

ABSTRACT

The embodiments in the present application provide a text-to-speech (TTS) conversion method comprising: obtaining text which is to be segmented into words to generate at least one character substring; generating audio units corresponding to the character substrings; and concatenating the audio units corresponding to the character substrings to output the final speech corresponding to the text. The method proposed in the present application is able to improve both the efficiency and accuracy of TTS conversion.

FIG. 1