

ABSTRACT

The disclosure relates to methods, apparatuses, devices, and storage medium for audio processing. The method includes: extracting, from first media content, background audio content and speech audio content corresponding to text content; generating, based on a request to replace first text in text content with second text, a second audio segment corresponding to the second text using timbre information associated with a first audio segment of speech audio content, which corresponding to first text; adjusting, based on the second audio segment, a third audio segment of the background audio content corresponding to first text; and generating second media content based on first media content, the second audio segment, and the adjusted third audio segment. Thus, the disclosure supports editing on the media content by modifying the text corresponding to the media content, and makes the modified audio content more realistic, improving the quality of the edited media content.