切换音频包队列;视频包队列用于放入混合后的媒体流中的视频包;切换音频包队列用于放入混合后的媒体流中的音频包;在对齐视频包队列和切换音频包队列后,将切换音频包队列中的音频包转移到音频包队列。

[0062] 本申请实施例中,在推流服务端创建音频包队列、视频包队列和切换音频包队列这三个队列;图1c是本申请实施例中通过三个队列实现连麦的结构示意图,如图1c所示,在第一用户和第二用户的连麦过程中,当收包线程接收到混合后的媒体流中的音频包时,将其放入切换音频包队列;当收包线程接收到混合后的媒体流中的视频包时,将其放入视频包队列;将视频包队列和切换音频包队列进行对齐,并在对齐视频包队列和切换音频包队列后,将切换音频包队列中的音频包转移到音频包队列;通过定时器的定时时刻触发任务线程池,当定时时刻到来时,通过任务线程池中的线程将视频包队列和音频包队列对应的音视频包发送出去,以此实现第一用户和第二用户的连麦。

[0063] 这里;通过对齐视频包队列和切换音频包队列,可确保两个用户在连麦过程中的音视频同步。

[0064] 在一些实施例中,对齐视频包队列和切换音频包队列,可以包括:在视频包队列放入多个视频包、且切换音频包队列放入多个音频包后,根据多个视频包的时间戳和多个音频包的时间戳,确定多个视频包的NTP时间和多个音频包的NTP时间;根据多个视频包的NTP时间和多个音频包的NTP时间,确定首次处于相同时刻的基准视频包的时间戳和基准音频包的时间戳;基准视频包为多个视频包中的其中一个视频包;基准音频包为多个音频包中的其中一个音频包;基于基准视频包的时间戳和基准音频包的时间戳,对齐视频包队列和切换音频包队列。

[0065] 示例性地,对于多个视频包中的任意一个视频包,该视频包的时间戳与该视频包的NTP时间表示不同时间轴上的相同时间点;同样地,对于多个音频包中的任意一个音频包,该音频包的时间戳与该音频包的NTP时间表示不同时间轴上的相同时间点。

[0066] 在一种实施方式中,视频包的时间戳对应的时间轴为视频时间轴;音频包的时间戳对应的时间轴为音频时间轴;NTP时间对应的时间轴为NTP时间轴。由于视频包的时间戳和音频包的时间戳是分开计数,且视频时间轴与音频时间轴的刻度也不相同;因而,在将视频包队列中的多个视频包与切换音频包队列中的多个音频包进行对齐的过程中,可将多个视频包的时间戳和多个音频包的时间戳映射到统一的NTP时间轴,得到多个视频包的NTP时间和多个音频包的NTP时间。

[0067] 示例性地,在得到多个视频包的NTP时间和多个音频包的NTP时间后,可以确定首次处于相同时刻(同一NTP时间)的基准视频包和基准音频包;再根据该相同时刻确定基准

视频包的时间戳和基准音频包的时间戳;进而,通过基准视频包的时间戳和基准音频包的时间戳,对齐视频包队列和切换音频包队列。

[0068] 在一种实施方式中,当视频包队列中放入5个视频包,且切换音频包队列中放入10个音频包时,假设根据5个视频包的NTP时间和10个音频包的NTP时间,确定首次处于相同时刻的是第3个视频包和第5个音频包;这里,第3个视频包为基准视频包,第5个音频包为基准音频包;此时,将前两个视频包和前四个音频包进行删除,即,根据第3个视频包的时间戳和第5个音频包的时间戳,对齐视频包队列和切换音频包队列。

[0069] 示例性地,在对齐视频包队列和切换音频包队列后,将切换音频包队列中的音频包转移到音频包队列;此时,视频包队列与音频包队列对齐,即,视频包队列中的视频包与音频包队列中的音频包是同步的。在视频包队列与音频包队列对齐后,当再次接收到混合后的媒体流中的音频包时,可以直接将该音频包放入音频包队列中。

[0070] 示例性地,可以预先通过定时器确定每个定时时刻,这里,相邻定时时刻的时间间隔可以根据实际情况进行设置,例如,可以为0.01S,0.02S等,本申请实施例对此不作限制。

[0071] 在一种实施方式中,可以在视频包队列与音频包队列中缓存设定时长的视频包和音频包;这样,每当定时时刻到来时,从设定时长的视频包和音频包中确定在该定时时刻对应的设定时间间隔内需要发送的视频包和音频包。这里,设定时长可以为2S,2.5S等。

[0072] 示例性地,在设定时长为2S、且相邻定时时刻的时间间隔为0.01S的情况下,当定时时刻到来时,从2S的视频包和音频包中确定该定时时刻对应的0.01S内需要发送的视频包和音频包。

[0073] 在一种实施方式中,在确定视频包队列和音频包队列在定时时刻对应的设定时间间隔内需要发送的视频包和音频包后,重新标识这些视频包和音频包的标识信息,能够保证直播连麦过程中画面的连续和同步。

[0074] 图1d是本申请实施例中对视频包和音频包进行同步的结构示意图,如图1d所示,Base audio ts为基准音频时间,用来维护表示首次处于相同时刻的第一个音频包的时间戳;Base video ts为基准视频时间,用来维护表示首次处于相同时刻的第一个视频包的时间戳;Base ntp为基准绝对ntp时间,用来维护首次处于相同时刻的绝对ntp时间;这里,Base ntp与第一个音频包的ntp时间以及一个视频包的ntp时间相同。Curr audio ts为当前音频时间,用来维护当前时刻的音频时间戳;Curr video ts为当前视频时间,用来维护当前时刻的视频时间戳。

[0075] 示例性地,根据Base audio ts和Base video ts,可以确定在第一个定时时刻对应的设定时间间隔内需要发送的视频包和音频包;在发送完成后,根据Curr audio ts和Curr video ts确定第二个定时时刻对应的设定时间间隔内需要发送的视频包和音频包。

[0076] 示例性地,可以预先在视频包队列与音频包队列中缓存2S的视频包和音频包;在相邻定时时刻的时间间隔为0.01S的情况下,假设Base ntp为12:00,根据Base audio ts和Base video ts,可以确定在ntp时间12:00至12:01这个时间间隔内需要发送的视频包和音频包,重新标识这些视频包和音频包的标识信息并进行发送;在发送完成后,将ntp时间12:01对应的音频时间戳作为Curr audio ts,将12:01对应的视频时间戳作为Curr video ts,再确定在ntp时间12:01至12:02这个时间间隔内需要发送的视频包和音频包,重新标识这些视频包和音频包的标识信息并进行发送,以此类推,直到将视频包队列的视频包与音频

包队列中音频包发送完毕。

[0077] 在一些实施例中,该方法还可以包括:在确定第一用户和第二用户进行连麦时,将第一用户和第二用户之间的媒体流传输过程与目标媒体流的推送过程进行分离。

[0078] 这里,可以将第一用户和第二用户之间的媒体流传输过程称为P2P (Peer to peer) 模式;目标媒体流的推送过程表示将目标用户的媒体流向对应的播放端的观看用户进行推送。

[0079] 本申请实施例中,可以将直播连麦中的p2p模式和目标媒体流的推送进行分离;也就是说,将第一用户和第二用户之间的媒体流交互过程与两个用户向播放端推送媒体流的过程进行分离;即,将负责这两部分的直播代码进行解耦,这样,不仅可以确保直播代码的快速部署和快速升级;还可以降低系统的维护成本,提高系统的稳定性。

[0080] 本申请实施例提出了一种媒体流的播放方法、装置、电子设备和计算机存储介质,该方法包括:在确定第一用户和第二用户进行连麦时,获取第一用户的第一媒体流和第二用户的第二媒体流;将第一媒体流和第二媒体流进行混频处理,得到混合后的媒体流;根据目标用户的媒体流的标识信息,重新标识混合后的媒体流的标识信息,得到目标媒体流;将目标媒体流推给目标用户对应的播放端进行播放;目标用户表示第一用户或第二用户;标识信息包括以下至少一项:序列号、时间戳和同步源标识;如此,该方法在确定两个用户进行连麦时,基于单个用户的媒体流的序列号和时间戳,重新标识两个用户混合后的媒体流的序列号和时间戳,能够解决因单个用户的媒体流与混合后的媒体流的序列号和时间戳不连续而造成的播放端音画不同步的问题;基于单个用户的媒体流的同步源标识,重新标识两个用户混合后的媒体流的同步源标识,能够解决因为单个用户的媒体流与混合后的媒体流的同步源标识不一致而造成的播放端画面中断的问题;进而,可以提高直播效果。

[0081] 为了能够更加体现本申请的目的,在本申请上述实施例的基础上,进行进一步的举例说明。

[0082] 图2是本申请实施例中的一种媒体流的播放方法的结构示意图,如图2所示,第一用户端C1推送第一媒体流到媒体服务器soup,第二用户端C2推送第二媒体流到媒体服务器soup;媒体服务器soup进行分流操作,在第一用户C1和第二用户C2未进行连麦的情况下,媒体服务器soup将第一媒体流直接发送给推流服务端PushService,推流服务端PushService发送给ffmpeg,进而转推给CDN;同样地,媒体服务器soup将第二媒体流直接发送给推流服务端PushService,推流服务端PushService发送给ffmpeg,进而转推给CDN。在第一用户C1和第二用户C2进行连麦的情况下,媒体服务器soup通过sip协议,控制第一媒体流发送给MCU混频服务器中的第一客户端soupClient1,控制第二媒体流发送给MCU混频服务器中的第二客户端soupClient2,这里,第一客户端用于接收第一用户的第一媒体流,第二客户端soupClient2用于接收第二用户的第二媒体流;通过MCU处理器对第一客户端soupClient1中的第一媒体流和第二客户端soupClient2中第二媒体流进行混频处理,得到混合后的媒体流;通过sip协议将混合后的媒体流转推给推流服务端PushService,在推流服务端PushService重新标识混合后的媒体流的标识信息后,发送给ffmpeg,进而转推给CDN。

[0083] 这里,第一信令服务器sipClient负责sip信令的解析和发送;第二信令服务器rtspServer表示使用实时流传输协议(real time streaming protocol,rtsp)的媒体服务端,ffmpeg通过第二信令服务器对媒体流进行编码和封装;描述会话的协议(Session

Description Protocol,SDP) 用于对媒体流的收发编码及端口信息进行描述。

[0084] CDN通过流媒体服务器(Simple RTMP Server,SRS),将重新标识的媒体流推给CDN的边缘节点,也就是转发给各地的本地服务器,这样,方便观看用户就近获取媒体流进行观看。

[0085] 相关技术中,推流服务端PushService在媒体服务器soup中集成;而本申请实施例中,由图2可以看出,将推流服务端PushService和媒体服务器soup这两个模块进行解耦,这样,第一用户端C1和第二用户C2之间仅通过媒体服务器soup便可进行媒体流的交互,即,实现P2P模式;第一用户端C1和第二用户C2的推流过程通过推流服务端PushService进行实现。进而,能够确保直播代码的快速部署和快速升级。

[0086] 图3为本申请实施例的媒体流的播放装置的组成结构示意图,如图3所示,该装置包括:获取模块300、混频模块301和播放模块302,其中:

[0087] 获取模块300,用于在确定第一用户和第二用户进行连麦时,获取第一用户的第一媒体流和第二用户的第二媒体流;

[0088] 混频模块301,用于将第一媒体流和第二媒体流进行混频处理,得到混合后的媒体流;

[0089] 播放模块302,用于根据目标用户的媒体流的标识信息,重新标识混合后的媒体流的标识信息,得到目标媒体流;将目标媒体流推给目标用户对应的播放端进行播放;目标用户表示第一用户或第二用户;标识信息包括以下至少一项:序列号、时间戳和同步源标识。

[0090] 在一些实施例中,该装置还包括同步模块303,同步模块303,在重新标识混合后的媒体流的标识信息之前,用于:

[0091] 创建音频包队列、视频包队列和切换音频包队列;视频包队列用于放入混合后的媒体流中的视频包;切换音频包队列用于放入混合后的媒体流中的音频包;

[0092] 在对齐视频包队列和切换音频包队列后,将切换音频包队列中的音频包转移到音频包队列。

[0093] 在一些实施例中,同步模块303,用于对齐视频包队列和切换音频包队列,包括:

[0094] 在视频包队列放入多个视频包、且切换音频包队列放入多个音频包后,根据多个视频包的时间戳和多个音频包的时间戳,确定多个视频包的NTP时间和多个音频包的NTP时间;

[0095] 根据多个视频包的NTP时间和多个音频包的NTP时间,确定首次处于相同时刻的基准视频包的时间戳和基准音频包的时间戳;基准视频包为多个视频包中的其中一个视频包;基准音频包为多个音频包中的其中一个音频包;

[0096] 基于基准视频包的时间戳和基准音频包的时间戳,对齐视频包队列和切换音频包队列。

[0097] 在一些实施例中,同步模块303,还用于:

[0098] 当定时时刻到来时,确定视频包队列和音频包队列在定时时刻对应的设定时间间隔内需要发送的视频包和音频包。

[0099] 在一些实施例中,播放模块302,还用于:

[0100] 在确定第一用户和第二用户进行连麦时,将第一用户和第二用户之间的媒体流传输过程与目标媒体流的推送过程进行分离。

[0101] 在实际应用中,上述获取模块300、混频模块301、播放模块302和同步模块303均可以由位于电子设备中的处理器实现,该处理器可以为ASIC、DSP、DSPD、PLD、FPGA、CPU、控制器、微控制器、微处理器中的至少一种。

[0102] 另外,在本实施例中的各功能模块可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能模块的形式实现。

[0103] 集成的单元如果以软件功能模块的形式实现并非作为独立的产品进行销售或使用,可以存储在一个计算机可读取存储介质中,基于这样的理解,本实施例的技术方案本质上或者说对相关技术做出贡献的部分或者该技术方案的全部或部分可以以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中,包括若干指令用以使得一台计算机设备(可以是个人计算机、服务器、或者网络设备等)或processor(处理器)执行本实施例方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(Read Only Memory,ROM)、随机存取存储器(Random Access Memory,RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0104] 具体来讲,本实施例中的一种媒体流的播放方法对应的计算机程序指令可以被存储在光盘、硬盘、U盘等存储介质上,当存储介质中的与一种媒体流的播放方法对应的计算机程序指令被一电子设备读取或被执行时,实现前述实施例的任意一种媒体流的播放方法。

[0105] 基于前述实施例相同的技术构思,参见图4,其示出了本申请实施例提供的电子设备400,可以包括:存储器401和处理器402;其中,

[0106] 存储器401,用于存储计算机程序和数据;

[0107] 处理器402,用于执行存储器中存储的计算机程序,以实现前述实施例的任意一种媒体流的播放方法。

[0108] 在实际应用中,上述存储器401可以是易失性存储器(volatile memory),例如RAM;或者非易失性存储器(non-volatile memory),例如ROM、快闪存储器(flash memory)、硬盘(Hard Disk Drive,HDD)或固态硬盘(Solid-State Drive,SSD);或者上述种类的存储器的组合,并向处理器402提供指令和数据。

[0109] 上述处理器402可以为ASIC、DSP、DSPD、PLD、FPGA、CPU、控制器、微控制器、微处理器中的至少一种。可以理解地,对于不同的媒体流的播放设备,用于实现上述处理器功能的电子器件还可以为其它,本申请实施例不作具体限定。

[0110] 在一些实施例中,本申请实施例提供的装置具有的功能或包含的模块可以用于执行上文方法实施例描述的方法,其具体实现可以参照上文方法实施例的描述,为了简洁,这里不再赘述。

[0111] 上文对各个实施例的描述倾向于强调各个实施例之间的不同之处,其相同或相似之处可以互相参考,为了简洁,本文不再赘述。

[0112] 本申请所提供的各方法实施例中所揭露的方法,在不冲突的情况下可以任意组合,得到新的方法实施例。

[0113] 本申请所提供的各产品实施例中所揭露的特征,在不冲突的情况下可以任意组合,得到新的产品实施例。

[0114] 本申请所提供的各方法或设备实施例中所揭露的特征,在不冲突的情况下可以任意组合,得到新的方法实施例或设备实施例。

[0115] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用硬件实施例、软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器和光学存储器等)上实施的计算机程序产品的形式。

[0116] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其它可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其它可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0117] 这些计算机程序指令也可装载到计算机或其它可编程数据处理设备上,使得在计算机或其它可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其它可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0118] 以上,仅为本申请的较佳实施例而已,并非用于限定本申请的保护范围。

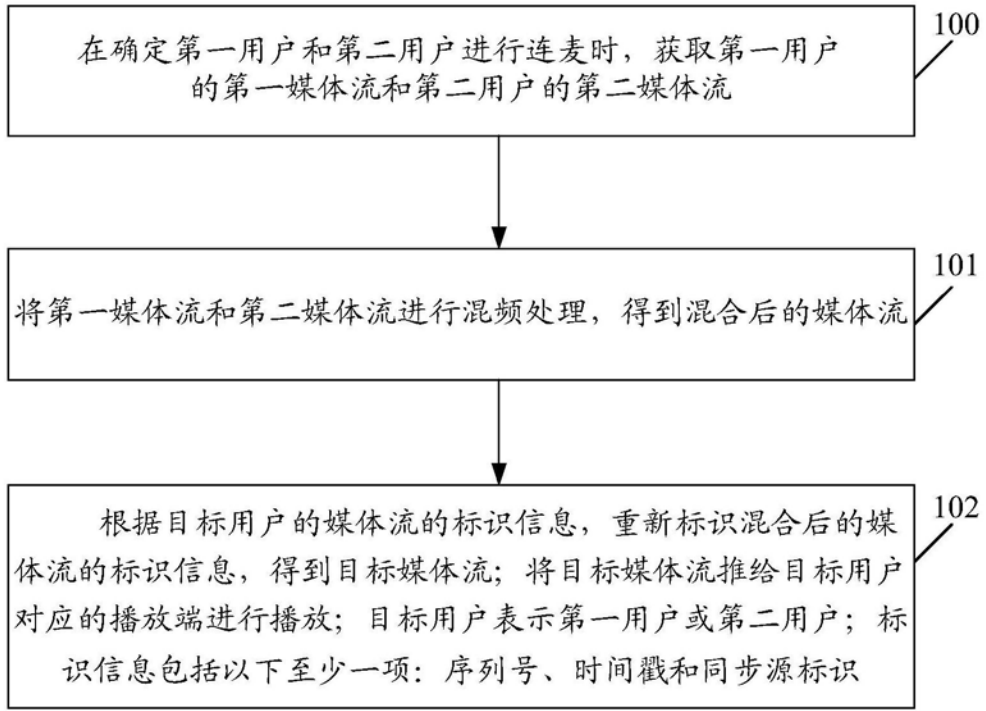


图1a

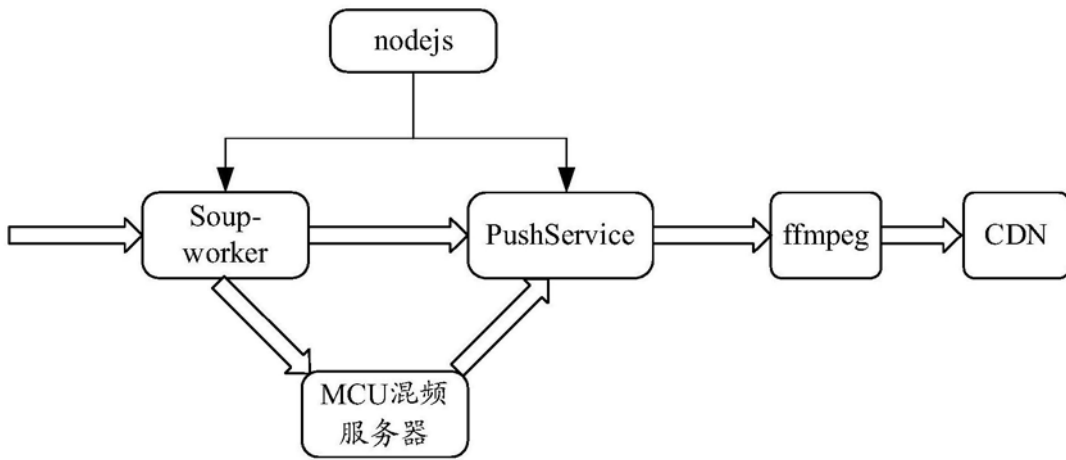


图1b

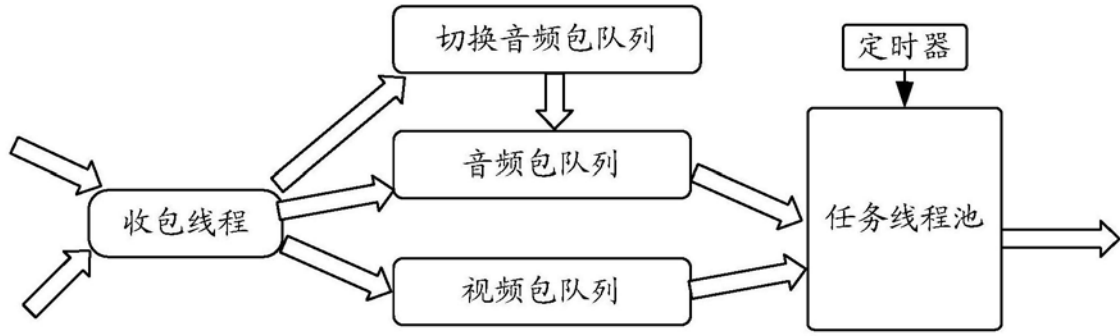


图1c

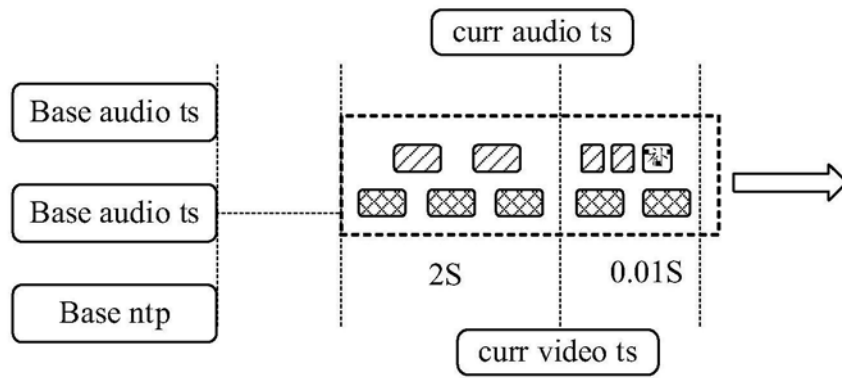


图1d

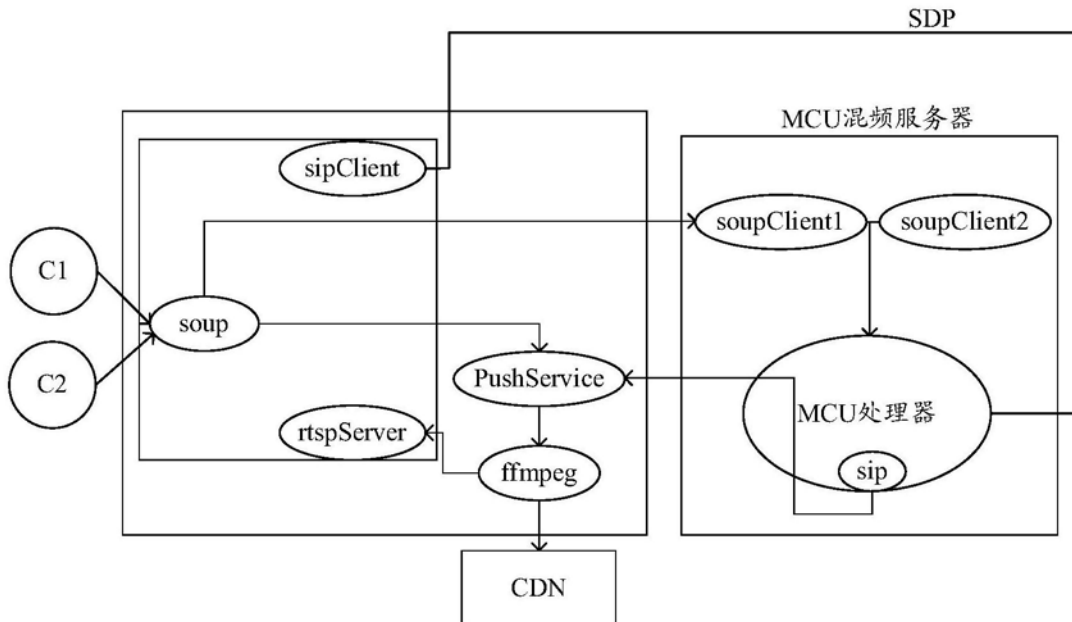


图2

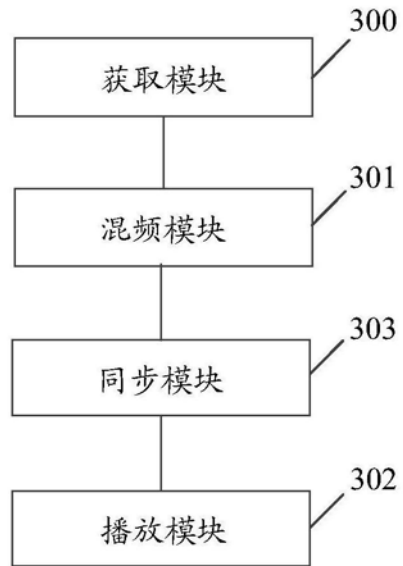


图3

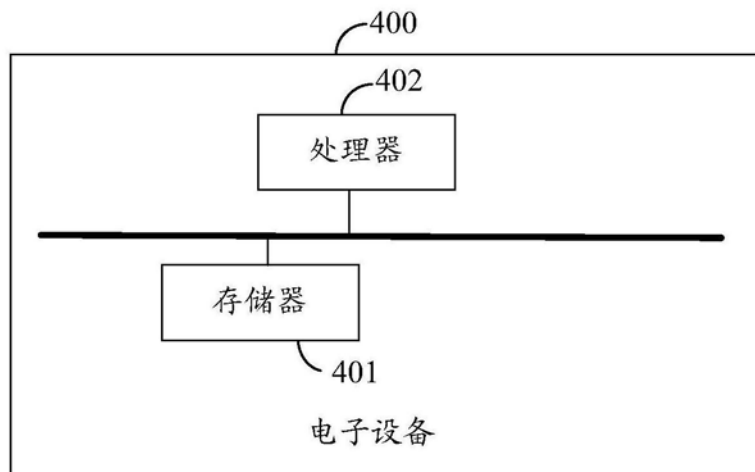


图4