(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 July 2015 (23.07.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/106688 A1

(51) International Patent Classification:
G06F 3/0487 (2013.01)

(21) International Application Number:
PCT/CN2015/070706

(22) International Filing Date:
14 January 2015 (14.01.2015)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
201410015324.2 14 January 2014 (14.01.2014) CN

(71) Applicant: **TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED** [CN/CN]; Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen, Guangdong 518000 (CN).

(72) Inventor: **LIN, Yu**; Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen, Guangdong 518000 (CN).

(74) Agent: **CENFO INTELLECTUAL PROPERTY AGENCY**; Room 210-212,2/F,Building "Golden"(Block 11,Industrial Village of Former Nantou Cheng), Design Industrial Park,No 3838,Nanshan Road, Nanshan District, Shenzhen, Guangdong 518052 (CN).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) Title: METHOD AND APPARATUS FOR VOICE ACCESS TO MULTIMEDIA INTERACTIVE WEBSITE

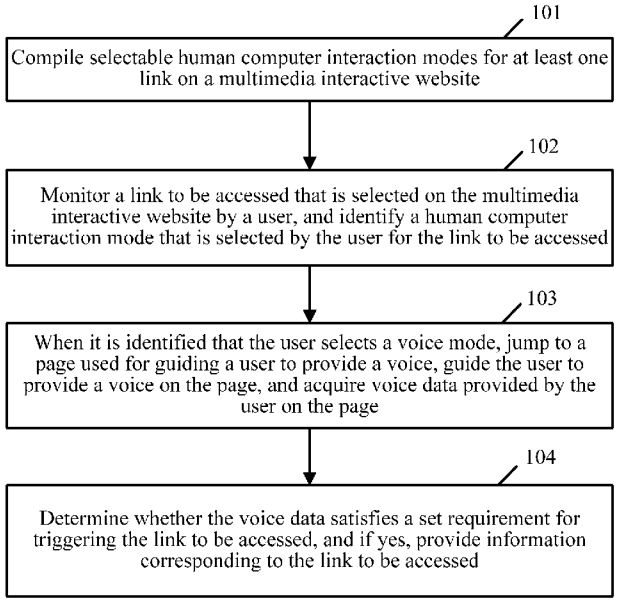

FIG. 1

(57) Abstract: A method and an apparatus for voice access to a multimedia interactive website. The method includes: compiling selectable human computer interaction modes for at least one link on the multimedia interactive website (101), where the selectable human computer interaction modes at least include: a voice mode and a non-voice mode; monitoring a link to be accessed that is selected on the multimedia interactive website by a user, and identifying a human computer interaction mode that is selected by the user for the link to be accessed (102); when it is identified that the user selects the voice mode, jumping to a page used for guiding a user to provide a voice, guiding the user to provide a voice on the page, and acquiring voice data provided on the page by the user (103); and determining whether the voice data satisfies a set requirement for triggering the link to be accessed, and if yes, providing information corresponding to the link to be accessed (104). By means of the method, voice interaction between a user and a multimedia website can be implemented.

WO 2015/106688 A1

Title of Invention: METHOD AND APPARATUS FOR VOICE ACCESS TO MULTIMEDIA INTERACTIVE WEBSITE

[1] PRIORITY CLAIM AND RELATED APPLICATION

[2] This application claims priority to Chinese Patent Application No. 201410015324.2, entitled "METHOD AND APPARATUS FOR VOICE ACCESS TO MULTIMEDIA INTERACTIVE WEBSITE" filed on January 14, 2014, which is incorporated by reference in its entirety.

[3] FIELD OF THE INVENTION

[4] The present application relates to network technologies, and in particular, to a method and an apparatus for voice access to a multimedia interactive website.

[5] BACKGROUND OF THE INVENTION

[6] At present, with the continuous development of website technologies, a great number of multimedia websites have emerged. Among these multimedia websites, a type of websites having a characteristic of strong human computer interaction is referred to as a multimedia interactive website in short.

[7] A user interacts with a multimedia interactive website by using many modes, which include, but are not limited to, the following types:

[8] 1. Click trigger mode: For example, a user clicks a link on a multimedia interactive website by using a mouse to enter a webpage corresponding to the link.

[9] 2. Scroll wheel trigger mode: For example, a user triggers a link on a multimedia interactive website by using a scroll wheel in a mouse to enter a webpage corresponding to the link.

[10] 3. Keyboard trigger mode: For example, a user triggers a link on a multimedia interactive website by using a keyboard to enter a webpage corresponding to the link.

[11] 4. Touch control mode: For example, a user selects a link on a multimedia interactive website by touch control to enter a webpage corresponding to the link.

[12] In the foregoing modes of interaction between a user and a multimedia interactive website, regardless of which mode is used, voice access of a user to a multimedia interactive website cannot be implemented.

[13] Therefore, a method for implementing voice access of a user to a multimedia interactive website is a technical problem that needs to be solved urgently at present.

[14] SUMMARY

[15] The present application provides a method and an apparatus for voice access to a multimedia interactive website, so as to implement voice access of a user to a multimedia interactive website.

[16] Technical solutions provided by the present application include:

[17] A method for voice access to a multimedia interactive website, where the multimedia interactive website has a human computer interaction function, includes:

[18] compiling selectable human computer interaction modes for at least one link on the multimedia interactive website, where the selectable human computer interaction modes at least include: a voice mode and a non-voice mode;

[19] monitoring a link to be accessed that is selected on the multimedia interactive website by a user, and identifying a human computer interaction mode that is selected by the user for the link to be accessed;

[20] when it is identified that the user selects the voice mode, jumping to a page used for guiding a user to provide a voice, guiding the user to provide a voice on the page, and acquiring voice data provided on the page by the user; and

[21] determining whether the voice data satisfies a set requirement for triggering the link to be accessed, and if yes, providing information corresponding to the link to be accessed.

[22] An apparatus for voice access to a multimedia interactive website, where the multimedia interactive website has a human computer interaction function, and the apparatus is applied to a client of a multimedia website, includes:

[23] a display unit, configured to display the multimedia interactive website, where at least one link on the multimedia interactive website is compiled with selectable human computer interaction modes, and the selectable human computer interaction modes at least include: a voice mode and a non-voice mode;

[24] a monitoring unit, configured to monitor a link to be accessed that is selected on the multimedia interactive website by a user;

[25] an identification unit, configured to identify a human computer interaction mode that is selected by the user for the link to be accessed;

[26] a guiding unit, configured to: when the identification unit identifies that the user selects the voice mode, jump to a page used for guiding a user to provide a voice, guide the user to provide a voice on the page, and acquire voice data provided on the page by the user; and

[27] a determination unit, configured to determine whether the voice data satisfies a set requirement for triggering the link to be accessed, and if yes, provide information corresponding to the link to be accessed.

[28] It can be seen from the above technical solutions that in the present disclosure, selectable human computer interaction modes are compiled by using a link on a multimedia interactive website, a link to be accessed that is selected on the multimedia interactive website by a user is monitored, and a human computer interaction mode that is selected by the user for the link to be accessed is identified; when it is identified

that the user selects a voice mode, a page for guiding a user to provide a voice is jumped to; the user is guided to provide a voice on the page, and voice data provided on the page by the user is acquired; and it is determined whether the voice data satisfies a set requirement for triggering the link to be accessed, and if yes, information corresponding to the link to be accessed is provided, which can implement interaction between a user and a multimedia interactive website by using a voice mode in place of a conventional mode such as a mouse and a keyboard, and improve the efficiency of accessing the multimedia interactive website by the user.

[29] BRIEF DESCRIPTION OF THE DRAWINGS

[30] FIG. 1 is a flowchart of a method according to an embodiment of the present invention;

[31] FIG. 2 is a schematic diagram of a link on a multimedia website according to an embodiment of the present invention;

[32] FIG. 3 is a schematic diagram of a page including a voice mode according to an embodiment of the present invention;

[33] FIG. 4 is a flowchart of the implementation of step 104 according to an embodiment of the present invention;

[34] FIG. 5 is a schematic diagram of a voice spectrum according to an embodiment of the present invention;

[35] FIG. 6 is a schematic diagram of another voice spectrum according to an embodiment of the present invention; and

[36] FIG. 7 is a structural diagram of an apparatus according to an embodiment of the present invention.

[37] DESCRIPTION OF EMBODIMENTS

[38] To make the objectives, technical solutions, and advantages of the present disclosure more clear, the present disclosure is further described in detail below with reference to the accompanying drawings and specific embodiments.

[39] The present application provides a method for voice access to a multimedia interactive website. Here, the multimedia interactive website has a human computer interaction function.

[40] Referring to FIG. 1, FIG. 1 is a flowchart of a method provided by the present disclosure. As shown in FIG. 1, the procedure is applicable to a client, which includes the following steps:

[41] Step 101: Compile selectable human computer interaction modes for at least one link on the multimedia interactive website.

[42] Here, the multimedia interactive website is generally a website having a characteristic of strong human computer interaction. In the present disclosure, a few improvements are made to an existing multimedia interactive website, that is, selectable

human computer interaction modes are compiled for all or some of links on the multimedia interactive website.

[43] As an embodiment of the present invention, the selectable human computer interaction modes at least include: a voice mode and a non-voice mode.

[44] As an embodiment of the present invention, the voice mode is implemented by using a prompt button used for indicating a microphone, whereas the non-voice mode is implemented by using a prompt button used for indicating no microphone. FIG. 2 shows selectable human computer interaction modes compiled on a link, for example, "I want to be the hero", on a multimedia interactive website.

[45] Step 102: Monitor a link to be accessed that is selected on the multimedia interactive website by a user, and identify a human computer interaction mode that is selected by the user for the link to be accessed.

[46] When the user enters the multimedia interactive website, the user selects, on the multimedia interactive website according to an interest of the user, the link to be accessed.

[47] After the user finishing selecting, on the multimedia interactive website, the link to be accessed, based on the description in step 101, if the selectable human computer interaction modes are compiled for the link to be accessed, the user selects, from the selectable human computer interaction modes, an interaction mode to be used by the user, for example, selects the voice mode or selects the non-voice mode.

[48] Step 103: When it is identified that the user selects the voice mode, jump to a page used for guiding a user to provide a voice, guide the user to provide a voice on the page, and acquire voice data provided on the page by the user.

[49] As an embodiment of the present invention, in step 103, the jumping to a page used for guiding a user to provide a voice includes:

[50] displaying, in the form of an animation, a new page that is associated with the link to be accessed and is used for guiding a user to provide a voice, where the page includes guiding information for guiding a user to provide a voice. As shown in FIG. 3, FIG. 3 shows, in the form of an animation, a new page that is associated with a link to be accessed and is used for guiding a user to provide a voice. In the new page, "Please speak into the microphone:" is the guiding information and guides the user to provide voice data into the microphone. In addition, in FIG. 3, a progress box is further included and used for indicating a state of the user. When the user provides voice data into the microphone, a progress in this progress box changes dynamically, or otherwise, when the user stops providing voice data into the microphone, the progress in this progress box remains unchanged.

[51] On the premise that a page that is jumped to includes the guiding information, in step 103, the guiding the user to provide a voice on the page may specifically be:

[52] guiding, by using the guiding information in the page, the user to provide the voice on the page.

[53] It should be noted that step 103 is performed on the premise that it is identified that the user selects the voice mode. However, if in the present disclosure, it is identified that the user selects the non-voice mode, the following steps may be performed:

[54] prompting the user to trigger the link based on the set non-voice mode.

[55] The non-voice mode is similar to an existing human computer interaction modes, and includes, but is not limited to, any one or any combination of the following modes:

[56] a mouse;

[57] a keyboard; and

[58] touch control.

[59] Step 104: Determine whether the voice data satisfies a set requirement for triggering the link to be accessed, and if yes, provide information corresponding to the link to be accessed.

[60] Preferably, as a preferred embodiment of the present invention, in step 104, when it is determined that the voice data does not satisfy the set requirement for triggering the link to be accessed, the user may be guided to provide a voice, and voice data provided on the page by the user is acquired, that is, the process returns to step 103.

[61] Here, the procedure shown in FIG. 1 is completed.

[62] It can be seen from the procedure shown in FIG. 1 that in the present disclosure, interaction between a user and a multimedia interactive website is implemented by using a voice mode in place of a conventional mode such as a mouse and a keyboard, which can improve the efficiency of accessing the multimedia interactive website by the user.

[63] Preferably, as an embodiment of the present invention, in the foregoing description, the selectable human computer interaction modes compiled for the link on the multimedia interactive website may be compiled by using a Flash tool.

[64] Based on that the selectable human computer interaction modes compiled for the link on the multimedia interactive website are compiled by using the Flash tool, in step 103 in the foregoing, the acquiring voice data provided on the page by the user may be:

[65] acquiring, by using a sound application interface SoundMixer API of the scripting language ActionScript in Flash, the voice data provided on the page by the user.

[66] In addition, in the present disclosure, as an embodiment of the present invention, the determining in step 104 may be implemented by using the procedure shown in FIG. 4.

[67] Referring to FIG. 4, FIG. 4 is a flowchart of the implementation of step 104 according to an embodiment of the present invention. As shown in FIG. 4, the procedure may include the following steps:

[68] Step 401: Extract a voice spectrum from the acquired voice data.

[69] In the present disclosure, the acquired voice data is binary data. The extraction of the

voice spectrum from the voice data formed of the binary data can be implemented by performing a fast Fourier transform (fft) on hidden Markov model (HMM) state-associated data.

[70] The extracted voice spectrum is formed of a left channel and a right channel. FIG. 5 shows a voice spectrum when fftmode is true and FIG. 6 shows another voice spectrum when fftmode is false.

[71] It can be seen from the voice spectrum shown in FIG. 5 or FIG. 6 that the voice spectrum may be formed of a left channel and a right channel.

[72] Step 402: Identify, according to pitch fluctuations of the voice in a left channel or in a right channel in the voice spectrum, the quantity of characters associated with the voice spectrum.

[73] In the present disclosure, for the quantities of characters identified from different channels in the voice spectrum, the results are same, and therefore, a single channel, for example, the left channel or the right channel can be selected to identify the quantity of characters associated with the voice spectrum.

[74] Specifically, in a voice spectrum, two voices of different pitch fluctuations are two different Chinese characters. Based on this, the quantity of characters associated with the voice spectrum is identified, according to a principle that two voices of different pitch fluctuations are two different Chinese characters, from the pitch fluctuations of the voice in the left channel or in the right channel in the voice spectrum.

[75] Step 403: Determine whether the quantity of characters identified in step 402 is the quantity of characters required by the link to be accessed, and if yes, determine that the voice data satisfies the set requirement for triggering the link to be accessed, or if not, determine that the voice data does not satisfy the set requirement for triggering the link to be accessed.

[76] If the quantity of characters required by the link to be accessed is 3, it is determined in step 403 whether the quantity of characters identified in step 402 is 3, and if yes, it is determined that the voice data satisfies the set requirement for triggering the link to be accessed, or if not, it is determined that the voice data does not satisfy the set requirement for triggering the link to be accessed.

[77] Therefore, the procedure shown in FIG. 4 is completed.

[78] It can be seen that, the determining whether the voice data satisfies a set requirement for triggering the link to be accessed in step 104 can be implemented by using the procedure shown in FIG. 4. However, it should be noted that the procedure shown in FIG. 4 is only an example of determining whether the voice data satisfies the set requirement for triggering the link to be accessed and is not intended to limit the present disclosure.

[79] A person skilled in the art may further use other modes to determine whether the

voice data satisfies the set requirement for triggering the link to be accessed. Two examples below are merely used for illustration:

[80] Example 1. Whether set voice data exists in the acquired voice data is identified based on an existing voice recognition system. If the set voice data is "I do", whether "I do" exists in the acquired voice data is identified based on the existing voice recognition system, and if yes, it is determined that the voice data satisfies the set requirement for triggering the link to be accessed, or if not, it is determined that the voice data does not satisfy the set requirement for triggering the link to be accessed.

[81] Example 2. Example 2 is mainly a combination of the procedure shown in FIG. 1 and the procedure shown in FIG. 4, that is, whether set voice data exists in the acquired voice data is first identified based on the existing voice recognition system, and the procedure show in FIG. 4 is performed only when it is identified that the set voice data exists in the acquired voice data, or otherwise, it is determined that the voice data does not satisfy the set requirement for triggering the link to be accessed. According to Example 2, the link to be accessed can be more accurately triggered according to a user requirement.

[82] Therefore, the description of the method provided by the present disclosure is completed.

[83] The apparatus provided by the present disclosure is described below:

[84] Referring to FIG. 7, FIG. 7 is a structural diagram of an apparatus according to an embodiment of the present invention. The apparatus is applied to a client of a multimedia interactive website, where the multimedia interactive website has a human computer interaction function. As shown in FIG. 7, the apparatus may include:

[85] a display unit, configured to display the multimedia interactive website, where at least one link on the multimedia interactive website is compiled with selectable human computer interaction modes, and the selectable human computer interaction modes at least include: a voice mode and a non-voice mode;

[86] a monitoring unit, configured to monitor a link to be accessed that is selected on the multimedia interactive website by a user;

[87] an identification unit, configured to identify a human computer interaction mode that is selected by the user for the link to be accessed;

[88] a guiding unit, configured to: when the identification unit identifies that the user selects the voice mode, jump to a page used for guiding a user to provide a voice, guide the user to provide a voice on the page, and acquire voice data provided on the page by the user; and

[89] a determination unit, configured to determine whether the voice data satisfies a set requirement for triggering the link to be accessed, and if yes, provide information corresponding to the link to be accessed.

[90] Preferably, as an embodiment of the present invention, the jumping to a page used for guiding a user to provide a voice includes:

[91] displaying, in the form of an animation, a new page that is associated with the link to be accessed and is used for guiding a user to provide a voice, where the page includes guiding information for guiding a user to provide a voice.

[92] Based on this, the guiding the user to provide a voice on the page includes:

[93] guiding, by using the guiding information in the new page, the user to provide the voice on the page.

[94] Preferably, in the present disclosure, the determining whether the voice data satisfies a set requirement for triggering the link to be accessed includes:

[95] extracting a voice spectrum from the voice data, where the voice spectrum is formed of a left channel and a right channel;

[96] identifying, according to pitch fluctuations of the voice in the left channel or in the right channel in the voice spectrum, the quantity of characters associated with the voice spectrum; and

[97] determining whether the quantity of characters is the quantity of characters required by the link to be accessed,

[98] and if yes, determining that the voice data satisfies the set requirement for triggering the link to be accessed, or

[99] if not, determining that the voice data does not satisfy the set requirement for triggering the link to be accessed.

[100] Preferably, in the present disclosure, the voice mode is implemented by using a prompt button used for indicating a microphone; and

[101] the non-voice mode is implemented by using a prompt button used for indicating no microphone.

[102] Preferably, in the present disclosure, the link on the multimedia interactive website is compiled by using Flash; and

[103] the voice data is acquired by using a sound application interface SoundMixer API of the scripting language ActionScript in Flash.

[104] Therefore, the description of the apparatus shown in FIG. 7 is completed.

[105] It can be seen from the above technical solutions that in the present disclosure, selectable human computer interaction modes are compiled for a link on a multimedia interactive website, a link to be accessed that is selected on the multimedia interactive website by a user is monitored; a human computer interaction mode that is selected by the user for the link to be accessed is identified; when it is identified that the user selects a voice mode, a page for guiding a user to provide a voice is jumped to; the user is guided to provide a voice on the page, and voice data provided on the page by the user is acquired; and it is determined whether the voice data satisfies a set requirement

for triggering the link to be accessed, and if yes, information corresponding to the link to be accessed is provided, which can implement interaction between a user and a multimedia interactive website by using a voice mode in place of a conventional mode such as a mouse and a keyboard, and improve the efficiency of accessing the multimedia interactive website by the user.

[106] The foregoing descriptions are merely preferred embodiments of the present invention, but are not intended to limit the present disclosure. Any modification, equivalent replacement, or improvement made within the spirit and principle of the present disclosure shall fall within the protection scope of the present disclosure.

Claims

[Claim 1] A method for voice access to a multimedia interactive website, wherein the multimedia interactive website has a human computer interaction function, and the method comprises:
compiling selectable human computer interaction modes for at least one link on the multimedia interactive website, wherein the selectable human computer interaction modes at least comprise: a voice mode and a non-voice mode;
monitoring a link to be accessed that is selected on the multimedia interactive website by a user, and identifying a human computer interaction mode that is selected by the user for the link to be accessed;
when it is identified that the user selects the voice mode, jumping to a page used for guiding a user to provide a voice, guiding the user to provide a voice on the page, and acquiring voice data provided on the page by the user; and
determining whether the voice data satisfies a set requirement for triggering the link to be accessed, and if yes, providing information corresponding to the link to be accessed.

[Claim 2] The method according to claim 1, wherein the jumping to a page used for guiding a user to provide a voice comprises:
displaying, in the form of an animation, a new page that is associated with the link to be accessed and is used for guiding a user to provide a voice, wherein the page comprises guiding information for guiding a user to provide a voice; and
the guiding the user to provide a voice on the page comprises:
guiding, by using the guiding information in the page, the user to provide the voice on the page.

[Claim 3] The method according to claim 1, wherein
the determining whether the voice data satisfies a set requirement for triggering the link to be accessed comprises:
extracting a voice spectrum from the voice data, wherein the voice spectrum is formed of a left channel and a right channel;
identifying, according to pitch fluctuations of the voice in the left channel or in the right channel in the voice spectrum, the quantity of characters associated with the voice spectrum; and
determining whether the quantity of characters is the quantity of characters required by the link to be accessed,

and if yes, determining that the voice data satisfies the set requirement for triggering the link to be accessed, or
if not, determining that the voice data does not satisfy the set requirement for triggering the link to be accessed.

[Claim 4] The method according to any one of claims 1 to 3, wherein,
the voice mode is implemented by using a prompt button used for indicating a microphone;
the non-voice mode is implemented by using a prompt button used for indicating no microphone.

[Claim 5] The method according to any one of claims 1 to 3, wherein the link on the multimedia interactive website is compiled by using Flash; and
the voice data is acquired by using a sound application interface SoundMixer API of the scripting language ActionScript in Flash.

[Claim 6] An apparatus for voice access to a multimedia interactive website, wherein the multimedia interactive website has a human computer interaction function and the apparatus is applied to a client of a multimedia website, comprising:
a display unit, configured to display the multimedia interactive website, wherein at least one link on the multimedia interactive website is compiled with selectable human computer interaction modes, and the selectable human computer interaction modes at least comprise: a voice mode and a non-voice mode;
a monitoring unit, configured to monitor a link to be accessed that is selected on the multimedia interactive website by a user;
an identification unit, configured to identify a human computer interaction mode that is selected by the user for the link to be accessed;
a guiding unit, configured to: when the identification unit identifies that the user selects the voice mode, jump to a page used for guiding a user to provide a voice, guide the user to provide a voice on the page, and acquire voice data provided on the page by the user; and
a determination unit, configured to determine whether the voice data satisfies a set requirement for triggering the link to be accessed, and if yes, provide information corresponding to the link to be accessed.

[Claim 7] The apparatus according to claim 6, wherein the jumping to a page used for guiding a user to provide a voice comprises:
displaying, in the form of an animation, a new page that is associated with the link to be accessed and is used for guiding a user to provide a voice, wherein the page comprises guiding information for guiding a

user to provide a voice; and
the guiding the user to provide a voice on the page comprises:
guiding, by using the guiding information in the new page, the user to provide the voice on the page.

[Claim 8] The apparatus according to claim 6, wherein the determining whether the voice data satisfies a set requirement for triggering the link to be accessed comprises:
extracting a voice spectrum from the voice data, wherein the voice spectrum is formed of a left channel and a right channel;
identifying, according to pitch fluctuations of the voice in the left channel or in the right channel in the voice spectrum, the quantity of characters associated with the voice spectrum; and
determining whether the quantity of characters is the quantity of characters required by the link to be accessed,
and if yes, determining that the voice data satisfies the set requirement for triggering the link to be accessed, or
if not, determining that the voice data does not satisfy the set requirement for triggering the link to be accessed.

[Claim 9] The apparatus according to any one of claims 6 to 8, wherein
the voice mode is implemented by using a prompt button used for indicating a microphone;
the non-voice mode is implemented by using a prompt button used for indicating no microphone.

[Claim 10] The apparatus according to any one of claims 6 to 8, wherein the link on the multimedia interactive website is compiled by using Flash; and
the voice data is acquired by using a sound application interface SoundMixer API of the scripting language ActionScript in Flash.

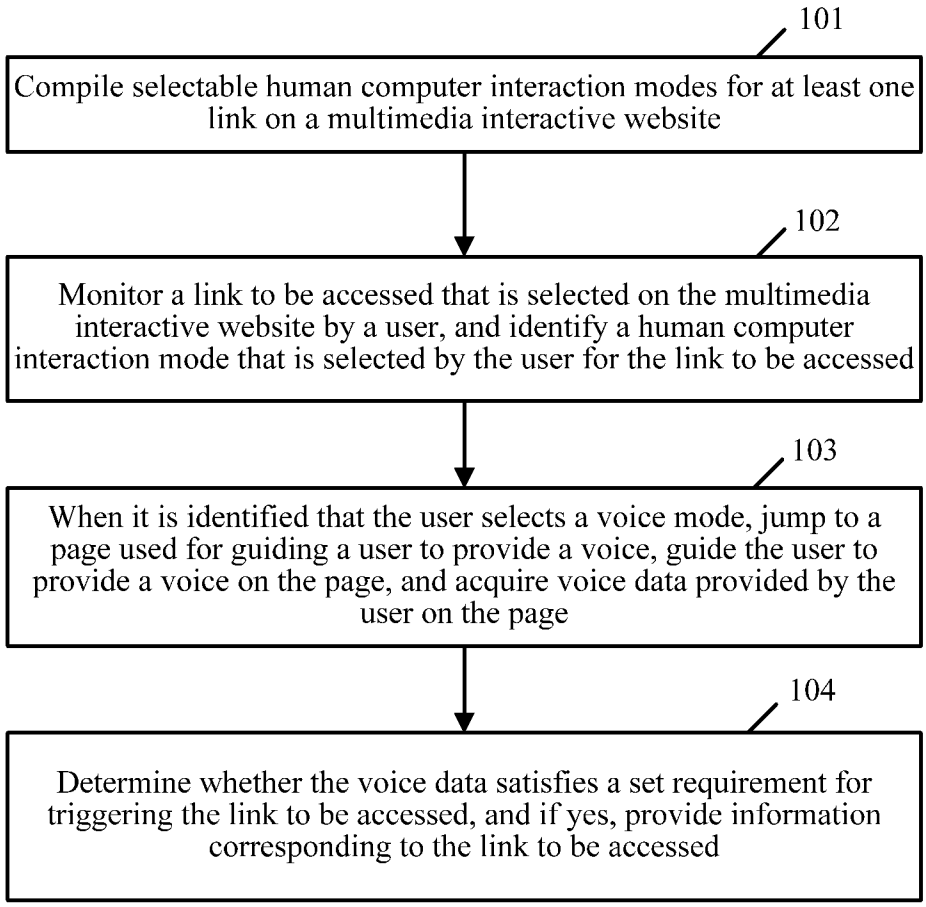
101
Compile selectable human computer interaction modes for at least one link on a multimedia interactive website
102
Monitor a link to be accessed that is selected on the multimedia interactive website by a user, and identify a human computer interaction mode that is selected by the user for the link to be accessed
103
When it is identified that the user selects a voice mode, jump to a page used for guiding a user to provide a voice, guide the user to provide a voice on the page, and acquire voice data provided by the user on the page
104
Determine whether the voice data satisfies a set requirement for triggering the link to be accessed, and if yes, provide information corresponding to the link to be accessed

FIG. 1

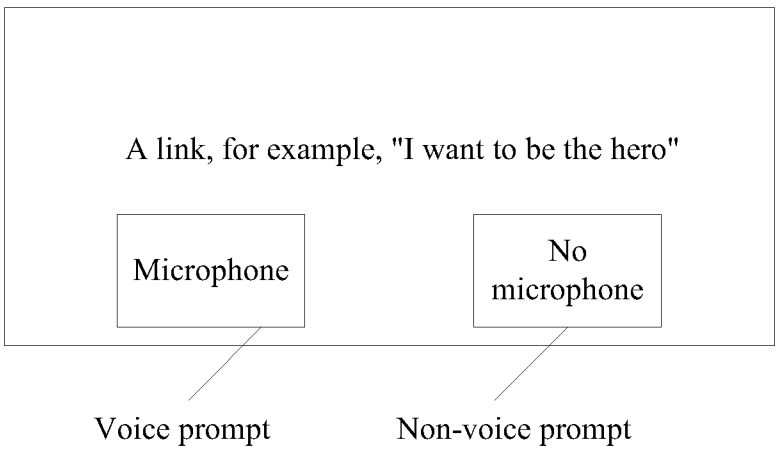
A link, for example, "I want to be the hero"
Microphone
No microphone
Voice prompt
Non-voice prompt

FIG. 2

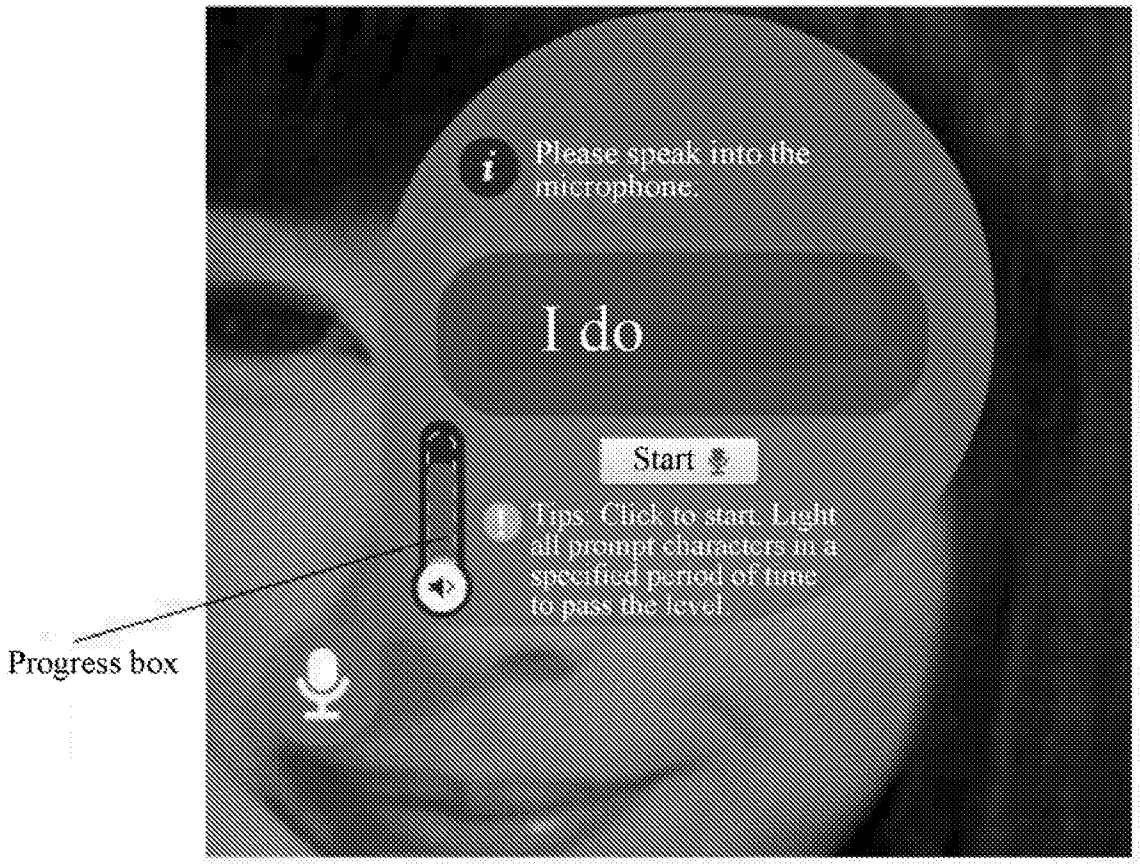
Please speak into the microphone
I do
Start
Tips: Click to start. Light all prompt characters in a specified period of time to pass the level.
Progress box

FIG. 3

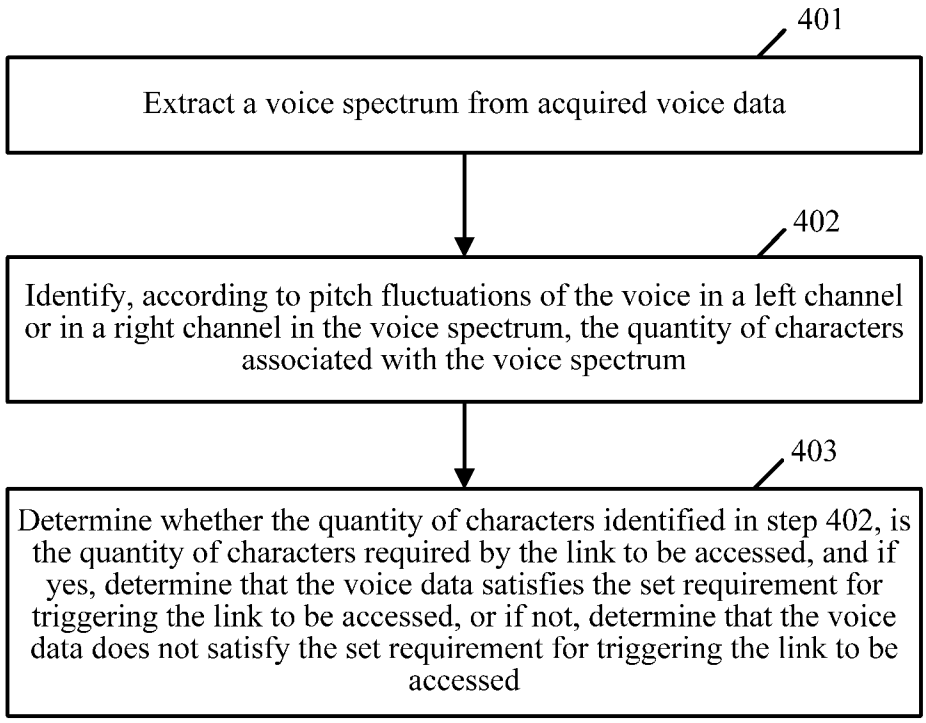
401
Extract a voice spectrum from acquired voice data
402
Identify, according to pitch fluctuations of the voice in a left channel or in a right channel in the voice spectrum, the quantity of characters associated with the voice spectrum
403
Determine whether the quantity of characters identified in step 402, is the quantity of characters required by the link to be accessed, and if yes, determine that the voice data satisfies the set requirement for triggering the link to be accessed, or if not, determine that the voice data does not satisfy the set requirement for triggering the link to be accessed

FIG. 4

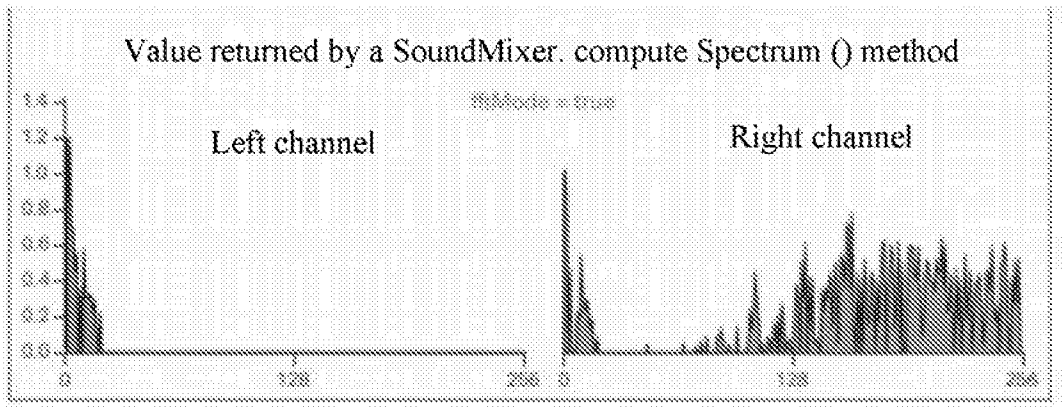
Value returned by a SoundMixer. compute Spectrum () method
Left channel
Right channel

FIG. 5

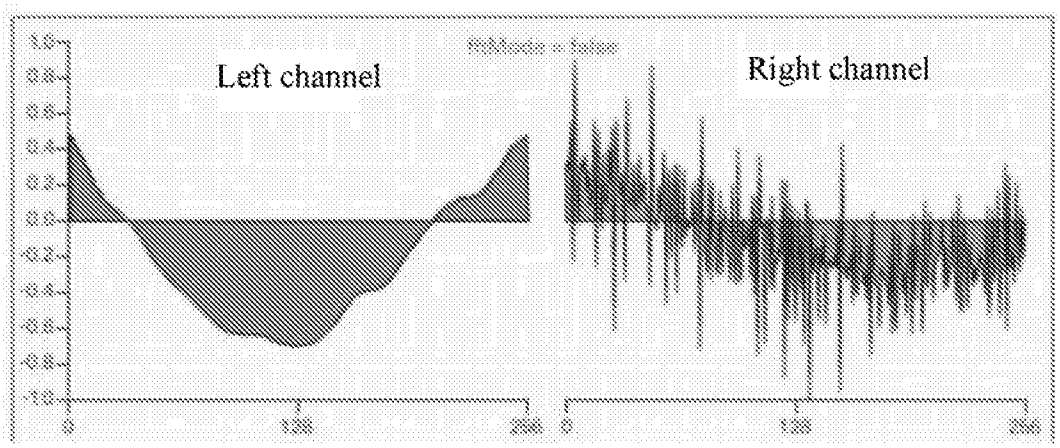
Left channel
Right channel

FIG. 6

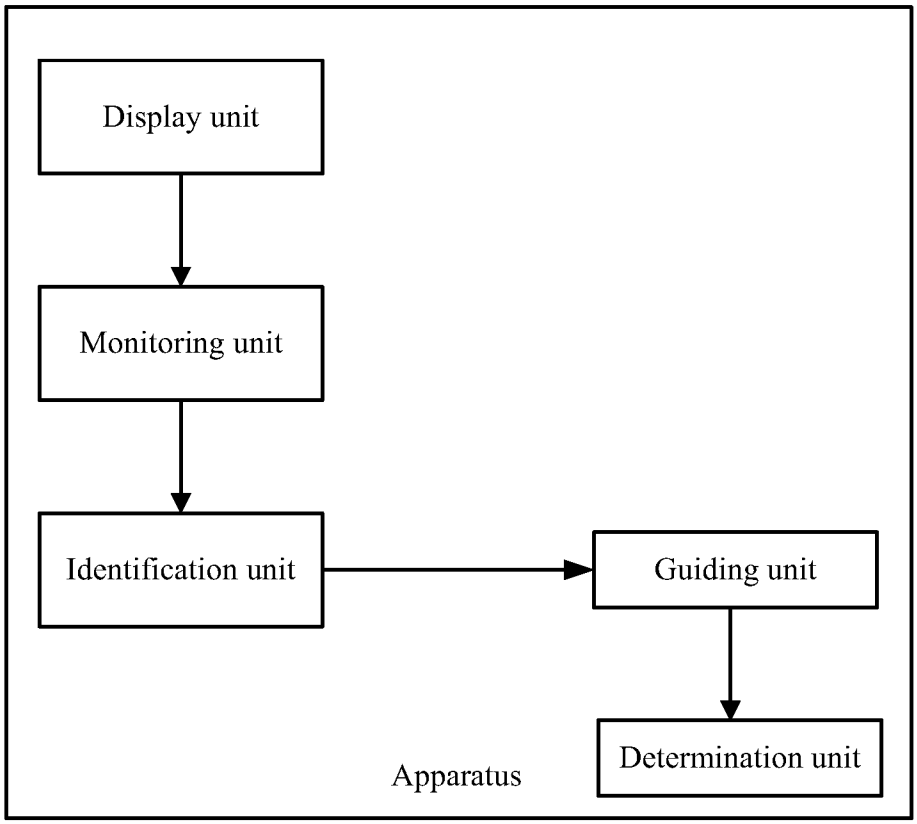
Display unit
Monitoring unit
Identification unit
Guiding unit
Determination unit
Apparatus

FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2015/070706

A. CLASSIFICATION OF SUBJECT MATTER

G06F 3/0487 (2013.01) 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)

G06F

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)

CNPAT,CNKI,WPI,EPODOC,GOOGLE voice, web, webpage, hyperlink, link, recognition, wave form, wave shape, character, text, number, quantity

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2013282381 A1 (NUANCE COMMUNICATIONS INC.) 24 October 2013 (2013-10-24) description, paragraphs [0022], [0025]-[0027], [0030], [0031], figure 1	1-2,4-7,9-10
X	CN 102567321 A (XU, DEWU) 11 July 2012 (2012-07-11) description, paragraphs [0011], [0012], [0014], [0015]	1-2,4-7,9-10
PX	CN 104123085 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 29 October 2014 (2014-10-29) claims 1-10	1-10
A	CN 101667188 A (LIU, XUEYING) 10 March 2010 (2010-03-10) the whole document	1-10
A	CN 102520792 A (JIANGSU QIYIDIAN NETWORK CO., LTD.) 27 June 2012 (2012-06-27) the whole document	1-10

☐ 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
26 March 2015	**17 April 2015**

Name and mailing address of the ISA/CN	Authorized officer
STATE INTELLECTUAL PROPERTY OFFICE OF THE P.R.CHINA(ISA/CN) **6,Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088, China**	**DONG,Libo**
Facsimile No. **(86-10)62019451**	Telephone No. **(86-10)61648110**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2015/070706

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
US 2013282381 A1	24 October 2013	EP 2653964 A2	23 October 2013
		CN 103377028 A	30 October 2013
CN 102567321 A	11 July 2012	None	
CN 104123085 A	29 October 2014	None	
CN 101667188 A	10 March 2010	None	
CN 102520792 A	27 June 2012	None	