WHAT IS CLAIMED IS:

1. A method for acquiring an online audio-visual content, separating tracks, integrating artificial intelligence and outputting an AI-generated medium, comprising the following steps:

a device server (11) of an audio-visual playback device (1) sending an audio-visual search request to an application programming interface (API) server (A1) of a streaming platform (A) by an API technique according to a keyword; after the API server (A1) performs a search according to the audio-visual search request to obtain a search result corresponding to the keyword, the device server (11) obtaining a plurality of link codes of the search result from the API server (A1) by the API technique, wherein each of the link codes corresponds to a different online audio-visual content (B, B'), and the device server (11) first excludes any of the link codes corresponding to the online audio-visual content in an invalid state;

the device server (11) obtaining one of the link codes, and the device server, or a generative artificial intelligence server (3) in signal communication with the device server (11), parsing the one of the link codes to acquire the corresponding online audio-visual content from the API server without downloading the online audio-visual content;

the device server (11) or the generative artificial intelligence server (3) separating a plurality of track contents (C1, C2, C3, C1', C2', C3', C4') from the online audio-visual content (B, B') through a sampling module (2), wherein the sampling module (2) includes a front-end neural network, the front-end neural network samples and separates a plurality of track audios from the online audio-visual content according to a frequency characteristic, and the track contents includes the track audios;

the generative artificial intelligence server (3) selecting part or all of the track contents (C1, C2, C3, C1', C2', C3', C4'), the selected track contents (C1, C2, C3, C1', C2', C3', C4') being aligned and positioned on a timeline, wherein a playback start time common to the track contents in a sampling process is marked as a start point of each of the track contents on the timeline, and the start points of the selected track contents are aligned with one another to complete the positioning;

the generative artificial intelligence server (3) generating at least one AI-generated content (D1, D2, D3, D1', D2', D3', D4', D5') according to the selected track contents (C1, C2, C3, C1', C2', C3', C4'), wherein the at least one AI-generated content includes at least one of an AI-generated image, an AI-generated animation and an AI-generated text, each of the at least one AI-generated content is an independently selectable object, and the at least one AI-generated content (D1, D2, D3, D1', D2', D3', D4', D5') is aligned with the track contents (C1, C2, C3, C1', C2', C3', C4') and positioned on the timeline;

the generative artificial intelligence server (3) mixing the selected track contents (C1, C2, C3, C1', C2', C3', C4') with the selected at least one AI-generated content (D1, D2, D3, D1', D2', D3', D4', D5'), outputting an AI-generated medium (E, E', E2') and transmitting the AI-generated medium back to the device server (11); and

the device server (11) playing the AI-generated medium (E, E', E2') on a playback interface (12).

2. The method as claimed in claim 1, wherein the audio-visual playback device (1) includes a searching and sorting interface (13) that is in signal communication with the device server (11); the searching and sorting interface (13) sorts the link codes in the audio-visual playback device (1) and displays the search result.

3. The method as claimed in claim 2, wherein the audio-visual playback device (1) further includes a database (15) that is in signal communication with the device server (11); after the searching and sorting interface (13) displays the search result, a user selects at least one of the link codes, the searching and sorting interface (13) adds the selected link code to a playback queue, the device server (11) or the generative artificial intelligence server (3) sequentially obtains the link codes from the playback queue; when the number of times that one of the link codes is obtained reaches a threshold, the device server (11) stores the one of the link codes in the database (15).

4. The method as claimed in claim 2, wherein the streaming platform (A) includes a plurality of streaming platforms (A) each having the application programming interface server (A1) that is in signal communication with the device server (11), the searching and sorting interface (13) has a single search mode and an integrated search mode; when the searching and sorting interface (13) is operated in the single search mode, the device server (11) sends the audio-visual search request to the application programming interface server (A1) of one of the streaming platforms (A), the audio-visual playback device (1) displays the search result of the one of the streaming platforms (A); when the searching and sorting interface (13) is operated in the integrated search mode, the device server (11) sends the audio-visual search request to the application programming interface servers (A1) of all the streaming platforms (A), and the device server (11) integrates and displays the link codes from the streaming platforms that correspond to the same online audio-visual content (B, B').

5. The method as claimed in claim 2, wherein before the searching and sorting interface (13) displays the search result, the device server (11) or the generative artificial intelligence server (3) first verifies whether the streaming platform (A) is activated before the streaming platform (A) acquires the online audio-visual content (B, B').

6. The method as claimed in claim 1, wherein before outputting the AI-generated medium, the generative artificial intelligence server (3) first performs an additional change to at least one part of the selected track contents (C1, C2, C3, C1', C2', C3', C4'), and the additional change includes one of style, tempo, melody, volume and pitch.

7. The method as claimed in claim 1, wherein the sampling module (2) further includes an object recognition model and an edge detection algorithm, and when the online audio-visual content (B, B') contains both audio and video, a main character image is captured by the object recognition model and the edge detection algorithm, and at least one of the track contents (C1, C2, C3, C1', C2', C3', C4') includes the captured main character image.

8. A system for acquiring an online audio-visual content, separating tracks, integrating artificial intelligence and outputting an AI-generated medium, in signal communication with a streaming platform, the streaming platform having an application programming interface (API) server, the system comprising:

an audio-visual playback device (1), having a device server (11) and a playback interface (12) that is in signal communication with the device server (11);

a sampling module (2) in signal communication with the device server (11), the sampling module comprising a front-end neural network; and

a generative artificial intelligence server (3) in signal communication with the device server (11) and the sampling module (2);

wherein the device server (11) sends an audio-visual search request to the API server (A1) by an API technique according to a keyword; after the API server (A1) performs a search according to the audio-visual search request to obtain a search result corresponding to the keyword, the device server (11) obtains a plurality of link codes of the search result from the API server (A1) by the API technique, each of the link codes corresponds to a different online audio-visual content (B, B'), and the device server (11) first excludes any of the link codes corresponding to the online audio-visual content in an invalid state;

wherein the device server (11) obtains one of the link codes, and the device server (11) or the generative artificial intelligence server (3) parses the one of the link codes to acquire the corresponding online audio-visual content from the API server without downloading the online audio-visual content;

wherein the device server (11) or the generative artificial intelligence server (3) separates a plurality of track contents (C1, C2, C3, C1', C2', C3', C4') from the online audio-visual content (B, B') through the sampling module (2), the front-end neural network samples and separates a plurality of track audios from the online audio-visual content according to a frequency characteristic, and the track contents includes the track audios;

wherein the generative artificial intelligence server (3) selects part or all of the track contents (C1, C2, C3, C1', C2', C3', C4'), the selected track contents (C1, C2, C3, C1', C2', C3', C4') are aligned and positioned on a timeline, a playback start time common to the track contents in a sampling process is marked as a start point of each of the track contents on the timeline, and the start points of the selected track contents are aligned with one another to complete the positioning;

wherein the generative artificial intelligence server (3) generates at least one AI-generated content (D1, D2, D3, D1', D2', D3', D4', D5') according to the selected track contents (C1, C2, C3, C1', C2', C3', C4'), the at least one AI-generated content includes at least one of an AI-generated image, an AI-generated animation and an AI-generated text, each of the at least one AI-generated content is an independently selectable object, and the at least one AI-generated content (D1, D2, D3, D1', D2', D3', D4', D5') is aligned with the track contents (C1, C2, C3, C1', C2', C3', C4') and positioned on the timeline;

wherein the generative artificial intelligence server (3) mixes the selected track contents (C1, C2, C3, C1', C2', C3', C4') with the at least one AI-generated content (D1, D2, D3, D1', D2', D3', D4', D5'), outputs an AI-generated medium (E, E', E2') and transmits the AI-generated medium back to the device server (11); and

wherein the device server (11) plays the AI-generated medium (E, E', E2') on a playback interface (12).

9. The system as claimed in claim 8, wherein the streaming platform (A) has an application programming interface server (A1), the device server (11) is in signal communication with the application programming interface server (A1), the audio-visual playback device (1) includes a searching and sorting interface (13) that is in signal communication with the device server (11); the searching and sorting interface (13) sorts the link codes in the audio-visual playback device (1) and displays the search result.

10. The system as claimed in claim 9, wherein the audio-visual playback device (1) further includes a database (15) that is in signal communication with the device server (11); after the searching and sorting interface (13) displays the search result, a user selects at least one of the link

codes, the searching and sorting interface (13) adds the selected link code to a playback queue, the device server (11) or the generative artificial intelligence server (3) sequentially obtains the link codes from the playback queue; when the number of times that one of the link codes is obtained reaches a threshold, the device server (11) stores the one of the link codes in the database (15).

11. The system as claimed in claim 9, wherein the streaming platform (A) includes a plurality of streaming platforms (A) each having the application programming interface server (A1) that is in signal communication with the device server (11), the searching and sorting interface (13) has a single search mode and an integrated search mode; when the searching and sorting interface (13) is operated in the single search mode, the device server (11) sends the audio-visual search request to the application programming interface server (A1) of one of the streaming platforms (A), the audio-visual playback device (1) displays the search result of the one of the streaming platforms (A); when the searching and sorting interface (13) is operated in the integrated search mode, the device server (11) sends the audio-visual search request to the application programming interface servers (A1) of all the streaming platforms (A), and the device server (11) integrates and displays the link codes from the streaming platforms that correspond to the same online audio-visual content (B, B').

12. The system as claimed in claim 9, wherein before the searching and sorting interface (13) displays the search result, the device server (11) or the generative artificial intelligence server (3) first verifies whether the streaming platform (A) is activated before the streaming platform (A) acquires the online audio-visual content (B, B').

13. The system as claimed in claim 8, wherein before outputting the AI-generated medium, the generative artificial intelligence server (3) first performs an additional change to at least one part of the selected track contents (C1, C2, C3, C1', C2', C3', C4'), and the additional change includes one of style, tempo, melody, volume and pitch.

14. The system as claimed in claim 8, wherein the sampling module (2) further includes an object recognition model and an edge detection algorithm, and when the online audio-visual content (B,B') contains both audio and video, a main character image is captured by the object recognition model and the edge detection algorithm, and at least one of the track contents (C1, C2, C3, C1', C2', C3', C4') includes the captured main character image.