Claims:

1. A subtitle processing method, comprising:

performing, in a process of editing multimedia material, speech recognition on an audio corresponding to the multimedia material, to obtain a subtitle text corresponding to the audio and timestamp information of audio fragments respectively corresponding to respective text elements included in the subtitle text, wherein each of the text elements is a word;

matching with respective material units in the multimedia material according to the timestamp information of the audio fragments respectively corresponding to the respective text elements, and determining material fragments in the multimedia material respectively matching with the respective text elements; wherein a time in an editing timeline of a material fragment matching with a text element is consistent with a time in the editing timeline of an audio fragment corresponding to the text element; and

synthesizing the respective text elements respectively with material fragments in a matching time period to obtain a target multimedia material with an animation effect in which the subtitle text jumps out word by word;

wherein the synthesizing the respective text elements respectively with the material fragments in the matching time period to obtain the target multimedia material with the animation effect in which the subtitle text jumps out word by word comprises:

applying a specified first subtitle animation style to the respective text elements in batches, and synthesizing the text elements with the first subtitle animation style and the material fragments in the matching time period, to obtain the target multimedia material with the animation effect in which the subtitle text jumps out word by word in the first subtitle animation style.

2. The method according to claim 1, wherein, in the editing timeline, a starting moment of a time period to which the material fragment matching with the text element belongs is consistent with a starting moment of the audio fragment corresponding to the text element; and in the editing timeline, an ending moment of the time period of the material fragment matching with the text element belongs is consistent with an ending moment of the audio fragment corresponding to the text element, or the ending moment of the time period to which the material fragment matching the text element belongs is later than the ending moment of the audio fragment corresponding to the text element.

3. The method according to claim 1 or 2, further comprising:

deleting, in response to a text deletion instruction, a corresponding text element from the subtitle text to obtain an updated subtitle text;

determining, according to timestamp information of audio fragments respectively corresponding to respective text elements in the updated subtitle text, the material fragments in the multimedia material matching with the respective text elements;

synthesizing the respective text elements included in the updated subtitle text respectively with material fragments in a matching time period, so that a subtitle with the animation effect in which the subtitle text jumps out word by word is re-added to the multimedia material.

4. The method according to any one of claims 1-3, further comprising:

inserting, in response to a text insertion instruction, a new text element into the subtitle text to obtain an updated subtitle text;

matching with the respective material units in the multimedia material according to timestamp information of audio fragments respectively corresponding to respective text elements in the updated subtitle text, and determining the material fragments in the multimedia material respectively matching with the respective text elements; wherein, the new text element is merged with an adjacent text element and shares timestamp information of an audio fragment corresponding to the adjacent text element with the adjacent text element;

synthesizing the respective text elements included in the updated subtitle text respectively with the material fragments in a matching time range, so that a subtitle with the animation effect in which the subtitle text jumps out word by word is re-added to the multimedia material.

5. The method according to claim 4, wherein, if a position where the new text element is inserted is at a very beginning of the subtitle text, the new text element is merged with a first text element in the subtitle text to share timestamp information of an audio fragment corresponding to the first text element; and

if the position where the new text element is inserted is in the middle or at the end of the subtitle text, the new text element is merged with an adjacent previous text element to share timestamp information of an audio fragment corresponding to the previous text element.

6. The method according to any one of claims 1-5, further comprising:

replacing, in response to a text replacement instruction, one or more text elements in the subtitle text with a replacement text to obtain an updated subtitle text;

determining, according to timestamp information of audio fragments respectively corresponding to respective text elements in the updated subtitle text, material fragments in the multimedia material matching with the respective text elements; wherein the replacement text corresponds to timestamp information of an audio fragment corresponding to the replaced text element;

synthesizing the respective text elements included in the updated subtitle text respectively with the material fragments in the matching time period, so that a subtitle with the animation effect in which the subtitle text jumps out word by word is re-added to the multimedia material.

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

applying, in response to a subtitle animation style switching instruction, a second subtitle animation style to the respective text elements in batches, and synthesizing the text elements with the second subtitle animation style and the material fragments in a matching time range to obtain the target multimedia material with the animation effect in which the subtitle text jumps out word by word in the second subtitle animation style.

8. The method according to any one of claims 1-7, wherein the audio corresponding to the multimedia material is an original audio included in the multimedia material or a background music added to the multimedia material.

9. An electronic device comprising a memory and a processor, wherein,

the memory is configured to store computer program instructions;

the processor is configured to execute the computer program instructions to cause the electronic device to implement a subtitle processing method, which comprises:

performing, in a process of editing multimedia material, speech recognition on an audio corresponding to the multimedia material, to obtain a subtitle text corresponding to the audio and timestamp information of audio fragments respectively corresponding to respective text elements included in the subtitle text, wherein each of the text elements is a word;

matching with respective material units in the multimedia material according to the timestamp information of the audio fragments respectively corresponding to the respective text elements, and determining material fragments in the multimedia material respectively matching with the respective text elements; wherein a time in an editing timeline of a material fragment matching with a text element is consistent with a time in the editing timeline of an audio fragment corresponding to the text element; and

synthesizing the respective text elements respectively with material fragments in a matching time period to obtain a target multimedia material with an animation effect in which the subtitle text jumps out word by word;

wherein the synthesizing the respective text elements respectively with the material fragments in the matching time period to obtain the target multimedia material with the animation effect in which the subtitle text jumps out word by word comprises:

applying a specified first subtitle animation style to the respective text elements in batches, and synthesizing the text elements with the first subtitle animation style and the material fragments in the matching time period, to obtain the target multimedia material with the animation effect in which the subtitle text jumps out word by word in the first subtitle animation style.

10. The electronic device according to claim 9, wherein, in the editing timeline, a starting moment of a time period to which the material fragment matching with the text element belongs is consistent with a starting moment of the audio fragment corresponding to the text element; and in the editing timeline, an ending moment of the time period of the material fragment matching with the text element belongs is consistent with an ending moment of the audio fragment corresponding to the text element, or the ending moment of the time period to which the material fragment matching the text element belongs is later than the ending moment of the audio fragment corresponding to the text element.

11. The electronic device according to claim 9 or 10, wherein the subtitle processing method further comprises:

deleting, in response to a text deletion instruction, a corresponding text element from the subtitle text to obtain an updated subtitle text;

determining, according to timestamp information of audio fragments respectively corresponding to respective text elements in the updated subtitle text, the material fragments in the multimedia material matching with the respective text elements;

synthesizing the respective text elements included in the updated subtitle text respectively with material fragments in a matching time period, so that a subtitle with the animation effect in which the subtitle text jumps out word by word is re-added to the multimedia material.

12. The electronic device according to any one of claims 9-11, wherein the subtitle processing method further comprises:

inserting, in response to a text insertion instruction, a new text element into the subtitle text to obtain an updated subtitle text;

matching with the respective material units in the multimedia material according to timestamp information of audio fragments respectively corresponding to respective text elements in the updated subtitle text, and determining the material fragments in the multimedia material respectively matching with the respective text elements; wherein, the new text element is merged with an adjacent text element and shares timestamp information of an audio fragment corresponding to the adjacent text element with the adjacent text element;

synthesizing the respective text elements included in the updated subtitle text respectively with the material fragments in a matching time range, so that a subtitle with the animation effect in which the subtitle text jumps out word by word is re-added to the multimedia material.

13. The electronic device according to claim 12, wherein, if a position where the new text element is inserted is at a very beginning of the subtitle text, the new text element is merged with a first text element in the subtitle text to share timestamp information of an audio fragment corresponding to the first text element; and

if the position where the new text element is inserted is in the middle or at the end of the subtitle text, the new text element is merged with an adjacent previous text element to share timestamp information of an audio fragment corresponding to the previous text element.

14. The electronic device according to any one of claims 9-13, wherein the subtitle processing method further comprises:

replacing, in response to a text replacement instruction, one or more text elements in the subtitle text with a replacement text to obtain an updated subtitle text;

determining, according to timestamp information of audio fragments respectively corresponding to respective text elements in the updated subtitle text, material fragments in the multimedia material matching with the respective text elements; wherein the replacement text corresponds to timestamp information of an audio fragment corresponding to the replaced text element;

synthesizing the respective text elements included in the updated subtitle text respectively with the material fragments in the matching time period, so that a subtitle with the animation effect in which the subtitle text jumps out word by word is re-added to the multimedia material.

15. The electronic device according to claim 9, wherein the subtitle processing method further comprises:

applying, in response to a subtitle animation style switching instruction, a second subtitle animation style to the respective text elements in batches, and synthesizing the text elements with the second subtitle animation style and the material fragments in a matching time range to obtain the target multimedia material with the animation effect in which the subtitle text jumps out word by word in the second subtitle animation style.

16. The electronic device according to any one of claims 9-15, wherein the audio corresponding to the multimedia material is an original audio included in the multimedia material or a background music added to the multimedia material.

17. A readable storage medium comprising computer program instructions, wherein,

upon at least one processor of an electronic device executing the computer program instructions, the computer program instructions cause the electronic device to implement a subtitle processing method, which comprises:

performing, in a process of editing multimedia material, speech recognition on an audio corresponding to the multimedia material, to obtain a subtitle text corresponding to the audio and timestamp information of audio fragments respectively corresponding to respective text elements included in the subtitle text, wherein each of the text elements is a word;

matching with respective material units in the multimedia material according to the timestamp information of the audio fragments respectively corresponding to the respective text elements, and

determining material fragments in the multimedia material respectively matching with the respective text elements; wherein a time in an editing timeline of a material fragment matching with a text element is consistent with a time in the editing timeline of an audio fragment corresponding to the text element; and

synthesizing the respective text elements respectively with material fragments in a matching time period to obtain a target multimedia material with an animation effect in which the subtitle text jumps out word by word;

wherein the synthesizing the respective text elements respectively with the material fragments in the matching time period to obtain the target multimedia material with the animation effect in which the subtitle text jumps out word by word comprises:

applying a specified first subtitle animation style to the respective text elements in batches, and synthesizing the text elements with the first subtitle animation style and the material fragments in the matching time period, to obtain the target multimedia material with the animation effect in which the subtitle text jumps out word by word in the first subtitle animation style.

18. The readable storage medium according to claim 17, wherein, in the editing timeline, a starting moment of a time period to which the material fragment matching with the text element belongs is consistent with a starting moment of the audio fragment corresponding to the text element; and in the editing timeline, an ending moment of the time period of the material fragment matching with the text element belongs is consistent with an ending moment of the audio fragment corresponding to the text element, or the ending moment of the time period to which the material fragment matching the text element belongs is later than the ending moment of the audio fragment corresponding to the text element.