

权利要求书

- 1、一种检索方法，所述方法包括：
确定文本和至少一个视频之间的第一相似度，所述文本用于表征检索条件；
确定所述文本的第一人物互动图和所述至少一个视频的第二人物互动图；
5 确定所述第一人物互动图和所述第二人物互动图之间的第二相似度；
根据所述第一相似度和所述第二相似度，从所述至少一个视频中确定出与所述检索条件相匹配的视频。
- 2、根据权利要求1所述的检索方法，其中，所述确定文本和至少一个视频之间的第一相似度，包括：
10 确定所述文本的段落特征；
确定所述至少一个视频的视频特征；
根据所述文本的段落特征和所述至少一个视频的视频特征，确定所述文本和所述至少一个视频之间的第一相似度。
- 3、根据权利要求2所述的检索方法，其中，所述段落特征包括句子特征和句子的数量；所述视频特征包括镜头特征和镜头的数量。
15 频特征包括镜头特征和镜头的数量。
- 4、根据权利要求1至3任一项所述的检索方法，其中，所述确定所述文本的第一人物互动图，包括：
检测所述文本中包含的人名；
在数据库中搜索到所述人名对应的人物的肖像，并提取所述肖像的图像特征，得到所述人物的
20 角色节点；
解析确定所述文本的语义树，基于所述语义树得到所述人物的运动特征，得到所述人物的动作节点；
将每个所述人物对应的角色节点和动作节点连接；
其中，所述人物的角色节点用肖像的图像特征表征；所述人物的动作节点采用语义树中的运动
25 特征表征。
- 5、根据权利要求4所述的检索方法，其中，所述方法还包括：
将连接同一动作节点的角色节点相互连接。
- 6、根据权利要求4或5所述的检索方法，其中，所述检测所述文本中包含的人名，包括：
将所述文本中的代词替换为所述代词所代表的所述人名。
- 30 7、根据权利要求1至6任一项所述的检索方法，其中，所述确定所述至少一个视频的第二人物互动图，包括：
检测出所述至少一个视频的每个镜头中的人物；
提取所述人物的人体特征与运动特征；
将所述人物的人体特征附加到所述人物的角色节点上，将所述人物的运动特征附加到所述人物

的动作节点上;

将每个人物对应的角色节点和动作节点相连。

8、根据权利要求7所述的检索方法,其中,所述确定所述至少一个视频的第二人物互动图,还包括:将同时出现在一个镜头中的一组人物作为同组人物,将所述同组人物中的人物的角色节点两两相连。

9、根据权利要求7或8所述的检索方法,其中,所述确定所述至少一个视频的第二人物互动图,还包括:

将一个镜头中的一位人物和其相邻镜头的每个人物的角色节点都相连。

10、根据权利要求1至9任一项所述的检索方法,其中,所述根据所述第一相似度和所述第二相似度,从所述至少一个视频中确定出与所述检索条件相匹配的视频,包括:

对每个视频的所述第一相似度和所述第二相似度加权求和,得到每个视频的相似程度值;

将相似程度值最高的视频,确定为与所述检索条件相匹配的视频。

11、根据权利要求1至10任一项所述的检索方法,其中,所述检索方法通过检索网络实现,所述方法还包括:

15 确定文本和训练样本集中的视频之间的第一相似度预测值,所述文本用于表征检索条件;

确定所述文本的第一人物互动图和所述训练样本集中的视频的第二人物互动图之间的第二相似度;

根据所述第一相似度预测值与所述第一相似度真值确定所述第一相似度的损失;

根据所述第二相似度预测值与所述第二相似度真值确定所述第二相似度的损失;

20 根据所述第一相似度的损失以及所述第二相似度的损失,结合损失函数确定总损失值;

根据所述总损失值调整所述检索网络的权重参数。

12、根据权利要求11所述的检索方法,所述检索网络包括第一子网络以及第二子网络;所述第一子网络用于确定文本与视频的第一相似度,所述第二子网络用于确定文本的第一人物互动图和所述视频的第二人物互动图之间的相似度;

25 所述根据所述总损失值调整所述检索网络的权重参数,包括:

基于所述总损失值调整所述第一子网络以及所述第二子网络的权重参数。

13、一种检索装置,所述装置包括:

第一确定模块,被配置为确定文本和至少一个视频之间的第一相似度,所述文本用于表征检索条件;

30 第二确定模块,被配置为确定所述文本的第一人物互动图和所述至少一个视频的第二人物互动图;确定所述第一人物互动图和所述第二人物互动图之间的第二相似度;

处理模块,被配置为根据所述第一相似度和所述第二相似度,从所述至少一个视频中确定出与所述检索条件相匹配的视频。

14、根据权利要求13所述的检索装置,其中,所述第一确定模块,被配置为:

确定所述文本的段落特征;

确定所述至少一个视频的视频特征;

根据所述文本的段落特征和所述至少一个视频的视频特征, 确定所述文本和所述至少一个视频之间的第一相似度。

- 5 15、根据权利要求 14 所述的检索装置, 其中, 所述段落特征包括句子特征和句子的数量; 所述视频特征包括镜头特征和镜头的数量。

16、根据权利要求 13 至 15 任一项所述的检索装置, 其中, 所述第二确定模块, 被配置为:

检测所述文本中包含的人名;

在数据库中搜索到所述人名对应的人物的肖像, 并提取所述肖像的图像特征, 得到所述人物的

- 10 角色节点;

解析确定所述文本的语义树, 基于所述语义树得到所述人物的运动特征, 得到所述人物的动作节点;

将每个所述人物对应的角色节点和动作节点连接;

其中, 所述人物的角色节点用肖像的图像特征表征; 所述人物的动作节点采用语义树中的运动特征表征。

- 15 特征表征。

17、根据权利要求 16 所述的检索装置, 其中, 所述第二确定模块, 还被配置为:

将连接同一动作节点的角色节点相互连接。

18、根据权利要求 16 或 17 所述的检索装置, 其中, 所述第二确定模块, 被配置为:

将所述文本中的代词替换为所述代词所代表的所述人名。

- 20 19、根据权利要求 13 至 18 任一项所述的检索装置, 其中, 所述第二确定模块, 被配置为:

检测出所述至少一个视频的每个镜头中的人物;

提取所述人物的人体特征与运动特征;

将所述人物的人体特征附加到所述人物的角色节点上, 将所述人物的运动特征附加到所述人物的动作节点上;

- 25 将每个人物对应的角色节点和动作节点相连。

20、根据权利要求 19 所述的检索装置, 其中, 所述第二确定模块, 还被配置为: 将同时出现在一个镜头中的一组人物作为同组人物, 将所述同组人物中的人物的角色节点两两相连。

21、根据权利要求 19 或 20 所述的检索装置, 其中, 所述第二确定模块, 还被配置为:

将一个镜头中的一位人物和其相邻镜头的每个人物的角色节点都相连。

- 30 22、根据权利要求 13 至 21 任一项所述的检索装置, 其中, 所述处理模块, 被配置为:

对每个视频的所述第一相似度和所述第二相似度加权求和, 得到每个视频的相似度值;

将相似度值最高的视频, 确定为与所述检索条件相匹配的视频。

23、根据权利要求 13 至 22 任一项所述的检索装置, 其中, 所述检索装置通过检索网络实现, 所述装置还包括:

训练模块，被配置为：

确定文本和训练样本集中的视频之间的第一相似度预测值，所述文本用于表征检索条件；

确定所述文本的第一人物互动图和所述训练样本集中的视频的第二人物互动图之间的第二相似度；

5 根据所述第一相似度预测值与所述第一相似度真值确定所述第一相似度的损失；

根据所述第二相似度预测值与所述第二相似度真值确定所述第二相似度的损失；

根据所述第一相似度的损失以及所述第二相似度的损失，结合损失函数确定总损失值；

根据所述总损失值调整所述检索网络的权重参数。

24、根据权利要求 23 所述的检索装置，所述检索网络包括第一子网络以及第二子网络；所述第
10 一子网络用于确定文本与视频的第一相似度，所述第二子网络用于确定文本的第一人物互动图和所
述视频的第二人物互动图之间的相似度；

所述训练模块，被配置为：

基于所述总损失值调整所述第一子网络以及所述第二子网络的权重参数。

25、一种检索装置，所述装置包括：存储器、处理器及存储在存储器上并可在处理器上运行的
15 计算机程序，所述处理器执行所述程序时实现权利要求 1 至 12 任一项所述的检索方法。

26、一种存储介质，所述存储介质存储有计算机程序，所述计算机程序被处理器执行时，能够
使得所述处理器执行权利要求 1 至 12 任一项所述的检索方法。

27、一种计算机程序，包括计算机可读代码，当所述计算机可读代码在电子设备中运行时，所
述电子设备中的处理器执行用于实现权利要求 1 至 12 中的任一项所述的检索方法。