

DATA PROCESSING METHOD, DATA PROCESSING APPARATUS, ELECTRONIC DEVICE AND STORAGE MEDIUM

Technical Field

5 The present disclosure relates to a technical field of computers, and particularly to a data processing method, a data processing apparatus, an electronic device and a storage medium.

Background

10 The ‘webcast’ is basically classified into two types. One type is to provide TV signal watching on the Internet, such as the live broadcast of various sports games and cultural activities. The principle of this type of live broadcast is to collect and convert TV (analog) signals into digital signals, input the digital signals to a computer, and then upload them to a website in real time for people to watch, which is equivalent to ‘Internet TV’. The other type is what people know as ‘webcast’, which is to set up independent signal collection devices (audio and video) on site,
15 introduce a director (a directing device or a platform), upload signals to a server through a network, and publish the signals to a website address for people to watch.

20 During a live broadcast, an anchor usually manually inputs some information to complete an interaction with the audience, such as inputting various information of red envelopes so as to send the red envelopes to the audience watching the live broadcast to improve an activation, editing a broadcast content and sending a broadcast reminder to the audience subscribing to the live broadcast, editing commodity information so as to recommend a commodity to the audience and so on. Taking the release of red envelopes as an example, the existing red envelopes release method is that an anchor enters corresponding information of red envelopes (e.g., an amount and number of red envelopes) in a live broadcast room to generate red envelopes for the audience in the live
25 broadcast room to collect.

 However, in this way, a user needs to manually input various information during the live broadcast, which is very troublesome for the user to operate.

Summary of the Disclosure

30 The embodiments of the present disclosure provide a data processing method to facilitate a

user's operation.

Accordingly, the embodiments of the present disclosure further provide a data processing apparatus, an electronic device and a storage medium, so as to ensure the implementation and application of the above system.

5 In order to solve the above problem, an embodiment of the present disclosure discloses data processing method, comprising: extracting speech data of a first user from live stream data; recognizing the speech data to determine characteristic information corresponding to a target object; determining the target object and corresponding descriptive information based on the characteristic information; and determining presentation information of the target object based on the target object
10 and the descriptive information.

 In order to solve the above problem, an embodiment of the present disclosure discloses a data processing method, comprising: extracting speech data of a first user from live stream data; recognizing the speech data to determine characteristic information corresponding to a virtual interaction object; determining descriptive information corresponding to the virtual interaction
15 object based on the characteristic information, wherein the descriptive information comprises an object quantity and an object resource; calling a virtual interactive plug-in, and configuring attribute information of the virtual interaction object based on the descriptive information; and adding the virtual interaction object to the live stream data for display.

 In order to solve the above problem, an embodiment of the present disclosure discloses a data
20 processing method, comprising: extracting speech data of a first user from live stream data; recognizing the speech data to determine characteristic information corresponding to a commodity object; determining a target commodity object and corresponding descriptive information based on the characteristic information, wherein the descriptive information comprises price information and quantity information of the target commodity object; determining link information of the target
25 commodity object based on the target commodity object and the descriptive information; and adding the link information of the target commodity object to the live stream data for display.

 In order to solve the above problem, an embodiment of the present disclosure discloses a data processing method, comprising: extracting speech data of a target user from data of a video conference; recognizing the speech data to determine characteristic information corresponding to a
30 target object; determining the target object and corresponding descriptive information based on the

characteristic information; determining presentation information of the target object based on the target object and the descriptive information; and adding the presentation information to the data of the video conference for display.

5 In order to solve the above problem, an embodiment of the present disclosure discloses a data processing method, comprising: extracting speech data of a target user from data of a live stream teaching activity; recognizing the speech data to determine characteristic information corresponding to a target object; determining the target object and corresponding descriptive information based on the characteristic information; determining presentation information of the target object based on the target object and the descriptive information; and adding the presentation information to the data
10 of a live stream teaching activity for display.

In order to solve the above problem, an embodiment of the present disclosure discloses a data processing apparatus, comprising: a speech data extraction module configured to extract speech data of a first user from live stream data; a characteristic information extraction module configured to recognize the speech data to determine characteristic information corresponding to a target object;
15 a characteristic information filtering module configured to determine the target object and corresponding descriptive information based on the characteristic information; and a presentation information extraction module configured to determine presentation information of the target object based on the target object and the descriptive information.

In order to solve the above problem, an embodiment of the present disclosure discloses a data
20 processing apparatus, comprising: a speech data acquisition module configured to extract speech data of a first user from live stream data; a characteristic information acquisition module configured to recognize the speech data to determine characteristic information corresponding to a virtual interaction object; a characteristic information processing module configured to determine descriptive information corresponding to the virtual interaction object based on the characteristic
25 information, wherein the descriptive information comprises an object quantity and an object resource; an interaction object generation module configured to call a virtual interactive plug-in and configure attribute information of the virtual interaction object based on the descriptive information; and an interaction object display module configured to add the virtual interaction object to the live stream data for display.

30 In order to solve the above problem, an embodiment of the present disclosure discloses a data

processing apparatus, comprising: a speech data determination module configured to extract speech data of a first user from live stream data; a characteristic information determination module configured to recognize the speech data to determine characteristic information corresponding to a commodity object; a commodity information filtering module configured to determine a target commodity object and corresponding descriptive information based on the characteristic information, wherein the descriptive information comprises price information and quantity information of the target commodity object; a link information determination module configured to determine link information of the target commodity object based on the target commodity object and the descriptive information; and a link information presentation module configured to add the link information of the target commodity object to the live stream data for display.

In order to solve the above problem, an embodiment of the present disclosure discloses a data processing apparatus, comprising: a speech data acquisition module configured to extract speech data of a target user from data of a video conference; a characteristic information acquisition module configured to recognize the speech data to determine characteristic information corresponding to a target object; a characteristic information screening module configured to determine the target object and corresponding descriptive information based on the characteristic information; a presentation information acquisition module configured to determine presentation information of the target object based on the target object and the descriptive information; and a presentation information output module configured to add the presentation information to the data of the video conference for display.

In order to solve the above problem, an embodiment of the present disclosure discloses a data processing apparatus, comprising: a speech data filtering module configured to extract speech data of a target user from data of a live stream teaching activity; a characteristic information recognition module configured to recognize the speech data to determine characteristic information corresponding to a target object; an object information screening module configured to determine the target object and corresponding descriptive information based on the characteristic information; an object information processing module configured to determine presentation information of the target object based on the target object and the descriptive information; and a presentation information sending module configured to add the presentation information to the data of a live stream teaching activity for display.

In order to solve the above problem, an embodiment of the present disclosure discloses an electronic device comprising: a processor; and a memory storing executable codes which, when executed, enables the processor to perform the method of one or more of the above embodiments.

5 In order to solve the above problem, an embodiment of the present disclosure discloses one or more machine-readable media storing executable codes which, when executed, enables a processor to perform the method of one or more of the above embodiments.

As compared with the prior art, the embodiments of the present disclosure have the following advantages:

In the embodiments of the present disclosure, the speech data of the first user can be extracted
10 from the live stream data, and recognized to determine the corresponding characteristic information. Then, the target object and the descriptive information are determined based on the characteristic information, and the corresponding presentation information is determined based on the target object and the descriptive information. In the embodiments of the present disclosure, the corresponding presentation information can be generated based on the speech data of the first user,
15 and the user does not need to manually input the corresponding data, thus facilitating the operation of the user.

Brief Description of the Drawings

Fig. 1 is a schematic diagram of a structure of a data processing system according to an
20 embodiment of the present disclosure;

Fig. 2 is a flow diagram of a data processing method according to an embodiment of the present disclosure;

Fig. 3 is a schematic diagram of displaying a red envelope at a second terminal according to an embodiment of the present disclosure;

25 Fig. 4 is a flow diagram of a data processing method according to another embodiment of the present disclosure;

Fig. 5 is a flow diagram of a data processing method according to still another embodiment of the present disclosure;

Fig. 6 is a flow diagram of a data processing method according to still another embodiment of
30 the present disclosure;

Fig. 7 is a flow diagram of a data processing method according to still another embodiment of the present disclosure;

Fig. 8 is a flow diagram of a data processing method according to still another embodiment of the present disclosure;

5 Fig. 9 is a schematic diagram of a structure of a data processing apparatus according to an embodiment of the present disclosure;

Fig. 10 is a schematic diagram of a structure of a data processing apparatus according to another embodiment of the present disclosure;

10 Fig. 11 is a schematic diagram of a structure of a data processing apparatus according to still another embodiment of the present disclosure;

Fig. 12 is a schematic diagram of a structure of a data processing apparatus according to still another embodiment of the present disclosure;

Fig. 13 is a schematic diagram of a structure of a data processing apparatus according to still another embodiment of the present disclosure;

15 Fig. 14 is a schematic diagram of a structure of an exemplary apparatus provided by an embodiment of the present disclosure.

Detailed Description of the Embodiments

20 In order that the above objectives, features and advantages of the present disclosure are more obvious and understandable, the present disclosure will be further explained in detail with reference to the drawings and the specific embodiments.

The embodiments of the present disclosure may be applied to a field of live broadcast, and particularly, may be applied to a scene of recognizing speech data in a live broadcast process based on an Automatic Speech Recognition (ASR) and Natural Language Processing (NLP), wherein the
25 ASR is a technology for converting a human speech into a text. The processing of speech data by the ASR may comprise: converting speech data into an audio with a waveform, segmenting the audio (e.g., segmenting the audio on a frame basis to obtain audio frames), then determining corresponding states of the segmented audios, determining phones based on at least one state, and determining words from the phones to form text information using the words. The NLP is an
30 interdisciplinary field of computer science, artificial intelligence and linguistics, and it aims to

enable a computer to process or ‘understand’ a natural language. The processing of text information by the NLP may comprise: acquiring text information, performing data cleaning processing for text information (e.g., the text information may be subjected to de-duplication, marking, noise reduction, alignment, etc.), then tokenizing the processed data, and extracting features of tokens to determine word vectors, wherein the feature of a token is related to the token and features of words before and after the token. After the word vectors are determined, the semantics of the token is determined based on the word vectors.

For example, this embodiment of the present disclosure may be applied to a scene of recognizing speech data in live stream data based on the ASR and the NLP, as illustrated in Fig. 1.

In this scene, a first user in a live broadcast transmits live stream data to a server through a first terminal, and the server assigns the live stream data to a second terminal of a second user watching the live broadcast, wherein the first user may be understood as an anchor of the live broadcast, the second user may be understood as the audience watching the live broadcast, and the first and second terminals may be electronic devices such as mobile phones and computers.

During the live broadcast, the first user may assign some red envelopes or products to the second users watching the live broadcast, so as to improve the activation of the second users. In the embodiment of the present disclosure, the red envelope may be sent by recognizing speech data of a user. Specifically, the first user may say, ‘let’s send a rich red envelope of 88 CNY to see which five persons can get it’. A processing end in the embodiment of the present disclosure may convert speech data of the first user in the live broadcast into text information based on the ASR, and perform a semantic recognition on the text information based on the NLP, so as to determine corresponding target object information (e.g., ‘red envelope’ contained in the text information) and descriptive information (e.g., ‘five’ and ‘88 CNY’ contained in the text information), wherein the target object information is used to determine a target object, the descriptive information is used to determine attribute information of the target object, and the target object may be an interaction object to complete an interaction between the users, such as an application program, a plug-in, a link, etc., which may be set depending on the scene. For example, in a live broadcast scene, the target object may be a red envelope plug-in, link information of a commodity object, etc., and the corresponding descriptive information may be a red envelope size, a red envelope amount, a commodity price, a commodity quantity, etc. After the target object and the descriptive information

are determined, corresponding presentation information (e.g., red envelope) is determined and presented to the second users watching the live broadcast. In the embodiment of the present disclosure, by recognizing the speech data of the first user, the corresponding target object and descriptive information are determined, then the presentation information of the target object is
5 determined and added to the live stream data for display, so that any user in the live broadcast does not need to manually input various information of the red envelope, thus facilitating the operation of the first user. In the embodiment of the present disclosure, the processing end may be either the first terminal or the server.

The data processing method in the embodiment of the present disclosure will be described as
10 follows through an example where the present disclosure is applied in a scene of red envelope release in a live broadcast process. The data processing method in the embodiment of the present disclosure may be executed by a processing end, as illustrated in Fig. 2, and the method comprises:
Step 202: extracting speech data of a first user from live stream data. The first user may be understood as an anchor of a live broadcast, and the number thereof may be one or more. In a case
15 where the processing end is a first terminal, the first terminal may extract an audio collected by a sound collection device as the speech data of the first user. In a case where the processing end is a second terminal, the second terminal may analyze the live stream data to obtain video data and audio data in the live stream data, and extract the audio data as the speech data of the first user. In addition, the live stream data may contain speech data of other users except the first user, and the
20 processing end may recognize the speech data of the first user by voiceprint to recognize the speech data of the first user, wherein the voiceprint is an acoustic spectrum carrying speech information and displayed by an electro-acoustic instrument. Whether a user is the anchor (first user) is determined by recognizing the voiceprint in the speech data.

Step 204: recognizing the speech data to determine characteristic information corresponding to
25 a target object. The characteristic information is information describing features of the target object, and for example may comprise a name of the target object (e.g., red envelope) and information describing attributes of the target object. After determining the speech data, the processing end may convert the speech data into text information, and perform a semantic recognition on the text information to determine characteristic words corresponding to the target object. Specifically, in an
30 optional embodiment, recognizing the speech data to determine characteristic information

corresponding to the target object comprises: performing a speech recognition on the speech data to determine corresponding text information; and performing a semantic analysis on the text information to determine the characteristic information corresponding to the target object. The characteristic information may be understood as characteristic words of the target object in the text information, and the characteristic words of the target object may comprise object characteristic words of the corresponding target object information and descriptive characteristic words of the corresponding descriptive information of the target object. A characteristic word library corresponding to a plurality of target objects may be preset at the processing end. The processing end may convert the speech data into the text information based on the ASR, and match the words in the text information with the characteristic words in the characteristic word library based on the NLP, so as to determine the characteristic words corresponding to the target object.

After determining the characteristic information corresponding to the target object, in step 206, the processing end may determine the target object and the corresponding descriptive information based on the characteristic information. The processing end may filter out characteristic words related to a task according to a correspondence between the characteristic words and the target object, and a correspondence between the characteristic words and the descriptive information. Specifically, in an optional embodiment, determining the target object and the corresponding descriptive information based on the characteristic information comprises: recognizing a first characteristic word in the characteristic information to determine the target object; and recognizing a second characteristic word in the characteristic information to determine the descriptive information corresponding to the target object. The first characteristic word may also be called as an object characteristic word, and the second characteristic word may also be called as a descriptive characteristic word. The characteristic word library at the processing end may be divided into a first word library containing the object characteristic words and a second word library containing the descriptive characteristic words. The processing end may match the characteristic words in the text information with characteristic words in the first word library and the second word library to determine the object characteristic words and the descriptive characteristic words, so as to determine the target object and the descriptive information. It should be noted that one target object may be corresponding to a plurality of characteristic words, and one piece of descriptive information may be corresponding to a plurality of characteristic words. For example, the object

characteristic words corresponding to the red envelope (target object) may comprise 'red envelope', 'rich red envelope', etc.; and the descriptive characteristic words corresponding to the number of red envelopes may comprise '5 red envelopes', '5 users', etc.

After determining the target object and the descriptive information, in step 208, the processing
5 end determines the presentation information of the target object based on the target object and the descriptive information. The descriptive information is used to set the attribute information of the target object, while the processing end may call the plug-in of the target object and set the attribute information of the target object based on the descriptive information, so as to determine the presentation information corresponding to the target object. Specifically, in an optional embodiment,
10 determining the presentation information of the target object based on the target object and the descriptive information comprises: setting attribute information of the target object based on the descriptive information; and determining presentation information of the target object based on the target object and the attribute information. After the target object is determined, the plug-in corresponding to the target object may be called, wherein the plug-in corresponding to the target
15 object is used to set the attribute information of the target object. The attribute information of the target object is configured based on the descriptive information to determine the presentation information of the target object. For example, the target object in the live broadcast scene may be understood as a corresponding red envelope plug-in, etc. In the embodiment of the present disclosure, the corresponding red envelope plug-in, etc., may be called and the attribute information
20 of the target object may be set in view of the descriptive information (an amount and number of the red envelopes, etc.), so as to determine the presentation information. It should be noted that in the embodiment of the present disclosure, the prize of the target object may be a coupon, a gold coin, a gift, etc. in addition to a red envelope.

The processing end may add the determined presentation information to the live stream data
25 for display. The processing end may add the presentation information to the live stream data and transmit the live stream data to the second terminal so as to be collected by the second user of the second terminal. Specifically, in a case where the processing end is a first terminal, the first terminal adds the presentation information to the live stream data and transmits the live stream data to a server, which then transmits the live stream data to the second terminal. In a case where the
30 processing end is a server, the server end adds the presentation information to the live stream data

and transmits the live stream data to the second terminal.

In the embodiment of the present disclosure, the presentation information may be directionally sent to a part of audience watching the live broadcast. Specifically, in an optional embodiment, adding the presentation information to the live stream data for display comprises: acquiring
5 permission levels of the second users; determining a second user meeting a permission condition; and adding the presentation information to live stream data related to the second user meeting the permission condition, so as to present the presentation information to the second user meeting the permission condition. The permission condition is used to filter the second user to determine a user who receives the presentation information. In an example, the permission condition may be set
10 based on a speech of the first user. Specifically, the descriptive information may further comprise the permission condition, based on which the processing end acquires the permission levels of the second users to filter a second user, and directionally sends the presentation information to a second user filtered out. For example, the permission level may be a user level, and the permission condition may be a level condition (e.g., level 3). The processing end may acquire the user levels of
15 second users, and filter out a second user whose user level reaches level 3, so as to add the presentation information to live stream data related to the second user filtered out. In another example, the permission condition may also be a condition preset at the processing end, so as to filter out a second user meeting the permission condition to receive the presentation information.

In the example illustrated in Fig. 3, the presentation information (e.g. red envelope) is sent to
20 the second user of the second terminal, and the second user may operate on the presentation information (e.g., clicking 'open') to receive the corresponding red envelope. Specifically, the processing end may receive operation data for the presentation information from the second user; analyze the operation data to determine a target user; and assign a target object corresponding to the presentation information to the target user. The second terminal may receive the operation data from
25 the second user and transmit it to the server for processing. In an optional embodiment, the descriptive information may comprise a collection rule, and the second user may perform corresponding operations based on the collection rule. After receiving the operation data, the processing end (the server or the first terminal) may analyze the operation data based on the collection rule to determine the target user, and assign the target object to the corresponding target
30 user. For example, the collection rule may be sending a specified word or sentence, and determining

the target user by determining whether the second user sends the specified word or sentence, so as to assign the corresponding red envelope to the target user. In addition, the collection rules may also be any other rule, such as a time rule. The processing end may sort the second user based on time information of the operation data, and filter out a second user whose ranking meets the descriptive information condition as the target user, so as to assign the corresponding red envelope to the target user. The collection rule may also be whether the second user share the live broadcast room, whether the second user send a designated gift, whether the user levels of the second user reach a preset level, etc., which is not limited thereto.

After the corresponding presentation information is generated, the presentation information may also be sent to the first user for correction. Specifically, in an optional embodiment, the processing end may present the presentation information; receive a modification operation on the presentation information by the first user to determine the modified presentation information; and add the modified presentation information to the live stream data for display. After the presentation information is determined, the presentation information may be presented to the first user, and a modification to the presentation information by the first user, such as a modification to the amount and number of red envelopes, may be received to determine the modified presentation information and the modified presentation information is presented to the second user.

Before the presentation information is added to the live stream data, a confirmation instruction from the first user may also be acquired to confirm that the presentation information is to be presented. Specifically, in an optional embodiment, the processing end may receive a confirmation instruction of the presentation information from the user, so as to add the presentation information to the live stream data for display based on the confirmation instruction. In an example, the confirmation instruction may be determined from the input information of the first user, and the processing end may present the presentation information to the first user and provide controls such as confirmation and cancellation. The first user may complete the confirmation of the presentation information by clicking the confirmation control, or cancel the presentation of the presentation information through the canceling control. In another example, the confirmation instruction may be determined based on the speech data of the first user, the processing end may present the presentation information to the first user, and the first user may control the sending of the presentation information through the speech data. For example, the first user may say ‘3, 2, 1, start’

to confirm the sending of the presentation information for presentation.

In the embodiment of the present disclosure, the speech data of the first user can be extracted from the live stream data, and recognized to determine the corresponding characteristic information. Then, the target object and the descriptive information are determined based on the characteristic
5 information, and the corresponding presentation information is determined based on the target object and the descriptive information. Next, the presentation information is added to the live stream data for display. In the embodiment of the present disclosure, the corresponding presentation information can be generated based on the speech data of the first user and added to the live stream data for display, so that the user does not need to manually input the corresponding data, thus
10 facilitating the operation of the user.

On the basis of the above embodiments, the present disclosure further discloses a data processing method applicable to a processing end, as illustrated in Fig. 4, comprising:

Step 402: extracting speech data of a first user from live stream data;

Step 404: performing a speech recognition on the speech data to determine corresponding text
15 information;

Step 406: performing a semantic analysis on the text information to determine characteristic information corresponding to a target object;

Step 408: recognizing a first characteristic word in the characteristic information to determine the target object;

20 Step 410: recognizing a second characteristic word in the characteristic information to determine descriptive information corresponding to the target object;

Step 412: setting attribute information of the target object based on the descriptive information;

Step 414: determining presentation information of the target object based on the target object and the attribute information;

25 Step 416: adding the presentation information to the live stream data for display. The processing end may also directionally present the presentation information to a user who meets a permission condition. Specifically, adding the presentation information to the live stream data for display comprises: acquiring permission levels of second users; determining a second user meeting the permission condition; adding the presentation information to live stream data related to the
30 second user meeting the permission condition, so as to present the presentation information to the

second user meeting the permission condition;

Step 418: receiving operation data for the presentation information from the second user;

Step 420: analyzing the operation data to determine a target user;

Step 422: assigning the target object corresponding to the presentation information to the target

5 user.

This embodiment of the present disclosure may be applied to a scene of recognizing speech data of a first user in a live broadcast process. Specifically, in an example of recognizing the speech data of the first user in the live broadcast process and releasing the red envelope, the first user may say 'let's send a rich red envelope of 88 CNY to see which five persons can get it' during the live
10 broadcast. The processing end may extract the speech data and convert the speech data into the text information based on the ASR. Then, based on the NLP, the words in the text information are determined, such as 'we', 'send', 'a', '88 CNY', 'rich red envelope' and 'five persons'. Next, the target object is determined by filtering the object characteristic words among the words of the text information, and the descriptive characteristic words are determined as the descriptive information.
15 For example, 'a rich red envelope' is determined as the object characteristic word, and 'five persons' and '88 CNY' are determined as the descriptive characteristic words. Next, the presentation information (red envelope) is determined based on the target object and the descriptive information. After determining the presentation information, the processing end may add the presentation information to the live stream data, send the live stream data to the second user of the
20 second terminal, receive the operation data from the second user, and filter out the target user based on the operation data. In an example, the second user may be sorted based on the time when they receive the red envelopes, and the top five second user (corresponding to the number of red envelopes) may be selected as the target users. After that, the target object is assigned to the target users. In the embodiment of the present disclosure, the target object can be quickly presented based
25 on the speech data of the first user, and the first user does not need to manually input various data of the target object, thus facilitating the operation of the first user.

On the basis of the above embodiments, the present disclosure further discloses a data processing method which is applicable to a processing end, and capable of completing a publication of a virtual interaction object by recognizing speech data of a user during a live broadcast, while the
30 user does not need to manually input various data of the virtual interaction object, thus facilitating

the operation of user. In a scene of red envelope release, the virtual interaction object may be understood as a red envelope. As illustrated in Fig. 5, the method comprises:

Step 502: extracting speech data of a first user from live stream data.

Step 504: recognizing the speech data to determine characteristic information corresponding to a virtual interaction object. In a scene of red envelope release, the virtual interaction object may be understood as a red envelope. The virtual interaction object refers to an object that can complete an information interaction between the first user and second user, such as an application program, an application plug-in, etc.

Step 506: determining descriptive information corresponding to the virtual interaction object based on the characteristic information, wherein the descriptive information comprises an object quantity and an object resource. In the scene of red envelope release, the object quantity may be understood as the number of red envelopes, and the object resource may be understood as the amount of red envelopes.

Step 508: calling a virtual interactive plug-in, and configuring attribute information of the virtual interaction object based on the descriptive information.

Step 510: adding the virtual interaction object to the live stream data for display.

The virtual interaction object in the embodiments of the present disclosure may comprise a coupon, a gold coin, etc. in addition to a red envelope. Specifically, in an optional embodiment, the virtual interaction object comprises at least one of a virtual red envelope, a virtual coupon and a virtual gold coin. The steps of this embodiment of the present disclosure are similar to those of the above embodiments. For details, please refer to implementations of the above embodiments, which will not be described in detail here.

In the embodiments of the present disclosure, for example in the scene of red envelope release, the first user may say ‘let’s send a rich red envelope of 88 CNY to see which five persons can get it’ during the live broadcast. The processing end may extract the speech data from the live stream data, convert the speech data into the text information based on the ASR, and then recognize the text information based on the NLP, so as to recognize the corresponding characteristic information of the virtual interaction object. Next, descriptive information (e.g., ‘five persons’ and ‘88 CNY’) corresponding to the virtual interaction object (e.g., the red envelope) is determined based on the characteristic information. Then, a virtual interactive plug-in (e.g., a red envelope plug-in) is called,

and attribute information of the virtual interaction object is configured based on the descriptive information. After that, the virtual interaction object is added to the live stream data and displayed to the second user. The second user may operate on the virtual interaction object to receive the corresponding virtual interaction objects.

5 In the embodiment of the present disclosure, the speech data of the first user can be extracted from the live stream data, and then recognized to determine the characteristic information corresponding to the virtual interaction object. Next, the descriptive information of the virtual interaction object is determined based on the characteristic information, and attribute information of the virtual interaction object is configured based on the descriptive information. After that, the
10 virtual interaction object is added to the live stream data for display. Based on the embodiment of the present disclosure, corresponding virtual interaction object can be generated based on the speech data of the first user and added to the live stream data for display, so that the user does not need to manually input various attribute data corresponding to the virtual interaction object, thus facilitating the operation of the user.

15 The data processing method of the embodiment of the present disclosure may be applied not only to the scene of red envelope release in the live broadcast process, but also to a scene of commodity recommendation in the live broadcast process, where the method of the embodiment of the present disclosure can recognize the speech data of the first user, and form and present the presentation information corresponding to the commodities to the second user for purchases on
20 selection. Specifically, the data processing method of the embodiment of the present disclosure may be applied to the processing end, as illustrated in Fig. 6, comprising:

Step 602: extracting speech data of a first user from live stream data.

Step 604: recognizing the speech data to determine characteristic information corresponding to a commodity object.

25 Step 606: determining a target commodity object and corresponding descriptive information based on the characteristic information, wherein the descriptive information comprises price information and quantity information of the target commodity object.

Step 608: determining link information of the target commodity object based on the target commodity object and the descriptive information. The link information of the commodity object
30 may comprise a link, price information and quantity information of the commodity object, and the

link information of commodity object may be corresponding to a shop, and the second user may click the link information to enter the corresponding shop to complete a transaction of the commodity object. The link information may comprise not only the descriptive information and the link of the commodity object, but also data such as images corresponding to the commodity object.

- 5 In an example, the link information of the commodity object in the embodiment of the present disclosure may be presented through a floating presentation window, and the processing end may add the link, the price information and the quantity information of the commodity object to the floating presentation window for display.

Step 610: adding the link information of the target commodity object to the live stream data for
10 display.

In the embodiment of the present disclosure, the speech data of the first user can be extracted from the live stream data, and recognized to determine the characteristic information corresponding to the commodity object. Next, the target commodity object and the descriptive information are determined based on the characteristic information. Next, the link of the commodity object is
15 acquired to determine the link information of the commodity object in view of the descriptive information. After that, the link information is added to the live stream data for display. The second user may click the link information of a commodity object to enter a corresponding shop to purchase the commodity object. According to the embodiment of the present disclosure, the link information of the commodity object can be generated based on the speech data of the first user, and
20 added to the live stream data for display, so that the user does not need to manually input the data corresponding to the commodity object, thus facilitating the operation of the user.

This embodiment of the present disclosure may be applied not only to a scene of anchor live broadcast, but also to a scene of enterprise video conference. For example, this embodiment of the present disclosure may be applied to a scene of enterprise video conference to recognize speech
25 data of an enterprise manager and extract corresponding characteristic information, so as to determine and present the presentation information corresponding to the target object to other users in the video conference, which facilitates the enterprise manager to release materials and incentives (e.g. red envelope incentives) to other users. Specifically, an embodiment of the present disclosure provides a data processing method, as illustrated in Fig. 7, comprising:

30 Step 702: extracting speech data of a target user from data of a video conference. In an

optional embodiment, live stream data of the video conference may comprise live stream data of a manager and live stream data of employees. The processing end may filter out and determine the live stream data of the manager from the live stream data based on role information of the users.

Step 704: recognizing the speech data to determine characteristic information corresponding to
5 a target object.

Step 706: determining the target object and corresponding descriptive information based on the characteristic information.

Step 708: determining presentation information of the target object based on the target object and the descriptive information.

10 Step 710: adding the presentation information to the data of the video conference for display. The steps of this embodiment of the present disclosure are similar to those of the above embodiments. For details, please refer to the implementations of the above embodiments, which will not be described in detail here.

In the embodiment of the present disclosure, for example in a live broadcast scene of video
15 conference of an enterprise, members of a video may comprise leaders and employees of the enterprise, and data of a video conference may comprise live stream data of management users (e.g., enterprise supervisors and leaders, etc.) and other users of the video conference. This embodiment of the present disclosure can extract live stream data of a management user (a target user) from the data of the video conference, extract speech data, recognize characteristic information in the speech
20 data, determine a target object and descriptive information based on the characteristic information, and determine and present the presentation information of the target object to other users (or specific users or user groups) in the video conference, e.g., issue a reward (e.g., a red envelope reward) to a certain user or users of a certain department in the video conference. According to this embodiment of the present disclosure, corresponding presentation information can be generated and
25 displayed based on the speech data of the management user, and the user does not need to manually input the data corresponding to the target object, thus facilitating the operation of the user.

This embodiment of the present disclosure may be applied not only to a scene of anchor live broadcast, but also to a scene of a live stream teaching activity. For example, this embodiment of the present disclosure may be applied to a scene of a live stream teaching activity to recognize
30 speech data of a teacher and extract corresponding characteristic information to determine

presentation information corresponding to an interaction object for students. It is convenient for the teacher to ask a question (present a question to the students in the form of a floating window) or give rewards (e.g., little red flowers) during the live stream teaching activity. Specifically, an embodiment of the present disclosure provides a data processing method, as illustrated in Fig. 8, comprising:

Step 802: extracting speech data of a target user from data of a live stream teaching activity. In an optional embodiment, the data of a live stream teaching activity may comprise live stream data of a teacher and live stream data of students, and a processing end may filter out the live stream data of the teacher from the data of a live stream teaching activity based on role information of users and determine the speech data of the teacher.

Step 804: recognizing the speech data to determine characteristic information corresponding to a target object.

Step 806: determining the target object and corresponding descriptive information based on the characteristic information.

Step 808: determining presentation information of the target object based on the target object and the descriptive information.

Step 810: adding the presentation information to the data of a live stream teaching activity for display.

The steps of this embodiment of the present disclosure are similar to those of the above embodiments. For details, please refer to the implementations of the above embodiments, which will not be described in detail here.

For example, in a scene of a live stream teaching activity, members of the live stream teaching activity may comprise a teacher and students, and data of a live stream teaching activity may comprise live stream data of the teacher and the students. According to the embodiment of the present disclosure, speech data of the teacher (target user) can be extracted from the data of a live stream teaching activity, and characteristic information in the speech data can be recognized. Next, a target object and descriptive information are determined based on the characteristic information, and presentation information of the target object is determined for the students. For example, the teacher may ask a question to the students through the speech data, and the question may be presented to the students in the form of a floating window (presentation information), thus

facilitating the students to answer the question. For another example, a presentation floating window containing a plurality of student logos (e.g., student avatars) may be preset, comments may be added to the student logos in the floating window based on the teacher's speech (e.g., little red flowers are added to the student avatars), and the presentation floating window may be presented to the students for incentives. According to the embodiment of the present disclosure, the corresponding presentation information can be generated and displayed based on the speech data of the teacher, and the user does not need to manually input the corresponding data, thus facilitating the operation of the user.

It should be noted that, for the sake of simple description, all the method embodiments are expressed as a series of action combinations, but those skilled in the art should know that the embodiments of this specification are not limited by the described action sequence, because according to the embodiments of this specification, some steps can be performed in other sequences or at the same time. Secondly, those skilled in the art should also know that the embodiments described in this specification are all preferred embodiments, and the actions and modules involved are not necessarily essential for the embodiments in this specification.

On the basis of the above embodiments, the present disclosure further provides a data processing apparatus, as illustrated in Fig. 9, comprising the following modules:

a speech data extraction module 902 configured to extract speech data of a first user from live stream data;

a characteristic information extraction module 904 configured to recognize the speech data to determine characteristic information corresponding to a target object;

a characteristic information filtering module 906 configured to determine the target object and corresponding descriptive information based on the characteristic information; and

a presentation information extraction module 908 configured to determine presentation information of the target object based on the target object and the descriptive information.

To sum up, in the embodiment of the present disclosure, the speech data of the first user can be extracted from the live stream data, and recognized to determine the corresponding characteristic information. Next, the target object and the descriptive information are determined based on the characteristic information, and the corresponding presentation information is determined based on the target object and the descriptive information. After that, the presentation information is added to

the live stream data for display. According to the embodiment of the present disclosure, the corresponding presentation information can be generated based on the speech data of the first user, and the presentation information can be added to the live stream data for display, so that the user does not need to manually input corresponding data, thus facilitating the operation of the user.

5 On the basis of the above embodiments, the present disclosure further provides a data processing apparatus, which may specifically comprise the following modules:

 a speech extraction module configured to extract speech data of a first user from live stream data;

 a speech conversion module configured to perform a speech recognition on the speech data to
10 determine corresponding text information;

 a text recognition module configured to perform a semantic analysis on the text information and determine characteristic information corresponding to a target object;

 an object information determination module configured to recognize a first characteristic word in the characteristic information to determine the target object; the first characteristic word may also
15 be called as an object characteristic word;

 a descriptive information determination module configured to recognize a second characteristic word in the characteristic information to determine descriptive information corresponding to the target object; the second characteristic word may also be called as a descriptive characteristic word;

 a target object configuration module configured to set attribute information of the target object
20 based on the descriptive information;

 a presentation information generation module configured to determine presentation information of the target object based on the target object and the attribute information;

 a presentation information adding module configured to add the presentation information to the live stream data for display;

25 an operation data receiving module configured to receive operation data for the presentation information from the second user;

 a target user filtering module configured to analyze the operation data to determine a target user; and

 an object assignment module configured to assign a target object corresponding to the
30 presentation information to the target user.

This embodiment of the present disclosure may be applied to a scene of recognizing speech data of a first user in a live broadcast process. Specifically, in an example of recognizing the speech data of the first user in the live broadcast process and releasing the red envelope, the first user may say 'let's send a rich red envelope of 88 CNY to see which five persons can get it' during the live broadcast. The processing end may extract the speech data and convert the speech data into the text information based on the ASR. Then, based on the NLP, the words in the text information are determined, such as 'we', 'send', 'a', '88 CNY', 'rich red envelope' and 'five persons'. Next, the target object is determined by filtering the object characteristic words among the words of the text information, and the descriptive characteristic words are determined as the descriptive information. For example, 'a rich red envelope' is determined as the object characteristic word, and 'five persons' and '88 CNY' are determined as the descriptive characteristic words. Next, the presentation information (red envelope) is determined based on the target object and the descriptive information. After determining the presentation information, the processing end may add the presentation information to the live stream data, send the live stream data to the second user of the second terminal, receive the operation data from the second user, and filter out the target user based on the operation data. In an example, the second user may be sorted based on the time when they receive the red envelopes, and the top five second user (corresponding to the number of red envelopes) may be selected as the target users. After that, the target object is assigned to the target users. In the embodiment of the present disclosure, the target object can be quickly presented based on the speech data of the first user, and the first user does not need to manually input various data of the target object, thus facilitating the operation of the first user.

On the basis of the above embodiments, the present disclosure further provides a data processing apparatus, as illustrated in Fig. 10, comprising the following modules:

a speech data acquisition module 1002 configured to extract speech data of a first user from live stream data;

a characteristic information acquisition module 1004 configured to recognize the speech data to determine characteristic information corresponding to a virtual interaction object;

a characteristic information processing module 1006 configured to determine descriptive information corresponding to the virtual interaction object based on the characteristic information, wherein the descriptive information comprises an object quantity and an object resource;

an interaction object generation module 1008 configured to call a virtual interactive plug-in and configure attribute information of the virtual interaction object based on the descriptive information; and

an interaction object display module 1010 configured to display the virtual interaction object;
5 the virtual interaction object may be added to the live stream data for display.

To sum up, in the embodiment of the present disclosure, the speech data of the first user can be extracted from the live stream data, and recognized to determine the characteristic information corresponding to the virtual interaction object. Next, the descriptive information of the virtual interaction object is determined based on the characteristic information, and attribute information of
10 the virtual interaction object is configured based on the descriptive information. After that, the virtual interaction object is added to the live stream data for display. According to the embodiment of the present disclosure, corresponding virtual interaction object can be generated based on the speech data of the first user and added to the live stream data for display, so that the user does not need to manually input various attribute data corresponding to the virtual interaction object, thus
15 facilitating the operation of the user.

On the basis of the above embodiments, the present disclosure further provides a data processing apparatus, as illustrated in Fig. 11, comprising the following modules:

a speech data determination module 1102 configured to extract speech data of a first user from live stream data;

20 a characteristic information determination module 1104 configured to recognize the speech data to determine characteristic information corresponding to a commodity object;

a commodity information filtering module 1106 configured to determine a target commodity object and corresponding descriptive information based on the characteristic information, wherein the descriptive information comprises price information and quantity information of the target
25 commodity object;

a link information determination module 1108 configured to determine link information of the target commodity object based on the target commodity object and the descriptive information; and

a link information presentation module 1110 configured to display the link information of the target commodity object; the link information of the target commodity object may be added to the
30 live stream data for display.

To sum up, in the embodiment of the present disclosure, the speech data of the first user can be extracted from the live stream data, and recognized to determine the characteristic information corresponding to the commodity object. Next, the target commodity object and the descriptive information are determined based on the characteristic information. Next, the link of the commodity object is acquired to determine the link information of the commodity object in view of the descriptive information. After that, the link information is added to the live stream data for display. The second user may click the link information of a commodity object to enter a corresponding shop to purchase the commodity object. According to the embodiment of the present disclosure, the link information of the commodity object can be generated based on the speech data of the first user, and added to the live stream data for display, so that the user does not need to manually input the data corresponding to the commodity object, thus facilitating the operation of the user.

On the basis of the above embodiments, the present disclosure further provides a data processing apparatus, as illustrated in Fig. 12, comprising the following modules:

- a speech data acquisition module 1202 configured to extract speech data of a target user from data of a video conference;
- a characteristic information acquisition module 1204 configured to recognize the speech data to determine characteristic information corresponding to a target object;
- a characteristic information screening module 1206 configured to determine the target object and corresponding descriptive information based on the characteristic information;
- a presentation information acquisition module 1208 configured to determine presentation information of the target object based on the target object and the descriptive information; and
- a presentation information output module 1210 configured to add the presentation information to the data of the video conference for display.

In the embodiment of the present disclosure, for example in a live broadcast scene of video conference of an enterprise, members of a video may comprise leaders and employees of the enterprise, and data of the video conference may comprise live stream data of management users (e.g., enterprise supervisors and leaders, etc.) and other users of the video conference. This embodiment of the present disclosure can extract live stream data of a management user (a target user) from the data of the video conference, extract speech data, recognize characteristic information in the speech data, determine a target object and descriptive information based on the

characteristic information, and determine and present the presentation information of the target object to other users (or specific users or user groups) in the video conference, e.g., issue a reward (e.g., a red envelope reward) to a certain user or users of a certain department in the video conference. According to this embodiment of the present disclosure, corresponding presentation
5 information can be generated and displayed based on the speech data of the management user, and the user does not need to manually input the data corresponding to the target object, thus facilitating the operation of the user.

On the basis of the above embodiments, the present disclosure provides a data processing apparatus, as illustrated in Fig. 13, comprising the following modules:

10 a speech data filtering module 1302 configured to extract speech data of a target user from data of a live stream teaching activity;

a characteristic information recognition module 1304 configured to recognize the speech data to determine characteristic information corresponding to a target object;

an object information screening module 1306 configured to determine the target object and
15 corresponding descriptive information based on the characteristic information;

an object information processing module 1308 configured to determine presentation information of the target object based on the target object and the descriptive information; and

a presentation information sending module 1310 configured to add the presentation information to the data of a live stream teaching activity for display.

20 To sum up, for example in a scene of a live stream teaching activity, members of the live stream teaching activity may comprise a teacher and students, and data of a live stream teaching activity may comprise live stream data of the teacher and the students. According to the embodiment of the present disclosure, speech data of the teacher (target user) can be extracted from the data of a live stream teaching activity, and characteristic information in the speech data can be
25 recognized. Next, a target object and descriptive information are determined based on the characteristic information, and presentation information of the target object is determined for the students. For example, the teacher may ask a question to the students through the speech data, and the question may be presented to the students in the form of a floating window (presentation information), thus facilitating the students to answer the question. For another example, a
30 presentation floating window containing a plurality of student logos (e.g., student avatars) may be

preset, comments may be added to the student logos in the floating window based on the teacher's speech (e.g., little red flowers are added to the student avatars), and the presentation floating window may be presented to the students for incentives. According to the embodiment of the present disclosure, the corresponding presentation information can be generated and displayed
5 based on the speech data of the teacher, and the user does not need to manually input the corresponding data, thus facilitating the operation of the user.

An embodiment of the present disclosure further provides a nonvolatile readable storage medium storing one or more modules (programs) which when applied to a device, enables the device to execute instructions of method steps in the embodiment of the present disclosure.

10 An embodiment of the present disclosure provides one or more machine-readable media storing instructions which when executed by one or more processors, enables an electronic device to perform the method of one or more of the above embodiments. In the embodiment of the present disclosure, the electronic device comprises a server, a terminal device, etc.

The embodiment of the present disclosure may be implemented as a device that uses any
15 suitable hardware, firmware, software, or any combination thereof to achieve a desired configuration, and the device may comprise electronic devices such as a server (cluster), a terminal, etc. Fig. 14 schematically illustrates an exemplary apparatus 1400 that may be configured to implement various embodiments described in the present disclosure.

For one embodiment, Fig. 14 illustrates an exemplary apparatus 1400 having one or more
20 processors 1402, a control module (chipset) 1404 coupled to at least one of processor(s) 1402, a memory 1406 coupled to the control module 1404, and a non-volatile memory (NVM)/ storage device 1408 coupled to the control module 1404, one or more input devices 1410 coupled to the control module 1404, and a network interface 1412 coupled to the control module 1404.

The processor 1402 may comprise one or more single-core or multi-core processors, and may
25 comprise any combination of general-purpose processors or special-purpose processors (e.g., graphics processors, application processors, baseband processors, etc.). In some embodiments, the apparatus 1400 may serve as a server, a terminal and other device as described in the embodiments of the present disclosure.

In some embodiments, the apparatus 1400 may comprise one or more computer-readable
30 media (e.g., the memory 1406 or the NVM/storage device 1408) having instructions 1414, and one

or more processors 1402 associated with the one or more computer-readable media and configured to execute the instructions 1414 to implement modules to perform actions described in the present disclosure.

For one embodiment, the control module 1404 may comprise any suitable interface controller
5 to provide any suitable interface to at least one of the processor(s) 1402 and/or any suitable device or component in communication with the control module 1404.

The control module 1404 may comprise a memory controller module to provide an interface to the memory 1406. The memory controller module may be a hardware module, a software module, and/or a firmware module.

10 The memory 1406 may be configured, for example, to load and store data and/or instructions 1414 for the apparatus 1400. For one embodiment, the memory 1406 may comprise any suitable volatile memory, such as a suitable DRAM. In some embodiments, the memory 1406 may comprise a Double-Data-Rate Fourth Generation Synchronous Dynamic Random Access Memory (DDR4 SDRAM).

15 For one embodiment, the control module 1404 may comprise one or more input/output controllers to provide an interface to the NVM/storage device 1408 and input/output device(s) 1410.

For example, the NVM/storage device 1408 may be configured to store data and/or instructions 1414. The NVM/storage device 1408 may comprise any suitable nonvolatile memory
20 (e.g., a flash memory) and/or any suitable nonvolatile storage device(s) (e.g., one or more hard disk drives (HDD), one or more compact disk (CD) drives, and/or one or more digital versatile disk (DVD) drives).

The NVM/storage device 1408 may comprise a storage resource as part of a device on which the apparatus 1400 is installed, or it may be accessed by the device without being part of the device.
25 For example, NVM/storage device 1408 may be accessed via the input/output device(s) 1410 over a network.

The input/output device(s) 1410 may provide an interface for the apparatus 1400 to communicate with any other suitable device, and may comprise a communication component, an audio component, a sensor component, etc. The network interface 1412 may provide an interface
30 for the apparatus 1400 for communication through one or more networks, and the apparatus 1400

may wirelessly communicate with one or more components of a wireless network according to any one of one or more wireless network standards and/or protocols, for example, performing a wireless communication by accessing a wireless network based on communication standards, such as WiFi, 2G, 3G, 4G, 5G, etc., or combinations thereof.

5 For one embodiment, at least one of the processor(s) 1402 may be logically packaged with one or more controllers (e.g., a memory controller module) of the control module 1404. For one embodiment, at least one of the processor(s) 1402 may be logically packaged with one or more controllers of the control module 1404 to form a System-In-Package (SiP). For one embodiment, at least one of the processor(s) 1402 may be logically integrated with one or more controllers of the
10 control module 1404 on a same mold. For one embodiment, at least one of the processor(s) 1402 may be logically integrated with one or more controllers of the control module 1404 on a same mold to form a System-On-Chip (SoC).

In various embodiments, the apparatus 1400 may be, but is not limited to, a terminal device such as a server, a desktop computing device or a mobile computing device (e.g., a laptop
15 computing device, a handheld computing device, a tablet computer, a netbook, etc.). In various embodiments, the apparatus 1400 may have more or fewer components and/or different architectures. For example, in some embodiments, the apparatus 1400 comprises one or more cameras, a keyboard, a Liquid Crystal Display (LCD) screen (including a touch screen display), a non-volatile memory port, a plurality of antennas, a graphics chip, an Application Specific
20 Integrated Circuit (ASIC), and a speaker.

In which, a main control chip may be used as a processor or a control module in a detection apparatus, sensor data, position information, etc. is stored in the memory or the NVM/storage device, a sensor group may serve as an input/output device, and the communication interface may comprise a network interface.

25 An embodiment of the present disclosure further provides an electronic device comprising: a processor; and a memory storing executable codes which, when executed, enables the processor to perform the method according to one or more embodiments of the present disclosure.

An embodiment of the present disclosure further provides one or more machine-readable media storing executable codes which, when executed, enables a processor to perform the method
30 according to one or more embodiments of the present disclosure.

As for the apparatus embodiments, because they are substantially similar to the method embodiments, the descriptions thereof are relatively simple, and relevant contents can be found in some descriptions of the method embodiments.

Each embodiment of this specification is described in a progressive manner. Each embodiment
5 lays an emphasis on its difference from other embodiments, and the same or similar parts of the embodiments can refer to each other.

The embodiments of the present disclosure are described with reference to flow diagrams and/or block diagrams of the method, the terminal device (system) and the computer program product according to the embodiments of the present disclosure. It should be understood that each
10 flow and/or block in a flow diagram and/or block diagram, and combinations of the flows and/or blocks in the flow diagram and/or block diagram can be implemented through computer program instructions. The computer program instructions may be provided to a general computer, a dedicated computer, an embedded processor or a processor of any other programmable data processing device, to produce a machine so that the instructions, which are executed through the
15 computer or the processor of any other programmable data processing device, produce means for realizing the functions specified in one or more flows in the flow diagram and one or more blocks in the block diagram.

These computer program instructions can also be stored in a computer readable memory capable of directing the computer or other programmable data processing devices to operate in a
20 specific manner, such that the instructions stored in the computer readable memory generate a manufactured article including an instruction device that implements the function(s) designated in one flow or a plurality of flows in the flow diagram and/or a block or a plurality of blocks in the block diagram.

These computer program instructions can also be loaded onto a computer or other
25 programmable data processing devices, such that a series of operation steps is executed on the computer or other programmable devices to generate the processing realized by the computer, therefore the instructions executed on the computer or other programmable devices provide the steps for implementing the function designated in one flow or a plurality of flows in the flow chart and/or a block or a plurality of blocks in the block diagram.

30 Although the preferred embodiments among the embodiments of the present disclosure have

been described, those skilled in the art can make additional variations and modifications to these embodiments once acquiring the basic inventive concepts. Therefore, the appended claims are intended to be interpreted as including the preferred embodiments and all variations and modifications falling within the scope of the embodiments of the present disclosure.

5 Finally, it should be noted that the relational terms herein, such as ‘first’ and ‘second’, are only used to distinguish one entity or operation from another entity or operation, and do not necessarily require or imply any such actual relationship or sequence between these entities or operations. Furthermore, the term ‘comprise’, ‘include’ or any other variant is intended to cover non-exclusive inclusions, so that a process, a method, an article or a terminal device comprising a series of
10 elements comprise not only those elements, but also other elements not explicitly listed, or further comprise inherent elements of such process, method, article or terminal device. In a case where there is no further limitation, the elements defined by a sentence ‘comprising a ...’ do not exclude other identical elements existing in the process, method, article or terminal device comprising the elements.

15 A data processing method, a data processing apparatus, an electronic device and a storage medium provided by the present disclosure are described above in detail. Specific examples are applied herein to demonstrate the principles and implementations of the present disclosure. The descriptions of the above embodiments are only used to help understand the method and the core ideas of the present disclosure. Meanwhile, according to the ideas of the present disclosure, those of
20 ordinary skills in the art can make some changes in the implementations and the application scopes. To sum up, the contents of this specification should not be construed as limitations to the present disclosure.

CLAIMS

1. A data processing method, comprising:
extracting speech data of a first user from live stream data;
5 recognizing the speech data to determine characteristic information corresponding to a target object;
determining the target object and corresponding descriptive information based on the characteristic information; and
determining presentation information of the target object based on the target object and the
10 descriptive information.
 2. The method according to claim 1, wherein recognizing the speech data to determine characteristic information corresponding to a target object comprises:
performing a speech recognition on the speech data to determine corresponding text
information;
15 performing a semantic analysis on the text information to determine the characteristic information corresponding to the target object.
 3. The method according to claim 1, wherein determining the target object and corresponding descriptive information based on the characteristic information comprises:
recognizing a first characteristic word in the characteristic information to determine the target
20 object;
recognizing a second characteristic word in the characteristic information to determine the descriptive information corresponding to the target object.
 4. The method according to claim 1, wherein determining presentation information of the target object based on the target object and the descriptive information comprises:
25 setting attribute information of the target object based on the descriptive information;
determining the presentation information of the target object based on the target object and the attribute information.
 5. The method according to claim 1, further comprising:
adding the presentation information to the live stream data for display.
 - 30 6. The method according to claim 5, wherein adding the presentation information to the live stream data for display comprises:
acquiring permission levels of second users;
determining a second user meeting a permission condition; and
adding the presentation information to live stream data related to the second user meeting the

permission condition, so as to present the presentation information to the second user meeting the permission condition.

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

receiving operation data for the presentation information from the second user;

5 analyzing the operation data to determine a target user; and

assigning a target object corresponding to the presentation information to the target user.

8. A data processing method, comprising:

extracting speech data of a first user from live stream data;

10 recognizing the speech data to determine characteristic information corresponding to a virtual interaction object;

determining descriptive information corresponding to the virtual interaction object based on the characteristic information, wherein the descriptive information comprises an object quantity and an object resource;

15 calling a virtual interactive plug-in, and configuring attribute information of the virtual interaction object based on the descriptive information; and

adding the virtual interaction object to the live stream data for display.

9. The method according to claim 8, wherein the virtual interaction object comprises at least one of a virtual red envelope, a virtual coupon and a virtual gold coin.

10. A data processing method, comprising:

20 extracting speech data of a first user from live stream data;

recognizing the speech data to determine characteristic information corresponding to a commodity object;

25 determining a target commodity object and corresponding descriptive information based on the characteristic information, wherein the descriptive information comprises price information and quantity information of the target commodity object;

determining link information of the target commodity object based on the target commodity object and the descriptive information; and

adding the link information of the target commodity object to the live stream data for display.

11. A data processing method, comprising:

30 extracting speech data of a target user from data of a video conference;

recognizing the speech data to determine characteristic information corresponding to a target object;

determining the target object and corresponding descriptive information based on the characteristic information;

determining presentation information of the target object based on the target object and the descriptive information; and

adding the presentation information to the data of the video conference for display.

12. A data processing method, comprising:

5 extracting speech data of a target user from data of a live stream teaching activity;

recognizing the speech data to determine characteristic information corresponding to a target object;

determining the target object and corresponding descriptive information based on the characteristic information;

10 determining presentation information of the target object based on the target object and the descriptive information; and

adding the presentation information to the data of the live stream teaching activity.

13. A data processing apparatus, comprising:

15 a speech data extraction module configured to extract speech data of a first user from live stream data;

a characteristic information extraction module configured to recognize the speech data to determine characteristic information corresponding to a target object;

a characteristic information filtering module configured to determine the target object and corresponding descriptive information based on the characteristic information; and

20 a presentation information extraction module configured to determine presentation information of the target object based on the target object and the descriptive information.

14. A data processing apparatus, comprising:

a speech data acquisition module configured to extract speech data of a first user from live stream data;

25 a characteristic information acquisition module configured to recognize the speech data to determine characteristic information corresponding to a virtual interaction object;

a characteristic information processing module configured to determine descriptive information corresponding to the virtual interaction object based on the characteristic information, wherein the descriptive information comprises an object quantity and an object resource;

30 an interaction object generation module configured to call a virtual interactive plug-in and configure attribute information of the virtual interaction object based on the descriptive information; and

an interaction object display module configured to add the virtual interaction object to the live stream data for display.

15. A data processing apparatus, comprising:

a speech data determination module configured to extract speech data of a first user from live stream data;

a characteristic information determination module configured to recognize the speech data to
5 determine characteristic information corresponding to a commodity object;

a commodity information filtering module configured to determine a target commodity object and corresponding descriptive information based on the characteristic information, wherein the descriptive information comprises price information and quantity information of the target commodity object;

10 a link information determination module configured to determine link information of the target commodity object based on the target commodity object and the descriptive information; and

a link information presentation module configured to add the link information of the target commodity object to the live stream data for display.

16. A data processing apparatus, comprising:

15 a speech data acquisition module configured to extract speech data of a target user from data of a video conference;

a characteristic information acquisition module configured to recognize the speech data to determine characteristic information corresponding to a target object;

a characteristic information screening module configured to determine the target object and
20 corresponding descriptive information based on the characteristic information;

a presentation information acquisition module configured to determine presentation information of the target object based on the target object and the descriptive information; and

a presentation information output module configured to add the presentation information to the data of the video conference for display.

25 17. A data processing apparatus, comprising:

a speech data filtering module configured to extract speech data of a target user from data of the live stream teaching activity;

a characteristic information recognition module configured to recognize the speech data to determine characteristic information corresponding to a target object;

30 an object information screening module configured to determine the target object and corresponding descriptive information based on the characteristic information;

an object information processing module configured to determine presentation information of the target object based on the target object and the descriptive information; and

a presentation information sending module configured to add the presentation information to

the data of the live stream teaching activity for display.

18. An electronic device comprising:

a processor; and

a memory storing executable codes which, when executed, enables the processor to perform

5 the method according to any one of claims 1 to 12.

19. A machine-readable media storing executable codes which, when executed, enables a processor to perform the method according to any one of claims 1 to 12.