

ONLINE KARAOKE ROOM IMPLEMENTATION METHOD, ELECTRONIC DEVICE, AND COMPUTER-READABLE STORAGE MEDIUM

[0001] This application claims the priority to the Chinese Patent Application No.
5 202011357981.7, titled "METHOD FOR IMPLEMENTING ONLINE KARAOKE ROOM,
ELECTRONIC DEVICE, AND COMPUTER-READABLE STORAGE MEDIUM", filed on
November 27, 2020 with the China National Intellectual Property Administration, the content
of which is incorporated herein by reference.

10 **TECHNICAL FIELD**

[0002] The present disclosure relates to the field of computer technology, and in particular,
to a method for implementing an online karaoke room, an electronic device, and a
computer-readable storage medium.

15 **BACKGROUND**

[0003] Conventional designs of online karaoke rooms utilize an asynchronous manner to
generate a chorus of two users. That is, user A first records his or her singing part on
terminal A, then transmits a synthesized work to terminal B, and then user B insert his or her
singing part via the terminal B to complete a final chorus work.

20 [0004] Accordingly, it is discovered that the conventional technology has at least an issue of
being incapable of achieving real-time antiphonal singing among multiple accounts.

SUMMARY

[0005] An objective of embodiments of the present disclosure is providing a method for
25 implementing an online karaoke room, an electronic device, and a computer-readable storage
medium. Real-time antiphonal singing can be achieved among multiple accounts.

[0006] In order to achieve the above objective, a method for implementing an online
karaoke room is provided according to a first aspect of the present disclosure. A first
account corresponding to a first terminal and a second account corresponding to a second

client are matched with a virtual room. The method comprises: in singing duration of the first account, playing locally stored first audio content by the first terminal, transmitting, by the first terminal, a first target audio and progress information of the first target audio to the second terminal, and playing the first target audio by the second terminal, where the first
5 audio content is audio content which is from an accompaniment audio and corresponds to the first account, and the first target audio is an audio composite of the first audio content locally stored at the first terminal and a first dry-sound audio collected by the first terminal; in singing duration of the second account, determining, by the second terminal, a current play progress according to the progress information received by the second terminal, starting, by
10 the second terminal, playing locally stored second audio content based on the current play progress, and transmitting a second target audio and progress information of the second target video to the first terminal, where the second audio content is audio content which is from the accompaniment audio and corresponds to the second account, and the second target audio is an audio composite of the second audio content locally stored at the second terminal and a
15 second dry-sound audio collected by the second terminal; and in the singing period of the second account, playing the second target audio by the first terminal in response to the first terminal receiving the second target audio transmitted by the second terminal.

[0007] In order to achieve the above objective, an electronic device is provided according to a second aspect of the present disclosure. The electronic device comprises: a memory
20 storing a computer program; and a processor, configured to execute the computer program to implement steps in the foregoing method which are performed by the first terminal or the second terminal.

[0008] In order to achieve the above objective, a computer-readable storage medium is provided according to a third aspect of the present disclosure. The computer-readable
25 storage medium stores a computer program, and the computer program when executed by a processor implements steps in the foregoing method which are performed by the first terminal or the second terminal.

[0009] Accordingly, the method for implementing the online karaoke room is provided according to embodiments of the present disclosure. The first account corresponding to the
30 first terminal and the second account corresponding to the second client are matched with the virtual room, and the method comprises following steps. In the singing duration of the first

account, the first terminal plays the first audio content that is locally stored, and transmits the first target audio and the progress information of the first target audio to the second terminal, and the second terminal plays the first target audio. The first audio content is from the accompaniment audio and corresponds to the first account, and the first target audio is the audio composite of the first audio content locally stored at the first terminal and the first dry-sound audio collected by the first terminal. In the singing duration of the second account, the second terminal determines the current play progress according to the received progress information, starts playing the second audio content, which is locally stored, based on the current play progress, and transmits the second target audio and the progress information of the second target video to the first terminal. The second audio content is from the accompaniment audio and corresponds to the second account, and the second target audio is the audio composite of the second audio content locally stored at the second terminal and the second dry-sound audio collected by the second terminal. In the singing period of the second account, the first terminal plays the second target audio by in response to receiving the second target audio sent by the second terminal.

[0010] Herein real-time alternate singing, i.e., an antiphonal singing mode, between the first account and the second account can be achieved in the virtual room when the first account and the second account are matched with the same virtual room. The first account corresponds to the first audio content, and the second account corresponds to the second audio content. When a user logging into the first account sings, the accompaniment audio played by the first terminal is the first audio content that is stored locally, and the audio played by the second terminal is the audio composite of the first audio content stored at the first terminal and the first dry-sound audio collected by the first terminal, i.e., the first target audio. Alignment between the accompaniment audio and the dry-sound audio is guaranteed at a side of the second terminal. When a user logging into the second account sings, the second terminal determines the current play progress according to the received progress information, and starts playing the second audio content, which is stored locally, based on the current play progress. The audio played by the first terminal is the audio composite of the second audio content stored at the second terminal and the second dry-sound audio collected by the second terminal, i.e., the second target audio. Alignment between the accompaniment audio and the dry-sound audio is guaranteed at a side of the first terminal. Thereby, the method of implementing the online karaoke room provided according to embodiments of the present

disclosure achieves the alignment between the accompaniment audio and the dry audio in playback at each of the first terminal and the second terminal, when the first account and the second account perform the antiphonal singing in real time. The electronic device and the computer-readable storage medium disclosed herein can also achieve the foregoing technical effects.

[0011] The foregoing general description and following detailed description are merely exemplary, and the present disclosure is not limited thereto.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] For clearer illustration of the technical solutions according to embodiments of the present disclosure or conventional techniques, hereinafter briefly described are the drawings to be applied in embodiments of the present disclosure or conventional techniques. Apparently, the drawings in the following descriptions are only some embodiments of the present disclosure, and other drawings may be obtained by those skilled in the art based on the provided drawings without creative efforts. The drawings are configured to provide further understanding on the present disclosure and constitute a part of the description. The drawings in conjunction with following embodiments are configured to explain the present disclosure, and do not constitute a limitation to the present disclosure.

[0013] Figure 1 is an architectural diagram of a system for implementing an online karaoke room according to an embodiment of the present disclosure.

[0014] Figure 2 is a flow chart of a method for implementing an online karaoke room according to an embodiment of the present disclosure.

[0015] Figure 3 is a schematic diagram of audio transmission and audio playback according to an embodiment of the present disclosure.

[0016] Figure 4 is a structural diagram of an electronic device according to an embodiment of the present disclosure.

DETAILED DESCRIPTION OF THE EMBODIMENTS

[0017] Hereinafter technical solutions in embodiments of the present disclosure are

described clearly and completely in conjunction with the drawings in embodiments of the present closure. Apparently, the described embodiments are only some rather than all of the embodiments of the present disclosure. Any other embodiments obtained based on the embodiments of the present disclosure by those skilled in the art without any creative effort
5 fall within the scope of protection of the present disclosure.

[0018] Hereinafter a system adopted by a method for implementing the online karaoke room is illustrated to facilitate understanding of the method. Reference is made to Figure 1, which is a structural diagram of a system for implementing an online karaoke room according to an embodiment of the present disclosure. As shown in Figure 1, the system comprises a
10 server 10 and multiple terminals 20, and each terminal 20 is connected with the server 10 via a network.

[0019] The server 10 is configured to create a virtual room and match multiple accounts to the same virtual room. Each virtual room is isolated from each other. The virtual room may simulate a room in the real world, and has isolation characteristics as the real room, so as
15 to implement the online karaoke room herein. When multiple accounts are matched with the same virtual room, the server 10 establishes an audio link between the terminal 20 corresponding to each account and the virtual room. That is, the accounts matched with the same virtual room are capable to communicate with each other in audio through the virtual room.

[0020] The terminal 20 may include a fixed terminal such as a PC (personal computer) and a mobile terminal such as a mobile phone. Each terminal 20 is provided with an audio collection apparatus such as a microphone and an audio output apparatus such as a
20 loudspeaker. The audio collection apparatus is configured to collect a dry-sound audio corresponding to a target song and a voice-call audio, and the audio output apparatus is configured to output an audio composite of a dry-sound audio and an accompaniment audio and output a voice-call audio. The audio link between the terminal 20, which corresponds to an account in a microphone-in-use state, and the virtual room is activated, and such terminal 20 transmits the collected dry-sound audio to the virtual room via the audio link. The audio
25 link between the terminal 20, which corresponds to an account in a microphone-out-of-use state, and the virtual room is deactivated. Activation and deactivation on the audio link
30 between each terminal 20 and the virtual room do not affect a process of playing songs.

[0021] Herein a method for implementing an online karaoke room is provided. Real-time antiphonal singing among multiple accounts can be achieved.

[0022] Reference is made to Figure 2, which is a flow chart of a method for implementing an online karaoke room according to an embodiment of the present disclosure. As shown in
5 Figure 2, the method comprises steps S101 to S103.

[0023] In step S101, in singing duration of the first account, the first terminal plays locally stored first audio content, and transmits a first target audio and progress information of the first target audio to the second terminal, and the second terminal plays the first target audio. The first audio content is from an accompaniment audio and corresponds to the first account,
10 and the first target audio is an audio composite of the first audio content locally stored at the first terminal and a first dry-sound audio collected by the first terminal.

[0024] Herein the first account corresponding to the first terminal and the second account corresponding to the second terminal are matched with the same virtual room. Audio links are established between the virtual room and the first terminal corresponding to the first
15 account and between the virtual room and the second terminal corresponding to the second account, respectively. The audio link implements audio transmission between the first terminal, or the second terminal, and the virtual room. Audio communication between the first account and the second account via the virtual room simulates audio communication in the real world.

[0025] This embodiment concerns a scenario of antiphonal singing between the first
20 account and the second account. In the antiphonal singing mode, both the first account and the second account are in a microphone-in-use state, and users logging to the first account and the second account sing different parts, respectively, of a target song. The first account corresponds to the first audio content, and the second account corresponds to the second audio
25 content. That is, in one embodiment, the method further comprises a following step. In a case that both the first account and the second account are in the microphone-in-use state, the first terminal determines the first audio content corresponding to the first account, and the second terminal determines the second audio content corresponding to the second account. On display interfaces of the first terminal and the second terminal, lyrics corresponding to the
30 first audio content are displayed in a first preset manner, and lyrics corresponding to the second audio content are displayed in a second preset manner. For example, the lyrics

corresponding to different accounts may be displayed in different colors, and a prompt is presented when the singing duration of a certain account is about to start.

[0026] In an embodiment, in the singing duration of the first account, the first terminal plays the first audio content, which is locally stored, as the accompaniment audio, collects the first dry-sound audio simultaneously, synthesizes the first target audio from the locally stored first audio content and the first dry-sound audio, and transmits the first target audio and progress information of the first target audio to the second terminal. The audio played by the second terminal is the first target audio. At the side of the second terminal, alignment between the accompaniment audio and the dry-sound audio is thereby guaranteed.

[0027] The first account may be in an in-ear-monitor mode or a loudspeaker mode. In the in-ear-monitor mode, the first terminal plays the first target audio through an in-ear monitor. That is, a user can hear the first dry-sound audio, which is sung by him or her, through the in-ear-monitor. In the loudspeaker mode, the first terminal plays the accompaniment audio, and needs to perform echo cancellation on an audio collected by its audio collection apparatus to obtain the first dry-sound audio. That is, in an embodiment, the method further comprises a following step. In a case that the first terminal is in the loudspeaker mode, the first terminal performs echo processing on its collected audio to obtain the first dry-sound audio, the second terminal performs echo processing on its collected audio to obtain the second dry-sound audio.

[0028] In step S102, in singing duration of the second account, the second terminal determines a current play progress according to the progress information that is received, starts playing locally stored second audio content based on the current play progress, and transmits a second target audio and progress information of the second target video to the first terminal. The second audio content is from the accompaniment audio and corresponds to the second account, and the second target audio is an audio composite of the second audio content locally stored at the second terminal and a second dry-sound audio collected by the second terminal.

[0029] In an embodiment, in the singing duration of the second account, the second terminal plays the second audio content, which is locally stored, as the accompaniment audio. The second terminal may determine the current play progress according to the previously received progress information of the first target audio, and start playing the locally stored

second audio content from the current play progress. Simultaneously, the second terminal collects the second dry-sound audio, synthesizes the second target audio from the locally stored second audio content and the collected second dry-sound audio, and sends the second target audio and the progress information of the second target audio to the first terminal via a server.

[0030] In step S103, in the singing duration of the second account, the first terminal plays the second target audio in response to receiving the second target audio sent by the second terminal.

[0031] Due to a round-trip delay between the first terminal and the second terminal, the first terminal has not yet received the second target audio sent by the second terminal when finishing playing the first audio content. At such time, the first terminal may continue playing the second audio content that is stored locally. That is, in an embodiment, the method further comprises a following step. In the singing duration of the second account, the first terminal plays the locally stored second audio content or a blank audio, when not receiving the second target audio sent by the second terminal. It is appreciated that the first terminal may play the blank audio when not receiving the second target audio sent by the second terminal, which is not limited herein. Further, the first terminal switches to playing the second target audio, upon receiving the second target audio sent by the second terminal. At the side of the first terminal, alignment between the accompaniment audio and the dry-sound audio is thereby guaranteed.

[0032] As an example, total duration of the target song is 180s, the first audio content occupies 0-30s, 60s-90s, and 120s-150s, the second audio content occupies 30-60s, 90s-120s, 150s-180s, and a round-trip delay between the first terminal and the second terminal is 0.6s. As shown in Figure 3, the black portion refers to playing the audio content that is locally stored, and the white portion refers to the audio transmitted to the counterpart, and the shaded portion refers to playing the audio transmitted by the counterpart. During each of 30s-30.6s, 90.6s-91.2s, and 151.2s-151.8s, the first terminal continues playing the locally stored second audio content or the blank audio. During each of 60.3s-60.9s and 120.9s - 121.5s, the second terminal continues playing the locally stored first audio content or the blank audio.

[0033] Herein the case in which the first account switches to playing what is sung by a user logging into the second account is illustrated. The case in which the second account

switches to playing what is sung by a user logging into the first account is similar, and hence details are not repeated herein.

[0034] In an embodiment, the step S103 comprises following steps. In the singing duration of the second account and before receiving the second target audio, the first terminal
5 applies a fade-out effect on the first audio content which is being played, and applies a fade-in effect on the second target audio in response to receiving the second target audio sent by the second terminal. In an embodiment, the first terminal plays the blank audio when waiting for the second terminal to transmit the second target audio. Accordingly the first terminal applies the fade-out effect on the first audio content when the first audio content is to be
10 played, and applies the fade-in effect on the second target audio content when starting to play the received second target audio, such that the blank audio played by the first terminal when waiting for the second terminal to transmit the second target audio would not be too abrupt. At the side of the first terminal, transition between the first audio content and the second target audio is thereby more natural and smoother.

[0035] Herein real-time alternate singing, i.e., an antiphonal singing mode, between the first account and the second account can be achieved in the virtual room when the first account and the second account are matched with the same virtual room. The first account corresponds to the first audio content, and the second account corresponds to the second audio content. When the user logging into the first account sings, the accompaniment audio played
20 by the first terminal is the first audio content that is stored locally, and the audio played by the second terminal is the audio composite of the first audio content stored at the first terminal and the first dry-sound audio collected by the first terminal, i.e., the first target audio. Alignment between the accompaniment audio and the dry-sound audio is guaranteed at the side of the second terminal. When the user logging into the second account sings, the
25 second terminal determines the current play progress according to the received progress information, and starts playing the second audio content, which is stored locally, based on the current play progress. The audio played by the first terminal is the audio composite of the second audio content stored at the second terminal and the second dry-sound audio collected by the second terminal, i.e., the second target audio. Alignment between the accompaniment
30 audio and the dry-sound audio is guaranteed at the side of the first terminal. Thereby, the method of implementing the online karaoke room provided according to embodiments of the present disclosure achieves the alignment between the accompaniment audio and the dry

audio in playback at each of the first terminal and the second terminal, when the first account and the second account perform the antiphonal singing in real time.

[0036] In an embodiment on a basis of the forgoing embodiments, the method further comprises a following step. A target audio is transmitted to a third terminal corresponding to
5 a third account in the virtual room, such that the third terminal is capable of playing the target audio. The target audio comprises the first target audio transmitted by the first terminal and the second target audio transmitted by the second terminal.

[0037] As a specific implementation, the first terminal transmits the first target audio to the second terminal via the server, and the second terminal transmits the second target audio to the
10 first terminal via the server. That is, the server can obtain the first target audio synthesized at the first terminal audio and obtain the second target audio synthesized at the second terminal. After receiving the target audio, the server transmits the target audio to the third terminal corresponding to the third account, and the third terminal plays the received antiphonal-singing audio in a normal playing mode. Thereby, it is achieved that the
15 antiphonal-singing audio is pushed to the third account.

[0038] In an embodiment on a basis of the forgoing embodiments, the method further comprises a following step. The first terminal and the second terminal download the accompaniment audio of the target song from the server, and store accompaniment audio as
the first audio content and the second audio content, respectively. In a case that both the first
20 account and the second account are in a microphone-out-of-use state, the first terminal and the second terminal play the locally stored accompaniment audio separately.

[0039] As a specific implementation, after any account in the virtual room selects to sing the target song, the terminals corresponding to all accounts in such virtual room download the accompaniment audio of the target song from the server. After the download is completed,
25 the server determines a moment of playing and informs all terminals of such moment, and all terminals start playing the locally stored accompaniment audio at such moment.

[0040] In an embodiment on a basis of the forgoing embodiments, the method further comprises a following step. In a case that the second account is in the microphone-out-of-use state and the first account switches from the microphone-out-of-use
30 state to the microphone-in-use state, the first terminal transmits the first synthesized audio and the progress information of the first target audio to the second terminal, and the second

terminal stops playing the locally stored accompaniment audio and starts playing the first synthesized audio. The first synthesized audio is an audio composite of the accompaniment audio locally stored at the first terminal and the first dry-sound audio collected by the first terminal.

5 **[0041]** As a specific implementation, during playback of the accompaniment audio of the target song, the first account is in a solo mode when switching to the microphone-in-use state. In the solo mode, the first terminal continues playing the locally stored accompaniment audio, and starts collecting the first dry-sound audio. The first terminal mixes the locally stored accompaniment audio and the first dry-sound audio to generate the first synthesized audio,
10 and transmits the first synthesized audio and the progress information thereof to the second terminal via the server. The second terminal stops playing the locally stored accompaniment audio, and starts playing the received first synthesized audio. That is, the accompaniment audio and the dry-sound audio played at the second terminal both come from the first terminal, i.e., the first synthesized audio played by the second terminal is synthesized at the first
15 terminal. At the side of the second terminal, alignment between the accompaniment audio and the dry-sound audio is thereby guaranteed.

[0042] In an embodiment, the second terminal stopping playing the locally stored accompaniment audio and starting playing the first synthesized audio comprises a following step. The second terminal applies the fade-out effect on the locally stored accompaniment
20 audio, and applies the fade-in effect on the first synthesized audio. As a specific implementation, the audio of which playback needs to be stopped can fade out while the audio of which playback needs to be started can fade in, such that switching in audio playing is more natural.

[0043] When the first account is in the microphone-in-use state and the second account is in the microphone-out-of-use state, the first account may be in the in-ear-monitor mode or the
25 loudspeaker mode. In the in-ear-monitor mode, the first terminal plays the first synthesized audio via the in-ear monitor, that is, the user can hear the first dry-sound audio sung by him or her through the in-ear monitor. In the loudspeaker mode, the first terminal plays the accompaniment audio, and needs to perform echo cancellation on an audio after the audio
30 collection apparatus thereof collects the audio, so as to obtain the first dry-sound audio.

[0044] In an embodiment on a basis of the forgoing embodiments, the method further

comprises a following step. In a case that the second account is in the microphone-out-of-use state and the first account switches from the microphone-in-use state to the microphone-out-of-use state, the second terminal determines the current play progress according to the received progress information, and starts playing the locally stored accompaniment audio based on the current play progress.

[0045] As a specific implementation, when the first account is in the microphone-in-use state and the second account is in the microphone-out-of-use state, the first synthesized audio transmitted by the first terminal to the second terminal each carries the progress information. After the first account switches to the microphone-out-of-use state, the first account and the second account are both in the microphone-out-of-use state. The first terminal continues playing the locally stored accompaniment audio. The second terminal may determine the current play progress according to the received progress information, and start to continue playing the locally stored accompaniment audio based on the current play progress.

[0046] In an embodiment on a basis of the forgoing embodiments, the method further comprises a following step. In a case that the first account switches from the microphone-in-use state to the microphone-out-of-use state and then the second account switches from the microphone-out-of-use state to the microphone-in-use state, the second terminal transmits the second synthesized audio and the progress information thereof to the first terminal, and the first terminal stops playing the locally stored accompaniment audio and starts playing the second synthesized audio. The second synthesized audio is an audio composite of the accompaniment audio locally stored at the second terminal and the second dry-sound audio collected by the second terminal.

[0047] As the specific implementation, the second account switches to the microphone-in-use state after the first account switches to the microphone-out-of-use state, so as to achieve successive singing of the same target song between different accounts. After switching to the microphone-in-use state, the second terminal collects the second dry-sound audio, mixes the locally stored accompaniment audio and the second dry-sound audio to generate the second synthesized audio, and transmits the second synthesized audio and the progress information thereof to the first terminal via the server. The first terminal stops playing the locally stored accompaniment audio, and starts playing the received second synthesized audio. The solo mode of the second account is similar to the aforementioned

solo mode of the first account, and details are not repeated herein.

[0048] In an embodiment on a basis of the forgoing embodiments, the method further comprises a following step. In a case that the first account is in the microphone-in-use state and the second account switches from the microphone-in-use state to the microphone-out-of-use state in the singing duration of the second account, the first terminal determines the current play progress according to the received progress information, starts playing the locally stored accompaniment audio based on the current play progress, and transmits the first synthesized audio carrying a progress of the accompaniment audio to the second terminal via the server. The second terminal stops playing the locally stored accompaniment audio and starts playing the first synthesized audio.

[0049] As a specific implementation, when both the first account and the second account are in the microphone-in-use state, the first account and the second account are in the antiphonal singing mode. In the singing duration of the first account, the first terminal plays the locally stored accompaniment audio, and the second terminal plays the first synthesized audio transmitted by the first terminal. In a case that the second account is switched to the micro-out-of-use state in such duration, the first account resumes the solo mode, and the audio played by the first terminal and the audio played by the second terminal both do not need to change. In the singing duration of the second account, the first terminal plays the second synthesized audio transmitted by the second terminal, and the second terminal plays the locally stored accompaniment audio. In a case that the second account is switched to the micro-out-of-use state in such duration, the first account resumes the solo mode. The first terminal determines the current play progress according to the received progress information of the second synthesized audio, starts playing the locally stored accompaniment audio based on the current play progress, starts collecting the first dry-sound audio, mixes the locally stored accompaniment audio and the first dry-sound audio to generate the first synthetic audio, transmits the first synthetic audio and the progress information thereof to the second terminal via the server. The second terminal stops playing the locally stored accompaniment audio, and starts playing the received first synthesized audio.

[0050] A server is further provided according to embodiments of the present disclosure. Reference is made to Figure 4, which is a structural diagram of an electronic device according to an embodiment of the present disclosure. As shown in Figure 4, the electronic

device 40 may include a processor 41 and a memory 42.

[0051] The processor 41 may comprise one or more processing cores, such as a 4-core processor, an 8-core processor, or the like. The processor 41 may adopt at least one hardware form among the DSP (digital signal processing), the FPGA (field programmable gate array), and the PLA (programmable logic array). The processor 41 may further include a main processor and a coprocessor. The main processor is a processor for processing data in a wake-up state, also called a CPU (central processing unit). The coprocessor is a low-power processor for processing data in a standby state. In some embodiments, the processor 41 may be integrated with a GPU (graphics processing unit), and the GPU is configured to render and draw content to be displayed on a display screen. In some embodiments, the processor 41 may further include an AI (artificial intelligence) processor, and the AI processor is configured to process computing operations related to machine learning.

[0052] The memory 42 may comprise one or more computer-readable storage media, and may be non-transitory. The memory 42 may include a high-speed random access memory and a non-volatile memory, such as one or more magnetic disk storage apparatuses and one or more flash memory storage apparatuses. Herein the memory 42 is at least configured to store a computer program 421 as follows. The computer program after being loaded and executed by the processor 41 can implement relevant steps, in the method according to any foregoing embodiment, which are performed by the first terminal or the second terminal. In addition, resources stored in the memory 42 may further store an operating system 422, data 423, and the like. Such storage may be temporary storage or permanent storage. The operating system 422 may include the Windows, the Unix, the Linux, and the like.

[0053] In some embodiments, the electronic device 40 may further comprise a display 43, an input/output interface 44, a communication interface 45, a sensor 46, a power supply 47 and a communication bus 48.

[0054] The structure of the server as shown in Figure 4 does not constitute a limitation on the server herein. In practice, the server may comprise more or less components than those as shown in Figure 4, or certain components may be combined in the server.

[0055] A computer-readable storage medium comprising program instructions is further provided according to another embodiment. The program instructions when executed by a

processor implements the steps of the method according to any forgoing embodiment.

[0056] The embodiments of the present disclosure are described in a progressive manner, and each embodiment places emphasis on the difference from other embodiments. Therefore, one embodiment can refer to other embodiments for the same or similar parts.

5 Since apparatuses disclosed in the embodiments correspond to methods disclosed in the embodiments, the description of the apparatuses is simple, and reference may be made to the relevant part of the methods. Those skilled in the art can make several improvements and modifications on the present disclosure without departing from principles of the present disclosure, and such improvements and modifications also fall within the protection scope of

10 the claims of the present disclosure.

[0057] It should be noted that, the relationship terms such as "first", "second" and the like are only used herein to distinguish one entity or operation from another, rather than to necessitate or imply that an actual relationship or order exists between the entities or operations. Furthermore, the terms such as "include", "comprise" or any other variants

15 thereof means to be non-exclusive. Therefore, a process, a method, an article or a device including a series of elements include not only the disclosed elements but also other elements that are not clearly enumerated, or further include inherent elements of the process, the method, the article or the device. Unless expressively limited, the statement "including a..." does not exclude the case that other similar elements may exist in the process, the method, the

20 article or the device other than enumerated elements.

CLAIMS

1. A method for implementing an online karaoke room, wherein a first account corresponding to a first terminal and a second account corresponding to a second client are
5 matched with a virtual room, and the method comprises:
- in singing duration of the first account,
- playing locally stored first audio content by the first terminal;
- transmitting, by the first terminal, a first target audio and progress information of the first target audio to the second terminal; and
10 playing the first target audio by the second terminal,
- wherein the first audio content is audio content which is from an accompaniment audio and corresponds to the first account, and the first target audio is an audio composite of the first audio content locally stored at the first terminal and a first dry-sound audio collected by the first terminal;
- 15 in singing duration of the second account,
- determining, by the second terminal, a current play progress according to the progress information received by the second terminal;
- starting, by the second terminal, playing locally stored second audio content based on the current play progress; and
- 20 transmitting a second target audio and progress information of the second target video to the first terminal;
- wherein the second audio content is audio content which is from the accompaniment audio and corresponds to the second account, and the second target audio is an audio composite of the second audio content locally stored at the second
25 terminal and a second dry-sound audio collected by the second terminal; and
- in the singing period of the second account,
- playing the second target audio by the first terminal, in response to the first terminal receiving the second target audio transmitted by the second terminal.

2. The method according to claim 1, further comprising:

playing, by the first terminal in the singing duration of the second account, the locally stored second audio content or a blank audio when the first terminal does not receive the second target audio transmitted by the second terminal.

5 3. The method according to claim 1, further comprising:

transmitting a target audio to a third terminal corresponding to a third account in the virtual room to enable third terminal to play the target audio, wherein the target audio comprises the first target audio transmitted by the first terminal and the second target audio transmitted by the second terminal.

10 4. The method according to claim 1, wherein the first terminal is in a loudspeaker mode, and the method further comprises:

performing, by the first terminal, echo processing on an audio collected by the first terminal to obtain the first dry-sound audio; and

15 performing, by the second terminal, echo processing on an audio collected by the second terminal to obtain the second dry-sound audio.

5. The method according to claim 1, wherein playing the second target audio, by the first terminal in the singing period of the second account, in response to the first terminal receiving the second target audio transmitted by the second terminal comprises:

in the singing period of the second account,

20 applying, by the first terminal, a fade-out effect on the played first audio content before receiving the second target audio; and

applying, by the first terminal, a fade-in effect on the second target audio in response to receiving the second target audio.

6. The method according to claim 1, further comprising:

25 downloading, by the first terminal and the second terminal, the accompaniment audio of the target song from the server;

storing, by the first terminal and the second terminal, the accompaniment audio as the first audio content and the second audio content, respectively; and

playing, by first terminal and the second terminal separately, the respective locally stored accompaniment audio, in response to both the first account and the second account being in a microphone-out-of-use state.

7. The method according to claim 1, further comprising:

5 in response to the second account being in a microphone-out-of-use state and the first account switching from the microphone-out-of-use state to a microphone-in-use state,

transmitting, by the first terminal, the first synthesized audio and the progress information of the first synthesized audio to the second terminal, wherein the first synthesized audio is an audio composite of the accompaniment audio locally stored at
10 the first terminal and the first dry-sound audio collected by the first terminal;

stopping, by the second terminal, playing the accompaniment audio locally stored at the second terminal; and

starting, by the second terminal, playing the first synthesized audio.

8. The method according to claim 7, wherein stopping, by the second terminal, playing
15 the accompaniment audio locally stored at the second terminal and starting, by the second terminal, playing the first synthesized audio comprises:

applying, by the second terminal, a fade-out effect on the accompaniment audio locally stored at the second terminal; and

applying, by the second terminal, a fade-in effect on the first synthesized audio.

20 9. The method according to claim 1, further comprising:

in response to the first account switching from a microphone-in-use state to a microphone-out-of-use state and the second account being in the microphone-out-of-use state,

determining the current play progress by the second terminal according to the received progress information; and

25 starting, by the second terminal based on the current play progress, playing the accompaniment audio locally stored at the second terminal.

10. The method according to claim 1, further comprising:

in response to the first account being switched from a microphone-in-use state to a

microphone-out-of-use state and then the second account being switched from the microphone-out-of-use state to the microphone-in-use state,

transmitting, by the second terminal, the second synthesized audio and progress information of the second synthesized audio to the first terminal, wherein the second synthesized audio is an audio composite of the accompaniment audio locally stored at the second terminal and the second dry-sound audio collected by the second terminal;

stopping, by the first terminal, playing the accompaniment audio locally stored at the first terminal; and

starting, by the first terminal, playing the second synthesized audio.

11. The method according to claim 1, further comprising:

in response to both the first account and the second account being in a microphone-in-use state,

determining, by the first terminal from the accompaniment audio, the first audio content corresponding to the first account; and

determining, by the second terminal from the accompaniment audio, the second audio content corresponding to the second account.

12. The method according to claim 1, further comprising:

in response to the first account being in a microphone-in-use state and the second account switching from the microphone-in-use state to the microphone-out-of-use state in the singing duration of the second account,

determining, by the first terminal, another current play progress based on progress information of the progress information of the second synthesized audio;

starting, by the first terminal based on the another current play progress, playing the accompaniment audio locally stored at the first terminal;

transmitting, by the first terminal, the first synthesized audio carrying an accompaniment progress to the second terminal via the server;

stopping, by the second terminal, playing the accompaniment audio locally stored at the second terminal; and

starting, by the second terminal, playing the first synthesized audio.

13. An electronic device, comprising:

a memory storing a computer program; and

a processor;

5 wherein the processor when executing the computer program is configured to perform steps, which are in the method according to any one of claims 1 to 12 and performed by the first terminal or the second terminal.

14. A computer-readable storage medium, storing a computer program, wherein:

10 the computer program when executed by a processor implements steps, which are in the method according to any one of claims 1 to 12 and performed by the first terminal or the second terminal.