

ABSTRACT

A question answering method and apparatus, an electronic device and a storage medium. The method comprises: acquiring a first modal feature and a second modal feature
5 corresponding to each piece of first modal information among multi-modal information (S110); for one piece of the first modal information, respectively carrying out semantic similarity matching on the first modal feature and a user question feature and on the second modal feature and the user question feature to correspondingly obtain a first similarity vector representation and a second similarity vector representation (S120); carrying out
10 information fusion on the first similarity vector representation and the second similarity vector representation to obtain a multi-modal similarity score (S130); and in response to one of the multi-modal similarity scores being greater than or equal to a preset threshold, determining corresponding first modal information to be an answer (S140). The method can solve the problem in the related art of low answering accuracy caused by neglecting other
15 modal information in a document.

(FIG. 1)