

CLAIMS

What is claimed is:

1. A ~~role separation~~ method, comprising:

5 acquiring sound source angle data corresponding to a speech data frame of a to-be-separated role collected by a speech collection device;

performing identification on the to-be-separated role based on the sound source angle data to obtain a first identification result of the to-be-separated role; and

separating a role based on the first identification result of the to-be-separated role.

10 2. The method according to claim 1, wherein after the acquiring the sound source angle data corresponding to the speech data frame of the to-be-separated role collected by the speech collection device, the method further comprises:

performing voice activity detection on the speech data frame of the to-be-separated role to obtain a speech data frame with a voice activity;

15 filtering and smoothing the speech data frame with the voice activity based on an energy spectrum of the speech data frame of the to-be-separated role to obtain a filtered and smoothed speech data frame; and

updating the sound source angle data based on the filtered and smoothed speech data frame to obtain updated sound source angle data.

20 3. The method according to claim 2, wherein the filtering and smoothing the speech data frame with the voice activity based on the energy spectrum of the speech data frame of the to-be-separated role to obtain the filtered and smoothed speech data frame comprises:

filtering and smoothing the speech data frame with the voice activity through a median filter based on spectral flatness of the energy spectrum of the speech data frame of the to-be-separated role to obtain the filtered and smoothed speech data frame.

30 4. The method according to claim 1, wherein the performing the identification on the to-be-separated role based on the sound source angle data to obtain the first identification result of the to-be-separated role comprises:

performing sequence clustering on the sound source angle data to obtain a sequence clustering result of the sound source angle data; and

determining that a role identifier corresponding to the sequence clustering result of the sound source angle data is the first identification result of the to-be-separated role.

5 5. The method according to claim 4, wherein the performing the sequence clustering on the sound source angle data to obtain the sequence clustering result of the sound source angle data comprises:

 determining a distance between the sound source angle data and a sound source angle sequence clustering center; and

10 determining the sequence clustering result of the sound source angle data based on the distance between the sound source angle data and the sound source angle sequence clustering center.

6. The method according to claim 1, wherein after the obtaining the first identification result of the to-be-separated role, the method further comprises:

15 performing voiceprint identification on a speech data frame of the to-be-separated role within a preset time period to obtain a second identification result of the to-be-separated role; and

20 in response to determining that ~~if~~ the first identification result is different from the second identification result, using the second identification result to correct the first identification result to obtain a final identification result of the to-be-separated role.

7. The method according to claim 1, wherein the speech collection device comprises a microphone array, ~~and~~

25 8. The method according to claim 7, wherein the acquiring the sound source angle data corresponding to the speech data frame of the to-be-separated role collected by the speech collection device ~~further~~ comprises:

 acquiring a covariance matrix of the speech data frame received by at least some microphones in the microphone array;

30 performing eigenvalue decomposition on the covariance matrix to obtain multiple eigenvalues;

 selecting a first quantity of largest eigenvalues from the multiple eigenvalues, and forming a speech signal sub-space based on eigenvectors corresponding to the selected eigenvalues; ~~wherein the first quantity is equivalent to an estimated quantity of sound~~

~~sources;~~ and

determining the sound source angle data based on the speech signal sub-space.

9. The method according to claim 8, wherein the first quantity is equivalent to an estimated quantity of sound sources.

810. The method according to claim 6, wherein after the obtaining the final identification result of the to-be-separated role, the method further comprises:

acquiring face image data of the to-be-separated role collected by an image collection device;

performing face recognition on the face image data to obtain a third identification result of the to-be-separated role; and

in response to determining that ~~if~~ the third identification result is different from the second identification result, using the third identification result to correct the second identification result to obtain the final identification result of the to-be-separated role.

~~9.——A role separation method, comprising:~~

~~sending to a cloud a role separation request carrying a speech data frame of a to-be-separated role, so that the cloud acquires sound source angle data corresponding to the speech data frame based on the role separation request, performs identification on the to-be-separated role based on the sound source angle data, and separates the role based on an identification result of the to-be-separated role; and~~

~~receiving a separation result of the role sent by the cloud based on the role separation request.~~

~~10.——A role separation method, comprising:~~

~~receiving from a speech collection device a role separation request carrying a speech data frame of a to-be-separated role;~~

~~acquiring sound source angle data corresponding to the speech data frame based on the role separation request;~~

~~performing identification on the to-be-separated role based on the sound source angle data to obtain an identification result of the to-be-separated role; and~~

~~separating the role based on the identification result of the to-be-separated role, and sending to the speech collection device a role separation result corresponding to the~~

~~role separation request.~~

~~11. —A meeting minutes recording method, comprising:
acquiring sound source angle data corresponding to a speech data frame of a
meeting role collected by a speech collection device disposed in a meeting room;
performing identification on the meeting role based on the sound source angle
data to obtain an identification result of the meeting role; and
recording meeting minutes of the meeting role based on the identification result
of the meeting role.~~

~~1211. An role display method, apparatus comprising:
one or more processors; and
one or more memories storing thereon computer-readable instructions that,
executable by the one or more processors, cause the one or more processors to perform
acts comprising:~~

acquiring sound source angle data corresponding to a speech data frame
of a role collected by a speech collection device;
performing identification on the role based on the sound source angle data
to obtain an identification result of the role; and
displaying identity data of the role on an interactive interface of the speech
collection device based on the identification result of the role.

~~1312. The apparatus method according to claim 1211, wherein after the acquiring
the sound source angle data corresponding to the speech data frame of the role collected
by the speech collection device, the method acts further comprises:~~
switching on a lamps of the speech collection device in a sound source direction
indicated by the sound source angle data.

~~1413. The apparatus method according to claim 112, wherein the acts further
comprising:~~
displaying a speaking action images or a speech waveform images of the role on
the interactive interface of the speech collection device.

154. One or more memories storing thereon computer-readable instructions that, executable by one or more processors, cause the one or more processors to perform acts~~A role separation apparatus~~, comprising:

~~a first acquisition module, configured to acquire~~ing sound source angle data corresponding to a speech data frame of a to-be-separated role collected by a speech collection device;

~~a first identification module, configured to perform~~ing identification on the to-be-separated role based on the sound source angle data to obtain a first identification result of the to-be-separated role; and

~~a separation module, configured to separate the to-be-separated~~separating a role based on the first identification result of the to-be-separated role.

1615. The one or more memories according to claim 14, wherein after the acquiring~~A role separation apparatus, comprising:~~

~~a first sending module, configured to send to a cloud a role separation request carrying a speech data frame of a to-be-separated role, so that the cloud acquires the~~ sound source angle data corresponding to the speech data frame of the to-be-separated role collected by the speech collection device, the acts further comprise:

performing voice activity detection on the speech data frame of~~based on the role separation request, performs identification on the to-be-separated role~~ to obtain a speech data frame with a voice activity;

filtering and smoothing the speech data frame with the voice activity based on an energy spectrum of the speech data frame~~based on the sound source angle data, and separates the role based on an identification result of the to-be-separated role~~ to obtain a filtered and smoothed speech data frame; and

updating the sound source angle based on the filtered and smoothed speech data frame to obtain updated sound source angle data.~~a first receiving module, configured to receive a separation result of the role sent by the cloud based on the role separation request.~~

16. The one or more memories according to claim 15, wherein the filtering and smoothing the speech data frame with the voice activity based on the energy spectrum of the speech data frame of the to-be-separated role to obtain the filtered and smoothed

speech data frame comprises:

filtering and smoothing the speech data frame with the voice activity through a median filter based on a spectral flatness of the energy spectrum of the speech data frame of the to-be-separated role to obtain the filtered and smoothed speech data frame.

5

17. The one or more memories according to claim 14, wherein the performing the identification on the to-be-separated role based on the sound source angle data to obtain the first identification result of the to-be-separated role comprises:

performing sequence clustering on the sound source angle data to obtain a sequence clustering result of the sound source angle data; and

10

determining that a role identifier corresponding to the sequence clustering result of the sound source angle data is the first identification result of the to-be-separated role.

18. The one or more memories according to claim 17, wherein the performing the sequence clustering on the sound source angle data to obtain the sequence clustering result of the sound source angle data comprises:

15

determining a distance between the sound source angle data and a sound source angle sequence clustering center; and

20

determining the sequence clustering result of the sound source angle data based on the distance between the sound source angle data and the sound source angle sequence clustering center.

19. The one or more memories according to claim 14, wherein after the obtaining the first identification result of the to-be-separated role, the acts further comprise:

25

performing voiceprint identification on a speech data frame of the to-be-separated role within a preset time period to obtain a second identification result of the to-be-separated role; and

30

in response to determining that the first identification result is different from the second identification result, using the second identification result to correct the first identification result to obtain a final identification result of the to-be-separated role.

20. The one or more memories according to claim 14, wherein:
the speech collection device comprises a microphone array; and
the acquiring the sound source angle data corresponding to the speech data frame
of the to-be-separated role collected by the speech collection device comprises:
5 acquiring a covariance matrix of the speech data frame received by at least some
microphones in the microphone array;
performing eigenvalue decomposition on the covariance matrix to obtain multiple
eigenvalues;
selecting a first quantity of largest eigenvalues from the multiple eigenvalues, and
10 forming a speech signal sub-space based on eigenvectors corresponding to the selected
eigenvalues; and
determining the sound source angle data based on the speech signal sub-space.

~~17.—A role separation apparatus, comprising:~~
15 ~~a second receiving module, configured to receive from a speech collection device~~
~~a role separation request carrying a speech data frame of a to-be-separated role;~~
~~a third acquisition module, configured to acquire sound source angle data~~
~~corresponding to the speech data frame based on the role separation request;~~
~~a second identification module, configured to perform identification on the to-be-~~
20 ~~separated role based on the sound source angle data to obtain an identification result of~~
~~the to-be-separated role; and~~
~~a second sending module, configured to separate the role based on the~~
~~identification result of the to-be-separated role, and send to the speech collection device~~
~~a role separation result corresponding to the role separation request.~~

25 ~~18.—A meeting minutes recording apparatus, comprising:~~
~~a fourth acquisition module, configured to acquire sound source angle data~~
~~corresponding to a speech data frame of a meeting role collected by a speech collection~~
~~device disposed in a meeting room;~~
~~a third identification module, configured to perform identification on the meeting~~
30 ~~role based on the sound source angle data to obtain an identification result of the meeting~~
~~role; and~~

~~a recording module, configured to record meeting minutes of the meeting role based on the identification result of the meeting role.~~

~~19. — A role display apparatus, comprising:~~

~~5 a fifth acquisition module, configured to acquire sound source angle data corresponding to a speech data frame of a role collected by a speech collection device;~~

~~a fourth identification module, configured to perform identification on the role based on the sound source angle data to obtain an identification result of the role; and~~

~~10 a first display module, configured to display identity data of the role on an interactive interface of the speech collection device based on the identification result of the role.~~

~~20. — An electronic device, comprising: a processor, a memory, a communications interface, and a communications bus, wherein the processor, the memory, and the communications interface communicate with each other through the communications bus;~~

~~the memory is configured to store at least one executable instruction that causes the processor to execute an operation corresponding to the role separation method according to any one of claims 1—8, or an operation corresponding to the role separation method according to claim 9, or an operation corresponding to the role separation method according to claim 10, or an operation corresponding to the meeting minutes recording method according to claim 11, or an operation corresponding to the role display method according to any one of claims 12—14.~~

~~21. — A computer storage medium having a computer program stored thereon, wherein the program, when executed by a processor, is caused to implement the role separation method according to any one of claims 1—8, or the role separation method according to claim 9, or the role separation method according to claim 10, or the meeting minutes recording method according to claim 11, or the role display method according to any one of claims 12—14.~~