(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) **EP 4 398 602 A1**

(12) **EUROPEAN PATENT APPLICATION**
published in accordance with Art. 153(4) EPC

(43) Date of publication:
10.07.2024 Bulletin 2024/28

(21) Application number: **22880085.0**

(22) Date of filing: **15.09.2022**

(51) International Patent Classification (IPC):
H04R 5/02 *(2006.01)*

(52) Cooperative Patent Classification (CPC):
G06F 18/211; H04R 5/02; H04S 7/00

(86) International application number:
PCT/CN2022/119134

(87) International publication number:
WO 2023/061145 (20.04.2023 Gazette 2023/16)

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
KH MA MD TN

(30) Priority: **12.10.2021 CN 202111186340**

(71) Applicant: **Huawei Technologies Co., Ltd.**
Shenzhen, Guangdong 518129 (CN)

(72) Inventors:
- **CUI, Liwei**
  Shenzhen, Guangdong 518129 (CN)
- **CAI, Shuanglin**
  Shenzhen, Guangdong 518129 (CN)
- **ZHANG, Jingjing**
  Shenzhen, Guangdong 518129 (CN)
- **OU, Bitao**
  Shenzhen, Guangdong 518129 (CN)
- **SUN, Rui**
  Shenzhen, Guangdong 518129 (CN)
- **XU, Jianfeng**
  Shenzhen, Guangdong 518129 (CN)
- **YANG, Jun**
  Shenzhen, Guangdong 518129 (CN)

(74) Representative: **Körber, Martin Hans**
Mitscherlich PartmbB
Patent- und Rechtsanwälte
Karlstraße 7
80333 München (DE)

(54) METHOD FOR CONSTRUCTING STEREO SPEAKER SYSTEM, AND RELATED APPARATUS

(57) This application provides a method for building a stereo speaker system and a related apparatus. The method relates to fields such as intelligent terminals and human-computer interaction. In the method, a user may be supported in moving a speaker to generate a first motion state, and when the first running state of the speaker matches a preset action, building of a stereo speaker system may be triggered. For example, the user may hold the speaker and draw a moving pattern, and when the drawn pattern is a specified pattern, a stereo speaker system may be built by using the speaker and another speaker. This manner of building is easy to operate for users. In addition, in this application, a corresponding condition is set to avoid falsely triggering building of a speaker system, and a corresponding condition is further set to avoid falsely building a speaker system. This improves accuracy of triggering building of a speaker system and accuracy of building of the speaker system. Further, in this application, before the user performs an operation, the user may be further prompted for performing an operation, so that the user is guided to create the speaker system.

EP 4 398 602 A1

(Cont. next page)

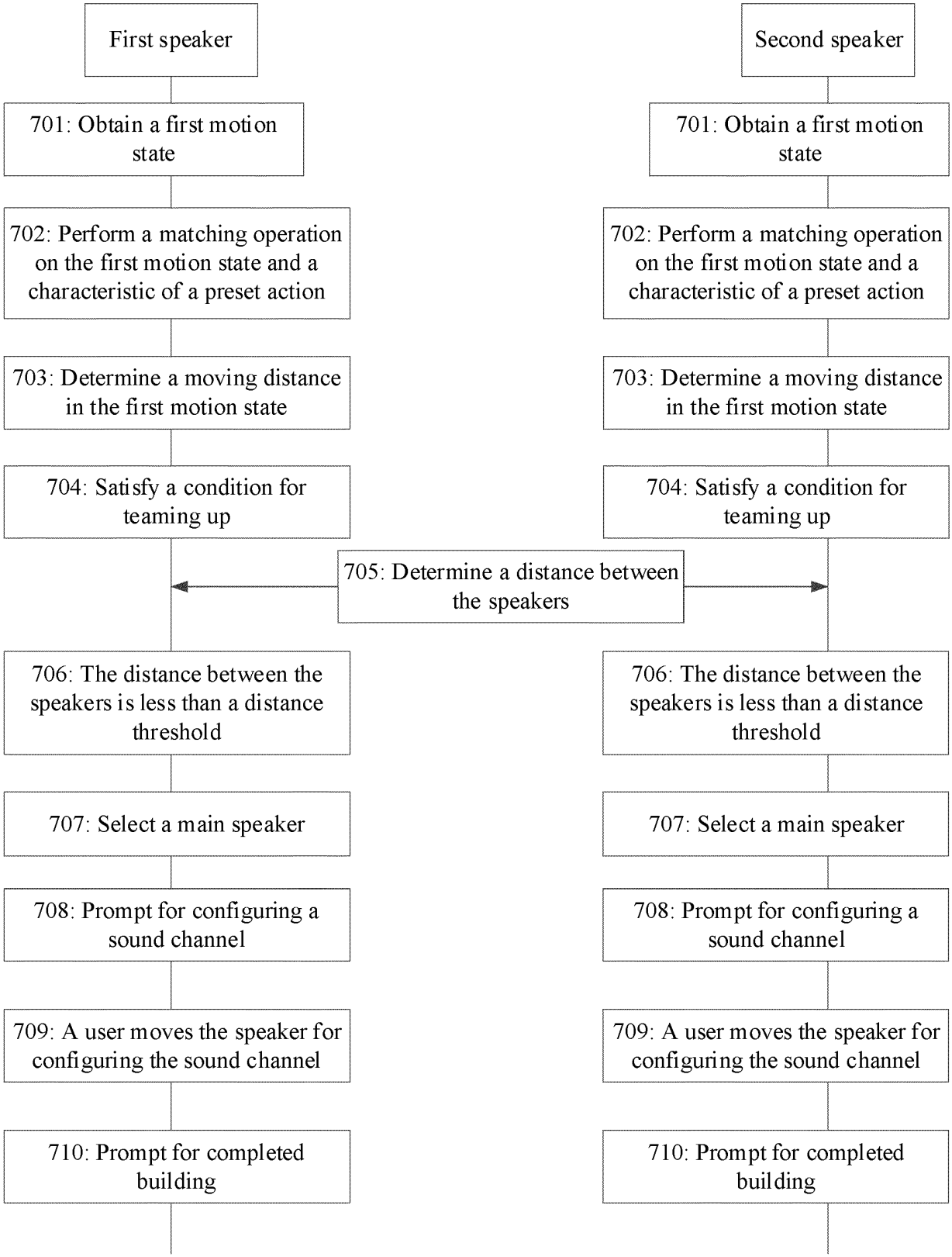
First speaker
Second speaker
701: Obtain a first motion state
701: Obtain a first motion state
702: Perform a matching operation on the first motion state and a characteristic of a preset action
702: Perform a matching operation on the first motion state and a characteristic of a preset action
703: Determine a moving distance in the first motion state
703: Determine a moving distance in the first motion state
704: Satisfy a condition for teaming up
704: Satisfy a condition for teaming up
705: Determine a distance between the speakers
706: The distance between the speakers is less than a distance threshold
706: The distance between the speakers is less than a distance threshold
707: Select a main speaker
707: Select a main speaker
708: Prompt for configuring a sound channel
708: Prompt for configuring a sound channel
709: A user moves the speaker for configuring the sound channel
709: A user moves the speaker for configuring the sound channel
710: Prompt for completed building
710: Prompt for completed building

FIG. 7

Description

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to Chinese Patent Application No. 202111186340.4, filed with the China National Intellectual Property Administration on October 12, 2021 and entitled "METHOD FOR BUILDING STEREO SPEAKER SYSTEM AND RELATED APPARATUS", which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

[0002] This application relates to the field of terminal technologies, and in particular, to a method for building a stereo speaker system and a related apparatus.

BACKGROUND

[0003] Speakers have become common electronic products in people's life. For convenience of enjoying high-quality audio anytime and anywhere, portable speakers have become favorite products of electronics enthusiasts.

[0004] When there are a plurality of speakers, a stereo speaker system may be built to create a stereo effect. For example, two speakers may be used as a left sound channel and a right sound channel respectively, and may provide a more stereoscopic effect than one speaker.

[0005] A conventional manner of building a stereo speaker system is as follows: First, a corresponding application (application, APP) is used to scan a nearby speaker, and if a plurality of speakers of a same type are found through scanning, the stereo speaker system may be built. As shown in FIG. 1, for example, two speakers are found through scanning, and a user may manually select, in an interface, a speaker to be set as a left sound channel and a speaker to be set as a right sound channel, to build a stereo speaker system.

[0006] Another implementation is to simultaneously press keys on speakers to trigger building of a stereo speaker system.

[0007] However, an operation manner of building a stereo speaker system is lack of variation, and a new method for building a stereo speaker system is needed.

SUMMARY

[0008] An objective of this application is to provide a method for building a stereo speaker system and a related apparatus, to resolve a problem in a conventional technology that an operation manner of building a stereo speaker system is lack of variation.

[0009] The foregoing objective and other objectives are achieved with features in the independent claims. Further implementations are embodied in the dependent claims, the description, and the accompanying drawings.

[0010] According to a first aspect, a method for building a stereo speaker system is provided, and is applied to a first speaker. The method includes:

obtaining a first motion state of the first speaker;
if the first motion state matches a characteristic of a preset action, searching for a second speaker; and
if the second speaker is found, building a stereo speaker system with the second speaker.

[0011] In this way, in this application, a user enables the speaker to generate a specific motion state, and when the motion state matches the characteristic of the preset action, the stereo speaker system can be built. Therefore, for the user, the stereo speaker system can be built simply by holding the speaker and performing a corresponding action, allowing for easy operation.

[0012] In a possible design, before the obtaining a first motion state of the first speaker, the method further includes:

prompting a user for the preset action that needs to be performed for building the stereo speaker system.

[0013] In this way, the prompt for the action brings convenience for the user to know an action needs to be performed to build the speaker system, and helps the user understand and implement an operation of building the speaker system.

[0014] In a possible design, the obtaining a first motion state of the first speaker includes:

generating a first acceleration sequence of the first speaker based on acceleration information of the first speaker, where the first acceleration sequence of the first speaker stores first indication information arranged in a time sequence, and the first indication information is used to express a correspondence between an acceleration and duration of the acceleration.

[0015] The characteristic of the preset action includes a first sequence template, and the method further includes:

performing a matching operation on the first sequence template and the first acceleration sequence of the first speaker; and
if a quantity of times the first acceleration sequence of the first speaker matches the first sequence template is greater than or equal to a first specified quantity of times, determining that the first motion state matches the characteristic of the preset action; or
if a quantity of times the first acceleration sequence of the first speaker matches the first sequence template is less than a first specified quantity of times, determining that the first motion state does not match the characteristic of the preset action.

[0016] In this way, in this application, whether the acquired first running state matches the characteristic of the preset action is determined in a manner of template matching. The manner of template matching is easy to implement, and the template included in the character-

istic of the preset action is expressed based on time domain information, and therefore can be applicable to different actions. An application scope is not limited.

[0017] In a possible design, the obtaining a first motion state of the first speaker includes:

generating a second acceleration sequence of the first speaker based on acceleration information of the first speaker, where the second acceleration sequence of the first speaker stores second indication information, and the second indication information is used to express an acceleration and frequency domain information of the acquired acceleration.

[0018] The characteristic of the preset action includes a second sequence template, and the method further includes:

performing a matching operation on the second sequence template and the second acceleration sequence of the first speaker; and

if a quantity of times the second acceleration sequence of the first speaker matches the second sequence template is greater than or equal to a second specified quantity of times, determining that the first motion state matches the characteristic of the preset action; or

if a quantity of times the second acceleration sequence of the first speaker matches the second sequence template is less than a second specified quantity of times, determining that the first motion state does not match the characteristic of the preset action.

[0019] In this way, in this application, whether the acquired first running state matches the characteristic of the preset action is determined in a manner of template matching. The manner of template matching is easy to implement, and the template included in the characteristic of the preset action is expressed based on frequency domain information, and therefore can be applicable to different actions. An application scope is not limited.

[0020] In a possible design, the method further includes:

determining a moving distance of the first speaker based on the first motion state of the first speaker.

[0021] Before the searching for a second speaker, the method further includes:

determining that the moving distance of the first speaker is greater than a specified distance.

[0022] In this way, in this application, building of the stereo speaker system is triggered only when the moving distance of the first speaker is large enough, to improve accuracy of triggering building of the stereo speaker system and avoid false triggering as much as possible.

[0023] In a possible design, the building a stereo speaker system with the second speaker includes:

determining, according to a main speaker selection rule, whether the first speaker is used as a main speaker in the stereo speaker system; and

configuring a sound channel for the first speaker.

[0024] In this way, in this application, the main speaker can be autonomously selected according to a specified rule, and a sound channel can be configured for each speaker autonomously.

[0025] In a possible design, the main speaker selection rule includes at least one of the following rules:

selecting a speaker connected to a network as the main speaker;

selecting a speaker configured with a network but not connected to the network as the main speaker;

selecting a speaker connected to an intelligent terminal device as the main speaker;

selecting a speaker that first matches the characteristic of the preset action as the main speaker; and

selecting a speaker with a largest media access control (media access control, MAC) address as the main speaker.

[0026] The selection rule is simple and easy to implement, and can select the main speaker as unambiguously as possible.

[0027] In a possible design, the configuring a sound channel for the first speaker includes:

prompting the user for configuring the sound channel for the first speaker;

obtaining a second motion state of the first speaker; and

configuring a sound channel corresponding to the second motion state as the sound channel of the first speaker.

[0028] In this way, in this application, the user may configure a sound channel for each speaker by performing a specific action. For example, one sound channel may be configured with one shake, and the other sound channel may be configured with two shakes, so that flexibility and operation convenience of sound channel configuration are improved.

[0029] In a possible design, the prompting the user for configuring the sound channel for the first speaker includes:

prompting, by using at least one of a sound effect, a light effect, and a screen display, the user for configuring the sound channel for the first speaker.

[0030] In this way, in this application, the user may be prompted by using a sound effect, a light effect, or a screen display that is easy to implement, to guide the user to complete the operation of building the stereo speaker system.

[0031] In a possible design, the second motion state includes at least one of the following parameters:

a quantity of times the first speaker is shaken, a speed of a shake, an acceleration of the shake, a direction of

the shake, a moving distance, a quantity of collisions of the first speaker, and a motion trail of the first speaker.

[0032] In this way, in this application, the user may be supported to manually control the speaker in a manner that is easy to operate to complete sound channel configuration.

[0033] In a possible design, the method further includes:

determining a position relationship with the second speaker; and

if it is determined that the position relationship with the second speaker is a specified position relationship, performing the operation of building the stereo speaker system with the second speaker.

[0034] In this way, in this application, accuracy of building the speaker system can be ensured based on the position relationship, to avoid falsely building the speaker system as much as possible.

[0035] In a possible design, the specified position relationship includes that a distance between the first speaker and the second speaker is less than a distance threshold.

[0036] In this way, in this application, the speaker system is built by using speakers close to each other, allowing for simple and easy implementation, and ensuring a stereo effect as it is ensured that the speaker system is built by using the speakers within a specific distance range.

[0037] According to a second aspect, an embodiment of this application further provides a first speaker. The speaker includes:

an obtaining module, configured to obtain a first motion state of the first speaker;

a searching module, configured to: if the first motion state matches a characteristic of a preset action, search for a second speaker; and

a system building module, configured to: if the second speaker is found, build a stereo speaker system with the second speaker.

[0038] In a possible design, the speaker further includes:

a prompting module, configured to: before the first motion state of the first speaker is obtained, prompt a user for the preset action that needs to be performed for building the stereo speaker system.

[0039] In a possible design, the obtaining module is specifically configured to:

generate a first acceleration sequence of the first speaker based on acceleration information of the first speaker, where the first acceleration sequence of the first speaker stores first indication information arranged in a time sequence, and the first indication information is used to express a correspondence between an acceleration and duration of the acceleration.

[0040] The characteristic of the preset action includes a first sequence template, and the speaker further includes:

a first matching module, configured to: perform a matching operation on the first sequence template and the first acceleration sequence of the first speaker; and

if a quantity of times the first acceleration sequence of the first speaker matches the first sequence template is greater than or equal to a first specified quantity of times, determine that the first motion state matches the characteristic of the preset action; or

if a quantity of times the first acceleration sequence of the first speaker matches the first sequence template is less than a first specified quantity of times, determine that the first motion state does not match the characteristic of the preset action.

[0041] In a possible design, the obtaining module is specifically configured to:

generate a second acceleration sequence of the first speaker based on acceleration information of the first speaker, where the second acceleration sequence of the first speaker stores second indication information, and the second indication information is used to express an acceleration and frequency domain information of the acquired acceleration.

[0042] The characteristic of the preset action includes a second sequence template, and the speaker further includes:

a second matching module, configured to: perform a matching operation on the second sequence template and the second acceleration sequence of the first speaker; and

if a quantity of times the second acceleration sequence of the first speaker matches the second sequence template is greater than or equal to a second specified quantity of times, determine that the first motion state matches the characteristic of the preset action; or

if a quantity of times the second acceleration sequence of the first speaker matches the second sequence template is less than a second specified quantity of times, determine that the first motion state does not match the characteristic of the preset action.

[0043] In a possible design, the speaker further includes:

a distance determining module, configured to determine a moving distance of the first speaker based on the first motion state of the first speaker.

[0044] The speaker further includes:

the distance determining module, configured to: before the search for the second speaker, determine that the moving distance of the first speaker is greater than a

specified distance.

[0045] In a possible design, the system building module is specifically configured to:

determine, according to a main speaker selection rule, whether the first speaker is used as a main speaker in the stereo speaker system; and
configure a sound channel for the first speaker.

[0046] In a possible design, the main speaker selection rule includes at least one of the following rules:

selecting a speaker connected to a network as the main speaker;
selecting a speaker configured with a network but not connected to the network as the main speaker;
selecting a speaker connected to an intelligent terminal device as the main speaker;
selecting a speaker that first matches the characteristic of the preset action as the main speaker; and
selecting a speaker with a largest media access control address MAC address as the main speaker.

[0047] In a possible design, the system building module is specifically configured to:

prompt the user for configuring the sound channel for the first speaker;
obtain a second motion state of the first speaker; and
configure a sound channel corresponding to the second motion state as the sound channel of the first speaker.

[0048] In a possible design, the system building module is specifically configured to:
prompt, by using at least one of a sound effect, a light effect, and a screen display, the user for configuring the sound channel for the first speaker.

[0049] In a possible design, the second motion state includes at least one of the following parameters:
a quantity of times the first speaker is shaken, a speed of a shake, an acceleration of the shake, a direction of the shake, a moving distance, a quantity of collisions of the first speaker, and a motion trail of the first speaker.

[0050] In a possible design, the system building module is further configured to:

determine a position relationship with the second speaker; and
if it is determined that the position relationship with the second speaker is a specified position relationship, perform the operation of building the stereo speaker system with the second speaker.

[0051] In a possible design, the specified position relationship includes that a distance between the first speaker and the second speaker is less than a distance threshold.

[0052] According to a third aspect, an embodiment of this application further provides a speaker. The speaker includes one or more processors, one or more memories, one or more loudspeakers, one or more microphones, and a communication module. The one or more microphones are configured to acquire a sound signal; the communication module is configured to communicate with another speaker; the one or more loudspeakers are configured to emit a sound signal; and the one or more processors are coupled to the one or more memories. The one or more memories are configured to store a computer-executable program code. The program code includes instructions, and when the one or more processors execute the instructions, the speaker is enabled to execute the technical solutions according to the first aspect and any possible design in the first aspect.

[0053] According to a fourth aspect, an embodiment of this application provides a chip. The chip includes a processor and an interface. The interface is configured to receive a code instruction, and transmit the received code instruction to the processor. The chip is coupled to a memory of a speaker, so that the processor performs the technical solutions according to the first aspect in embodiments of this application and any possible design in the first aspect. In embodiments of this application, "coupling" means that two components are directly or indirectly combined with each other.

[0054] According to a fifth aspect, an embodiment of this application provides a speaker system. The speaker system includes one or more speakers. At least one speaker is the speaker according to the second aspect or the third aspect, and the speaker may perform all or some steps performed by the first speaker in the first aspect.

[0055] According to a sixth aspect, an embodiment of this application provides a speaker system. The speaker system includes a first speaker and at least one second speaker. The first speaker and the second speaker are disposed at different positions; the first speaker and the second speaker can communicate with each other; and the first speaker is the speaker according to the second aspect or the third aspect (the speaker may be the first speaker in the first aspect and the technical solution of any possible design in the first aspect).

[0056] According to a seventh aspect, an embodiment of this application provides a computer-readable storage medium. The computer-readable storage medium includes a computer program. When the computer program is run on a computer, the computer is enabled to perform the technical solution in the first aspect or any possible design in the first aspect in embodiments of this application.

[0057] According to an eighth aspect, an embodiment of this application provides a program product. When the computer program product is run on a computer, the computer is enabled to perform the technical solution in the first aspect or any possible design in the first aspect in embodiments of this application.

[0058] In addition, for technical effects brought by any possible design manner in the second aspect to the eighth aspect, refer to technical effects brought by different design manners in a related section describing the method. Details are not described herein.

BRIEF DESCRIPTION OF DRAWINGS

[0059]

FIG. 1 is a schematic diagram of an application scenario according to an embodiment of this application;
FIG. 2 is a first schematic diagram of a structure of a speaker according to an embodiment of this application;
FIG. 3 is a second schematic diagram of a structure of a speaker according to an embodiment of this application;
FIG. 4A is a third schematic diagram of a structure of a speaker according to an embodiment of this application;
FIG. 4B is a schematic diagram of a speaker determining a user orientation according to an embodiment of this application;
FIG. 5 is a schematic diagram of an operation that can be performed by a user by holding a speaker according to an embodiment of this application;
FIG. 6 is a schematic diagram of customizing a preset action according to an embodiment of this application;
FIG. 7 is a schematic flowchart of a method for building a stereo speaker system according to an embodiment of this application;
FIG. 8 is a schematic diagram of a first sequence sub-template according to an embodiment of this application;
FIG. 9 is a schematic diagram of a process of matching with a characteristic of a preset action according to an embodiment of this application;
FIG. 10 is a schematic diagram of sound channel configuration when there are two speakers according to an embodiment of this application;
FIG. 11 is a schematic diagram of sound channel configuration when there are three speakers according to an embodiment of this application;
FIG. 12 is a schematic diagram of sound channel configuration when there are six speakers according to an embodiment of this application;
FIG. 13 is a schematic diagram of sound channel configuration when there are twelve speakers according to an embodiment of this application;
FIG. 14 is another schematic flowchart of a method for building a stereo speaker system according to an embodiment of this application;
FIG. 15 is still another schematic flowchart of a method for building a stereo speaker system according to an embodiment of this application;
FIG. 16 is a schematic diagram of an application scenario in which a stereo speaker system obtains a network resource according to an embodiment of this application;
FIG. 17 is a schematic diagram of another application scenario in which a stereo speaker system obtains a network resource according to an embodiment of this application; and
FIG. 18 is another schematic diagram of a structure of a speaker according to an embodiment of this application.

DESCRIPTION OF EMBODIMENTS

[0060] Technical solutions in embodiments of this application are described below clearly and completely with reference to the accompanying drawings in embodiments of this application.

[0061] "A plurality of" in embodiments of this application means two or more. It should be noted that, in the descriptions of embodiments of this application, terms such as "first" and "second" are merely used for distinguishing in descriptions, but should not be understood as indicating or implying relative importance, or should not be understood as indicating or implying a sequence.

[0062] Terms used in the following embodiments are merely intended to describe specific embodiments, but are not intended to limit this application. As used in the description and the appended claims of this application, the terms "one", "a", "the", "this" of singular forms are intended to further include expressions such as "one or more", unless otherwise specified in the context clearly. It should be further understood that, in embodiments of this application, "one or more" means one, two, or more; and "and/or" describes an association relationship between associated objects, and represents that three relationships may exist. For example, A and/or B may represent a case in which only A exists, both A and B exist, or only B exists, where A and B may be singular or plural. The character "/" usually indicates an "or" relationship between associated objects.

[0063] Reference to "an embodiment", "some embodiments", or the like described in this specification indicates that one or more embodiments of this application includes/include a specific feature, structure, or characteristic described with reference to those embodiments. Therefore, statements such as "in an embodiment", "in some embodiments", "in some other embodiments", and "in other embodiments" that appear at different places in this specification do not necessarily mean reference to a same embodiment. Instead, the statements mean "one or more but not all of the embodiments", unless otherwise particularly emphasized in another manner. The terms "include", "have", and their variants all mean "include but are not limited to", unless otherwise particularly emphasized in another manner.

[0064] FIG. 1 is a schematic diagram of an application scenario according to an embodiment of this application. FIG. 1 shows building of a stereo speaker system by

using each of speakers used by a plurality of people in a gathering. As shown in FIG. 1, it is assumed that there are two speakers (including a speaker 1 and a speaker 2). A user A holds the speaker 1, and a user B holds the speaker 2. The user A and the user B may each shake respective speakers at the same time, so that the speaker 1 and the speaker 2 may be used to build a stereo speaker system. Therefore, the stereo speaker system may be built with a shake, increasing fun and convenience of building the system.

[0065] FIG. 2 is a function block diagram of a speaker according to an embodiment of this application. In some embodiments, a speaker 100 may include one or more input devices (input device) 101, one or more output devices (output device) 102, and one or more processors (processor) 103. The input device 101 may detect various types of input signals (which may be referred to as inputs for short), and the output device 102 may provide various types of output information (which may be referred to as outputs for short). The processor 103 may receive an input signal from the one or more input devices 101, generate output information in response to the input signal, and output the output information by using the one or more output devices 102.

[0066] In some embodiments, the one or more input devices 101 may detect various types of inputs and provide signals (for example, input signals) corresponding to the detected inputs; and then the one or more input devices 101 may provide the input signals for the one or more processors 103. In some examples, the one or more input devices 101 may include any component or part that can detect an input signal. For example, the input device 101 may include an audio sensor (for example, one or more microphones), an acceleration sensor, a distance sensor, an optical or visual sensor (for example, a camera, a visible light sensor, or an invisible light sensor), an optical proximity sensor, a touch sensor, a pressure sensor, a mechanical device (for example, a watch crown, a switch, a button, or a key), a temperature sensor, a communication device (for example, a wired or wireless communication apparatus), and the like; or the input device 101 may be some combinations of the foregoing components. In this embodiment of this application, data acquired by the acceleration sensor may be used to determine whether a user has shaken or touched a speaker, to build a stereo speaker.

[0067] In some embodiments, the one or more output devices 102 may provide various types of outputs. For example, the one or more output devices 102 may receive one or more signals (for example, an output signal provided by the one or more processors 103) and provide an output corresponding to the signal. In some examples, the output device 102 may include any appropriate component or part configured to provide an output. For example, the output device 102 may include an audio output device (for example, one or more loudspeakers), a visual output device (for example, one or more lights or displays), a tactile output device, a communication device (for example, a wired or wireless communication device), and the like; or the output device 102 may be some combinations of the foregoing components.

[0068] In some embodiments, the one or more processors 103 may be coupled to the input device 101 and the output device 102. The processor 103 may communicate with the input device 101 and the output device 102. For example, the one or more processors 103 may receive an input signal (for example, an input signal corresponding to an input detected by the input device 101) from the input device 101. The one or more processors 103 may parse the received input signal to determine whether to provide one or more corresponding outputs in response to the input signal. If it is determined to provide the one or more corresponding outputs in response to the input signal, the one or more processors 103 may send an output signal to the output device 102 to provide an output.

[0069] FIG. 3 is a function block diagram of a speaker 300 according to another embodiment of this application. The speaker 300 may be an example of the speaker 100 described in FIG. 2. As shown in FIG. 3, the speaker 300 includes a microphone 301, a loudspeaker 302, a processor 303, a memory 304, a communication module 305, a sensor module 306, and a light 307. It can be understood that the components shown in FIG. 3 do not constitute a specific limitation on the speaker 300. The speaker 300 may further include more or fewer components than those shown in the figure, or some components may be combined, or some components may be split, or there may be a different component arrangement.

[0070] The processor 303 may include one or more processing units. For example, the processor 303 may include an application processor (application processor, AP), a modem processor, a graphics processing unit (graphics processing unit, GPU), an image signal processor (image signal processor, ISP), a controller, a memory, a video codec, a digital signal processor (digital signal processor, DSP), a baseband processor, and/or a neural-network processing unit (neural-network processing unit, NPU). Different processing units may be independent components, or may be integrated into one or more processors. The controller may be a nerve center and a command center of the speaker 300. The controller may generate an operation control signal based on an instruction operation code and a time sequence signal, to complete control of instruction reading and instruction execution. In some other embodiments, a memory may be further disposed in the processor 303, and is configured to store instructions and data. In some embodiments, the memory in the processor 303 is a cache memory. The memory may store instructions or data that has just been used or is cyclically used by the processor 303. If the processor 303 needs to use the instructions or the data again, the processor 303 may directly invoke the instructions or the data from the memory. This avoids repeated access, reduces waiting time of the processor 303, and improves system efficiency. The processor 303

may run a software code/module for the method for building a stereo speaker system provided in some embodiments of this application, to implement a speaker control function.

[0071] The microphone 301, also referred to as a "mike" or "mic", is configured to: acquire a sound signal (for example, acquire a sound made by a user), and convert the sound signal into an electrical signal. In some embodiments, one or more microphones 301, for example, a microphone array, may be disposed on the speaker 300. In some other embodiments, in addition to acquiring the sound signal, the microphone 301 may further implement a noise reduction function for the sound signal, or may further identify a source of the sound signal, implement a directional recording function, and the like.

[0072] The loudspeaker 302, also referred to as a "loudspeaker unit", is configured to convert an electrical audio signal into a sound signal. The speaker 300 may play a sound signal, for example, music, by using the loudspeaker 302.

[0073] In some embodiments, the microphone 301 and the loudspeaker 302 are coupled to the processor 303. For example, after receiving a sound signal, the microphone 301 sends the sound signal or an electrical audio signal converted from the sound signal to the processor 303. The processor 303 determines whether to respond to the sound signal or the electrical audio signal. If it is determined to respond to the sound signal or the electrical audio signal, a corresponding output signal is output, for example, music is played by using the loudspeaker 302.

[0074] The memory 304 may be configured to store a computer-executable program code. The executable program code includes instructions. By running the instructions stored in the memory, the processor 303 performs various functional applications and data processing of the speaker 300. The memory may include a high-speed random access memory, or may include a non-volatile memory, for example, at least one magnetic disk memory, flash memory, or universal flash storage (universal flash storage, UFS). This is not limited in this embodiment of this application. In some embodiments, the memory 304 may store information, for example, a "wake word". In some other embodiments, the memory 304 may alternatively store audio information (for example, a song, a comic dialog, or a storytelling show).

[0075] The communication module 305 may be a wireless communication module (for example, Bluetooth or radio). The speaker 300 is connected to another device, for example, another speaker, a mobile phone, or a television, by using the communication module 305.

[0076] The sensor module 306 may include a barometric pressure sensor 306A, a temperature sensor 306B, an acceleration sensor 306C, and the like. It should be understood that FIG. 3 merely lists several examples of sensors. In actual application, the speaker 300 may alternatively include more or fewer sensors, or a sensor listed above is replaced with another sensor with a same or similar function. This is not limited in this embodiment of this application.

[0077] The barometric pressure sensor 306A is configured to measure barometric pressure. In some embodiments, the processor 303 may be coupled to the barometric pressure sensor 306A, and use a barometric pressure value measured by the barometric pressure sensor 306A to assist in calculation, for example, calculation of an attenuation coefficient of a sound.

[0078] The temperature sensor 306B is configured to detect a temperature. In some embodiments, the processor 303 may be coupled to the temperature sensor 306B, and use a temperature value measured by the temperature sensor 306B to assist in calculation, for example, calculation of an attenuation coefficient of a sound.

[0079] The acceleration sensor 306C is configured to acquire acceleration information of the speaker when a user shakes the speaker, so that a stereo speaker system can be built based on the acceleration information.

[0080] In some embodiments, the speaker 300 may include a display (or a display screen), or may not include a display. The display may be configured to display a display interface of an application, for example, display a currently played song. The display includes a display panel. The display panel may be a liquid crystal display (liquid crystal display, LCD), an organic light-emitting diode (organic light-emitting diode, OLED), an active-matrix organic light-emitting diode (active-matrix organic light-emitting diode, AMOLED), a flexible light-emitting diode (flex light-emitting diode, FLED), a mini-LED, a micro-LED, a micro-OLED, a quantum dot light-emitting diode (quantum dot light-emitting diode, QLED), or the like. In some embodiments, a touch sensor may be disposed in the display to form a touchscreen. This is not limited in this embodiment of this application. The touch sensor is configured to detect a touch operation performed on or near the touch sensor. The touch sensor may transfer the detected touch operation to the processor 303 to determine a type of the touch event. The processor 303 may provide, through the display, a visual output related to the touch operation.

[0081] In some embodiments, FIG. 3 may further include more components such as a battery and a USB interface. Details are not described in this embodiment of this application.

[0082] FIG. 4A is a schematic diagram of a structure of a speaker according to an embodiment of this application. A speaker 400 may be an example of the speaker described in FIG. 2 or FIG. 3. As shown in FIG. 4A, the speaker 400 may include a base 401 and a housing 402.

[0083] In some embodiments, the base 401 has a supporting function. For example, the base 401 may support the housing 402 and a component (for example, a processor, a microphone, or a loudspeaker) surrounded by the housing 402. In some examples, the base 401 may be made of metal, plastic, ceramic, or any other material that has a supporting function; or a combination of these

materials.

[0084] In some embodiments, one or more loudspeakers 406 may be supported on the base 401. For example, the base 401 may support a fixing part 404, and one or more loudspeakers 406 may be disposed on the fixing part 404. In some examples, the base 401 may support the fixing part 404 by using a support column 405 or in another manner. The fixing part 404 may have any shape, for example, may be circular or square. In some embodiments, the one or more loudspeakers 406 may be arranged on the fixing part 404 in a specific manner. For example, the one or more loudspeakers 406 may be evenly distributed on a peripheral of the fixing part 404. For example, a distance between each loudspeaker is the same. In some embodiments, the one or more loudspeakers 406 may be coupled to a processor 403. The processor 403 may output an audio signal by using the one or more loudspeakers 406.

[0085] In some embodiments, the housing 402 may be any cubic shape, for example, a cylinder, a cube, or a square. The housing 402 may surround components such as the processor 403, the fixing part 404, and the one or more loudspeakers 406. The housing 402 may include a single housing component, or more than two housing components. For example, the housing 402 may include an upper housing 402a and a lateral housing 402b. One or more housing components may be made of metal, plastic, ceramic, crystal, or a combination of these materials, or may be any other housing component appropriate for being disposed on the speaker. In some embodiments, the lateral housing 402b may be a housing with a mesh structure. For example, the mesh may have a shape, for example, a circular hole, a square hole, or a hexagonal hole. The housing with the mesh structure can function for, for example, decoration, dust-proof, and protection of components (for example, the loudspeaker and the microphone) inside the housing, and the housing with the mesh structure can reduce blocking of a sound output from the loudspeaker.

[0086] In some embodiments, the upper housing 402a may be a housing with a mesh structure, or without a mesh structure. An input device, for example, a switch, a button, or a key, may be disposed on the upper housing 402a. For example, the switch is used to turn on or off the speaker. The button or the key may function to volume up or down, or the like. In some other embodiments, a display screen 409 (for example, a touch display screen) may be disposed on the upper housing 402a, and may be used to receive an input, provide a visual output, and the like. For example, a name of a currently played song, a name of a singer, and the like may be displayed on the display screen 409. Clearly, the speaker may be disposed with no display screen. This is not limited in this embodiment of this application.

[0087] In some embodiments, the upper housing 402a may be connected to a fixing part 407. One or more microphones 408 may be disposed on the fixing part 407. The fixing part 407 may have any shape, for example, may be circular or square. In some embodiments, the one or more microphones 408 may be arranged on the fixing part 407 in a specific manner. For example, the one or more microphones 408 may be evenly distributed on a peripheral of the fixing part 407. For example, a distance between each microphone is the same. For another example, a central angle a corresponding to every two adjacent microphones (for example, an included angle between straight lines that connect the two microphones and a central point of the fixing part 407) may be fixed, for example, may be 30 degrees or 60 degrees.

[0088] In some embodiments, the one or more microphones 408 may be coupled to the processor 403. The processor 403 may obtain an input signal (for example, a sound signal from a user) by using the one or more microphones 408.

[0089] In the following embodiments of this application, as an example, the application scenario in FIG. 1 is used, and the speaker 1 and/or the speaker 2 in FIG. 1 are/is the speaker 400 shown in FIG. 4A. For ease of description below, one of the speaker 1 and the speaker 2 is referred to as a main speaker, and the other is referred to as an auxiliary speaker. In some embodiments, the main speaker and the auxiliary speaker may be used as a set. For example, the main speaker is used to play a left sound channel, and the auxiliary speaker is used to play a right sound channel; or the main speaker is used to play a right sound channel, and the auxiliary speaker is used to play a left sound channel. That is, the main speaker and the auxiliary speaker can cooperate to implement a stereo effect of audio. In some embodiments, whether a speaker is a main speaker or an auxiliary speaker may be set before the speaker is delivered from a factory, or may be customized by a user (for example, the speaker receives an input operation through a touch display screen, and the input operation is used to select the speaker to be a main speaker or an auxiliary speaker). In this embodiment of this application, speaker selection may alternatively be performed according to a specific rule.

[0090] In some embodiments, structures of the main speaker and the auxiliary speaker may be the same. For example, both the main speaker and the auxiliary speaker have the structure shown in FIG. 4A. In some other embodiments, alternatively, structures of the main speaker and the auxiliary speaker may not be completely the same. For example, a display screen may be disposed on the main speaker, but no display screen is disposed on the auxiliary speaker. In other embodiments, functions of some components in the main speaker and the auxiliary speaker may not be completely the same. For example, a processor of the main speaker may be configured to calculate a delay (for example, a time difference between a first time length and a second time length, where the first time length may be a time length required for a sound reaching a user from the main speaker, and the second time length may be a time length required for a sound reaching a user from the auxiliary

speaker), a loudness gain, and the like, while a processor of the auxiliary speaker does not have this function.

[0091] In some embodiments, a memory of the main speaker and/or the auxiliary speaker may store an audio file (for example, a song, a comic dialog, or a storytelling show), and the main speaker and the auxiliary speaker may play the stored audio file. For example, the main speaker may receive an input (for example, receive an input operation through a touch display screen, or receive a speech input through a microphone). The input may be used to start the main speaker and/or the auxiliary speaker, or may be used to control the main speaker and the auxiliary speaker to play, switch a song, and the like. In some embodiments, one or more microphones of the main speaker acquire a sound signal (for example, a sound signal from a user), and the processor identifies a "wake word+play a song" included in the sound signal. When determining that the song does not exist in the memory, the processor may download the song from a network side, or output prompt information (for example, speech information) to prompt the user that the song does not exist.

[0092] In some other embodiments, the main speaker and/or the auxiliary speaker may be connected to another electronic device (for example, a mobile phone or a television), and may be connected in a wired or wireless manner. For example, the main speaker is connected to a mobile phone (for example, through a Bluetooth connection). The mobile phone may send an audio signal to the main speaker, so that the main speaker and the auxiliary speaker play the audio signal (for example, after receiving the audio signal, the main speaker may send the audio signal to the auxiliary speaker). For example, a music play application (for example, KUGOU Music) is running on the mobile phone, and a song "All the Way to North" is being played. The mobile phone may send an audio signal of the song to the main speaker, so that the main speaker and the auxiliary speaker play the audio signal. In some other embodiments, after the main speaker is connected to the mobile phone, a user may control, by using the main speaker, the mobile phone to perform a corresponding operation. The foregoing example is still used. The user gives a sound signal "Xiao Bai, play the song Listen to Your Mother" in a room. The main speaker acquires the sound signal, and may pause playing All the Way to North and output prompt information "Searching for Listen to Your Mother". For example, the main speaker may search a local memory for whether the song Listen to Your Mother exists. If the song does not exist, the main speaker may download the song from the network side, or the main speaker may send an instruction to the mobile phone, where the instruction is used to instruct the mobile phone to play the song Listen to Your Mother. After receiving the instruction, the mobile phone downloads or live streams the song, and sends an audio signal of the song to the main speaker, so that the main speaker and the auxiliary speaker play the audio signal of the song (that is, Listen to Your Mother).

[0093] In some embodiments, both the main speaker and the auxiliary speaker may enable a "wake word" automatic identification function. The main speaker is used as an example. After the main speaker enables a "wake word" automatic identification function, all or some components (for example, the one or more microphones and processors) of the main speaker are in an enabled state. The sound signal from the user in the room is received by the one or more microphones of the main speaker. The one or more microphones send the received sound signal to the processor. When determining that the sound signal includes the "wake word", the processor enables another component (for example, one or more loudspeakers). In some embodiments, the "wake word" may be set by default when the speaker is delivered from a factory, or may be customized by the user. For example, the "wake word" may be "Xiao Bai", "Xiao Yin", "Xiao Yi", or the like.

[0094] In some other embodiments, both the main speaker and the auxiliary speaker may enable a "wake word+play a song" automatic identification function. The main speaker is used as an example. After the main speaker enables a "wake word+play a song" automatic identification function, all or some components (for example, the one or more microphones and processors) of the main speaker are in an enabled state. The sound signal from the user in the room is received by the one or more microphones of the main speaker. The one or more microphones send the received sound signal to the processor. When determining that the sound signal includes the "wake word+play a song", the processor enables another component (for example, one or more loudspeakers). For example, the user says "Xiao Bai, play All the Way to North" in the room. The microphone of the main speaker acquires the sound signal, and then sends the sound signal to the processor. The processor identifies that the wake word, Xiao Bai, included in the sound signal, and playing a song that is also included. The processor enables another component (for example, the one or more loudspeakers).

[0095] In some embodiments, the main speaker may receive an input operation through an input device (for example, a touchscreen on the main speaker), or receive an input operation through another device connected to the main speaker, for example, a mobile phone. To enable the "wake word" automatic identification function or the "wake word+play a song" automatic identification function in response to the input operation, the main speaker may send an instruction to the auxiliary speaker, where the instruction is used to instruct the auxiliary speaker to enable the "wake word" automatic identification function or the "wake word+play a song" automatic identification function.

[0096] In some embodiments, the user may be in any position in the room, and a distance between the main speaker and the user is different from a distance between the auxiliary speaker and the user. The main speaker and auxiliary speaker enable the "wake word" automatic

identification function or the "wake word+play a song" automatic identification function. The microphones of the main speaker and the auxiliary speaker acquire a sound signal. When the main speaker and the auxiliary speaker determine that the sound signal includes "wake word" or "wake word+play a song", a position of the user may be determined, and then sound parameters of the main speaker and the auxiliary speaker are controlled based on the position of the user. For example, the sound parameters may include delays, loudness gains, and the like of the main speaker and the auxiliary speaker. Therefore, in this embodiment, only when the main speaker and the auxiliary speaker identify the "wake word" or "wake word+play a song" included in the acquired sound signal, the sound parameters of the main speaker and the auxiliary speaker are adjusted based on the position of the user.

[0097] The structure shown in FIG. 4A is used as an example, a process in which the main speaker and the auxiliary speaker determine the position of the user may include: The main speaker acquires a sound signal 1. The auxiliary speaker acquires a sound signal 2. The main speaker determines that the sound signal 1 includes a "wake word", and the auxiliary speaker determines that the sound signal 2 includes a "wake word". Clearly, to improve accuracy, the auxiliary speaker may further send the sound signal 2 or the "wake word" included in the sound signal 2 to the main speaker, and the main speaker determines that the "wake word" in the sound signal 1 and the sound signal 2 is a same wake word. The main speaker may determine, based on the sound signal 1, a first direction/orientation of the user relative to the main speaker. For example, the first direction/orientation may be represented as a first angle between the user and an x-axis in a coordinate system constructed by the main speaker. The auxiliary speaker may determine, based on the sound signal 2, a second direction/orientation of the user relative to the auxiliary speaker. For example, the second direction/orientation may be represented as a second angle between the user and an x-axis in a coordinate system constructed by the auxiliary speaker. The auxiliary speaker may send the second angle to the main speaker, and the main speaker determines the position of the user based on the first angle, the second angle, and a distance D between the main speaker and the auxiliary speaker. Specifically, a process in which the main speaker and the auxiliary speaker each construct a coordinate system and the main speaker and the auxiliary speaker determine the position of the user is described in detail below.

[0098] The structure shown in FIG. 4A is still used as an example. The main speaker may determine, based on the sound signal 1, the first direction/orientation of the user relative to the main speaker in a plurality of manners. For example, a microphone array positioning technology (for example, an orientation of a sound source is estimated based on a difference of time at which sound signals are received by at least two microphones in a microphone array on the main speaker), a steered-beamformer (steered-beamformer) positioning method, a high-resolution spectral analysis (high-resolution spectral analysis)-based positioning method, or a time-delay estimation (time-delay estimation, TDE)-based sound source positioning technology is used. This is not limited in this embodiment of this application. The microphone array positioning technology is used as an example. A process in which the main speaker determines, based on the sound signal 1, the first direction/orientation of the user relative to the main speaker may include: A microphone array 408 of the main speaker acquires a sound signal, where it is assumed that the sound signal acquired by a microphone 408-1 and a microphone 408-2 is relatively strong, the main speaker may calculate the first orientation of a sound source, that is, the user relative to the main speaker based on a first moment t1 at which the microphone 408-1 acquires the sound signal, a second moment t2 at which the microphone 408-2 acquires the sound signal, and a distance L1 between the microphone 408-1 and the microphone 408-2 (the distance may be stored in the main speaker after delivery). As shown in FIG. 4B, the main speaker may determine an included angle A between the user and the microphone 408-1 based on (t1-t2)*c, L1, and a trigonometric function relationship. The included angle A may be used as the first orientation of the user relative to the main speaker. Alternatively, because the included angle A is the included angle between the user and the microphone 408-1, the main speaker may convert coordinates of the included angle A into the coordinate system constructed by the main speaker, to obtain an included angle B. The included angle B may also be used as the first orientation of the user relative to the main speaker. Structures of the auxiliary speaker and the main speaker may be the same. Therefore, a process in which the auxiliary speaker determines the second orientation of the user relative to the auxiliary speaker may be similar to the foregoing process.

[0099] In some other embodiments, in a process in which the user continuously gives sound signals, the main speaker and the auxiliary speaker may continuously acquire the sound signals in real time (the sound signals may not include the "wake word" or the "wake word+play a song"), then determine the position of the user, and adjust the sound parameters of the main speaker and the auxiliary speaker based on the position of the user until a sound signal including the "wake word" or the "wake word+play a song" is detected. Then, the main speaker and the auxiliary speaker are controlled to play an audio signal based on the adjusted sound parameters (for example, delays and loudness gains of the main speaker and the auxiliary speaker).

[0100] In some embodiments, the main speaker may detect the distance D from the auxiliary speaker, to determine whether to build a stereo speaker system. The distance may be a linear distance between the main speaker and the auxiliary speaker. After detecting the distance D, the main speaker may send the distance D

to the auxiliary speaker, and the auxiliary speaker does not need to detect the distance D; or the auxiliary speaker may alternatively detect the distance D from the main speaker for future use. Clearly, the auxiliary speaker may detect the distance D from the main speaker, and then send the distance D to the main speaker. In other words, the main speaker does not need to detect the distance D, or the like.

[0101] The main speaker is used as an example. As an example, the main speaker may detect the distance from the auxiliary speaker by using a distance sensor. The distance sensor may be a laser distance sensor, an infrared distance sensor, or the like. For example, the distance sensor on the main speaker emits infrared light with a specific frequency, and the infrared light is reflected from the auxiliary speaker, and the main speaker receives the light emitted from the auxiliary speaker. The main speaker may calculate the distance between the main speaker and the auxiliary speaker based on first time at which the infrared light is transmitted and second time at which the reflected light is received. As another example, the main speaker may alternatively reach the goal of measuring the distance between the main speaker and the auxiliary speaker by communicating with the auxiliary speaker. For example, the main speaker sends a detection signal to the auxiliary speaker, the auxiliary speaker sends a feedback signal to the main speaker after receiving the detection signal, and the main speaker receives the feedback signal. The main speaker may determine the distance between the main speaker and the auxiliary speaker based on second time at which the feedback signal is received and first time at which the detection signal is sent. As still another example, the main speaker may alternatively receive an input operation by using the input device (for example, the touchscreen on the main speaker), and the input operation is used to input the distance between the main speaker and the auxiliary speaker.

[0102] Similarly, the distance between the main speaker and the auxiliary speaker may alternatively be determined by the main speaker and the auxiliary speaker by using the microphone array positioning technology (for example, an orientation of a sound source is estimated based on a difference of time at which sound signals are received by at least two microphones in a microphone array on the main speaker), the steered-beamformer (steered-beamformer) positioning method, the high-resolution spectral analysis (high-resolution spectral analysis)-based positioning method, the time-delay estimation (time-delay estimation, TDE)-based sound source positioning technology, or the like. For example, the main speaker may emit a sound, and the auxiliary speaker determines the distance between the auxiliary speaker and the main speaker based on the microphone array. When the main speaker and the auxiliary speaker are not determined, at least one speaker of the two speakers may emit a sound to measure the distance.

[0103] The following embodiments describe possible implementations of building a stereo speaker system.

[0104] First, for ease of understanding, some key terms used in embodiments of this application are explained.

(1) Motion status: information about how a speaker moves. For example, a first motion state and a second motion state mentioned in the following embodiments are parameters for describing how a speaker moves. The motion status may be data acquired by an acceleration sensor, or may be information obtained by processing data acquired by an acceleration sensor.

(2) First indication information: used to describe a correspondence between an acceleration and duration of the acceleration. For example, defining a same acceleration means that if a change between accelerations is less than a preset change, the accelerations are a same acceleration, and it should be noted that different directions of accelerations indicate different accelerations. For example, an acceleration range [1, 2] is considered as a same acceleration, with a change between accelerations within a range of 0.5 based on a reference of 1.5. Then, duration of each acceleration is collected, and each acceleration and the duration corresponding to the acceleration may be expressed by using corresponding first indication information. For example, in a possible implementation, the first indication information is (A, T), where A represents an acceleration, and T represents duration corresponding to the acceleration A. In this example, the duration T of the acceleration A is described through explicit expression. In another embodiment, the duration T of the acceleration A may alternatively be described through implicit expression. For example, if the acceleration sensor acquires data at regular intervals, it may be specified that the duration of each acceleration is expressed by using the interval (for example, 5 ms). Therefore, the first indication information may alternatively be expressed as (A). That is, the duration is 5 ms by default.

It should be noted that the acceleration sensor may simultaneously acquire acceleration information in three directions (that is, a direction X, a direction Y, and a direction Z). (A, T) may be represented as (X1, Xt1, Y1, Yt1, Z1, Zt1) or (X1, Y1, Z1, Xt1, Yt1, Zt1). X1, Y1, and Z1 each represent an acquired acceleration from each direction, Xt1 represents duration of the acceleration X1, Yt1 represents duration of the acceleration Y1, and Zt1 represents duration of the acceleration Z1.

(3) First acceleration sequence (time sequence information): used to describe a manner of organizing the first indication information, that is, a time-domain signal of the motion status. When there are a plurality of accelerations, the first indication information may be sorted in a time sequence, to obtain a first accel-

eration sequence of a speaker. For example, the first acceleration sequence is S={(X1, Xt1, Y1, Yt1, Z1, Zt1), (X2, Xt2, Y2, Yt2, Z2, Zt2), ..., (Xn, Xtn, Yn, Ytn, Zn, Ztn)}, where n is a positive integer. Clearly, the first indication information may alternatively be sorted by different acceleration directions, for example, S={(X1, Xt1, X2, Xt2, ..., Xn, Xtn), (Y1, Yt1, Y2, Yt2, ..., Yn, Ytn), ..., (Z1, Zt1, Z2, Zt2, ..., Zn, Ztn)}. Therefore, a change of the acceleration and the duration corresponding to the acceleration over time can be expressed based on the first acceleration sequence, to obtain the motion status of the speaker expressed in the time sequence.

(4) Second indication information: frequency domain information of each point in the motion status of the speaker.

In embodiments of this application, the motion status may not only be expressed based on the first acceleration sequence as described above, but may also be expressed based on frequency domain information of an acceleration acquired by the acceleration sensor. In embodiments of this application, frequency domain information of each acceleration may be referred to as the second indication information. An expression manner of the frequency domain information is that, for example, if collected frequency domain information of an acceleration A is expressed as P, second indication information of the acceleration A is expressed as (A, P). For example, it is assumed that duration of sampling every single time is 9 s, and an obtained sampling result of an acceleration in a direction X is (1,2,3,4,1,2,3,5,1,2,3,4,1,2,3,5, ...). The sampling result is converted into frequency domain, to obtain frequency domain information corresponding to the direction X. For a direction Y and a direction Z, refer to a processing manner for the direction X. Details are not described again in this application.

(5) Second acceleration sequence (frequency domain information): used to describe a manner of organizing the second indication information, and used to express the motion status of the speaker in the frequency domain (that is, a frequency-domain signal of the motion status). When there are a plurality of pieces of second indication information, the second acceleration sequence that is S={Xp, Yp, Zp} may be obtained. Xp, Yp, and Zp are second indication information in three directions X, Y, and Z respectively. In the foregoing example, the acceleration in the direction X includes (1,2,3,4,1,2,3,5,1,2,3,4,1,2,3,5, ...), and the second acceleration sequence corresponding to the direction X is Xp={(1,4), (2,4), (3,4), (4,8), (5,8)}. Second acceleration sequences corresponding to the direction Y and the direction Z may be obtained by deduction.

(6) Preset action and a characteristic of the preset action: To improve accuracy of triggering building of a stereo speaker system and avoid that any movement of a speaker triggers an operation of building the stereo speaker, the preset action may be predefined in embodiments of this application. Only when a motion status of the speaker matches the characteristic of the preset action, it is considered that the stereo speaker system needs to be built, so that false building can be avoided. For example, as shown in FIG. 5, the preset action may be holding the speaker and drawing a specific pattern (for example, an 8-shaped pattern, a wavy line, or a pentagonal star). For example, the preset action may be shaking the speaker from side to side, shaking the speaker up or down, or generating a collision of the speaker. A manner of generating the collision may be a collision with another speaker, or may be a collision with another object, as long as the speaker has a collision.

[0105] Clearly, it should be noted that, during implementation, the preset action may be configured based on an actual requirement, or even the user may customize the preset action. Either is applicable to embodiments of this application.

[0106] For ease of determining whether an operation performed by the user on the speaker is the preset action, in embodiments of this application, the characteristic of the preset action is used to describe the preset action.

[0107] A manner of obtaining the characteristic of the preset action may be implemented as follows: After the preset action is predefined, the speaker is operated based on the preset action (for example, the speaker is held and an 8-shaped pattern is drawn), and the motion status of the speaker is acquired (for example, the foregoing first acceleration sequence or second acceleration sequence is acquired), to obtain the characteristic of the preset action and store the characteristic in a memory of the speaker.

[0108] To facilitate different user groups to conveniently control the speaker to build a stereo speaker system, in embodiments of this application, characteristics, of a preset action, applicable to different user groups may be constructed for the same preset action. For example, if the preset action is drawing the number 8, data on the elderly drawing the number 8 may be acquired, to obtain a characteristic applicable to the elderly and the number 8 drawn by the elderly; and data on children drawing the number 8 may be acquired, to obtain a characteristic applicable to children and the number 8 drawn by children.

[0109] In addition, in embodiments of this application, the user may alternatively customize the preset action. For example, as shown in FIG. 6, a user triggers a service logic of customizing the preset action by using a key on the speaker. Then, the speaker may prompt the user for customizing the preset action for the speaker. For example, the customized preset action is drawing a pentagonal star. The user holds the speaker and draws a pentagonal star, while the speaker acquires the motion status of the speaker as a sample of the customized preset action,

and counts the performed customized preset action. It is assumed that the user holds the speaker and draws a pentagonal star three times, and the speaker acquires three samples. In this case, a preset counting requirement is satisfied, and sample acquisition ends. Then, the three samples are analyzed (for example, an average value is obtained), to obtain a characteristic of the customized preset action, and the characteristic is stored in the memory. Then, the user holds the speaker and draws a pentagonal star, and matching is performed on the characteristic of the preset action stored in the memory. If the matching is successful, building of a stereo speaker system is triggered.

[0110] (7) First sequence template and second sequence template: Both templates are characteristics of a preset action, and are used for performing matching on the acquired motion status of the speaker. In (3) and (5) described above, the time sequence information obtained by sorting and analyzing the acceleration information is the first acceleration sequence, and the frequency domain information is the second acceleration sequence. Therefore, in embodiments of this application, the first sequence template corresponding to the first acceleration sequence is provided, and the second sequence template corresponding to the second acceleration sequence is provided. When the motion status of the speaker is expressed by using the first acceleration sequence, the first sequence template and the first acceleration sequence are used for matching. When the motion status of the speaker is expressed by using the second acceleration sequence, the second sequence template and the second acceleration sequence are used for matching. Then, it is determined, based on a matching result, whether to trigger an operation of building the stereo speaker system.

[0111] FIG. 7 is a schematic flowchart of a method for building a stereo speaker system according to an embodiment of this application. Two speakers are used as an example, and the method includes the following steps.

[0112] To help a user build a stereo speaker system, after the user moves a speaker, the speaker may prompt the user for a preset action, for example, with a voice prompt: "Please hold the speaker and draw the number 8 to build a stereo speaker system", that needs to be performed for building the stereo speaker system. Alternatively, if the speaker has a display screen, the display screen may output a text "Please hold the speaker and draw the number 8 to build a stereo speaker system" for prompting. A specific manner of prompting is not limited in this application.

[0113] Step 701: The user moves a first speaker, and the first speaker obtains a first motion state of the first speaker; and similarly, the user moves a second speaker, and the second speaker also obtains a first motion state of the second speaker.

[0114] It should be noted that, timing of the user moving the first speaker and the second speaker is not limited. To be specific, the user may move the first speaker before moving the second speaker, may move the second speaker before moving the first speaker, or may move the second speaker and the first speaker at the same time. In addition, the first speaker and the second speaker may be moved by a same user or different users.

[0115] An extracted motion state is, for example, the first acceleration sequence S={(X1, Xt1, Y1, Yt1, Z1, Zt1), (X2, Xt2, Y2, Yt2, Z2, Zt2), ..., (Xn, Xtn, Yn, Ytn, Zn, Ztn)} described above.

[0116] The extracted motion state is, for example, the second acceleration sequence obtained through frequency domain analysis on the acceleration information mentioned above, for example, the second acceleration sequence Xp={(1,4), (2,4), (3,4), (4,8), (5,8)} in the direction X.

[0117] Step 702: The first speaker determines whether the first motion state of the first speaker matches a characteristic of a preset action; and similarly, the second speaker determines whether the first motion state of the second speaker matches a characteristic of a preset action.

[0118] It should be noted that the preset action for the first speaker and the preset action for the second speaker may be the same or may be different. For example, the preset actions of the first speaker and the second speaker may both be drawing the number 8; or the preset action for the first speaker may be drawing the number 8, and the preset action for the second speaker may be drawing a pentagonal star.

[0119] For ease of understanding, the first speaker is used as an example herein to describe whether the first motion state matches the characteristic of the preset action. The defined characteristic of the preset action is represented by a first sequence template. The first sequence template includes first sequence sub-templates corresponding to three directions X, Y, and Z respectively. FIG. 8 is a schematic diagram of a first sequence sub-template corresponding to a direction X. The first sequence sub-template corresponding to the direction X presents a run chart of accelerations in the direction X changing over time (by default, duration of each acceleration is a sampling interval). When an acceleration sensor outputs an acceleration value, a first acceleration sequence is obtained. The first acceleration sequence includes first acceleration subsequences corresponding to the three directions X, Y, and Z respectively. FIG. 8 is a schematic diagram of a first acceleration subsequence corresponding to the direction X. FIG. 8 shows a trend of accelerations in the direction X changing over time (by default, duration of each acceleration is a sampling interval).

[0120] When an acceleration value is acquired, the first sequence sub-template corresponding to the direction X and the first acceleration subsequence acquired in the direction X are used for template matching. As shown in FIG. 8, it can be learned that three first sequence sub-templates (for example, A1, A2, and A3 in FIG. 8) are matched. Similarly, for the direction Y, a first sequence

sub-template corresponding to the direction Y and a first acceleration subsequence corresponding to the direction Y may be used for template matching; and for the direction Z, a first sequence sub-template corresponding to the direction Z and a first acceleration subsequence corresponding to the direction Z may be used for matching. Details are not described herein again.

[0121] To improve accuracy of triggering building of the stereo speaker system, for the direction X, when a quantity of times the first sequence sub-template corresponding to the direction X is matched is greater than a specified quantity of times, it is determined that the first acceleration subsequence corresponding to the direction X matches a characteristic of a preset action corresponding to the direction X.

[0122] Similarly, for the direction Y, when a quantity of times the first sequence sub-template corresponding to the direction Y is matched is greater than a specified quantity of times, it is determined that the first acceleration subsequence corresponding to the direction Y matches a characteristic of a preset action corresponding to the direction Y For the direction Z, when a quantity of times the first sequence sub-template corresponding to the direction Z is matched is greater than a specified quantity of times, it is determined that the first acceleration subsequence corresponding to the direction Z matches a characteristic of a preset action corresponding to the direction Z.

[0123] When the characteristics of the preset actions in the three directions X, Y, and Z are all matched, it is determined that the first motion state of the first speaker matches the characteristic of the preset action; otherwise, the first motion state of the first speaker does not match the characteristic of the preset action.

[0124] In some other embodiments, the acceleration sensor may not acquire an acceleration within a period of time. In this case, during implementation, to accurately determine whether the characteristic of the preset action is matched, a time length threshold t' may be set. The direction X is used as an example. For example, for the direction X, when the quantity of times the first sequence sub-template corresponding to the direction X is matched within the time length threshold t' is greater than the specified quantity of times, it is determined that the first acceleration subsequence corresponding to the direction X matches the characteristic of the preset action corresponding to the direction X; otherwise, the first acceleration subsequence corresponding to the direction X does not match the characteristic of the preset action corresponding to the direction X. As shown in FIG. 9, first, the first sequence sub-template corresponding to the direction X is matched within a time period A1, and a quantity of times of matching is counted as 1. Then, if the first sequence sub-template corresponding to the direction X is not acquired or is not matched within time t (t>t'), the first sequence sub-template that corresponds to the direction X and that is previously matched is discarded, and the quantity of times of matching is re-denoted as 0. Then, the quantity of times of matching is recounted until the quantity of times of matching is greater than the specified quantity of times, and it is determined that the characteristic of the preset action corresponding to the direction X is matched.

[0125] In addition, a threshold of a time difference between two consecutive times the first sequence sub-template corresponding to the direction X is matched may be further limited. In this case, it is specified that if the time difference between two consecutive times the first sequence sub-template corresponding to the direction X is matched is less than the time difference threshold, the quantity of times the first sequence sub-template corresponding to the direction X is matched is added up; otherwise, the quantity of times is recounted. Same processing is further performed for the direction Y and the direction Z, and details are not described herein again.

[0126] It should be noted that, in this embodiment of this application, matching may be separately performed in the three directions X, Y, and Z, or direct matching on the first sequence template may be performed. For example, the first sequence template is represented as P={(Xp1, Xpt1, Yp1, Ypt1, Zp1, Zpt1), (Xp2, Xpt2, Yp2, Ypt2, Zp2, Zpt2), ..., (Xpm, Xptm, Ypm, Yptm, Zpm, Zptm)}, where m is a positive integer. It is assumed that the first acceleration sequence is S={(X1, Xt1, Y1, Yt1, Z1, Zt1), (X2, Xt2, Y2, Yt2, Z2, Zt2), ..., (Xn, Xtn, Yn, Ytn, Zn, Ztn)}. In this case, a matching manner is that matching between (X1, Xt1, Y1, Yt1, Z1, Zt1) and (Xp1, Xpt1, Yp1, Ypt1, Zp1, Zpt1) is performed; matching between (Xp2, Xpt2, Yp2, Ypt2, Zp2, Zpt2) and (X2, Xt2, Y2, Yt2, Z2, Zt2) is performed; and so on. If the first sequence template P is matched, where P={(Xp1, Xpt1, Yp1, Ypt1, Zp1, Zpt1), (Xp2, Xpt2, Yp2, Ypt2, Zp2, Zpt2), ..., (Xpm, Xptm, Ypm, Yptm, Zpm, Zptm)}, it is denoted that the first sequence template is matched once. A specific matching manner is not limited in this application. For details, refer to a matching manner for a time-domain signal, or a difference between each pair of corresponding elements may be determined for matching. If the difference is within a preset difference range, it is determined that matching is successful. For example, when matching between (X1, Xt1, Y1, Yt1, Z1, Zt1) and (Xp1, Xpt1, Yp1, Ypt1, Zp1, Zpt1) is performed, as shown in Table 1, differences between X1 and Xp1, Xt1 and Xpt1, Y1 and Yp1, Yt1 and Ypt1, Z1 and Zp1, and Zt1 and Zpt1 are calculated separately, and a largest difference (assumed to be $\alpha 6$) is used. If an absolute value of the largest difference $\alpha 6$ is less than or equal to a preset difference, it is determined that (X1, Xt1, Y1, Yt1, Z1, Zt1) matches (Xp1, Xpt1, Yp1, Ypt1, Zp1, Zpt1); otherwise, (X1, Xt1, Y1, Yt1, Z1, Zt1) does not match (Xp1, Xpt1, Yp1, Ypt1, Zp1, Zpt1).

Table 1

X1	Xt1	Y1	Yt1	Z1	Zt1
Xp1	Xpt1	Yp1	Ypt1	Zp1	Zpt1

(continued)

α1	α2	α3	α4	α5	α6

[0127] An implementation of determining, when the first acceleration sequence is used to express the first motion state of the speaker, whether the first motion matches the characteristic of the preset action is described above. A matching manner used when a frequency-domain signal, that is, the second acceleration sequence, is used to express the first motion state of the speaker is described below. Similar to the first acceleration sequence, in this embodiment of this application, the second acceleration sequence has a corresponding second sequence template. In an implementation, if a quantity of times the second acceleration sequence matches the second sequence template is greater than or equal to a second specified quantity of times, it is determined that the first motion state matches the characteristic of the preset action; or if a quantity of times the second acceleration sequence matches the second sequence template is less than a second specified quantity of times, it is determined that the first motion state does not match the characteristic of the preset action.

[0128] For example, the second sequence template may include second sequence sub-templates corresponding to the directions X, Y, and Z. The direction X is used as an example. Matching between a second sequence sub-template corresponding to the direction X and frequency domain information corresponding to the direction X is performed. If a quantity of times the second sequence sub-template is matched is greater than the second specified quantity of times, it is determined that matching is successful in the direction X; otherwise, the frequency domain information corresponding to the direction X does not match the second sequence sub-template corresponding to the direction X. When matching is successful in the direction X, similarly, if matching is also successful in the direction Y and the direction Z, it is determined that the first motion state matches the characteristic of the preset action; otherwise, if matching is unsuccessful in any of the directions, it is determined that the first motion state does not match the characteristic of the preset action.

[0129] Similarly, when the acceleration sensor has acquired no data within a period of time (the period of time is greater than the time length threshold t'), or the second sequence template is not matched, the matching fails, and the motion state of the speaker does not match the characteristic of the preset action.

[0130] Similarly, it is specified that if a time difference between two consecutive times the second sequence template is matched is less than a time difference threshold, the quantity of times the second sequence template is matched is added up; otherwise, if a time difference between two consecutive times the second sequence template is matched is not less than a time difference threshold, the matching fails, and the quantity of times is recounted.

[0131] In some embodiments, when the first motion state of the first speaker matches the characteristic of the preset action, an operation of building the stereo speaker system may be triggered. To further avoid false triggering, in this embodiment of this application, whether a moving distance of the speaker is long enough may be further used as another trigger condition, and step 703 shown in FIG. 7 may be implemented.

[0132] Step 703: Determine a moving distance of the first speaker in the first motion state of the first speaker, and determine a moving distance of the second speaker in the first motion state of the second speaker.

[0133] It should be noted that timing for performing step 702 and step 703 is not limited.

[0134] Step 704: If the first motion state of the first speaker matches the characteristic of the preset action, and the moving distance of the first speaker is greater than a specified distance, trigger the operation of building the stereo speaker system, where to be specific, the first speaker searches for a surrounding speaker; and similarly, if the first motion state of the second speaker matches the characteristic of the preset action, and the moving distance of the second speaker is greater than a specified distance, trigger the operation of building the stereo speaker system, where to be specific, the second speaker searches for a surrounding speaker.

[0135] A specific manner of searching may be that both the first speaker and the second speaker scan device information of a peer end, for example, information about a system version and signal strength. During implementation, for example, the first speaker and the second speaker separately broadcast respective device information, and then the first speaker may find the device information of the second speaker through a search, and the second speaker may find the device information of the first speaker through a search.

[0136] When a distance between the two speakers is long or there is an obstacle between the two speakers, it is actually not suitable to build the stereo speaker system. Therefore, for ease of accurately building the stereo speaker system, a condition for building the stereo speaker system may be added in this embodiment of this application.

[0137] In step 705, after finding the second speaker through a search, the first speaker may determine a position relationship between the first speaker and the second speaker; and similarly, after finding the first speaker through a search, the second speaker may determine the position relationship between the second speaker and the first speaker.

[0138] In step 706, if the position relationship between the first speaker and the second speaker is a specified position relationship, the stereo speaker system is built by using the first speaker and the second speaker. In addition, the first speaker and the second speaker may further prompt the user for speakers that can build a ster-

eo speaker system. For example, the first speaker and the second speaker may flash a light, or control to generate a specific light effect to prompt the user that the first speaker and the second speaker may be used to build a stereo speaker system. In addition to the light effect prompt, a sound effect prompt may be used, which is also applicable to this embodiment of this application. For example, audio "I can be used to build a stereo speaker system" is output. In addition, a display screen prompt may be used. Clearly, in another embodiment, a combined prompt may be used. For example, a light effect+sound effect prompt or a sound effect+display screen prompt is also applicable to this embodiment of this application.

[0139] During implementation, the position relationship may be expressed by a distance. The first speaker may broadcast the device information of the first speaker, so that a surrounding speaker can perceive the first speaker; and similarly, the second speaker broadcasts the device information of the second speaker, so that a surrounding speaker can perceive the second speaker. After obtaining the device information of the second speaker, the first speaker determines a distance from the second speaker. Similarly, after finding the first speaker through a search, the second speaker may determine the distance between the second speaker and the first speaker. When the distance between the two speakers is less than a distance threshold D, it is determined that the two speakers satisfy a condition for teaming up. Then, the two speakers start to build a stereo speaker system.

[0140] Building the stereo speaker system includes the following parts: First, a main speaker and an auxiliary speaker are selected (step 707 may be implemented). Second, a role of each speaker is determined, that is, a sound channel configuration of each speaker is determined (step 708 may be implemented).

[0141] Step 707: Select the main speaker from the two speakers according to a main speaker selection rule.

[0142] The main speaker selection rule includes at least one of the following rules. In a possible implementation, when there are a plurality of rules included in the main speaker selection rule, the following rules are sorted by priority (clearly, the priorities may also be sorted based on an actual requirement, which is also applicable to this embodiment of this application).

[0143] First priority: selecting a speaker connected to a network as the main speaker;

second priority: selecting a speaker configured with a network but not connected to the network as the main speaker;

third priority: selecting a speaker connected to an intelligent terminal device as the main speaker;

fourth priority: selecting a speaker that first matches the characteristic of the preset action as the main speaker; and

fifth priority: selecting a speaker with a largest MAC address as the main speaker.

[0144] To be specific, first, a speaker connected to a network is selected as the main speaker, so that a resource needed to be played by the speaker can be obtained through the network. If no speaker is connected to a network, a speaker configured with a network is selected as the main speaker. If no speaker is configured with a network, a speaker that can communicate with a terminal device, for example, a mobile phone, is selected as the main speaker. The main speaker selected in the foregoing manners can obtain, by communicating with a network, a resource to be played.

[0145] If none of the foregoing conditions is satisfied, a speaker that first matches the characteristic of the preset action may be selected as the main speaker. For example, when a stereo speaker system is built with a shake, a speaker that is first shaken is used as the main speaker.

[0146] In addition, the speaker that first matches the characteristic of the preset action is selected as the main speaker. Clearly, it should be noted that a selection rule is that a speaker that can communicate with a network is preferentially selected as the main speaker. If there is no such a speaker, it should be ensured that all speakers select a same speaker as the main speaker. This is not limited in this application.

[0147] After the main speaker is selected, a user may be prompted, by using a sound effect or a light effect, for a speaker that is the main speaker.

[0148] It should be noted that, if no surrounding speaker is found, the search may be repeated for a plurality of times, and if no surrounding speaker is found during these plurality of times, the operation of building the stereo speaker system may be ended, and the user may be further prompted for a failure of the building and a cause of the failure. For example, the cause is that no surrounding speaker is found.

[0149] Step 708: Each speaker may determine a sound channel for the speaker based on position information of the speaker, and prompt the user for the sound channel.

[0150] During implementation, each speaker may obtain a relative position relationship of the speaker in the stereo speaker system. FIG. 10 to FIG. 13 are schematic diagrams of the relative position relationship, and include several cases for description.

[0151] As shown in FIG. 10, when the first speaker is on a left side and the second speaker is on a right side, the first speaker is used as a left sound channel, and the second speaker is used as a right sound channel. Then, the first speaker prompts, through audio and/or a light effect, that the first speaker is the left sound channel and the second speaker is the right sound channel.

[0152] As shown in FIG. 11, if three speakers are included, a first speaker determines that the first speaker is a left sound channel, a second speaker determines that the second speaker is a right sound channel, and a third speaker determines that the third speaker is a subwoofer.

[0153] As shown in FIG. 12, if six speakers are includ-

ed, a first speaker determines that the first speaker is a left sound channel, a second speaker is a right sound channel, a third speaker is a subwoofer, a fourth speaker is a center-channel speaker, a fifth speaker is a rear left sound channel, and a sixth speaker is a rear right sound channel.

[0154] As shown in FIG. 13, if 12 speakers are included, a first speaker determines that the first speaker is a left sound channel, a second speaker is a right sound channel, a third speaker is a subwoofer, a fourth speaker is a center-channel speaker, a fifth speaker is a rear left sound channel, a sixth speaker is a rear right sound channel, a seventh speaker is a center left sound channel, an eighth speaker is a center right sound channel, and a ninth speaker to a twelfth speaker at positions shown in FIG. 13 are all surround sound speakers.

[0155] Clearly, in some other embodiments, each speaker may not determine a sound channel for the speaker, but prompt the user for configuring the sound channel for each speaker. In addition, even after the sound channel for each speaker is configured, the user may be prompted for confirming a sound channel configuration or modifying the sound channel configured for each speaker.

[0156] Step 709: The user may configure a sound channel for each speaker by moving the speaker. For example, the first speaker obtains a second motion state of the first speaker, and then configures a sound channel corresponding to the second motion state as a sound channel of the first speaker. Similarly, as shown in FIG. 7, the second speaker obtains a second motion state of the second speaker, and then configures a sound channel corresponding to the second motion state as a sound channel of the second speaker.

[0157] It should be noted that the second motion states of the first speaker and the second speaker are different. Different second motion states may correspond to different sound channel configurations. For example, during implementation, the second motion state may be described by using at least one of the following parameters:

(1) A quantity of shakes: For example, one shake indicates being used as a left sound channel, two shakes indicate being used as a right sound channel, and so on. Different quantities of shakes indicate different sound channel configurations.

(2) A speed of a shake: For example, a speed range may be set, where a first speed range corresponds to a left sound channel, a second speed range corresponds to a right sound channel, and so on. Different speed ranges indicate different sound channel configurations.

(3) An acceleration of a shake: Similarly, in addition to a speed, a sound channel configuration may alternatively be determined based on an acceleration. For example, a slight shake corresponds to a first acceleration range, and a sound channel is a left sound channel, while a violent shake is represented by a second acceleration range, corresponding to a right sound channel, and so on. During implementation, a plurality of acceleration ranges may be stored in a speaker, and an acceleration of a shake performed by a user is compared with an acceleration range, to determine an acceleration range to which the acceleration of the shake belongs, and then determine a sound channel to which the speaker belongs.

(4) A direction of a shake: For example, moving a speaker to the left indicates a left sound channel, moving the speaker to the right indicates a right sound channel, moving the speaker forward indicates a subwoofer, and so on, as long as different moving directions correspond to different sound channel configurations.

(5) A moving distance: For example, a movement within a first distance range indicates a left sound channel, and a movement within a second distance range indicates a right sound channel.

(6) A quantity of collisions: For example, one collision indicates a left sound channel, two collisions indicate a right sound channel, and so on.

(7) A motion trail or the like: For example, a trail resembling the number 1 indicates a left sound channel, a trail resembling the number 2 indicates a right sound channel, and a trail resembling a circle indicates a surround sound.

[0158] Clearly, during implementation, the foregoing parameters may not only be used alone, but also may be used in combination. For example, when a quantity of shakes and an acceleration of a shake are combined, one slight shake indicates a left sound channel, two slight shakes indicate a right sound channel, one violent shake indicates a rear left sound channel, and two violent shakes indicate a rear right sound channel. Any feasible combination manner is also applicable to this embodiment of this application.

[0159] For another example, a sound channel configuration of a speaker may alternatively be determined by determining whether a user has shaken the speaker or has touched the speaker. During implementation, whether an operation on the speaker is a shake or a touch may be determined based on a change frequency or magnitude of an acceleration with reference to a moving distance. In a possible implementation, it is considered as a touch when an acceleration is greater than g/2 and a distance between two devices is less than a distance threshold (for example, 5 cm) when building is triggered; otherwise, it is determined that the two devices are shaken.

[0160] In some other embodiments, the user may alternatively customize second motion states corresponding to different sound channel configurations.

[0161] In some embodiments, when each speaker configures a sound channel for the speaker, the speaker may prompt the user for a sound channel configured by

the speaker, so that the user can make a confirmation.

[0162] Step 710: Prompt the user that the stereo speaker system is built up, and prompt a sound channel configuration of each speaker, so that the user knows how the system is built up.

[0163] In conclusion, a case in which two speakers are used to build a speaker system is described above in detail. When more speakers are used for building, three speakers are used as an example to describe a manner of building.

Embodiment 1

[0164] It is assumed that, in this embodiment, after a first speaker and a second speaker are used to build a stereo speaker system according to the procedure shown in FIG. 7, a third speaker joins the speakers. It is assumed that the first speaker is a main speaker, as shown in FIG. 14:

> Step 1401: A user shakes the third speaker, and the third speaker obtains a first motion state of the third speaker.
>
> Step 1402: After determining that the first motion state matches a characteristic of a preset action and that a moving distance in the first motion state is greater than a specified distance, the third speaker searches for a nearby speaker.

[0165] The user may move at the same time, or move the first speaker and the third speaker successively within a specific time range, so that the first speaker and the third speaker can find each other through a search.

[0166] Step 1403: The first speaker is found, and therefore the first speaker interacts with the third speaker, to determine a distance between the first speaker and the third speaker.

[0167] In a possible implementation, the first speaker may notify the third speaker that the first speaker has been used to build a stereo speaker system, and therefore, step 1404 may be implemented.

[0168] Step 1404: If the distance between the first speaker and the third speaker is less than a distance threshold, the third speaker prompts the user for configuring a sound channel for the third speaker. To be specific, when learning that the third speaker has joined the existing stereo speaker system, the third speaker may prompt the user for configuring the sound channel for the third speaker.

[0169] Step 1405: After the sound channel is configured, the third speaker prompts the user for a sound channel configuration result, that is, prompts for a sound channel that has been configured.

[0170] In other possible implementations, the first speaker may determine a relative position relationship between the speakers, automatically determine a sound channel configuration of the third speaker based on the position relationship, and notify the third speaker of the sound channel configuration, and then the third speaker prompts the user for confirming the sound channel configuration. If the user confirms to agree on the sound channel configuration, the third speaker plays content based on the sound channel configuration; otherwise, if the user does not agree, the user may customize a sound channel configuration of the third speaker, and the sound channel configuration of the third speaker is completed.

[0171] Clearly, in some embodiments, because the third speaker joins the system, a sound channel configuration of each speaker may change. In this case, the user is also supported to re-customize the sound channel configuration of each speaker, or the first speaker redetermines the sound channel configuration of each speaker based on the relative position relationship between the speakers.

[0172] In addition, because the third speaker joins the system, the main speaker may be reselected, or the main speaker may remain unchanged. In a possible implementation, as described above, when a speaker is selected based on the foregoing priority, if a priority of the first speaker is high, the first speaker may still be used as the main speaker; or if a priority of the first speaker is low, whether to change the third speaker into the main speaker is considered. For example, if the first speaker is a speaker already connected to a network, an operation of selecting the main speaker may not be performed when the third speaker joins the system; or if the first speaker cannot be connected to a network, after the third speaker joins the system, the main speaker may be reselected.

[0173] It should be further noted that, in this embodiment, the third speaker requests to join the built stereo speaker system (that is, the speaker system built by using the first speaker and the second speaker). During implementation, the built speaker system may require that another speaker (for example, the third speaker) is allowed to join within a first time length threshold after the building is completed, or a first time length threshold may not be set. When the first time length threshold is not set, a new speaker may join the built speaker system at any time. The first time length threshold may be set by the user.

Embodiment 2

[0174] It is assumed that, in this embodiment, a first speaker, a second speaker, and a third speaker are moved by a user almost at the same time (for example, within a second time length threshold range) to build a stereo speaker system. As shown in FIG. 15, an implementation is as follows:

> Step 1501: The first speaker, the second speaker, and the third speaker are moved by the user, and each speaker obtains a first motion state of the speaker.
>
> Step 1502: After determining that the first motion state of each speaker matches a characteristic of a preset action for the speaker and that a moving dis-

tance of the speaker is greater than a specified distance, the speaker searches for a nearby speaker.
Step 1503: The speakers interact with each other to determine distances between the speakers.
Step 1504: If the distances between the speakers are all less than a distance threshold, each speaker prompts the user for configuring a sound channel for the speaker, and selects a main speaker.
Step 1505: After the user configures the sound channel, each speaker prompts the user for the sound channel configured for the speaker, and the building is completed.

[0175] After the building is completed, as shown in FIG. 16, if the main speaker is connected to a network, the main speaker may respond to a speech of the user, and obtain, through the network, an audio resource indicated by the user. After the audio resource is obtained, the audio resource is sent to each auxiliary speaker for play.

[0176] In addition, as shown in FIG. 17, if the main speaker is not connected to a network but is connected to a mobile phone of the user, the mobile phone may obtain the audio resource from the network, and then the main speaker distributes the audio resource to each auxiliary speaker for play.

[0177] Based on a same inventive concept, an embodiment of this application further provides a first speaker. As shown in FIG. 18, the speaker includes:

an obtaining module 1801, configured to obtain a first motion state of the first speaker;
a searching module 1802, configured to: if the first motion state matches a characteristic of a preset action, search for a second speaker; and
a system building module 1803, configured to: if the second speaker is found, build a stereo speaker system with the second speaker.

[0178] In a possible design, the speaker further includes:
a prompting module, configured to: before the first motion state of the first speaker is obtained, prompt a user for the preset action that needs to be performed for building the stereo speaker system.

[0179] In a possible design, the obtaining module is specifically configured to:
generate a first acceleration sequence of the first speaker based on acceleration information of the first speaker, where the first acceleration sequence of the first speaker stores first indication information arranged in a time sequence, and the first indication information is used to express a correspondence between an acceleration and duration of the acceleration.

[0180] The characteristic of the preset action includes a first sequence template, and the speaker further includes:

a first matching module, configured to: perform a matching operation on the first sequence template and the first acceleration sequence of the first speaker; and
if a quantity of times the first acceleration sequence of the first speaker matches the first sequence template is greater than or equal to a first specified quantity of times, determine that the first motion state matches the characteristic of the preset action; or
if a quantity of times the first acceleration sequence of the first speaker matches the first sequence template is less than a first specified quantity of times, determine that the first motion state does not match the characteristic of the preset action.

[0181] In a possible design, the obtaining module is specifically configured to:
generate a second acceleration sequence of the first speaker based on acceleration information of the first speaker, where the second acceleration sequence of the first speaker stores second indication information, and the second indication information is used to express an acceleration and frequency domain information of the acquired acceleration.

[0182] The characteristic of the preset action includes a second sequence template, and the speaker further includes:

a second matching module, configured to: perform a matching operation on the second sequence template and the second acceleration sequence of the first speaker; and
if a quantity of times the second acceleration sequence of the first speaker matches the second sequence template is greater than or equal to a second specified quantity of times, determine that the first motion state matches the characteristic of the preset action; or
if a quantity of times the second acceleration sequence of the first speaker matches the second sequence template is less than a second specified quantity of times, determine that the first motion state does not match the characteristic of the preset action.

[0183] In a possible design, the speaker further includes:
a distance determining module, configured to determine a moving distance of the first speaker based on the first motion state of the first speaker.

[0184] The speaker further includes:
the distance determining module, configured to: before the search for the second speaker, determine that the moving distance of the first speaker is greater than a specified distance.

[0185] In a possible design, the system building module is specifically configured to:

determine, according to a main speaker selection

rule, whether the first speaker is used as a main speaker in the stereo speaker system; and
configure a sound channel for the first speaker.

[0186] In a possible design, the main speaker selection rule includes at least one of the following rules:

selecting a speaker connected to a network as the main speaker;
selecting a speaker configured with a network but not connected to the network as the main speaker;
selecting a speaker connected to an intelligent terminal device as the main speaker;
selecting a speaker that first matches the characteristic of the preset action as the main speaker; and
selecting a speaker with a largest media access control address MAC address as the main speaker.

[0187] In a possible design, the system building module is specifically configured to:

prompt the user for configuring the sound channel for the first speaker;
obtain a second motion state of the first speaker; and
configure a sound channel corresponding to the second motion state as the sound channel of the first speaker.

[0188] In a possible design, the system building module is specifically configured to:
prompt, by using at least one of a sound effect, a light effect, and a screen display the user for configuring the sound channel for the first speaker.

[0189] In a possible design, the second motion state includes at least one of the following parameters:
a quantity of times the first speaker is shaken, a speed of a shake, an acceleration of the shake, a direction of the shake, a moving distance, a quantity of collisions of the first speaker, and a motion trail of the first speaker.

[0190] In a possible design, the system building module is further configured to:

determine a position relationship with the second speaker; and
if it is determined that the position relationship with the second speaker is a specified position relationship, perform the operation of building the stereo speaker system with the second speaker.

[0191] In a possible design, the specified position relationship includes that a distance between the first speaker and the second speaker is less than a distance threshold.

[0192] The implementations in this application may be combined in any manner to achieve different technical effects.

[0193] In the foregoing embodiments provided in this application, the method provided in embodiments of this application is described from the perspective of a speaker (a main speaker and/or an auxiliary speaker) as an execution subject. To implement functions in the method provided in embodiments of this application, a terminal device may include a hardware structure and/or a software module, and the functions are implemented by the hardware structure, the software module, or a combination of the hardware structure and the software module. Whether a specific function in the functions is performed by the hardware structure, the software module, or the combination of the hardware structure and the software module depends on particular application and design constraints of the technical solutions.

[0194] According to the context, the term "when" or "after" used in the foregoing embodiments may be interpreted as meaning "if", "after", "in response to determining", or "in response to detecting". Similarly, according to the context, the phrase "when it is determined that" or "if (a stated condition or event) is detected" may be interpreted as meaning "if it is determined that", "in response to determining", "when (the stated condition or event) is detected", or "in response to detecting (the stated condition or event)". In addition, in the foregoing embodiments, relationship terms such as first and second are used to distinguish one entity from another entity, but do not limit any actual relationship and sequence between these entities.

[0195] All or some of the foregoing embodiments may be implemented by using software, hardware, firmware, or any combination thereof. When software is used for implementation, all or some of embodiments may be implemented in a form of a computer program product. The computer program product includes one or more computer instructions. When the computer program instructions are loaded and executed on a computer, the procedures or functions according to embodiments of the present invention are all or partially generated. The computer may be a general-purpose computer, a dedicated computer, a computer network, or another programmable apparatus. The computer instructions may be stored in a computer-readable storage medium or may be transmitted from a computer-readable storage medium to another computer-readable storage medium. For example, the computer instructions may be transmitted from a website, computer, server, or data center to another website, computer, server, or data center in a wired (for example, by using a coaxial cable, an optical fiber, or a digital subscriber line) or wireless (for example, via infrared, radio, or microwaves) manner. The computer-readable storage medium may be any usable medium accessible by the computer, or a data storage device, for example, a server or a data center, integrating one or more usable media. The usable medium may be a magnetic medium (for example, a floppy disk, a hard disk, or a magnetic tape), an optical medium (for example, a DVD), a semiconductor medium (for example, a solid-state drive (solid-state drive, SSD)), or the like.

[0196] It should be noted that a part of this patent ap-

plication document includes copyright-protected content. The copyright owner reserves the copyright except cases in which copies are made by using patent files or documented content of the patent files in the patent office.

Claims

1. A method for building a stereo speaker system, applied to a first speaker, wherein the method comprises:

   obtaining a first motion state of the first speaker;
   if the first motion state matches a characteristic of a preset action, searching for a second speaker; and
   if the second speaker is found, building a stereo speaker system with the second speaker.

2. The method according to claim 1, wherein before the obtaining a first motion state of the first speaker, the method further comprises:
   prompting a user for the preset action that needs to be performed for building the stereo speaker system.

3. The method according to claim 1 or 2, wherein the obtaining a first motion state of the first speaker comprises:

   generating a first acceleration sequence of the first speaker based on acceleration information of the first speaker, wherein the first acceleration sequence of the first speaker stores first indication information arranged in a time sequence, and the first indication information is used to express a correspondence between an acceleration and duration of the acceleration; and
   the characteristic of the preset action comprises a first sequence template, and the method further comprises:

      performing a matching operation on the first sequence template and the first acceleration sequence of the first speaker; and
      if a quantity of times the first acceleration sequence of the first speaker matches the first sequence template is greater than or equal to a first specified quantity of times, determining that the first motion state matches the characteristic of the preset action; or
      if a quantity of times the first acceleration sequence of the first speaker matches the first sequence template is less than a first specified quantity of times, determining that the first motion state does not match the characteristic of the preset action.

4. The method according to claim 1 or 2, wherein the obtaining a first motion state of the first speaker comprises:

   generating a second acceleration sequence of the first speaker based on acceleration information of the first speaker, wherein the second acceleration sequence of the first speaker stores second indication information, and the second indication information is used to express an acceleration and frequency domain information of the acquired acceleration; and
   the characteristic of the preset action comprises a second sequence template, and the method further comprises:

      performing a matching operation on the second sequence template and the second acceleration sequence of the first speaker; and
      if a quantity of times the second acceleration sequence of the first speaker matches the second sequence template is greater than or equal to a second specified quantity of times, determining that the first motion state matches the characteristic of the preset action; or
      if a quantity of times the second acceleration sequence of the first speaker matches the second sequence template is less than a second specified quantity of times, determining that the first motion state does not match the characteristic of the preset action.

5. The method according to claim 3 or 4, wherein the method further comprises:

   determining a moving distance of the first speaker based on the first motion state of the first speaker; and
   before the searching for a second speaker, the method further comprises:
   determining that the moving distance of the first speaker is greater than a specified distance.

6. The method according to any one of claims 1 to 5, wherein the building a stereo speaker system with the second speaker comprises:

   determining, according to a main speaker selection rule, whether the first speaker is used as a main speaker in the stereo speaker system; and configuring a sound channel for the first speaker.

7. The method according to claim 6, wherein the main speaker selection rule comprises at least one of the following rules:

selecting a speaker connected to a network as the main speaker;
selecting a speaker configured with a network but not connected to the network as the main speaker;
selecting a speaker connected to an intelligent terminal device as the main speaker;
selecting a speaker that first matches the characteristic of the preset action as the main speaker; and
selecting a speaker with a largest media access control address MAC address as the main speaker.

8. The method according to claim 6, wherein the configuring a sound channel for the first speaker comprises:

prompting the user for configuring the sound channel for the first speaker;
obtaining a second motion state of the first speaker; and
configuring a sound channel corresponding to the second motion state as the sound channel of the first speaker.

9. The method according to claim 8, wherein the prompting the user for configuring the sound channel for the first speaker comprises:
prompting, by using at least one of a sound effect, a light effect, and a screen display, the user for configuring the sound channel for the first speaker.

10. The method according to claim 8, wherein the second motion state comprises at least one of the following parameters:
a quantity of times the first speaker is shaken, a speed of a shake, an acceleration of the shake, a direction of the shake, a moving distance, a quantity of collisions of the first speaker, and a motion trail of the first speaker.

11. The method according to any one of claims 1 to 10, wherein the method further comprises:

determining a position relationship with the second speaker; and
if it is determined that the position relationship with the second speaker is a specified position relationship, performing the operation of building the stereo speaker system with the second speaker.

12. The method according to claim 11, wherein the specified position relationship comprises that a distance between the first speaker and the second speaker is less than a distance threshold.

13. A first speaker, wherein the speaker comprises:

an obtaining module, configured to obtain a first motion state of the first speaker;
a searching module, configured to: if the first motion state matches a characteristic of a preset action, search for a second speaker; and
a system building module, configured to: if the second speaker is found, build a stereo speaker system with the second speaker.

14. The speaker according to claim 13, wherein the speaker further comprises:
a prompting module, configured to: before the first motion state of the first speaker is obtained, prompt a user for the preset action that needs to be performed for building the stereo speaker system.

15. The speaker according to claim 13 or 14, wherein the obtaining module is specifically configured to:

generate a first acceleration sequence of the first speaker based on acceleration information of the first speaker, wherein the first acceleration sequence of the first speaker stores first indication information arranged in a time sequence, and the first indication information is used to express a correspondence between an acceleration and duration of the acceleration; and
the characteristic of the preset action comprises a first sequence template, and the speaker further comprises:

a first matching module, configured to: perform a matching operation on the first sequence template and the first acceleration sequence of the first speaker; and
if a quantity of times the first acceleration sequence of the first speaker matches the first sequence template is greater than or equal to a first specified quantity of times, determine that the first motion state matches the characteristic of the preset action; or
if a quantity of times the first acceleration sequence of the first speaker matches the first sequence template is less than a first specified quantity of times, determine that the first motion state does not match the characteristic of the preset action.

16. The speaker according to claim 13 or 14, wherein the obtaining module is specifically configured to:

generate a second acceleration sequence of the first speaker based on acceleration information of the first speaker, wherein the second acceleration sequence of the first speaker stores second indication information, and the second indi-

cation information is used to express an acceleration and frequency domain information of the acquired acceleration; and

the characteristic of the preset action comprises a second sequence template, and the speaker further comprises:

a second matching module, configured to: perform a matching operation on the second sequence template and the second acceleration sequence of the first speaker; and

if a quantity of times the second acceleration sequence of the first speaker matches the second sequence template is greater than or equal to a second specified quantity of times, determine that the first motion state matches the characteristic of the preset action; or

if a quantity of times the second acceleration sequence of the first speaker matches the second sequence template is less than a second specified quantity of times, determine that the first motion state does not match the characteristic of the preset action.

17. The speaker according to claim 15 or 16, wherein the speaker further comprises:

a distance determining module, configured to determine a moving distance of the first speaker based on the first motion state of the first speaker; and

the speaker further comprises:

the distance determining module, configured to: before the search for the second speaker, determine that the moving distance of the first speaker is greater than a specified distance.

18. The speaker according to any one of claims 13 to 17, wherein the system building module is specifically configured to:

determine, according to a main speaker selection rule, whether the first speaker is used as a main speaker in the stereo speaker system; and configure a sound channel for the first speaker.

19. The speaker according to claim 18, wherein the main speaker selection rule comprises at least one of the following rules:

selecting a speaker connected to a network as the main speaker;

selecting a speaker configured with a network but not connected to the network as the main speaker;

selecting a speaker connected to an intelligent terminal device as the main speaker;

selecting a speaker that first matches the characteristic of the preset action as the main speaker; and

selecting a speaker with a largest media access control address MAC address as the main speaker.

20. The speaker according to claim 18, wherein the system building module is specifically configured to:

prompt the user for configuring the sound channel for the first speaker;

obtain a second motion state of the first speaker; and

configure a sound channel corresponding to the second motion state as the sound channel of the first speaker.

21. The speaker according to claim 20, wherein the system building module is specifically configured to:

prompt, by using at least one of a sound effect, a light effect, and a screen display, the user for configuring the sound channel for the first speaker.

22. The speaker according to claim 20, wherein the second motion state comprises at least one of the following parameters:

a quantity of times the first speaker is shaken, a speed of a shake, an acceleration of the shake, a direction of the shake, a moving distance, a quantity of collisions of the first speaker, and a motion trail of the first speaker.

23. The speaker according to any one of claims 13 to 22, wherein the system building module is further configured to:

determine a position relationship with the second speaker; and

if it is determined that the position relationship with the second speaker is a specified position relationship, perform the operation of building the stereo speaker system with the second speaker.

24. The speaker according to claim 23, wherein the specified position relationship comprises that a distance between the first speaker and the second speaker is less than a distance threshold.

25. A first speaker, comprising one or more processors, one or more memories, one or more microphones, one or more loudspeakers, and a communication module, wherein

the one or more microphones are configured to

acquire a sound signal;
the communication module is configured to communicate with another speaker;
the one or more loudspeakers are configured to emit a sound signal; and
the one or more memories are configured to store program instructions, wherein the program instructions are executed by the one or more processors, so that the speaker performs the method according to any one of claims 1 to 12.

26. A chip, wherein the chip comprises a processor and an interface;

the interface is configured to: receive code instructions, and transmit the received code instructions to the processor; and
the processor is configured to run the received code instructions sent through the interface, to perform the method according to any one of claims 1 to 12.

27. A computer-readable storage medium, wherein the computer-readable storage medium stores a computer program, and the computer program comprises program instructions; and when the program instructions are executed by a computer, the computer is enabled to perform the method according to any one of claims 1 to 12.

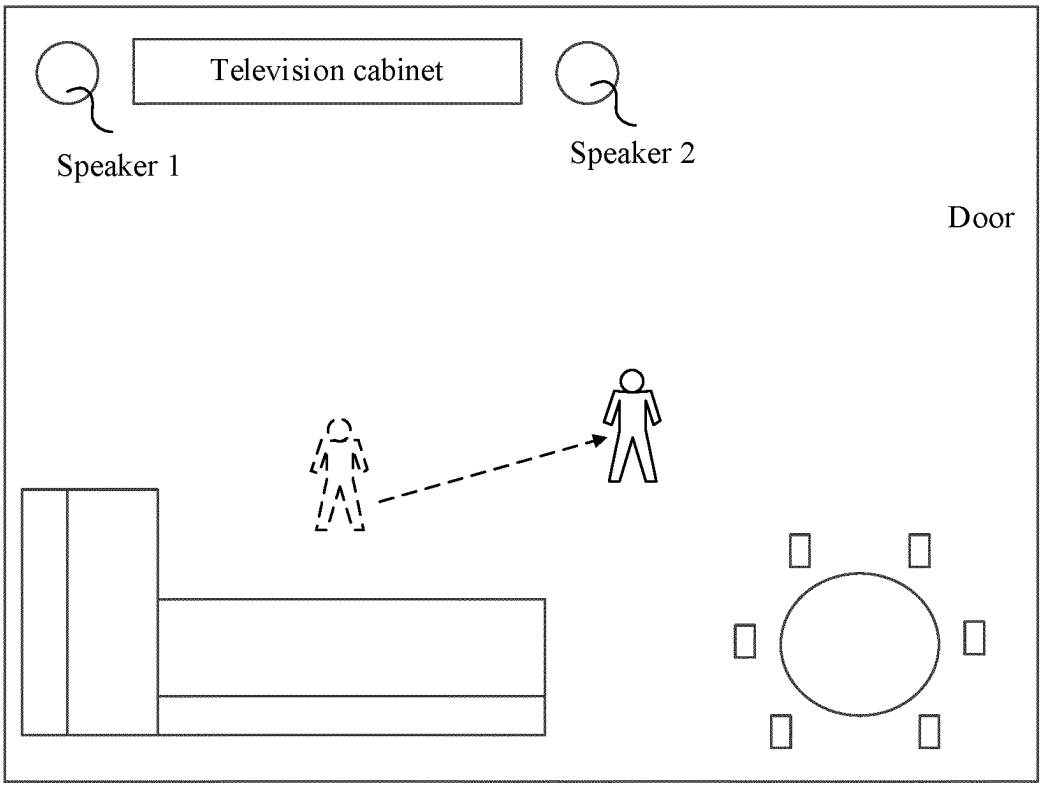
Television cabinet
Speaker 1
Speaker 2
Door

FIG. 1

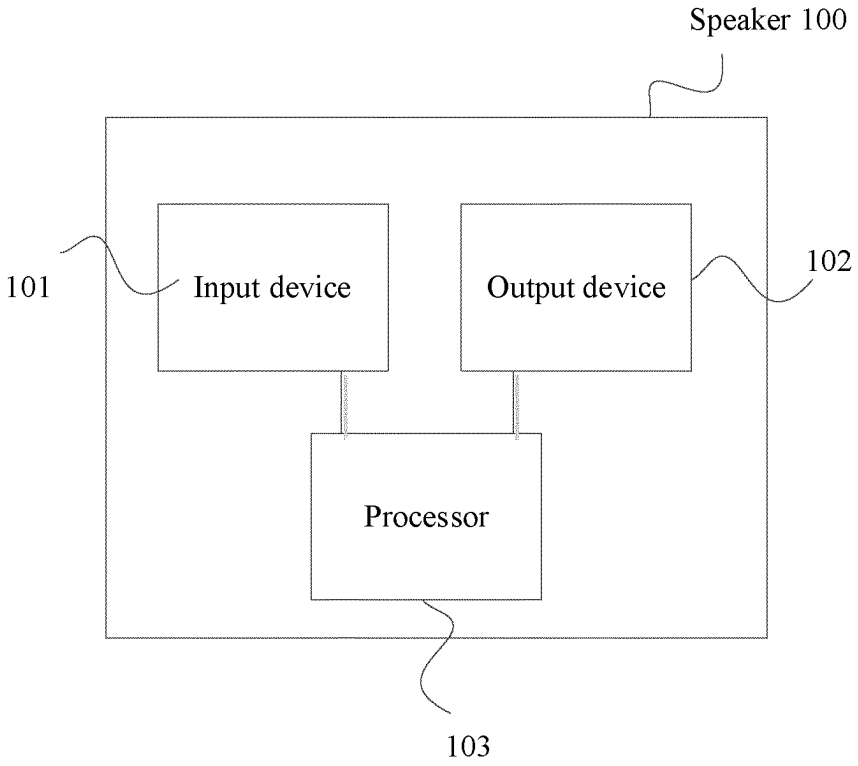
Speaker 100
101
Input device
Output device
102
Processor
103

FIG. 2

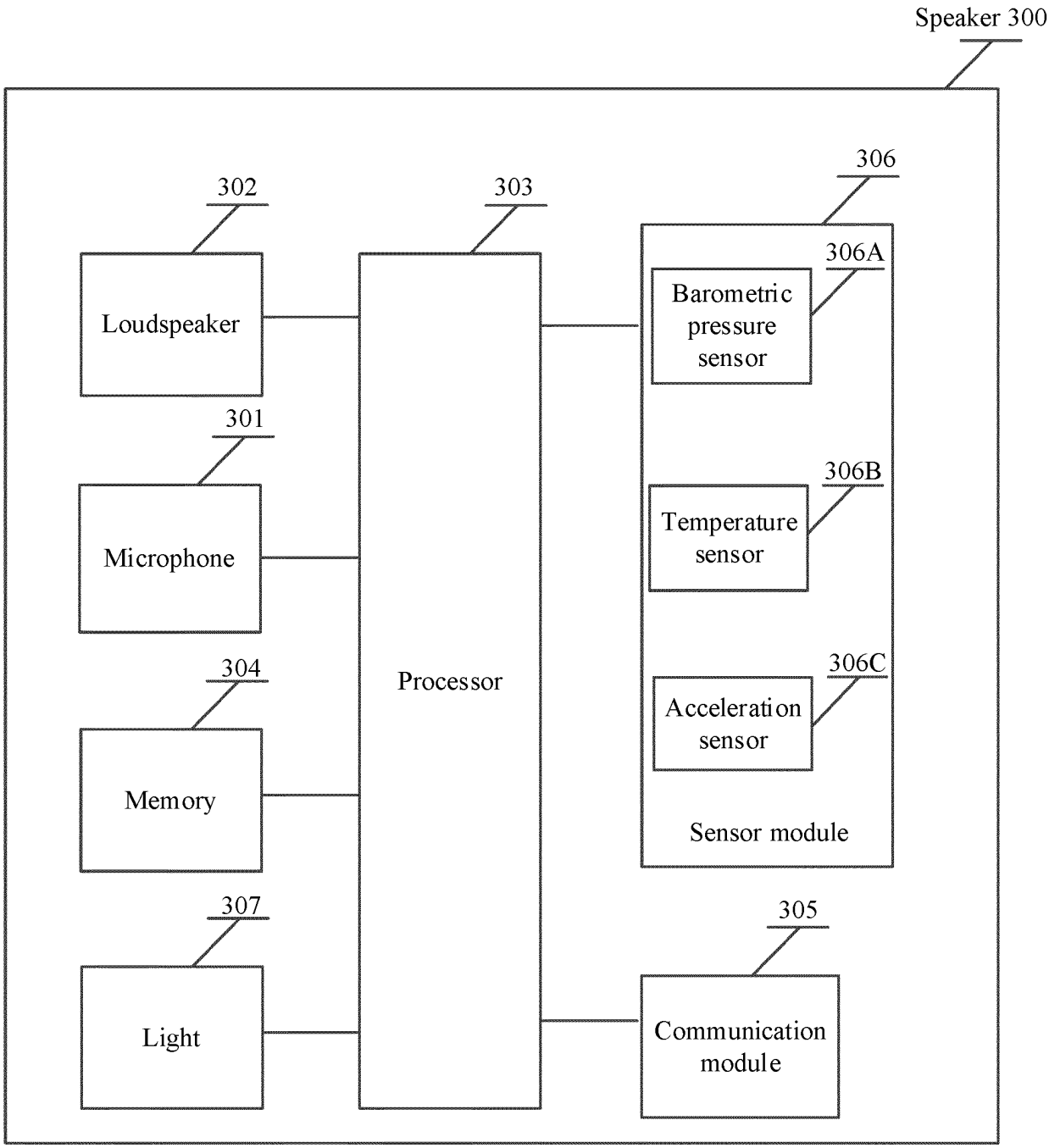
Speaker 300
302
303
306
306A
Loudspeaker
Barometric pressure sensor
301
306B
Microphone
Temperature sensor
304
Processor
306C
Acceleration sensor
Memory
Sensor module
307
305
Light
Communication module

FIG. 3

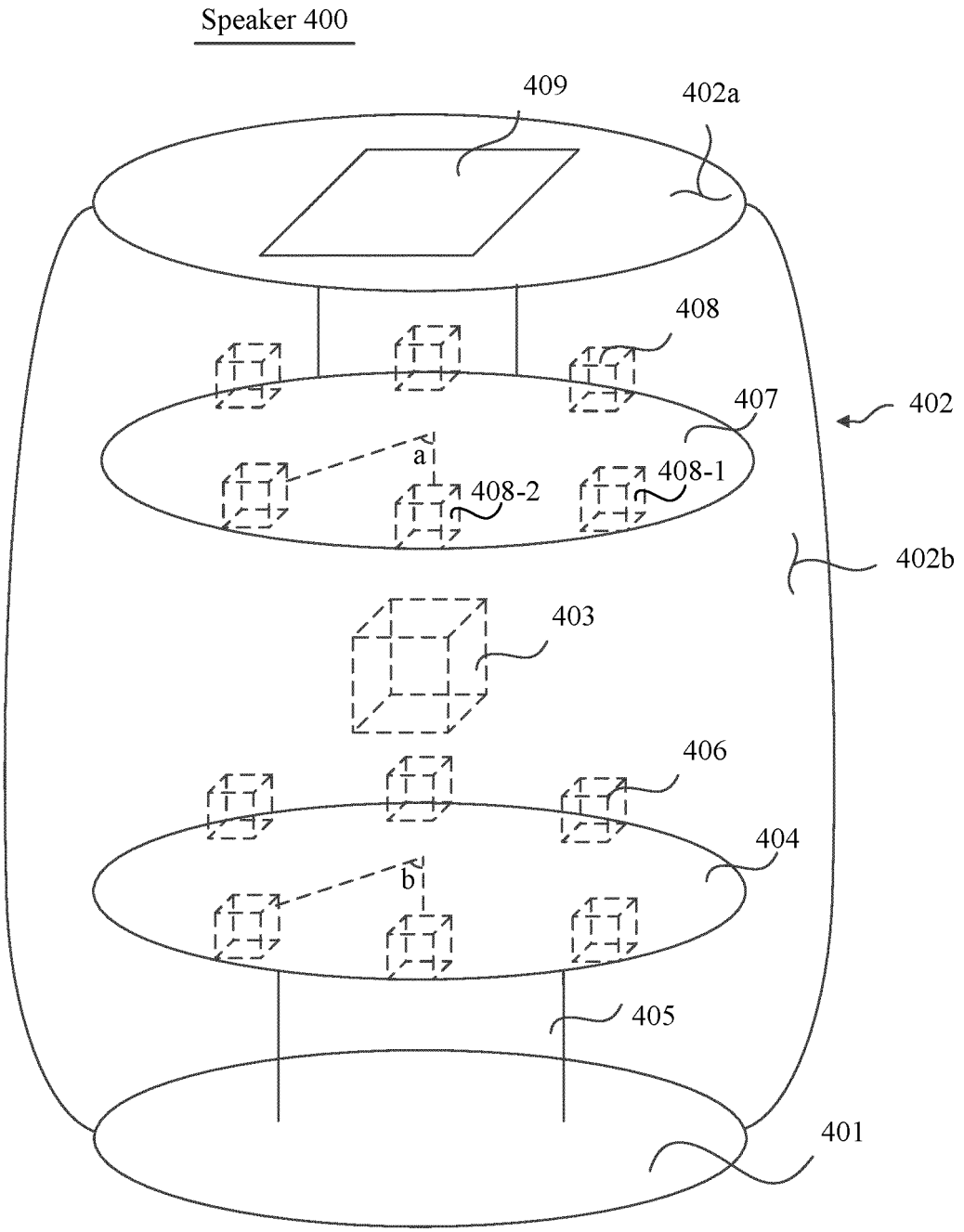
Speaker 400
409
402a
408
407
402
a
408-1
408-2
402b
403
406
404
b
405
401

FIG. 4A

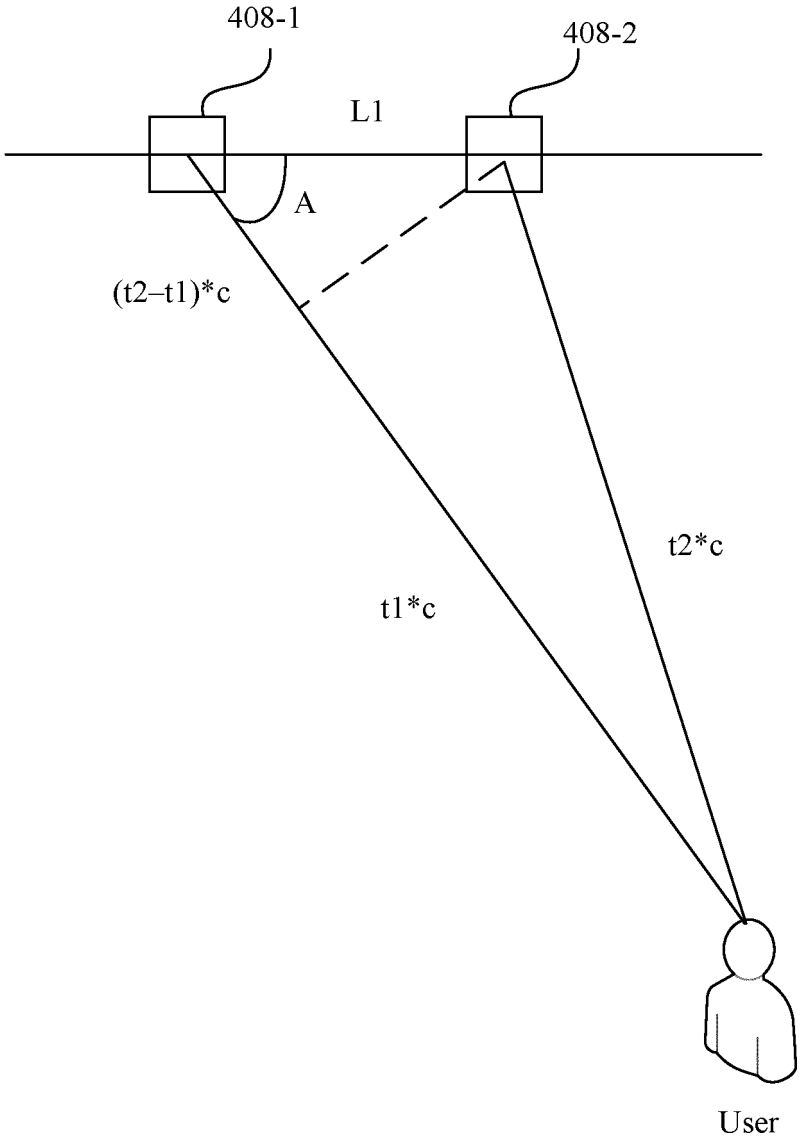

FIG. 4B

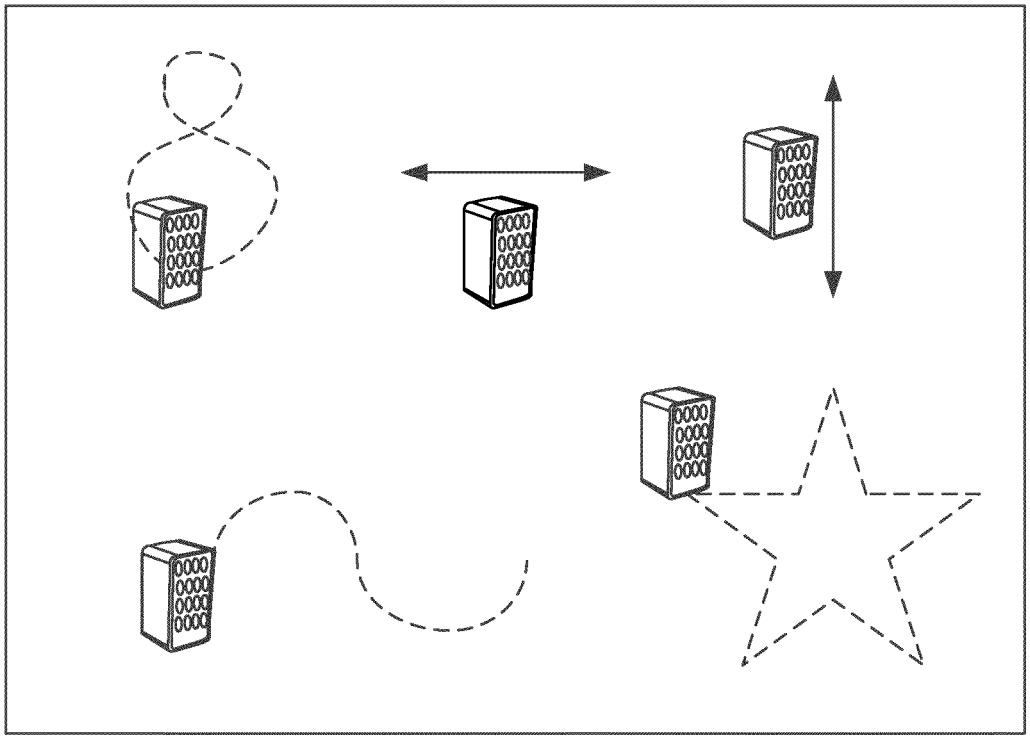

FIG. 5

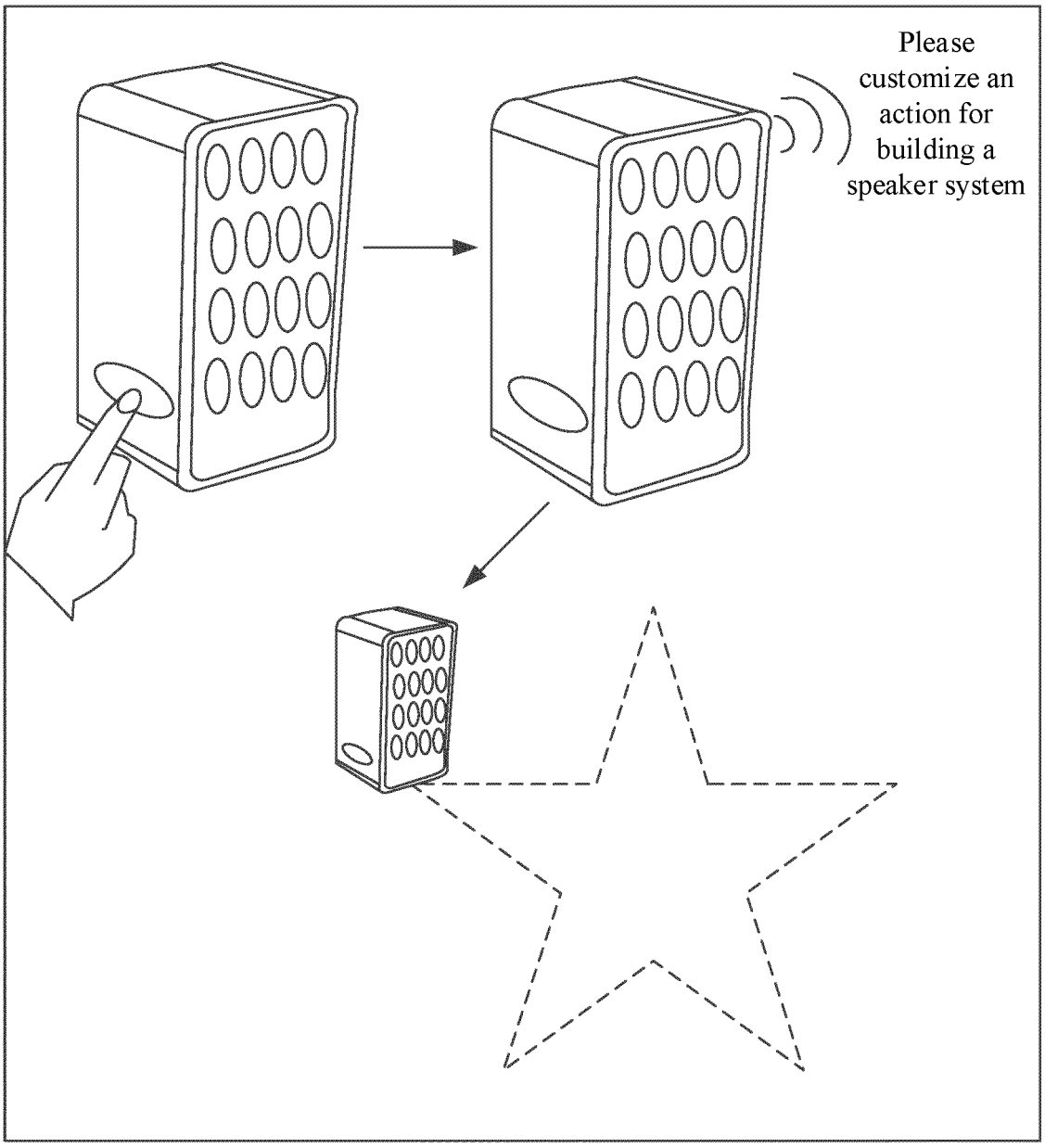

FIG. 6

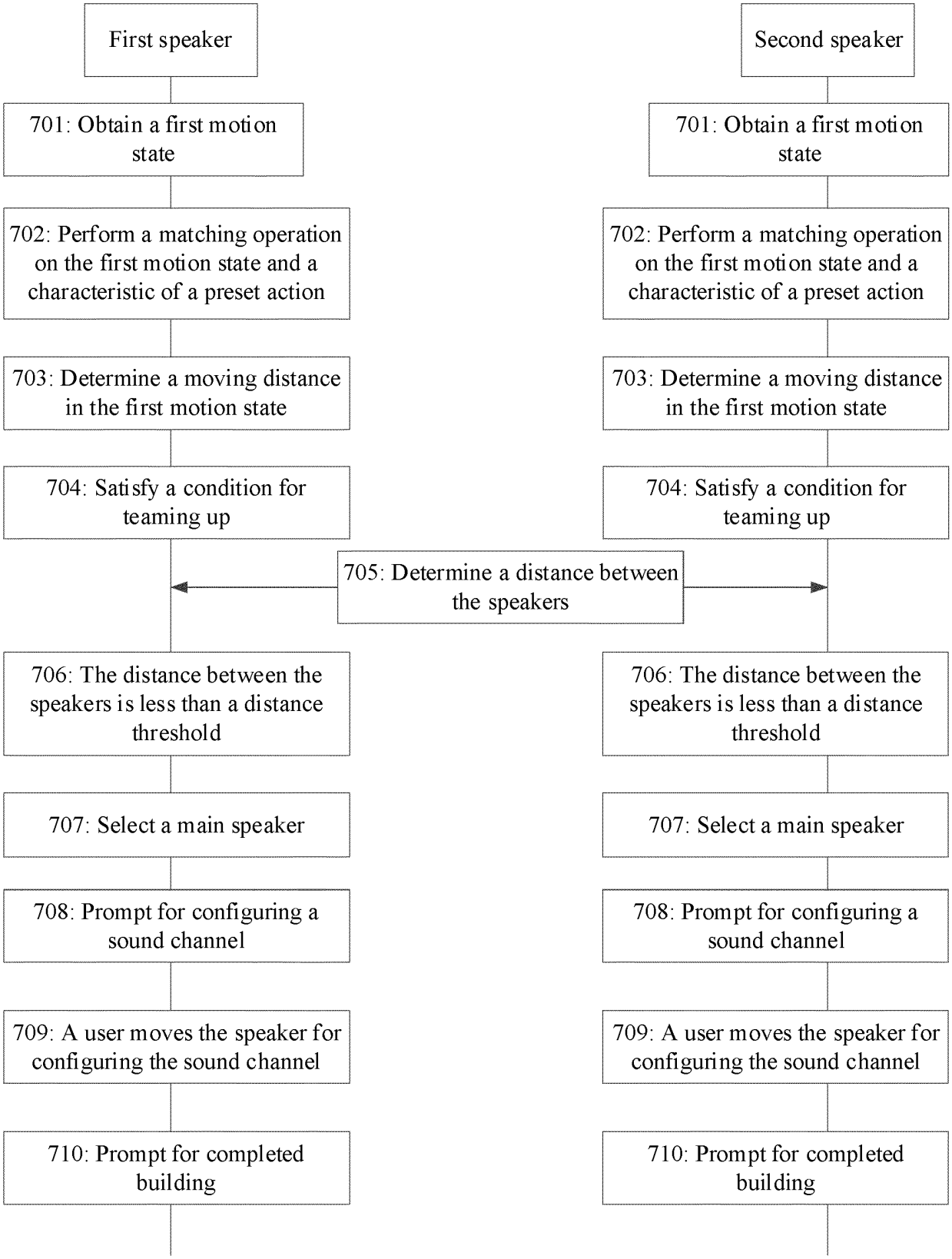
First speaker
Second speaker
701: Obtain a first motion state
701: Obtain a first motion state
702: Perform a matching operation on the first motion state and a characteristic of a preset action
702: Perform a matching operation on the first motion state and a characteristic of a preset action
703: Determine a moving distance in the first motion state
703: Determine a moving distance in the first motion state
704: Satisfy a condition for teaming up
704: Satisfy a condition for teaming up
705: Determine a distance between the speakers
706: The distance between the speakers is less than a distance threshold
706: The distance between the speakers is less than a distance threshold
707: Select a main speaker
707: Select a main speaker
708: Prompt for configuring a sound channel
708: Prompt for configuring a sound channel
709: A user moves the speaker for configuring the sound channel
709: A user moves the speaker for configuring the sound channel
710: Prompt for completed building
710: Prompt for completed building

FIG. 7

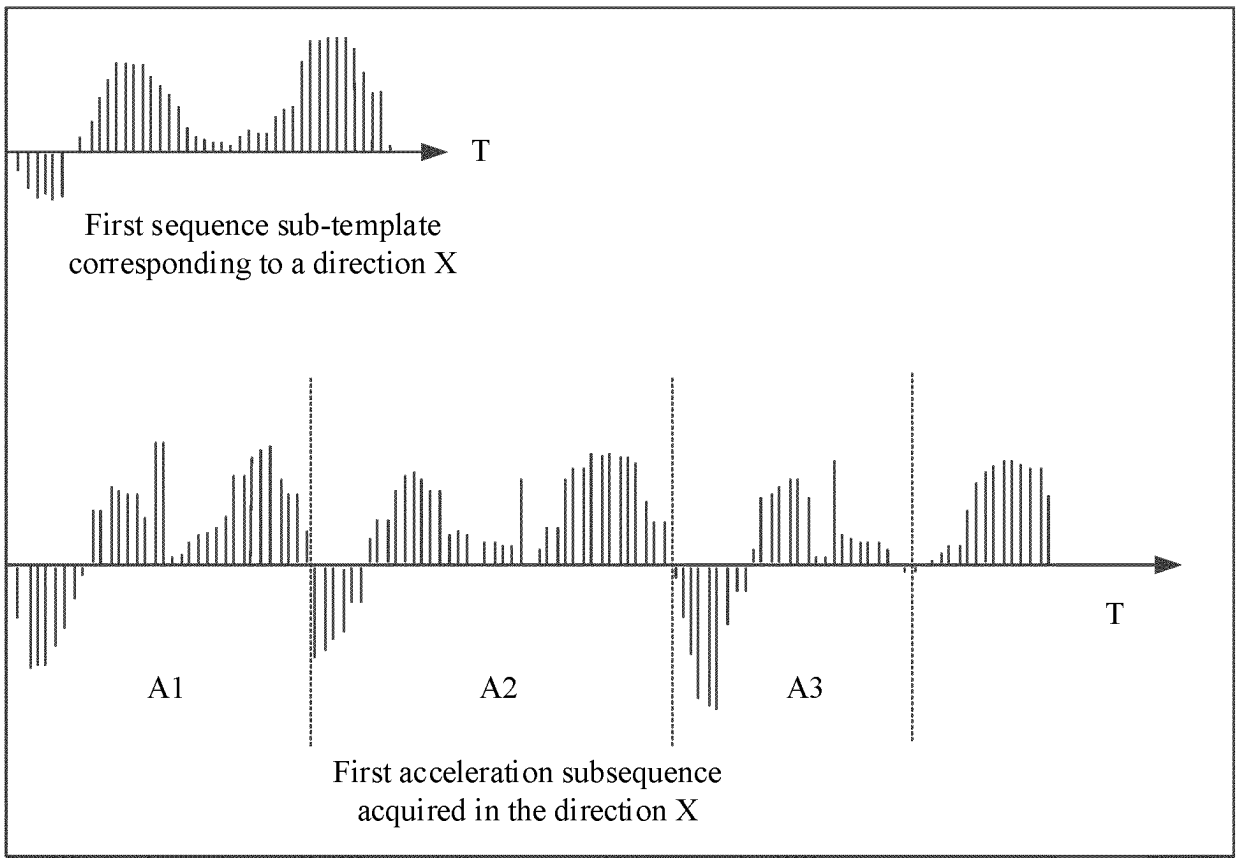
T
First sequence sub-template
corresponding to a direction X
T
A1
A2
A3
First acceleration subsequence
acquired in the direction X

FIG. 8

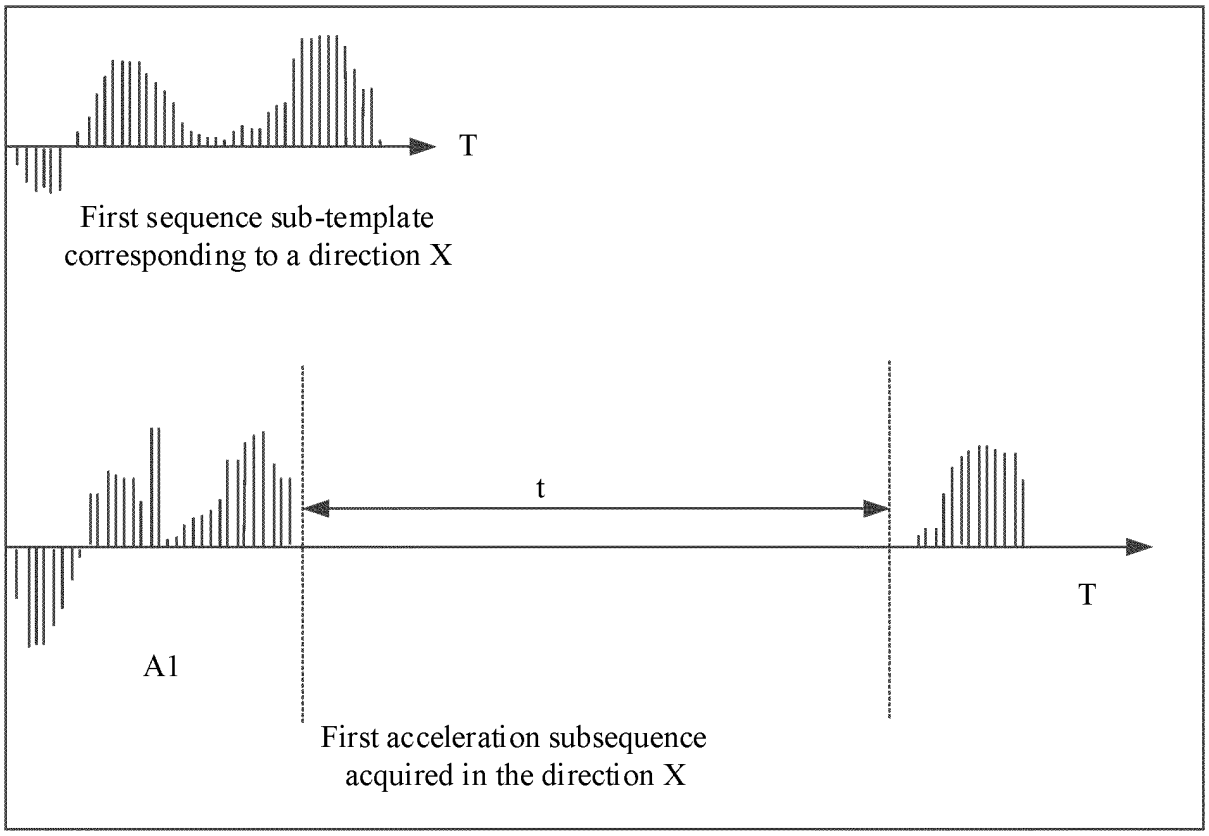
T
First sequence sub-template corresponding to a direction X
t
T
A1
First acceleration subsequence acquired in the direction X

FIG. 9

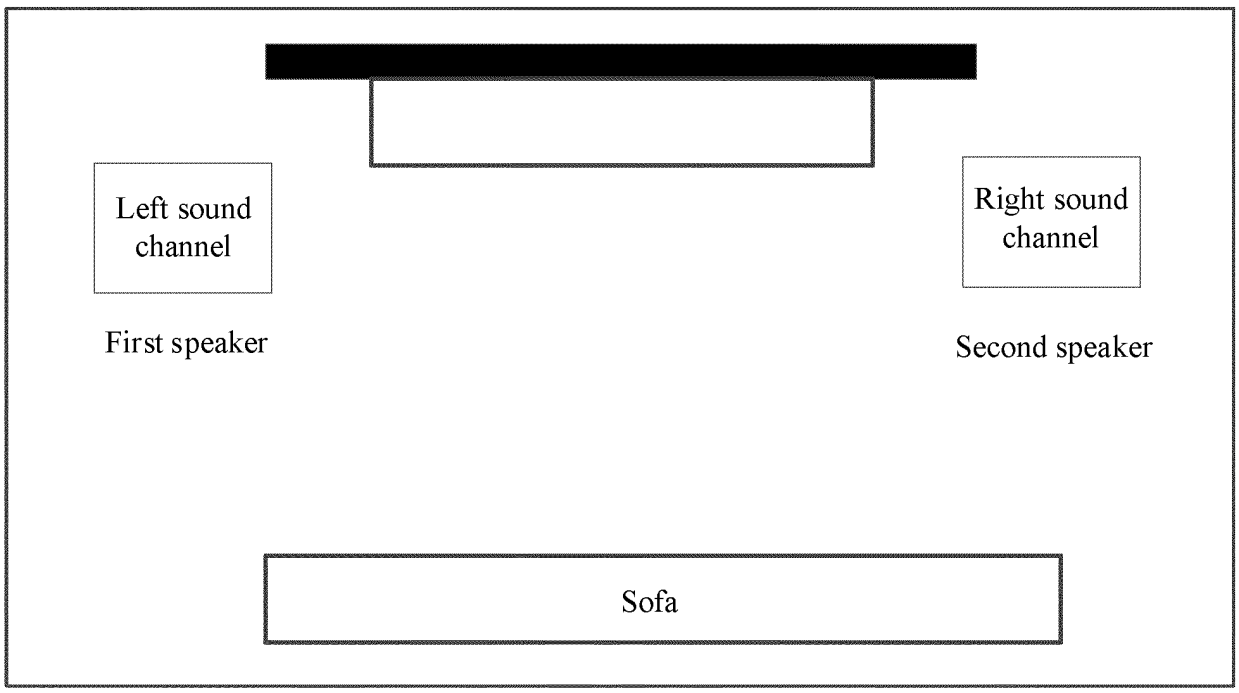
Left sound channel
Right sound channel
First speaker
Second speaker
Sofa

FIG. 10

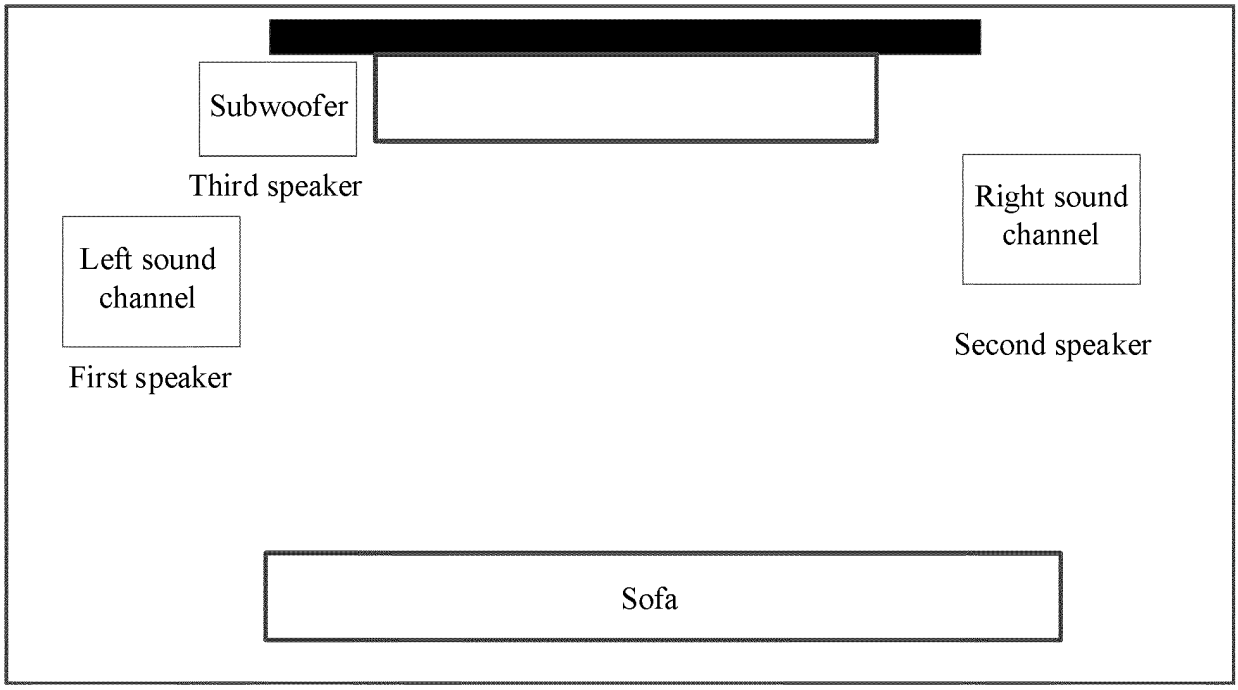
Subwoofer
Third speaker
Left sound channel
First speaker
Right sound channel
Second speaker
Sofa

FIG. 11

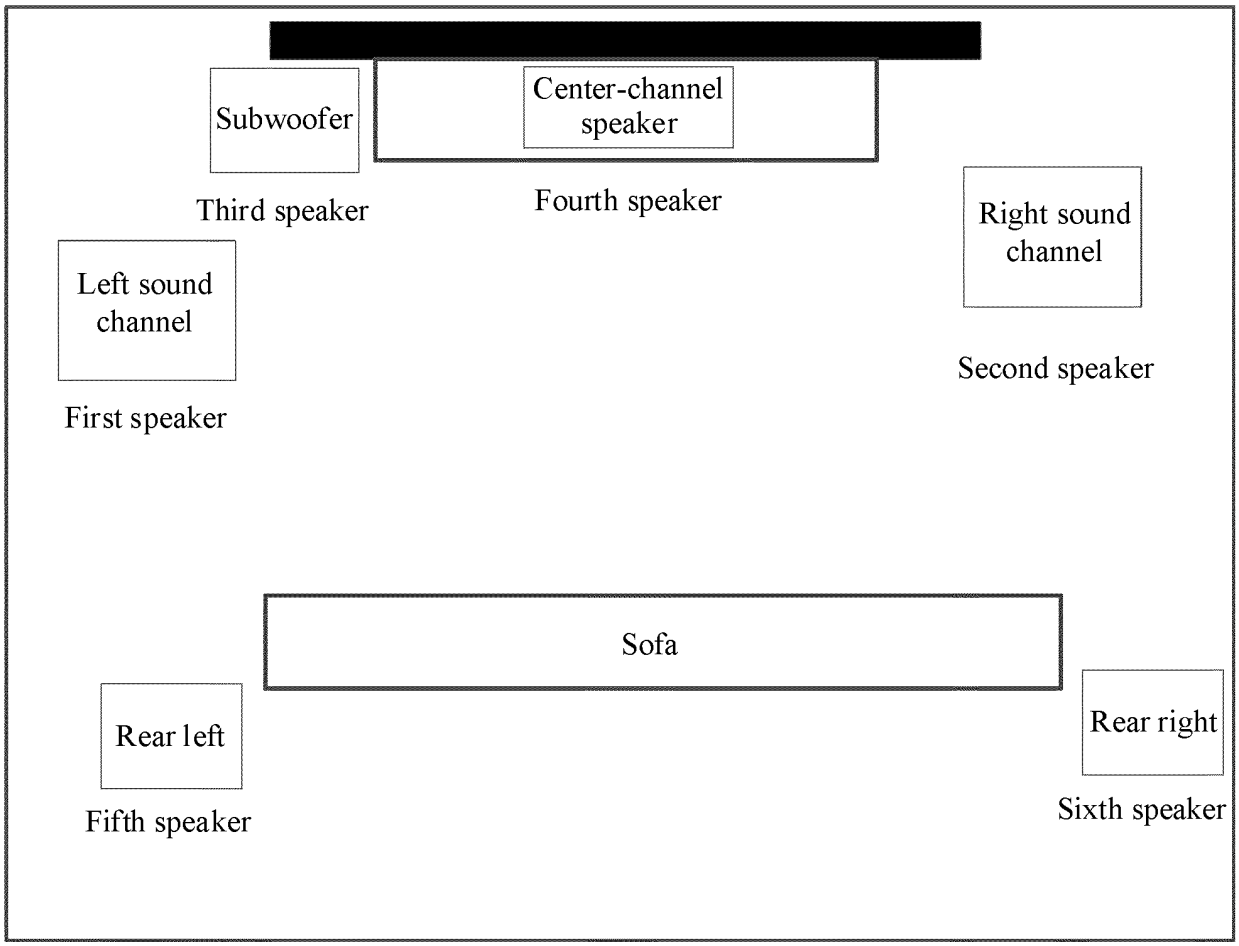
Subwoofer
Center-channel speaker
Third speaker
Fourth speaker
Right sound channel
Left sound channel
Second speaker
First speaker
Sofa
Rear left
Rear right
Fifth speaker
Sixth speaker

FIG. 12

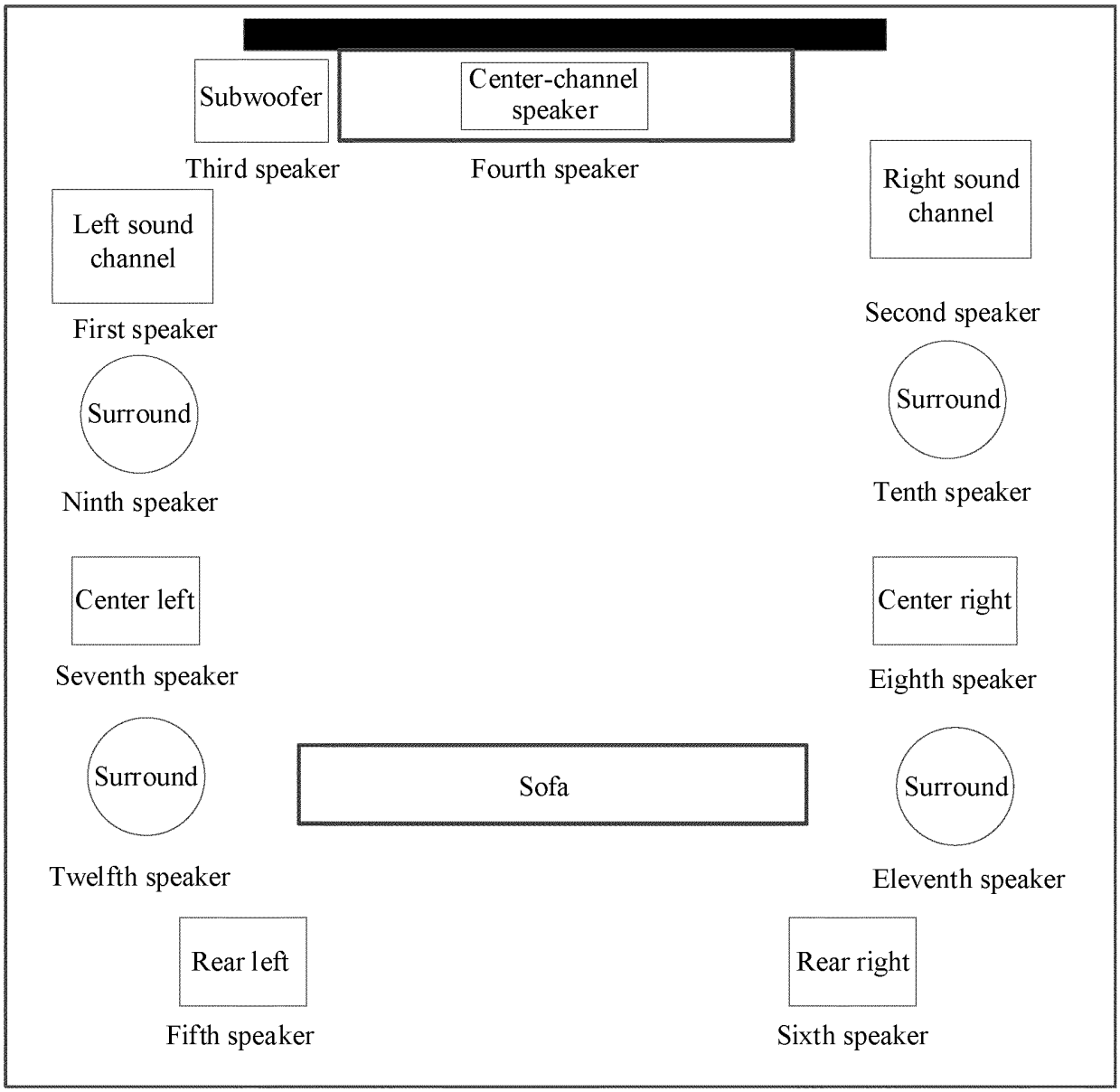
Subwoofer
Center-channel speaker
Third speaker
Fourth speaker
Right sound channel
Left sound channel
Second speaker
First speaker
Surround
Surround
Ninth speaker
Tenth speaker
Center left
Center right
Seventh speaker
Eighth speaker
Surround
Sofa
Surround
Twelfth speaker
Eleventh speaker
Rear left
Rear right
Fifth speaker
Sixth speaker

FIG. 13

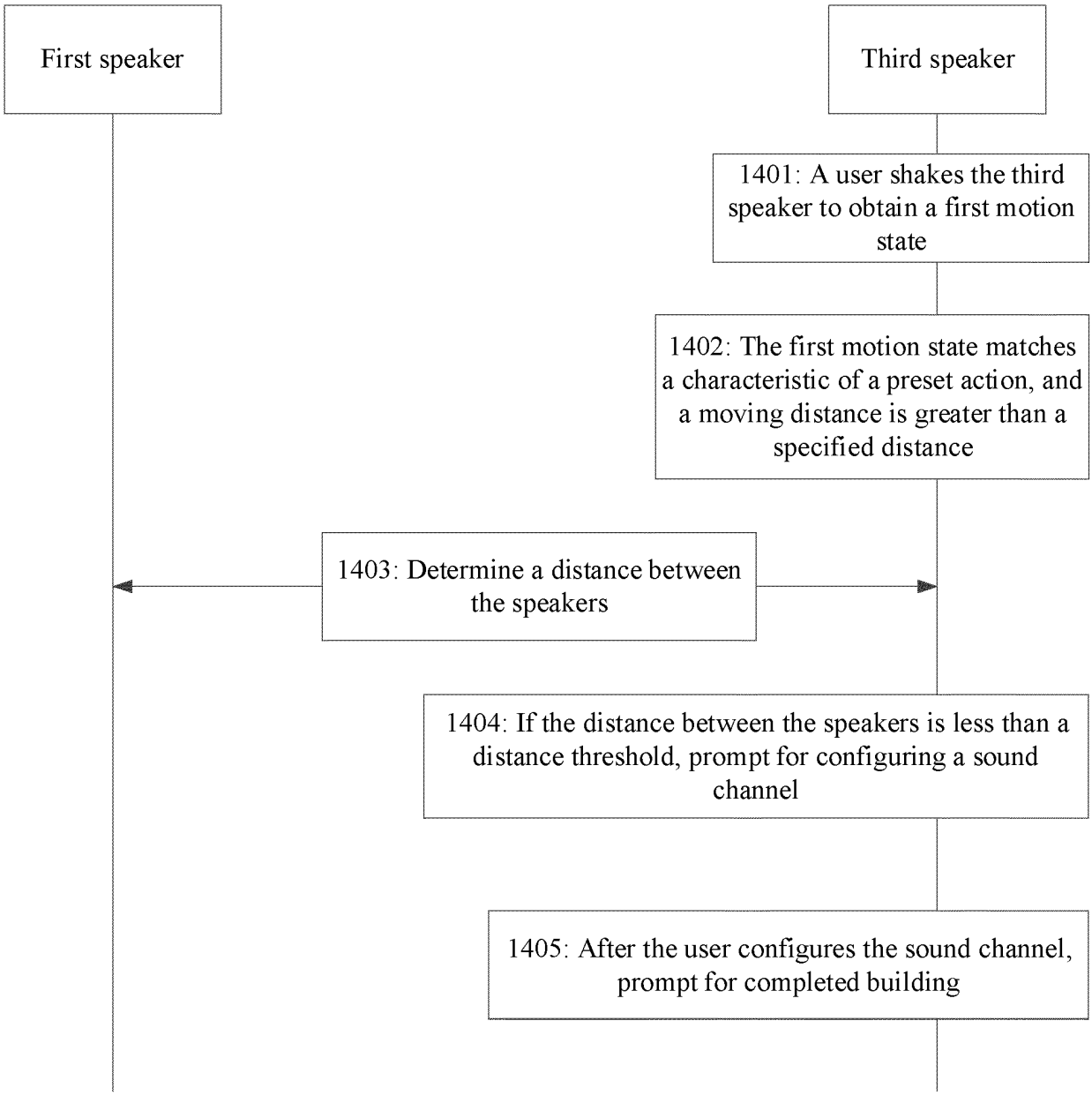
First speaker
Third speaker
1401: A user shakes the third speaker to obtain a first motion state
1402: The first motion state matches a characteristic of a preset action, and a moving distance is greater than a specified distance
1403: Determine a distance between the speakers
1404: If the distance between the speakers is less than a distance threshold, prompt for configuring a sound channel
1405: After the user configures the sound channel, prompt for completed building

FIG. 14

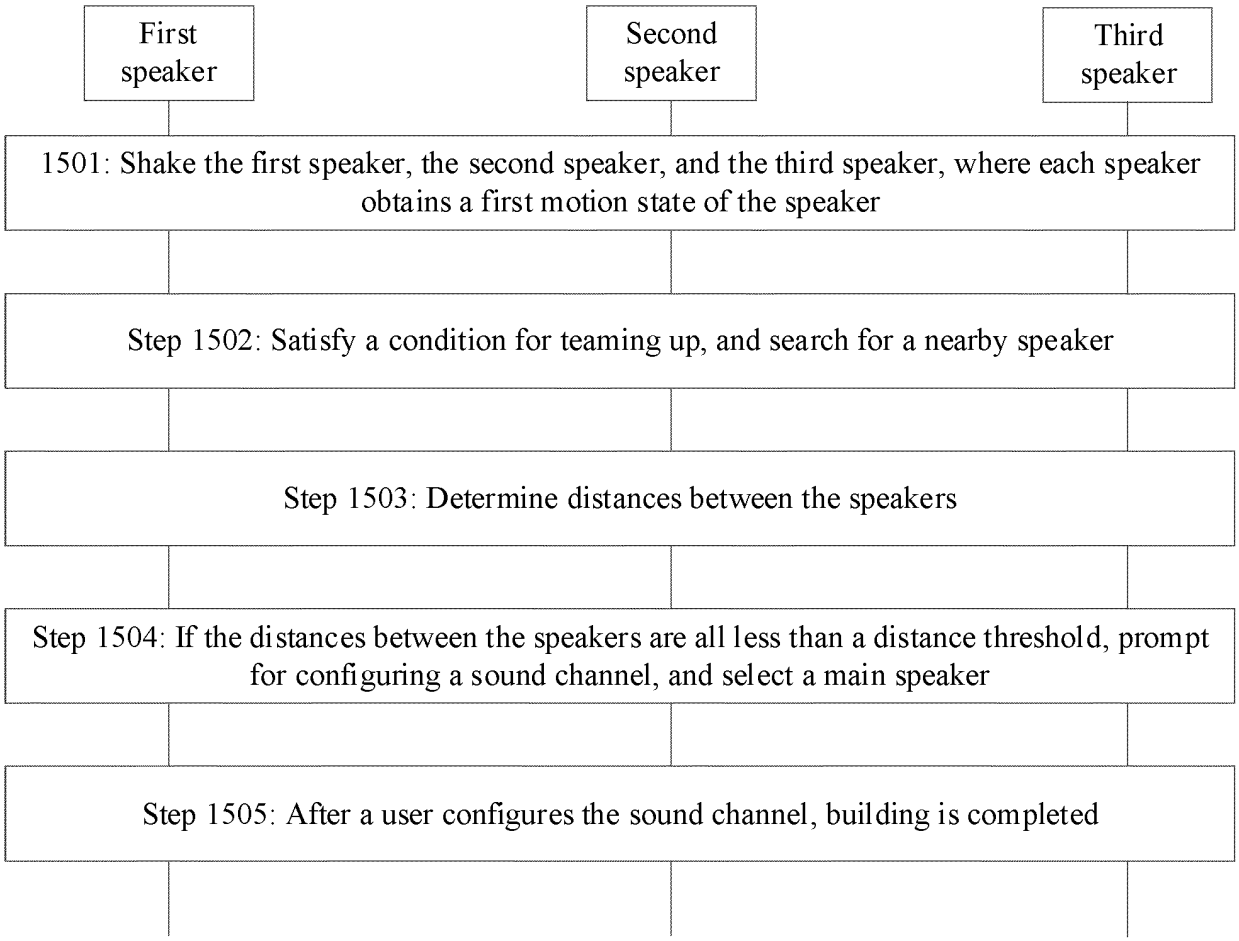
First speaker
Second speaker
Third speaker
1501: Shake the first speaker, the second speaker, and the third speaker, where each speaker obtains a first motion state of the speaker
Step 1502: Satisfy a condition for teaming up, and search for a nearby speaker
Step 1503: Determine distances between the speakers
Step 1504: If the distances between the speakers are all less than a distance threshold, prompt for configuring a sound channel, and select a main speaker
Step 1505: After a user configures the sound channel, building is completed

FIG. 15

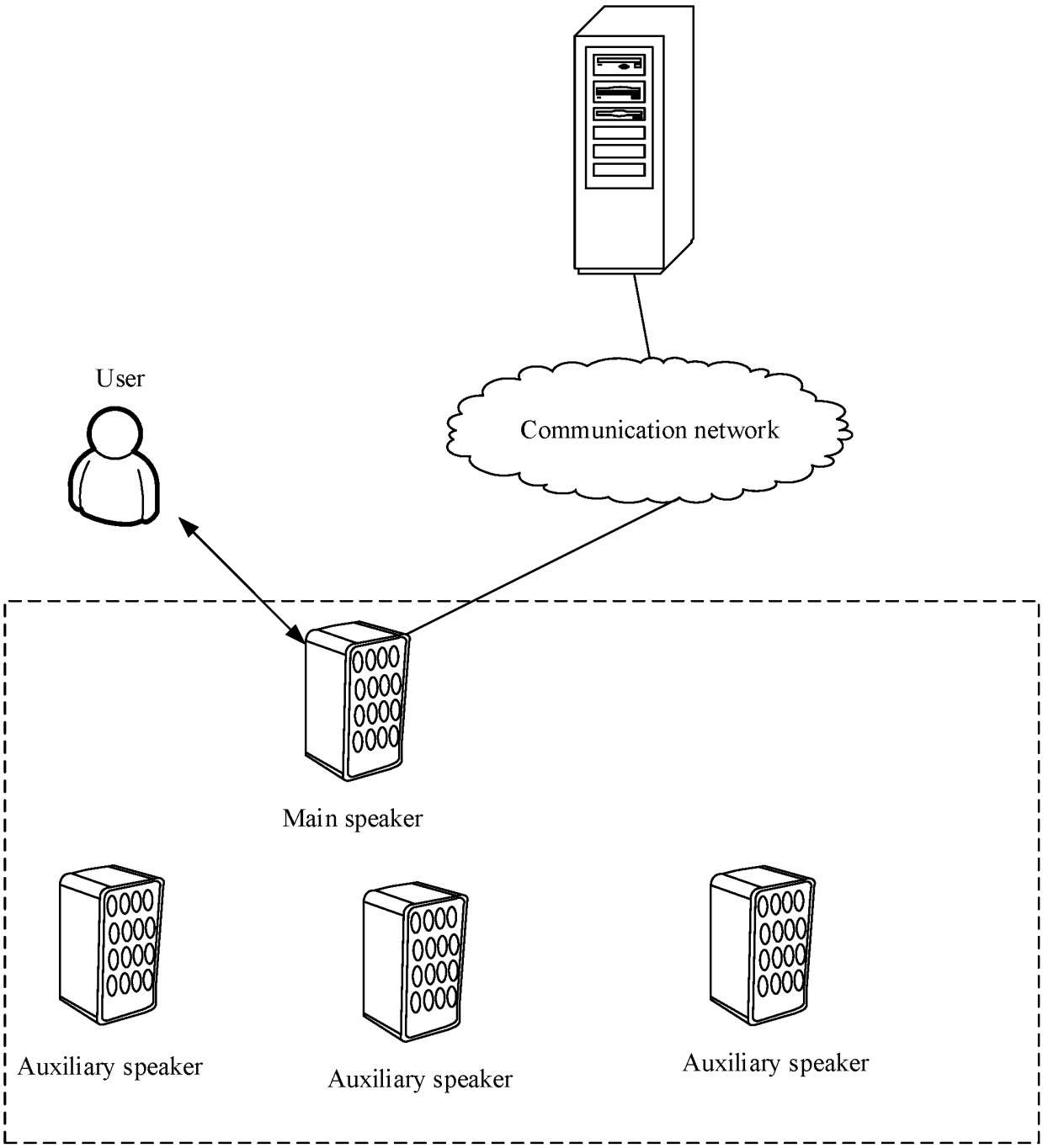
User
Communication network
Main speaker
Auxiliary speaker
Auxiliary speaker
Auxiliary speaker

FIG. 16

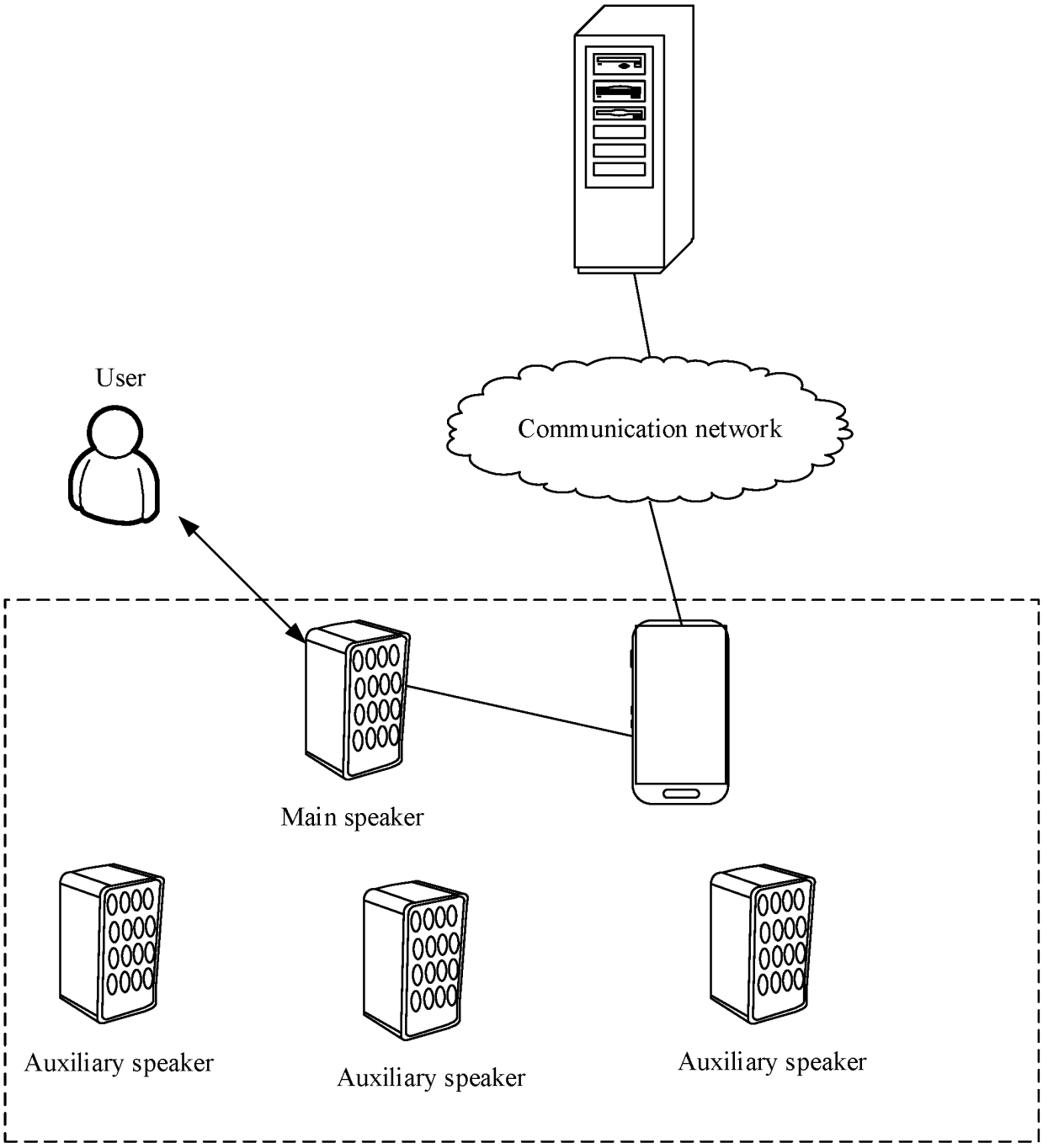
User
Communication network
Main speaker
Auxiliary speaker
Auxiliary speaker
Auxiliary speaker

FIG. 17

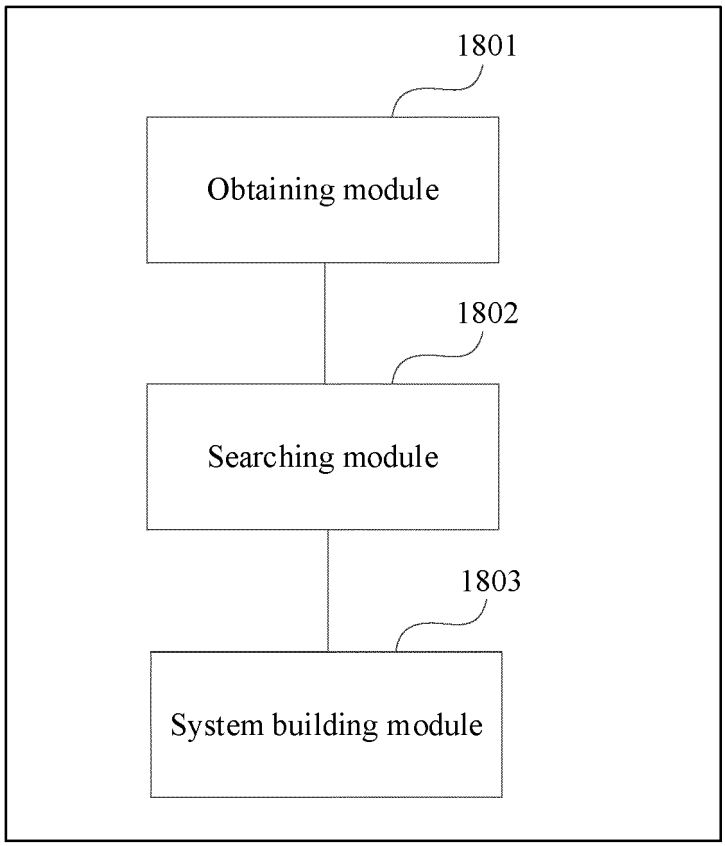
1801
Obtaining module
1802
Searching module
1803
System building module

FIG. 18

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2022/119134

A. CLASSIFICATION OF SUBJECT MATTER

H04R 5/02(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04R5/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, ENTXTC, CNKI: 音箱, 音响, 扬声器, 外放, 回放, 音频, 声音, 输出, 碰, 撞, 摇, 晃, 转动, 翻转, 敲, 运动, 蓝牙, 无线, 近场, 配对, 连接, 立体声, tws, 环绕, H04R5/+, H04W76/+; VEN, ENTXT, USTXT, EPTXT, WOTXT: (sound or musical or speaker) 2w box, play w out, play w back, output, hit, crash+, bump, shake, rotate, knock, pair+, connect+, stereo, tws, surround, H04R5/+, H04W76/+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	CN 115129368 A (ALIBABA SINGAPORE HOLDINGS LIMITED) 30 September 2022 (2022-09-30) description, paragraphs 0110-0115	1, 27
X	CN 113377325 A (SHENZHEN YAYU TECHNOLOGY CO., LTD.) 10 September 2021 (2021-09-10) description, paragraphs 0058-0061 and 0079, and figures 2-4	1-27
A	CN 108684023 A (JIANG PEILING) 19 October 2018 (2018-10-19) entire document	1-27
A	CN 207427432 U (SHENZHEN ESON TECHNOLOGY CO., LTD.) 29 May 2018 (2018-05-29) entire document	1-27
A	US 2016165343 A1 (JAZZ HIPSTER CORP.) 09 June 2016 (2016-06-09) entire document	1-27

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 October 2022	**14 December 2022**
Name and mailing address of the ISA/CN **China National Intellectual Property Administration (ISA/CN)** **No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China** Facsimile No. **(86-10)62019451**	Authorized officer Telephone No.

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2022/119134

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 115129368 A	30 September 2022	None	
CN 113377325 A	10 September 2021	None	
CN 108684023 A	19 October 2018	None	
CN 207427432 U	29 May 2018	None	
US 2016165343 A1	09 June 2016	None	

Form PCT/ISA/210 (patent family annex) (January 2015)

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- CN 202111186340 **[0001]**