

一种线下线上实时合唱方法、设备及介质

本申请要求于 2023 年 11 月 27 日提交中国专利局、申请号为 202311602552.5、发明名称为“一种线下线上实时合唱方法、设备及介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5 技术领域

本发明涉及通信技术领域，特别涉及一种线下线上实时合唱方法、设备及介质。

背景技术

- 10 当前，市面上的某些合唱应用软件可以实现用户在线参与合唱，也即当前参与合唱的用户仅限于线上的方式。而对于线下演唱场景，如明星演唱会、推流演唱会等，演唱者和台下观众可以线下参与合唱，而线上用户仅可以通过终端设备实时观看线下演唱会，因此还缺少一种实现线上用户和线下用户实时合唱的方案。
- 15 综上，如何实现线上用户和线下用户实时合唱是目前有待解决的问题。

发明内容

有鉴于此，本发明的目的在于提供一种线下线上实时合唱方法、设备及介质，能够实现线上用户和线下用户实时合唱。其具体方案如下：

- 20 第一方面，本申请公开了一种线下线上实时合唱方法，应用于云服务器，包括：
- 利用拉流转推服务网络从内容分发网络中拉取线下演唱场景中的实时音视频流，并对所述实时音视频流进行进度打标得到当前打标进度，以及将所述实时音视频流和所述当前打标进度推送至实时音视频服务网络；
- 25 利用所述实时音视频服务网络将所述实时音视频流和所述当前打标进度下发至目标合唱用户终端，以便所述目标合唱用户终端播放所述实时音视频流，并利用所述当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，得到携带有进度信息的标记后人声音频流；

通过所述实时音视频服务网络获取所述目标合唱用户终端上传的所述标记后人声音频流,并利用实时音视频拉流服务网络将所述实时音视频流、所述当前打标进度和所述标记后人声音频流拉取至云端合流服务网络,以利用所述云端合流服务网络基于所述当前打标进度和所述进度信息对所述实时音视频流和所述标记后人声音频流进行合并,并将合并后音视频流推送至所述内容分发网络进行分发。

可选的,所述利用拉流转推服务网络从内容分发网络中拉取线下演唱场景中的实时音视频流,并对所述实时音视频流进行进度打标得到当前打标进度,包括:

10 在当前时间间隔达到预设时间间隔阈值时,利用拉流转推服务网络基于时间戳从前到后的顺序从当前目标音视频流中拉取与所述预设时间间隔阈值对应的实时音视频流;所述当前目标音视频流为内容分发网络中与线下演唱场景对应的当前未被拉取的音视频流;

确定与所述当前拉流操作对应的当前进度序列号,并利用所述当前进度序列号对所述实时音视频流进行进度打标得到当前打标进度。

可选的,所述在当前时间间隔达到预设时间间隔阈值时,利用拉流转推服务网络基于时间戳从前到后的顺序从当前目标音视频流中拉取与所述预设时间间隔阈值对应的实时音视频流,包括:

20 基于首次拉流时间戳和所述预设时间间隔阈值确定与每次拉流操作对应的目标拉流时间戳;其中,将所述拉流转推服务网络从所述内容分发网络中首次拉取到线下演唱场景中实时音视频流的时间作为所述首次拉流时间戳;

每当当前时间达到所述目标拉流时间戳时,利用拉流转推服务网络基于时间戳从前到后的顺序从当前目标音视频流中拉取与所述预设时间间隔阈值对应实时音视频流。

25 可选的,所述确定与所述当前拉流操作对应的当前进度序列号,并利用所述当前进度序列号对所述实时音视频流进行进度打标得到当前打标进度,包括:

-3-

基于上一拉流时间戳对应的历史打标进度和所述预设时间间隔阈值确定与所述当前拉流操作对应的当前进度序列号；其中，若当前拉流操作为首拉流操作，则所述当前进度序列号为预设的初始值；

将所述当前进度序列号作为所述实时音视频流的当前打标进度。

- 5 可选的，所述利用拉流转推服务网络从内容分发网络中拉取线下演唱场景中的实时音视频流，并对所述实时音视频流进行进度打标得到当前打标进度，包括：

- 10 利用拉流转推服务网络持续从内容分发网络中拉取线下演唱场景中的实时音视频流，并在当前时间间隔达到预设时间间隔阈值时，利用预先设置的进度序列号对所述实时音视频流进行进度打标得到当前打标进度。

可选的，所述线下线上实时合唱方法，还包括：

获取其他合唱用户终端上传的线上人声音频流；

- 15 利用所述实时音视频服务网络将所述其他合唱用户终端上传的线上人声音频流转发至所述目标合唱用户终端，以便所述目标合唱用户终端同时播放所述实时音视频流和所述其他合唱用户终端上传的线上人声音频流。

可选的，所述将合并后音视频流推送至所述内容分发网络进行分发，包括：

- 20 将合并后音视频流推送至所述内容分发网络的各内容分发网络节点，以便所述内容分发网络节点在接收到线上观众用户终端发送的数据拉取请求后向所述线上观众用户终端下发所述合并后音视频流；其中，所述线上观众用户终端基于当前网络通信状况从各所述内容分发网络节点中筛选出一个所述内容分发网络节点并向其发送所述数据拉取请求。

第二方面，本申请公开了一种线下线上实时合唱方法，应用于目标合唱用户终端，包括：

- 25 从云服务器中的实时音视频服务网络中拉取并播放线下演唱场景中的实时音视频流，并获取所述实时音视频流的当前打标进度；其中，所述云服务器利用所述拉流转推服务网络从内容分发网络中拉取所述实时音视频流，并对所述实时音视频流进行进度打标得到所述当前打标进度；

利用所述当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，得到携带有进度信息的标记后人声音频流；

5 将所述标记后人声音频流上传至所述实时音视频服务网络，以便实时音视频拉流服务网络将从所述实时音视频服务网络拉取到的所述实时音视频流、所述当前打标进度和所述标记后人声音频流发送至云端合流服务网络，并通过所述云端合流服务网络基于所述当前打标进度和所述进度信息对所述实时音视频流和所述标记后人声音频流进行合并，将合并后音视频流推送至所述内容分发网络进行分发。

10 可选的，所述利用所述当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，包括：

通过本地线上合唱应用开启本地音频采集设备，以通过所述本地音频采集设备采集本地音频数据，并对所述本地音频数据进行去噪处理，以得到人声音频流；

15 利用所述当前打标进度对通过所述本地线上合唱应用获取到的所述人声音频流进行标记。

第三方面，本申请公开了一种电子设备，包括：

存储器，用于保存计算机程序；

处理器，用于执行所述计算机程序，以实现前述公开的线下线上实时合唱方法的步骤。

20 第四方面，本申请公开了一种计算机可读存储介质，用于存储计算机程序；其中，所述计算机程序被处理器执行时实现前述公开的线下线上实时合唱方法的步骤。

25 可见，本申请公开了一种线下线上实时合唱方法，应用于云服务器，具体包括：利用拉流转推服务网络从内容分发网络中拉取线下演唱场景中的实时音视频流，并对所述实时音视频流进行进度打标得到当前打标进度，以及将所述实时音视频流和所述当前打标进度推送至实时音视频服务网络；利用所述实时音视频服务网络将所述实时音视频流和所述当前打标进度下发至目标合唱用户终端，以便所述目标合唱用户终端播放所述实时音视频流，并利用所述当前打标进度对本地线上合唱应用实时获取到的人声

音频流进行标记，得到携带有进度信息的标记后人声音频流；通过所述实时音视频服务网络获取所述目标合唱用户终端上传的所述标记后人声音频流，并利用实时音视频拉流服务网络将所述实时音视频流、所述当前打标进度和所述标记后人声音频流拉取至云端合流服务网络，以利用所述云端合流服务网络基于所述当前打标进度和所述进度信息对所述实时音视频流和所述标记后人声音频流进行合并，并将合并后音视频流推送至所述内容分发网络进行分发。

由此可见，拉流转推服务网络在从内容分发网络中拉取到线下演唱场景中的实时音视频流后，则对实时音视频流进行进度打标得到当前打标进度，再将实时音视频流和当前打标进度推送至实时音视频服务网络，也即本申请中的拉流推转服务提供音视频流的拉取、进度打标以及推送的功能。进一步的，实时音视频服务网络将实时音视频流和当前打标进度下发至目标合唱用户终端，以便目标合唱用户终端播放获取到的实时音视频流，并利用当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，得到携带有进度信息的标记后人声音频流；可以理解的是，通过利用当前打标进度对本地采集的人声音频流进行标记，能够将线下的实时音视频流和线上的人声音频流进行关联。目标合唱用户终端将标记后人声音频流上传至实时音视频服务网络后，再利用实时音视频拉流服务网络将实时音视频流、当前打标进度和标记后人声音频流拉取至云端合流服务网络，由于此时的实时音视频流和人声音频流均已标记有相应的进度，因此云端合流服务网络则基于当前打标进度和进度信息对实时音视频流和标记后人声音频流进行精准合并得到合并后音视频流，最后将合并后音视频流推送回内容分发网络进行分发。通过上述方案实现了线上用户和线下用户实时合唱的效果。

附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术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发明的实施例，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，还可以根据提供的附图获得其他的附图。

图1为本申请公开的线下线上实时合唱方案所适用的系统框架示意图；

图2为本申请公开的一种云服务器端的线下线上实时合唱方法流程图；

图3为本申请公开的一种具体的线下线上实时合唱方法流程图；

5 图4为本申请公开的一种合唱用户终端的线下线上实时合唱方法流程图；

图5为本申请公开的一种电子设备结构图。

具体实施方式

10 下面将结合本申请实施例中的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例仅仅是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

15 当前，对于线下演唱场景，如明星演唱会、推流演唱会等，演唱者和台下观众可以线下参与合唱，而线上用户仅可以通过终端设备实时观看线下演唱会，因此还缺少一种实现线上用户和线下用户实时合唱的方案。为此，本申请实施例公开了一种线下线上实时合唱方法、设备及介质，能够实现线上用户和线下用户实时合唱。

20 本申请的线下线上实时合唱方案中，采用的系统框架可以参见图1所示，具体可以包括云服务器01、目标合唱用户终端02和线上观众用户终端03。

其中，云服务器01具体包括拉流转推服务网络011、实时音视频服务网络012、实时音视频拉流服务网络013和云端合流服务网络014。

25 拉流转推服务网络011的主要功能是提供内容拉取并推送，本申请中还提供进度打标功能。需要指出的是，线下演唱场景中安装有音视频采集设备和推流设备，用于实时采集线下演唱场景中的音视频流并推流至内容分发网络（Content Delivery Network，即CDN），因此拉流转推服务网络011具体为从内容分发网络中拉取线下演唱场景中的实时音视频流，并对实时

音视频流进行进度打标得到当前打标进度，然后将实时音视频流和当前打标进度推送至实时音视频服务网络012。

实时音视频服务网络012用于提供低延迟实时音视频服务，具体包括但不限于腾讯云TRTC（Tencent RTC，其中RTC为Real Time Communication，实时音视频），声网RTC等。实时音视频服务网络012与拉流转推服务网络011、目标合唱用户终端02和实时音视频拉流服务网络013之间均有数据交互。其中，实时音视频服务网络012具体用于接收拉流转推服务网络011推送的实时音视频流和当前打标进度；实时音视频服务网络012还用于将实时音视频流和当前打标进度下发至目标合唱用户终端02，以及获取目标合唱用户终端02上传的标记后人声音频流；实时音视频服务网络012还用于响应实时音视频拉流服务网络013的拉流操作。

实时音视频拉流服务网络013用于将实时音视频服务网络012中获取到的实时音视频流、当前打标进度和标记后人声音频流拉取至云端合流服务网络014。

云端合流服务网络014用于基于当前打标进度和标记后人声音频流携带的进度信息对实时音视频流和标记后人声音频流进行合并，并将合并后音视频流推送至所述内容分发网络进行分发，从而实现线上观众用户终端03能够收听到线上用户和线下用户实时合唱的效果。

目标合唱用户终端02为使用线上合唱应用软件通过网络形式在线参与合唱的用户终端，目标合唱用户所持的终端设备可以是智能手机、平板、电脑等设备。目标合唱用户终端02用于从实时音视频服务网络012中获取并播放线下的实时音视频流，并利用从实时音视频服务网络012中获取到的当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，得到携带有进度信息的标记后人声音频流，再将标记后人声音频流推送至实时音视频服务网络012。

线上观众用户终端03为使用线上合唱应用软件通过网络形式收听合唱的用户终端，观众用户所持的终端设备可以是智能手机、平板、电脑等设备。线上观众用户终端03用于从内容分发网络中拉取合并后音视频流，从而收听到线上用户和线下用户实时合唱的效果。

参见图2所示，本申请实施例公开了一种线下线上实时合唱方法，应用于云服务器，该方法包括：

- 5 步骤 S11：利用拉流转推服务网络从内容分发网络中拉取线下演唱场景中的实时音视频流，并对所述实时音视频流进行进度打标得到当前打标进度，以及将所述实时音视频流和所述当前打标进度推送至实时音视频服务网络。

本实施例中，拉流转推服务网络在从内容分发网络中拉取到线下演唱场景中的实时音视频流后，则对实时音视频流进行进度打标得到当前打标进度，并将实时音视频流和当前打标进度推送至实时音视频服务网络。其中，在一种具体实施方式中，实时音视频流和当前打标进度可以作为两个独立的数据分别推送至实时音视频服务网络；在另一种具体实施方式中，实时音视频流可以携带有当前打标进度，然后作为一个整体推送至实时音视频服务网络。

- 15 需要注意的是，拉流转推服务网络的主要功能是提供内容拉取并推送，本申请实施例中拉流转推服务中还提供进度打标功能，用于对拉取到的实时音频流进行进度打标。

此外，还需要指出的是，线下演唱场景中安装有音视频采集设备和推流设备，用于实时采集活动现场的视频画面和声音，以得到音视频流，然后推流至内容分发网络。在相关方案中，内容分发网络在获取到线下演唱场景中的实时音视频流后，则可以进行后续的内容分发，即将实时音视频流推送至各内容分发网络节点，以便内容分发网络节点在接收到线上观众用户终端发送的数据拉取请求后向线上观众用户终端下发实时音视频流，从而使得线上观众用户终端可以实时观看线下演唱会。

- 25 在具体实施方式中，所述利用拉流转推服务网络从内容分发网络中拉取线下演唱场景中的实时音视频流，并对所述实时音视频流进行进度打标得到当前打标进度，包括：利用拉流转推服务网络持续从内容分发网络中拉取线下演唱场景中的实时音视频流，并在当前时间间隔达到预设时间间隔阈值时，利用预先设置的进度序列号对实时音视频流进行进度打标得到

当前打标进度。也即，本实施例中，拉流转推服务会持续的从内容分发网络中拉取线下演唱场景中的实时音视频流，并且拉流转推服务还需对拉取到的实时音视频流进行进度打标，具体的，每当当前时间间隔达到预设时间间隔阈值时，则利用预先设置的进度序列号对当前拉取到的实时音视频流进行进度打标得到当前打标进度。例如，假设以 20ms 作为一帧数据，相应的将预设时间间隔阈值设置为 20ms，那么则理解为拉流转推服务每隔 20ms 对当前拉取到的实时音视频流进行一次进度打标。在具体实施方式中，进度序列号也可以根据预设时间间隔阈值设定，即假设第一次拉到实时音视频流后作为起始时间 0，后续每 20ms 持续累加进度，从而得到 0、20、40、60、... 的进度序列号。此外，也可以根据打标操作的次数设置进度序列号，例如每进行一次进度打标操作，则进度序列号累计 1 次，从而得到 1、2、3... 的进度序列号。本实施例对进度序列号的设置方式不做限制。

步骤 S12：利用所述实时音视频服务网络将所述实时音视频流和所述当前打标进度下发至目标合唱用户终端，以便所述目标合唱用户终端播放所述实时音视频流，并利用所述当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，得到携带有进度信息的标记后人声音频流。

本实施例中，实时音视频服务网络将拉流转推服务推送的实时音视频流和当前打标进度下发至目标合唱用户终端，以便目标合唱用户终端播放获取到的实时音视频流，并利用当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，得到携带有进度信息的标记后人声音频流；可以理解的是，通过利用当前打标进度对本地采集的人声音频流进行标记，可以将线下的实时音视频流和线上的人声音频流进行关联。

其中，本地线上合唱应用具体可以是全民 k 歌或 QQ 音乐等应用软件，用户可随意点击加入合唱以及退出合唱。

进一步的，上述方法还包括：获取其他合唱用户终端上传的线上人声音频流；利用所述实时音视频服务网络将所述其他合唱用户终端上传的线上人声音频流转发至所述目标合唱用户终端，以便所述目标合唱用户终端同时播放所述实时音视频流和所述其他合唱用户终端上传的线上人声音频

流。也即，目标合唱用户终端除了会播放线下的实时音视频流外，还会同时播放其他合唱用户终端上传的线上人声音频流，也即合唱用户会听到线下合唱用户和其他线上合唱用户的声音。在具体实施方式中，参与合唱的用户会通过实时音视频服务网络的低延迟 RTC 通道拉取其他线上合唱者的线上人声音频流并播放。也即，参与合唱的终端用户之间通过实时音视频服务网络提供的通道进行音频流的互传。

步骤 S13：通过所述实时音视频服务网络获取所述目标合唱用户终端上传的所述标记后人声音频流，并利用实时音视频拉流服务网络将所述实时音视频流、所述当前打标进度和所述标记后人声音频流拉取至云端合流服务网络，以利用所述云端合流服务网络基于所述当前打标进度和所述进度信息对所述实时音视频流和所述标记后人声音频流进行合并，并将合并后音视频流推送至所述内容分发网络进行分发。

本实施例中，目标合唱用户终端会将标记后人声音频流上传至实时音视频服务网络后，再利用实时音视频拉流服务网络将实时音视频流、当前打标进度和标记后人声音频流拉取至云端合流服务网络。由于此时的实时音视频流和人声音频流均已标记有相应的进度，因此云端合流服务网络则基于当前打标进度和进度信息对实时音视频流和标记后人声音频流进行精准合并得到合并后音视频流，从而实现音视频流的精准对齐。最后云端合流服务将合并后音视频流推送回内容分发网络进行分发，从而实现其他观众能够听到线下线上一同合唱的效果。

可见，本申请公开了一种线下线上实时合唱方法，应用于云服务器，具体包括：利用拉流转推服务网络从内容分发网络中拉取线下演唱场景中的实时音视频流，并对所述实时音视频流进行进度打标得到当前打标进度，以及将所述实时音视频流和所述当前打标进度推送至实时音视频服务网络；利用所述实时音视频服务网络将所述实时音视频流和所述当前打标进度下发至目标合唱用户终端，以便所述目标合唱用户终端播放所述实时音视频流，并利用所述当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，得到携带有进度信息的标记后人声音频流；通过所述实时音视频服务网络获取所述目标合唱用户终端上传的所述标记后人声音频

流，并利用实时音视频拉流服务网络将所述实时音视频流、所述当前打标进度和所述标记后人声音频流拉取至云端合流服务网络，以利用所述云端合流服务网络基于所述当前打标进度和所述进度信息对所述实时音视频流和所述标记后人声音频流进行合并，并将合并后音视频流推送至所述内容分发网络进行分发。

由此可见，拉流转推服务网络在从内容分发网络中拉取到线下演唱场景中的实时音视频流后，则对实时音视频流进行进度打标得到当前打标进度，再将实时音视频流和当前打标进度推送至实时音视频服务网络，也即本申请中的拉流推转服务提供音视频流的拉取、进度打标以及推送的功能。

进一步的，实时音视频服务网络将实时音视频流和当前打标进度下发至目标合唱用户终端，以便目标合唱用户终端播放获取到的实时音视频流，并利用当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，得到携带有进度信息的标记后人声音频流；可以理解的是，通过利用当前打标进度对本地采集的人声音频流进行标记，能够将线下的实时音视频流和线上的人声音频流进行关联。目标合唱用户终端将标记后人声音频流上传至实时音视频服务网络后，再利用实时音视频拉流服务网络将实时音视频流、当前打标进度和标记后人声音频流拉取至云端合流服务网络，由于此时的实时音视频流和人声音频流均已标记有相应的进度，因此云端合流服务网络则基于当前打标进度和进度信息对实时音视频流和标记后人声音频流进行精准合并得到合并后音视频流，最后将合并后音视频流推送回内容分发网络进行分发。通过上述方案实现了线上用户和线下用户实时合唱的效果。

参见图 3 所示，本申请实施例公开了一种具体的线下线上实时合唱方法，相对于上一实施例，本实施例对技术方案作了进一步的说明和优化。具体包括：

步骤 S21：在当前时间间隔达到预设时间间隔阈值时，利用拉流转推服务网络基于时间戳从前到后的顺序从当前目标音视频流中拉取与所述预

设时间间隔阈值对应的实时音视频流；所述当前目标音视频流为内容分发网络中与线下演唱场景对应的当前未被拉取的音视频流。

本实施例中，拉流转推服务也可以间断的从内容分发网络中拉取线下演唱场景中的实时音视频流，但为了保证实时性，减少时延，时间间隔阈值不易设置太大，例如，可以设置为每隔 20ms 从内容分发网络拉取一次。因此，在当前时间间隔达到预设时间间隔阈值时，则利用拉流转推服务网络基于时间戳从前到后的顺序从当前目标音视频流中拉取与预设时间间隔阈值对应的实时音视频流。

可以理解的是，为了避免拉取重复的音视频流且保证音视频流的连续性，本实施例需要按照时间戳从前到后的顺序从内容分发网络中当前未被拉取的音视频流中拉取与预设时间间隔阈值对应的实时音视频流。也即，拉流转推服务每次均从内容分发网络中拉取 20ms 的音视频流，假设一段音视频流按照毫秒为单位记为 0~100ms，那么第一次则拉取 0~19ms 的内容，第二次则拉取 20~39ms 的内容，以此类推。

在具体实施方式中，上述在当前时间间隔达到预设时间间隔阈值时，利用拉流转推服务网络基于时间戳从前到后的顺序从当前目标音视频流中拉取与所述预设时间间隔阈值对应的实时音视频流，包括：基于首次拉流时间戳和所述预设时间间隔阈值确定与每次拉流操作对应的目标拉流时间戳；其中，将所述拉流转推服务网络从所述内容分发网络中首次拉取到线下演唱场景中实时音视频流的时间作为所述首次拉流时间戳；每当当前时间达到所述目标拉流时间戳时，利用拉流转推服务网络基于时间戳从前到后的顺序从当前目标音视频流中拉取与所述预设时间间隔阈值对应实时音视频流。

也即，由于拉流转推服务每隔预设时间间隔阈值进行一次拉流操作，那么在确定出拉流转推服务网络首次从内容分发网络中拉取到线下演唱场景中实时音视频流的时间（即首次拉流时间戳）后，则可根据首次拉流时间戳和预设时间间隔阈值确定出与每次拉流操作对应的目标拉流时间戳。如此一来，每当当前时间达到目标拉流时间戳时，则利用拉流转推服务网

络基于时间戳从前到后的顺序从当前目标音视频流中拉取与预设时间间隔阈值对应的实时音视频流。

步骤 S22: 确定与所述当前拉流操作对应的当前进度序列号, 并利用所述当前进度序列号对所述实时音视频流进行进度打标得到当前打标进度, 将所述实时音视频流和所述当前打标进度推送至实时音视频服务网络。

本实施例中, 拉流转推服务每进行一次拉流操作并拉取到实时音视频流后, 则利用与当前拉流操作对应的当前进度序列号对实时音视频流进行进度打标得到当前打标进度, 然后将实时音视频流和当前打标进度推送至实时音视频服务网络。

10 在具体实施方式中, 上述确定与所述当前拉流操作对应的当前进度序列号, 并利用所述当前进度序列号对所述实时音视频流进行进度打标得到当前打标进度, 包括: 基于上一拉流时间戳对应的历史打标进度和所述预设时间间隔阈值确定与所述当前拉流操作对应的当前进度序列号; 其中, 若当前拉流操作为首次拉流操作, 则所述当前进度序列号为预设的初始值; 15 将所述当前进度序列号作为所述实时音视频流的当前打标进度。也即, 本实施例可以根据拉流的时间设置打标进度序列号, 具体的, 确定上一拉流时间戳对应的历史打标进度, 然后在此基础上加上预设时间间隔阈值以得到与当前拉流操作对应的当前进度序列号。例如, 假设上一拉流时间戳对应的历史打标进度为 40, 预设拉流时间间隔阈值为 20, 那么当前拉流操作 20 对应的当前进度序列号则为 60。在拉流时间间隔阈值不变的情况下, 打标进度则随着时间的进行, 按照 20ms 的速度逐步递增。若当前为首次拉流操作, 则当前进度序列号为预设的初始值, 例如可设置为 0。

步骤 S23: 利用所述实时音视频服务网络将所述实时音视频流和所述当前打标进度下发至目标合唱用户终端, 以便所述目标合唱用户终端播放 25 所述实时音视频流, 并利用所述当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记, 得到携带有进度信息的标记后人声音频流。

步骤S24: 通过所述实时音视频服务网络获取所述目标合唱用户终端上传的所述标记后人声音频流, 并利用实时音视频拉流服务网络将所述实时音视频流、所述当前打标进度和所述标记后人声音频流拉取至云端合流服

务网络，以利用所述云端合流服务网络基于所述当前打标进度和所述进度信息对所述实时音视频流和所述标记后人声音频流进行合并。

步骤S25: 将合并后音视频流推送至所述内容分发网络的各内容分发网络节点，以便所述内容分发网络节点在接收到线上观众用户终端发送的数据拉取请求后向所述线上观众用户终端下发所述合并后音视频流；其中，
5 所述线上观众用户终端基于当前网络通信状况从各所述内容分发网络节点中筛选出一个所述内容分发网络节点并向其发送所述数据拉取请求。

本实施例中，云端合流服务网络在对实时音视频流和标记后人声音频流进行合并后，则将合并后音视频流推送至内容分发网络的各内容分发网络节点。
10 内容分发网络节点在接收到线上观众用户终端发送的数据拉取请求后则向线上观众用户终端下发该合并后音视频流。可以理解的是，线上观众用户终端基于当前网络通信状况从各内容分发网络节点中筛选出一个内容分发网络节点并向其发送数据拉取请求，在具体实施例中，线上观众用户终端会通过离自己最近的内容分发网络节点拉取该合并后音视频流，
15 从而实现了线上观众能够听到线下用户和线上用户一起合唱的效果。

其中，关于上述步骤 S23 和 S24 更加具体的处理过程可以参考前述实施例中公开的相应内容，在此不再进行赘述。

可见，本申请实施例中，拉流转推服务网络可以每隔预设时间间隔阈值完成一次拉流以及进度打标操作。具体的，由于拉流转推服务每隔预设时间间隔阈值进行一次拉流操作，那么在确定出拉流转推服务网络的首次拉流时间戳后，则可根据首次拉流时间戳和预设时间间隔阈值确定出与每次拉流操作对应的目标拉流时间戳。如此一来，每当当前时间达到目标拉流时间戳时，则执行一次拉流操作。在这个过程中，为了避免拉取重复的音视频流且保证音视频流的连续性，本实施例需要按照时间戳从前到后的
20 顺序从内容分发网络中当前未被拉取的音视频流中拉取与预设时间间隔阈值对应的实时音视频流。另外，线上观众用户终端会通过离自己最近的内容分发网络节点拉取该合并后音视频流，从而实现了线上观众能够听到线下用户和线上用户一起合唱的效果。
25

参见图4所示，本申请实施例公开了一种线下线上实时合唱方法，应用于目标合唱用户终端，该方法包括：

步骤 S31：从云服务器中的实时音视频服务网络中拉取并播放线下演唱场景中的实时音视频流，并获取所述实时音视频流的当前打标进度；其中，所述云服务器利用所述拉流转推服务网络从内容分发网络中拉取所述实时音视频流，并对所述实时音视频流进行进度打标得到所述当前打标进度。

本实施例中，目标合唱用户终端即线上参与合唱的用户终端，目标合唱用户终端从云服务器中的实时音视频服务网络中拉取并播放线下演唱场景中的实时音视频流，并获取实时音视频流的当前打标进度。需要指出的是，云服务器具体利用拉流转推服务网络从内容分发网络中拉取实时音视频流，并对实时音视频流进行进度打标得到当前打标进度，然后拉流转推服务再将实时音视频流和当前打标进度推送至实时音视频服务网络。

步骤 S32：利用所述当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，得到携带有进度信息的标记后人声音频流。

本实施例中，目标合唱用户终端利用获取到的当前打标进度对本地线上合唱应用获取到的由本地音频采集设备采集的人声音频流进行标记，得到携带有进度信息的标记后人声音频流。其中，本地音频采集设备具体可以为麦克风。

在具体实施方式中，上述利用所述当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，包括：通过本地线上合唱应用开启本地音频采集设备，以通过所述本地音频采集设备采集本地音频数据，并对所述本地音频数据进行去噪处理，以得到人声音频流；利用所述当前打标进度对通过所述本地线上合唱应用获取到的所述人声音频流进行标记。可以理解的是，当用户确认参加合唱后，本地线上合唱应用会开启本地音频采集设备，即麦克风，然后通过麦克风采集本地的音频数据，并且在这个过程中可以对本地音频数据进行去噪处理，以去除周围环境音和噪音，以得到人声音频流。然后再利用当前打标进度对通过本地线上合唱应用获取到的人声音频流进行标记，通过利用第一进度信息对本地采集的线上人声

数据进行标记。通过利用当前打标进度对本地采集的人声音频流进行标记，能够将线下的实时音视频流和线上的人声音频流进行关联。

步骤 S33: 将所述标记后人声音频流上传至所述实时音视频服务网络，以便实时音视频拉流服务网络将从所述实时音视频服务网络拉取到的所述
5 实时音视频流、所述当前打标进度和所述标记后人声音频流发送至云端合流服务网络，并通过所述云端合流服务网络基于所述当前打标进度和所述进度信息对所述实时音视频流和所述标记后人声音频流进行合并，将合并后音视频流推送至所述内容分发网络进行分发。

本实施例中，目标合唱用户终端将标记后人声音频流上传至实时音视
10 频服务网络，进一步的，由实时音视频拉流服务网络同时将实时音视频服务平台中的实时音视频流、当前打标进度和标记后人声音频流拉取至云端合流服务网络，由云端合流服务网络基于当前打标进度和进度信息对实时音视频流和标记后人声音频流进行精准合并得到合并后音视频流，从而实现线下实时音视频流和显示人声音频流的精准对齐，最后云端合流服务网
15 络将合并后音视频流推送至内容分发网络进行分发，从而实现其他观众能够听到线上用户和线下用户一起合唱的效果。

可见，目标合唱用户终端即线上参与合唱的用户终端，目标合唱用户终端从云服务器中的实时音视频服务网络中拉取并播放线下演唱场景中的实时音视频流，并获取实时音视频流的当前打标进度。进一步的，目标合唱
20 用户终端利用获取到的当前打标进度对本地线上合唱应用获取到的由本地音频采集设备采集的人声音频流进行标记，得到携带有进度信息的标记后人声音频流。最后，目标合唱用户终端将标记后人声音频流上传至实时音视频服务网络，进一步的，由实时音视频拉流服务网络同时将实时音视频服务平台中的实时音视频流、当前打标进度和标记后人声音频流拉取至
25 云端合流服务网络，由云端合流服务网络基于当前打标进度和进度信息对实时音视频流和标记后人声音频流进行精准合并得到合并后音视频流，从而实现线下实时音视频流和显示人声音频流的精准对齐，最后云端合流服务网络将合并后音视频流推送至内容分发网络进行分发，从而实现其他观众能够听到线上用户和线下用户一起合唱的效果。

图5为本申请实施例提供的一种电子设备的结构示意图。具体可以包括：至少一个处理器21、至少一个存储器22、电源23、通信接口24、输入输出接口25和通信总线26。其中，所述存储器22用于存储计算机程序，所述计算机程序由所述处理器21加载并执行，以实现前述任一实施例公开的由电子设备执行的线下线上实时合唱方法中的相关步骤。

本实施例中，电源23用于为电子设备20上的各硬件设备提供工作电压；通信接口24能够为电子设备20创建与外界设备之间的数据传输通道，其所遵循的通信协议是能够适用于本申请技术方案的任意通信协议，在此不对其进行具体限定；输入输出接口25，用于获取外界输入数据或向外界输出数据，其具体的接口类型可以根据具体应用需要进行选取，在此不进行具体限定。

其中，处理器21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21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21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21可以在集成有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21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。

另外，存储器22作为资源存储的载体，可以是只读存储器、随机存储器、磁盘或者光盘等，其上所存储的资源包括操作系统221、计算机程序222及数据223等，存储方式可以是短暂存储或者永久存储。

其中，操作系统221用于管理与控制电子设备20上的各硬件设备以及计算机程序222，以实现处理器21对存储器22中海量数据223的运算与处理，其可以是Windows、Unix、Linux等。计算机程序222除了包括能够用于完

成前述任一实施例公开的由电子设备20执行的线下线上实时合唱方法的计算机程序之外，还可以进一步包括能够用于完成其他特定工作的计算机程序。数据223除了可以包括电子设备接收到的由外部设备传输进来的数据，也可以包括由自身输入输出接口25采集到的数据等。

5

进一步的，本申请实施例还公开了一种计算机可读存储介质，所述存储介质中存储有计算机程序，所述计算机程序被处理器加载并执行时，实现前述任一实施例公开的线下线上实时合唱方法步骤。

10 本说明书中各个实施例采用递进的方式描述，每个实施例重点说明的都是与其它实施例的不同之处，各个实施例之间相同或相似部分互相参见即可。

本领域技术人员还可以进一步意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、计算机软件或者二者的结合来实现，为了清楚地说明硬件和软件的可互换性，在上述说明中已经按照功能一般性地描述了各示例的组成及步骤。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。本领域技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

20 结合本文中所公开的实施例描述的方法或算法的步骤可以直接用硬件、处理器执行的软件模块，或者二者的结合来实施。软件模块可以置于随机存储器（Random Access Memory，即RAM）、内存、只读存储器（Read-Only Memory，即ROM）、电可编程ROM、电可擦除可编程ROM、寄存器、硬盘、可移动磁盘、只读光盘（Compact Disc Read-Only Memory，即CD-ROM）、或技术领域内所公知的任意其它形式的存储介质中。

25 最后，还需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，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、物品或者设备不仅包括那些

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、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

- 5 以上对本发明所提供的一种线下线上实时合唱方法、设备及存储介质进行了详细介绍，本文中应用了具体个例对本发明的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本发明的方法及其核心思想；同时，对于本领域的一般技术人员，依据本发明的思想，在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对本发明的限制。
- 10

权 利 要 求

- 1、一种线下线上实时合唱方法，其特征在于，应用于云服务器，包括：
利用拉流转推服务网络从内容分发网络中拉取线下演唱场景中的实时音视频流，并对所述实时音视频流进行进度打标得到当前打标进度，以及
5 将所述实时音视频流和所述当前打标进度推送至实时音视频服务网络；
利用所述实时音视频服务网络将所述实时音视频流和所述当前打标进度下发至目标合唱用户终端，以便所述目标合唱用户终端播放所述实时音视频流，并利用所述当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，得到携带有进度信息的标记后人声音频流；
10 通过所述实时音视频服务网络获取所述目标合唱用户终端上传的所述标记后人声音频流，并利用实时音视频拉流服务网络将所述实时音视频流、所述当前打标进度和所述标记后人声音频流拉取至云端合流服务网络，以利用所述云端合流服务网络基于所述当前打标进度和所述进度信息对所述实时音视频流和所述标记后人声音频流进行合并，并将合并后音视频流推
15 送至所述内容分发网络进行分发。
- 2、根据权利要求1所述的线下线上实时合唱方法，其特征在于，所述利用拉流转推服务网络从内容分发网络中拉取线下演唱场景中的实时音视频流，并对所述实时音视频流进行进度打标得到当前打标进度，包括：
在当前时间间隔达到预设时间间隔阈值时，利用拉流转推服务网络基
20 于时间戳从前到后的顺序从当前目标音视频流中拉取与所述预设时间间隔阈值对应的实时音视频流；所述当前目标音视频流为内容分发网络中与线下演唱场景对应的当前未被拉取的音视频流；
确定与所述当前拉流操作对应的当前进度序列号，并利用所述当前进度序列号对所述实时音视频流进行进度打标得到当前打标进度。
- 25 3、根据权利要求2所述的线下线上实时合唱方法，其特征在于，所述在当前时间间隔达到预设时间间隔阈值时，利用拉流转推服务网络基于时间戳从前到后的顺序从当前目标音视频流中拉取与所述预设时间间隔阈值对应的实时音视频流，包括：

基于首次拉流时间戳和所述预设时间间隔阈值确定与每次拉流操作对应的目标拉流时间戳；其中，将所述拉流转推服务网络从所述内容分发网络中首次拉取到线下演唱场景中实时音视频流的时间作为所述首次拉流时间戳；

5 每当当前时间达到所述目标拉流时间戳时，利用拉流转推服务网络基于时间戳从前到后的顺序从当前目标音视频流中拉取与所述预设时间间隔阈值对应实时音视频流。

4、根据权利要求2所述的线下线上实时合唱方法，其特征在于，所述确定与所述当前拉流操作对应的当前进度序列号，并利用所述当前进度序列号对所述实时音视频流进行进度打标得到当前打标进度，包括：

基于上一拉流时间戳对应的历史打标进度和所述预设时间间隔阈值确定与所述当前拉流操作对应的当前进度序列号；其中，若当前拉流操作为首次拉流操作，则所述当前进度序列号为预设的初始值；

将所述当前进度序列号作为所述实时音视频流的当前打标进度。

15 5、根据权利要求1所述的线下线上实时合唱方法，其特征在于，所述利用拉流转推服务网络从内容分发网络中拉取线下演唱场景中的实时音视频流，并对所述实时音视频流进行进度打标得到当前打标进度，包括：

利用拉流转推服务网络持续从内容分发网络中拉取线下演唱场景中的实时音视频流，并在当前时间间隔达到预设时间间隔阈值时，利用预先设置的进度序列号对所述实时音视频流进行进度打标得到当前打标进度。

20 6、根据权利要求1所述的线下线上实时合唱方法，其特征在于，还包括：

获取其他合唱用户终端上传的线上人声音频流；

25 利用所述实时音视频服务网络将所述其他合唱用户终端上传的线上人声音频流转发至所述目标合唱用户终端，以便所述目标合唱用户终端同时播放所述实时音视频流和所述其他合唱用户终端上传的线上人声音频流。

7、根据权利要求1至6任一项所述的线下线上实时合唱方法，其特征在于，所述将合并后音视频流推送至所述内容分发网络进行分发，包括：

将合并后音视频流推送至所述内容分发网络的各内容分发网络节点，以便所述内容分发网络节点在接收到线上观众用户终端发送的数据拉取请求后向所述线上观众用户终端下发所述合并后音视频流；其中，所述线上观众用户终端基于当前网络通信状况从各所述内容分发网络节点中筛选出一个所述内容分发网络节点并向其发送所述数据拉取请求。

8、一种线下线上实时合唱方法，其特征在于，应用于目标合唱用户终端，包括：

从云服务器中的实时音视频服务网络中拉取并播放线下演唱场景中的实时音视频流，并获取所述实时音视频流的当前打标进度；其中，所述云服务器利用所述拉流转推服务网络从内容分发网络中拉取所述实时音视频流，并对所述实时音视频流进行进度打标得到所述当前打标进度；

利用所述当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，得到携带有进度信息的标记后人声音频流；

将所述标记后人声音频流上传至所述实时音视频服务网络，以便实时音视频拉流服务网络将从所述实时音视频服务网络拉取到的所述实时音视频流、所述当前打标进度和所述标记后人声音频流发送至云端合流服务网络，并通过所述云端合流服务网络基于所述当前打标进度和所述进度信息对所述实时音视频流和所述标记后人声音频流进行合并，将合并后音视频流推送至所述内容分发网络进行分发。

9、根据权利要求 8 所述的线下线上实时合唱方法，其特征在于，所述利用所述当前打标进度对本地线上合唱应用实时获取到的人声音频流进行标记，包括：

通过本地线上合唱应用开启本地音频采集设备，以通过所述本地音频采集设备采集本地音频数据，并对所述本地音频数据进行去噪处理，以得到人声音频流；

利用所述当前打标进度对通过所述本地线上合唱应用获取到的所述人声音频流进行标记。

10、一种电子设备，其特征在于，包括：

存储器，用于保存计算机程序；

-23-

处理器，用于执行所述计算机程序，以实现如权利要求 1 至 9 任一项所述的线下线上实时合唱方法的步骤。

- 11、一种计算机可读存储介质，其特征在于，用于存储计算机程序；其中，所述计算机程序被处理器执行时实现如权利要求 1 至 9 任一项所述
- 5 的线下线上实时合唱方法的步骤。