(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116150704 A
(43)申请公布日 2023.05.23

(21)申请号 202310434950.4

(22)申请日 2023.04.21

(71)申请人 广东工业大学
地址 510050 广东省广州市越秀区东风东路729号

(72)发明人 赖培源 戴青云 刘庆

(74)专利代理机构 佛山粤进知识产权代理事务所(普通合伙) 44463
专利代理师 耿鹏

(51)Int.Cl.
G06F 18/25(2023.01)
G06F 18/23(2023.01)
G06F 18/2411(2023.01)
G06F 40/30(2020.01)

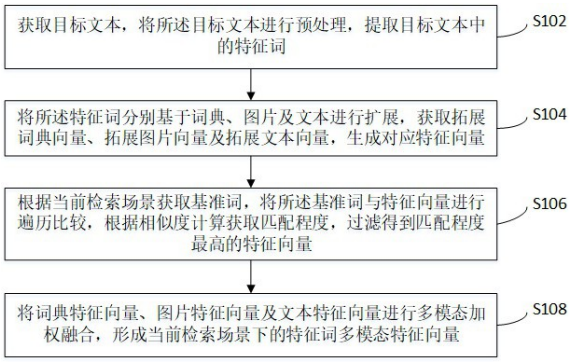
权利要求书3页 说明书9页 附图5页

(54)发明名称

基于语义相似度匹配的多模态融合表征方法及系统

(57)摘要

本发明公开了一种基于语义相似度匹配的多模态融合表征方法及系统,包括:获取目标文本,进行预处理提取目标文本中的特征词;将特征词分别基于词典、图片及文本进行扩展,获取若干拓展词典向量、拓展图片向量及拓展文本向量,生成对应特征向量;根据当前检索场景获取基准词,与特征向量进行遍历比较,根据相似度计算获取匹配程度,过滤得到匹配程度最高的特征向量;将词典特征向量、图片特征向量及文本特征向量进行多模态加权融合,形成当前检索场景下的特征词多模态特征向量。本方法通过多语义过滤及多模态特征表征,有效提高科技成果等复杂文本的量化表征,提升推荐及聚类系统性能。



CN 116150704 A

1.一种基于语义相似度匹配的多模态融合表征方法,其特征在于,包括以下步骤:

获取目标文本,将所述目标文本进行预处理,提取目标文本中的特征词;

将所述特征词分别基于词典、图片及文本进行扩展,获取拓展词典向量、拓展图片向量及拓展文本向量,生成对应特征向量;

根据当前检索场景获取基准词,将所述基准词与特征向量进行遍历比较,根据相似度计算获取匹配程度,过滤得到匹配程度最高的特征向量;

将词典特征向量、图片特征向量及文本特征向量进行多模态加权融合,形成当前检索场景下的特征词多模态特征向量。

2.根据权利要求1所述的一种基于语义相似度匹配的多模态融合表征方法,其特征在于,获取目标文本,将所述目标文本进行预处理,提取目标文本中的特征词,具体为:

获取科技成果的描述文本作为目标文本,将所述目标文本进行分词,获取对应分词结果,在所述分词结果中去停用词后对文本进行表示,利用词嵌入模型生成对应的词向量;

将原始特征空间中的词向量进行空间映射,转换到低维特征空间,判断不同词向量在目标文本中的出现频率,根据预设频率阈值进行词向量的筛选;

若词向量的出现频率大于等于预设频率阈值,则将对应词向量作为关键词,若词向量的出现频率小于预设频率阈值,则视为低频词进行滤除;

获取各关键词的位置信息,对标题位置及非标题位置设置不同的权重值,根据所述各关键词的位置信息获取关键词的位置权重;

确定科技成果对应目标文本的类别信息,根据所述类别信息利用大数据手段检索类别语料,获取某一关键词在对应类别语料的出现频率,获取关键词的类别权重;

基于关键词的位置权重及类别权重进行特征词筛选,将符合预设标准的关键词作为目标文本中的特征词。

3.根据权利要求1所述的一种基于语义相似度匹配的多模态融合表征方法,其特征在于,将特征词基于词典进行扩展,获取若干拓展词典向量,生成对应特征向量,具体为:

获取目标文本的特征词的词向量,通过所述特征词的词向量分析语义信息,基于预设词典进行语义拓展,

根据预设词典生成拓展词数据集合,通过聚类方法对所述拓展词数据集合进行分析,利用特征词的词向量作为初始聚类中心;

获取拓展词数据集合中各拓展词向量到初始聚类中心的欧式距离,将各拓展词向量归于最近的初始聚类中心构成聚类结果;

当拓展词数据集合中所有拓展词向量聚类结束后,在特征词的词向量对应的各个聚类结果中进行距离均值计算,获取新的聚类中心,当迭代次数达到预设标准后,结束聚类操作;

根据最后一次迭代运算获取各个特征词的词向量对应的聚类结果,在各个聚类结果中,获取对应的拓展词典语义,生成语义的词向量;

根据当前检索场景获取基准词,根据基准词与语义词向量的相似度计算语义的词向量的匹配程度,筛选符合标准的拓展词典向量,并生成词典特征向量。

4.根据权利要求1所述的一种基于语义相似度匹配的多模态融合表征方法,其特征在于,将特征词基于图片进行扩展,获取拓展图片向量,生成对应特征向量,具体为:

根据特征词构建检索任务获取拓展图片集合，将拓展图片集合中的拓展图片数据进行预处理；

基于注意力机制优化的ResNet50网络构建图片特征提取模型，将预处理后的拓展图片数据导入图片特征提取模型；

通过卷积获取拓展图片数据的特征，对特征进行平均池化实现特征的压缩，对压缩后的特征进行激励，预测各通道的重要性，并利用注意力机制获取各通道的权重；

对特征通道进行加权，对拓展图片数据的特征进行重新标定，输出拓展图片数据的特征，根据拓展图片数据的特征获取拓展图片向量，与基准词进行相似度计算获取拓展图片的匹配程度；

筛选符合标准的拓展图片向量，生成图片特征向量。

5.根据权利要求1所述的一种基于语义相似度匹配的多模态融合表征方法，其特征在于，将特征词基于文本进行扩展，获取拓展文本向量，生成对应特征向量，具体为：

获取目标文本的特征词，根据所述特征词的出现频率、位置特征及首次出现到末次出现的距离特征获取目标文本的特征序列；

基于所述特征序列利用相似度进行数据检索，获取拓展文本集合，将拓展文本集合中的拓展文本数据进行预处理；

基于LSTM网络构建文本特征提取模型，将预处理后的拓展文本数据导入文本特征提取模型，设文本数据的长度为T，通过LSTM单元进行特征表示，经过T个时间步骤后，文本特征提取模型输出表征文本特征的隐向量；

通过所述隐向量获取对应拓展文本向量，与基准词进行相似度计算获取拓展图片的匹配程度，筛选符合标准的拓展文本向量，生成文本特征向量。

6.根据权利要求1所述的一种基于语义相似度匹配的多模态融合表征方法，其特征在于，将词典特征向量、图片特征向量及文本特征向量进行多模态加权融合，形成当前检索场景下的特征词多模态特征向量，具体为：

对生成的词典特征向量、图片特征向量及文本特征向量进行降维，对三个向量进行拼接，获取融合后多模态特征向量；

将融合后多模态特征向量与对应的科技成果构建表征三元组，为科技成果推荐生成数据基础。

7.一种基于语义相似度匹配的多模态融合表征系统，其特征在于，该系统包括：存储器、处理器，所述存储器中包括一种基于语义相似度匹配的多模态融合表征方法程序，所述一种基于语义相似度匹配的多模态融合表征方法程序被所述处理器执行时实现如下步骤：

获取目标文本，将所述目标文本进行预处理，提取目标文本中的特征词；

将所述特征词分别基于词典、图片及文本进行扩展，获取拓展词典向量、拓展图片向量及拓展文本向量，生成对应特征向量；

根据当前检索场景获取基准词，将所述基准词与特征向量进行遍历比较，根据相似度计算获取匹配程度，过滤得到匹配程度最高的特征向量；

将词典特征向量、图片特征向量及文本特征向量进行多模态加权融合，形成当前检索场景下的特征词多模态特征向量。

8.根据权利要求7所述的一种基于语义相似度匹配的多模态融合表征系统，其特征在

于,将特征词基于词典进行扩展,获取若干拓展词典向量,生成对应特征向量,具体为:

获取目标文本的特征词的词向量,通过所述特征词的词向量分析语义信息,基于预设词典进行语义拓展,

根据预设词典生成拓展词数据集合,通过聚类方法对所述拓展词数据集合进行分析,利用特征词的词向量作为初始聚类中心;

获取拓展词数据集合中各拓展词向量到初始聚类中心的欧式距离,将各拓展词向量归于最近的初始聚类中心构成聚类结果;

当拓展词数据集合中所有拓展词向量聚类结束后,在特征词的词向量对应的各个聚类结果中进行距离均值计算,获取新的聚类中心,当迭代次数达到预设标准后,结束聚类操作;

根据最后一次迭代运算获取各个特征词的词向量对应的聚类结果,在各个聚类结果中,获取对应的拓展词典语义,生成语义的词向量;

根据当前检索场景获取基准词,根据基准词与语义词向量的相似度计算语义的词向量的匹配程度,筛选符合标准的拓展词典向量,并生成词典特征向量。

9.根据权利要求7所述的一种基于语义相似度匹配的多模态融合表征系统,其特征在于,将特征词基于图片进行扩展,获取拓展图片向量,生成对应特征向量,具体为:

根据特征词构建检索任务获取拓展图片集合,将拓展图片集合中的拓展图片数据进行预处理;

基于注意力机制优化的ResNet50网络构建图片特征提取模型,将预处理后的拓展图片数据导入图片特征提取模型;

通过卷积获取拓展图片数据的特征,对特征进行平均池化实现特征的压缩,对压缩后的特征进行激励,预测各通道的重要性,并利用注意力机制获取各通道的权重;

对特征通道进行加权,对拓展图片数据的特征进行重新标定,输出拓展图片数据的特征,根据拓展图片数据的特征获取拓展图片向量,与基准词进行相似度计算获取拓展图片的匹配程度;

筛选符合标准的拓展图片向量,生成图片特征向量。

10.根据权利要求7所述的一种基于语义相似度匹配的多模态融合表征系统,其特征在于,将特征词基于文本进行扩展,获取拓展文本向量,生成对应特征向量,具体为:

获取目标文本的特征词,根据所述特征词的出现频率、位置特征及首次出现到末次出现的距离特征获取目标文本的特征序列;

基于所述特征序列利用相似度进行数据检索,获取拓展文本集合,将拓展文本集合中的拓展文本数据进行预处理;

基于LSTM网络构建文本特征提取模型,将预处理后的拓展文本数据导入文本特征提取模型,设文本数据的长度为T,通过LSTM单元进行特征表示,经过T个时间步骤后,文本特征提取模型输出表征文本特征的隐向量;

通过所述隐向量获取对应拓展文本向量,与基准词进行相似度计算获取拓展图片的匹配程度,筛选符合标准的拓展文本向量,生成文本特征向量。

基于语义相似度匹配的多模态融合表征方法及系统

技术领域

[0001] 本发明涉及人工智能领域，更具体的，涉及一种基于语义相似度匹配的多模态融合表征方法及系统。

背景技术

[0002] 随着科学技术的高速发展，科研成果数量呈现爆炸式的增长。据相关统计，国内外论文总数已超过3亿篇，每天还有近万篇新的学术论文、专利、研究报告、项目成果被公开发表。海量的科技成果数据给科技创新活动提供了丰富的数据资源，然而这些数据专业性强，分类难度大，面临语义信息抽取困难，关联关系难以挖掘，相关信息无法扩充等问题，为科技成果的智能分析和查询带来了全新的挑战，也是成果转化平台对接中亟需解决的技术难题。

[0003] 在科技成果转化平台中，成果的特征实体提取是所有数据处理的核心基础，包括成果推荐、成果模糊检索，成果聚类，成果拓展等，都离不开精准的特征提取。而直接从成果的描述文本中提取实体词，面临提取精度低，词义特征不明确等多重问题，因此针对复杂文本、多语义等特征，如何提供一种基于多模态的特征提取方法是亟不可待需要解决的问题。

发明内容

[0004] 为了解决上述至少一个技术问题，本发明提出了一种基于语义相似度匹配的多模态融合表征方法及系统。

[0005] 本发明第一方面提供了一种基于语义相似度匹配的多模态融合表征方法，包括：

获取目标文本，将所述目标文本进行预处理，提取目标文本中的特征词；

将所述特征词分别基于词典、图片及文本进行扩展，获取拓展词典向量、拓展图片向量及拓展文本向量，生成对应特征向量；

根据当前检索场景获取基准词，将所述基准词与特征向量进行遍历比较，根据相似度计算获取匹配程度，过滤得到匹配程度最高的特征向量；

将词典特征向量、图片特征向量及文本特征向量进行多模态加权融合，形成当前检索场景下的特征词多模态特征向量。

[0006] 本方案中，获取目标文本，将所述目标文本进行预处理，提取目标文本中的特征词，具体为：

获取科技成果的描述文本作为目标文本，将所述目标文本进行分词，获取对应分词结果，在所述分词结果中去停用词后对文本进行表示，利用词嵌入模型生成对应的词向量；

将原始特征空间中的词向量进行空间映射，转换到低维特征空间，判断不同词向量在目标文本中的出现频率，根据预设频率阈值进行词向量的筛选；

若词向量的出现频率大于等于预设频率阈值，则将对应词向量作为关键词，若词向量的出现频率小于预设频率阈值，则视为低频词进行滤除；

获取各关键词的位置信息，对标题位置及非标题位置设置不同的权重值，根据所述各关键词的位置信息获取关键词的位置权重；

确定科技成果对应目标文本的类别信息，根据所述类别信息利用大数据手段检索类别语料，获取某一关键词在对应类别语料的出现频率，获取关键词的类别权重；

基于关键词的位置权重及类别权重进行特征词筛选，将符合预设标准的关键词作为目标文本中的特征词。

[0007] 本方案中，将特征词基于词典进行扩展，获取若干拓展词典向量，生成对应特征向量，具体为：

获取目标文本的特征词的词向量，通过所述特征词的词向量分析语义信息，基于预设词典进行语义拓展，

根据预设词典生成拓展词数据集合，通过聚类方法对所述拓展词数据集合进行分析，利用特征词的词向量作为初始聚类中心；

获取拓展词数据集合中各拓展词向量到初始聚类中心的欧式距离，将各拓展词向量归于最近的初始聚类中心构成聚类结果；

当拓展词数据集合中所有拓展词向量聚类结束后，在特征词的词向量对应的各个聚类结果中进行距离均值计算，获取新的聚类中心，当迭代次数达到预设标准后，结束聚类操作；

根据最后一次迭代运算获取各个特征词的词向量对应的聚类结果，在各个聚类结果中，获取对应的拓展词典语义，生成语义的词向量；

根据当前检索场景获取基准词，根据基准词与语义词向量的相似度计算语义的词向量的匹配程度，筛选符合标准的拓展词典向量，并生成词典特征向量。

[0008] 本方案中，将特征词基于图片进行扩展，获取拓展图片向量，生成对应特征向量，具体为：

根据特征词构建检索任务获取拓展图片集合，将拓展图片集合中的拓展图片数据进行预处理，采用双线性插值算法统一图片尺寸，如800×800；

基于注意力机制优化的ResNet50网络构建图片特征提取模型，将预处理后的拓展图片数据导入图片特征提取模型；

通过卷积获取拓展图片数据的特征，对特征进行平均池化实现特征的压缩，对压缩后的特征进行激励，预测各通道的重要性，并利用注意力机制获取各通道的权重；

对特征通道进行加权，对拓展图片数据的特征进行重新标定，输出拓展图片数据的特征，根据拓展图片数据的特征获取拓展图片向量，与基准词进行相似度计算获取拓展图片的匹配程度；

筛选符合标准的拓展图片向量，生成图片特征向量。

[0009] 本方案中，将特征词基于文本进行扩展，获取拓展文本向量，生成对应特征向量，具体为：

获取目标文本的特征词，根据所述特征词的出现频率、位置特征及首次出现到末次出现的距离特征获取目标文本的特征序列；

基于所述特征序列利用相似度进行数据检索，获取拓展文本集合，将拓展文本集合中的拓展文本数据进行预处理；

基于LSTM网络构建文本特征提取模型，将预处理后的拓展文本数据导入文本特征提取模型，设文本数据的长度为T，通过LSTM单元进行特征表示，经过T个时间步骤后，文本特征提取模型输出表征文本特征的隐向量；

通过所述隐向量获取对应拓展文本向量，与基准词进行相似度计算获取拓展图片的匹配程度，筛选符合标准的拓展文本向量，生成文本特征向量。

[0010] 本方案中，将词典特征向量、图片特征向量及文本特征向量进行多模态加权融合，形成当前检索场景下的特征词多模态特征向量，具体为：

对生成的词典特征向量、图片特征向量及文本特征向量按照预设的维度采用预训练模型进行降维，如whitenning模型，对三个向量进行拼接，获取融合后多模态特征向量；

将融合后多模态特征向量与对应的科技成果构建表征三元组，为科技成果推荐生成数据基础。

[0011] 本发明第二方面还提供了一种基于语义相似度匹配的多模态融合表征系统，该系统包括：存储器、处理器，所述存储器中包括一种基于语义相似度匹配的多模态融合表征方法程序，所述一种基于语义相似度匹配的多模态融合表征方法程序被所述处理器执行时实现如下步骤：

获取目标文本，将所述目标文本进行预处理，提取目标文本中的特征词；

将所述特征词分别基于词典、图片及文本进行扩展，获取拓展词典向量、拓展图片向量及拓展文本向量，生成对应特征向量；

根据当前检索场景获取基准词，将所述基准词与特征向量进行遍历比较，根据相似度计算获取匹配程度，过滤得到匹配程度最高的特征向量；

将词典特征向量、图片特征向量及文本特征向量进行多模态加权融合，形成当前检索场景下的特征词多模态特征向量。

[0012] 本发明公开了一种基于语义相似度匹配的多模态融合表征方法及系统，包括：获取目标文本，进行预处理提取目标文本中的特征词；将特征词分别基于词典、图片及文本进行扩展，获取若干拓展词典向量、拓展图片向量及拓展文本向量，生成对应特征向量；根据当前检索场景获取基准词，与特征向量进行遍历比较，根据相似度计算获取匹配程度，过滤得到匹配程度最高的特征向量；将词典特征向量、图片特征向量及文本特征向量进行多模态加权融合，形成当前检索场景下的特征词多模态特征向量。本方法通过多语义过滤及多模态特征表征，有效提高科技成果等复杂文本的量化表征，提升推荐及聚类系统性能。

附图说明

[0013] 图1示出了本发明一种基于语义相似度匹配的多模态融合表征方法的流程图；

图2示出了本发明基于词典进行扩展获取词典特征向量的方法流程图；

图3示出了本发明基于图片进行扩展获取图片特征向量的方法流程图；

图4示出了本发明基于文本进行扩展获取文本特征向量的方法流程图；

图5示出了本发明生成科技成果多模态融合表征的示意图；

图6示出了本发明一种基于语义相似度匹配的多模态融合表征系统的框图。

具体实施方式

[0014] 为了能够更清楚地理解本发明的上述目的、特征和优点，下面结合附图和具体实施方式对本发明进行进一步的详细描述。需要说明的是，在不冲突的情况下，本申请的实施例及实施例中的特征可以相互组合。

[0015] 在下面的描述中阐述了很多具体细节以便于充分理解本发明，但是，本发明还可以采用其他不同于在此描述的其他方式来实施，因此，本发明的保护范围并不受下面公开的具体实施例的限制。

[0016] 图1示出了本发明一种基于语义相似度匹配的多模态融合表征方法的流程图。

[0017] 如图1所示，本发明第一方面提供了一种基于语义相似度匹配的多模态融合表征方法，包括：

S102，获取目标文本，将所述目标文本进行预处理，提取目标文本中的特征词；

S104，将所述特征词分别基于词典、图片及文本进行扩展，获取拓展词典向量、拓展图片向量及拓展文本向量，生成对应特征向量；

S106，根据当前检索场景获取基准词，将所述基准词与特征向量进行遍历比较，根据相似度计算获取匹配程度，过滤得到匹配程度最高的特征向量；

S108，将词典特征向量、图片特征向量及文本特征向量进行多模态加权融合，形成当前检索场景下的特征词多模态特征向量。

[0018] 需要说明的是，获取科技成果的描述文本作为目标文本，将所述目标文本进行分词，获取对应分词结果，常用的分词工具有JIEBA分词、LTP等；去除文本中的冗余词汇，只保留有实际含义的词，通常在分词之后利用停用词典进行匹配，识别并过滤停用词，在所述分词结果中去停用词后对文本进行表示，利用Word2Vec词嵌入模型生成对应的词向量，将原始特征空间中的词向量进行空间映射，转换到低维特征空间，判断不同词向量在目标文本中的出现频率，根据预设频率阈值进行词向量的筛选；若词向量的出现频率大于等于预设频率阈值，则将对应词向量作为关键词，若词向量的出现频率小于预设频率阈值，则视为低频词进行滤除；获取各关键词的位置信息，对标题位置及非标题位置设置不同的权重值，例如标题位置权重设为1，非标题位置权重根据位置不同设为0.8及0.5，根据所述各关键词的位置信息获取关键词的位置权重；确定科技成果对应目标文本的类别信息，根据所述类别信息利用大数据手段检索海量的科技成果数据，提取其关键词作为类别语料，获取某一关键词在对应类别语料的出现频率，获取关键词的类别权重；基于关键词的位置权重及类别权重进行特征词筛选，将符合预设标准的关键词作为目标文本中的特征词。

[0019] 图2示出了本发明基于词典进行扩展获取词典特征向量的方法流程图。

[0020] 根据本发明实施例，将特征词基于词典进行扩展，获取若干拓展词典向量，生成对应特征向量，具体为：

S202，获取目标文本的特征词的词向量，通过所述特征词的词向量分析语义信息，基于预设词典进行语义拓展，

S204，根据预设词典生成拓展词数据集合，通过聚类方法对所述拓展词数据集合进行分析，利用特征词的词向量作为初始聚类中心；

S206，获取拓展词数据集合中各拓展词向量到初始聚类中心的欧式距离，将各拓展词向量归于最近的初始聚类中心构成聚类结果；

S208，当拓展词数据集合中所有拓展词向量聚类结束后，在特征词的词向量对应的各个聚类结果中进行距离均值计算，获取新的聚类中心，当迭代次数达到预设标准后，结束聚类操作；

S210，根据最后一次迭代运算获取各个特征词的词向量对应的聚类结果，在各个聚类结果中，获取对应的拓展词典语义，生成语义的词向量；

S212，根据当前检索场景获取基准词，根据基准词与语义词向量的相似度计算语义的词向量的匹配程度，筛选符合标准的拓展词典向量，并生成词典特征向量。

[0021] 需要说明的是，当拓展词数据集合中各拓展词向量划分结束后，求取聚类结果 $M^{(t)}$ 中每个类簇的均值作为新的聚类中心 $C_n^{(t+1)}$，设 Z_n 为第 n 个类簇的样本总数，x_{ni} 为该类簇的第个拓展词向量，聚类中心点求取方法具体为：

$$C_n^{(t+1)}=\frac{1}{Z_n}\sum_{i=1}^{Z_n}x_{ni};$$

选定标准测度函数 σ_c 及最大迭代次数 T_{max}，若 $|C_n^{(t+1)}-C_n^{(t)}|\leq\sigma_c$ 或迭代次数大于等于 T_{max}，结束聚类流程，取最后一次运算结果为最终聚类结果，选取最后一次运算结果为最终聚类结果，否则令 $t=t+1$，继续迭代聚类，t 为迭代次数。

[0022] 图3示出了本发明基于图片进行扩展获取图片特征向量的方法流程图。

[0023] 根据本发明实施例，将特征词基于图片进行扩展，获取拓展图片向量，生成对应特征向量，具体为：

S302，根据特征词构建检索任务获取拓展图片集合，将拓展图片集合中的拓展图片数据进行预处理；

S304，基于注意力机制优化的ResNet50网络构建图片特征提取模型，将预处理后的拓展图片数据导入图片特征提取模型；

S306，通过卷积获取拓展图片数据的特征，对特征进行平均池化实现特征的压缩，对压缩后的特征进行激励，预测各通道的重要性，并利用注意力机制获取各通道的权重；

S308，对特征通道进行加权，对拓展图片数据的特征进行重新标定，输出拓展图片数据的特征，根据拓展图片数据的特征获取拓展图片向量，与基准词进行相似度计算获取拓展图片的匹配程度；

S310，筛选符合标准的拓展图片向量，生成图片特征向量。

[0024] 需要说明的是，根据当前检索场景获取基准词，其基准词为固定基准词，例如特征词为“苹果”其对应的语义包括：语义1：一种水果、语义2：一家企业、语义3：一部动画片，当前检索场景的基准词为“西瓜”时，语义123分别和西瓜向量相比，选择相似度最高的语义1作为特征向量；

所述ResNet50网络通过残差学习加速了CNN训练过程，有效避免了梯度消失和梯度爆炸问题，；另外，本发明通过ResNet50作为主干网络，并引入通道注意力机制，提取图片的深度特征。利用欧式距离或余弦计算等相似度计算获取匹配程度，根据匹配程度筛选特征向量。

[0025] 图4示出了本发明基于文本进行扩展获取文本特征向量的方法流程图。

[0026] 根据本发明实施例，将特征词基于文本进行扩展，获取拓展文本向量，生成对应特征向量，具体为：

S402，获取目标文本的特征词，根据所述特征词的出现频率、位置特征及首次出现到末次出现的距离特征获取目标文本的特征序列；

S404，基于所述特征序列利用相似度进行数据检索，获取拓展文本集合，将拓展文本集合中的拓展文本数据进行预处理；

S406，基于LSTM网络构建文本特征提取模型，将预处理后的拓展文本数据导入文本特征提取模型，设文本数据的长度为T，通过LSTM单元进行特征表示，经过T个时间步骤后，文本特征提取模型输出表征文本特征的隐向量；

S408，通过所述隐向量获取对应拓展文本向量，与基准词进行相似度计算获取拓展图片的匹配程度，筛选符合标准的拓展文本向量，生成文本特征向量。

[0027] 需要说明的是，LSTM利用其特殊的门控结构，主要包括输入门、输出门和遗忘门，有选择地影响每个时刻的状态。输入门是控制当前单元的输入，输出门是控制当前LSTM单元的输出，遗忘门是控制上一时刻单元中存储的历史信息，本发明通过LSTM网络捕捉文本的关键信息；

对生成的词典特征向量、图片特征向量及文本特征向量进行降维，对三个向量进行拼接，获取融合后多模态特征向量；将融合后多模态特征向量与对应的科技成果构建表征三元组，为科技成果推荐生成数据基础，具体步骤如图5所示。

[0028] 图6示出了本发明一种基于语义相似度匹配的多模态融合表征系统的框图。

[0029] 本发明第二方面还提供了一种基于语义相似度匹配的多模态融合表征系统6，该系统包括：存储器61、处理器62，所述存储器中包括一种基于语义相似度匹配的多模态融合表征方法程序，所述一种基于语义相似度匹配的多模态融合表征方法程序被所述处理器执行时实现如下步骤：

获取目标文本，将所述目标文本进行预处理，提取目标文本中的特征词；

将所述特征词分别基于词典、图片及文本进行扩展，获取拓展词典向量、拓展图片向量及拓展文本向量，生成对应特征向量；

根据当前检索场景获取基准词，将所述基准词与特征向量进行遍历比较，根据相似度计算获取匹配程度，过滤得到匹配程度最高的特征向量；

将词典特征向量、图片特征向量及文本特征向量进行多模态加权融合，形成当前检索场景下的特征词多模态特征向量。

[0030] 需要说明的是，获取科技成果的描述文本作为目标文本，将所述目标文本进行分词，获取对应分词结果，常用的分词工具有JIEBA分词、LTP等；去除文本中的冗余词汇，只保留有实际含义的词，通常在分词之后利用停用词典进行匹配，识别并过滤停用词，在所述分词结果中去停用词后对文本进行表示，利用Word2Vec词嵌入模型生成对应的词向量，将原始特征空间中的词向量进行空间映射，转换到低维特征空间，判断不同词向量在目标文本中的出现频率，根据预设频率阈值进行词向量的筛选；若词向量的出现频率大于等于预设频率阈值，则将对应词向量作为关键词，若词向量的出现频率小于预设频率阈值，则视为低频词进行滤除；获取各关键词的位置信息，对标题位置及非标题位置设置不同的权重值，例

如标题位置权重设为1，非标题位置权重根据位置不同设为0.8及0.5，根据所述各关键词的位置信息获取关键词的位置权重；确定科技成果对应目标文本的类别信息，根据所述类别信息利用大数据手段检索海量的科技成果数据，提取其关键词作为类别语料，获取某一关键词在对应类别语料的出现频率，获取关键词的类别权重；基于关键词的位置权重及类别权重进行特征词筛选，将符合预设标准的关键词作为目标文本中的特征词。

[0031] 根据本发明实施例，将特征词基于词典进行扩展，获取若干拓展词典向量，生成对应特征向量，具体为：

获取目标文本的特征词的词向量，通过所述特征词的词向量分析语义信息，基于预设词典进行语义拓展，

根据预设词典生成拓展词数据集合，通过聚类方法对所述拓展词数据集合进行分析，利用特征词的词向量作为初始聚类中心；

获取拓展词数据集合中各拓展词向量到初始聚类中心的欧式距离，将各拓展词向量归于最近的初始聚类中心构成聚类结果；

当拓展词数据集合中所有拓展词向量聚类结束后，在特征词的词向量对应的各个聚类结果中进行距离均值计算，获取新的聚类中心，当迭代次数达到预设标准后，结束聚类操作；

根据最后一次迭代运算获取各个特征词的词向量对应的聚类结果，在各个聚类结果中，获取对应的拓展词典语义，生成语义的词向量；

根据当前检索场景获取基准词，根据基准词与语义词向量的相似度计算语义的词向量的匹配程度，筛选符合标准的拓展词典向量，并生成词典特征向量。

[0032] 需要说明的是，当拓展词数据集合中各拓展词向量划分结束后，求取聚类结果$M^{(t)}$中每个类簇的均值作为新的聚类中心$C_n^{(t+1)}$，设Z_n为第n个类簇的样本总数，x_{ni}为该类簇的第i个拓展词向量，聚类中心点求取方法具体为：

$$C_n^{(t+1)}=\frac{1}{Z_n}\sum_{i=1}^{Z_n}x_{ni};$$

选定标准测度函数σ_c及最大迭代次数T_{max}，若$|C_n^{(t+1)}-C_n^{(t)}|\le\sigma_c$或迭代次数大于等于$T_{max}$，结束聚类流程，取最后一次运算结果为最终聚类结果，选取最后一次运算结果为最终聚类结果，否则令$t=t+1$，继续迭代聚类，t为迭代次数。

[0033] 根据本发明实施例，将特征词基于图片进行扩展，获取拓展图片向量，生成对应特征向量，具体为：

根据特征词构建检索任务获取拓展图片集合，将拓展图片集合中的拓展图片数据进行预处理；

基于注意力机制优化的ResNet50网络构建图片特征提取模型，将预处理后的拓展图片数据导入图片特征提取模型；

通过卷积获取拓展图片数据的特征，对特征进行平均池化实现特征的压缩，对压缩后的特征进行激励，预测各通道的重要性，并利用注意力机制获取各通道的权重；

对特征通道进行加权，对拓展图片数据的特征进行重新标定，输出拓展图片数据

的特征，根据拓展图片数据的特征获取拓展图片向量，与基准词进行相似度计算获取拓展图片的匹配程度；

筛选符合标准的拓展图片向量，生成图片特征向量。

[0034] 需要说明的是，根据当前检索场景获取基准词，其基准词为固定基准词，例如特征词为“苹果”其对应的语义包括：语义1：一种水果、语义2：一家企业、语义3：一部动画片，当前检索场景的基准词为“西瓜”时，语义123分别和西瓜向量相比，选择相似度最高的语义1作为特征向量；

所述ResNet50网络通过残差学习加速了CNN训练过程，有效避免了梯度消失和梯度爆炸问题，；另外，本发明通过ResNet50作为主干网络，并引入通道注意力机制，提取图片的深度特征。利用欧式距离或余弦计算等相似度计算获取匹配程度，根据匹配程度筛选特征向量。

[0035] 根据本发明实施例，将特征词基于文本进行扩展，获取拓展文本向量，生成对应特征向量，具体为：

获取目标文本的特征词，根据所述特征词的出现频率、位置特征及首次出现到末次出现的距离特征获取目标文本的特征序列；

基于所述特征序列利用相似度进行数据检索，获取拓展文本集合，将拓展文本集合中的拓展文本数据进行预处理；

基于LSTM网络构建文本特征提取模型，将预处理后的拓展文本数据导入文本特征提取模型，设文本数据的长度为T，通过LSTM单元进行特征表示，经过T个时间步骤后，文本特征提取模型输出表征文本特征的隐向量；

通过所述隐向量获取对应拓展文本向量，与基准词进行相似度计算获取拓展图片的匹配程度，筛选符合标准的拓展文本向量，生成文本特征向量。

[0036] 需要说明的是，LSTM利用其特殊的门控结构，主要包括输入门、输出门和遗忘门，有选择地影响每个时刻的状态。输入门是控制当前单元的输入，输出门是控制当前LSTM单元的输出，遗忘门是控制上一时刻单元中存储的历史信息，本发明通过LSTM网络捕捉文本的关键信息；

对生成的词典特征向量、图片特征向量及文本特征向量进行降维，对三个向量进行拼接，获取融合后多模态特征向量；将融合后多模态特征向量与对应的科技成果构建表征三元组，为科技成果推荐生成数据基础。

[0037] 本发明第三方面还提供一种计算机可读存储介质，所述计算机可读存储介质中包括一种基于语义相似度匹配的多模态融合表征方法程序，所述一种基于语义相似度匹配的多模态融合表征方法程序被处理器执行时，实现如上述任一项所述的一种基于语义相似度匹配的多模态融合表征方法的步骤。

[0038] 在本申请所提供的几个实施例中，应该理解到，所揭露的设备和方法，可以通过其它的方式实现。以上所描述的设备实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，如：多个单元或组件可以结合，或可以集成到另一个系统，或一些特征可以忽略，或不执行。另外，所显示或讨论的各组成部分相互之间的耦合、或直接耦合、或通信连接可以是通过一些接口，设备或单元的间接耦合或通信连接，可以是电性的、机械的或其它形式的。

[0039] 上述作为分离部件说明的单元可以是、或也可以不是物理上分开的，作为单元显示的部件可以是、或也可以不是物理单元；既可以位于一个地方，也可以分布到多个网络单元上；可以根据实际的需要选择其中的部分或全部单元来实现本实施例方案的目的。

[0040] 另外，在本发明各实施例中的各功能单元可以全部集成在一个处理单元中，也可以是各单元分别单独作为一个单元，也可以两个或两个以上单元集成在一个单元中；上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能单元的形式实现。

[0041] 本领域普通技术人员可以理解：实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成，前述的程序可以存储于计算机可读取存储介质中，该程序在执行时，执行包括上述方法实施例的步骤；而前述的存储介质包括：移动存储设备、只读存储器(ROM，Read-Only Memory)、随机存取存储器(RAM，Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0042] 或者，本发明上述集成的单元如果以软件功能模块的形式实现并作为独立的产品销售或使用时，也可以存储在一个计算机可读取存储介质中。基于这样的理解，本发明实施例的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备(可以是个人计算机、服务器、或者网络设备等)执行本发明各个实施例所述方法的全部或部分。而前述的存储介质包括：移动存储设备、ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0043] 以上所述，仅为本发明的具体实施方式，但本发明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应以所述权利要求的保护范围为准。

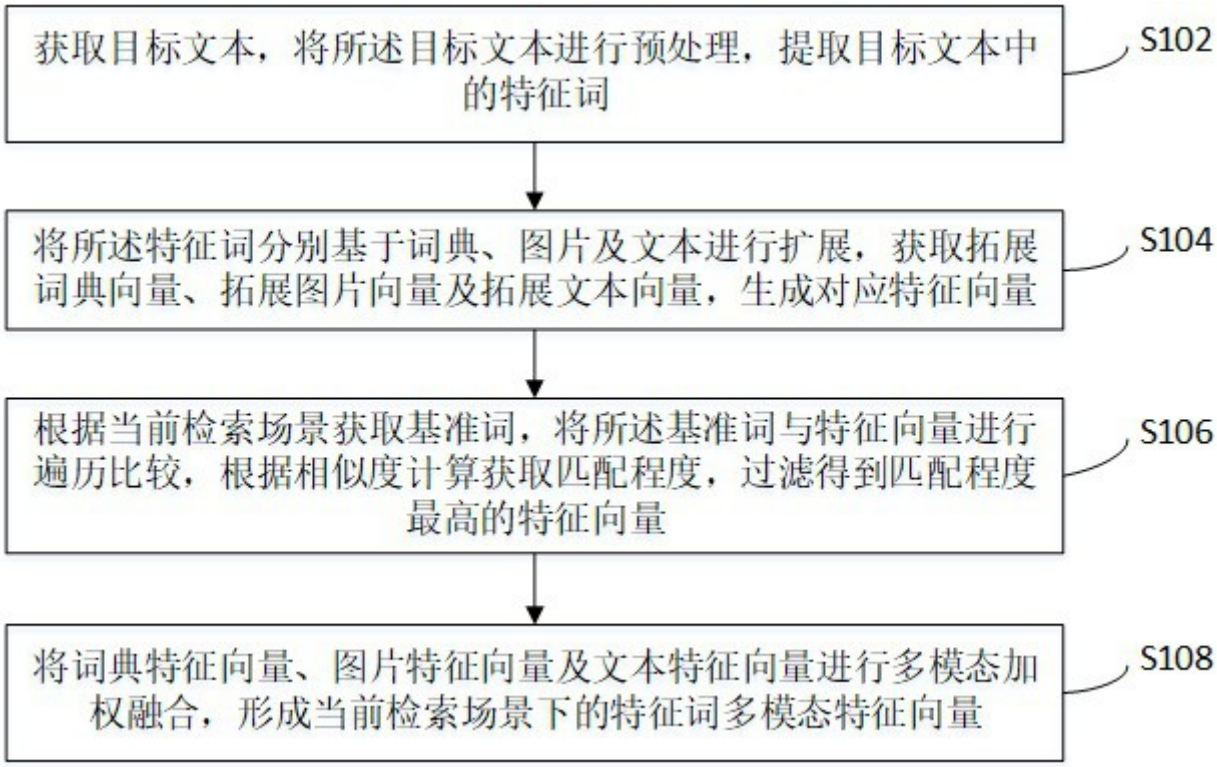

获取目标文本，将所述目标文本进行预处理，提取目标文本中的特征词
S102
将所述特征词分别基于词典、图片及文本进行扩展，获取拓展词典向量、拓展图片向量及拓展文本向量，生成对应特征向量
S104
根据当前检索场景获取基准词，将所述基准词与特征向量进行遍历比较，根据相似度计算获取匹配程度，过滤得到匹配程度最高的特征向量
S106
将词典特征向量、图片特征向量及文本特征向量进行多模态加权融合，形成当前检索场景下的特征词多模态特征向量
S108

图 1

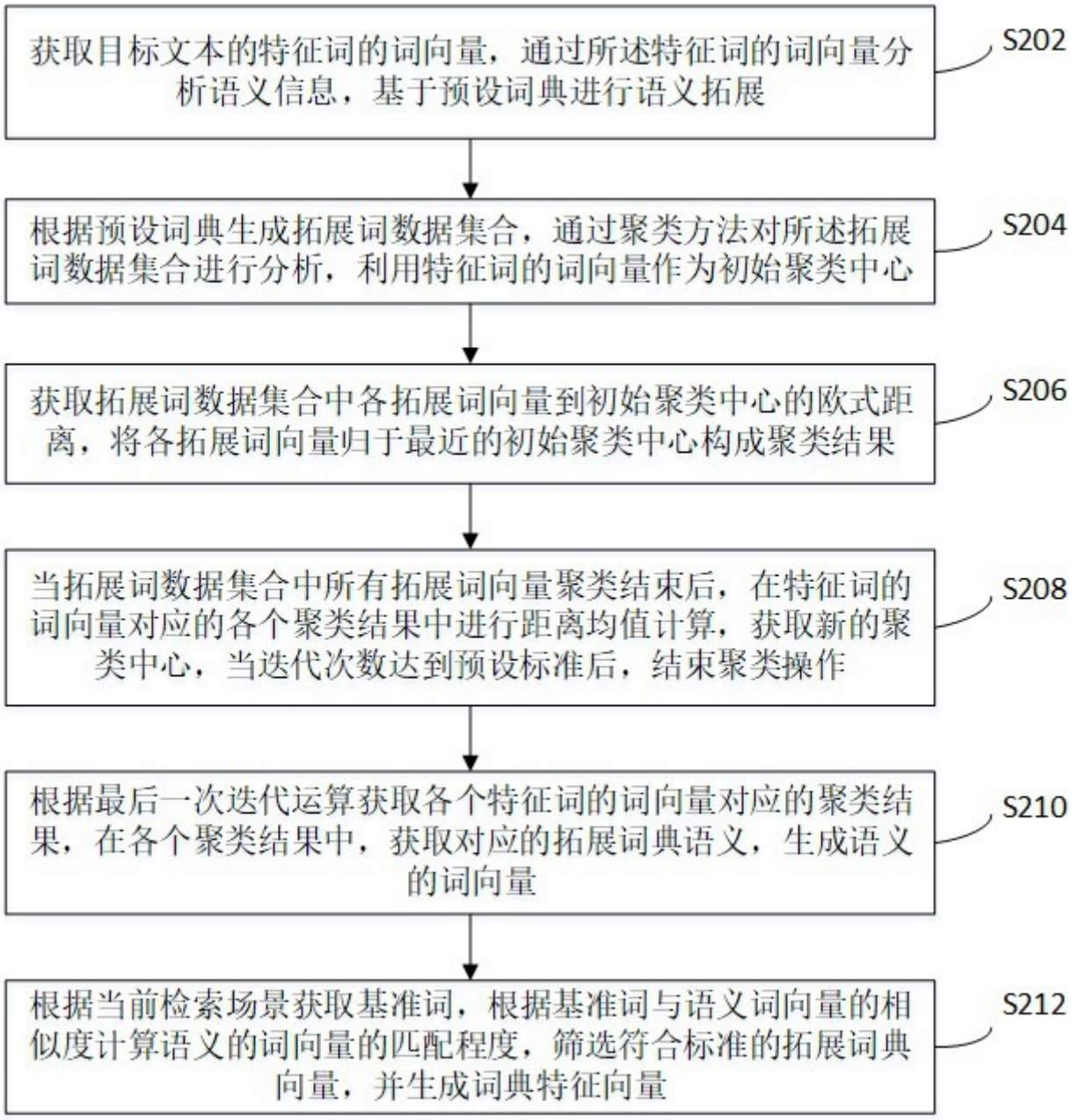

获取目标文本的特征词的词向量，通过所述特征词的词向量分析语义信息，基于预设词典进行语义拓展
S202
根据预设词典生成拓展词数据集合，通过聚类方法对所述拓展词数据集合进行分析，利用特征词的词向量作为初始聚类中心
S204
获取拓展词数据集合中各拓展词向量到初始聚类中心的欧式距离，将各拓展词向量归于最近的初始聚类中心构成聚类结果
S206
当拓展词数据集合中所有拓展词向量聚类结束后，在特征词的词向量对应的各个聚类结果中进行距离均值计算，获取新的聚类中心，当迭代次数达到预设标准后，结束聚类操作
S208
根据最后一次迭代运算获取各个特征词的词向量对应的聚类结果，在各个聚类结果中，获取对应的拓展词典语义，生成语义的词向量
S210
根据当前检索场景获取基准词，根据基准词与语义词向量的相似度计算语义的词向量的匹配程度，筛选符合标准的拓展词典向量，并生成词典特征向量
S212

图 2

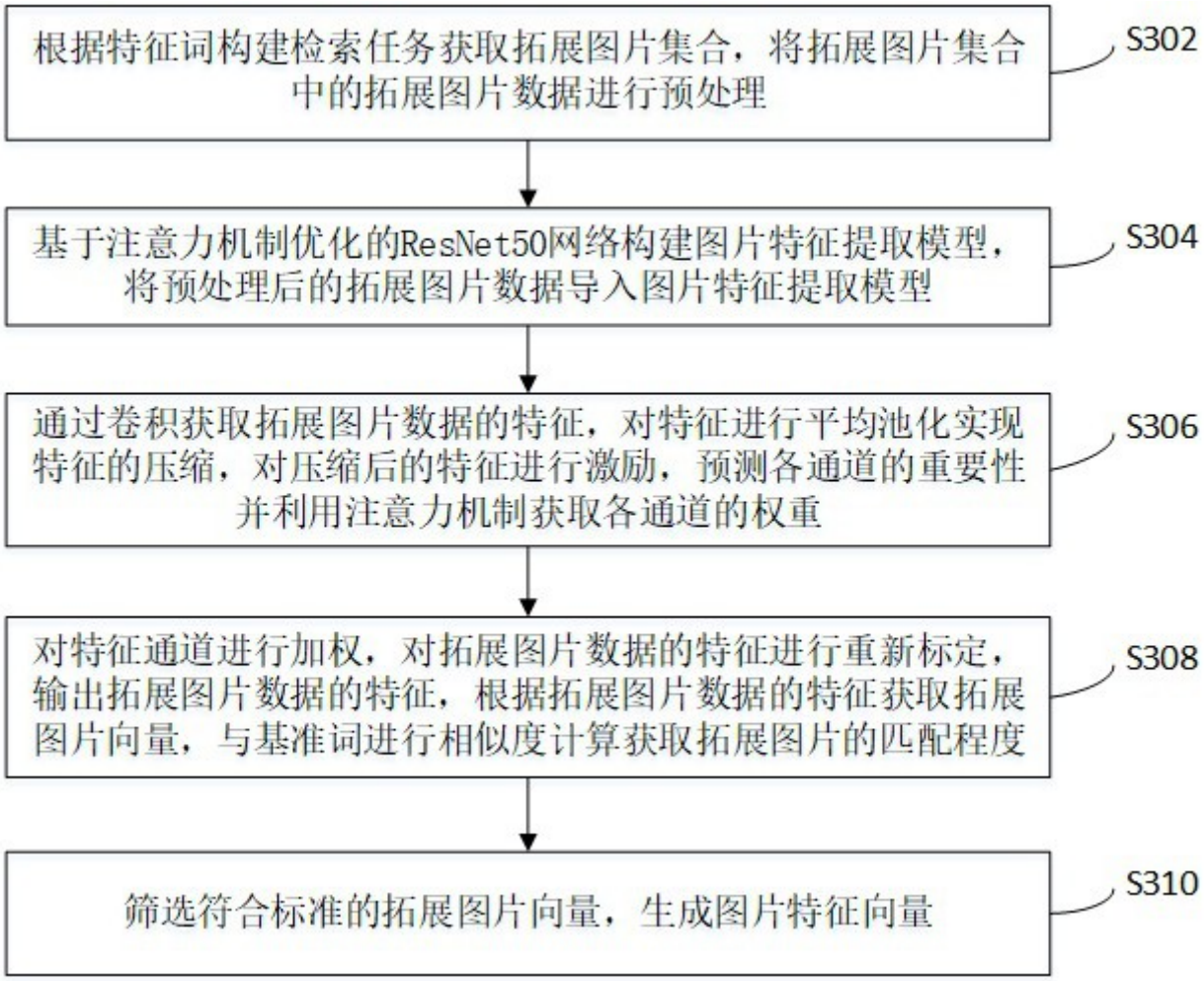

根据特征词构建检索任务获取拓展图片集合，将拓展图片集合中的拓展图片数据进行预处理
S302
基于注意力机制优化的ResNet50网络构建图片特征提取模型，将预处理后的拓展图片数据导入图片特征提取模型
S304
通过卷积获取拓展图片数据的特征，对特征进行平均池化实现特征的压缩，对压缩后的特征进行激励，预测各通道的重要性并利用注意力机制获取各通道的权重
S306
对特征通道进行加权，对拓展图片数据的特征进行重新标定，输出拓展图片数据的特征，根据拓展图片数据的特征获取拓展图片向量，与基准词进行相似度计算获取拓展图片的匹配程度
S308
筛选符合标准的拓展图片向量，生成图片特征向量
S310

图 3

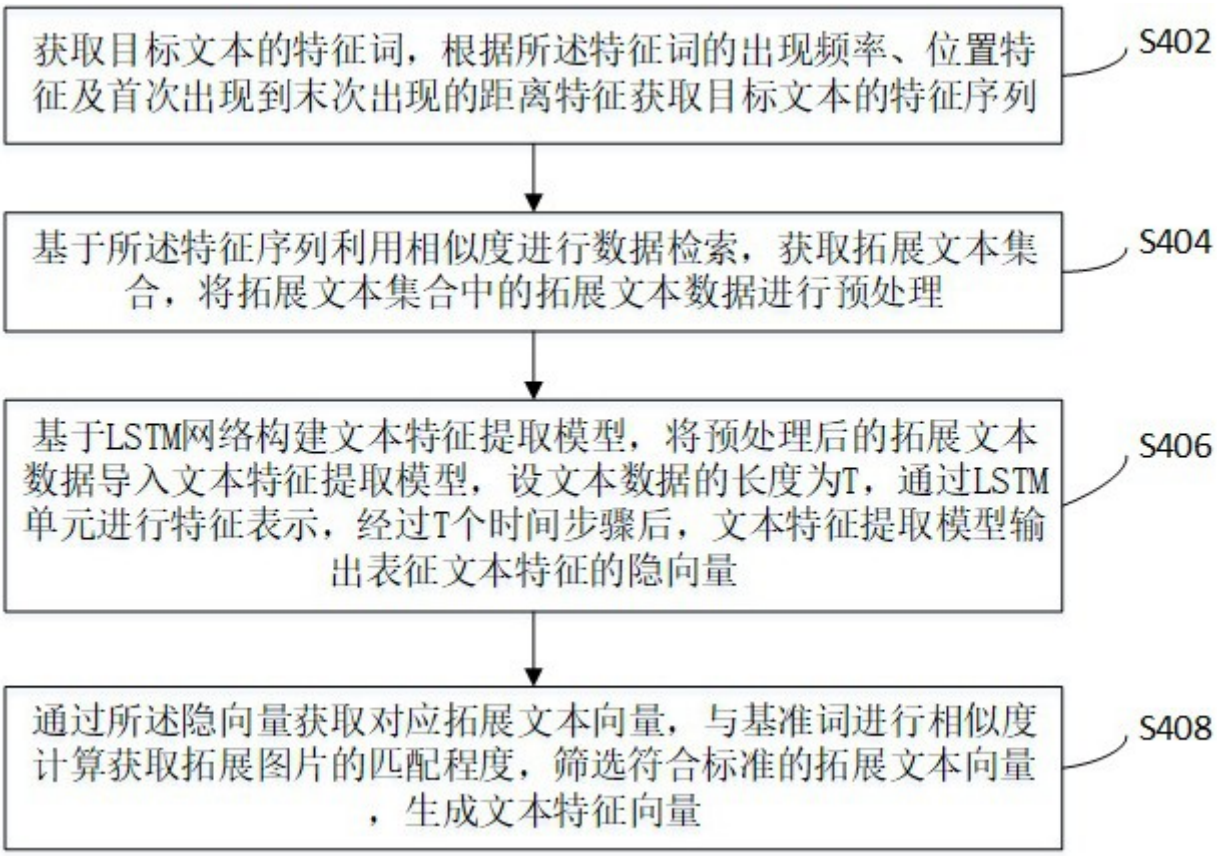

获取目标文本的特征词，根据所述特征词的出现频率、位置特征及首次出现到末次出现的距离特征获取目标文本的特征序列
S402
基于所述特征序列利用相似度进行数据检索，获取拓展文本集合，将拓展文本集合中的拓展文本数据进行预处理
S404
基于LSTM网络构建文本特征提取模型，将预处理后的拓展文本数据导入文本特征提取模型，设文本数据的长度为T，通过LSTM单元进行特征表示，经过T个时间步骤后，文本特征提取模型输出表征文本特征的隐向量
S406
通过所述隐向量获取对应拓展文本向量，与基准词进行相似度计算获取拓展图片的匹配程度，筛选符合标准的拓展文本向量，生成文本特征向量
S408

图 4

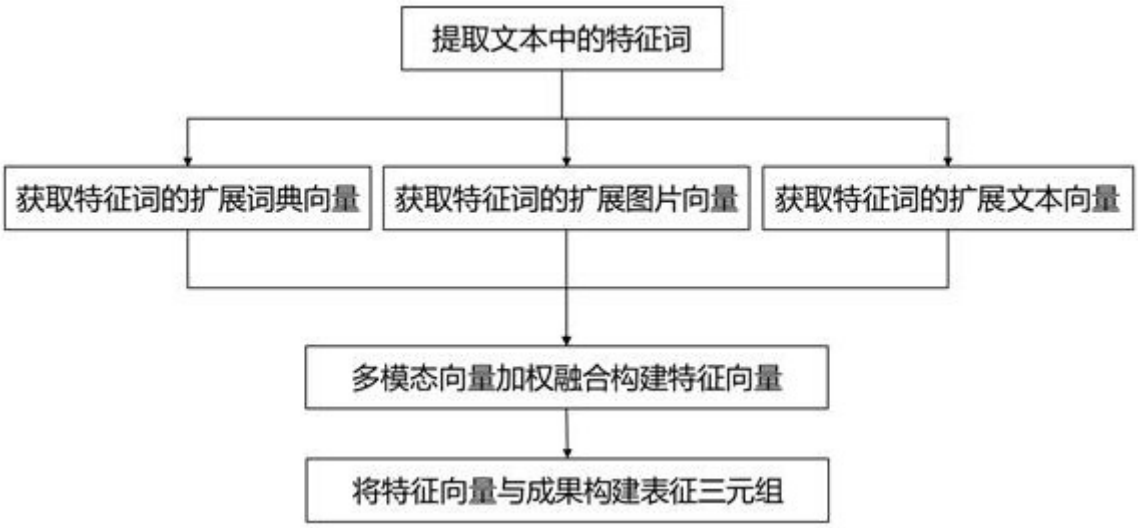

提取文本中的特征词
获取特征词的扩展词典向量
获取特征词的扩展图片向量
获取特征词的扩展文本向量
多模态向量加权融合构建特征向量
将特征向量与成果构建表征三元组

图 5

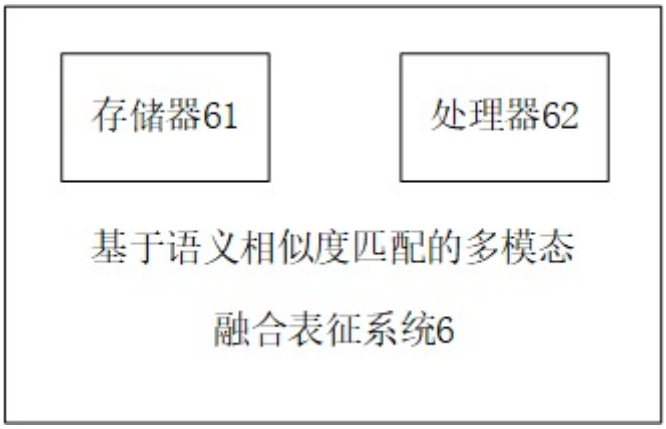


图 6