

ABSTRACT

An audio synthesis method, a device and a computer-readable storage medium. The method comprises: acquiring song audio to be processed and corresponding song information (S101); performing human voice separation processing on the song audio to be processed, so as to obtain human voice audio (S102); on the basis of the human voice audio, determining target timbre information from a plurality of pieces of candidate timbre information (S103); acquiring a text template, and on the basis of the text template and the song information, generating text to be processed (S104); on the basis of the target timbre information, performing audio synthesis processing on the text to be processed, so as to obtain audio to be synthesized (S105); and performing synthesis processing on the audio to be synthesized and the song audio to be processed, so as to obtain synthesized audio (S106). The quality of synthesized audio which is obtained by means of the method is not limited to the artificial broadcasting level, such that the quality of synthesized audio is higher, and the playing effect is better.