



(12) 发明专利

(10) 授权公告号 CN 111373393 B

(45) 授权公告日 2022.05.31

(21) 申请号 201780097137.5

(22) 申请日 2017.11.24

(65) 同一申请的已公布的文献号
申请公布号 CN 111373393 A

(43) 申请公布日 2020.07.03

(85) PCT国际申请进入国家阶段日
2020.05.22

(86) PCT国际申请的申请数据
PCT/CN2017/112956 2017.11.24

(87) PCT国际申请的公布数据
W02019/100348 ZH 2019.05.31

(73) 专利权人 华为技术有限公司
地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72) 发明人 付宇新 温丰 薛常亮

(74) 专利代理机构 北京龙双利达知识产权代理有限公司 11329
专利代理师 王龙华 章慷

(51) Int. Cl.
G06F 16/583 (2019.01)
G06F 16/55 (2019.01)

G06V 10/44 (2022.01)
G06V 10/46 (2022.01)
G06V 10/54 (2022.01)
G06V 10/764 (2022.01)

(56) 对比文件

CN 103235955 A, 2013.08.07
CN 106855883 A, 2017.06.16
CN 104615676 A, 2015.05.13
US 2015269191 A1, 2015.09.24
US 2010088295 A1, 2010.04.08
US 9697234 B1, 2017.07.04
CN 104424226 A, 2015.03.18

张祯伟等.改进视觉词袋模型的快速图像检索方法.《计算机系统应用》.2016,(第12期),
张旺等.基于按类视觉词袋模型与标签投票的图像检索.《微型电脑应用》.2015,(第03期),
金铭等.一种基于视觉词袋模型的图像检索方法.《计算机应用与软件》.2017,(第04期),
吕林涛等.基于高层语义视觉词袋的色情图像过滤模型.《计算机应用》.2011,(第07期),

(续)

审查员 王辞

权利要求书4页 说明书21页 附图3页

(54) 发明名称

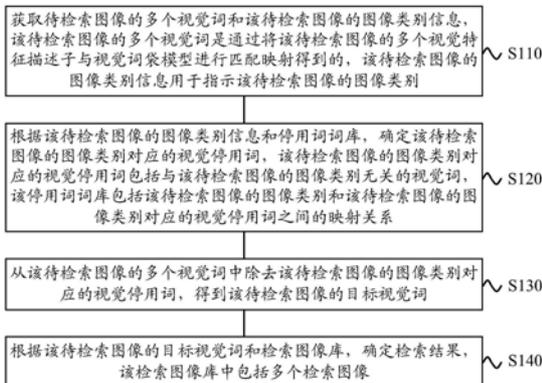
图像检索方法和装置以及图像库的生成方法和装置

(57) 摘要

一种图像检索方法和装置以及图像库的生成方法和装置,该图像检索方法包括:获取待检索图像的多个视觉词和该待检索图像的图像类别信息;根据该待检索图像的图像类别信息和停用词词库,确定该待检索图像的图像类别对应的视觉停用词,该待检索图像的图像类别对应的视觉停用词包括与该待检索图像的图像类别无关的视觉词;从该待检索图像的多个视觉词中除去该待检索图像的图像类别对应的视觉停用词,得到该待检索图像的目标视觉词(S130);根据该待

检索图像的目标视觉词和检索图像库,确定检索结果,该检索图像库中包括多个检索图像(S140)。能够提高图像检索的效率和精确度。

100



CN 111373393 B

[接上页]

(56) 对比文件

Kiana Hajebi等.Stopping Rules for
Bag-of-Words Image Search and Its

Application in Appearance-Based
Localization.《arXiv.org》.2013,

1. 一种图像检索方法,其特征在于,包括:

获取待检索图像的多个视觉词和所述待检索图像的图像类别信息,所述待检索图像的多个视觉词是通过将所述待检索图像的多个视觉特征描述子与视觉词袋模型中的视觉词进行匹配映射得到的,所述待检索图像的图像类别信息用于指示所述待检索图像的图像类别;

根据所述待检索图像的图像类别信息和停用词词库,确定所述待检索图像的图像类别对应的视觉停用词,所述待检索图像的图像类别对应的视觉停用词包括与所述待检索图像的图像类别无关的视觉词,所述停用词词库包括所述待检索图像的图像类别和所述待检索图像的图像类别对应的视觉停用词之间的映射关系;

从所述待检索图像的多个视觉词中除去所述待检索图像的图像类别对应的视觉停用词,得到所述待检索图像的目标视觉词;

根据所述待检索图像的目标视觉词和检索图像库,确定检索结果,所述检索图像库中包括多个检索图像。

2. 根据权利要求1所述的方法,其特征在于,所述检索图像库中包括所述多个检索图像和所述多个检索图像中每个检索图像对应的目标视觉词的映射关系,所述每个检索图像对应的目标视觉词是从所述每个检索图像对应的多个视觉词中除去所述每个检索图像的图像类别对应的视觉停用词后得到的。

3. 根据权利要求1或2所述的方法,其特征在于,在所述根据所述待检索图像的图像类别信息和停用词词库,确定所述待检索图像的图像类别对应的视觉停用词之前,所述方法还包括:

获取训练图像库中每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和正样本图像集合信息,所述每个训练图像的多个视觉词是通过将所述每个训练图像的多个视觉特征描述子与所述视觉词袋模型中的视觉词进行匹配映射得到的,所述每个训练图像的图像类别信息用于指示所述每个训练图像的图像类别,所述正样本图像集合信息用于指示至少一个正样本图像集合,所述正样本图像集合包括人工标注的所述训练图像库中的多个相似的训练图像;

根据所述每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和正样本图像集合信息,生成所述停用词词库。

4. 根据权利要求3所述的方法,其特征在于,所述根据所述每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和正样本图像集合信息,生成所述停用词词库,包括:

根据所述每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和所述正样本图像集合信息,确定所述训练图像库的多种图像类别与所述训练图像库的多个视觉词之间的相关性,所述训练图像库的多种图像类别包括所述每个训练图像的图像类别,所述训练图像库的多个视觉词包括所述每个训练图像的多个视觉词;

根据所述训练图像库的多种图像类别与所述训练图像库的多个视觉词之间的相关性,生成所述停用词词库。

5. 根据权利要求4所述的方法,其特征在于,所述根据所述每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和所述正样本图像集合信息,确定所述训练图像库

的多种图像类别与所述训练图像库的多个视觉词之间的相关性,包括:

根据第一事件发生的概率、第二事件发生的概率和所述第一事件与所述第二事件同时发生的概率,确定第一图像类别与第一视觉词之间的相关性,所述训练图像库的多种图像类别包括所述第一图像类别,所述训练图像库的多个视觉词包括所述第一视觉词,所述第一事件表示所述训练图像库中的第一训练图像的多个视觉词和所述训练图像库中的第二训练图像的多个视觉词均包括所述第一视觉词,所述第一训练图像的图像类别为所述第一图像类别,所述第二事件表示所述第一训练图像和所述第二训练图像属于相同的正样本图像集合。

6. 根据权利要求2所述的方法,其特征在于,所述根据所述待检索图像的目标视觉词和检索图像库,确定检索结果,包括:

确定所述待检索图像的目标视觉词与所述检索图像库中的检索图像的目标视觉词之间的相似度;

将与所述待检索图像的目标视觉词的相似度大于第一预设值的至少一个检索图像,确定为所述检索结果。

7. 一种图像库的生成方法,其特征在于,包括:

获取训练图像库中每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和正样本图像集合信息,所述每个训练图像的多个视觉词是通过将所述每个训练图像的多个视觉特征描述子与视觉词袋模型中的视觉词进行匹配映射得到的,所述每个训练图像的图像类别信息用于指示所述每个训练图像的图像类别,所述正样本图像集合信息用于指示至少一个正样本图像集合,所述正样本图像集合包括人工标注的所述训练图像库中的多个相似的训练图像;

根据所述每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和所述正样本图像集合信息,生成停用词词库,所述停用词词库包括所述每个训练图像的图像类别和与所述每个训练图像的图像类别对应的视觉停用词之间的映射关系,所述每个训练图像的图像类别对应的视觉停用词包括与所述每个训练图像的图像类别无关的视觉词。

8. 根据权利要求7所述的方法,其特征在于,所述根据所述每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和所述正样本图像集合信息,生成停用词词库,包括:

根据所述每个训练图像对应的多个视觉词、所述每个训练图像的图像类别信息和所述正样本图像集合信息,确定所述训练图像库的多种图像类别与所述训练图像库的多个视觉词之间的相关性,所述训练图像库的多种图像类别包括所述每个训练图像的图像类别,所述训练图像库的多个视觉词包括所述每个训练图像的多个视觉词;

根据所述训练图像库的多种图像类别与所述训练图像库的多个视觉词之间的相关性,生成所述停用词词库。

9. 一种图像检索装置,其特征在于,包括:

获取单元,用于获取待检索图像的多个视觉词和所述待检索图像的图像类别信息,所述待检索图像的多个视觉词是通过将所述待检索图像的多个视觉特征描述子与视觉词袋模型中的视觉词进行匹配映射得到的,所述待检索图像的图像类别信息用于指示所述待检索图像的图像类别;

处理单元,用于根据所述获取单元获取的所述待检索图像的图像类别信息和停用词词库,确定所述待检索图像的图像类别对应的视觉停用词,所述待检索图像的图像类别对应的视觉停用词包括与所述待检索图像的图像类别无关的视觉词,所述停用词词库包括所述待检索图像的图像类别和所述待检索图像的图像类别对应的视觉停用词之间的映射关系;

从所述获取单元获取的所述待检索图像的多个视觉词中除去所述待检索图像的图像类别对应的视觉停用词,得到所述待检索图像的目标视觉词;

检索单元,用于根据所述处理单元得到的所述待检索图像的目标视觉词和检索图像库,确定检索结果,所述检索图像库中包括多个检索图像。

10. 根据权利要求9所述的装置,其特征在于,所述检索图像库中包括所述多个检索图像和所述多个检索图像中每个检索图像对应的目标视觉词的映射关系,所述每个检索图像对应的目标视觉词是从所述每个检索图像对应的多个视觉词中除去所述每个检索图像的图像类别对应的视觉停用词后得到的。

11. 根据权利要求9或10所述的装置,其特征在于,所述装置还包括生成单元,

所述获取单元还用于在所述根据所述待检索图像的图像类别信息和停用词词库,确定所述待检索图像的图像类别对应的视觉停用词之前,获取训练图像库中每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和正样本图像集合信息,所述每个训练图像的多个视觉词是通过将所述每个训练图像的多个视觉特征描述子与所述视觉词袋模型中的视觉词进行匹配映射得到的,所述每个训练图像的图像类别信息用于指示所述每个训练图像的图像类别,所述正样本图像集合信息用于指示至少一个正样本图像集合,所述正样本图像集合包括人工标注的所述训练图像库中的多个相似的训练图像;

所述生成单元用于根据所述每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和正样本图像集合信息,生成所述停用词词库。

12. 根据权利要求11所述的装置,其特征在于,所述生成单元具体用于:

根据所述每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和所述正样本图像集合信息,确定所述训练图像库的多种图像类别与所述训练图像库的多个视觉词之间的相关性,所述训练图像库的多种图像类别包括所述每个训练图像的图像类别,所述训练图像库的多个视觉词包括所述每个训练图像的多个视觉词;

根据所述训练图像库的多种图像类别与所述训练图像库的多个视觉词之间的相关性,生成所述停用词词库。

13. 根据权利要求12所述的装置,其特征在于,所述生成单元具体用于:

根据第一事件发生的概率、第二事件发生的概率和所述第一事件与所述第二事件同时发生的概率,确定第一图像类别与第一视觉词之间的相关性,所述训练图像库的多种图像类别包括所述第一图像类别,所述训练图像库的多个视觉词包括所述第一视觉词,所述第一事件表示所述训练图像库中的第一训练图像的多个视觉词和所述训练图像库中的第二训练图像的多个视觉词均包括所述第一视觉词,所述第一训练图像的图像类别为所述第一图像类别,所述第二事件表示所述第一训练图像和所述第二训练图像属于相同的正样本图像集合。

14. 根据权利要求10所述的装置,其特征在于,所述检索单元具体用于:

确定所述待检索图像的目标视觉词与所述检索图像库中的检索图像的目标视觉词之

间的相似度；

将与所述待检索图像的目标视觉词的相似度大于第一预设值的至少一个检索图像，确定为所述检索结果。

15. 一种图像库的生成装置，其特征在于，包括：

获取单元，用于获取训练图像库中每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和正样本图像集合信息，所述每个训练图像的多个视觉词是通过将所述每个训练图像的多个视觉特征描述子与视觉词袋模型中的视觉词进行匹配映射得到的，所述每个训练图像的图像类别信息用于指示所述每个训练图像的图像类别，所述正样本图像集合信息用于指示至少一个正样本图像集合，所述正样本图像集合包括人工标注的所述训练图像库中的多个相似的训练图像；

生成单元，用于根据所述获取单元获取的所述每个训练图像的多个视觉词、所述每个训练图像的图像类别信息和所述正样本图像集合信息，生成停用词词库，所述停用词词库包括所述每个训练图像的图像类别和与所述每个训练图像的图像类别对应的视觉停用词之间的映射关系，所述每个训练图像的图像类别对应的视觉停用词包括与所述每个训练图像的图像类别无关的视觉词。

16. 根据权利要求15所述的装置，其特征在于，所述生成单元具体用于：

根据所述每个训练图像对应的多个视觉词、所述每个训练图像的图像类别信息和所述正样本图像集合信息，确定所述训练图像库的多种图像类别与所述训练图像库的多个视觉词之间的相关性，所述训练图像库的多种图像类别包括所述每个训练图像的图像类别，所述训练图像库的多个视觉词包括所述每个训练图像的多个视觉词；

根据所述训练图像库的多种图像类别与所述训练图像库的多个视觉词之间的相关性，生成所述停用词词库。

17. 一种图像检索装置，所述装置包括存储器、处理器、通信接口及存储在所述存储器上并可在所述处理器上运行的计算机程序，其中，所述存储器、所述处理器以及所述通信接口之间通过内部连接通路互相通信，其特征在于，所述处理器执行所述计算机程序时执行上述权利要求1至权利要求6中任一项所述的方法。

18. 一种图像库的生成装置，所述装置包括存储器、处理器、通信接口及存储在所述存储器上并可在所述处理器上运行的计算机程序，其中，所述存储器、所述处理器以及所述通信接口之间通过内部连接通路互相通信，其特征在于，所述处理器执行所述计算机程序时执行上述权利要求7或权利要求8所述的方法。

19. 一种计算机可读介质，用于存储计算机程序，其特征在于，所述计算机程序包括用于执行上述权利要求1至权利要求6中任一项所述的方法的指令。

20. 一种计算机可读介质，用于存储计算机程序，其特征在于，所述计算机程序包括用于执行上述权利要求7或权利要求8所述的方法的指令。

图像检索方法和装置以及图像库的生成方法和装置

技术领域

[0001] 本申请涉及图像检索技术领域,更具体地,涉及图像检索技术领域中图像检索方法和装置以及图像库的生成方法和装置。

背景技术

[0002] 视觉词袋(bag of visual words,BoVW)模型被广泛应用到图像检索领域,视觉词袋模型包括多个视觉词,该多个视觉词是对从多个图像提取的多个视觉特征描述子进行聚类得到的,该多个视觉词中的每个视觉词为一个聚类中心。

[0003] 现有的图像检索过程中,首先获取待检索图像的多个视觉特征描述子,将该多个视觉特征描述子与该视觉词袋模型中的视觉词进行匹配映射,得到该待检索图像的多个视觉词,该多个视觉词用于表示该待检索图像,根据该待检索图像的多个视觉词,计算该待检索图像与检索图像库中的检索图像的相似度,将该检索图像库中与该待检索图像的相似度最高的至少一个图像作为图像检索结果输出。

[0004] 然而,当该待检索图像的内容杂乱繁多,或该待检索图像蕴含的信息量较大的时候,该待检索图像的多个视觉词的数量较大,因此,在进行图像检索的时候,效率比较低,精确度也比较差。

发明内容

[0005] 本申请提供一种图像检索方法和装置以及图像处理方法和装置,有利于提高图像检索的效率和精确度。

[0006] 第一方面,本申请提供了一种图像检索方法,该方法包括:

[0007] 获取待检索图像的多个视觉词和该待检索图像的图像类别信息,该待检索图像的多个视觉词是通过将该待检索图像的多个视觉特征描述子与视觉词袋模型中的视觉词进行匹配映射得到的,该待检索图像的图像类别信息用于指示该待检索图像的图像类别;

[0008] 根据该待检索图像的图像类别信息和停用词词库,确定该待检索图像的图像类别对应的视觉停用词,该待检索图像的图像类别对应的视觉停用词包括与该待检索图像的图像类别无关的视觉词,该停用词词库包括该待检索图像的图像类别和该待检索图像的图像类别对应的视觉停用词之间的映射关系;

[0009] 从该待检索图像的多个视觉词中除去该待检索图像的图像类别对应的视觉停用词,得到该待检索图像的目标视觉词;

[0010] 根据该待检索图像的目标视觉词和检索图像库,确定检索结果,该检索图像库中包括多个检索图像。

[0011] 本申请实施例提供的图像检索方法,该待检索图像的目标视觉词是通过从该待检索图像的多个视觉词中除去该待检索图像的图像类别对应的视觉停用词后得到的,即从该待检索图像的多个视觉词中除去了对于辨识该检索图像无显著作用、或影响图像识别的视觉词,也就是说,该待检索图像的目标视觉词对于识别该待检索图像的作用比较显著。因

此,通过该待检索图像的目标视觉词与检索图像库进行检索,有利于提高图像检索的效率和精确度。

[0012] 在一种可能的实现方式中,该检索图像库中还包括该多个检索图像中每个检索图像对应的目标视觉词,该每个检索图像对应的目标视觉词是从该每个检索图像对应的多个视觉词中除去该每个检索图像的图像类别对应的视觉停用词后得到的。

[0013] 本申请实施例提供的图像检索方法,该检索图像库中存储的检索图像的目标视觉词是通过从检索图像的多个视觉词中除去了该检索图像的图像类别对应的视觉停用词后得到的,有利于减少该检索图像库内存的占用率。

[0014] 此外,根据待检索图像的目标视觉词与检索图像库中检索图像的目标视觉词的相似度,确定检索结果,有利于提高图像检索的效率和精确度。

[0015] 应理解,该图像检索装置可以为具有计算和存储功能的第一设备,该第一设备例如可以为计算机,或者该图像检索装置可以为第一设备中的功能模块,本申请实施例对此不作限定。

[0016] 还应理解,本申请实施例中的图像的视觉特征点,是指图像经比例、旋转、平移、视角等变换还能保持一致性的像素点,即图像中最容易识别的像素点,例如角点或者纹理丰富的边缘点。图像的视觉特征点质量好坏将直接影响图像检索的效率和精度。

[0017] 可选地,图像的视觉特征点的类型可以包括尺度不变特征转换(scale-invariant feature transform,SIFT)、ORB、加速稳健特征(speeded up robust feature,SURF)、加速分割测试获得特征(features from accelerated segment test,FAST)等,本申请实施例对此不作限定。

[0018] 可选地,图像的视觉特征点可以为一个或多个,本申请实施例对此不作限定。

[0019] 还应理解,本申请实施例中的图像的视觉特征描述子,是指通过数学特征描述图像的视觉特征点。

[0020] 例如,以ORB为例,获取图像的视觉特征描述子的主要步骤包括:在图像的视觉特征点附近随机选取若干个像素对,通过比较每个像素对中两个像素之间的大小关系得到0或1的编码;利用视觉特征点方向的信息,将视觉特征点旋转得到鲁棒的二进制向量视觉特征描述子。

[0021] 可选地,图像的视觉特征描述子可以为一个或多个,本申请实施例对此不作限定。

[0022] 还应理解,本申请实施例中的视觉词袋模型包括多个视觉词,该多个视觉词中的每个视觉词是对从多个图像中提取的视觉特征描述子进行聚类后得到的一个聚类中心。

[0023] 还应理解,本申请实施例中的图像的视觉词,是指通过将该图像的视觉特征描述子和视觉词袋模型中的视觉词进行匹配映射,得到的该视觉词袋模型中与该视觉特征描述子距离最近的视觉词。

[0024] 可选地,图像的视觉词可以为一个或多个,本申请实施例对此不作限定。

[0025] 还应理解,本申请实施例中根据不同的分类方法对多个图像进行分类,能够得到每个图像的图像类别。

[0026] 作为一个可选实施例,如果按照场景对图像进行分类,则图像的图像类别可以包括森林场景、郊区场景、室内场景等。

[0027] 作为另一个可选实施例,如果按照天气对图像进行分类,则图像的图像类别可以

包括晴天、雨天、雪天等。

[0028] 还应理解,由于同一个图像中出现的不同视觉词可能对辨识该图像具有不同的作用,不同图像中出现的相同视觉词可能对辨识这两个图像具有相同的作用,因此,本申请实施例中图像类别对应的视觉停用词是指对于辨识某种图像类别的图像无显著作用、或影响图像识别的视觉词,即与该图像类别的图像无关的视觉词。

[0029] 应理解,本申请实施例中所述的与某种图像类别无关的视觉停用词,是指与该种图像类别的图像的相关性低于预设阈值的视觉词。

[0030] 可选地,图像类别对应的视觉停用词可以包括一个或多个视觉词,本申请实施例对此不作限定。

[0031] 例如,在森林场景和郊区场景下,几乎每个图像都包含大量的树木,从图像中的树木提取的特征点对辨识该图像是森林场景还是郊区场景的辨识度较低,因此,树木可以为森林类或郊区类的视觉停用词。

[0032] 又例如,在雨天,图像中会留下雨水降落的痕迹,从图像中的雨水提取的特征点也会对该图像的多个视觉词造成污染,因此,雨水可以为雨天的视觉停用词。

[0033] 还应理解,本申请实施例中的图像的目标视觉词包括从该图像的多个视觉词中除去该图像的图像类别对应的视觉停用词后的视觉词。

[0034] 可选地,图像的目标视觉词可以包括一个或多个视觉词,本申请实施例对此不作限定。

[0035] 还应理解,本申请实施例中的正样本图像集合中包括人工标注的可以被认为是相似度高或相同的图像。

[0036] 例如,在不同场景拍摄同一物体的两个图像,例如,雨天的学校和雪天的学校。

[0037] 例如,在不同时刻拍摄同一场景的两个图像,例如回环检测中同一场景的当前位姿和历史位姿。

[0038] 在一种可能的实现方式中,在该根据该待检索图像的图像类别信息和停用词词库,确定该待检索图像的图像类别对应的视觉停用词之前,该方法还包括:获取训练图像库中每个训练图像的多个视觉词、该每个训练图像的图像类别信息和正样本图像集合信息,该每个训练图像的多个视觉词是通过将该每个训练图像的多个视觉特征描述子与该视觉词袋模型中的视觉词进行匹配映射得到的,该每个训练图像的图像类别信息用于指示该每个训练图像的图像类别,该正样本图像集合信息用于指示至少一个正样本图像集合,该正样本图像集合包括人工标注的该训练图像库中的多个相似的训练图像;根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和正样本图像集合信息,生成该停用词词库。

[0039] 在一种可能的实现方式中,该根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和正样本图像集合信息,生成该停用词词库,包括:根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,确定该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,该训练图像库的多种图像类别包括该每个训练图像的图像类别,该训练图像库的多个视觉词包括该每个训练图像的多个视觉词;根据该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,生成该停用词词库。

[0040] 作为一个可选实施例,该图像检索装置可以将该训练图像库对应的多个视觉词中与该训练图像库对应每种图像类别的相关性最小的至少一个视觉词,作为该每种图像类别对应的视觉停用词。

[0041] 作为另一个可选实施例,该图像检索装置可以将该训练图像库对应的多个视觉词中与该训练图像库对应每种图像类别的相关性小于第一预设阈值的至少一个视觉词,作为该每种图像类别对应的视觉停用词。

[0042] 可选地,每种图像类别对应的视觉停用词可以包括一个或多个视觉词,本申请实施例对此不作限定。

[0043] 可选地,该停用词词库可以包括该每种图像类别和与该每种图像类别对应的视觉停用词的映射关系。

[0044] 应理解,由于同一个图像中出现的不同视觉词可能对辨识该图像具有不同的作用,不同图像中出现的相同视觉词可能对辨识这两个图像具有相同的作用。

[0045] 在一种可能的实现方式中,该根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,确定该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,包括:根据第一事件发生的概率、第二事件发生的概率和该第一事件与该第二事件同时发生的概率,确定第一图像类别与第一视觉词之间的相关性,该训练图像库的多种图像类别包括该第一图像类别,该训练图像库的多个视觉词包括该第一视觉词,该第一事件表示该训练图像库中的第一训练图像的多个视觉词和该训练图像库中的第二训练图像的多个视觉词均包括该第一视觉词,该第一训练图像的图像类别为该第一图像类别,该第二事件表示该第一训练图像和该第二训练图像属于相同的正样本图像集合。

[0046] 例如,假设该训练图像集中共有P个训练图像,M种图像类别,L个视觉词,该P个训练图像中第一类别的训练图像共有N个,则第一视觉词和该第一图像类别的训练图像的相关性可以通过公式(1)至公式(6)确定:

$$[0047] \quad p(x) = \frac{\sum_{i=1}^N \text{count}(x_i)}{P * N} \quad (1)$$

$$[0048] \quad p(y) = \frac{\sum_{i=1}^N \text{count}(y_i)}{P * N} \quad (2)$$

$$[0049] \quad p(x, y) = \frac{\sum_{i=1}^N \text{count}(x_i, y_i)}{P * N} \quad (3)$$

$$[0050] \quad PMI(x, y) = \ln \frac{p(x, y)}{p(x)p(y)} \quad (4)$$

$$[0051] \quad H(y) = \ln \frac{P * N}{\text{count}(y)} \quad (5)$$

$$[0052] \quad RATE_{PMI}(x, y) = \frac{PMI(x, y)}{H(y)} \quad (6)$$

[0053] 其中, x 表示第一事件,该第一事件为该 N 个训练图像中的第一训练图像的多个视觉词和该 P 个训练图像中除该第一训练图像外的第二训练图像的多个视觉词均包括该 L 个视觉词中的第一视觉词, x_i 表示第一类别的训练图像中第 i 个图像发生的第一事件; y 表示第二事件,该第二事件为该第一训练图像和该第二训练图像属于同一个正样本图像集合, y_i 表示第一类别的训练图像中第 i 个图像发生的第二事件, $\text{count}(x_i)$ 表示第一类别的训练图像中第 i 个图像发生第一事件的次数, $\text{count}(y)$ 表示第二事件发生的次数, $\text{count}(y_i)$ 表示第一类别的训练图像中第 i 个图像发生第二事件的次数, $\text{count}(x_i, y_i)$ 表示第一类别的训练图像中第 i 个图像中第一事件和第二事件同时发生的次数, $p(x)$ 表示第一事件发生的概率, $p(y)$ 表示第二事件发生的概率, $p(x, y)$ 表示第一事件与第二事件同时发生的概率, $\text{PMI}(x, y)$ 表示第一事件和第二事件的点互信息量, $H(y)$ 表示第二事件的信息熵, $\text{RATE}_{\text{PMI}}(x, y)$ 表示第一事件和第二事件的点互信息率即该第一视觉词和该第一图像类别的相关性,其中, P 、 L 、 M 、 N 均为大于1的正整数。

[0054] 可选地,该图像检索装置可以自己生成该检索图像库,或者可以从图像库的生成装置获取该检索图像库,本申请实施例对此不作限定。

[0055] 可选地,该检索图像库可以为根据该多个训练图像训练得到的,或者该检索图像库可以为根据该图像检索装置在本次检索之前检索的历史待检索图像训练得到的,或者可以为根据其他图像训练得到的,本申请实施例对此不作限定。

[0056] 作为一个可选实施例,该图像检索装置可以根据该多个训练图像中每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该停用词词库,生成该检索图像库。也就是说,该检索图像库是根据该多个训练图像训练得到的。

[0057] 具体而言,该图像检索装置可以根据该每个训练图像的图像类别信息和该停用词词库,确定该每个训练图像的图像类别对应的视觉停用词,从该每个训练图像的多个视觉词中除去该每个训练图像的图像类别对应的视觉停用词,得到该每个训练图像的目标视觉词,并将该每个训练图像的目标视觉词添加至该检索图像库。

[0058] 可选地,该图像检索装置还可以使用本申请实施例中的停用词词库获得该每个训练图像的目标视觉词,或者可以通过其他方式获得该每个训练图像的目标视觉词,本申请实施例对此不作限定。

[0059] 例如,电子商务中,该检索图像库包括用户提供的所有商品图像的目标视觉词,该检索图像为用户想要购买的商品图像。

[0060] 作为另一个可选实施例,该图像检索装置可以将 $S140$ 之前检索的历史待检索图像的目标视觉词添加至该检索图像库,以生成该检索图像库。

[0061] 例如,回环检测中,检索图像库包括所有历史位姿图像,该待检索图像为当前位姿图像。

[0062] 本申请实施例提供的图像检索方法,在回环检测场景中,能够通过保存历史出现过的场景,利用当前图像进行检索识别回环,构建当前位姿和历史位姿的一个约束,通过优化减小整体误差,以得到全局一致的地图。在电子商务场景中,在不清楚商品名称时,用户提交商品的图像,系统根据该商品的图像进行检索,并返回相似度较高的图像作为检索结果。

[0063] 可选地,该图像检索装置可以根据多种方式从该检索图像库中选出与该待检索图

像最相似的至少一个检索图像作为检索结果输出,本申请实施例对此不作限定。

[0064] 作为一个可选实施例,该图像检索装置可以计算该待检索图像的多个视觉词与该检索图像库中的检索图像的多个视觉词之间的相似度,将与该待检索图像的相似度最高的至少一个检索图像确定为检索结果。

[0065] 作为另一个可选实施例,该图像检索装置可以计算该待检索图像的多个视觉词与该检索图像库中的检索图像的多个视觉词之间的相似度,将相似度大于第二预设阈值的至少一个检索图像确定为检索结果。

[0066] 本申请实施例提供的图像检索方法,该检索图像库中存储的检索图像的目标视觉词是通过从检索图像的多个视觉词中除去了该检索图像的图像类别对应的视觉停用词后得到的,有利于减少该检索图像库内存的占用率。

[0067] 此外,根据待检索图像的目标视觉词与检索图像库中检索图像的目标视觉词的相似度,确定与该待检索图像相似的至少一个检索图像,以得到检索结果,有利于提高图像检索的效率和精确度。

[0068] 第二方面,本申请提供了一种图像处理方法,该方法包括:

[0069] 获取训练图像库中每个训练图像的多个视觉词、该每个训练图像的图像类别信息和正样本图像集合信息,该每个训练图像的多个视觉词是通过将该每个训练图像的多个视觉特征描述子与该视觉词袋模型中的视觉词进行匹配映射得到的,该每个训练图像的图像类别信息用于指示该每个训练图像的图像类别,该正样本图像集合信息用于指示至少一个正样本图像集合,该正样本图像集合包括人工标注的该训练图像库中的多个相似的训练图像。

[0070] 根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,生成停用词词库,该停用词词库包括该每个训练图像的图像类别和与该每个图像的图像类别对应的视觉停用词之间的映射关系,该每个训练图像的图像类别对应的视觉停用词包括与该每个训练图像的图像类别无关的视觉词。

[0071] 本申请实施例提供的数据库的生成方法,通过获取训练图像库中每个训练图像的多个视觉词、该每个训练图像的图像类别信息和正样本图像集合信息,并根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,生成停用词词库,有利于提高图像检索的效率和精确度。

[0072] 应理解,该图像库的生成装置可以为具有计算和存储功能的第二设备,该第二设备例如可以为计算机,或者该图像库生成装置可以为该第二设备中的功能模块,本申请实施例对此不作限定。

[0073] 可选地,该第二设备与第一方面中的第一设备可以为相同的设备或不同的设备,本申请实施例对此不作限定。

[0074] 可选地,该第二设备与该第二设备相同时,该图像库的生成装置与第一方面中的图像检索装置为同一个设备中的不同功能模块,或者该图像库的生成装置为该图像检索装置中的功能模块。

[0075] 在一种可能的实现方式中,该根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,生成该停用词词库,包括:根据该每个训练图像对应的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,确定该

训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,该训练图像库的多种图像类别包括该每个训练图像的图像类别,该训练图像库的多个视觉词包括该每个训练图像的多个视觉词;根据该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,生成该停用词词库。

[0076] 在一种可能的实现方式中,该根据该每个训练图像对应的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,确定该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,包括:根据第一事件发生的概率、第二事件发生的概率和该第一事件与该第二事件同时发生的概率,确定第一图像类别与第一视觉词之间的相关性,该训练图像库的多种图像类别包括该第一图像类别,该训练图像库的多个视觉词包括该第一视觉词,该第一事件表示该训练图像库中的第一训练图像的多个视觉词和该训练图像库中的第二训练图像的多个视觉词均包括该第一视觉词,该第一训练图像的图像类别为该第一图像类别,该第二事件表示该第一训练图像和该第二训练图像属于相同的正样本图像集合。

[0077] 在一种可能的实现方式中,在该根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,生成停用词词库之后,该方法还包括:根据该每个训练图像的图像类别信息和该停用词词库,确定该每个训练图像的图像类别对应的视觉停用词,从该每个训练图像的多个视觉词中除去该每个训练图像的图像类别对应的视觉停用词,得到该每个训练图像的目标视觉词,并将该每个训练图像的目标视觉词添加至该检索图像库。

[0078] 在一种可能的实现方式中,该获取训练图像库中每个训练图像的多个视觉词,包括:获取该每个训练图像,提取该每个训练图像的多个视觉特征描述子,该多个视觉特征描述子用于描述该每个训练图像的多个视觉特征点,该多个视觉特征描述子与该多个视觉特征点一一对应,获取视觉词袋模型,将视觉词袋模型中与该多个视觉特征描述子中的每个视觉特征描述子的距离最近的多个视觉词,确定为该每个训练图像的多个视觉词。

[0079] 第三方面,本申请提供了一种图像检索装置,用于执行上述第一方面或第一方面的任意可能的实现方式中的方法。

[0080] 第四方面,本申请提供了一种图像处理装置,用于执行上述第二方面或第二方面的任意可能的实现方式中的方法。

[0081] 第五方面,本申请提供了一种图像检索装置,该装置包括:存储器、处理器、通信接口及存储在该存储器上并可在该处理器上运行的计算机程序,其特征在于,该处理器执行该计算机程序时执行上述第一方面或第一方面的任意可能的实现方式中的方法。

[0082] 第六方面,本申请提供了一种图像处理装置,该装置包括:存储器、处理器、通信接口及存储在该存储器上并可在该处理器上运行的计算机程序,其特征在于,该处理器执行该计算机程序时执行上述第二方面或第二方面的任意可能的实现方式中的方法。

[0083] 第七方面,本申请提供了一种计算机可读介质,用于存储计算机程序,该计算机程序包括用于执行第一方面或第一方面的任意可能的实现方式中的方法的指令。

[0084] 第八方面,本申请提供了一种计算机可读介质,用于存储计算机程序,该计算机程序包括用于执行第二方面或第二方面的任意可能的实现方式中的方法的指令。

[0085] 第九方面,本申请提供了一种包含指令的计算机程序产品,当其在计算机上运行

时,使得计算机执行上述第一方面或第一方面的任意可能的实现方式中的方法。

[0086] 第十方面,本申请提供了一种包含指令的计算机程序产品,当其在计算机上运行时,使得计算机执行上述第二方面或第二方面的任意可能的实现方式中的方法。

[0087] 第十一方面,本申请提供了一种芯片,包括:输入接口、输出接口、至少一个处理器、存储器,所述输入接口、输出接口、所述处理器以及所述存储器之间通过内部连接通路互相通信,所述处理器用于执行所述存储器中的代码,当所述代码被执行时,所述处理器用于执行上述第一方面或第一方面的任意可能的实现方式中的方法。

[0088] 第十二方面,本申请提供了一种芯片,包括:输入接口、输出接口、至少一个处理器、存储器,所述输入接口、输出接口、所述处理器以及所述存储器之间通过内部连接通路互相通信,所述处理器用于执行所述存储器中的代码,当所述代码被执行时,所述处理器用于执行上述第二方面或第二方面的任意可能的实现方式中的方法。

附图说明

[0089] 图1是本申请实施例的图像检索方法的示意性流程图;

[0090] 图2是本申请实施例的图像库的生成方法的示意性框图;

[0091] 图3是本申请实施例的图像检索装置的示意性框图;

[0092] 图4是本申请实施例的图像库的生成装置的示意性框图;

[0093] 图5是本申请实施例的另一图像检索装置的示意性框图;

[0094] 图6是本申请实施例的另一图像库的生成装置的示意性框图。

具体实施方式

[0095] 下面将结合附图,对本申请中的技术方案进行描述。

[0096] 为了清楚起见,首先对本申请中所使用的术语作以解释。

[0097] 1、图像的视觉特征点

[0098] 图像的视觉特征点是指图像经比例、旋转、平移、视角等变换还能保持一致性的像素点,即图像中最容易识别的像素点,例如角点或者纹理丰富的边缘点。图像的视觉特征点质量好坏将直接影响图像检索的效率和精度。

[0099] 可选地,图像的视觉特征点的类型可以包括尺度不变特征转换(scale-invariant feature transform,SIFT)、ORB、加速稳健特征(speeded up robust feature,SURF)、加速分割测试获得特征(features from accelerated segment test,FAST)等,本申请实施例对此不作限定。

[0100] 可选地,图像的视觉特征点可以为一个或多个,本申请实施例对此不作限定。

[0101] 例如,以ORB为例,对该图像进行FAST角点的提取的主要步骤包括:计算该图像中每个像素与其邻域像素亮度之间的差异,如果像素与它邻域的像素差别较大,则它更可能是角点;然后通过非极大值抑制,在一定区域内仅保留响应极大值的角点,避免角点集中的问题;针对FAST角点不具有方向性和尺度的弱点,添加尺度和旋转的描述。尺度不变性通过构建图像金字塔,对图像进行不同层次的降采样,获得不同分辨率的图像来实现。旋转不变性则通过灰度质心法实现,即通过计算图像块灰度值的质心和几何中心连接得到的方向向量来作为特征点方向的描述。

[0102] 2、图像的视觉特征描述子

[0103] 图像的视觉特征描述子是指通过数学特征描述图像的视觉特征点。

[0104] 例如,以ORB为例,获取图像的视觉特征描述子的主要步骤包括:在图像的视觉特征点附近随机选取若干个像素对,通过比较每个像素对中两个像素之间的大小关系得到0或1的编码;利用视觉特征点方向的信息,将视觉特征点旋转得到鲁棒的二进制向量视觉特征描述子。

[0105] 可选地,图像的视觉特征描述子可以为一个或多个本申请实施例对此不作限定。

[0106] 3、视觉词袋模型

[0107] 视觉词袋模型包括多个视觉词,该多个视觉词中的每个视觉词是对从多个图像中提取的视觉特征描述子进行聚类后得到的一个聚类中心。

[0108] 4、图像的视觉词

[0109] 图像的视觉词是指通过将该图像的视觉特征描述子和视觉词袋模型中的视觉词进行匹配映射,得到的该视觉词袋模型中与该视觉特征描述子距离最近的视觉词。

[0110] 可选地,图像的视觉词可以为一个或多个,本申请实施例对此不作限定。

[0111] 5、图像的图像类别

[0112] 根据不同的分类方法对多个图像进行分类,能够得到每个图像的图像类别。

[0113] 作为一个可选实施例,如果按照场景对图像进行分类,则图像的图像类别可以包括森林场景、郊区场景、室内场景等。

[0114] 作为另一个可选实施例,如果按照天气对图像进行分类,则图像的图像类别可以包括晴天、雨天、雪天等。

[0115] 6、图像类别对应的视觉停用词

[0116] 由于同一个图像中出现的不同视觉词可能对辨识该图像具有不同的作用,不同图像中出现的相同视觉词可能对辨识这两个图像具有相同的作用,图像类别对应的视觉停用词是指对于辨识某种图像类别的图像无显著作用、或影响图像识别的视觉词,即与该图像类别的图像无关的视觉词。

[0117] 应理解,本申请实施例中所述的与某种图像类别无关的视觉停用词,是指与该种图像类别的图像的相关性低于预设阈值的视觉词。

[0118] 可选地,图像类别对应的视觉停用词可以包括一个或多个视觉词,本申请实施例对此不作限定。

[0119] 例如,在森林场景和郊区场景下,几乎每个图像都包含大量的树木,从图像中的树木提取的特征点对辨识该图像是森林场景还是郊区场景的辨识度较低,因此,树木可以为森林类或郊区类的视觉停用词。

[0120] 又例如,在雨天,图像中会留下雨水降落的痕迹,从图像中的雨水提取的特征点也会对该图像的单词表示造成污染,因此,雨水可以为雨天的视觉停用词。

[0121] 7、图像的目标视觉词

[0122] 图像的目标视觉词包括从该图像的多个视觉词中除去该图像的图像类别对应的视觉停用词后的视觉词。

[0123] 可选地,图像的目标视觉词可以包括一个或多个视觉词,本申请实施例对此不作限定。

[0124] 8、正样本图像集合

[0125] 正样本图像集合中包括人工标注的可以被认为是相似度高或相同的图像。

[0126] 例如,在不同场景拍摄同一物体的两个图像,例如,雨天的学校和雪天的学校。

[0127] 例如,在不同时刻拍摄同一场景的两个图像,例如回环检测中同一场景的当前位姿和历史位姿。

[0128] 本申请实施例可应用的场景包括即时定位与地图构建(simultaneous localization and mapping,SLAM)中的回环检测(loop closure)、电子商务中的商品图像检索等。

[0129] 回环检测通过保存历史出现过的场景,利用当前图像进行检索识别回环,构建当前位姿和历史位姿的一个约束,通过优化减小整体误差,以得到全局一致的地图。

[0130] 电子商务中在不清楚商品名称时,用户提交商品的图像,系统根据该商品的图像进行检索,并返回相似度较高的图像作为检索结果。

[0131] 图1示出了本申请实施例提供的图像检索方法100的示意性流程图。该方法可以由图像检索装置执行。

[0132] 应理解,该图像检索装置可以为具有计算和存储功能的第一设备,该第一设备例如可以为计算机,或者该图像检索装置可以为第一设备中的功能模块,本申请实施例对此不作限定。

[0133] S110,获取待检索图像的多个视觉词和该待检索图像的图像类别信息,该待检索图像的多个视觉词是通过将该待检索图像的多个视觉特征描述子与视觉词袋模型中的视觉词进行匹配映射得到的,该待检索图像的图像类别信息用于指示该待检索图像的图像类别。

[0134] S120,根据该待检索图像的图像类别信息和停用词词库,确定该待检索图像的图像类别对应的视觉停用词,该待检索图像的图像类别对应的视觉停用词包括与该待检索图像的图像类别无关的视觉词,该停用词词库包括该待检索图像的图像类别和该待检索图像的图像类别对应的视觉停用词之间的映射关系。

[0135] S130,从该待检索图像的多个视觉词中除去该待检索图像的图像类别对应的视觉停用词,得到该待检索图像的目标视觉词。

[0136] S140,根据该待检索图像的目标视觉词和检索图像库,确定检索结果,该检索图像库中包括多个检索图像。

[0137] 可选地,S110中,该图像检索装置可以通过多种方式获取该待检索图像的多个视觉词,本申请实施例对此不作限定。

[0138] 作为一个可选实施例,该图像检索装置可以获取待检索图像,提取该待检索图像的多个视觉特征描述子,该多个视觉特征描述子用于描述该待检索图像的多个视觉特征点,该多个视觉特征描述子与该多个视觉特征点一一对应,获取视觉词袋模型,将视觉词袋模型中与该多个视觉特征描述子中的每个视觉特征描述子的距离最近的多个视觉词,确定为该待检索图像的多个视觉词。

[0139] 可选地,该视觉词袋模型可以为现有训练好的视觉词袋模型,或者可以为该图像检索装置通过对训练图集中的训练图像的视觉特征描述子进行聚类得到的,本申请实施例对此不作限定。

[0140] 可选地,该图像检索装置可以通过多种方式获取该待检索图像,例如通过摄像头拍摄、本地磁盘读取、网络下载,或其他方式,本申请实施例对此不作限定。

[0141] 可选地,该图像检索装置获取到的该待检索图像可以为经过去畸变、去噪、或其他预处理操作后的图像,本申请实施例对此不作限定。

[0142] 可选地,S110中,该图像检索装置可以通过多种方式获取该待检索图像的图像类别信息,本申请实施例对此不作限定。

[0143] 作为一个可选实施例,该图像检索装置可以根据该待检索图像和图像分类模型,确定该待检索图像的图像类别信息,该图像分类模型包括该待检索图像与该待检索图像的图像类别的映射关系。

[0144] 作为另一个可选实施例,该图像检索装置可以根据该待检索图像和预设的分类算法,确定该待检索图像的图像类别信息。

[0145] 作为又一个可选实施例,该图像检索装置可以获取人工标注的该待检索图像的图像类别信息。

[0146] 可选地,该待检索图像的图像类别信息可以为一个或多个比特,即通过该1个或多个比特指示该待检索图像的图像类别,本申请实施例对此不作下限定。

[0147] 作为一个可选实施例,该待检索图像的图像类别信息可以为2个比特,例如,该2比特为“00”时指示该待检索图像为第一类图像,该2比特为“01”时指示该待检索图像为第二类图像,该2比特为“10”时指示该待检索图像为第三类图像,该2比特为“11”时指示该待检索图像为第四类图像。

[0148] 可选地,在S120之前,该图像检索装置可以获取该停用词词库。

[0149] 可选地,该停用词词库可以包括多种图像类别中每种图像类别的标识和与该每种图像类别的标识对应的视觉停用词的映射关系。

[0150] 可选地,该图像检索装置可以自己生成该停用词词库,或者可以从图像库的生成装置获取该停用词词库,本申请实施例对此不限定。

[0151] 可选地,该图像检索装置可以获取训练图像库中每个训练图像的多个视觉词、该每个训练图像的图像类别信息和正样本图像集合信息,该每个训练图像的多个视觉词是通过将该每个训练图像的多个视觉特征描述子与该视觉词袋模型中的视觉词进行匹配映射得到的,该每个训练图像的图像类别信息用于指示该每个训练图像的图像类别,该正样本图像集合信息用于指示至少一个正样本图像集合,该正样本图像集合包括人工标注的该训练图像库中的多个相似的训练图像;根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,生成停用词词库,该停用词词库包括该每个训练图像的图像类别和与该每个图像的图像类别对应的视觉停用词之间的映射关系,该每个训练图像的图像类别对应的视觉停用词包括与该每个训练图像的图像类别无关的视觉词。

[0152] 具体地,该图像检索装置根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,生成该停用词词库,可以为根据该每个训练图像对应的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,确定该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,该训练图像库的多种图像类别包括该每个训练图像的图像类别,该训练图像库的多个视觉词包括该每个训练图像的多个视觉词;根据该训练图像库的多种图像类别与该训练图像库的多个视觉词之间

的相关性,生成该停用词词库。

[0153] 作为一个可选实施例,该图像检索装置可以根据第一事件发生的概率、第二事件发生的概率和该第一事件与该第二事件同时发生的概率,确定第一图像类别与第一视觉词之间的相关性,该训练图像库的多种图像类别包括该第一图像类别,该训练图像库的多个视觉词包括该第一视觉词,该第一事件表示该训练图像库中的第一训练图像的多个视觉词和该训练图像库中的第二训练图像的多个视觉词均包括该第一视觉词,该第一训练图像的图像类别为该第一图像类别,该第二事件表示该第一训练图像和该第二训练图像属于相同的正样本图像集合。

[0154] 例如,假设该训练图像集中共有P个训练图像,M种图像类别,L个视觉词,该P个训练图像中第一类别的训练图像共有N个,则第一视觉词和该第一图像类别的训练图像的相关性可以通过公式(1)至公式(6)确定:

$$[0155] \quad p(x) = \frac{\sum_{i=1}^N \text{count}(x_i)}{P * N} \quad (1)$$

$$[0156] \quad p(y) = \frac{\sum_{i=1}^N \text{count}(y_i)}{P * N} \quad (2)$$

$$[0157] \quad p(x, y) = \frac{\sum_{i=1}^N \text{count}(x_i, y_i)}{P * N} \quad (3)$$

$$[0158] \quad PMI(x, y) = \ln \frac{p(x, y)}{p(x)p(y)} \quad (4)$$

$$[0159] \quad H(y) = \ln \frac{P * N}{\text{count}(y)} \quad (5)$$

$$[0160] \quad RATE_{PMI}(x, y) = \frac{PMI(x, y)}{H(y)} \quad (6)$$

[0161] 其中,x表示第一事件,该第一事件为该N个训练图像中的第一训练图像的多个视觉词和该P个训练图像中除该第一训练图像外的第二训练图像的多个视觉词均包括该L个视觉词中的第一视觉词, x_i 表示第一类别的训练图像中第i个图像发生的第一事件;y表示第二事件,该第二事件为该第一训练图像和该第二训练图像属于同一个正样本图像集合, y_i 表示第一类别的训练图像中第i个图像发生的第二事件, $\text{count}(x_i)$ 表示第一类别的训练图像中第i个图像发生第一事件的次数, $\text{count}(y_i)$ 表示第一类别的训练图像中第i个图像发生第二事件的次数, $\text{count}(y)$ 表示第二事件发生的次数, $\text{count}(x_i, y_i)$ 表示第一类别的训练图像中第i个图像中第一事件和第二事件同时发生的次数, $p(x)$ 表示第一事件发生的概率, $p(y)$ 表示第二事件发生的概率, $p(x, y)$ 表示第一事件与第二事件同时发生的概率, $PMI(x, y)$ 表示第一事件和第二事件的点互信息量, $H(y)$ 表示第二事件的信息熵, $RATE_{PMI}(x, y)$ 表示第一事件和第二事件的点互信息率即该第一视觉词和该第一图像类别的相关性,其中,P、L、M、N均为大于1的正整数。

[0162] 可选地,该图像检索装置可以通过多种方式获取正样本图像集合信息,本申请实

施例对此不作限定。

[0163] 作为一个可选实施例,该图像检索装置可以获取该每个训练图像中携带的一个或多个比特,根据该每个训练图像的一个或多个比特,获取该正样本图像集合信息。例如,若该多个训练图像中的第一训练图像和第二训练图像携带相同的比特,则确定该第一训练图像和该第二训练图像属于同一个正样本图像集合。

[0164] 作为另一个可选实施例,该图像检索装置可以获取第一信息,该第一信息包括多个正样本图像集合中每个正样本图像集合的标识和该每个正样本图像集合包括的训练图像的标识之间的映射关系,该图像检索装置可以根据该第一信息,获取该正样本图像集合信息。

[0165] 可选地,该图像检索装置可以通过多种方式从该训练图像库的多个视觉词中确定该每种图像类别对应的视觉停用词,本申请实施例对此不作限定。

[0166] 作为一个可选实施例,该图像检索装置可以将该训练图像库的多个视觉词中与每种图像类别的相关性最小的至少一个视觉词,作为该每种图像类别对应的视觉停用词。

[0167] 作为另一个可选实施例,该图像检索装置可以将该训练图像库的多个视觉词中与该每种图像类别的相关性小于第一预设阈值的至少一个视觉词,作为该每种图像类别对应的视觉停用词。

[0168] 可选地,每种图像类别对应的视觉停用词可以为一个或多个视觉词,本申请实施例对此不作限定。

[0169] 可选地,在S140之前,该图像检索装置可以获取该检索图像库。

[0170] 可选地,该检索图像库中包括多个检索图像和该多个检索图像中每个检索图像对应的目标视觉词,该每个检索图像对应的目标视觉词是从该每个检索图像对应的多个视觉词中除去该每个检索图像的图像类别对应的视觉停用词后得到的。

[0171] 可选地,该图像检索装置可以自己生成该检索图像库,或者可以从图像库的生成装置获取该检索图像库,本申请实施例对此不作限定。

[0172] 可选地,该检索图像库可以为根据该多个训练图像训练得到的,或者该检索图像库可以为根据该图像检索装置在本次检索之前检索的历史待检索图像训练得到的,或者可以为根据其他图像训练得到的,本申请实施例对此不作限定。

[0173] 作为一个可选实施例,该图像检索装置可以根据该多个训练图像中每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该停用词词库,生成该检索图像库。也就是说,该检索图像库是根据该多个训练图像训练得到的。

[0174] 具体而言,该图像检索装置可以根据该每个训练图像的图像类别信息和该停用词词库,确定该每个训练图像的图像类别对应的视觉停用词,从该每个训练图像的多个视觉词中除去该每个训练图像的图像类别对应的视觉停用词,得到该每个训练图像的目标视觉词,并将该每个训练图像的目标视觉词添加至该检索图像库。

[0175] 可选地,该图像检索装置还可以使用本申请实施例中的停用词词库获得该每个训练图像的目标视觉词,或者可以通过其他方式获得该每个训练图像的目标视觉词,本申请实施例对此不作限定。

[0176] 例如,电子商务中,该检索图像库包括用户提供的所有商品图像的目标视觉词,该检索图像为用户想要购买的商品图像。

[0177] 作为另一个可选实施例,该图像检索装置可以将S140之前检索的历史待检索图像的目标视觉词添加至该检索图像库,以生成该检索图像库。

[0178] 例如,回环检测中,检索图像库包括所有历史位姿图像,该待检索图像为当前位姿图像。

[0179] 本申请实施例提供的图像检索方法,在回环检测场景中,能够通过保存历史出现过的场景,利用当前图像进行检索识别回环,构建当前位姿和历史位姿的一个约束,通过优化减小整体误差,以得到全局一致的地图。在电子商务场景中,在不清楚商品名称时,用户提交商品的图像,系统根据该商品的图像进行检索,并返回相似度较高的图像作为检索结果。

[0180] 可选地,该图像检索装置可以根据多种方式从该检索图像库中选出与该待检索图像最相似的至少一个检索图像作为检索结果输出,本申请实施例对此不作限定。

[0181] 作为一个可选实施例,该图像检索装置可以计算该待检索图像的多个视觉词与该检索图像库中的检索图像的多个视觉词之间的相似度,将与该待检索图像的相似度最高的至少一个检索图像确定为检索结果。

[0182] 作为另一个可选实施例,该图像检索装置可以计算该待检索图像的多个视觉词与该检索图像库中的检索图像的多个视觉词之间的相似度,将相似度大于第二预设阈值的至少一个检索图像确定为检索结果。

[0183] 本申请实施例提供的图像检索方法,该待检索图像的目标视觉词是通过从该待检索图像的多个视觉词中除去该待检索图像的图像类别对应的视觉停用词后得到的,即从该待检索图像的多个视觉词中除去对于辨识该检索图像无显著作用、或影响图像识别的视觉词,也就是说,该待检索图像的目标视觉词对于识别该待检索图像的作用比较显著。因此,通过该待检索图像的目标视觉词与检索图像库进行检索,有利于提高图像检索的效率和精确度。

[0184] 此外,该检索图像库中存储的检索图像的目标视觉词是通过从检索图像的多个视觉词中除去了该检索图像的图像类别对应的视觉停用词后得到的,有利于减少该检索图像库内存的占用率。

[0185] 图2是本申请实施例提供的图像库的生成方法200的示意性流程图,该方法200可以由图像库的生成装置执行,本申请实施例对此不作限定。

[0186] 应理解,该图像库的生成装置可以为具有计算和存储功能的第二设备,该第二设备例如可以为计算机,或者该图像库生成装置可以为该第二设备中的功能模块,本申请实施例对此不作限定。

[0187] 可选地,该第二设备与图1中所述的第一设备可以为相同的设备或不同的设备,本申请实施例对此不作限定。

[0188] 可选地,该第二设备与该第二设备相同时,该图像库的生成装置与图1中所述的图像检索装置为同一个设备中的不同功能模块,或者该图像库的生成装置为该图像检索装置中的功能模块。

[0189] S210,获取训练图像库中每个训练图像的多个视觉词、该每个训练图像的图像类别信息和正样本图像集合信息,该每个训练图像的多个视觉词是通过将该每个训练图像的多个视觉特征描述子与该视觉词袋模型中的视觉词进行匹配映射得到的,该每个训练图像

的图像类别信息用于指示该每个训练图像的图像类别,该正样本图像集合信息用于指示至少一个正样本图像集合,该正样本图像集合包括人工标注的该训练图像库中的多个相似的训练图像。

[0190] S220,根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,生成停用词词库,该停用词词库包括该每个训练图像的图像类别和与该每个图像的图像类别对应的视觉停用词之间的映射关系,该每个训练图像的图像类别对应的视觉停用词包括与该每个训练图像的图像类别无关的视觉词。

[0191] 本申请实施例提供的数据库的生成方法,通过获取训练图像库中每个训练图像的多个视觉词、该每个训练图像的图像类别信息和正样本图像集合信息,并根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,生成停用词词库,该停用词词库用于获得待检索图像的目标视觉词,有利于提高图像检索的效率和精确度。

[0192] 可选地,S210中,该图像库的生成装置可以通过多种方式获取该训练图像的多个视觉词,本申请实施例对此不作限定。

[0193] 作为一个可选实施例,该图像训练装置可以获取训练图像,提取该训练图像的多个视觉特征描述子,该多个视觉特征描述子用于描述该训练图像的多个视觉特征点,该多个视觉特征描述子与该多个视觉特征点一一对应,获取视觉词袋模型,将视觉词袋模型中与该多个视觉特征描述子中的每个视觉特征描述子的距离最近的多个视觉词,确定为该训练图像的多个视觉词。

[0194] 可选地,该视觉词袋模型也可以为现有训练好的视觉词袋模型,或者可以为该图像库的生成装置自己通过对该多个图像对应的多个视觉特征描述子进行聚类得到的,本申请实施例对此不作限定。

[0195] 可选地,该图像库的生成装置可以通过多种方式获取训练图像,例如通过摄像头拍摄、本地磁盘读取、网络下载,或其他方式,本申请实施例对此不作限定。

[0196] 可选地,该图像库的生成装置获取到的该多个图像可以为经过去畸变、去噪、或其他预处理操作后的图像,本申请实施例对此不作限定。

[0197] 可选地,S210中,该图像库的生成装置可以通过多种方式获取训练图像的图像类别信息,本申请实施例对此不作限定。

[0198] 作为一个可选实施例,该图像库的生成装置可以根据该训练图像和图像分类模型,确定该训练图像的图像类别信息,该图像分类模型包括该训练图像与该训练图像的图像类别的映射关系。

[0199] 作为另一个可选实施例,该图像库的生成装置可以根据该训练图像和预设的分类算法,确定该训练图像的图像类别信息。

[0200] 作为又一个可选实施例,该图像库的生成装置可以获取人工标注的该训练图像的图像类别信息。

[0201] 可选地,该训练图像的图像类别信息可以为一个或多个比特,即通过该1个或多个比特指示该训练图像的图像类别,本申请实施例对此不作下限定。

[0202] 作为一个可选实施例,该训练图像的图像类别信息可以为2个比特,例如,该2比特为“00”时指示该训练图像为第一类图像,该2比特为“01”时指示该训练图像为第二类图像,

该2比特为“10”时指示该训练图像为第三类图像,该2比特为“11”时指示该训练图像为第四类图像。

[0203] 可选地,S220中,该图像库的生成装置根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,生成该停用词词库,可以为根据该每个训练图像对应的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,确定该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,该训练图像库的多种图像类别包括该每个训练图像的图像类别,该训练图像库的多个视觉词包括该每个训练图像的多个视觉词;根据该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,生成该停用词词库。

[0204] 作为一个可选实施例,该图像库的生成装置可以根据第一事件发生的概率、第二事件发生的概率和该第一事件与该第二事件同时发生的概率,确定第一图像类别与第一视觉词之间的相关性,该训练图像库的多种图像类别包括该第一图像类别,该训练图像库的多个视觉词包括该第一视觉词,该第一事件表示该训练图像库中的第一训练图像的多个视觉词和该训练图像库中的第二训练图像的多个视觉词均包括该第一视觉词,该第一训练图像的图像类别为该第一图像类别,该第二事件表示该第一训练图像和该第二训练图像属于相同的正样本图像集合。

[0205] 例如,假设该训练图像集中共有P个训练图像,M种图像类别,L个视觉词,该P个训练图像中第一类别的训练图像共有N个,则第一视觉词和该第一图像类别的训练图像的相关性可以通过上述公式(1)至公式(6)确定。

[0206] 可选地,该图像库的生成装置可以通过多种方式获取正样本图像集合信息,本申请实施例对此不作限定。

[0207] 作为一个可选实施例,该图像库的生成装置可以获取该每个训练图像中携带的一个或多个比特,根据该每个训练图像的一个或多个比特,获取该正样本图像集合信息。例如,若该多个训练图像中的第一训练图像和第二训练图像携带相同的比特,则确定该第一训练图像和该第二训练图像属于同一个正样本图像集合。

[0208] 作为另一个可选实施例,该图像库的生成装置可以获取第一信息,该第一信息包括多个正样本图像集合中每个正样本图像集合的标识和该每个正样本图像集合包括的训练图像的标识之间的映射关系,该图像库的生成装置可以根据该第一信息,获取该正样本图像集合信息。

[0209] 可选地,该图像库的生成装置可以通过多种方式从该训练图像库的多个视觉词中确定该每种图像类别对应的视觉停用词,本申请实施例对此不作限定。

[0210] 作为一个可选实施例,该图像库的生成装置可以将该训练图像库的多个视觉词中与每种图像类别的相关性最小的至少一个视觉词,作为该每种图像类别对应的视觉停用词。

[0211] 作为另一个可选实施例,该图像库的生成装置可以将该训练图像库的多个视觉词中与该每种图像类别的相关性小于第一预设阈值的至少一个视觉词,作为该每种图像类别对应的视觉停用词。

[0212] 可选地,每种图像类别对应的视觉停用词可以为一个或多个视觉词,本申请实施例对此不作限定。

[0213] 可选地,在S220之后,该图像库的生成装置可以根据该每个训练图像的图像类别信息和该停用词词库,确定该每个训练图像的图像类别对应的视觉停用词,从该每个训练图像的多个视觉词中除去该每个训练图像的图像类别对应的视觉停用词,得到该每个训练图像的目标视觉词,并将该每个训练图像的目标视觉词添加至该检索图像库。

[0214] 本申请实施例提供的图像检索方法,该检索图像库中存储的检索图像的目标视觉词是通过从检索图像的多个视觉词中除去了该检索图像的图像类别对应的视觉停用词后得到的,有利于减少该检索图像库内存的占用率。

[0215] 上面结合图1和图2详细描述了本申请实施例提供的图像检索方法和图像处理方法,下面将结合图3至图6描述本申请实施例提供的图像检索装置和图像处理装置。

[0216] 图3是本申请实施例提供的图像检索装置300的示意性框图。该装置300包括:

[0217] 获取单元310,用于获取待检索图像的多个视觉词和该待检索图像的图像类别信息,该待检索图像的多个视觉词是通过将该待检索图像的多个视觉特征描述子与视觉词袋模型中的视觉词进行匹配映射得到的,该待检索图像的图像类别信息用于指示该待检索图像的图像类别;

[0218] 处理单元320,用于根据该获取单元310获取的该待检索图像的图像类别信息和停用词词库,确定该待检索图像的图像类别对应的视觉停用词,该待检索图像的图像类别对应的视觉停用词包括与该待检索图像的图像类别无关的视觉词,该停用词词库包括该待检索图像的图像类别和该待检索图像的图像类别对应的视觉停用词之间的映射关系;从该获取单元310获取的该待检索图像的多个视觉词中除去该待检索图像的图像类别对应的视觉停用词,得到该待检索图像的目标视觉词;

[0219] 检索单元330,用于根据该处理单元320得到的该待检索图像的目标视觉词和检索图像库,确定检索结果,该检索图像库中包括多个检索图像。

[0220] 可选地,该检索图像库中包括该多个检索图像和该多个检索图像中每个检索图像对应的目标视觉词的映射关系,该每个检索图像对应的目标视觉词是从该每个检索图像对应的多个视觉词中除去该每个检索图像的图像类别对应的视觉停用词后得到的。

[0221] 可选地,该装置还包括生成单元,该获取单元还用于在该根据该待检索图像的图像类别信息和停用词词库,确定该待检索图像的图像类别对应的视觉停用词之前,获取训练图像库中每个训练图像的多个视觉词、该每个训练图像的图像类别信息和正样本图像集合信息,该每个训练图像的多个视觉词是通过将该每个训练图像的多个视觉特征描述子与该视觉词袋模型中的视觉词进行匹配映射得到的,该每个训练图像的图像类别信息用于指示该每个训练图像的图像类别,该正样本图像集合信息用于指示至少一个正样本图像集合,该正样本图像集合包括人工标注的该训练图像库中的多个相似的训练图像;该生成单元用于根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和正样本图像集合信息,生成该停用词词库。

[0222] 可选地,该生成单元具体用于:根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,确定该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,该训练图像库的多种图像类别包括该每个训练图像的图像类别,该训练图像库的多个视觉词包括该每个训练图像的多个视觉词;根据该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,生成该停用词词库。

[0223] 可选地,该生成单元具体用于:根据第一事件发生的概率、第二事件发生的概率和该第一事件与该第二事件同时发生的概率,确定第一图像类别与第一视觉词之间的相关性,该训练图像库的多种图像类别包括该第一图像类别,该训练图像库的多个视觉词包括该第一视觉词,该第一事件表示该训练图像库中的第一训练图像的多个视觉词和该训练图像库中的第二训练图像的多个视觉词均包括该第一视觉词,该第一训练图像的图像类别为该第一图像类别,该第二事件表示该第一训练图像和该第二训练图像属于相同的正样本图像集合。

[0224] 可选地,该检索单元具体用于:确定该待检索图像的目标视觉词与该检索图像库中的检索图像的目标视觉词之间的相似度;将与该待检索图像的目标视觉词的相似度大于第一预设值的至少一个检索图像,确定为该检索结果。

[0225] 应理解,这里的图像检索装置300以功能单元的形式体现。这里的术语“单元”可以指应用特有集成电路(application specific integrated circuit,ASIC)、电子电路、用于执行一个或多个软件或固件程序的处理器(例如共享处理器、专有处理器或组处理器等)和存储器、合并逻辑电路和/或其它支持所描述的功能的合适组件。在一个可选例子中,本领域技术人员可以理解,图像检索装置300可以具体为上述方法100和方法200实施例中的图像检索装置,图像检索装置300可以用于执行上述方法100和方法200实施例中与图像检索装置对应的各个流程和/或步骤,为避免重复,在此不再赘述。

[0226] 图4示出了本申请实施例提供的图像库的生成装置400的示意性框图,该装置400包括:

[0227] 获取单元410,用于获取训练图像库中每个训练图像的多个视觉词、该每个训练图像的图像类别信息和正样本图像集合信息,该每个训练图像的多个视觉词是通过将该每个训练图像的多个视觉特征描述子与该视觉词袋模型中的视觉词进行匹配映射得到的,该每个训练图像的图像类别信息用于指示该每个训练图像的图像类别,该正样本图像集合信息用于指示至少一个正样本图像集合,该正样本图像集合包括人工标注的该训练图像库中的多个相似的训练图像;

[0228] 生成单元420,用于根据该获取单元410获取的该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,生成停用词词库,该停用词词库包括该每个训练图像的图像类别和与该每个训练图像的图像类别对应的视觉停用词之间的映射关系,该每个训练图像的图像类别对应的视觉停用词包括与该每个训练图像的图像类别无关的视觉词。

[0229] 可选地,该生成单元具体用于:根据该每个训练图像对应的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,确定该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,该训练图像库的多种图像类别包括该每个训练图像的图像类别,该训练图像库的多个视觉词包括该每个训练图像的多个视觉词;根据该训练图像库的多种图像类别与该训练图像库的多个视觉词之间的相关性,生成该停用词词库。

[0230] 可选地,该生成单元具体用于:根据第一事件发生的概率、第二事件发生的概率和该第一事件与该第二事件同时发生的概率,确定第一图像类别与第一视觉词之间的相关性,该训练图像库的多种图像类别包括该第一图像类别,该训练图像库的多个视觉词包括

该第一视觉词,该第一事件表示该训练图像库中的第一训练图像的多个视觉词和该训练图像库中的第二训练图像的多个视觉词均包括该第一视觉词,该第一训练图像的图像类别为该第一图像类别,该第二事件表示该第一训练图像和该第二训练图像属于相同的正样本图像集合

[0231] 可选地,该生成单元还用于在该根据该每个训练图像的多个视觉词、该每个训练图像的图像类别信息和该正样本图像集合信息,生成停用词词库之后,根据该每个训练图像的图像类别信息和该停用词词库,确定该每个训练图像的图像类别对应的视觉停用词,从该每个训练图像的多个视觉词中除去该每个训练图像的图像类别对应的视觉停用词,得到该每个训练图像的目标视觉词,并将该每个训练图像的目标视觉词添加至该检索图像库。

[0232] 可选地,该获取单元具体用于获取该每个训练图像,提取该每个训练图像的多个视觉特征描述子,该多个视觉特征描述子用于描述该每个训练图像的多个视觉特征点,该多个视觉特征描述子与该多个视觉特征点一一对应,获取视觉词袋模型,将视觉词袋模型中与该多个视觉特征描述子中的每个视觉特征描述子的距离最近的多个视觉词,确定为该每个训练图像的多个视觉词。

[0233] 应理解,这里的图像库的生成装置400以功能单元的形式体现。这里的术语“单元”可以指ASIC、电子电路、用于执行一个或多个软件或固件程序的处理器(例如共享处理器、专有处理器或组处理器等)和存储器、合并逻辑电路和/或其它支持所描述的功能的合适组件。在一个可选例子中,本领域技术人员可以理解,图像库的生成装置400可以具体为上述方法100和方法100实施例中的图像库的生成装置,图像库的生成装置400可以用于执行上述方法100和方法200实施例中与图像库的生成装置对应的各个流程和/或步骤,为避免重复,在此不再赘述。

[0234] 图5示出了本申请实施例提供的图像检索装置500的示意性框图,该图像检索装置500可以为图1和图2中所述的图像检索装置,该图像检索装置可以采用如图5所示的硬件架构。该图像检索装置可以包括处理器510、通信接口520和存储器530,该处理器510、通信接口520和存储器530通过内部连接通路互相通信。图3中的处理单元320和检索单元330所实现的相关功能可以由处理器510来实现,获取单元310所实现的相关功能可以由处理器510控制通信接口520来实现。

[0235] 该处理器510可以包括是一个或多个处理器,例如包括一个或多个中央处理单元(central processing unit,CPU),在处理器是一个CPU的情况下,该CPU可以是单核CPU,也可以是多核CPU。

[0236] 该通信接口520用于发送和/或接收数据。该通信接口可以包括发送接口和接收接口,发送接口用于发送数据,接收接口用于接收数据。

[0237] 该存储器530包括但不限于是随机存取存储器(random access memory,RAM)、只读存储器(read-only memory,ROM)、可擦除可编程存储器(erasable programmable read only memory,EPR0M)、只读光盘(compact disc read-only memory,CD-ROM),该存储器530用于存储相关指令及数据。

[0238] 存储器530用于存储图像检索装置的程序代码和数据,可以为单独的器件或集成在处理器510中。

[0239] 具体地,所述处理器510用于控制通信接口与其它装置,例如图像库的生成装置进行数据传输。具体可参见方法实施例中的描述,在此不再赘述。

[0240] 可以理解的是,图5仅仅示出了图像检索装置的简化设计。在实际应用中,图像检索装置还可以分别包含必要的其他元件,包括但不限于任意数量的通信接口、处理器、控制器、存储器等,而所有可以实现本申请的图像检索装置都在本申请的保护范围之内。

[0241] 在一种可能的设计中,图像检索装置500可以被替换为芯片装置,例如可以为可用于图像检索装置中的芯片,用于实现图像检索装置中处理器510的相关功能。该芯片装置可以为实现相关功能的现场可编程门阵列,专用集成芯片,系统芯片,中央处理器,网络处理器,数字信号处理电路,微控制器,还可以采用可编程控制器或其他集成芯片。该芯片中,可选的可以包括一个或多个存储器,用于存储程序代码,当所述代码被执行时,使得处理器实现相应的功能。

[0242] 图6示出了本申请实施例提供的图像库的生成装置600的示意性框图,该图像库的生成装置600可以为图1和图2中所述的图像库的生成装置,该图像库的生成装置可以采用如图6所示的硬件架构。该图像库的生成装置可以包括处理器610、通信接口620和存储器630,该处理器610、通信接口620和存储器630通过内部连接通路互相通信。图4中的生成单元420所实现的相关功能可以由处理器610来实现,获取单元410所实现的相关功能可以由处理器610控制通信接口620来实现。

[0243] 该处理器610可以包括是一个或多个处理器,例如包括一个或多个中央处理单元(central processing unit,CPU),在处理器是一个CPU的情况下,该CPU可以是单核CPU,也可以是多核CPU。

[0244] 该通信接口620用于发送和/或接收数据。该通信接口可以包括发送接口和接收接口,发送接口用于发送数据,接收接口用于接收数据。

[0245] 该存储器630包括但不限于是随机存取存储器(random access memory,RAM)、只读存储器(read-only memory,ROM)、可擦除可编程存储器(erasable programmable read only memory,EPR0M)、只读光盘(compact disc read-only memory,CD-ROM),该存储器630用于存储相关指令及数据。

[0246] 存储器630用于存储图像库的生成装置的程序代码和数据,可以为单独的器件或集成在处理器610中。

[0247] 具体地,所述处理器610用于控制通信接口与其它装置,例如图像检索装置进行数据传输。具体可参见方法实施例中的描述,在此不再赘述。

[0248] 可以理解的是,图6仅仅示出了图像库的生成装置的简化设计。在实际应用中,图像库的生成装置还可以分别包含必要的其他元件,包括但不限于任意数量的通信接口、处理器、控制器、存储器等,而所有可以实现本申请的图像库的生成装置都在本申请的保护范围之内。

[0249] 在一种可能的设计中,图像库的生成装置600可以被替换为芯片装置,例如可以为可用于图像库的生成装置中的芯片,用于实现图像库的生成装置中处理器610的相关功能。该芯片装置可以为实现相关功能的现场可编程门阵列,专用集成芯片,系统芯片,中央处理器,网络处理器,数字信号处理电路,微控制器,还可以采用可编程控制器或其他集成芯片。该芯片中,可选的可以包括一个或多个存储器,用于存储程序代码,当所述代码被执行时,

使得处理器实现相应的功能。

[0250] 本领域普通技术人员可以意识到,结合本文中所公开的实施例描述的各示例的单元及算法步骤,能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本申请的范围。

[0251] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统、装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0252] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统、装置和方法,可以通过其它的方式实现。例如,以上所描述的装置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,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的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,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0253] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者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者全部单元来实现本实施例方案的目的。

[0254] 另外,在本申请各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。

[0255] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用,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,本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案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)执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0256] 以上所述,仅为本申请的具体实施方式,但本申请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申请揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本申请的保护范围之内。因此,本申请的保护范围应以所述权利要求的保护范围为准。

100

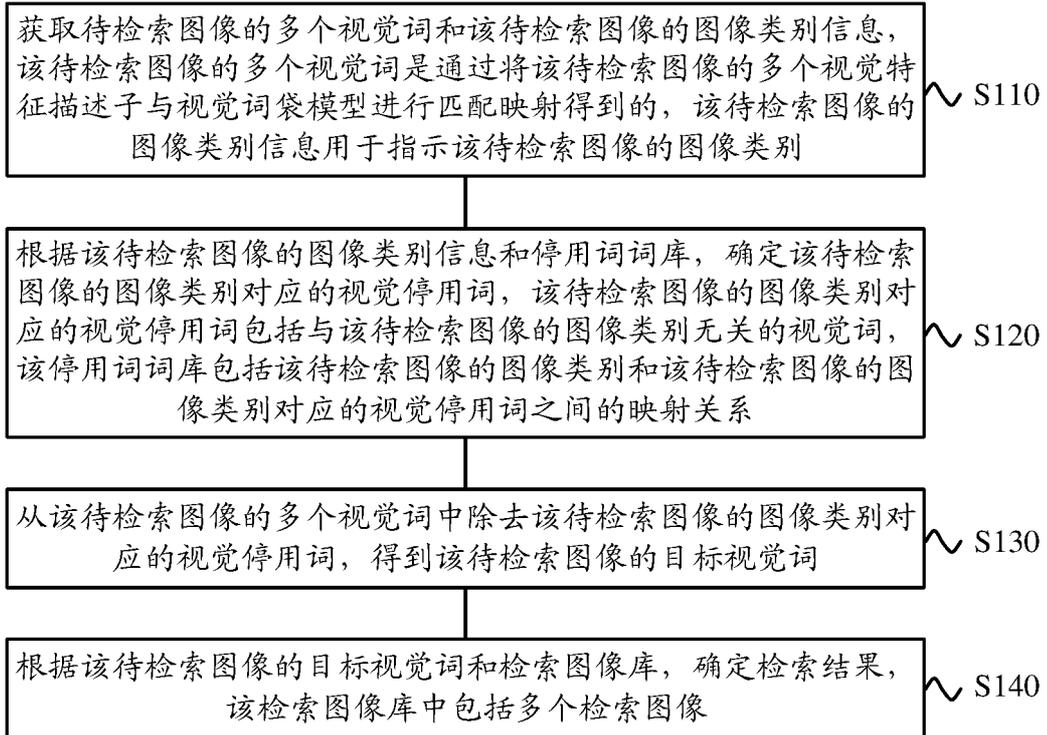


图1

200

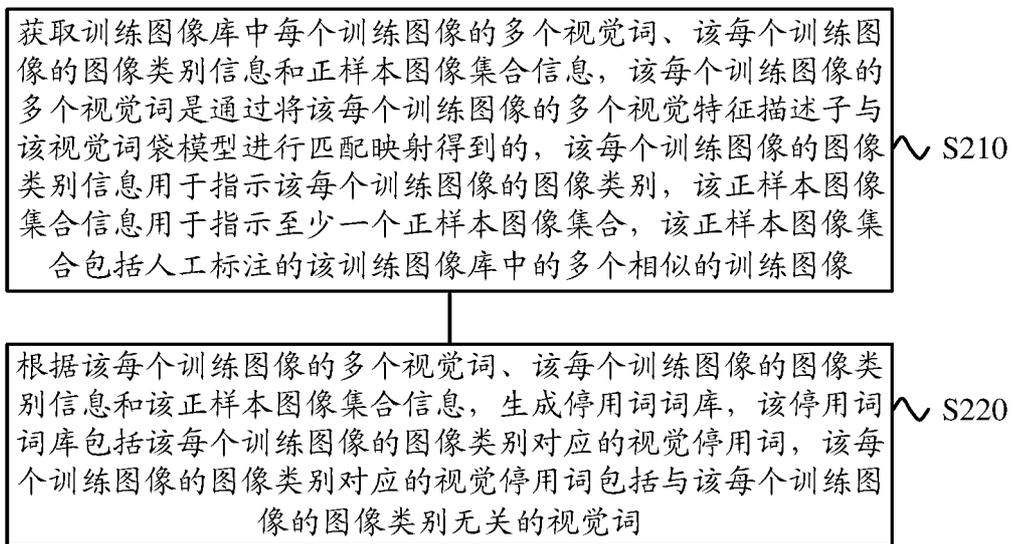


图2

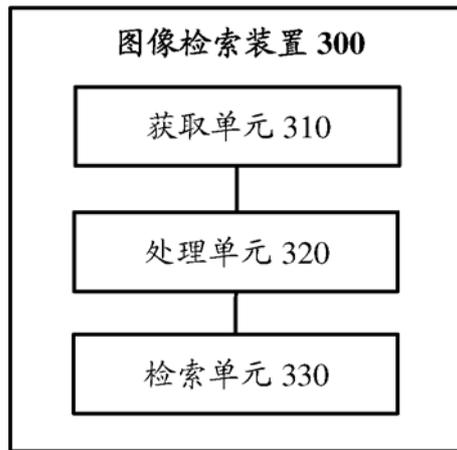


图3

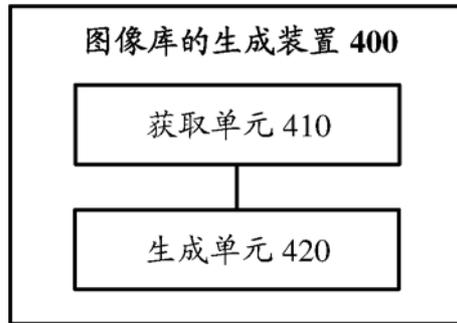


图4

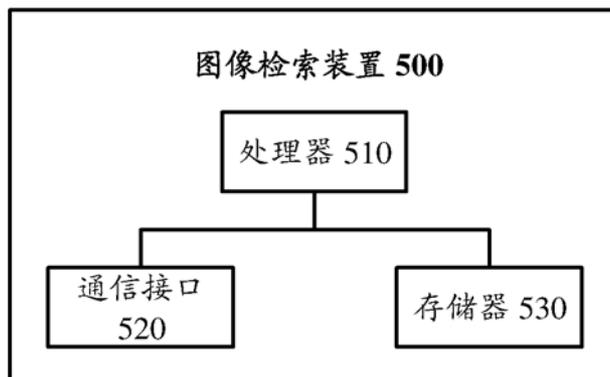


图5

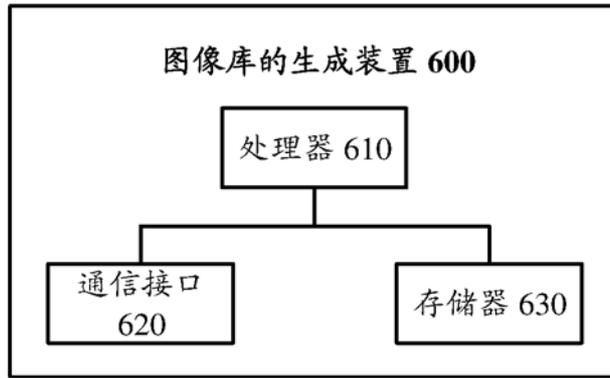


图6