



(12)发明专利申请

(10)申请公布号 CN 110598037 A

(43)申请公布日 2019.12.20

(21)申请号 201910900623.7

G06N 3/08(2006.01)

(22)申请日 2019.09.23

(71)申请人 腾讯科技(深圳)有限公司

地址 518057 广东省深圳市南山区高新区
科技中一路腾讯大厦35层

(72)发明人 陈震鸿 颜强 赵猛

(74)专利代理机构 深圳翼盛智成知识产权事务
所(普通合伙) 44300

代理人 蔡艾莹

(51)Int.Cl.

G06F 16/583(2019.01)

G06F 16/58(2019.01)

G06F 16/535(2019.01)

G06F 16/538(2019.01)

G06N 3/04(2006.01)

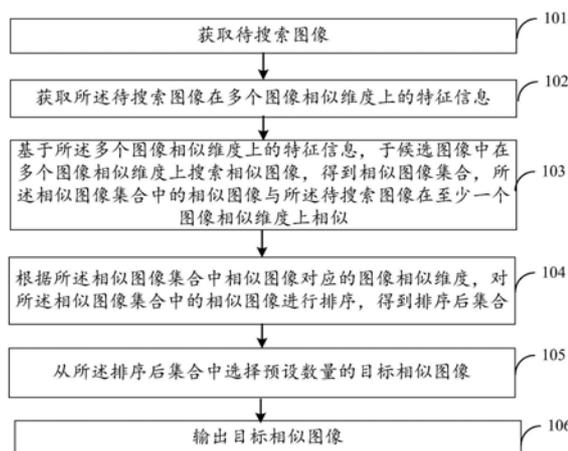
权利要求书3页 说明书27页 附图9页

(54)发明名称

一种图像搜索方法、装置和存储介质

(57)摘要

本申请实施例公开了一种图像搜索方法、装置和存储介质,通过获取待搜索图像;获取待搜索图像在多个图像相似维度上的特征信息;可以基于多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合,相似图像集合中的相似图像与待搜索图像在至少一个图像相似维度上相似;再根据相似图像集合中相似图像对应的图像相似维度,对相似图像集合中的相似图像进行排序,得到排序后集合;从排序后集合中选择预设数量的目标相似图像;输出目标相似图像,通过上述方案,可以对待搜索图像在多个图像相似维度上进行相似图像的搜索,提升搜索结果的多样性,提升图像搜索能力。



1. 一种图像搜索方法,其特征在于,包括:

获取待搜索图像;

获取所述待搜索图像在多个图像相似维度上的特征信息;

基于所述多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合,所述相似图像集合中的相似图像与所述待搜索图像在至少一个图像相似维度上相似;

根据所述相似图像集合中相似图像对应的图像相似维度,对所述相似图像集合中的相似图像进行排序,得到排序后集合;

从所述排序后集合中选择预设数量的目标相似图像;

输出目标相似图像。

2. 根据权利要求1所述的图像搜索方法,其特征在于,所述图像相似维度包括:图像承载对象相似维度、图像语义相似维度以及图像操作行为相似维度;

所述基于所述多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像包括:

基于图像承载对象相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像承载对象相似维度上相似的相似图像;

基于图像语义相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像语义相似维度上相似的相似图像;

基于图像操作行为相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像操作行为相似维度上相似的相似图像。

3. 根据权利要求2所述的图像搜索方法,其特征在于,图像承载对象相似维度包括:图像特征相似子维度;

获取待搜索图像在图像承载对象相似维度上的特征信息,包括:

通过特征提取模型的局部特征提取块,提取所述待搜索图像的局部特征;

通过对应于所述局部特征提取块设置的注意力模块,学习所述局部特征提取块提取的局部特征的特征权重;

基于所述特征权重对所述局部特征提取块提取的局部特征进行处理,得到注意力局部特征,基于所述注意力局部特征得到所述待搜索图像的图像特征;

所述基于图像承载对象相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像承载对象相似维度上相似的相似图像,包括:

获取候选图像基于所述特征提取模型提取的图像特征;

从所述候选图像中,搜索图像特征与所述待搜索图像的图像特征相似的相似图像。

4. 根据权利要求2所述的图像搜索方法,其特征在于,所述图像语义相似维度包括:图像语义相似子维度;

获取待搜索图像在图像语义相似维度上的特征信息,包括:

识别所述待搜索图像的语义信息;

所述基于图像语义相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像语义相似维度上相似的相似图像包括:

获取候选图像的语义信息,从所述候选图像中,搜索语义信息与所述待搜索图像的语

义信息相似的相似图像。

5. 根据权利要求2所述的图像搜索方法,其特征在於,所述图像语义相似维度包括:图像对话关系相似子维度;

获取待搜索图像在图像语义相似维度上的特征信息,包括:

识别所述待搜索图像的语义信息;

所述基于图像语义相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像语义相似维度上相似的相似图像包括:

确定与所述待搜索图像的语义信息在对话关系上相似的目标语义信息;

获取候选图像的语义信息;

从所述候选图像中,搜索语义信息与所述目标语义信息相似的相似图像。

6. 根据权利要求2所述的图像搜索方法,其特征在於,所述图像操作行为相似维度包括:单用户图像操作行为相似子维度;

获取待搜索图像在图像操作行为相似维度上的特征信息,包括:

获取所述待搜索图像在第一历史图像标识号文本中的词向量,其中,所述第一历史图像标识号文本为第一历史操作图像集合中的图像的标识号组成的文本,所述第一历史操作图像集合中的图像为同一用户在历史时间段内进行了操作的图像,且所述第一用户历史操作图像集合包括所述待搜索图像;

所述基于图像操作行为相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像操作行为相似维度上相似的相似图像,包括:

获取候选图像在第二历史图像标识号文本中的词向量,其中,所述第二历史图像标识号文本为第二历史操作图像集合中的图像的标识号组成的文本,所述第二历史操作图像集合中的图像为同一用户在历史时间段内进行了操作的图像,且所述第二用户历史操作图像集合包括所述候选图像;

从所述候选图像中,搜索词向量与所述待搜索图像的词向量相似的相似图像。

7. 根据权利要求2所述的图像搜索方法,其特征在於,所述图像操作行为相似维度包括多用户图像操作行为相似子维度;

获取待搜索图像在图像操作行为相似维度上的特征信息,包括:

获取待搜索图像的认识信息;

所述基于图像操作行为相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像操作行为相似维度上相似的相似图像,包括:

基于所述待搜索图像的认识信息,获取包含待搜索图像的目标二部图对应的图像相似矩阵,其中,所述目标二部图的一边顶点表示用户,另一边顶点表示所述候选图像中所述用户的历史操作图像,所述图像相似矩阵包括所述历史操作图像之间的图像相似度;

在所述图像相似矩阵中,搜索与所述待搜索图像的图像相似度不低于预设相似度阈值的相似图像。

8. 根据权利要求1-7任一项所述的图像搜索方法,其特征在於,所述根据所述相似图像集合中相似图像对应的图像相似维度,对所述相似图像集合中的相似图像进行排序,包括:

根据所述相似图像在对应的图像相似维度下的相似度,以及各图像相似维度对应的相似度权重,计算所述相似图像集合中相似图像的综合相似度;

根据所述相似图像的综合相似度,对所述相似图像集合中的相似图像进行排序。

9. 一种图像搜索装置,其特征在于,包括:

第一获取单元,用于获取待搜索图像;

第二获取单元,用于获取所述待搜索图像在多个图像相似维度上的特征信息;

搜索单元,用于基于所述多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合,所述相似图像集合中的相似图像与所述待搜索图像在至少一个图像相似维度上相似;

排序单元,用于根据所述相似图像集合中相似图像对应的图像相似维度,对所述相似图像集合中的相似图像进行排序,得到排序后集合;

选择单元,用于从所述排序后集合中选择预设数量的目标相似图像;

输出单元,用于输出目标相似图像。

10. 一种存储介质,其上存储有计算机程序,其特征在于,当计算机程序在计算机上运行时,使得所述计算机执行如权利要求1至8中任一项所述的图像搜索方法。

一种图像搜索方法、装置和存储介质

技术领域

[0001] 本申请涉及计算机技术领域,具体涉及一种图像搜索方法、装置和存储介质。

背景技术

[0002] 目前,以图搜图作为一种图像搜索功能,在一些即时通讯软件和搜索引擎中被提供给用户使用,对于以图搜图,需要能够根据输入的图像搜索目标图像。

发明内容

[0003] 有鉴于此,本申请实施例提供了一种图像搜索方法、装置和存储介质,可以提升图像搜索的搜索能力以及提升搜索结果的多样性。

[0004] 本申请实施例提供了一种图像搜索方法,该方法包括:

[0005] 获取待搜索图像;

[0006] 获取所述待搜索图像在多个图像相似维度上的特征信息;

[0007] 基于所述多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合,所述相似图像集合中的相似图像与所述待搜索图像在至少一个图像相似维度上相似;

[0008] 根据所述相似图像集合中相似图像对应的图像相似维度,对所述相似图像集合中的相似图像进行排序,得到排序后集合;

[0009] 从所述排序后集合中选择预设数量的目标相似图像;

[0010] 输出目标相似图像。

[0011] 在一个实施例中,所述图像相似维度包括:图像承载对象相似维度、图像语义相似维度以及图像操作行为相似维度;

[0012] 所述基于所述多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像包括:

[0013] 基于图像承载对象相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像承载对象相似维度上相似的相似图像;

[0014] 基于图像语义相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像语义相似维度上相似的相似图像;

[0015] 基于图像操作行为相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像操作行为相似维度上相似的相似图像。

[0016] 在一个实施例中,图像承载对象相似维度包括:图像特征相似子维度;

[0017] 获取待搜索图像在图像承载对象相似维度上的特征信息,包括:

[0018] 通过特征提取模型的局部特征提取块,提取所述待搜索图像的局部特征;

[0019] 通过对应于所述局部特征提取块设置的注意力模块,学习所述局部特征提取块提取的局部特征的特征权重;

[0020] 基于所述特征权重对所述局部特征提取块提取的局部特征进行处理,得到注意力

局部特征,基于所述注意力局部特征得到所述待搜索图像的图像特征;

[0021] 所述基于图像承载对象相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像承载对象相似维度上相似的相似图像,包括:

[0022] 获取候选图像基于所述特征提取模型提取的图像特征;

[0023] 从所述候选图像中,搜索图像特征与所述待搜索图像的图像特征相似的相似图像。

[0024] 在一个实施例中,所述图像承载对象相似维度包括:人脸相似子维度;

[0025] 获取待搜索图像在图像承载对象相似维度上的特征信息,包括:

[0026] 检测所述待搜索图像中是否存在人脸;

[0027] 若检测到所述待搜索图像中存在人脸,则通过人脸特征提取模型对所述待搜索图像进行人脸特征的提取;

[0028] 所述基于图像承载对象相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像承载对象相似维度上相似的相似图像,包括:

[0029] 获取候选图像基于所述人脸特征提取模型提取的人脸特征;

[0030] 从所述候选图像中,搜索人脸特征与所述待搜索图像的人脸特征相似的相似图像。

[0031] 在一个实施例中,所述图像语义相似维度包括图像语义相似子维度;

[0032] 获取待搜索图像在图像语义相似维度上的特征信息,包括:

[0033] 识别所述待搜索图像的语义信息;

[0034] 所述基于图像语义相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像语义相似维度上相似的相似图像包括:

[0035] 获取候选图像的语义信息,从所述候选图像中,搜索语义信息与所述待搜索图像的语义信息相似的相似图像。

[0036] 在一个实施例中,所述图像语义相似维度包括图像对话关系相似子维度;

[0037] 获取待搜索图像在图像语义相似维度上的特征信息,包括:

[0038] 识别所述待搜索图像的语义信息;

[0039] 所述基于图像语义相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像语义相似维度上相似的相似图像包括:

[0040] 确定与所述待搜索图像的语义信息在对话关系上相似的目标语义信息;

[0041] 获取候选图像的语义信息;

[0042] 从所述候选图像中,搜索语义信息与所述目标语义信息相似的相似图像。

[0043] 在一个实施例中,所述图像语义相似子维度包括:图像字符相似次级子维度;

[0044] 所述识别所述待搜索图像的语义信息,包括:

[0045] 通过图像文字识别技术,识别所述待搜索图像上记载的文字信息;

[0046] 所述获取候选图像的语义信息,从所述候选图像中,搜索语义信息与所述待搜索图像的语义信息相似的相似图像包括:

[0047] 获取候选图像上记载的文字信息,从所述候选图像中,搜索文字信息与所述待搜索图像的文字信息相似的相似图像。

[0048] 在一个实施例中,所述图像语义相似子维度包括:图像标签相似次级子维度;

- [0049] 所述识别所述待搜索图像的语义信息,包括:
- [0050] 若所述待搜索图像存在标签,则识别所述待搜索图像的标签;
- [0051] 若所述待搜索图像不存在标签,则通过标签分类模型对所述待搜索图像的标签进行预测,根据预测结果获取所述待搜索图像对应的标签;
- [0052] 所述获取候选图像的语义信息,从所述候选图像中,搜索语义信息与所述待搜索图像的语义信息相似的相似图像,包括:
- [0053] 获取所述候选图像的标签;
- [0054] 从所述候选图像中,搜索标签与所述待搜索图像的标签相似的相似图像。
- [0055] 在一个实施例中,所述图像语义相似子维度包括:图像描述信息相似次级子维度;
- [0056] 所述识别所述待搜索图像的语义信息,包括:
- [0057] 通过图像理解模型对所述待搜索图像进行分析,以生成所述待搜索图像的图像描述信息;
- [0058] 所述获取候选图像的语义信息,从所述候选图像中,搜索语义信息与所述待搜索图像的语义信息相似的相似图像,包括:
- [0059] 获取所述候选图像基于所述图像理解模型生成的图像描述信息;
- [0060] 从所述候选图像中,搜索图像描述信息与所述待搜索图像的图像描述信息相似的相似图像。
- [0061] 在一个实施例中,所述图像操作行为相似维度包括:单用户图像操作行为相似子维度;
- [0062] 获取待搜索图像在图像操作行为相似维度上的特征信息,包括:
- [0063] 获取所述待搜索图像在第一历史图像标识号文本中的词向量,其中,所述第一历史图像标识号文本为第一历史操作图像集合中的图像的标识号组成的文本,所述第一历史操作图像集合中的图像为同一用户在历史时间段内进行了操作的图像,且所述第一用户历史操作图像集合包括所述待搜索图像;
- [0064] 所述基于图像操作行为相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像操作行为相似维度上相似的相似图像,包括:
- [0065] 获取候选图像在第二历史图像标识号文本中的词向量,其中,所述第二历史图像标识号文本为第二历史操作图像集合中的图像的标识号组成的文本,所述第二历史操作图像集合中的图像为同一用户在历史时间段内进行了操作的图像,且所述第二用户历史操作图像集合包括所述候选图像;
- [0066] 从所述候选图像中,搜索词向量与所述待搜索图像的词向量相似的相似图像。
- [0067] 在一个实施例中,所述图像操作行为相似维度包括多用户图像操作行为相似子维度;
- [0068] 获取待搜索图像在图像操作行为相似维度上的特征信息,包括:
- [0069] 获取待搜索图像的识别信息;
- [0070] 所述基于图像操作行为相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像操作行为相似维度上相似的相似图像,包括:
- [0071] 基于所述待搜索图像的识别信息,获取包含待搜索图像的目标二部图对应的图像相似度矩阵,其中,所述目标二部图的一边顶点表示用户,另一边顶点表示所述候选图像中

所述用户的历史操作图像,所述图像相似度矩阵包括所述历史操作图像之间的图像相似度;

[0072] 在所述图像相似度矩阵中,搜索与所述待搜索图像的图像相似度不低于预设相似度阈值的相似图像。

[0073] 在一个实施例中,所述根据所述相似图像集合中相似图像对应的图像相似维度,对所述相似图像集合中的相似图像进行排序,包括:

[0074] 根据所述相似图像在对应的图像相似维度下的相似度,以及各图像相似维度对应的相似度权重,计算所述相似图像集合中相似图像的综合相似度;

[0075] 根据所述相似图像的综合相似度,对所述相似图像集合中的相似图像进行排序。

[0076] 在一个实施例中,该方法还包括:

[0077] 在用户历史操作行为数据的累积量满足预设条件时,不执行所述根据所述相似图像集合中相似图像对应的图像相似维度,对所述相似图像集合中的相似图像进行排序,得到排序后集合的步骤;其中,所述用户历史操作行为数据包括:基于所述待搜索图像进行图像搜索的用户,在历史时间段内进行图像搜索产生的行为数据;

[0078] 根据所述用户历史操作行为数据,分析用户对所述相似图像集合中相似图像的用户偏好,基于所述用户偏好对所述相似图像进行排序,得到排序后集合。

[0079] 本实施例还提供一种图像搜索装置,该装置包括:

[0080] 第一获取单元,用于获取待搜索图像;

[0081] 第二获取单元,用于获取所述待搜索图像在多个图像相似维度上的特征信息;

[0082] 搜索单元,用于基于所述多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合,所述相似图像集合中的相似图像与所述待搜索图像在至少一个图像相似维度上相似;

[0083] 排序单元,用于根据所述相似图像集合中相似图像对应的图像相似维度,对所述相似图像集合中的相似图像进行排序,得到排序后集合;

[0084] 选择单元,用于从所述排序后集合中选择预设数量的目标相似图像;

[0085] 输出单元,用于输出目标相似图像。

[0086] 在一个实施例中,图像相似维度包括:图像承载对象相似维度、图像语义相似维度以及图像操作行为相似维度;

[0087] 搜索单元包括:

[0088] 第一搜索子单元,用于基于图像承载对象相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像承载对象相似维度上相似的相似图像;

[0089] 第二搜索子单元,用于基于图像语义相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像语义相似维度上相似的相似图像;

[0090] 第三搜索子单元,用于基于图像操作行为相似维度上的特征信息,于候选图像中搜索与所述待搜索图像在图像操作行为相似维度上相似的相似图像。

[0091] 在一个实施例中,图像承载对象相似维度包括:图像特征相似子维度;

[0092] 第二获取单元,包括:

[0093] 图像特征提取子单元,用于通过特征提取模型的局部特征提取块,提取所述待搜索图像的局部特征;

- [0094] 学习子单元,用于通过对应于所述局部特征提取块设置的注意力模块,学习所述局部特征提取块提取的局部特征的特征权重;
- [0095] 处理子单元,用于基于所述特征权重对所述局部特征提取块提取的局部特征进行处理,得到注意力局部特征,基于所述注意力局部特征得到所述待搜索图像的图像特征;
- [0096] 第一搜索子单元,用于:
- [0097] 获取候选图像基于所述特征提取模型提取的图像特征;
- [0098] 从所述候选图像中,搜索图像特征与所述待搜索图像的图像特征相似的相似图像。
- [0099] 在一个实施例中,所述图像承载对象相似维度包括:人脸相似子维度;
- [0100] 第二获取单元,包括:
- [0101] 人脸检测子单元,用于检测所述待搜索图像中是否存在人脸;
- [0102] 人脸特征提取子单元,用于若检测到所述待搜索图像中存在人脸,则通过人脸特征提取模型对所述待搜索图像进行人脸特征的提取;
- [0103] 第一搜索子单元,用于:
- [0104] 获取候选图像基于所述人脸特征提取模型提取的人脸特征;
- [0105] 从所述候选图像中,搜索人脸特征与所述待搜索图像的人脸特征相似的相似图像。
- [0106] 在一实施例中,所述图像语义相似维度包括图像语义相似子维度;
- [0107] 第二获取单元,用于识别所述待搜索图像的语义信息;
- [0108] 第二搜索子单元,用于获取候选图像的语义信息,从所述候选图像中,搜索语义信息与所述待搜索图像的语义信息相似的相似图像。
- [0109] 在一实施例中,所述图像语义相似维度包括图像对话关系相似子维度;
- [0110] 第二获取单元,用于识别所述待搜索图像的语义信息;
- [0111] 第二搜索子单元,用于:
- [0112] 确定与所述待搜索图像的语义信息在对话关系上相似的目标语义信息;
- [0113] 获取候选图像的语义信息;
- [0114] 从所述候选图像中,搜索语义信息与所述目标语义信息相似的相似图像。
- [0115] 在一实施例中,所述图像语义相似子维度包括:图像字符相似次级子维度;
- [0116] 第二获取单元,用于通过图像文字识别技术,识别所述待搜索图像上记载的文字信息;
- [0117] 第二搜索子单元,用于获取候选图像上记载的文字信息,从所述候选图像中,搜索文字信息与所述待搜索图像的文字信息相似的相似图像。
- [0118] 在一实施例中,所述图像语义相似子维度包括:图像标签相似次级子维度;
- [0119] 第二获取单元,用于若所述待搜索图像存在标签,则识别所述待搜索图像的标签;若所述待搜索图像不存在标签,则通过标签分类模型对所述待搜索图像的标签进行预测,根据预测结果获取所述待搜索图像对应的标签;
- [0120] 第二搜索子单元,用于:
- [0121] 获取所述候选图像的标签;
- [0122] 从所述候选图像中,搜索标签与所述待搜索图像的标签相似的相似图像。

- [0123] 在一实施例中,所述图像语义相似子维度包括:图像描述信息相似次级子维度;
- [0124] 第二获取单元,用于通过图像理解模型对所述待搜索图像进行分析,以生成所述待搜索图像的图像描述信息;
- [0125] 第二搜索子单元,用于:
- [0126] 获取所述候选图像基于所述图像理解模型生成的图像描述信息;
- [0127] 从所述候选图像中,搜索图像描述信息与所述待搜索图像的图像描述信息相似的相似图像。
- [0128] 在一实施例中,所述图像操作行为相似维度包括:单用户图像操作行为相似子维度;
- [0129] 第二获取单元,用于获取所述待搜索图像在第一历史图像标识号文本中的词向量,其中,所述第一历史图像标识号文本为第一历史操作图像集合中的图像的标识号组成的文本,所述第一历史操作图像集合中的图像为同一用户在历史时间段内进行了操作的图像,且所述第一用户历史操作图像集合包括所述待搜索图像;
- [0130] 第三搜索子单元,用于:
- [0131] 获取候选图像在第二历史图像标识号文本中的词向量,其中,所述第二历史图像标识号文本为第二历史操作图像集合中的图像的标识号组成的文本,所述第二历史操作图像集合中的图像为同一用户在历史时间段内进行了操作的图像,且所述第二用户历史操作图像集合包括所述候选图像;
- [0132] 从所述候选图像中,搜索词向量与所述待搜索图像的词向量相似的相似图像。
- [0133] 在一实施例中,所述图像操作行为相似维度包括多用户图像操作行为相似子维度;
- [0134] 第二获取单元,用于获取待搜索图像的认识信息;
- [0135] 第三搜索子单元,用于:
- [0136] 基于所述待搜索图像的认识信息,获取包含待搜索图像的目标二部图对应的图像相似度矩阵,其中,所述目标二部图的一边顶点表示用户,另一边顶点表示所述候选图像中所述用户的历史操作图像,所述图像相似度矩阵包括所述历史操作图像之间的图像相似度;
- [0137] 在所述图像相似度矩阵中,搜索与所述待搜索图像的图像相似度不低于预设相似度阈值的相似图像。
- [0138] 在一实施例中,排序单元,包括:
- [0139] 加权子单元,用于根据所述相似图像在对应的图像相似维度下的相似度,以及各图像相似维度对应的相似度权重,计算所述相似图像集合中相似图像的综合相似度;
- [0140] 排序子单元,用于根据所述相似图像的综合相似度,对所述相似图像集合中的相似图像进行排序。
- [0141] 在一实施例中,排序单元,还用于:
- [0142] 在用户历史操作行为数据的累积量满足预设条件时,不执行所述根据所述相似图像集合中相似图像对应的图像相似维度,对所述相似图像集合中的相似图像进行排序,得到排序后集合的步骤;其中,所述用户历史操作行为数据包括:基于所述待搜索图像进行图像搜索的用户,在历史时间段内进行图像搜索产生的行为数据;

[0143] 根据所述用户历史操作行为数据,分析用户对所述相似图像集合中相似图像的用户偏好,基于所述用户偏好对所述相似图像进行排序,得到排序后集合。

[0144] 本发明实施例还提供一种存储介质,其上存储有计算机程序,当计算机程序在计算机上运行时,使得所述计算机执行如本发明实施例所述的图像搜索方法。

[0145] 本发明实施例公开了一种图像搜索方法、装置和存储介质,通过获取待搜索图像;获取待搜索图像在多个图像相似维度上的特征信息;可以基于多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合,相似图像集合中的相似图像与待搜索图像在至少一个图像相似维度上相似;再根据相似图像集合中相似图像对应的图像相似维度,对相似图像集合中的相似图像进行排序,得到排序后集合;从排序后集合中选择预设数量的目标相似图像;输出目标相似图像,通过上述方案,可以对待搜索图像在多个图像相似维度上进行相似图像的搜索,提升搜索结果的多样性,提升图像搜索能力。

附图说明

[0146] 为了更清楚地说明本发明实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的一些实施例,对于本领域技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0147] 图1a是本发明实施例提供的图像搜索系统的结构示意图;

[0148] 图1b是本发明实施例提供的图像搜索方法的流程图;

[0149] 图2a是本发明实施例提供的一种特征提取模型的示意图;

[0150] 图2b是图2a中的局部特征提取块改进后的结构示意图;

[0151] 图2c是本发明实施例提供的一种人脸特征识别的流程图;

[0152] 图3是本发明实施例提供的一种标签分类模型的示意图;

[0153] 图4a是本发明实施例提供的一种图像搜索方法的流程图;

[0154] 图4b是本发明实施例提供的一种图像搜索页面的显示示意图;

[0155] 图4c是本发明实施例提供的目标相似图像的显示示意图;

[0156] 图4d是本发明实施例提供的图像搜索的框架图;

[0157] 图5是本发明实施例提供的图像搜索装置的结构示意图;

[0158] 图6是本发明实施例提供的计算机设备的结构示意图;

[0159] 图7是本发明实施例提供的分布式系统100应用于区块链系统的一个可选的结构示意图;

[0160] 图8是本发明实施例提供的区块链结构的一个可选的示意图。

具体实施方式

[0161] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0162] 本发明实施例提供一种图像搜索方法、装置、计算机设备和存储介质。具体地，本发明实施例提供适用于第一计算机设备的图像搜索装置(为了区分可以称为第一图像搜索装置)中，以及适用于第二计算机设备的图像搜索装置(为了区分可以称为第二图像搜索装置)。其中，第一计算机设备可以为终端等设备，该终端可以为手机、平板电脑、笔记本电脑等设备。其中，第二计算机设备可以为服务器等网络侧设备。

[0163] 比如，第一图像搜索装置可以集成在终端中，第二图像搜索装置可以集成在服务器，该服务器可以是单台服务器，也可以是由多个服务器组成的服务器集群。

[0164] 本发明实施例将以第一计算机设备为终端，第二计算机设备为服务器为例，来介绍图像搜索方法。

[0165] 参考图1a，本发明实施例提供了图像搜索系统包括终端10和服务器20等；终端10与服务器20之间通过网络连接，比如，通过有线或无线网络连接等，其中，图像搜索装置集成在终端中，比如，以客户端的形式集成在终端中。

[0166] 其中，终端10，可以用于显示应用程序客户端的图像搜索页面，图像搜索页面为供用户以图像搜索图像的搜索页面，其中，图像搜索页面包括：待搜索图像选择控件；当检测到用户针对待搜索图像选择控件的触发操作时，显示待搜索图像的选择页面；基于用户针对选择页面的待搜索图像确定操作，显示搜索结果列表，搜索结果列表包括搜索到的与待搜索图像相似的目标相似图像。

[0167] 其中，终端10触发搜索相似图像的方式有多种，比如，终端10可以触发服务器20在多个图像相似维度上搜索与待搜索图像相似的目标相似图像。具体地，终端10可以向服务器20发送待搜索图像，以触发服务器对相似图像的搜索；服务器20可以具体用于：

[0168] 获取待搜索图像；获取待搜索图像在多个图像相似维度上的特征信息；基于多个图像相似维度上的特征信息，于候选图像中在多个图像相似维度上搜索相似图像，得到相似图像集合，相似图像集合中的相似图像与待搜索图像在至少一个图像相似维度上相似；根据相似图像集合中相似图像对应的图像相似维度，对相似图像集合中的相似图像进行排序，得到排序后集合；从排序后集合中选择预设数量的目标相似图像；输出目标相似图像。

[0169] 其中，服务器输出目标图像具体可以为服务器向终端10发送目标相似图像。

[0170] 本发明中，可选的，在服务器上执行的对于目标相似图像的具体搜索过程，也可以由终端执行。

[0171] 以下分别进行详细说明。需说明的是，以下实施例的描述顺序不作为对实施例优选顺序的限定。

[0172] 本发明实施例将从图像搜索装置的角度进行描述，该图像搜索装置具体可以集成在终端或服务器中。

[0173] 本发明实施例提供的一种图像搜索方法，该方法可以由终端的处理器执行，如图1b所示，该图像搜索方法的具体流程可以如下：

[0174] 101、获取待搜索图像；

[0175] 对于服务器侧而言，该待搜索图像可以是终端发送的，具体的，步骤“获取待搜索图像”可以包括：接收终端发送的待搜索图像。

[0176] 对于终端侧而言，可以基于用户针对选择页面的待搜索图像确定操作，从本地存储的图像(包括但不限于通过网络下载的图像以及终端拍摄的图像)中获取待搜索图像，或

者实时拍摄照片作为待搜索图像。

[0177] 可选的,本实施例中,提供图像搜索功能的应用程序客户端包括但不限于即时通讯客户端和图像搜索引擎,待搜索图像可以包括应用在即时通讯客户端中的互动图像。

[0178] 其中,待搜索图像的图像类型可以有多种,例如可以是静态图像或动态图像,可以是纯图案的图像,也可以是包含文字和图案的图像。譬如,待搜索图像可以包括静态表情、动态表情等聊天互动表情。

[0179] 102、获取待搜索图像在多个图像相似维度上的特征信息;

[0180] 本实施例中,特征信息可以表示待搜索图像的特征,其信息类型不限,可以是特征向量、或文本信息等等。

[0181] 本实施例中,不同图像相似维度上,待搜索图像的特征信息可能不同,为了提取特征信息,可以基于各图像相似维度的特征信息提取需要,对待搜索图像进行预处理。

[0182] 可选的,步骤“获取待搜索图像在多个图像相似维度上的特征信息”,可以包括:

[0183] 基于待搜索图像在多个图像相似维度上的特征提取需要,对待搜索图像进行预处理;

[0184] 对预处理之后待搜索图像,提取在多个图像相似维度上的特征信息。

[0185] 其中,对待搜索图像的预处理包括但不限于:对于动态图像,以动态图像中的预设帧图像代替帧数不固定的动态图像;对待搜索图像进行格式转换,将待搜索图像的格式转换为目标格式;根据不同的特征提取算法的需要,对待搜索图像进行等比例缩放、裁剪、以及生成灰度图等操作。

[0186] 其中,预设帧可以根据需要选择,例如可以是第一帧、或中间帧、或最后一帧等等,本实施例对此没有限定。

[0187] 其中,目标格式可以根据实际需要选择,例如目标格式为PNG格式、或JPEG格式等等。

[0188] 本实施例的图像相似维度的数量为多个,比如,图像相似维度,可以包括图像承载对象相似维度、图像语义相似维度以及图像操作行为相似维度。

[0189] 其中,在图像承载对象相似维度上相似,可以理解为图像的承载对象相似,图像的承载对象包括可以是图像中的所有对象,包括主体对象和非主体对象。待搜索图像在图像承载对象相似维度上的特征信息,可以为待搜索图像的承载对象的特征信息。在一个实施例中,待搜索图像的承载对象可以是,待搜索图像上的主体对象。

[0190] 其中,在图像语义相似维度上相似,可以理解为图像的语义信息相似,或者图像的语义信息对应的对话关系相似。待搜索图像在图像语义相似维度上的特征信息,可以为待搜索图像的语义信息。在一个实施例中,待搜索图像的语义信息可以是直接或间接从待搜索图像得到的信息,例如,待搜索图像的语义信息,可以为待搜索图像上的字符信息,可以为待搜索图像的标签,可以为对待搜索图像进行“看图说话”,得到的待搜索图像的图像描述信息等。

[0191] 其中,在图像操作行为相似维度上相似,可以理解为用户针对图像的操作行为相似,其中,用户针对图像的操作行为包括但不限于用户对图像的搜索行为,用户从搜索到的图像中选择图像的选择行为,以及用户将图像添加至图像收藏集合中的收藏行为等等。待搜索图像在图像操作行为相似维度上的特征信息,可以为待搜索图像对应的可以体现用户

搜索意向的信息。

[0192] 103、基于多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合,相似图像集合中的相似图像与待搜索图像在至少一个图像相似维度上相似;

[0193] 其中,候选图像的图像类型也可以有多种,例如可以是静态图像或动态图像,可以是纯图案的图像,也可以是包含文字和图案的图像。譬如,候选图像可以包括静态表情、动态表情等聊天互动表情。

[0194] 本实施例中,候选图像的可以包括多种来源,例如,候选图像来自服务器(和/或终端)的图像数据库、候选图像来自于与自服务器(和/或终端)连接的其他装置中的图像数据库。对于聊天互动表情而言,该图像数据库可以为存储全量表情的表情库。

[0195] 在一个示例中,对于各候选图像,可以分别对应存储有相似图像信息集合,该相似图像信息集合包括:与候选图像相似的候选图像的认识信息,以及对应的图像相似维度。

[0196] 例如,可以通过对各候选图像,在多个图像相似维度上,与其他所有候选图像进行相似度分析,确定与各候选图像分别相似的候选图像,以及对应的图像相似维度,从而得到各候选图像的相似图像信息集合。

[0197] 例如,还可以在每一次图像搜索过程结束后,将待搜索图像添加进候选图像的集合中(若候选图像中存在待搜索图像,则不必添加),基于该次图像搜索过程中待搜索图像对应的搜索结果,确定待搜索图像的相似图像信息集合。

[0198] 对于存在相似图像信息集合的情况,待搜索图像在多个图像相似维度上的特征信息可以是待搜索图像的认识信息。其中,步骤“基于多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合”,可以包括:

[0199] 基于待搜索图像的认识信息,确定待搜索图像的相似图像信息集合;

[0200] 基于相似图像信息集合中的信息,确定候选图像中在多个图像相似维度上分别与待搜索图像相似的相似图像,得到相似图像集合。

[0201] 104、根据相似图像集合中相似图像对应的图像相似维度,对相似图像集合中的相似图像进行排序,得到排序后集合;

[0202] 在一个实施例中,在步骤103,候选图像中在多个图像相似维度上搜索相似图像时,可以确定各个相似图像在对应的相似图像维度下,与待搜索图像的相似度。

[0203] 本实施例可以为不同的相似图像维度设置对应的相似度权重,以相似度权重和相似度综合评价各个相似图像与待搜索图像的综合相似度。例如,步骤“根据相似图像集合中相似图像对应的图像相似维度,对相似图像集合中的相似图像进行排序”,可以包括:

[0204] 根据相似图像在对应的图像相似维度下的相似度,以及各图像相似维度对应的相似度权重,计算相似图像集合中相似图像的综合相似度;

[0205] 根据相似图像的综合相似度,对相似图像集合中的相似图像进行排序。

[0206] 例如,假设多个图像相似维度为图像相似维度1-3,相似度权重分别为a-c,假设相似图像X对应的图像相似维度为图像相似维度1和2,相似图像X在图像相似维度1和2下,与待搜索图像的相似度分别为A和B,则相似图像X的综合相似度为 $A*a+B*b$ 。

[0207] 其中,对相似图像集合中的相似图像进行排序,可以是对相似图像集合中的相似图像按照综合相似度进行降序排列。

[0208] 105、从排序后集合中选择预设数量的目标相似图像；

[0209] 其中，预设数量可以根据实际的需要设置，例如可以设置为300等数值。从排序后集合中选择预设数量的目标相似图像，可以是排序后集合中选择排列在前的预设数量的目标相似图像。

[0210] 106、输出目标相似图像。

[0211] 其中，对于服务器而言，输出目标相似图像，可以包括：向终端发送目标相似图像。终端接收目标相似图像后，可以按照目标相似图像的排列顺序进行显示。

[0212] 本实施例中，可以从图像的承载对象、图像语义、以及图像操作行为三个角度进行相似图像的分析。

[0213] 本实施例中，可以在步骤103之前，对候选图像进行预处理，该预处理包括但不限于：对于动态图像，以动态图像中的预设帧图像代替帧数不固定的动态图像；对候选图像进行格式转换，将候选图像的格式转换为目标格式；根据不同的特征提取算法的需要，对候选图像进行等比例缩放、裁剪、以及生成灰度图等操作。其中，预设帧可以根据需要选择，例如可以是第一帧、或中间帧、或最后一帧等等，本实施例对此没有限定。本实施例的目标格式可以根据实际需要选择，例如目标格式为PNG格式、或JPEG格式等等。可选的，待搜索图像的目标格式和候选图像的目标格式可以相同。

[0214] 其中，对于待搜索图像的相似图像，可以包括以下三种搜索方案：

[0215] (1) 在图像承载对象相似维度上进行搜索

[0216] 其中，步骤“基于多个图像相似维度上的特征信息，于候选图像中在多个图像相似维度上搜索相似图像”，可以包括：

[0217] 基于待搜索图像在图像承载对象相似维度上的特征信息，于候选图像中搜索与待搜索图像在图像承载对象相似维度上相似的相似图像。具体的，可以于区块链的共享账本中存储的候选图像中，搜索与待搜索图像在图像承载对象相似维度上相似的相似图像。

[0218] 该方案中，待搜索图像在图像承载对象相似维度上的特征信息，可以是待搜索图像中的承载对象的图像特征，如图像向量特征。

[0219] 在一个实施例中，图像向量特征可以通过特征提取模型提取。可选的，获取待搜索图像在图像承载对象相似维度上的特征信息，可以包括：基于特征提取模型提取待搜索图像的图像特征。

[0220] 其中，该特征提取模型可以为神经网络模型，该神经网络模型的类型可以有多种，包括但不限于残差神经网络(Residual Neural Network, ResNet)模型，VGG(Visual Geometry Group)神经网络模型，以及Inception深度模型等等。

[0221] 进一步的，该实施例中可以在特征提取模型的基础上，结合注意力机制，提升图像特征的有效性。可选的，步骤“获取待搜索图像在图像承载对象相似维度上的特征信息”，可以包括：

[0222] 通过特征提取模型的局部特征提取块，提取待搜索图像的局部特征；

[0223] 通过对应于局部特征提取块设置的注意力模块，学习局部特征提取块提取的局部特征的特征权重；

[0224] 基于特征权重对局部特征提取块提取的局部特征进行处理，得到注意力局部特征，基于注意力局部特征得到待搜索图像的图像特征。

[0225] 对应地,基于图像承载对象相似维度上的特征信息,于候选图像中搜索与待搜索图像在图像承载对象相似维度上相似的相似图像,包括:

[0226] 获取候选图像基于特征提取模型提取的图像特征;

[0227] 从候选图像中,搜索图像特征与待搜索图像的图像特征相似的相似图像。

[0228] 具体的,从候选图像中,搜索图像特征与待搜索图像的图像特征相似的相似图像,可以包括:基于候选图像的图像特征和待搜索图像的图像特征,计算待搜索图像与各候选图像的相似度,获取相似度不低于第一相似度阈值的候选图像为相似图像。其中,第一相似度阈值可以根据实际需要设置,例如设置为0.5。

[0229] 本实施例中,对于待搜索图像和候选图像的图像特征,都可以通过上述结合了注意力机制的特征提取模型进行提取的。下面以特征提取模型为ResNet残差神经网络模型,图像特征为ResNet向量为例,结合图2a和图2b对图像特征的提取过程进行示例说明。

[0230] 参考图2a,图2a示出了未结合注意力机制的特征提取模型,对图像特征(ResNet向量)进行提取的示意图。ResNet向量提取是指,采用ResNet模型将输入的图像转化为向量特征表示。

[0231] 如图2a所示,ResNet模型中包含多个相同的ResNet Block(残差块,可以理解为上述描述的局部特征提取块)。参考图2a中右侧的ResNet Block的细节图可知,这种Block采用卷积层(Convolution layers)进行局部特征抽取,然后再采用恒等映射(identity mapping)的方式来进行残差学习,有很强的特征捕捉能力。

[0232] 但是,这种Block对图像的所有局部区域进行了无差别的特征抽取,没有考虑到图像中主体的特征比背景更重要。针对这个不足,本申请改进了ResNet的Block结构,参考图2b,本实施例在原始ResNet Block的基础上,加入了注意力层(Attention layers,即上述实施例中描述的注意力模块),本实施例的每个ResNet Block均对应设置有注意力层,注意力层可以学习其对应的ResNet Block提取的不同局部特征的特征权重:提高图像的主体内容所对应的特征权重,降低图像的背景等区域的特征权重;并利用自身所学到的特征权重对卷积层所捕捉的局部特征进行加权,使得从特征提取模型输出的向量特征更有效。

[0233] 可以理解的是,本实施例中,从特征提取模型输出的向量特征的维度可以根据实际需要设置,不仅限于图2a中所示的128维。

[0234] 在实际的图像搜索中,用户可能会搜索很多与影视明星、游戏角色、以及知名形象等相似的相似图像或相似表情。这类表情的重要共同特征之一就是有人脸图像。而一般的特征提取模型,如ResNet,其实并不擅长获取人脸特征。

[0235] 相比于ResNet等模型,使用专业的人脸特征提取模型如FaceNet模型,获取的人脸特征的精度更高。

[0236] 在另一个实施例中,图像承载对象相似维度包括:人脸相似子维度;对应的特征信息包括:人脸特征。该人脸特征也可以是向量特征,通过人脸特征提取模型进行提取。

[0237] 可选的,获取待搜索图像在图像承载对象相似维度上的特征信息,可以包括:

[0238] 检测待搜索图像中是否存在人脸;

[0239] 若检测到待搜索图像中存在人脸,则通过人脸特征提取模型对待搜索图像进行人脸特征的提取。

[0240] 对应地,步骤“基于图像承载对象相似维度上的特征信息,于候选图像中搜索与待

搜索图像在图像承载对象相似维度上相似的相似图像”，可以包括：

[0241] 获取候选图像基于人脸特征提取模型提取的人脸特征；

[0242] 从候选图像中，搜索人脸特征与待搜索图像的人脸特征相似的相似图像。

[0243] 具体的，步骤“从候选图像中，搜索人脸特征与待搜索图像的人脸特征相似的相似图像”，可以包括：基于候选图像的人脸特征和待搜索图像的人脸特征，计算待搜索图像与各候选图像的相似度，获取相似度不低于第二相似度阈值的候选图像为相似图像。其中，第二相似度阈值可以根据实际需要设置，本实施例对此没有限制。

[0244] 其中，候选图像的人脸特征可以实时提取，也可以预先提取并存储为离线特征。获取候选图像基于人脸特征提取模型提取的人脸特征，可以包括：基于人脸特征提取模型对候选图像提取人脸特征，或者，从存储的离线特征中获取候选图像基于人脸特征提取模型提取的人脸特征。

[0245] 可选的，本实施例中人脸特征提取模型可以有多种，包括但不限于FaceNet模型。下面以人脸特征提取模型为FaceNet模型，人脸特征为FaceNet向量为例，结合图2c，对表情的FaceNet向量提取进行示例说明。

[0246] 参考图2c，对输入的表情进行人脸检测，判断表情中是否存在人脸，若表情上存在人脸，则采用的FaceNet模型对表情进行人脸特征的提取，若表情上不存在人脸，则结束人脸特征提取过程。

[0247] (2) 在图像语义相似维度上进行搜索

[0248] 方案(1)中是从图像的承载对象角度来搜索相似图像，但是无法准确捕捉图像如表情的文本语义。例如，当用户搜索“早安”的问候表情时，可根据ResNet向量检索到相似风格的“早安”、“午安”、“晚安”等问候表情，但不能通过ResNet向量的相似度来衡量哪个表情与待搜索表情的语义更相近。鉴于此，本实施例还提出，使用语义信息来衡量图像如表情的语义相似度。

[0249] 本方案中，步骤“基于多个图像相似维度上的特征信息，于候选图像中在多个图像相似维度上搜索相似图像”，可以包括：

[0250] 基于图像语义相似维度上的特征信息，于候选图像中搜索与待搜索图像在图像语义相似维度上相似的相似图像。具体的，可以于区块链的共享账本中存储的候选图像中，搜索与待搜索图像在图像语义相似维度上相似的相似图像。

[0251] 可选的，在一个实施例中，图像语义相似维度包括图像语义相似子维度；对应的特征信息包括图像的语义信息。可选的，语义信息包括但不限于图像上的字符信息，图像的标签携带的信息，通过对图像的理解得到的图像的描述信息。

[0252] 该实施例中，获取待搜索图像在图像语义相似维度上的特征信息，包括：识别待搜索图像的语义信息。

[0253] 对应地，步骤“基于图像语义相似维度上的特征信息，于候选图像中搜索与待搜索图像在图像语义相似维度上相似的相似图像”，可以包括：

[0254] 获取候选图像的语义信息，从候选图像中，搜索语义信息与待搜索图像的语义信息相似的相似图像。

[0255] 在该实施例的一个示例中，图像语义相似子维度包括：图像字符相似次级子维度；对应地，语义信息包括待搜索图像上记载的文字信息。

[0256] 其中,步骤“识别待搜索图像的语义信息”,可以包括:

[0257] 通过图像文字识别技术,识别待搜索图像上记载的文字信息。

[0258] 对应地,步骤“获取候选图像的语义信息,从候选图像中,搜索语义信息与待搜索图像的语义信息相似的相似图像”,可以包括:

[0259] 获取候选图像上记载的文字信息,从候选图像中,搜索文字信息与待搜索图像的文字信息相似的相似图像。

[0260] 具体的,步骤“获取候选图像上记载的文字信息,从候选图像中,搜索文字信息与待搜索图像的文字信息相似的相似图像”,可以包括:

[0261] 获取候选图像上记载的文字信息;

[0262] 计算待搜索图像的文字信息与各候选图像的文字信息的相似度;

[0263] 将相似度不低于第三相似度阈值的候选图像,确定为待搜索图像的相似图像。

[0264] 本实施例中,图像文字识别技术包括但不限于OCR (Optical Character Recognition,光学字符识别) 技术。对于两个图像的文字信息的相似度,可以通过如下的公式计算:

$$[0265] \quad \text{Similarity}_{OCR} = 1 - \frac{\text{Edict_Distance} (OCR_1, OCR_2)}{(\max(\text{len}(OCR_1), \text{len}(OCR_2)) + \text{const})}$$

[0266] 在该公式中,OCR1和OCR2分别表示两个图像(如待搜索图像和候选图像)上通过OCR技术识别出的文字信息,Edict_Distance函数用于计算两个OCR之间的编辑距离;len函数用于统计OCR的长度;const表示常数,避免出现分母为零的情况。

[0267] 虽然图像中的文字可以用来衡量图像的相似度,但存在大量的图像没有文字,因此无法根据上述方法获取到语义相似度。由于图像生产方如表情生产方会为大部分表情提供标签即Tag,该Tag中会记载一些图像的相关信息,所以本示例中将图像的Tag也作为图像的语义信息的一种,用于进行相似图像的搜索。

[0268] 在该实施例的另一个示例中,图像语义相似子维度包括:图像标签相似次级子维度;对应地,语义信息包括待搜索图像的标签。

[0269] 其中,步骤“识别待搜索图像的语义信息”,可以包括:

[0270] 若待搜索图像存在标签,则识别待搜索图像的标签;

[0271] 若待搜索图像不存在标签,则通过标签分类模型对待搜索图像的标签进行预测,根据预测结果获取待搜索图像对应的标签。

[0272] 对应地,步骤“获取候选图像的语义信息,从候选图像中,搜索语义信息与待搜索图像的语义信息相似的相似图像”,可以包括:

[0273] 获取候选图像的标签;

[0274] 从候选图像中,搜索标签与待搜索图像的标签相似的相似图像。

[0275] 其中,本实施例中的标签分类模型的种类不限,可以是任意可用的分类模型,参考图3,图3中示出了基于ResNet的标签分类模型的结构示意图。该标签分类模型在使用前,可以基于训练数据进行训练,训练数据为标注了Tag的预设数量的图像。其中,每个图像的Tag的数量不限,可以是一个或多个。

[0276] 本实施例可以通过图3的标签分类模型为待搜索图像或候选图像预测标签。

[0277] 本实施例中,候选图像的标签可以是在本次的相似图像搜索过程中,通过标签分类模型实时预测得到的,也可以是在本次的相似图像搜索过程前,已经通过标签分类模型预测并进行存储的。

[0278] 本实施例的待搜索图像和候选图像可以有多个Tag,可以理解的是,这些Tag可能会比较相似(例如,“宋XX的呆萌表情”与“宋XX的搞笑表情”是两个相似的Tag),本实施中,从候选图像中,搜索标签与待搜索图像的标签相似的相似图像,可以计算候选图像的标签和待搜索图像的标签的相似度,将相似度不低于第四相似度阈值的候选图像,确定为待搜索图像。其中,标签的相似度的计算方式不限,可以基于标签中的相同/相似字符的数量等确定。

[0279] 可选的,本实施例设计了一种两阶段计算Tag相似度的方法:

[0280] 先定义两个Tag之间的相似度计算公式。假定,待搜索图像的一个Tag表示为 $T_A = \{w_1, \dots, w_i, \dots, w_m\}$,候选图像的一个Tag表示为 $T_B = \{w_1, \dots, w_i, \dots, w_n\}$ 。其中, w_i 表示Tag中的第*i*个字符。则两个Tag的相似度计算如下:

$$[0281] \quad \text{Sim}_{A,B} = \frac{2 * |T_A \cap T_B|}{|T_A| + |T_B|}$$

[0282] 然后,定义两个图像的所有Tag之间的相似度公式。假定待搜索图像X的Tag表示为 $[(A, \text{weight}_A), (B, \text{weight}_B)]$,候选图像Y的Tag表示为 $[(C, \text{weight}_C), (D, \text{weight}_D)]$,其中,A、B、C、D分别表示Tag,weight表示Tag的权重。然后,确定待搜索图像的标签,与候选图像的标签中相似的标签。

[0283] (假定通过计算,标签B和C相似,其他任意Tag之间不相似),则可将待搜索图像X的所有Tag表示为向量 Tag_X :

$$[0284] \quad \text{Tag}_X = [\text{weight}_A, (\text{Sim}_{B,C} * \text{weight}_B * \text{weight}_C), 0]$$

[0285] 将候选图像Y的所有Tag表示为向量 Tag_Y :

$$[0286] \quad \text{Tag}_Y = [0, (\text{Sim}_{B,C} * \text{weight}_B * \text{weight}_C), \text{weight}_D]$$

[0287] 最后,计算 Tag_X 和 Tag_Y 的余弦相似度,作为待搜索图像X和候选图像Y的Tag相似度。

[0288] 在一个实施例中,图像语义相似子维度包括:图像描述信息相似次级子维度;对应的语义信息包括待搜索图像的描述信息。

[0289] 步骤“识别待搜索图像的语义信息”,可以包括:

[0290] 通过图像理解模型对待搜索图像进行分析,以生成待搜索图像的图像描述信息。

[0291] 对应地,步骤“获取候选图像的语义信息,从候选图像中,搜索语义信息与待搜索图像的语义信息相似的相似图像”,可以包括:

[0292] 获取候选图像基于图像理解模型生成的图像描述信息;

[0293] 从候选图像中,搜索图像描述信息与待搜索图像的图像描述信息相似的相似图像。

[0294] 本实施例中,图像理解模型起到了“看图说话”的作用,该图像理解模型的种类有很多,例如可以为Image Caption模型。

[0295] 可选的,步骤“从候选图像中,搜索图像描述信息与待搜索图像的图像描述信息相似的相似图像”,可以包括:

[0296] 对待搜索图像的图像描述信息,与候选图像的图像描述信息,进行余弦相似度的

计算,将余弦相似度不低于第五相似度阈值的候选图像,确定为相似图像。

[0297] 其中,候选图像的图像描述信息可以通过图像理解模型实时获取的,也可以是预先获取之后,存储在候选图像的信息集合中的。

[0298] 在实际中,用户搜索图像如表情时,会有对话场景的需求。例如当用户搜索“谢谢”的表情,很可能希望搜索到的是“不用谢”的表情。本实施例的语义相关,还可以是语义所属的对话关系相关。

[0299] 在本实施例中,图像语义相似维度包括图像对话关系相似子维度;对应的,图像在图像语义相似维度上的特征信息包括:意义信息。

[0300] 可选的,获取待搜索图像在图像语义相似维度上的特征信息,包括:识别待搜索图像的语义信息。

[0301] 对应地,“基于图像语义相似维度上的特征信息,于候选图像中搜索与待搜索图像在图像语义相似维度上相似的相似图像”,可以包括:

[0302] 确定与待搜索图像的语义信息在对话关系上相似的目标语义信息;

[0303] 获取候选图像的语义信息;

[0304] 从候选图像中,搜索语义信息与目标语义信息相似的相似图像。

[0305] 其中,候选图像的语义信息包括但不限于:待搜索图像上的文字信息,待搜索图像的标签,以及待搜索图像的描述信息。获取这些语义信息的具体方案可以参考前述内容中的相关描述,在此不进行赘述。

[0306] 在一个示例中,与待搜索图像的语义信息在对话关系上相似的目标语义信息,可以理解为与待搜索图像的语义信息符合对话场景的目标语义信息,进一步的,本实施例中,可以通过对话识别模型,确定与语义信息符合对话场景的目标语义信息。

[0307] 其中,对话识别模型包括但不限于Bert对话表情挖掘模型。

[0308] 本实施例中,可以先获取Bert预训练模型,以及外部公开的对话数据集;该对话数据集中可以各种场景中的对话,例如包含日常生活中常见的对话场景中的对话,比如谢谢和不用谢,对不起和没关系等等;然后利用对话数据集对Bert预训练模型进行微调训练得到上述使用的Bert对话表情挖掘模型。

[0309] 本实施例中,可以在使用外部的对话数据集之前,先对对话数据集进行预处理,该预处理包括:去除无效数据(例如非用户数据)、重复数据、噪音数据、异常数据,以及有针对性地去除低频数据(针对OCR识别的方案,例如引用次数较低的数据)等数据清洗操作。

[0310] 在一个示例中,步骤“基于图像语义相似维度上的特征信息,于候选图像中搜索与待搜索图像在图像语义相似维度上相似的相似图像”,可以包括:

[0311] 通过对话识别模型,确定与待搜索图像记载的文字信息符合对话场景的目标文字信息;

[0312] 获取候选图像上记载的文字信息;

[0313] 从候选图像中,搜索文字信息与目标文字信息相似的相似图像。

[0314] 在一个示例中,步骤“基于图像语义相似维度上的特征信息,于候选图像中搜索与待搜索图像在图像语义相似维度上相似的相似图像”,可以包括:

[0315] 确定与待搜索图像的标签上的标签信息符合对话场景的目标标签信息;

[0316] 获取候选图像上的标签信息;

[0317] 从候选图像中,搜索标签信息与目标标签信息相似的相似图像。

[0318] 在一个示例中,步骤“基于图像语义相似维度上的特征信息,于候选图像中搜索与待搜索图像在图像语义相似维度上相似的相似图像”,可以包括:

[0319] 确定与待搜索图像的描述信息符合对话场景的目标描述信息;

[0320] 获取候选图像的描述信息;

[0321] 从候选图像中,搜索描述信息与目标描述信息相似的相似图像。

[0322] (3) 在图像操作行为相似维度上进行搜索

[0323] 其中,步骤“基于多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像”,可以包括:

[0324] 基于图像操作行为相似维度上的特征信息,于候选图像中搜索与待搜索图像在图像操作行为相似维度上相似的相似图像。可选的,可以于区块链的共享账本存储的候选图像中,搜索与待搜索图像在图像操作行为相似维度上相似的相似图像。

[0325] 用户在应用程序中使用图像搜索功能时,应用程序如即时通讯客户端的后台日志系统能够记录用户在每个session中搜索和点击的图像。其中,对于Session:是指在一段指定时间内,用户在某个系统或平台上发生的一系列行为。以搜索图像为例,在指定时间段内,用户连续搜索的多个表情属于同一个Session。

[0326] 同个session中的表情通常体现了用户在短时间内的搜索意向,因此很容易产生相似表情。鉴于此,本申请可以利用日志数据,基于session进行相似图像的搜索。

[0327] 在一个实施例中,图像操作行为相似维度包括:单用户图像操作行为相似子维度;

[0328] 获取待搜索图像在图像操作行为相似维度上的特征信息,包括:

[0329] 获取待搜索图像在第一历史图像标识号文本中的词向量,其中,第一历史图像标识号文本为第一历史操作图像集合中的图像的标识号组成的文本,第一历史操作图像集合中的图像为同一用户在历史时间段内进行了操作的图像,且第一用户历史操作图像集合包括待搜索图像。

[0330] 对应地,步骤“基于图像操作行为相似维度上的特征信息,于候选图像中搜索与待搜索图像在图像操作行为相似维度上相似的相似图像”,可以包括:

[0331] 获取候选图像在第二历史图像标识号文本中的词向量,其中,第二历史图像标识号文本为第二历史操作图像集合中的图像的标识号组成的文本,第二历史操作图像集合中的图像为同一用户在历史时间段内进行了操作的图像,且第二用户历史操作图像集合包括候选图像;

[0332] 从候选图像中,搜索词向量与待搜索图像的词向量相似的相似图像。

[0333] 其中,第一历史操作图像集合可以从第二历史操作图像集合中确定,具体的,可以在本实施例之前,将每一个候选图像看作待搜索图像,从所有的第二历史操作图像集合中确定包含该候选图像的第一历史操作图像集合,并针对各个第一历史操作图像集合,确定每一个候选图像的词向量。

[0334] 其中,获取待搜索图像在第一历史图像标识号文本中的词向量,可以包括:

[0335] 确定候选图像中是否存在待搜索图像,若存在,则从候选图像的词向量中获取待搜索图像的词向量。

[0336] 本实施例中的历史时间段为泛指,表示的是以历史的一段时间段,对于该时间段

的起点和终点并没有限制。第一历史操作图像集合和第二历史操作图像集合的历史时间段可以相同或不同。

[0337] 其中,对于第一历史操作图像集合,可以基于发送待搜索图像的用户的日志数据得到。

[0338] 本实施例中,图像的标识号可以是图像ID,获取待搜索图像在第一历史图像标识号文本中的词向量,具体可以包括:

[0339] 获取包含待搜索图像的若干第一历史操作图像集合,其中,对于同一个第一历史操作图像集合而言,对图像进行操作的用户是相同的,对于不同的第一历史操作图像集合而言,对应的用户可以不同;

[0340] 将第一历史操作图像集合中的图像映射成独特的ID,将第一历史操作图像集合中的图像转换为一行ID文本;其中,同一图像对应的ID是相同的,不同图像对应的ID是不同的;

[0341] 将每一行ID文本作为一个语句,输入到预设语言模型中进行训练,得到第一历史操作图像集合中每个图像ID的词向量表示;

[0342] 从得到的词向量中,获取待搜索图像的词向量。

[0343] 对于候选图像的词向量,其获取过程与待搜索图像的词向量的获取过程类似,在此不再赘述。

[0344] 步骤“从候选图像中,搜索词向量与待搜索图像的词向量相似的相似图像”,可以包括:

[0345] 计算候选图像的词向量与待搜索图像的词向量的相似度;

[0346] 将相似度不低于第六相似度阈值的候选图像,确定为待搜索图像的相似图像。

[0347] 其中,该相似度可以为余弦相似度,预设语言模型可以为任意可用的语言模型,包括但不限于Word2vec语言模型。

[0348] 上述方案中,从单个用户的图像操作行为角度进行了分析,本实施例还可以从多个用户的图像操作行为角度进行分析。SimRank的基本思想是,如果两个用户相似,那么与这两个用户相关联的表情也相似;如果两个表情相似,那么与这两个表情相关联的用户也相似。本实施例还可以基于SimRank的基本思想,提出多用户图像操作行为角度的相似图像搜索。

[0349] 可选的,图像操作行为相似维度包括:多用户图像操作行为相似子维度;

[0350] 获取待搜索图像在图像操作行为相似维度上的特征信息,包括:获取待搜索图像的认识信息。其中,该认识信息可以是图像的ID。

[0351] 对应地,步骤“基于图像操作行为相似维度上的特征信息,于候选图像中搜索与待搜索图像在图像操作行为相似维度上相似的相似图像”,可以包括:

[0352] 基于待搜索图像的认识信息,获取包含待搜索图像的目标二部图对应的图像相似度矩阵,其中,目标二部图的一边顶点表示用户(该用户为泛指,不限于发送待搜索图像的用户,可以包括使用本实施例的图像搜索方法进行图像搜索的用户),另一边顶点表示候选图像中用户的历史操作图像,图像相似度矩阵包括历史操作图像之间的图像相似度;

[0353] 在图像相似度矩阵中,搜索与待搜索图像的图像相似度不低于预设相似度阈值的相似图像。

[0354] 对于二部图,设 $G=(V,E)$ 是一个无向图,如果顶点 V 可分割为两个互不相交的子集 (A,B) ,并且图中的每条边 (i,j) 所关联的两个顶点 i 和 j 分别属于这两个不同的顶点集 A 和 B ,则称图 G 为一个二部图。

[0355] 本实施例中的二部图基于用户和用户操作过的图像构建,二部图其中一边的顶点是用户ID,另一边的顶点是图像ID;本实施例中,若某个用户对某个图像产生操作行为包括搜索、点击、以及收藏等行为,则该用户ID与该图像ID存在一条边。目标二部图是包含待搜索图像的二部图。

[0356] 可选的,获取包含待搜索图像的目标二部图对应的图像相似度矩阵,可以是实时计算目标二部图的图像相似度矩阵,也可以是从存储的图像相似度矩阵中获取目标二部图对应的图像相似度矩阵。

[0357] 本实施例中,二部图的每条边有不同的权重,图像被用户操作过的次数越多,权重越大;不同操作行为的权重也不同,例如点击行为的权重大于搜索行为,收藏行为的权重大于点击行为,本实施例中的图像相似度矩阵基于每条边的权重计算得到。

[0358] 本实施例中,计算二部图对应的图像相似度矩阵,可以通过SimRank算法对二部图结构进行计算,得到二部图的表情相似度矩阵。

[0359] 根据上述描述可知,本实施例中有多个图像相似维度,每个图像相似维度下还存在多个子维度,有些子维度下还存在次级子维度,为了进一步提升排序的准确性,本实施例中,对于各个子维度和次级子维度,设置有对应的相似度权重。

[0360] 步骤“根据相似图像在对应的图像相似维度下的相似度,以及各图像相似维度对应的相似度权重,计算相似图像集合中相似图像的综合相似度”,可以包括:

[0361] 根据相似图像在各个图像相似子维度和图像相似次级子维度下的相似度,以及各图像相似子维度和图像相似次级子维度对应的相似度权重,计算相似图像集合中相似图像的综合相似度。

[0362] 本实施例中,在用户使用图像搜索一段时间后,会累积一定的用户历史操作行为数据,在用户历史操作行为数据的累积量满足预设条件时,不执行根据相似图像集合中相似图像对应的图像相似维度,对相似图像集合中的相似图像进行排序,得到排序后集合的步骤;而是根据用户历史操作行为数据,分析用户对相似图像集合中相似图像的用户偏好,基于用户偏好对相似图像进行排序,得到排序后集合。

[0363] 其中,可以基于用户历史操作行为数据,采用LambdaRank算法对相似图像进行排序。用户历史操作行为数据的累积量满足预设条件包括但不限于:用户历史操作行为数据的累积时长不低于预设时长阈值,或者用户历史操作行为数据的数据量不低于预设数据量阈值。

[0364] 其中,用户历史操作行为数据包括:基于待搜索图像进行图像搜索的用户,在历史时间段内进行图像搜索产生的行为数据,该行为数据包括但不限于:图像被搜索出来的搜索次数,图像被搜索后显示的时间长度,图像被搜索并显示出来后用户的点击次数,图像的曝光次数等等,以及通过这些数据衍生的衍生数据,例如点击次数与搜索次数的比例等等。其中,该历史时间段为泛指,表示历史的一段时间段,时间段的起点和起点没有限制。

[0365] 本实施例中,从图像、语义、图像操作行为等多个角度充分考虑了图像之间的相似性,极大地提高了相似表情的多样性,能够更好地满足用户多样化的需求。基于卷积神经网络

络的深度相似匹配模型是前沿的相似表情匹配方案,而本方案在此基础上提出了改进:采用注意力模块辅助卷积神经网络捕捉更重要的特征,提高了本实施例的相似匹配能力。

[0366] 根据上面实施例所描述的方法,以下将举例作进一步详细说明。

[0367] 在本实施例中,将以该第一图像搜索装置具体集成终端、第二图像搜索装置具体集成在服务器为例进行说明。

[0368] 如图4a所示,一种图像搜索方法,具体流程包括:

[0369] 401、显示应用程序客户端的图像搜索页面,图像搜索页面为供用户以图像搜索图像的搜索页面,其中,图像搜索页面包括:待搜索图像选择控件;

[0370] 其中,应用程序客户端可以是即时通讯客户端,或者搜索引擎等应用程序的客户端,本实施例对此没有限制。

[0371] 其中,步骤“显示应用程序的图像搜索页面”,可以包括:

[0372] 显示应用程序客户端的功能页面,其中,功能页面包括:搜索功能控件;

[0373] 当检测到用户针对搜索功能控件的触发操作时,显示内容搜索页面,其中,内容搜索页面包括:图像搜索控件;

[0374] 当检测到用户针对图像搜索控件的触发操作时,显示图像搜索页面。

[0375] 例如,参考图4b,图4b中,应用程序客户端的功能页面如“发现”页面上,包含有多个功能控件,如扫码功能控件“扫一扫”、内容发布控件“朋友圈”、小程序功能控件“小程序”、图像搜索控件“搜一搜”等功能控件;当用户对图像搜索控件“搜一搜”进行触发操作如点击、滑动等操作时,终端会显示内容搜索页面,内容搜索页面中包含若干用于搜索不同对象的搜索控件,包括但不限于图像搜索控件“表情”,文章搜索控件“文章”,公众号搜索控件“公众号”,小程序搜索控件“小程序”等等。当用户对内容搜索页面中的图像搜索控件,如“表情”控件进行操作如点击、滑动等操作时,触发终端显示如图4b所示的表情搜索页面,该表情搜索页面上显示有图标“+”表示的待搜索图像选择控件。

[0376] 402、当检测到用户针对待搜索图像选择控件的触发操作时,显示待搜索图像的选择页面;

[0377] 如图4c所示,当检测到用户针对选择页面的“+”控件的触发操作,如点击等操作时,显示待搜索图像的选择页面,该选择页面上显示有多个待选图像,如表情1、表情2以及表情3 等等。

[0378] 403、基于用户针对选择页面的待搜索图像确定操作,向服务器发送用户在选择页面上选择的候选图像,该被选择的候选图像为待搜索图像;

[0379] 如图4c所示,当检测到用户针对选择页面上的某个待选图像如表情1的选定操作,则将用户选定的待选图像如表情1作为待搜索图像,向服务器发送该待搜索图像。

[0380] 404、服务器接收待搜索图像;

[0381] 405、服务器获取待搜索图像在多个图像相似维度上的特征信息;

[0382] 参考图4d,图4d中是本实施例的图像搜索的一个可选的框架图,本实施例从图像、语义、用户的图像操作行为等三个角度进行特征提取,是图4d所示的相似图像搜索框架的核心内容。

[0383] 本实施例中的相似图像搜索框架,包括以下四个阶段:数据预处理阶段,特征提取阶段,召回阶段,排序阶段。

[0384] 本实施例的数据预处理阶段,包括对候选图像和待搜索图像的图像预处理,对搜索日志数据和外部的对话数据集的预处理(如数据清洗),这些预处理的描述可以参考上述实施例中的相关描述,在此不进行赘述。

[0385] 本实施例的特征提取阶段,包括对候选图像的特征提取和对待搜索图像的特征提取。本实施例中,在步骤401之前,服务器对候选图像可以通过以下的方式获取多个图像相似维度上的7种特征信息,并保存为离线的特征信息以便后续使用:

[0386] 对候选图像进行图像特征如ResNet向量的提取;

[0387] 对候选图像进行人脸检测,若候选图像上存在人脸,则对候选图像进行人脸特征如FaceNet向量的提取;

[0388] 对候选图像进行OCR识别,确定候选图像上记载的文字信息;

[0389] 对候选图像进行Tag识别,确定候选图像的Tag,其中,若候选图像上不存在Tag,则通过标签分类模型对候选图像的Tag进行预测;

[0390] 对候选图像进行OCR识别,确定候选图像上记载的文字信息,对候选图像进行对话表情挖掘,确定候选图像的文字信息的对话场景;

[0391] 对候选图像基于单用户的图像操作行为,确定候选图像的词向量;

[0392] 对候选图像进行Simrank相似挖掘多用户的图像操作行为,确定候选图像对应的二部图,基于候选图像的二部图,确定候选图像对应的图像相似度矩阵。

[0393] 本实施例中,服务器获取待搜索图像在多个图像相似维度上的特征信息包括:

[0394] 对待搜索图像提取图像特征如ResNet向量;

[0395] 对待搜索图像进行人脸检测,若待搜索图像上存在人脸,则对待搜索图像进行人脸特征如FaceNet向量的提取;

[0396] 对待搜索图像进行OCR识别,以确定待搜索图像上记载的文字信息;

[0397] 确定候选图像中是否存在待搜索图像,若存在,从候选图像的词向量中获取待搜索图像的词向量;

[0398] 确定候选图像的认识信息。

[0399] 406、服务器基于多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合,相似图像集合中的相似图像与待搜索图像在至少一个图像相似维度上相似;

[0400] 本实施例的搜索框架中的召回阶段,包括两个方面:构建检索系统和检索相似图像。其中,构建检索系统可以预先完成。参考图4d,相似图像搜索框架可以为图像特征向量、人脸特征向量、基于对单用户的图像操作行为的挖掘得到的词向量构建向量检索系统,为光学字符识别得到的特征和标签特征构建文本检索系统、为Bert对话表情挖掘的结果以及Simrank相似挖掘的结果构建KV(Key-value)键值对检索系统。

[0401] 其中,在KV检索系统中,可以基于Bert对话表情挖掘的结果以及Simrank相似挖掘的结果,存储每个候选图像和对应的挖掘结果的对应关系,例如存储每个候选图像和其相似图像的对应关系。

[0402] 在进行相似图像的搜索时,当候选图像中存在待搜索图像时,可以根据KV检索系统中存储的候选图像与挖掘结果的对应关系,确定待搜索图像的相似图像。

[0403] 在步骤406中,可以基于待搜索图像的特征信息以及候选图像的离线的特征信息,

进行相似图像的搜索,具体的搜索过程可以参考上述实施例中的描述,在此不进行赘述。

[0404] 407、服务器根据相似图像集合中相似图像对应的图像相似维度,对相似图像集合中的相似图像进行排序,得到排序后集合;

[0405] 步骤407中,进行排序时,对于每一个相似图像,可以根据各图像相似维度对应的不同权重,对各个图像相似维度下的相似图像与待搜索图像的相似度进行加权求和,得到各个相似图像的综合相似度;基于综合相似度对相似图像进行降序排列,得到排序后集合。

[0406] 本实施例中,在应用程序客户端的冷启动阶段,可以采用类似步骤407的方案进行排序,在(终端的当前用户的)用户历史操作行为数据(图4d中的用户行为数据)的累积量满足预设条件后,可以基于该用户历史操作行为数据,采用LambdaRank算法对相似图像进行排序,得到排序后集合。

[0407] 408、服务器从排序后集合中选择预设数量的目标相似图像;

[0408] 其中,预设数量可以根据实际需要设置,例如,预设数量可以为300张,400张等等。

[0409] 409、服务器向终端发送目标相似图像。

[0410] 可以理解的是,为了便于终端按照综合相似度从高到低的顺序显示目标相似图像,服务器向终端发送目标相似图像时,还可以将目标相似图像的顺序信息一同发给终端,以便于终端按照综合相似度的高低顺序显示目标相似图像。

[0411] 410、终端接收目标相似图像,显示搜索结果列表,搜索结果列表包括搜索到的与待搜索图像相似的目标相似图像。

[0412] 为了更好地实施以上方法,相应的,本发明实施例还提供一种图像搜索装置,其中,该图像搜索装置可以集成在终端或服务器中。

[0413] 参考图5,该图像搜索装置包括:

[0414] 第一获取单元501,用于获取待搜索图像;

[0415] 第二获取单元502,用于获取待搜索图像在多个图像相似维度上的特征信息;

[0416] 搜索单元503,用于基于多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合,相似图像集合中的相似图像与待搜索图像在至少一个图像相似维度上相似;

[0417] 排序单元504,用于根据相似图像集合中相似图像对应的图像相似维度,对相似图像集合中的相似图像进行排序,得到排序后集合;

[0418] 选择单元505,用于从排序后集合中选择预设数量的目标相似图像;

[0419] 输出单元506,用于输出目标相似图像。

[0420] 在一个实施例中,图像相似维度包括:图像承载对象相似维度、图像语义相似维度以及图像操作行为相似维度;

[0421] 搜索单元包括:

[0422] 第一搜索子单元,用于基于图像承载对象相似维度上的特征信息,于候选图像中搜索与待搜索图像在图像承载对象相似维度上相似的相似图像;

[0423] 第二搜索子单元,用于基于图像语义相似维度上的特征信息,于候选图像中搜索与待搜索图像在图像语义相似维度上相似的相似图像;

[0424] 第三搜索子单元,用于基于图像操作行为相似维度上的特征信息,于候选图像中搜索与待搜索图像在图像操作行为相似维度上相似的相似图像。

- [0425] 在一个实施例中,图像承载对象相似维度包括:图像特征相似子维度;
- [0426] 第二获取单元,包括:
- [0427] 图像特征提取子单元,用于通过特征提取模型的局部特征提取块,提取所述待搜索图像的局部特征;
- [0428] 学习子单元,用于通过对应于所述局部特征提取块设置的注意力模块,学习所述局部特征提取块提取的局部特征的特征权重;
- [0429] 处理子单元,用于基于所述特征权重对所述局部特征提取块提取的局部特征进行处理,得到注意力局部特征,基于所述注意力局部特征得到所述待搜索图像的图像特征;
- [0430] 第一搜索子单元,用于:
- [0431] 获取候选图像基于特征提取模型提取的图像特征;
- [0432] 从候选图像中,搜索图像特征与待搜索图像的图像特征相似的相似图像。
- [0433] 在一个实施例中,图像承载对象相似维度包括:人脸相似子维度;
- [0434] 第二获取单元,包括:
- [0435] 人脸检测子单元,用于检测待搜索图像中是否存在人脸;
- [0436] 人脸特征提取子单元,用于若检测到待搜索图像中存在人脸,则通过人脸特征提取模型对待搜索图像进行人脸特征的提取;
- [0437] 第一搜索子单元,用于:
- [0438] 获取候选图像基于人脸特征提取模型提取的人脸特征;
- [0439] 从候选图像中,搜索人脸特征与待搜索图像的人脸特征相似的相似图像。
- [0440] 在一实施例中,图像语义相似维度包括图像语义相似子维度;
- [0441] 第二获取单元,用于识别待搜索图像的语义信息;
- [0442] 第二搜索子单元,用于获取候选图像的语义信息,从候选图像中,搜索语义信息与待搜索图像的语义信息相似的相似图像。
- [0443] 在一实施例中,图像语义相似维度包括图像对话关系相似子维度;
- [0444] 第二获取单元,用于识别待搜索图像的语义信息;
- [0445] 第二搜索子单元,用于:
- [0446] 确定与待搜索图像的语义信息在对话关系上相似的目标语义信息;
- [0447] 获取候选图像的语义信息;
- [0448] 从候选图像中,搜索语义信息与目标语义信息相似的相似图像。
- [0449] 在一实施例中,图像语义相似子维度包括:图像字符相似次级子维度;
- [0450] 第二获取单元,用于通过图像文字识别技术,识别待搜索图像上记载的文字信息;
- [0451] 第二搜索子单元,用于获取候选图像上记载的文字信息,从候选图像中,搜索文字信息与待搜索图像的文字信息相似的相似图像。
- [0452] 在一实施例中,图像语义相似子维度包括:图像标签相似次级子维度;
- [0453] 第二获取单元,用于若待搜索图像存在标签,则识别待搜索图像的标签;若待搜索图像不存在标签,则通过标签分类模型对待搜索图像的标签进行预测,根据预测结果获取待搜索图像对应的标签;
- [0454] 第二搜索子单元,用于:
- [0455] 获取候选图像的标签;

- [0456] 从候选图像中,搜索标签与待搜索图像的标签相似的相似图像。
- [0457] 在一实施例中,图像语义相似子维度包括:图像描述信息相似次级子维度;
- [0458] 第二获取单元,用于通过图像理解模型对待搜索图像进行分析,以生成待搜索图像的图像描述信息;
- [0459] 第二搜索子单元,用于:
- [0460] 获取候选图像基于图像理解模型生成的图像描述信息;
- [0461] 从候选图像中,搜索图像描述信息与待搜索图像的图像描述信息相似的相似图像。
- [0462] 在一实施例中,图像操作行为相似维度包括:单用户图像操作行为相似子维度;
- [0463] 第二获取单元,用于获取待搜索图像在第一历史图像标识号文本中的词向量,其中,第一历史图像标识号文本为第一历史操作图像集合中的图像的标识号组成的文本,第一历史操作图像集合中的图像为同一用户在历史时间段内进行了操作的图像,且第一用户历史操作图像集合包括待搜索图像;
- [0464] 第三搜索子单元,用于:
- [0465] 获取候选图像在第二历史图像标识号文本中的词向量,其中,第二历史图像标识号文本为第二历史操作图像集合中的图像的标识号组成的文本,第二历史操作图像集合中的图像为同一用户在历史时间段内进行了操作的图像,且第二用户历史操作图像集合包括候选图像;
- [0466] 从候选图像中,搜索词向量与待搜索图像的词向量相似的相似图像。
- [0467] 在一实施例中,图像操作行为相似维度包括多用户图像操作行为相似子维度;
- [0468] 第二获取单元,用于获取待搜索图像的认识信息;
- [0469] 第三搜索子单元,用于:
- [0470] 基于待搜索图像的认识信息,获取包含待搜索图像的目标二部图对应的图像相似度矩阵,其中,目标二部图的一边顶点表示用户,另一边顶点表示候选图像中用户的历史操作图像,图像相似度矩阵包括历史操作图像之间的图像相似度;
- [0471] 在图像相似度矩阵中,搜索与待搜索图像的图像相似度不低于预设相似度阈值的相似图像。
- [0472] 在一实施例中,排序单元,包括:
- [0473] 加权子单元,用于根据相似图像在对应的图像相似维度下的相似度,以及各图像相似维度对应的相似度权重,计算相似图像集合中相似图像的综合相似度;
- [0474] 排序子单元,用于根据相似图像的综合相似度,对相似图像集合中的相似图像进行排序。
- [0475] 在一实施例中,排序单元,还用于:
- [0476] 在用户历史操作行为数据的累积量满足预设条件时,不执行根据相似图像集合中相似图像对应的图像相似维度,对相似图像集合中的相似图像进行排序,得到排序后集合的步骤;其中,用户历史操作行为数据包括:基于待搜索图像进行图像搜索的用户,在历史时间段内进行图像搜索产生的行为数据;而是根据用户历史操作行为数据,分析用户对相似图像集合中相似图像的用户偏好,基于用户偏好对相似图像进行排序,得到排序后集合。
- [0477] 此外,本发明实施例还提供一种计算机设备,该计算机设备可以为终端或者服务

器,如图6所示,其示出了本发明实施例所涉及的计算机设备的结构示意图,具体来讲:

[0478] 该计算机设备可以包括一个或者一个以上处理核心的处理器601、一个或一个以上计算机可读存储介质的存储器602、电源603和输入单元604等部件。本领域技术人员可以理解,图6中示出的计算机设备结构并不构成对计算机设备的限定,可以包括比图示更多或更少的部件,或者组合某些部件,或者不同的部件布置。其中:

[0479] 处理器601是该计算机设备的控制中心,利用各种接口和线路连接整个计算机设备的各个部分,通过运行或执行存储在存储器602内的软件程序和/或模块,以及调用存储在存储器602内的数据,执行计算机设备的各种功能和处理数据,从而对计算机设备进行整体监控。可选的,处理器601可包括一个或多个处理核心;优选的,处理器601可集成应用处理器和调制解调处理器,其中,应用处理器主要处理操作系统、用户界面和应用程序等,调制解调处理器主要处理无线通信。可以理解的是,上述调制解调处理器也可以不集成到处理器601中。

[0480] 存储器602可用于存储软件程序以及模块,处理器601通过运行存储在存储器602的软件程序以及模块,从而执行各种功能应用以及数据处理。存储器602可主要包括存储程序区和存储数据区,其中,存储程序区可存储操作系统、至少一个功能所需的应用程序(比如声音播放功能、图像播放功能等)等;存储数据区可存储根据计算机设备的使用所创建的数据等。此外,存储器602可以包括高速随机存取存储器,还可以包括非易失性存储器,例如至少一个磁盘存储器件、闪存器件、或其他易失性固态存储器件。相应地,存储器602还可以包括存储器控制器,以提供处理器601对存储器602的访问。

[0481] 计算机设备还包括给各个部件供电的电源603,优选的,电源603可以通过电源管理系统与处理器601逻辑相连,从而通过电源管理系统实现管理充电、放电、以及功耗管理等功能。电源603还可以包括一个或一个以上的直流或交流电源、再充电系统、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。

[0482] 该计算机设备还可包括输入单元604,该输入单元604可用于接收输入的数字或字符信息,以及产生与用户设置以及功能控制有关的键盘、鼠标、操作杆、光学或者轨迹球信号输入。

[0483] 尽管未示出,计算机设备还可以包括显示单元等,在此不再赘述。具体在本实施例中,计算机设备中的处理器601会按照如下的指令,将一个或一个以上的应用程序的进程对应的可执行文件加载到存储器602中,并由处理器601来运行存储在存储器602中的应用程序,从而实现各种功能,如下:

[0484] 获取待搜索图像;

[0485] 获取待搜索图像在多个图像相似维度上的特征信息;

[0486] 基于多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合,相似图像集合中的相似图像与待搜索图像在至少一个图像相似维度上相似;

[0487] 根据相似图像集合中相似图像对应的图像相似维度,对相似图像集合中的相似图像进行排序,得到排序后集合;

[0488] 从排序后集合中选择预设数量的目标相似图像;

[0489] 输出目标相似图像。

[0490] 本发明实施例涉及的系统可以是由客户端、多个节点(接入网络中的任意形式的计算机设备,如服务器、终端)通过网络通信的形式连接形成的分布式系统。

[0491] 以分布式系统为区块链系统为例,参见图7,图7是本发明实施例提供的分布式系统100应用于区块链系统的一个可选的结构示意图,由多个节点200(接入网络中的任意形式的计算设备,如服务器、用户终端)和客户端300形成,节点之间形成组成的点对点(P2P, Peer To Peer)网络,P2P协议是一个运行在传输控制协议(TCP,Transmission Control Protocol)协议之上的应用层协议。在分布式系统中,任何机器如服务器、终端都可以加入而成为节点,节点包括硬件层、中间层、操作系统层和应用层。本实施例中,候选图像(以及候选图像的各种特征信息,如图像特征、人脸特征、文字信息、描述信息以及标签等),可以通过节点被存储在区块链系统的共享账本中,计算机设备(例如终端或服务器)可以基于共享账本存储的记录数据获取候选图像,以及其特征信息。

[0492] 参见图7示出的区块链系统中各节点的功能,涉及的功能包括:

[0493] 1) 路由,节点具有的基本功能,用于支持节点之间的通信。

[0494] 节点除具有路由功能外,还可以具有以下功能:

[0495] 2) 应用,用于部署在区块链中,根据实际业务需求而实现特定业务,记录实现功能相关的数据形成记录数据,在记录数据中携带数字签名以表示任务数据的来源,将记录数据发送到区块链系统中的其他节点,供其他节点在验证记录数据来源以及完整性成功时,将记录数据添加到临时区块中。

[0496] 例如,应用实现的业务包括:

[0497] 2.1) 钱包,用于提供进行电子货币的交易的交易的功能,包括发起交易(即,将当前交易的交易记录发送给区块链系统中的其他节点,其他节点验证成功后,作为承认交易有效的响应,将交易的记录数据存入区块链的临时区块中;当然,钱包还支持查询电子货币地址中剩余的电子货币;

[0498] 2.2) 共享账本,用于提供账目数据的存储、查询和修改等操作的功能,将对账目数据的操作的记录数据发送到区块链系统中的其他节点,其他节点验证有效后,作为承认账目数据有效的响应,将记录数据存入临时区块中,还可以向发起操作的节点发送确认。

[0499] 2.3) 智能合约,计算机化的协议,可以执行某个合约的条款,通过部署在共享账本上的用于在满足一定条件时而执行的代码实现,根据实际的业务需求代码用于完成自动化的交易,例如查询买家所购买商品的物流状态,在买家签收货物后将买家的电子货币转移到商户的地址;当然,智能合约不仅限于执行用于交易的合约,还可以执行对接收的信息进行处理的合约。

[0500] 3) 区块链,包括一系列按照产生的先后时间顺序相互接续的区块(Block),新区块一旦加入到区块链中就不会再被移除,区块中记录了区块链系统中节点提交的记录数据。

[0501] 参见图8,图8是本发明实施例提供的区块结构(Block Structure)一个可选的示意图,每个区块中包括本区块存储交易记录的哈希值(本区块的哈希值)、以及前一区块的哈希值,各区块通过哈希值连接形成区块链。另外,区块中还可以包括有区块生成时的时间戳等信息。区块链(Blockchain),本质上是一个去中心化的数据库,是一串使用密码学方法相关联产生的数据块,每一个数据块中包含了相关的信息,用于验证其信息的有效性(防伪)和生成下一个区块。

[0502] 本领域普通技术人员可以理解,上述实施例的各种方法中的全部或部分步骤可以通过指令来完成,或通过指令控制相关的硬件来完成,该指令可以存储于计算机可读存储介质中,并由处理器进行加载和执行。

[0503] 为此,本申请实施例提供一种存储介质,其中存储有多条指令,该指令能够被处理器进行加载,以执行本申请实施例所提供的图像搜索方法中的步骤。例如,该指令可以执行如下步骤:

[0504] 获取待搜索图像;

[0505] 获取待搜索图像在多个图像相似维度上的特征信息;

[0506] 基于多个图像相似维度上的特征信息,于候选图像中在多个图像相似维度上搜索相似图像,得到相似图像集合,相似图像集合中的相似图像与待搜索图像在至少一个图像相似维度上相似;

[0507] 根据相似图像集合中相似图像对应的图像相似维度,对相似图像集合中的相似图像进行排序,得到排序后集合;

[0508] 从排序后集合中选择预设数量的目标相似图像;

[0509] 输出目标相似图像。

[0510] 其中,该存储介质可以包括:只读存储器(ROM,Read Only Memory)、随机存取记忆体(RAM,Random Access Memory)、磁盘或光盘等。

[0511] 由于该存储介质中所存储的指令,可以执行本发明实施例所提供的任一种图像搜索方法中的步骤,因此,可以实现本发明实施例所提供的任一种图像搜索方法所能实现的有益效果,详见前面的实施例,在此不再赘述。

[0512] 以上对本申请实施例所提供的图像搜索方法、装置、计算机设备和存储介质进行了详细介绍,本文中应用了具体个例对本申请的原理及实施方式进行了阐述,以上实施例的说明只是用于帮助理解本申请的方法及其核心思想;同时,对于本领域的技术人员,依据本申请的思想,在具体实施方式及应用范围上均会有改变之处,综上,本说明书内容不应理解为对本申请的限制。

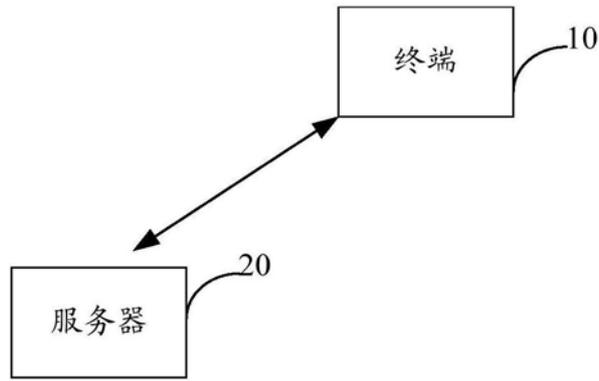


图1a

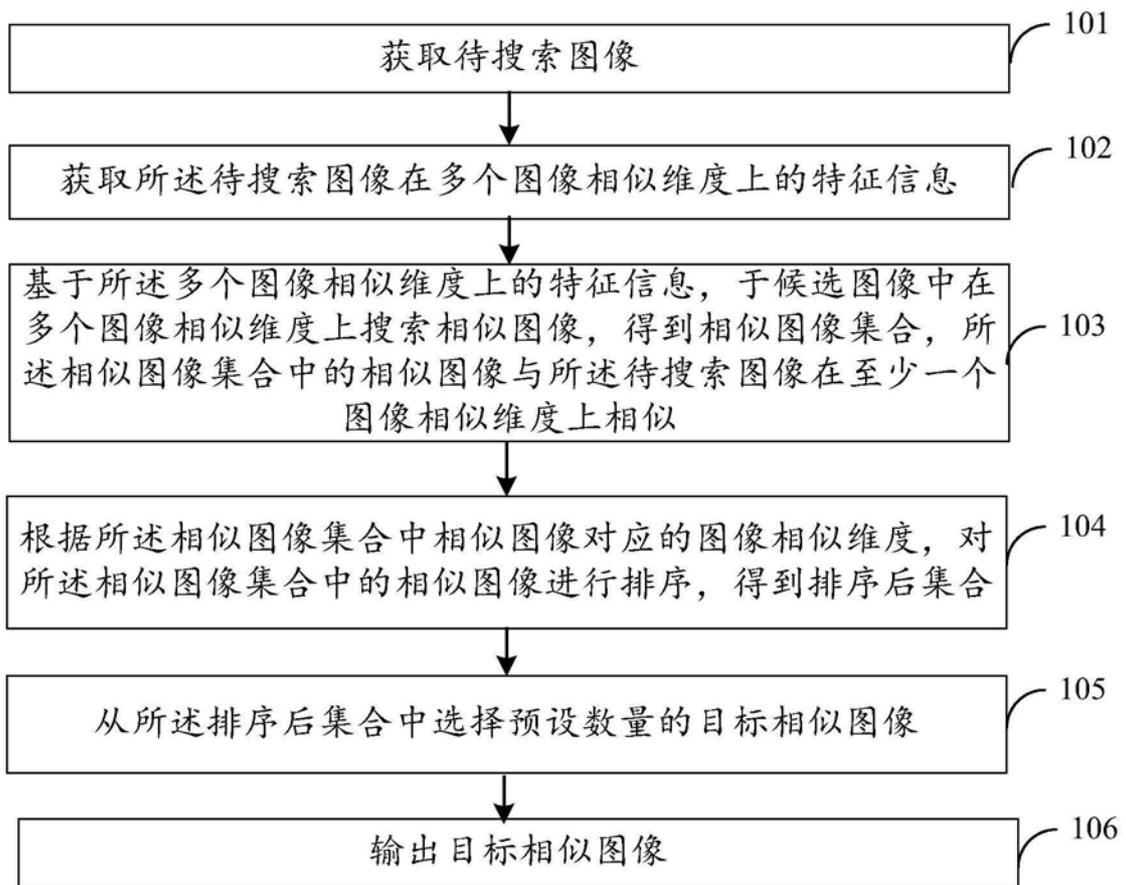


图1b

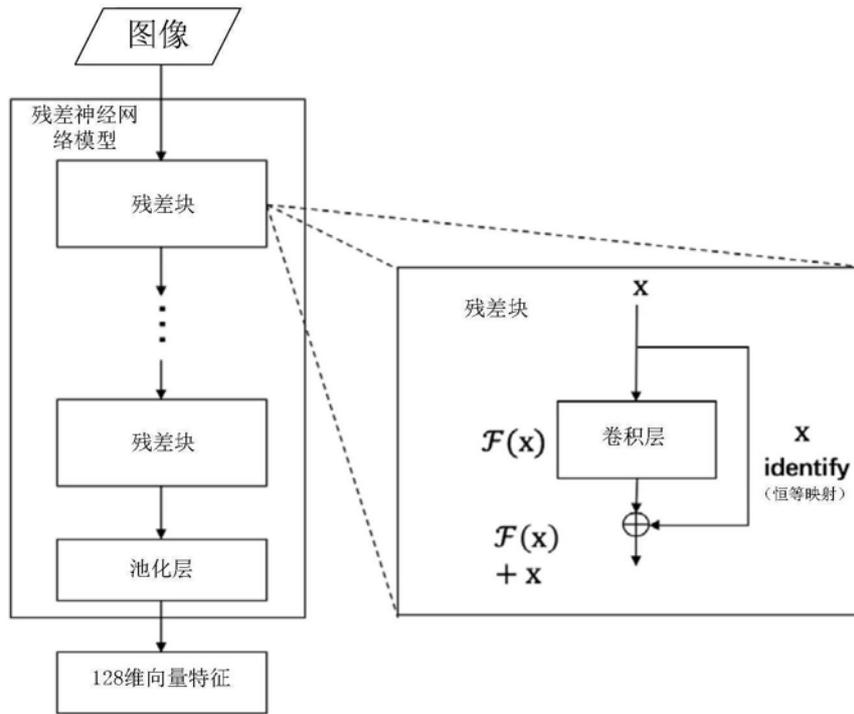


图2a

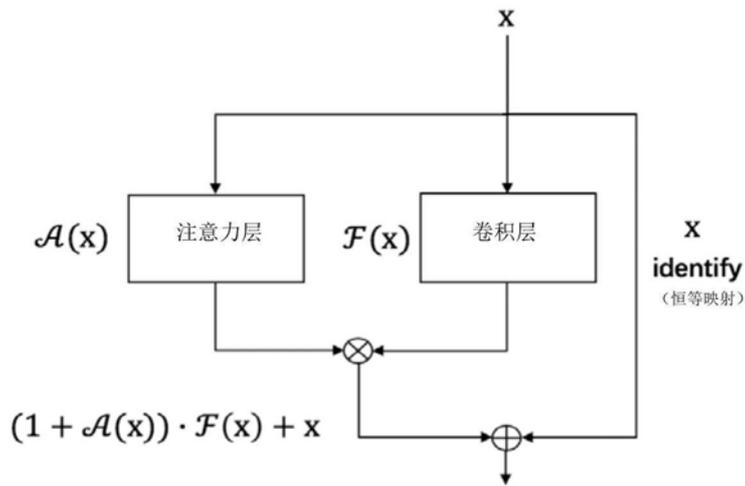


图2b

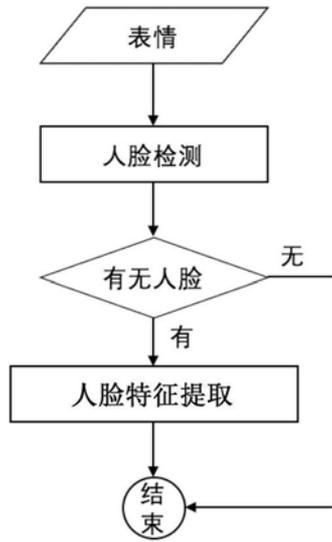


图2c

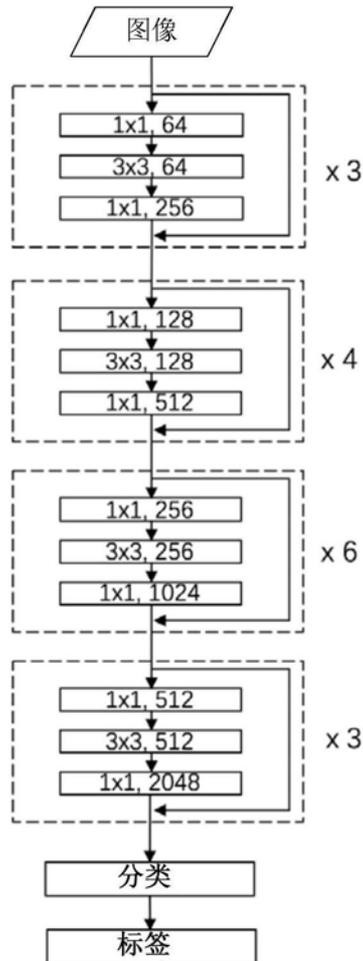


图3

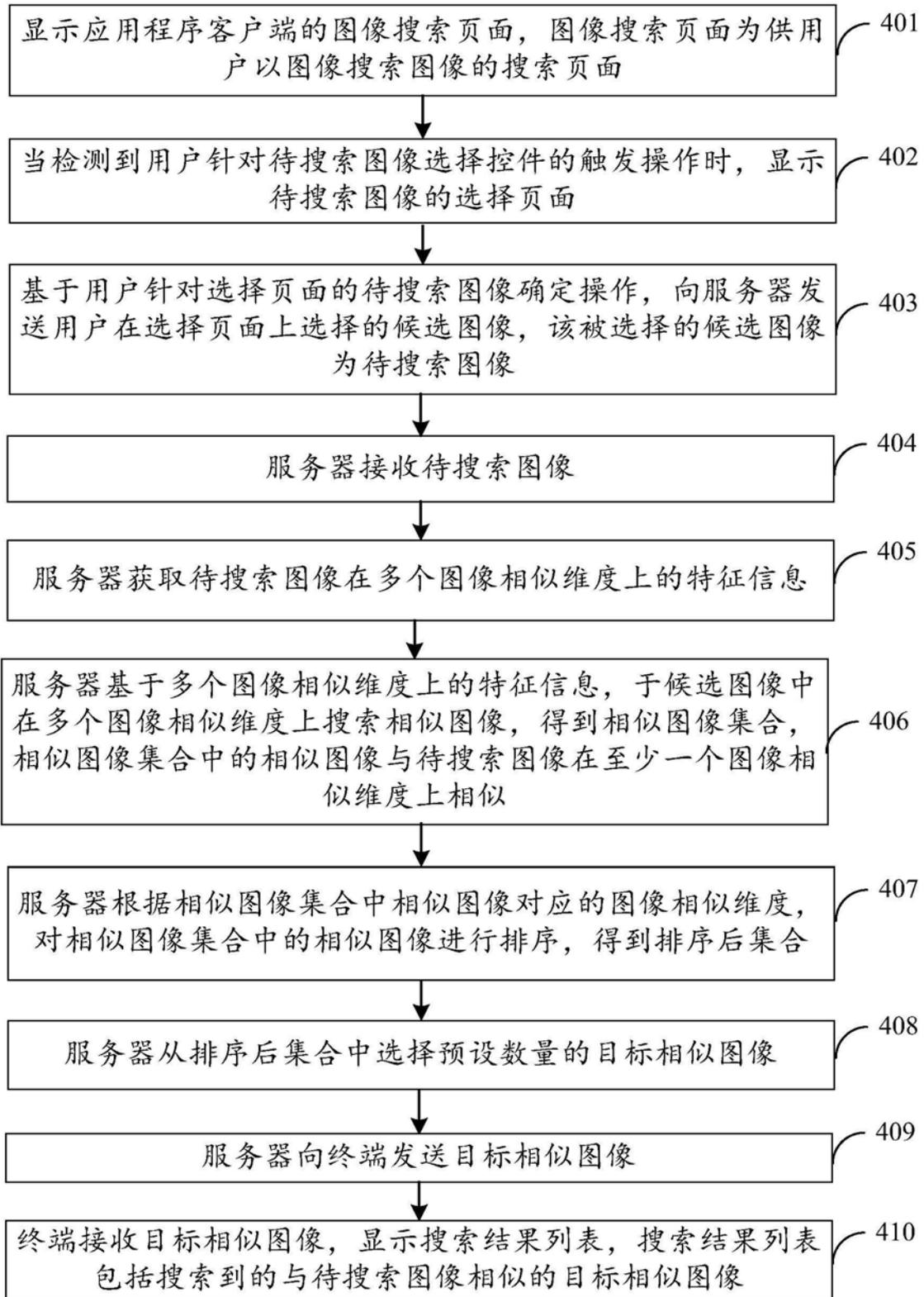


图4a

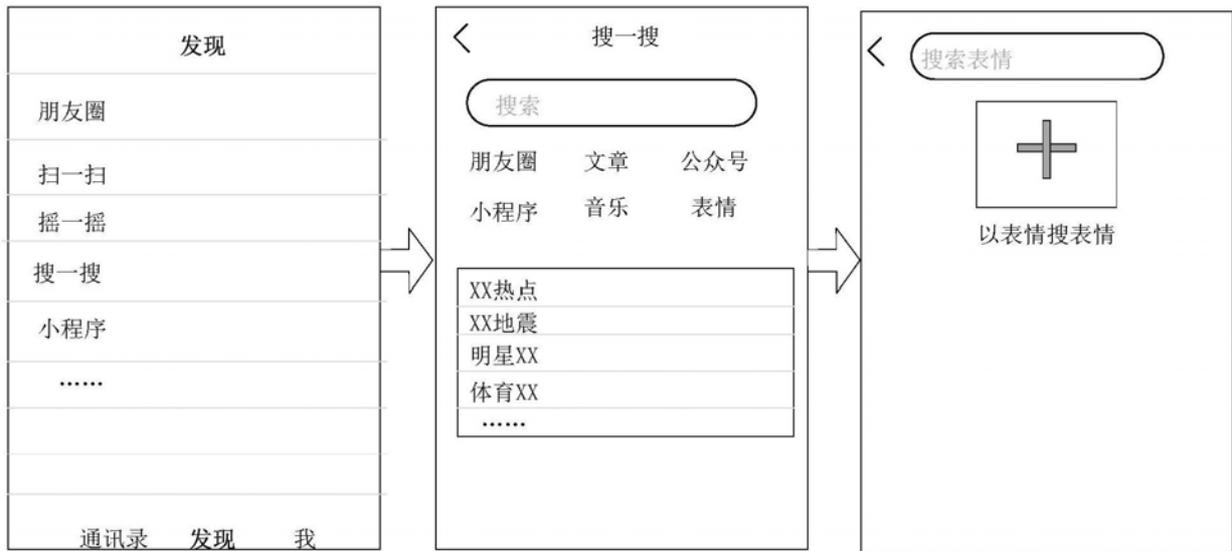


图4b

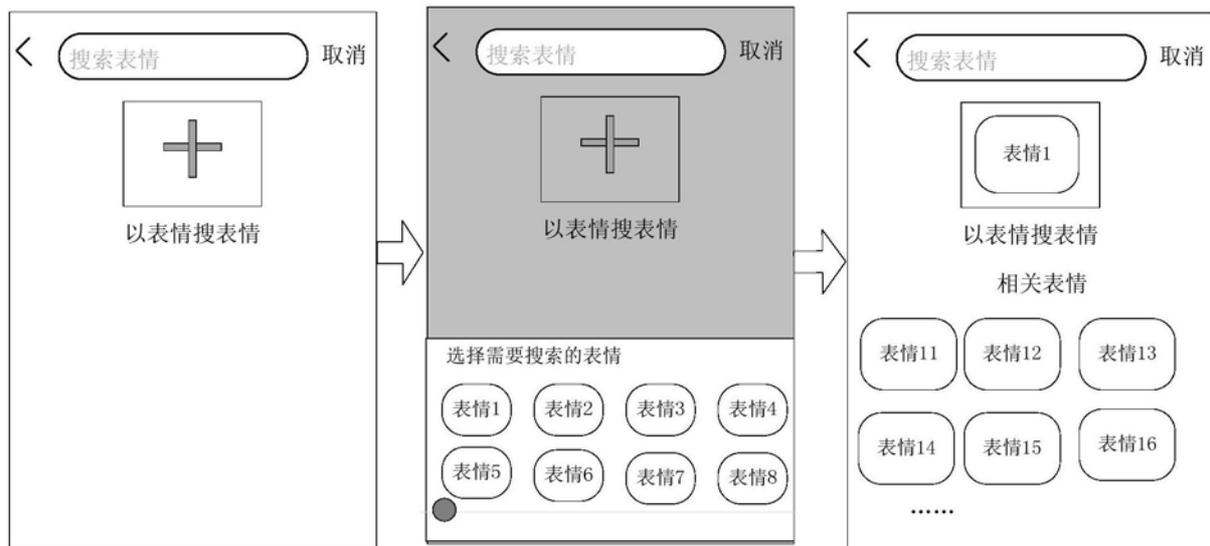


图4c



图4d

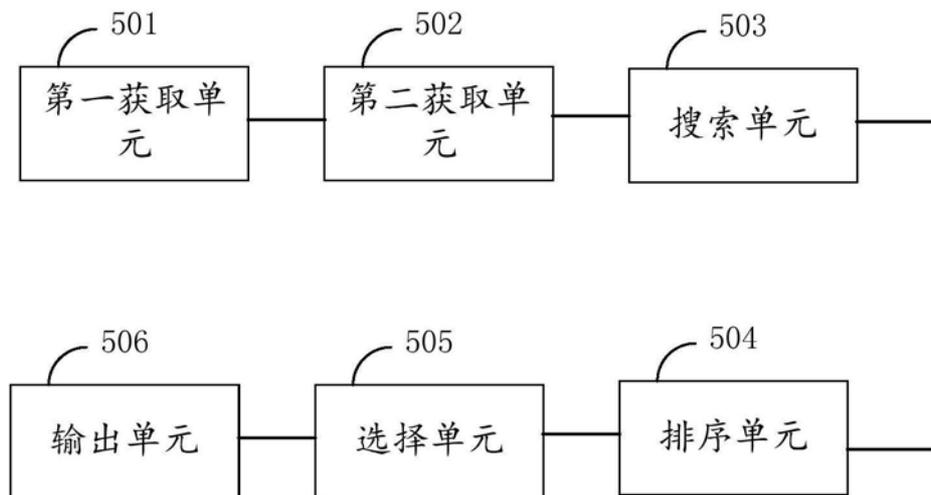


图5

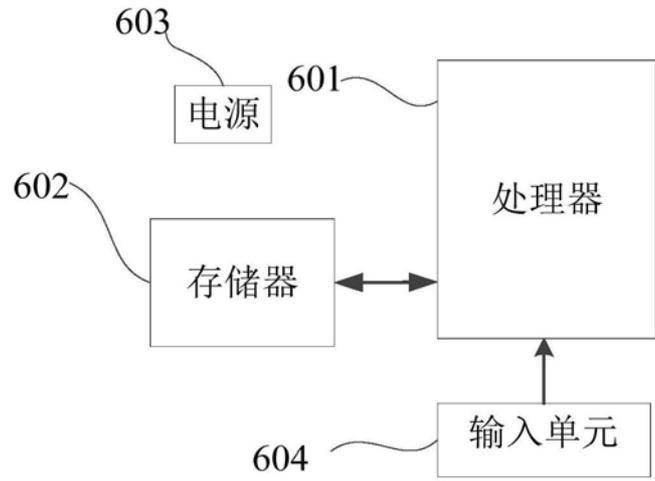


图6

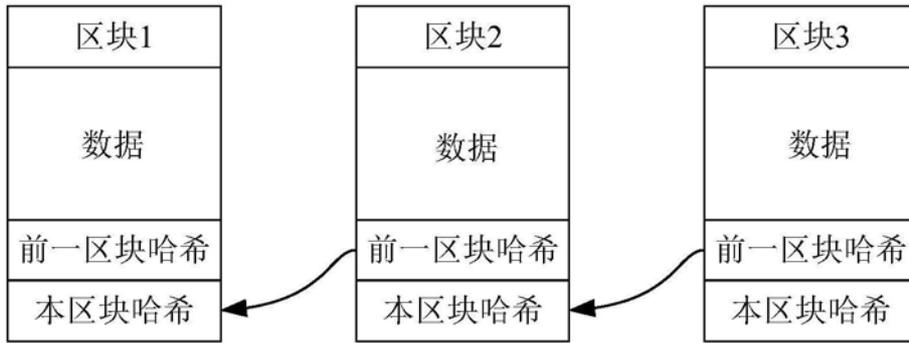


图8