CLAIMS

1. A method for processing a chorus audio, comprising:

obtaining a plurality of dry audios respectively from a plurality of performers singing a same target song;

performing time alignment on the obtained plurality of dry audios;

performing virtual sound image localization on the plurality of dry audios after the time alignment, to locate the dry audios to a plurality of virtual sound images; wherein the plurality of virtual sound images are located in a pre-configured virtual sound image coordinate system; wherein, the virtual sound image coordinate system is configured as: a human head is a center of the virtual sound image coordinate system; a midpoint of a line segment where a left ear and a right ear are located is a coordinate origin; a positive direction of a first axis represents a straight front direction of the human head; a positive direction of a second axis represents a lateral direction of the human head from the left ear to the right ear, and a positive direction of a third axis represents a straight above direction of the human head; a distance between each virtual sound image and the coordinate origin is within a set distance range, and an elevation angle of each virtual sound image relative to a plane formed by the first axis and the second axis is within a set angle range;

generating a chorus audio based on the plurality of dry audios after the virtual sound image localization; and

synthesizing, in a case that a lead audio based on the target song is obtained, the lead audio, the chorus audio and an accompaniment, to output a chorus-effect audio,

wherein the performing virtual sound image localization on the plurality of dry audios after the time alignment comprises: grouping the plurality of dry audios after the time alignment based on a quantity of virtual sound images, wherein a quantity of groups is equal to the quantity of the virtual sound images; and locating each group of dry audios respectively to one of the virtual sound images, so that each group of dry audios corresponds to a different one of the virtual sound images, and

wherein in a case that a quantity of the dry audios is greater than or equal to a predetermined threshold, a same dry audio is in only one group; or in a case that the quantity

of the dry audios is less than the predetermined threshold, a same dry audio is in a plurality of groups.

2. The method according to claim 1, wherein the performing time alignment on the obtained plurality of dry audios comprises:

determining a reference audio corresponding to the target song;

for each of the obtained plurality of dry audios, extracting an audio feature of the dry audio and an audio feature of the reference audio, wherein the audio feature is a fingerprint feature or fundamental frequency feature;

determining a time instant corresponding to a maximum similarity between the audio feature of the dry audio and the audio feature of the reference audio as an audio alignment time; and

performing the time alignment on the dry audio based on the audio alignment time.

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

performing band-pass filtering on the each of the plurality of dry audios respectively, to obtain bass data; and

the generating the chorus audio based on the plurality of dry audios after the virtual sound image localization comprises:

generating the chorus audio based on the plurality of dry audios after the virtual sound image localization and the bass data.

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

performing reverberation simulation on the each of the obtained plurality of dry audios respectively; and

the generating the chorus audio based on the plurality of dry audios after the virtual sound image localization comprises:

generating the chorus audio based on the plurality of dry audios after the virtual

sound image localization and the plurality of dry audios after the reverberation simulation.

5. The method according to claim 4, wherein the performing the reverberation simulation on each of the obtained plurality of dry audios respectively comprises:

performing the reverberation simulation on each of the obtained plurality of dry audios respectively by using a cascade of a comb filter and an all-pass filter.

6. The method according to claim 1, wherein after the performing virtual sound image localization on the plurality of dry audios after the time alignment, the method further comprises:

performing reverberation simulation respectively on each of the plurality of dry audios after the virtual sound image localization; and

the generating the chorus audio based on the plurality of dry audios after the virtual sound image localization comprises:

generating the chorus audio based on the plurality of dry audios after both the virtual sound image localization and the reverberation simulation.

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

performing dual-channel simulation on each of the obtained plurality of dry audios respectively; and

the generating the chorus audio based on the plurality of dry audios after the virtual sound image localization comprises:

generating the chorus audio based on the plurality of dry audios after the virtual sound image localization and the plurality of dry audios after the dual-channel simulation.

8. The method according to claim 7, wherein after the performing the dual-channel simulation on each of the obtained plurality of dry audios respectively, the method further comprises:

performing reverberation simulation on the plurality of dry audios after the dual-channel simulation; and

the generating the chorus audio based on the plurality of dry audios after the virtual sound image localization and the plurality of dry audios after the dual-channel simulation comprises:

generating the chorus audio based on the plurality of dry audios after the virtual sound image localization and the plurality of dry audios after both the dual-channel simulation and the reverberation simulation.

9. The method according to claim 1, wherein,

among the virtual sound images, an elevation angle of a virtual sound image located behind the human head relative to the plane formed by the first axis and the second axis is greater than an elevation angle of a virtual sound image located in front of the human head relative to the plane formed by the first axis and the second axis; or

the virtual sound images are evenly distributed in a full circle on the plane formed by the first axis and the second axis.

10. A method for processing a chorus audio, comprising:

obtaining a plurality of dry audios respectively from a plurality of performers singing a same target song;

performing time alignment on the obtained plurality of dry audios;

performing virtual sound image localization on the plurality of dry audios after the time alignment, to locate the dry audios to a plurality of virtual sound images; wherein the plurality of virtual sound images are located in a pre-configured virtual sound image coordinate system; wherein, the virtual sound image coordinate system is configured as: a human head is a center of the virtual sound image coordinate system; a midpoint of a line segment where a left ear and a right ear are located is a coordinate origin; a positive direction of a first axis represents a straight front direction of the human head; a positive direction of a second axis represents a lateral direction of the human head from the left ear to the right ear, and a positive direction of

a third axis represents a straight above direction of the human head; a distance between each virtual sound image and the coordinate origin is within a set distance range, and an elevation angle of each virtual sound image relative to a plane formed by the first axis and the second axis is within a set angle range;

generating a chorus audio based on the plurality of dry audios after the virtual sound image localization; and

synthesizing, in a case that a lead audio based on the target song is obtained, the lead audio, the chorus audio and an accompaniment, to output a chorus-effect audio,

wherein the synthesizing the lead audio, the chorus audio and the accompaniment comprises:

performing volume adjustment on the lead audio and the chorus audio respectively to cause a volume of the lead audio to be equal to or greater than a volume of the chorus audio; and

synthesizing the accompaniment with the lead audio and the chorus audio after the volume adjustment.

11. The method according to claim 10, wherein the synthesizing the lead audio, the chorus audio and the accompaniment further comprises:

performing reverberation simulation on the lead audio and the chorus audio; and

the synthesizing the accompaniment with the lead audio and the chorus audio after the volume adjustment further comprises:

synthesizing the accompaniment with the lead audio and the chorus audio after the volume adjustment and the reverberation simulation.

12. A device for processing a chorus audio, comprising:

a memory, configured to store a computer program; and

a processor, configured to implement, when executing the computer program, the method for processing a chorus audio according to any of claims 1 to 9.

13. A computer-readable storage medium storing a computer program, wherein the computer program, when executed by a processor, implements the method for processing a chorus audio according to any of claims 1 to 9.