

TEXT-BASED SPEECH SYNTHESIS METHOD AND DEVICE, COMPUTER DEVICE, AND NON-TRANSITORY COMPUTER-READABLE STORAGE MEDIUM

Abstract

The application provides a text-based speech synthesis method and device, and computer device. The text-based speech synthesis method includes that: first, a target text to be recognized is obtained (101); and then, each character in the target text is discretely characterized to generate a feature vector corresponding to each character (102); secondly, the feature vector is input into a pre-trained spectrum conversion model, to obtain a Mel-spectrum corresponding to each character in the target text output by the spectrum conversion model (103); finally, the Mel-spectrum is converted to speech to obtain speech corresponding to the target text (104). In this way, during speech synthesis, there is no need to mark every character in the text in pinyin, which effectively reduces the workload in the speech synthesis process and provides an effective solution for the pronunciation problem in the speech synthesis process. The application has a wider application scope in the field of artificial intelligence.

Figure 1