TEXT PROCESSING METHOD, MODEL TRAINING METHOD AND RELATED APPARATUSES

ABSTRACT

The present disclosure provides a text processing method, model training method and related apparatuses. The text processing method includes: obtaining a first phonetic representation of reference acoustic information and a phoneme sequence of text information; processing, in an autoregressive manner, the first phonetic representation of the reference acoustic information and the phoneme sequence to obtain a second phonetic representation of the text information; processing, in a non-autoregressive manner, the second phonetic representation of the text information, the phoneme sequence and first acoustic features of the reference acoustic information to obtain second acoustic features of the text information; and obtaining speech information corresponding to the text information and the reference acoustic information based on the second acoustic features of the text information.

FIG. 2