

ABSTRACT

The present disclosure provides a method and apparatus for an automated operation of a cloud phone, relating to the field of cloud computing technology, specifically the field of cloud phone technology. A specific implementation of the method includes: obtaining a task sequence of the cloud phone; for a current task in the task sequence, performing the following operation steps: generating a current operation instruction by using a vision-language model to reason about the current task and current interface information of the cloud phone; after executing the current operation instruction on the cloud phone, loading next interface information; in response to the current task being the last task of the task sequence, determining that the automated operation of the cloud phone is completed; and in response to the current task not being the last task of the task sequence, taking a next task in the task sequence as a new current task, taking the next interface information as new current interface information, and continuing to perform the operation steps.

Fig.2