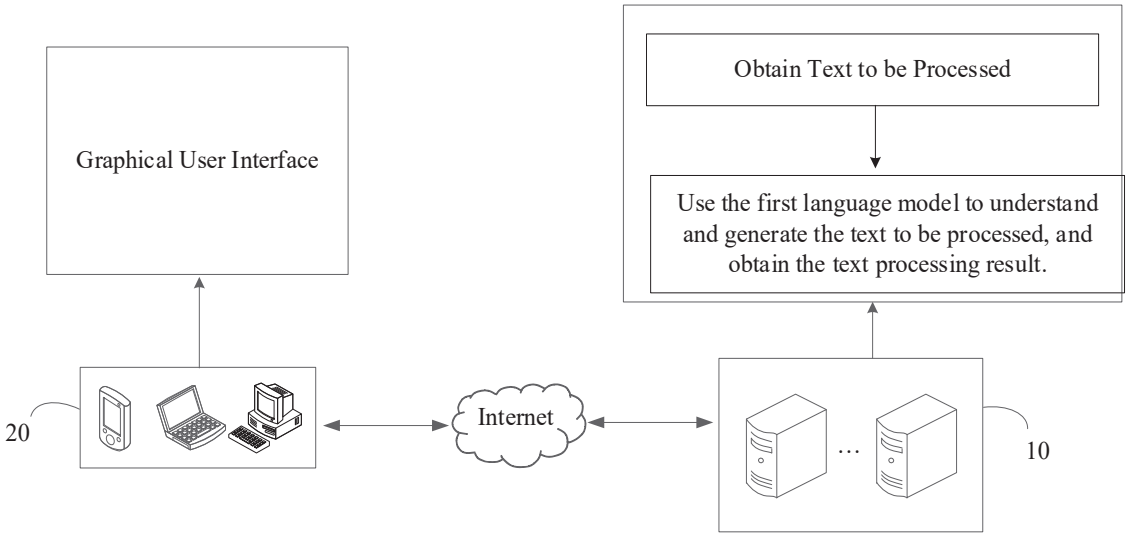
Graphical User Interface
Obtain Text to be Processed
Use the first language model to understand and generate the text to be processed, and obtain the text processing result.
20
Internet
10

FIG. 1

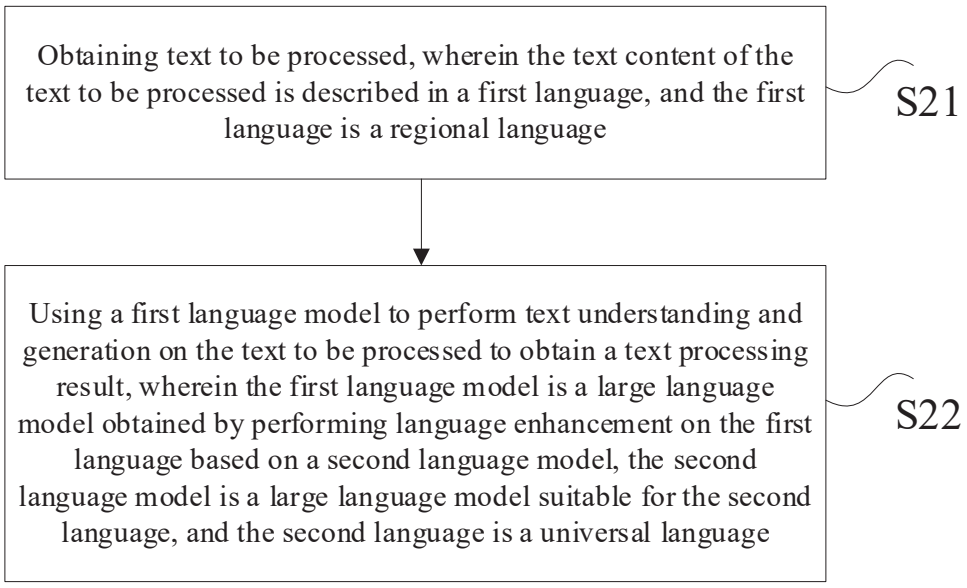
Obtaining text to be processed, wherein the text content of the text to be processed is described in a first language, and the first language is a regional language
S21
Using a first language model to perform text understanding and generation on the text to be processed to obtain a text processing result, wherein the first language model is a large language model obtained by performing language enhancement on the first language based on a second language model, the second language model is a large language model suitable for the second language, and the second language is a universal language
S22

FIG. 2

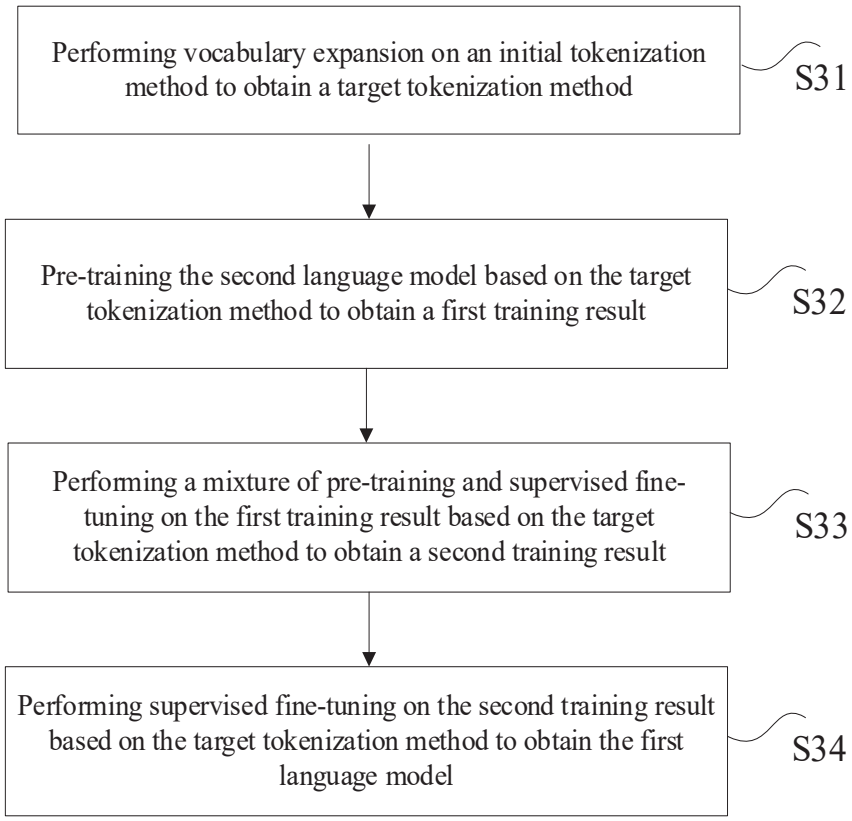
Performing vocabulary expansion on an initial tokenization method to obtain a target tokenization method
S31
Pre-training the second language model based on the target tokenization method to obtain a first training result
S32
Performing a mixture of pre-training and supervised fine-tuning on the first training result based on the target tokenization method to obtain a second training result
S33
Performing supervised fine-tuning on the second training result based on the target tokenization method to obtain the first language model
S34

FIG. 3

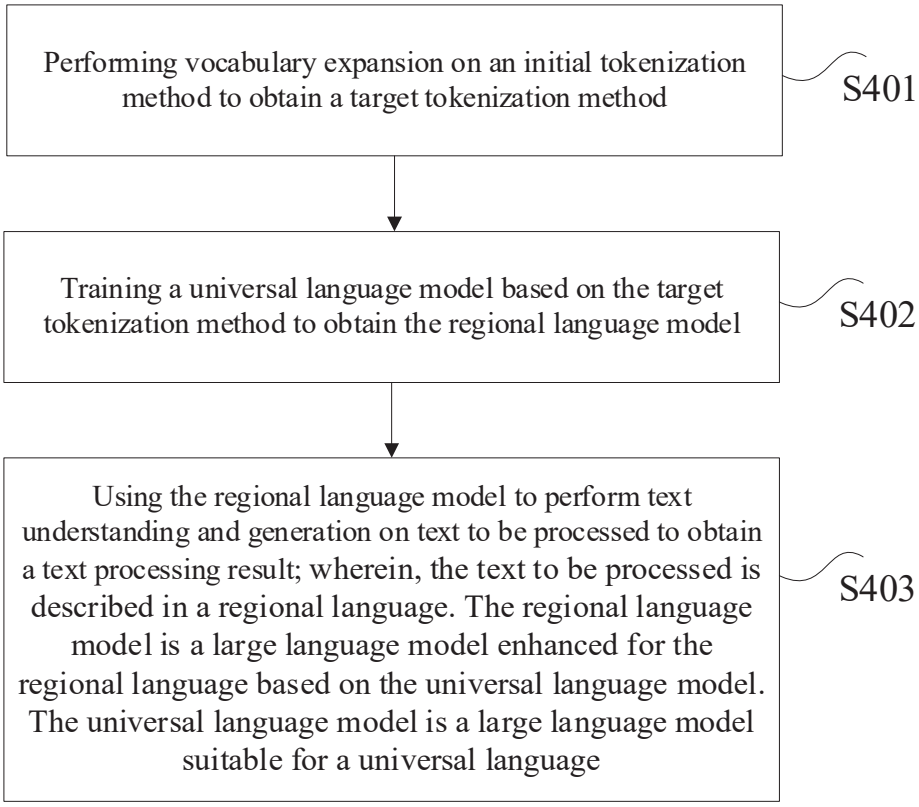
Performing vocabulary expansion on an initial tokenization method to obtain a target tokenization method
S401
Training a universal language model based on the target tokenization method to obtain the regional language model
S402
Using the regional language model to perform text understanding and generation on text to be processed to obtain a text processing result; wherein, the text to be processed is described in a regional language. The regional language model is a large language model enhanced for the regional language based on the universal language model. The universal language model is a large language model suitable for a universal language
S403

FIG. 4

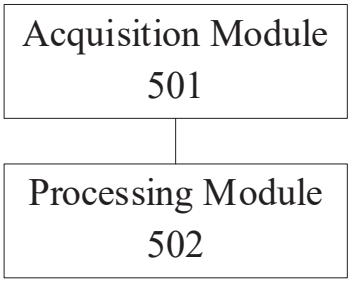
Acquisition Module 501
Processing Module 502

FIG. 5

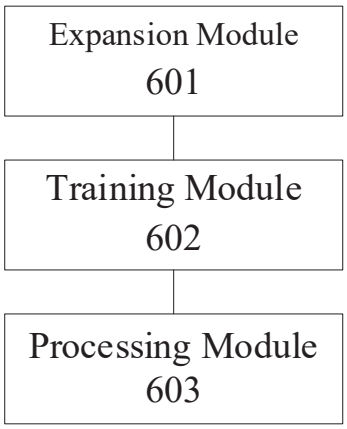
Expansion Module
601
Training Module
602
Processing Module
603

FIG. 6

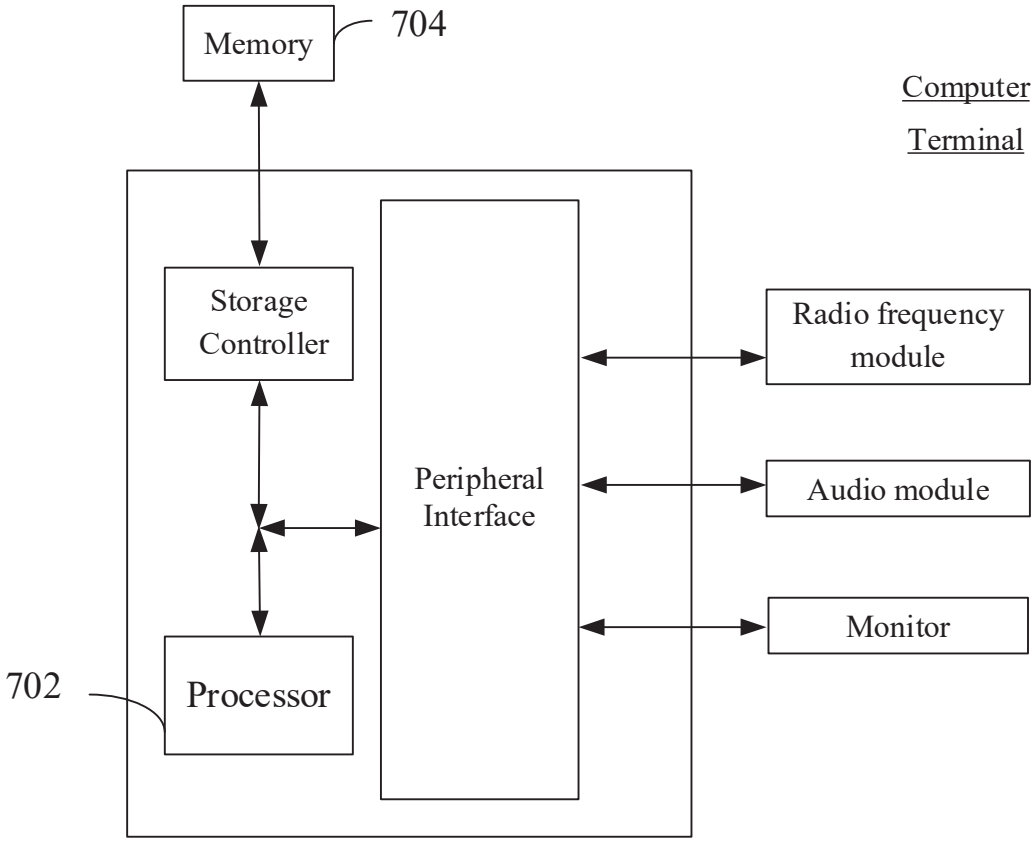
Memory
704
Computer
Terminal
Storage
Controller
Peripheral
Interface
Radio frequency
module
Audio module
Monitor
Processor
702

FIG. 7