

REAL-TIME ONLINE AND OFFLINE CHORUS METHOD, DEVICE, AND MEDIUM

[0001] This application claims the priority to Chinese Patent Application No. 202311602552.5, titled "REAL-TIME ONLINE AND OFFLINE CHORUS METHOD,
5 DEVICE, AND MEDIUM", filed on November 27, 2023 with the China National Intellectual Property Administration, which is incorporated herein by reference in its entirety.

FIELD

[0002] The present disclosure relates to the technical field of communications, and in
10 particular to a method for offline and online real-time chorus, a device, and a medium.

BACKGROUND

[0003] Currently, with conventional chorus application software on the market, users can participate in online chorus, that is, the users can participate in the chorus only in an online
15 manner. In offline singing scenarios such as star concerts and stream-pushing concerts, performers and audiences in a venue can participate in chorus in an offline manner, while online users can only watch offline concerts in real time on terminal devices. Therefore, a solution for real-time chorus between online users and offline users is still absent.

[0004] In summary, how to achieve real-time chorus between online users and offline users
20 is an issue to be addressed at present.

SUMMARY

[0005] In view of this, an objective of the present disclosure is to provide a method for offline and online real-time chorus, a device, and a medium, to achieve real-time chorus
25 between online users and offline users. The solutions according to the present disclosure are described as follows.

[0006] In a first aspect, a method for offline and online real-time chorus, applied to a cloud server, is provided in the present disclosure. The method includes:

by using a stream pulling-pushing service network, pulling a real-time audio-video

stream in an offline singing scenario from a content delivery network, performing progress marking on the real-time audio-video stream to obtain a current marked progress, and pushing the real-time audio-video stream and the current marked progress to a real-time audio-video service network;

5 delivering, by using the real-time audio-video service network, the real-time audio-video stream and the current marked progress to a target chorus user terminal, where the target chorus user terminal plays the real-time audio-video stream and marks, based on the current marked progress, a human voice audio stream acquired in real time through a local online chorus application to obtain a marked human voice audio stream carrying progress
10 information;

 acquiring, by using the real-time audio-video service network, the marked human voice audio stream uploaded by the target chorus user terminal;

 pulling, by using a real-time audio-video stream pulling service network, the real-time audio-video stream, the current marked progress, and the marked human voice audio stream
15 to a cloud stream mixing service network; and

 by using the cloud stream mixing service network, mixing the real-time audio-video stream and the marked human voice audio stream based on the current marked progress and the progress information, and pushing the mixed audio-video stream to the content delivery network for distribution.

20 **[0007]** In an embodiment, by using the stream pulling-pushing service network, pulling the real-time audio-video stream in the offline singing scenario from the content delivery network and performing progress marking on the real-time audio-video stream to obtain the current marked progress includes:

 upon a current time interval reaching a preset time interval threshold, pulling, by using
25 the stream pulling-pushing service network, the real-time audio-video stream within the preset time interval threshold from a current target audio-video stream in chronological order, where the current target audio-video stream is an audio-video stream that is not pulled from the content delivery network in the offline singing scenario; and

 determining a current progress sequence number corresponding to a current stream
30 pulling operation, and performing progress marking on the real-time audio-video stream

based on the current progress sequence number to obtain the current marked progress.

[0008] In an embodiment, upon the current time interval reaching the preset time interval threshold, pulling, by using the stream pulling-pushing service network, the real-time audio-video stream within the preset time interval threshold from the current target audio-video stream in chronological order includes:

determining, based on a first stream pulling timestamp and the preset time interval threshold, a target stream pulling timestamp corresponding to the current stream pulling operation, where the first stream pulling timestamp is a time instant when the stream pulling-pushing service network first pulls the real-time audio-video stream in the offline singing scenario from the content delivery network for the first stream pulling timestamp; and

upon a current time instant reaching the target stream pulling timestamp, pulling, by using the stream pulling-pushing service network, the real-time audio-video stream within the preset time interval threshold from the current target audio-video stream in chronological order.

[0009] In an embodiment, the determining the current progress sequence number corresponding to the current stream pulling operation, and performing progress marking on the real-time audio-video stream based on the current progress sequence number to obtain the current marked progress includes:

determining the current progress sequence number corresponding to the current stream pulling operation based on the preset time interval threshold and a historical marked progress corresponding to a previous stream pulling timestamp, where in a case that the current stream pulling operation is a first stream pulling operation implemented first by the stream pulling-pushing service network, the current progress sequence number is a preset initial value; and

determining the current progress sequence number as the current marked progress of the real-time audio-video stream.

[0010] In an embodiment, by using the stream pulling-pushing service network, pulling the real-time audio-video stream in the offline singing scenario from the content delivery network and performing progress marking on the real-time audio-video stream to obtain the current marked progress includes:

by using the stream pulling-pushing service network, continuously pulling the real-time audio-video stream in the offline singing scenario from the content delivery network, and performing progress marking, upon a current time interval reaching a preset time interval threshold, on the real-time audio-video stream based on a preset progress sequence number to obtain the current marked progress.

[0011] In an embodiment, the method for offline and online real-time chorus further includes:

acquiring an online human voice audio stream uploaded by a chorus user terminal other than the target chorus user terminal; and

forwarding, by using the real-time audio-video service network, the online human voice audio stream uploaded by the chorus user terminal to the target chorus user terminal, to cause the target chorus user terminal to simultaneously play the real-time audio-video stream and the online human voice audio stream uploaded by the chorus user terminal.

[0012] In an embodiment, the pushing the mixed audio-video stream to the content delivery network for distribution includes:

pushing the mixed audio-video stream to content delivery network nodes in the content delivery network, to cause one content delivery network node among the content delivery network nodes to deliver the mixed audio-video stream to an online audience user terminal upon receiving a data pulling request from the online audience user terminal, where the online audience user terminal selects the one content delivery network node from the content delivery network nodes based on current network communication conditions and sends the data pulling request to the selected content delivery network node.

[0013] In a second aspect, a method for offline and online real-time chorus, applied to a target chorus user terminal, is provided in the present disclosure. The method includes:

from a real-time audio-video service network in a cloud server, pulling and playing a real-time audio-video stream in an offline singing scenario, and acquiring a current marked progress of the real-time audio-video stream, where by using a stream pulling-pushing service network in the cloud server, the real-time audio-video stream is pulled from a content delivery network, and progress marking is performed on the real-time audio-video stream to obtain the current marked progress;

marking, based on the current marked progress, a human voice audio stream acquired in real time through a local online chorus application, to obtain a marked human voice audio stream carrying progress information;

5 uploading the marked human voice audio stream to the real-time audio-video service network;

 sending, by using a real-time audio-video stream pulling service network, the real-time audio-video stream, the current marked progress, and the marked human voice audio stream from the real-time audio-video service network to a cloud stream mixing service network; and

10 by using the cloud stream mixing service network, mixing the real-time audio-video stream and the marked human voice audio stream based on the current marked progress and the progress information, and pushing the mixed audio-video stream to the content delivery network for distribution.

[0014] In an embodiment, the marking, based on the current marked progress, the human voice audio stream acquired in real time through the local online chorus application includes:

15 activating a local audio acquisition device through the local online chorus application to acquire local audio data by using the local audio acquisition device, and denoising the local audio data to obtain the human voice audio stream; and

 marking, based on the current marked progress, the human voice audio stream acquired through the local online chorus application.

20 **[0015]** In a third aspect, an electronic device is provided in the present disclosure. The electronic device includes:

 a memory, configured to store a computer program; and

 a processor, configured to, when executing the computer program, perform the method for offline and online real-time chorus according to the foregoing embodiments.

25 **[0016]** In a fourth aspect, a computer-readable storage medium, storing a computer program, is provided in the present disclosure. The computer program, when being executed by a processor, causes the processor to perform the method for offline and online real-time chorus according to the foregoing embodiments.

[0017] It can be seen that a method for offline and online real-time chorus, applied to a

cloud server, is provided in the present disclosure. The method includes: by using a stream pulling-pushing service network, pulling a real-time audio-video stream in an offline singing scenario from a content delivery network, performing progress marking on the real-time audio-video stream to obtain a current marked progress, and pushing the real-time audio-video stream and the current marked progress to a real-time audio-video service network; delivering, by using the real-time audio-video service network, the real-time audio-video stream and the current marked progress to a target chorus user terminal, where the target chorus user terminal plays the real-time audio-video stream and marks, based on the current marked progress, a human voice audio stream acquired in real time through a local online chorus application to obtain a marked human voice audio stream carrying progress information; acquiring, by using the real-time audio-video service network, the marked human voice audio stream uploaded by the target chorus user terminal; pulling, by using a real-time audio-video stream pulling service network, the real-time audio-video stream, the current marked progress, and the marked human voice audio stream to a cloud stream mixing service network; and by using the cloud stream mixing service network, mixing the real-time audio-video stream and the marked human voice audio stream based on the current marked progress and the progress information, and pushing the mixed audio-video stream to the content delivery network for distribution.

[0018] It can be seen that after pulling the real-time audio-video stream in the offline singing scenario from the content delivery network, the stream pulling-pushing service network performs progress marking on the real-time audio-video stream to obtain the current marked progress, and then pushes the real-time audio-video stream and the current marked progress to the real-time audio-video service network. That is, the stream pulling-pushing service network according to the present disclosure provides functions of pulling, performing progress marking, and pushing audio-video streams. Further, the real-time audio-video service network delivers the real-time audio-video stream and the current marked progress to the target chorus user terminal, so that the target chorus user terminal plays the acquired real-time audio-video stream, and marks, based on the current marked progress, the human voice audio stream acquired in real time through the local online chorus application to obtain the marked human voice audio stream carrying the progress information. It can be understood that through marking the locally acquired human voice audio stream based on the current marked progress, the offline real-time audio-video stream is associated with the online human voice

audio stream. After the target chorus user terminal uploads the marked human voice audio stream to the real-time audio-video service network, the real-time audio-video stream pulling service network pulls the real-time audio-video stream, the current marked progress, and the marked human voice audio stream to the cloud stream mixing service network. Since both the
5 real-time audio-video stream and the human voice audio stream are currently marked with the progress, the cloud stream mixing service network accurately mixes the real-time audio-video stream and the marked human voice audio stream based on the current marked progress and the progress information to obtain the mixed audio-video stream, and finally pushes the mixed audio-video stream to the content delivery network for distribution. Through the above
10 solution, real-time chorus between online users and offline users is achieved.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] In order to more clearly illustrate technical solutions according to the embodiments of the present disclosure or in the conventional technology, drawings referred to in the
15 embodiments or in the conventional technology are briefly described hereinafter. Apparently, the drawings in the following descriptions only show some embodiments of the present disclosure, and other drawings may be obtained by those skilled in the art based on the provided drawings without any creative effort.

[0020] FIG. 1 is a schematic diagram of a system framework to which a solution for offline
20 and online real-time chorus according to the present disclosure is applicable;

[0021] FIG. 2 is a flowchart of a method for offline and online real-time chorus applied to a cloud server according to the present disclosure;

[0022] FIG. 3 is a flowchart of a method for offline and online real-time chorus according to an embodiment of the present disclosure;

25 **[0023]** FIG. 4 is a flowchart of a method for offline and online real-time chorus applied to a chorus user terminal according to the present disclosure; and

[0024] FIG. 5 is a structural diagram of an electronic device according to the present disclosure.

DETAILED DESCRIPTION

[0025] The technical solutions in the embodiments of the present disclosure are described clearly and completely hereinafter in conjunction with the drawings in the embodiments of the present disclosure. Apparently, the described embodiments are only a part rather than all of the embodiments of the present disclosure. All other embodiments obtained by those skilled in the art based on the embodiments of the present disclosure without any creative work fall within the protection scope of the present disclosure.

[0026] At present, in offline singing scenarios such as star concerts and streaming concerts, performers and audiences in a venue can participate in chorus in an offline manner, while online users can only watch offline concerts in real time on terminal devices. Therefore, a solution for real-time chorus between online users and offline users is still absent. To address this issue, a method for offline and online real-time chorus, a device, and a medium are provided in embodiments of the present disclosure, to achieve real-time chorus between online users and offline users.

[0027] In a solution for offline and online real-time chorus according to the present disclosure, a system framework shown in FIG. 1 is adopted. The system framework may include a cloud server 01, a target chorus user terminal 02, and an online audience user terminal 03.

[0028] The cloud server 01 may include a stream pulling-pushing service network 011, a real-time audio-video service network 012, a real-time audio-video stream pulling service network 013, and a cloud stream mixing service network 014.

[0029] A main function of the stream pulling-pushing service network 011 is pulling and pushing content. The stream pulling-pushing service network 011 according to the present disclosure further provides a progress marking function. It should be noted that an audio-video acquisition device and a stream pushing device are arranged in an offline singing scenario to acquire an audio-video stream in real time in the offline singing scenario and pushing the audio-video stream to a content delivery network (CDN). Therefore, the stream pulling-pushing service network 011 may pull the real-time audio-video stream in the offline singing scenario from the content delivery network, perform progress marking on the real-time audio-video stream to obtain a current marked progress, and then push the real-time audio-video stream and the current marked progress to the real-time audio-video service

network 012.

[0030] The real-time audio-video service network 012 provides low-latency real-time audio and video services and includes but is not limited to Tencent RTC (TRTC, where RTC is short for real time communication, for real-time audios and videos), Agora RTC and the like. The real-time audio-video service network 012 performs data interactions with the stream pulling-pushing service network 011, the target chorus user terminal 02, and the real-time audio-video stream pulling service network 013. The real-time audio-video service network 012 receives the real-time audio-video stream and the current marked progress pushed by the stream pulling-pushing service network 011. The real-time audio-video service network 012 further delivers the real-time audio-video stream and the current marked progress to the target chorus user terminal 02, and acquires a marked human voice audio stream uploaded by the target chorus user terminal 02. The real-time audio-video service network 012 further responds to stream pulling operations of the real-time audio-video stream pulling service network 013.

[0031] The real-time audio-video stream pulling service network 013 pulls the real-time audio-video stream, the current marked progress, and the marked human voice audio stream acquired from the real-time audio-video service network 012 to the cloud stream mixing service network 014.

[0032] The cloud stream mixing service network 014 mixes the real-time audio-video stream and the marked human voice audio stream based on the current marked progress and progress information carried by the marked human voice audio stream, and pushes the mixed audio-video stream to the content delivery network for distribution, thereby enabling the online audience user terminal 03 to hear real-time chorus between online users and offline users.

[0033] The target chorus user terminal 02 is a user terminal on which online chorus application software is adopted by a target chorus user to participate in chorus online through a network. A terminal device used by the target chorus user may be a smartphone, a tablet, a computer or the like. The target chorus user terminal 02 acquires and plays the offline real-time audio-video stream from the real-time audio-video service network 012, marks, based on the current marked progress acquired from the real-time audio-video service network 012, the human voice audio stream acquired in real time through the local online

chorus application software, to obtain the marked human voice audio stream carrying progress information, and then pushes the marked human voice audio stream to the real-time audio-video service network 012.

[0034] The online audience user terminal 03 is a user terminal on which the online chorus application software is adopted by an audience user to listen to chorus through a network. A terminal device adopted by the audience user may be a smartphone, a tablet, a computer or the like. The online audience user terminal 03 pulls the mixed audio-video stream from the content delivery network, thereby enabling the audience user to hear real-time chorus between online users and offline users.

[0035] Referring to FIG. 2, a method for offline and online real-time chorus, applied to a cloud server, is provided in an embodiment of the present disclosure. The method includes the following steps S11 to S13.

[0036] In step S11, by using a stream pulling-pushing service network, a real-time audio-video stream in an offline singing scenario is pulled from a content delivery network, progress marking is performed on the real-time audio-video stream to obtain a current marked progress, and the real-time audio-video stream and the current marked progress are pushed to a real-time audio-video service network.

[0037] In the embodiment, after pulling the real-time audio-video stream in the offline singing scenario from the content delivery network, the stream pulling-pushing service network performs progress marking on the real-time audio-video stream to obtain the current marked progress, and pushes the real-time audio-video stream and the current marked progress to the real-time audio-video service network. In an implementation, the real-time audio-video stream and the current marked progress are pushed to the real-time audio-video service network as two independent pieces of data. In another implementation, the real-time audio-video stream carries the current marked progress, and the real-time audio-video stream and the current marked progress are pushed as a whole to the real-time audio-video service network.

[0038] It should be noted that the main function of the stream pulling-pushing service network is pulling and pushing content. The stream pulling-pushing service network according to the embodiments of the present disclosure further provides a progress marking function for performing progress marking on the pulled real-time audio-video stream.

[0039] Furthermore, it should be noted that an audio-video acquisition device and a stream pushing device are arranged in the offline singing scenario, to acquire video images and sounds in real time at an event site to obtain an audio-video stream, and to push the audio-video stream to the content delivery network. In related solutions, after obtaining the real-time audio-video stream in the offline singing scenario, the content delivery network may perform subsequent content distribution. That is, the real-time audio-video stream is pushed to various content delivery network nodes, so that the content delivery network nodes deliver the real-time audio-video stream to online audience user terminals upon receiving data pulling requests from the online audience user terminals, thereby enabling users to watch an offline concert in real time on the online audience user terminals.

[0040] In an implementation, by using the stream pulling-pushing service network, pulling the real-time audio-video stream in the offline singing scenario from the content delivery network and performing progress marking on the real-time audio-video stream to obtain the current marked progress includes: by using the stream pulling-pushing service network, continuously pulling the real-time audio-video stream in the offline singing scenario from the content delivery network, and performing progress marking, upon a current time interval reaching a preset time interval threshold, on the real-time audio-video stream based on a preset progress sequence number to obtain the current marked progress. That is, in the embodiment, the stream pulling-pushing service network continuously pulls the real-time audio-video stream in the offline singing scenario from the content delivery network, and performs progress marking on the pulled real-time audio-video stream. In an embodiment, each time when the current time interval reaches the preset time interval threshold, the stream pulling-pushing service network performs progress marking on the currently pulled real-time audio-video stream based on the preset progress sequence number to obtain the current marked progress. For example, it is assumed that 20-ms data is a frame and the preset time interval threshold is set to 20ms. It can be understood that the stream pulling-pushing service network performs progress marking on the currently pulled real-time audio-video stream every 20ms. In an implementation, a progress sequence number may be determined based on the preset time interval threshold. For example, assuming that a time instant when the real-time audio-video stream is first pulled is determined as a starting time instant 0 and progress is continuously accumulated every 20ms thereafter, progress sequence numbers of 0, 20, 40, 60, ... are determined. Alternatively, the progress sequence number may be determined

based on the number of marking operations. For example, each time a progress marking operation is performed, the progress sequence number is incremented by 1, thereby obtaining progress sequence numbers 1, 2, 3, The manner of setting the progress sequence number is not limited in the embodiment.

5 **[0041]** In step S12, by using the real-time audio-video service network, the real-time audio-video stream and the current marked progress are delivered to a target chorus user terminal. The target chorus user terminal plays the real-time audio-video stream and marks, based on the current marked progress, a human voice audio stream acquired in real time through a local online chorus application, to obtain a marked human voice audio stream
10 carrying progress information.

[0042] In the embodiment, the real-time audio-video service network delivers the real-time audio-video stream and the current marked progress pushed by the stream pulling-pushing service network to the target chorus user terminal, so that the target chorus user terminal plays the acquired real-time audio-video stream and marks, based on the current marked progress,
15 the human voice audio stream acquired in real time through the local online chorus application, to obtain the marked human voice audio stream carrying the progress information. It can be understood that the offline real-time audio-video stream is associated with the online human voice audio stream through marking the locally acquired human voice audio stream based on the current marked progress.

20 **[0043]** The local online chorus application may be application software such as WeSing or JOOX Music, and users can freely click to join and exit a chorus.

[0044] Further, the above method includes: acquiring an online human voice audio stream uploaded by a chorus user terminal other than the target chorus user terminal; and forwarding, by using the real-time audio-video service network, the online human voice audio stream
25 uploaded by the chorus user terminal to the target chorus user terminal, so that the target chorus user terminal simultaneously plays the real-time audio-video stream and the online human voice audio stream uploaded by the chorus user terminal. That is, in addition to playing the offline real-time audio-video stream, the target chorus user terminal simultaneously plays the online human voice audio stream uploaded by the chorus user
30 terminal other than the target chorus user terminal, meaning that a chorus user can hear voices of both an offline chorus user and an online chorus user. In an implementation, an online

human voice audio stream of the online chorus user is pulled through a low-latency RTC channel of the real-time audio-video service network to the chorus user and is played by the chorus user. That is, an audio stream is exchanged between terminal users participating in the chorus through the low-latency RTC channel of the real-time audio-video service network.

5 **[0045]** In step S13, by using the real-time audio-video service network, the marked human voice audio stream uploaded by the target chorus user terminal is acquired; by using a real-time audio-video stream pulling service network, the real-time audio-video stream, the current marked progress, and the marked human voice audio stream are pulled to a cloud stream mixing service network; and by using the cloud stream mixing service network, the
10 real-time audio-video stream and the marked human voice audio stream are mixed based on the current marked progress and the progress information, and the mixed audio-video stream is pushed to the content delivery network for distribution.

[0046] In the embodiment, after the target chorus user terminal uploads the marked human voice audio stream to the real-time audio-video service network, the real-time audio-video
15 stream pulling service network pulls the real-time audio-video stream, the current marked progress, and the marked human voice audio stream to the cloud stream mixing service network. Since both the real-time audio-video stream and the human voice audio stream are currently marked with the progress, the cloud stream mixing service network accurately mixes the real-time audio-video stream and the marked human voice audio stream based on
20 the current marked progress and the progress information to obtain the mixed audio-video stream, thereby achieving accurate audio-video stream alignment. Finally, the cloud stream mixing service pushes the mixed audio-video stream to the content delivery network for distribution, thereby enabling other audiences to hear chorus of offline and online users.

[0047] It can be seen that a method for offline and online real-time chorus, applied to a
25 cloud server, is provided in the present disclosure. The method includes: by using a stream pulling-pushing service network, pulling a real-time audio-video stream in an offline singing scenario from a content delivery network, performing progress marking on the real-time audio-video stream to obtain a current marked progress, and pushing the real-time audio-video stream and the current marked progress to a real-time audio-video service
30 network; delivering, by using the real-time audio-video service network, the real-time audio-video stream and the current marked progress to a target chorus user terminal, where

the target chorus user terminal plays the real-time audio-video stream and marks, based on the current marked progress, a human voice audio stream acquired in real time through a local online chorus application to obtain a marked human voice audio stream carrying progress information; acquiring, by using the real-time audio-video service network, the marked human voice audio stream uploaded by the target chorus user terminal; pulling, by using a real-time audio-video stream pulling service network, the real-time audio-video stream, the current marked progress, and the marked human voice audio stream to a cloud stream mixing service network; and by using the cloud stream mixing service network, mixing the real-time audio-video stream and the marked human voice audio stream based on the current marked progress and the progress information, and pushing the mixed audio-video stream to the content delivery network for distribution.

[0048] It can be seen that after pulling the real-time audio-video stream in the offline singing scenario from the content delivery network, the stream pulling-pushing service network performs progress marking on the real-time audio-video stream to obtain the current marked progress, and then pushes the real-time audio-video stream and the current marked progress to the real-time audio-video service network. That is, the stream pulling-pushing service network according to the present disclosure provides functions of pulling, performing progress marking, and pushing audio-video streams. Further, the real-time audio-video service network delivers the real-time audio-video stream and the current marked progress to the target chorus user terminal, so that the target chorus user terminal plays the acquired real-time audio-video stream, and marks, based on the current marked progress, the human voice audio stream acquired in real time through the local online chorus application to obtain the marked human voice audio stream carrying the progress information. It can be understood that through marking the locally acquired human voice audio stream based on the current marked progress, the offline real-time audio-video stream is associated with the online human voice audio stream. After the target chorus user terminal uploads the marked human voice audio stream to the real-time audio-video service network, the real-time audio-video stream pulling service network pulls the real-time audio-video stream, the current marked progress, and the marked human voice audio stream to the cloud stream mixing service network. Since both the real-time audio-video stream and the human voice audio stream are currently marked with the progress, the cloud stream mixing service network accurately mixes the real-time audio-video stream and the marked human voice audio stream based on the current marked progress and

the progress information to obtain the mixed audio-video stream, and finally pushes the mixed audio-video stream to the content delivery network for distribution. Through the above solution, real-time chorus between online users and offline users is achieved.

[0049] Referring to FIG. 3, a method for offline and online real-time chorus is provided in an embodiment of the present disclosure. Compared with the forgoing embodiment, the technical solution according to this embodiment is further described and optimized. The method includes the following steps S21 to S25.

[0050] In step S21, upon a current time interval reaching a preset time interval threshold, by using the stream pulling-pushing service network, the real-time audio-video stream within the preset time interval threshold is pulled from a current target audio-video stream in chronological order. The current target audio-video stream is an audio-video stream that is not pulled from the content delivery network in the offline singing scenario.

[0051] In the embodiment, the stream pulling-pushing service network may intermittently pull the real-time audio-video stream in the offline singing scenario from the content delivery network. However, to ensure real-time performance and reduce latency, the preset time interval threshold is preferably set to a small value. For example, it may be configured that the real-time audio-video stream is pulled from the content delivery network every 20ms. Therefore, upon the current time interval reaching the preset time interval threshold, the stream pulling-pushing service network pulls the real-time audio-video stream within the preset time interval threshold from the current target audio-video stream in chronological order.

[0052] It can be understood that, to avoid pulling duplicate audio-video streams and to ensure continuity of audio-video streams, in the embodiment, it is required that the real-time audio-video stream within the preset time interval threshold is pulled in chronological order from an audio-video stream that is not pulled from the content delivery network. For example, assuming that the stream pulling-pushing service network pulls a 20-ms audio-video stream from the content delivery network each time and an audio-video stream is recorded in milliseconds from 0ms to 100ms, content from 0ms to 19ms is pulled first, then content from 20ms to 39ms is pulled, and so forth.

[0053] In an implementation, upon the current time interval reaching the preset time interval threshold, pulling, by using the stream pulling-pushing service network, the real-time

audio-video stream within the preset time interval threshold from the current target audio-video stream in chronological order includes: determining, based on a first stream pulling timestamp and the preset time interval threshold, a target stream pulling timestamp corresponding to the current stream pulling operation, where the first stream pulling timestamp is a time instant when the stream pulling-pushing service network first pulls the real-time audio-video stream in the offline singing scenario from the content delivery network for the first stream pulling timestamp; and upon a current time instant reaching the target stream pulling timestamp, pulling, by using the stream pulling-pushing service network, the real-time audio-video stream within the preset time interval threshold from the current target audio-video stream in chronological order.

[0054] That is, since the stream pulling-pushing service network implements a stream pulling operation every preset time interval threshold, after the time instant (that is, the first stream pulling timestamp) when the stream pulling-pushing service network first pulls the real-time audio-video stream in the offline singing scenario from the content delivery network is determined, the target stream pulling timestamp corresponding to the current stream pulling operation is determined based on the first stream pulling timestamp and the preset time interval threshold. In this way, upon the current time instant reaching the target stream pulling timestamp, the stream pulling-pushing service network pulls, from the current target audio-video stream, the real-time audio-video stream within the preset time interval threshold in chronological order.

[0055] In step S22, a current progress sequence number corresponding to a current stream pulling operation is determined, progress marking is performed on the real-time audio-video stream based on the current progress sequence number to obtain a current marked progress, and the real-time audio-video stream and the current marked progress are pushed to the real-time audio-video service network.

[0056] In the embodiment, each time the stream pulling-pushing service network implements a stream pulling operation to pull a real-time audio-video stream, the stream pulling-pushing service network performs progress marking on the real-time audio-video stream based on the current progress sequence number corresponding to the stream pulling operation to obtain the current marked progress and then pushes the real-time audio-video stream and the current marked progress to the real-time audio-video service network.

[0057] In an implementation, the determining the current progress sequence number corresponding to the current stream pulling operation, and performing progress marking on the real-time audio-video stream based on the current progress sequence number to obtain the current marked progress includes: determining the current progress sequence number
5 corresponding to the current stream pulling operation based on the preset time interval threshold and a historical marked progress corresponding to a previous stream pulling timestamp, where in a case that the current stream pulling operation is a first stream pulling operation implemented first by the stream pulling-pushing service network, the current progress sequence number is a preset initial value; and determining the current progress
10 sequence number as the current marked progress of the real-time audio-video stream. That is, in the embodiment, a marking progress sequence number is determined based on a stream pulling timestamp. Specifically, the historical marked progress corresponding to the previous stream pulling timestamp is determined, and then the preset time interval threshold is added thereto to obtain the current progress sequence number corresponding to the current stream
15 pulling operation. For example, assuming that the historical marked progress corresponding to the previous stream pulling timestamp is 40 and the preset time interval threshold for stream pulling is 20, the current progress sequence number corresponding to the current stream pulling operation is 60. When the preset time interval threshold for stream pulling remains unchanged, marking progress is gradually incremented at a rate of 20ms over time. In a case
20 that the current stream pulling operation is the first stream pulling operation, the current progress sequence number is the preset initial value which, for example, may be set to 0.

[0058] In step S23, by using the real-time audio-video service network, the real-time audio-video stream and the current marked progress are delivered to a target chorus user terminal. The target chorus user terminal plays the real-time audio-video stream and marks,
25 based on the current marked progress, a human voice audio stream acquired in real time through a local online chorus application, to obtain a marked human voice audio stream carrying progress information.

[0059] In step S24, by using the real-time audio-video service network, the marked human voice audio stream uploaded by the target chorus user terminal is acquired; by using a
30 real-time audio-video stream pulling service network, the real-time audio-video stream, the current marked progress, and the marked human voice audio stream are pulled to a cloud stream mixing service network; and by using the cloud stream mixing service network, the

real-time audio-video stream and the marked human voice audio stream are mixed based on the current marked progress and the progress information.

[0060] In step S25, the mixed audio-video stream is pushed to content delivery network nodes in the content delivery network, to cause one content delivery network node among the content delivery network nodes to deliver the mixed audio-video stream to an online audience user terminal upon receiving a data pulling request from the online audience user terminal. The online audience user terminal selects the one content delivery network node from the content delivery network nodes based on current network communication conditions and sends the data pulling request to the selected content delivery network node.

[0061] In the embodiment, after mixing the real-time audio-video stream and the marked human voice audio stream, the cloud stream mixing service network pushes the mixed audio-video stream to the content delivery network nodes of the content delivery network. Upon receiving the data pulling request from the online audience user terminal, the one content delivery network node among the content delivery network nodes delivers the mixed audio-video stream to the online audience user terminal. It can be understood that the online audience user terminal selects the one content delivery network node from the content delivery network nodes based on the current network communication conditions and sends the data pulling request to the selected content delivery network node. In an embodiment, the online audience user terminal pulls the mixed audio-video stream through the content delivery network node closest to the online audience user terminal, thereby enabling online audiences to hear chorus between offline users and online users.

[0062] Specific processing procedures of steps S23 and S24 described above may be referred to corresponding descriptions provided in foregoing embodiments, and details are not repeated herein.

[0063] It can be seen that in the embodiments of the present disclosure, the stream pulling-pushing service network may implement one stream pulling operation and one progress marking operation every preset time interval threshold. In an embodiment, since the stream pulling-pushing service network implements one stream pulling operation every preset time interval threshold, after the first stream pulling timestamp of the stream pulling-pushing service network is determined, the target stream pulling timestamp corresponding to each stream pulling operation is determined based on the first stream pulling timestamp and the

preset time interval threshold. In this way, one stream pulling operation is implemented upon the current time instant reaching the target stream pulling timestamp. In the process, to avoid pulling duplicate audio-video streams and to ensure continuity of audio-video streams, in the embodiment, the real-time audio-video stream within the preset time interval threshold is pulled in chronological order from an audio-video stream that is not pulled from the content delivery network. In addition, the online audience user terminal pulls the mixed audio-video stream through the content delivery network node closest to the online audience user terminal, thereby enabling online audiences to hear chorus between offline users and online users.

[0064] Referring to FIG. 4, a method for offline and online real-time chorus, applied to a target chorus user terminal, is provided in an embodiment of the present disclosure. The method includes the following steps S31 to S33.

[0065] In step S31, from a real-time audio-video service network in a cloud server, a real-time audio-video stream in an offline singing scenario is pulled and played, and a current marked progress of the real-time audio-video stream is acquired, where by using a stream pulling-pushing service network in the cloud server, the real-time audio-video stream is pulled from a content delivery network, and progress marking is performed on the real-time audio-video stream to obtain the current marked progress.

[0066] In the embodiment, the target chorus user terminal is a user terminal participating in chorus online. The target chorus user terminal pulls and plays the real-time audio-video stream in the offline singing scenario, and acquires the current marked progress of the real-time audio-video stream from the real-time audio-video service network in the cloud server. It should be noted that by using the stream pulling-pushing service network in the cloud server, the real-time audio-video stream is pulled from the content delivery network, progress marking is performed on the real-time audio-video stream to obtain the current marked progress, and then the real-time audio-video stream and the current marked progress are pushed to the real-time audio-video service network.

[0067] In step S32, a human voice audio stream acquired in real time through a local online chorus application is marked based on the current marked progress, to obtain a marked human voice audio stream carrying progress information.

[0068] In the embodiment, the target chorus user terminal marks, based on the acquired current marked progress, the human voice audio stream acquired by the local online chorus

application from a local audio acquisition device, to obtain the marked human voice audio stream carrying the progress information. The local audio acquisition device may be a microphone.

[0069] In an implementation, the marking, based on the current marked progress, the human voice audio stream acquired in real time through the local online chorus application includes: activating the local audio acquisition device through the local online chorus application to acquire local audio data by using the local audio acquisition device, and denoising the local audio data to obtain the human voice audio stream; and marking, based on the current marked progress, the human voice audio stream acquired through the local online chorus application. It can be understood that after a user confirms participation in chorus, the local online chorus application activates the local audio acquisition device, that is, the microphone, to acquire the local audio data by using the microphone. In the process, the denoising may be performed on the local audio data to remove ambient sounds and noise to obtain the human voice audio stream. Then, the human voice audio stream acquired through the local online chorus application is marked based on the current marked progress. Online human voice data acquired locally is marked based on first progress information, and the human voice audio stream is marked based on the current marked progress, so that the offline real-time audio-video stream is associated with the online human voice audio stream.

[0070] In step S33, the marked human voice audio stream is uploaded to the real-time audio-video service network; by using a real-time audio-video stream pulling service network, the real-time audio-video stream, the current marked progress, and the marked human voice audio stream are sent from the real-time audio-video service network to a cloud stream mixing service network; and by using the cloud stream mixing service network, the real-time audio-video stream and the marked human voice audio stream are mixed based on the current marked progress and the progress information, and the mixed audio-video stream is pushed to the content delivery network for distribution.

[0071] In the embodiment, the target chorus user terminal uploads the marked human voice audio stream to the real-time audio-video service network. Further, the real-time audio-video stream pulling service network simultaneously pulls the real-time audio-video stream, the current marked progress, and the marked human voice audio stream from the real-time audio-video service network to the cloud stream mixing service network. The cloud stream

mixing service network accurately mixes the real-time audio-video stream and the marked human voice audio stream based on the current marked progress and the progress information to obtain the mixed audio-video stream, thereby achieving precise alignment of the offline real-time audio-video stream and the online human voice audio stream. Finally, the cloud stream mixing service network pushes the mixed audio-video stream to the content delivery network for distribution, thereby enabling other audiences to hear chorus between online users and offline users.

[0072] It can be seen that the target chorus user terminal, that is, the user terminal participating in chorus online, pulls and plays the real-time audio-video stream in the offline singing scenario, and acquires the current marked progress of the real-time audio-video stream from the real-time audio-video service network in the cloud server. Further, the target chorus user terminal marks, based on the acquired current marked progress, the human voice audio stream acquired through the local online chorus application from the local audio acquisition device, to obtain the marked human voice audio stream carrying the progress information. Finally, the target chorus user terminal uploads the marked human voice audio stream to the real-time audio-video service network. Further, the real-time audio-video stream pulling service network simultaneously pulls the real-time audio-video stream, the current marked progress, and the marked human voice audio stream from the real-time audio-video service network to the cloud stream mixing service network. The cloud stream mixing service network accurately mixes the real-time audio-video stream and the marked human voice audio stream based on the current marked progress and the progress information to obtain the mixed audio-video stream, thereby achieving precise alignment of the offline real-time audio-video stream and the online human voice audio stream. Finally, the cloud stream mixing service network pushes the mixed audio-video stream to the content delivery network for distribution, thereby enabling other audiences to hear chorus between online users and offline users.

[0073] FIG. 5 is a schematic structural diagram of an electronic device according to an embodiment of the present disclosure. The electronic device may include at least one processor 21, at least one memory 22, a power supply 23, a communication interface 24, an input/output interface 25, and a communication bus 26. The memory 22 is configured to store a computer program. The computer program is loaded and executed by the processor 21, to perform the method for offline and online real-time chorus executed by the electronic device

as described in any one of the above embodiments.

[0074] In the embodiment, the power supply 23 is configured to provide an operating voltage to hardware devices of the electronic device 20. The communication interface 24 is capable of creating a data transmission channel between the electronic device 20 and an external device. The communication interface 24 uses any communication protocol applicable to the technical solution of the present disclosure, which is not specifically limited herein. The input/output interface 25 is configured to receive data from the external or output data to the external. An interface type of the input/output interface 25 may be determined based on actual requirements, which is not specifically limited herein.

[0075] The processor 21 may include one or more processing cores, such as a 4-core processor, an 8-core processor, or the like. The processor 21 may be implemented in at least one hardware form of digital signal processing (DSP), a field-programmable gate array (FPGA), and a programmable logic array (PLA). The processor 21 may include a main processor and a coprocessor. The main processor is configured to process data in an active state, which is also referred to as a central processing unit (CPU). The coprocessor is a low-power processor and is configured to process data in a standby state. In some embodiments, the processor 21 may be integrated with a graphics processing unit (GPU). The GPU is configured to render and draw a content to be displayed on a display screen. In some embodiments, the processor 21 may further include an artificial intelligence (AI) processor. The AI processor is configured to process computing operations related to machine learning.

[0076] In addition, the memory 22 serves as a carrier for resource storage, and may be a read-only memory, a random access memory, a magnetic disk, an optical disk, or the like. Resources stored on the memory may include an operating system 221, a computer program 222, data 223, and the like, which may be temporarily stored or permanently stored.

[0077] The operating system 221 is configured to manage and control the hardware devices of the electronic device 20 and the computer program 222, in order to enable the processor 21 to perform computing and processing on the massive data 223 in the memory 22. The operating system 221 may be Windows, Unix, Linux, or the like. In addition to the computer program configured to perform the method for offline and online real-time chorus executed by the electronic device 20 according to any one of the above embodiments, the computer program 222 may further include a computer program configured to implement other specific

tasks. The data 223 may include data received by the electronic device from the external device and data acquired by the input/output interface 25 of the electronic device.

[0078] Further, a computer-readable storage medium is provided according to an embodiment of the present disclosure. A computer program is stored in the storage medium and is loaded and executed by a processor, to perform the method for offline and online real-time chorus described in any one of the above embodiments.

[0079] The above embodiments are described in a progressive manner. Each of the embodiments is mainly focused on describing its differences from other embodiments, and reference may be made among these embodiments with respect to the same or similar parts.

[0080] Those skilled in the art should further appreciate that the units and algorithm steps of the examples described in conjunction with the embodiments disclosed herein may be implemented by electronic hardware, computer software, or a combination thereof. In order to clearly illustrate inter-replacement between the hardware and software, the components and steps of each example have been generally described according to their functions in the above description. Whether these functions are executed by hardware or software depends on the specific applications and design constraints of the technical solution. Those skilled in the art may implement the described function by using different methods for each particular application, but such implementation should not be considered as beyond the scope of the present disclosure.

[0081] Steps of the method or algorithm described in the embodiments herein may be directly implemented by hardware, a software module executed by a processor, or a combination thereof. The software module may be arranged in a random access memory (RAM), a memory, a read only memory (ROM), an electrically programmable ROM, an electrically erasable programmable ROM, a register, a hard disk, a removable disk, a compact disc read-only memory (CD-ROM), or any other forms of storage medium known in the art.

[0082] It should be further noted that relation terms such as "first" and "second" herein are only used to distinguish one entity or operation from another entity or operation, and does not necessarily require or imply that there is an actual relation or sequence between these entities or operations. Furthermore, the term "comprise", "include", or any other variation thereof is intended to cover a non-exclusive inclusion, so that a process, a method, an article, or a device including a set of elements includes not only those elements, but also other elements not

expressly listed or elements inherent in such process, method, article, or device. In addition, unless expressively limited otherwise, the statement "comprising (including) a(n)..." does not exclude existence of other identical elements in the process, method, article or device.

[0083] The method for offline and online real-time chorus, the device, and the storage
5 medium according to the present disclosure are introduced in detail above. The specific
examples are used herein to illustrate the principles and implementations of the present
disclosure. Description of the above embodiments is only for helping understanding of the
method and core idea thereof of the present disclosure. In addition, modifications may be
made by those of ordinary skill in the art on a specific implementation and application scope
10 based on the idea of the present disclosure. In summary, the description should not be
construed as imposing limitations to the present disclosure.

CLAIMS

1. A method for offline and online real-time chorus, applied to a cloud server, comprising:

5 by using a stream pulling-pushing service network, pulling a real-time audio-video stream in an offline singing scenario from a content delivery network, performing progress marking on the real-time audio-video stream to obtain a current marked progress, and pushing the real-time audio-video stream and the current marked progress to a real-time audio-video service network;

10 delivering, by using the real-time audio-video service network, the real-time audio-video stream and the current marked progress to a target chorus user terminal device, wherein the target chorus user terminal device plays the real-time audio-video stream and marks, based on the current marked progress, a human voice audio stream acquired in real time through a local online chorus application to obtain a marked human voice audio stream carrying progress
15 information;

acquiring, by using the real-time audio-video service network, the marked human voice audio stream uploaded by the target chorus user terminal device;

pulling, by using a real-time audio-video stream pulling service network, the real-time audio-video stream, the current marked progress, and the marked human voice audio stream
20 to a cloud stream mixing service network; and

by using the cloud stream mixing service network, mixing the real-time audio-video stream and the marked human voice audio stream based on the current marked progress and the progress information, and pushing the mixed audio-video stream to the content delivery network for distribution.

25 2. The method for offline and online real-time chorus according to claim 1, wherein, by using the stream pulling-pushing service network, pulling the real-time audio-video stream in the offline singing scenario from the content delivery network and performing progress marking on the real-time audio-video stream to obtain the current marked progress comprises:

30 upon a current time interval reaching a preset time interval threshold, pulling, by using

the stream pulling-pushing service network, the real-time audio-video stream within the preset time interval threshold from a current target audio-video stream in chronological order, wherein the current target audio-video stream is an audio-video stream that is not pulled from the content delivery network in the offline singing scenario; and

- 5 determining a current progress sequence number corresponding to a current stream pulling operation, and performing progress marking on the real-time audio-video stream based on the current progress sequence number to obtain the current marked progress.

3. The method for offline and online real-time chorus according to claim 2, wherein,
10 upon the current time interval reaching the preset time interval threshold, pulling, by using the stream pulling-pushing service network, the real-time audio-video stream within the preset time interval threshold from the current target audio-video stream in chronological order comprises:

15 determining, based on a first stream pulling timestamp and the preset time interval threshold, a target stream pulling timestamp corresponding to the current stream pulling operation, wherein the first stream pulling timestamp is a time instant when the stream pulling-pushing service network first pulls the real-time audio-video stream in the offline singing scenario from the content delivery network for the first stream pulling timestamp; and

20 upon a current time instant reaching the target stream pulling timestamp, pulling, by using the stream pulling-pushing service network, the real-time audio-video stream within the preset time interval threshold from the current target audio-video stream in chronological order.

4. The method for offline and online real-time chorus according to claim 2, wherein the
25 determining the current progress sequence number corresponding to the current stream pulling operation, and performing progress marking on the real-time audio-video stream based on the current progress sequence number to obtain the current marked progress comprises:

30 determining the current progress sequence number corresponding to the current stream pulling operation based on the preset time interval threshold and a historical marked progress

corresponding to a previous stream pulling timestamp, wherein in a case that the current stream pulling operation is a first stream pulling operation implemented first by the stream pulling-pushing service network, the current progress sequence number is a preset initial value; and

5 determining the current progress sequence number as the current marked progress of the real-time audio-video stream.

10 5. The method for offline and online real-time chorus according to claim 1, wherein, by using the stream pulling-pushing service network, pulling the real-time audio-video stream in the offline singing scenario from the content delivery network and performing progress marking on the real-time audio-video stream to obtain the current marked progress comprises:

15 by using the stream pulling-pushing service network, continuously pulling the real-time audio-video stream in the offline singing scenario from the content delivery network, and performing progress marking, upon a current time interval reaching a preset time interval threshold, on the real-time audio-video stream based on a preset progress sequence number to obtain the current marked progress.

 6. The method for offline and online real-time chorus according to claim 1, further comprising:

20 acquiring an online human voice audio stream uploaded by a chorus user terminal device other than the target chorus user terminal device; and

25 forwarding, by using the real-time audio-video service network, the online human voice audio stream uploaded by the chorus user terminal device to the target chorus user terminal device, to cause the target chorus user terminal device to simultaneously play the real-time audio-video stream and the online human voice audio stream uploaded by the chorus user terminal device.

 7. The method for offline and online real-time chorus according to any one of claims 1 to 6, wherein the pushing the mixed audio-video stream to the content delivery network for

distribution comprises:

pushing the mixed audio-video stream to content delivery network nodes in the content delivery network, to cause one content delivery network node among the content delivery network nodes to deliver the mixed audio-video stream to an online audience user terminal device upon receiving a data pulling request from the online audience user terminal device, wherein the online audience user terminal device selects the one content delivery network node from the content delivery network nodes based on current network communication conditions and sends the data pulling request to the selected content delivery network node.

8. A method for offline and online real-time chorus, applied to a target chorus user terminal device, comprising:

from a real-time audio-video service network in a cloud server, pulling and playing a real-time audio-video stream in an offline singing scenario, and acquiring a current marked progress of the real-time audio-video stream, wherein by using a stream pulling-pushing service network in the cloud server, the real-time audio-video stream is pulled from a content delivery network, and progress marking is performed on the real-time audio-video stream to obtain the current marked progress;

marking, based on the current marked progress, a human voice audio stream acquired in real time through a local online chorus application, to obtain a marked human voice audio stream carrying progress information;

uploading the marked human voice audio stream to the real-time audio-video service network;

sending, by using a real-time audio-video stream pulling service network, the real-time audio-video stream, the current marked progress, and the marked human voice audio stream from the real-time audio-video service network to a cloud stream mixing service network; and

by using the cloud stream mixing service network, mixing the real-time audio-video stream and the marked human voice audio stream based on the current marked progress and the progress information, and pushing the mixed audio-video stream to the content delivery network for distribution.

9. The method for offline and online real-time chorus according to claim 8, wherein the marking, based on the current marked progress, the human voice audio stream acquired in real time through the local online chorus application comprises:

5 activating a local audio acquisition device through the local online chorus application to acquire local audio data by using the local audio acquisition device, and denoising the local audio data to obtain the human voice audio stream; and

 marking, based on the current marked progress, the human voice audio stream acquired through the local online chorus application.

10 10. An electronic device, comprising:

 a memory, configured to store a computer program; and

 a processor, configured to perform, when executing the computer program, the method for offline and online real-time chorus according to any one of claims 1 to 9.

15 11. A computer-readable storage medium, configured to store a computer program, wherein the computer program, when being executed by a processor, causes the processor to perform the method for offline and online real-time chorus according to any one of claims 1 to 9.