Claims

1. A media retrieval method, comprising:

acquiring a media feature of a query media as a first media feature which contains a plurality of first media feature units;

performing, according to each individual of the first media feature unit, a first ranking on a plurality of existing media, and taking, according to a result of the first ranking, first k existing media as a first candidate media set, where k is a positive integer; and

performing, according to the plurality of the first media feature units arranged in sequence, a second ranking on the first candidate media set, and taking, according to a result of the second ranking, first n first candidate media as a retrieval result, where n is a positive integer.

2. The method according to claim 1, further comprising:

acquiring a media feature of the existing media as a second media feature in advance which contains a plurality of second media feature units; and

indexing the second media feature to obtain a feature index of the existing media in advance.

3. The method according to claim 2, wherein the step of performing, according to each individual of the first media feature unit, a first ranking on a plurality of existing media comprises:

performing, according to each individual of the first media feature unit, a term frequency–inverse document frequency (TF-IDF) ranking on a plurality of existing media.

4. The method according to claim 3, wherein the step of performing, according to each individual of the first media feature unit, a term frequency–inverse document frequency (TF-IDF) ranking on a plurality of existing media comprises:

matching the feature index of the existing media with the first media feature unit to perform the TF-IDF ranking on the existing media.

5. The method according to claim 4, wherein the step of obtaining a feature index of the existing media in advance comprises:

acquiring a forward index and/or an inverted index of the existing media in advance.

6. The method according to claim 4, wherein the step of matching the feature index of the existing media with the first media feature unit comprises:

making an exact match for the feature index of the existing media and the first media feature unit.

7. The method according to claim 2, wherein the step of performing, according to the plurality of the first media feature units arranged in sequence, a second ranking on the first candidate media set comprises:

obtaining a similarity matrix of the media in the first candidate media set according to the feature index of the existing media and the first media feature, and ranking the media in the first candidate media set on the basis of a straight line in the similarity matrix.

8. The method according to claim 7, wherein the step of acquiring a media feature of a query media as a first media feature comprises: acquiring multiple types of first media features of the query media;

the step of acquiring the media feature of the existing media as a second media feature in advance comprises: acquiring multiple types of second media feature of the existing media; and

the step of obtaining a similarity matrix of the media in the first candidate media set according to the feature index of the existing media and the first media feature comprises: determining the similarity matrix according to the multiple types of the second media features and the multiple types of the first media features.

9. The method according to claim 8, wherein each type of the first media feature includes a plurality of the first media feature units, and each type of the second media feature includes a plurality of the second media feature units;

the step of determining the similarity matrix according to the multiple types of the second media features and the multiple types of the first media features comprises:

determining the unit similarity of the same type of the second media feature unit and the first media feature unit respectively so as to obtain multiple types of the unit similarities; determining the average or minimum value of the multiple types of the unit similarities and determining the similarity matrix according to the average or minimum value of the multiple types of the unit similarities.

10. The method according to claim 2, further comprising:

slicing the query media and the existing media according to a preset time span to obtain a plurality of segments of the sub query media and segments of the sub existing media, and extracting the media features of the plurality of segments of the sub query media and segments of the sub existing media respectively to obtain a plurality of the first sub media features and the second sub media features of the same length.

11. The method according to claim 2, further comprising:

slicing, before the first ranking is performed, the obtained first media feature of the query media and the second media feature of the existing media according to a preset length to obtain a plurality of the first sub media features and the second sub media features of the same length.

12. The method according to claim 7, wherein, the plurality of the first media feature units are arranged in a temporal order in the first media feature, and the plurality of the second media feature units are arranged in a temporal order in the second media feature.

13. The method according to claim 12, further comprising:

determining duplicate segments of the query media and the media in the retrieval result according to the straight line in the similarity matrix.

14. A media retrieval apparatus, comprising:

a media feature acquisition module configured to acquire a media feature of a media as a first media feature which contains a plurality of first media feature units;

a first ranking module configured to perform, according to each individual of the first media feature unit, a first ranking on a plurality of existing media, and to take, according to a result of the first ranking, first k existing media as a first candidate media set, where k is a positive integer; and

a second ranking module configured to perform, according to the plurality of the first media feature units arranged in sequence, a second ranking on the first candidate media set, and to take, according to a result of the second ranking, first n first candidate media as a retrieval result, where n is a positive integer.

15. The apparatus according to claim 14, further comprising:

a feature indexing module configured to acquire a media feature of the existing media as a second media feature in advance which contains a plurality of second media feature units, and index the second media feature to obtain a feature index of the existing media in advance.

16. The apparatus according to claim 15, wherein the first ranking module is further configured to perform, according to each individual of the first media feature unit, a term frequency–inverse document frequency (TF-IDF) ranking on a plurality of existing media.

17. The apparatus according to claim 15, wherein the second ranking module further is further configured to obtain a similarity matrix of the media in the first candidate media set according to the feature index of the existing media and the first media feature, and to rank the media in the first candidate media set on the basis of a straight line in the similarity matrix.

18. The apparatus according to claim 17, wherein the media feature acquisition module feature further comprises an acquisition module for multiple types of first media features configured to acquiring multiple types of first media features of the query media;

the feature indexing module further comprises an acquisition module for multiple types of second media features configured to acquire multiple types of second media feature of the existing media; and

the second ranking module further comprises a second determination sub-module for unit similarity configured to determining the similarity matrix according to the multiple types of the second media features and the multiple types of the first media features.

19. A hardware device for media retrieval, comprising:

a memory, configured to store non-transitory computer-readable instructions; and

a processor, configured to perform the media retrieval method according to any of claims 1 to 13 when the computer-readable instructions are executed by the processor.

20. A computer-readable storage medium, storing non-transitory computer-readable instructions to perform the media retrieval method according to any of claims 1 to 13 when the non-transitory computer-readable instructions are executed by a computer.