CLAIMS:

1. A method of audio processing, comprising:

obtaining a first media content input by a user, the first media content comprising a first audio content; and

providing a second media content based on a selection of a style by the user, the second media content comprising a second audio content generated based on the first audio content, the second audio content having the same timbre as the first audio content, and the second audio content having at least one audio attribute corresponding to the style.

2. The method of claim 1, wherein the at least one audio attribute comprises at least one of pitch, or rhythm.

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

displaying a selection panel, wherein the selection panel provides a set of candidate audio effects; and

receiving a selection of a first audio effect in the set of audio effects by the user, the first audio effect corresponding to the style.

4. The method of claim 1 or 2, wherein the first media content comprises a first video content, and the second media content is generated by:

generating the second audio content based on the first audio content;

adjusting a play speed of a visual content of the first video content based on the second audio content; and

generating, based on the second audio content and the adjusted visual content, a second video content as the second media content.

5. The method of claim 4, wherein adjusting the play speed of the visual content of the first video content based on the second audio content comprises:

determining an audio portion in the second audio content corresponding to content and a video portion in the visual content corresponding to the content; and

adjusting a play speed of the video portion, so that the video portion is synchronous with the audio portion.

6. The method of claim 1 or 4, wherein obtaining the first media content input by the user comprises:

obtaining the first media content recorded by the user; and

obtaining the first media content uploaded by the user.

7. The method of claim 1 or 5, wherein the second audio content is generated by:

extracting the first audio content from the first media content; and

processing the first audio content by using a model to generate the second audio content, wherein the model is trained based on sample data corresponding to the style.

8. The method of claim 7, wherein the model comprises:

a speech recognition module configured to determine a text content corresponding to the first audio content;

a style conversion module configured to convert the text content into a first feature corresponding to the style; and

a speech generation module configured to generate an intermediate audio content based on the first feature and a second feature for characterizing a timbre of the first audio content.

9. The method of claim 8, wherein the speech generation module comprises a diffusion model.

10. The method of claim 8, wherein the first feature indicates at least a stress conversion style corresponding to the style.

11. An apparatus for audio processing, comprising:

an obtaining module configured to obtain a first media content input by a user, the first media content comprising a first audio content; and

a providing module configured to provide a second media content based on a selection of a style by the user, the second media content comprising a second audio content generated based on the first audio content, the second audio content having the same timbre as the first audio content, and the second audio content having at least one audio attribute corresponding to the style.

12. An electronic device, comprising:

at least one processing unit;

at least one memory coupled to the at least one processing unit and storing instructions for execution by the at least one processing unit, wherein the instructions, when executed by the at least one processing unit, cause the electronic device to perform the method according to any one of claims 1 to 10.

13. A computer readable storage medium, on which a computer program is stored, wherein the computer program is executable by a processor to implement the method according to any one of claims 1 to 10.