



(12)发明专利

(10)授权公告号 CN 104346370 B

(45)授权公告日 2018.10.23

(21)申请号 201310328673.5

(22)申请日 2013.07.31

(65)同一申请的已公布的文献号
申请公布号 CN 104346370 A

(43)申请公布日 2015.02.11

(73)专利权人 阿里巴巴集团控股有限公司
地址 英属开曼群岛大开曼资本大厦一座四
层847号邮箱

(72)发明人 刘瑞涛 张洪明 茹新峰

(74)专利代理机构 北京润泽恒知识产权代理有
限公司 11319

代理人 苏培华

(51)Int.Cl.
G06F 17/30(2006.01)

(56)对比文件

US 2005162523 A1,2005.07.28,

US 2012117072 A1,2012.05.10,

CN 102385578 A,2012.03.21,

CN 103164539 A,2013.06.19,

审查员 李诺

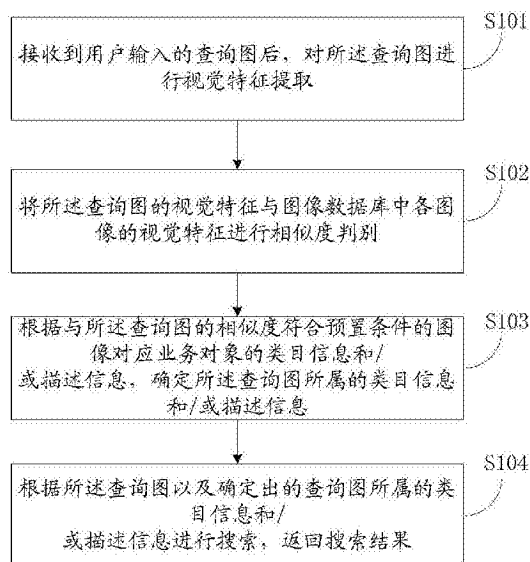
权利要求书2页 说明书10页 附图2页

(54)发明名称

图像搜索、获取图像文本信息的方法及装置

(57)摘要

本申请公开了图像搜索、获取图像文本信息的方法及装置,其中,所述图像搜索方法包括:接收到输入的查询图后,对所述查询图进行视觉特征提取;将所述查询图的视觉特征与图像数据库中各图像的视觉特征进行相似度判别;根据与所述查询图的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,确定所述查询图所属的类目信息和/或描述信息;根据所述查询图以及确定出的查询图所属的类目信息和/或描述信息进行搜索,返回搜索结果。通过本申请,能够自动获取图像的类目等信息,降低对用户交互的依赖。



1. 一种图像搜索方法,其特征在于,包括:

接收到用户输入的查询图后,提取所述查询图的主体内容区域,并基于所述主体内容区域进行视觉特征提取;所述主体内容区域为所述查询图中与业务对象相关的内容所在的区域;

将所述查询图的视觉特征与图像数据库中各图像的视觉特征进行相似度判别;所述图像数据库中保存有多个图像对应的视觉特征信息以及对应数据对象所属类目,所述多个图像为业务对象的主图;其中,如果所提取出的查询图的视觉特征为多种,则通过以下方式进行相似度判别:按照预先为各种特征设置的先后顺序,进行级联式的分层计算;其中,在每层的计算中,仅基于其中一种特征进行相似度判别,并将符合层内预置条件的图像集合输入到下一层,以便基于下一种特征进行相似度判别;

根据与所述查询图的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,确定所述查询图所属的类目信息和/或描述信息;

根据所述查询图以及确定出的查询图所属的类目信息和/或描述信息进行搜索,返回搜索结果。

2. 根据权利要求1所述的方法,其特征在于,如果所述查询图的主体内容是服饰类内容,则所述提取所述查询图的主体内容区域,包括:

利用人脸检测技术对所述查询图进行人脸检测,检测出人脸区域的位置及面积;

根据所述人脸区域的位置、面积以及预置的人脸区域与躯干区域的比例关系,确定躯干区域的位置及面积;

根据所述躯干区域的位置及面积,提取所述查询图的主体内容区域。

3. 根据权利要求1所述的方法,其特征在于,所述基于所述主体内容区域进行视觉特征提取包括:

对所述查询图进行全局特征和/或局部特征的提取。

4. 根据权利要求3所述的方法,其特征在于,所述全局特征包括全局边缘视觉特征和/或全局颜色分布特征,所述局部特征包括局部旋转不变特征。

5. 根据权利要求1至4任一项所述的方法,其特征在于,所述根据与所述查询图的相似度符合预置条件的图像对应业务对象的类目信息,确定所述查询图所属的类目信息,包括:

根据图像数据库中保存的各个图像的类目信息,确定所述相似度符合预置条件的各个图像分别对应的类目;

将出现次数最多的类目确定为所述查询图所属的类目信息。

6. 根据权利要求5所述的方法,其特征在于,所述根据与所述查询图的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,确定所述查询图所属的描述信息,包括:

提取所述相似度符合预置条件的图像中所述出现次数最多的类目对应的图像的描述信息,通过分析该描述信息获取所述查询图的描述信息。

7. 一种获取图像文本信息的方法,其特征在于,包括:

获取待确定类目信息的目标图像,提取所述目标图像的主体内容区域,并基于所述主体内容区域进行视觉特征提取;所述主体内容区域为所述目标图像中与业务对象相关的内容所在的区域;

将所述目标图像的视觉特征与图像数据库中各图像的视觉特征进行相似度判别;所述图像数据库中保存有多个图像对应的视觉特征信息以及对应数据对象所属类目,所述多个图像为业务对象的主图;其中,如果所提取出的目标图像的视觉特征为多种,则通过以下方式进行相似度判别:按照预先为各种特征设置的先后顺序,进行级联式的分层计算;其中,在每层的计算中,仅基于其中一种特征进行相似度判别,并将符合层内预置条件的图像集合输入到下一层,以便基于下一种特征进行相似度判别;

根据与所述目标图像的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,获取所述目标图像所属的类目信息和/或描述信息。

8. 根据权利要求7所述的方法,其特征在于,所述根据与所述目标图像的相似度符合预置条件的图像对应业务对象的类目信息,确定所述目标图像所属的类目信息,包括:

根据图像数据库中保存的各个图像的类目信息,确定所述相似度符合预置条件的各个图像分别对应的类目;

将出现次数最多的类目确定为所述目标图像所属的类目信息。

9. 一种图像搜索装置,其特征在于,包括:

特征提取单元,用于接收到用户输入的查询图后,提取所述查询图的主体内容区域,并基于所述主体内容区域进行视觉特征提取;所述主体内容区域为所述查询图中与业务对象相关的内容所在的区域;

相似度判别单元,用于将所述查询图的特征与图像数据库中各图像的视觉特征进行相似度判别;所述图像数据库中保存有多个图像对应的视觉特征信息以及对应数据对象所属类目,所述多个图像为业务对象的主图;其中,如果所提取出的查询图的视觉特征为多种,则通过以下方式进行相似度判别:按照预先为各种特征设置的先后顺序,进行级联式的分层计算;其中,在每层的计算中,仅基于其中一种特征进行相似度判别,并将符合层内预置条件的图像集合输入到下一层,以便基于下一种特征进行相似度判别;

预测单元,用于根据与所述查询图的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,确定所述查询图所属的类目信息和/或描述信息;

搜索结果返回单元,用于根据所述查询图以及确定出的查询图所属的类目信息和/或描述信息进行搜索,返回搜索结果。

10. 一种获取图像文本信息的装置,其特征在于,包括:

特征提取单元,用于获取待确定类目信息的目标图像,提取所述目标图像的主体内容区域,并基于所述主体内容区域进行视觉特征提取;所述主体内容区域为所述目标图像中与业务对象相关的内容所在的区域;

相似度判别单元,用于将所述目标图像的视觉特征与图像数据库中各图像的视觉特征进行相似度判别;所述图像数据库中保存有多个图像对应的视觉特征信息以及对应数据对象所属类目,所述多个图像为业务对象的主图;其中,如果所提取出的目标图像的视觉特征为多种,则通过以下方式进行相似度判别:按照预先为各种特征设置的先后顺序,进行级联式的分层计算;其中,在每层的计算中,仅基于其中一种特征进行相似度判别,并将符合层内预置条件的图像集合输入到下一层,以便基于下一种特征进行相似度判别;

预测单元,用于根据与所述目标图像的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,获取所述目标图像所属的类目信息和/或描述信息。

图像搜索、获取图像文本信息的方法及装置

技术领域

[0001] 本申请涉及图像处理技术领域,特别是涉及图像搜索、获取图像文本信息的方法及装置。

背景技术

[0002] 随着互联网上的图像数据信息日渐庞大,用户对网上图像搜索的需求也在不断增长,这使得各种基于Web的图像搜索引擎应运而生。图像搜索,就是通过搜索图像文本或者视觉特征,为用户提供互联网上相关图形图像资料检索服务的专业搜索引擎系统。

[0003] 图像搜索引擎按照搜索的图像范围可以分为两大类,一种是综合类图像搜索,另一种是垂直图像搜索。前者针对全网图像进行相似搜索,而后者则主要针对部分类目(如服装、鞋类产品)进行搜索。目前,电子商务交易平台等专业网站中的站内图像搜索引擎主要属于垂直图像搜索,根据用户上传的查询图进行搜索,返回与之相同或相似的业务对象的图像。

[0004] 最初,专业网站的站内图像搜索一般要以网站自身数据库中的某一幅图像作为查询图进行搜索。例如,某电子商务交易平台中,其图像数据库中保存有卖方用户上传的多幅业务对象的图像,同时还保存有各图像对应的业务对象所属的类目信息和相应的款式信息(包括颜色、形状等)等图像信息;用户可以选择其中一个业务对象的图像作为查询图,这样站内搜索引擎就可以根据查询图的类目信息和相应的颜色、形状等图像信息进行搜索,返回与查询图相同或高度相似的业务对象的图像。

[0005] 以上这种方式能够获得较为理想的搜索结果,但是,对于来源为网站图像数据库以外的图像(例如,在日常生活中用户使用手机拍摄的图像等)而言,由于无法事先获得查询图的相关描述信息,从而导致搜索结果图像相似度较低,召回率不高。当然,为了获得较优的搜索结果,可以要求用户在输入查询图的同时,还提供查询图中的主体内容所属的类目、款式信息等描述信息。但是,这会使得搜索结果过分依赖用户输入的描述信息,从用户角度而言,搜索过程变得繁琐,并且由于用户可能并不了解网站图像数据库对各类目的定义,或者自己对类目的理解存在偏差等原因,使得所输入的描述信息不一定准确,最终导致系统很可能返回错误的搜索结果。

发明内容

[0006] 本申请提供了图像搜索、获取图像文本信息的方法及装置,能够自动获取图像的文本信息,降低对用户交互的依赖。

[0007] 本申请提供了如下方案:

[0008] 一种图像搜索方法,包括:

[0009] 接收到输入的查询图后,对所述查询图进行视觉特征提取;

[0010] 将所述查询图的视觉特征与图像数据库中各图像的视觉特征进行相似度判别;

[0011] 根据与所述查询图的相似度符合预置条件的图像对应业务对象的类目信息和/或

描述信息,确定所述查询图所属的类目信息和/或描述信息;

[0012] 根据所述查询图以及确定出的查询图所属的类目信息和/或描述信息进行搜索,返回搜索结果。

[0013] 一种获取图像文本信息的方法,包括:

[0014] 获取待确定类目信息的目标图像,对所述目标图像进行视觉特征提取;

[0015] 将所述目标图像的视觉特征与图像数据库中各图像的视觉特征进行相似度判别;

[0016] 根据与所述目标图像的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,获取所述目标图像所属的类目信息和/或描述信息。

[0017] 一种图像搜索装置,包括:

[0018] 特征提取单元,用于接收到输入的查询图后,对所述查询图进行视觉特征提取;

[0019] 相似度判别单元,用于将所述查询图的特征与图像数据库中各图像的视觉特征进行相似度判别;

[0020] 预测单元,用于根据与所述查询图的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,确定所述查询图所属的类目信息和/或描述信息;

[0021] 搜索结果返回单元,用于根据所述查询图以及确定出的查询图所属的类目信息和/或描述信息进行搜索,返回搜索结果。

[0022] 一种获取图像文本信息的装置,包括:

[0023] 特征提取单元,用于获取待确定类目信息的目标图像,对所述目标图像进行视觉特征提取;

[0024] 相似度判别单元,用于将所述目标图像的视觉特征与图像数据库中各图像的视觉特征进行相似度判别;

[0025] 预测单元,用于根据与所述目标图像的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,获取所述目标图像所属的类目信息和/或描述信息。

[0026] 根据本申请提供的具体实施例,本申请公开了以下技术效果:

[0027] 通过本申请实施例,如果用户需要进行图像搜索,则可以仅提交查询图,不需要同时提交查询图所属的类目、描述信息等其他信息,并且提交的查询图可以是图像数据库之外的任意图像;在接收到查询图之后,系统可以首先根据查询图的视觉特征以及图像数据库中各图像的视觉特征预测出该查询图所属的类目和/或描述信息,进而可以结合查询图以及类目和/或描述信息,共同得出与查询图属于同种类目和/或在款式、色彩等视觉特征上高度相似的图像集合作为搜索结果。这样,可以使得在不需要用户提供类目或描述信息的情况下,就能为用户提供符合用户需求的搜索结果,并且,通过查询图特征的对比预测得到的类目、描述信息等更为客观、准确,可以降低对用户输入信息的依赖。

[0028] 另外,根据获取图像文本信息的方法及装置,可以根据用户提交的目标图像,自动根据目标图像的视觉特征以及图像数据库中图像的视觉特征,预测出目标图像所属的类目和/或描述信息。这样,在需要使用到目标图像的文本信息的应用中,就不再需要用户手动进行文本信息的输入,即使用户输入了文本信息,还可以根据预测出的信息对用户输入的信息进行验证,避免出现通过文本进行作弊等现象。

[0029] 当然,实施本申请的任一产品并不一定需要同时达到以上所述的所有优点。

附图说明

[0030] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例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申请的一些实施例,对于本领域普通技术人员来讲,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。

[0031] 图1是本申请实施例提供的图像搜索方法的流程图;

[0032] 图2是本申请实施例提供的获取图像文本信息的方法的流程图;

[0033] 图3是本申请实施例提供的图像搜索装置的示意图;

[0034] 图4是本申请实施例提供的获取图像文本信息的装置的示意图。

具体实施方式

[0035] 下面将结合本申请实施例中的附图,对本申请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申请一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员所获得的所有其他实施例,都属于本申请保护的范围。

[0036] 在本申请实施例中,在用户需要进行图像搜索时,可以使用网站图像数据库以外的图像作为查询图,例如,用户用手机拍摄的一张照片,或者其他网站、本地文件夹中的一幅图片等等,并且只需要输入查询图,而不用手动指定类目、描述信息(商品的属性、关键字、主颜色等款式信息)等。从搜索引擎角度而言,在接收到用户输入的查询图之后,可以首先分析出该查询图可能所属的类目,另外还可以预测出该查询图的描述信息等,进而就可以基于这些信息为用户提供搜索结果。下面对具体的实现方式进行详细地介绍。

[0037] 首先需要说明的是,在本申请实施例中,为了能够确定查询图所属的类目和/或描述信息,主要采用的方式是,将查询图与数据库中的图像进行比对,由于数据库中的图像自身都带有所属类目以及描述信息,因此,如果能在数据库中找到一些与查询图相似的图像,则可以根据这些图像所属的类目确定出当前的查询图所属的类目,进而还可以确定出当前的查询图的描述信息。

[0038] 而为了便于在数据库中查找与当前查询图相似的图像,可以首先在线下对图像数据库中的图像进行视觉特征提取,并将每个图像对应的视觉特征保存到图像数据库中。具体实现时,可以从各个图像中进行视觉特征提取,在数据库中保存各个图像对应的视觉特征,以便当用户提交上查询图时,同样从查询图中提取视觉特征,然后将查询图的视觉特征与数据库中各个图像的视觉特征进行比对,找到与查询图相似的图像。当然,对于具体的图像而言,一般除了主体内容之外,还可能存在背景等其他内容,但只有主体内容能够体现出图像主要显示的内容。例如,某图像主要想展现某服饰,则只有其中人物的躯干部分属于该图像的主体内容区域。因此,在从各个图像中提取视觉特征之前,还可以首先对各个图像的主体内容区域进行检测,然后再基于主体内容区域进行特征视觉提取,这样可以避免图像的背景等对相似度判别准确度的影响。

[0039] 其中,关于具体如何进行主体内容区域检测,以及具体提取哪些视觉特征,后文中会有详细地介绍。这里需要指出的是,对于电子商务交易平台,图像数据库中的图像一般是

卖方用户上传的业务对象的图像,而卖方用户可能会为同一个业务对象上传多幅图像,其中一幅是主要图像(简称主图),本申请实施例中,可以仅针对业务对象的主图进行视觉特征提取。另外,由于系统中的卖方用户众多,并且也总是在不断上传新的业务对象的图像,因此,具体实现时,可以提取数据库中每天(也可以是其他时间长度)新增的业务对象的主图进行特征提取。当然,由于这些图像均是由用户上传的,因此可能存在图像质量(像素、清晰度等)不符合要求的情况,因此,还可以预先对图像质量进行判断,如果满足要求,再进行主体内容区域的检测,以及视觉特征的提取。这样,系统可以周期性的(比如每天)将计算好的图像特征推送到线上分布式图像数据库中,用于进行对查询图类目的预测,还可以用于后续搜索。

[0040] 这样,用户上传一张查询图后,系统就可以首先提取查询图的视觉特征,并将提取到的查询图视觉特征输入至线上实时分析器中。该分析器就可以根据相应的查询图视觉特征进行类目推断,同时还可以提取该类目相应的款式等描述信息。进而,就可以利用这些信息在线上分布式索引中进行查询,将得到的结果图像按照一定的规则进行排序,返回给用户。下面就基于图像搜索的过程对具体的实现进行介绍。

[0041] 参见图1,本申请实施例首先提供了一种图像搜索方法,该方法可以包括以下步骤:

[0042] S101:接收到输入的查询图后,对所述查询图进行视觉特征提取;

[0043] 对于提取查询图视觉特征的具体实现,与线下提取数据库中各图像的视觉特征的方式是相同的,并且提取的视觉特征类型也是对应的,因此,这里一并进行介绍。

[0044] 在具体实现时,提取的图像视觉特征可以是图像的颜色直方图、纹理、形状等全局特征,后续基于这些全局特征进行图像相似性的计算,还可以进行图像的搜索。然而,虽然这种全局特征能够对图像进行全局描述,但是在图像的细节方面却不能很好的进行区分。因此,在优选的实施方式中,可以采用全局特征(颜色、边缘)与局部特征相结合的方式对图像进行描述,进而基于这些特征,从数据库的图像中查找与查询图相似的图像,这样能够从全局和局部来对图像进行描述,提高预测的准确率。当然,在对准确率要求不高的情况下,也可以只提取全局特征或者局部特征。

[0045] 其中,具体在提取全局特征时,可以包括全局边缘视觉特征和/或全局颜色分布特征。具体在提取局部特征时,可以包括局部旋转不变特征。具体实现时,可以仅提取其中的任意一种视觉特征信息,或者,也可以同时提取其中的任意两种或者三种,等等。也就是说,在本申请实施例中,在提取视觉特征的数量上并没有特殊的限制,即使仅提取其中一种特征,也可以达到预测图像所述类目信息等的目的,同时还可以减少存储空间的压力。当然,如果这几种特征都进行提取,则一幅查询图可以提取出三种不同类型的特征,类似的,对于数据库中的各个用于建立索引的图像,同样可以提取这三种不同类型的特征,保存到数据库中。需要说明的是,无论是全局特征还是局部特征,都可以通过已有技术中的一些标准来提取,这里不进行详述。

[0046] 如前文所述,在进行视觉特征提取之前,还可以首先对查询图的主体内容区域进行检测,然后基于主体内容区域进行视觉特征提取,这样可以进一步提高预测以及搜索的准确率。其中,具体在进行主体内容区域检测时,通常可以采用图像分割、显著性检测等方法进行提取,这类方法依赖于图像色彩分布信息,计算量较大,可能会影响系统的性能,并

且当图像场景复杂时,可能无法准确分割出主体内容区域,对后续处理过程造成影响。因此,在本申请实施例中,针对主要内容是由模特展示服饰类内容的查询图,可以采用人脸检测的方式,来进一步确定图像的主体内容区域。

[0047] 优先地,可以首先利用人脸检测技术对查询图进行人脸检测,检测出人脸区域的位置及面积,然后根据人脸区域的位置、面积以及预置的人脸区域与躯干区域的比例关系,确定躯干区域的位置及面积,最后根据躯干区域的位置及面积,提取查询图的主题内容区域。

[0048] 例如,首先对输入图像(可以是查询图,也可以是数据库中的图像)进行人脸检测,若检测到人脸,则得到人脸圆形区域以及区域中心点坐标 $center(x,y)$;否则将整张图像作为服饰主体区域输出;

[0049] 接下来,由先验知识得知,人体躯干部分可以看成是一个矩形区域 $Rect$,其长和宽与人脸圆形区域的直径 R 有一个比例关系,根据该关系可以得到 $Rect$ 的长与宽。例如,根据实际情况可以推荐如下参数:

[0050] 长 $=3.5*R$;宽 $=2.5*R$;

[0051] 这样,由人脸区域中心点坐标 $center(x,y)$ 和矩形区域 $Rect$ 的长和宽,可以得到人体躯干矩形区域左上角的点 $P1(x,y)$,并根据点 $P1(x,y)$ 和 $Rect$ 的长和宽得到相应的服饰主体区域的坐标。

[0052] S102:将所述查询图的视觉特征与图像数据库中各图像的视觉特征进行相似度判别;

[0053] 在得到查询图的视觉特征之后,由于预先也对数据库中的图像进行了视觉特征提取,因此,就可以将查询图的视觉特征与图像数据库中各图像的视觉特征进行相似度判别。其中,如果仅提取一种类型的视觉特征,并且具体的视觉特征一般是以一定维度的向量来表示,因此,可以直接通过计算向量间的距离,来表示两个图像之间的相似度。但是,在实际应用中,有可能针对同一个图像提取多种不同类型的视觉特征,例如前文所述的既可能包括全局特征,又可能包括局部特征,而全局特征也可能有多种,等等。而在这种具有多种不同类型的视觉特征的情况下,如果想要计算两个图像之间的相似性,则一般情况下可以基于分类器进行。也就是说,人工的选取一批训练样本进行分类器模型的训练,然后再将查询图的各种视觉特征输入到分类器中,输出该查询图所属的类目。然而,这种方式一般具有以下弊端:

[0054] 第一,人工选取训练样本,人工成本过大,且选取过程带有主观性,对分类器模型的训练结果造成不利影响;

[0055] 第二,在实际的系统中,样本的分布往往呈现巨大的不均衡性,其表现在部分类目样本数量较多,而部分类目样本数量较少。这种样本分布的不均衡性会对分类器的训练过程造成较大的影响,使得最后训练得到的分类器模型不能很好的区分各类样本;

[0056] 第三,图像数据库中的图像数据量巨大,且图像场景十分复杂。因此,选取训练样本的数量较为困难:选取数量较少,不能很好的描述各类样本;选取数量较多,则对分类器模型的构建提出了巨大的挑战,需要较多的资源进行分类器的构建;

[0057] 第四,基于分类器的物体类目预测系统在正式上线后,需要定期进行分类器模型的更新,而这又涉及到训练样本的重新选取。整个过程耗费较多资源,不利于快速实时进行

系统更新。

[0058] 基于以上使用分类器预测类目时的缺点,本申请实施例提出了一种级联式的Re-Search图像相似性计算方法,也即,按照预先为各种视觉类型的特征设置的先后顺序,进行级联式的分层计算;其中,在每层的计算中,仅基于其中一种类型的特征进行相似度判别,并将符合层内预置条件的图像集合输入到下一层、以便基于下一种类型的特征进行相似度判别。

[0059] 例如,假设一幅查询图包含三种不同类型的视觉特征:全局边缘特征、全局颜色分布特征和局部旋转不变特征,并且图像数据库中的各幅图像也分别具有上述三种类型的特征。并且,假设预先设定的各种视觉特征的顺序为:全局颜色分布特征、全局边缘特征、局部旋转不变特征,则具体在进行比对时,可以如下进行:

[0060] 步骤一:将查询图的全局颜色分布特征和图像数据库中每张图像的全局颜色特征按照相似性度量方法1计算相似度,并选择出在部分相似度较高的图像集合进入下一步判别;

[0061] 步骤二:将查询图的全局边缘特征与步骤一中得到的图像集合中的每一幅图像的全局边缘特征按照相似性度量方法2计算相似度,并选择出部分相似度较高的图像集合进入下一步判别;

[0062] 步骤三:将查询图的局部旋转不变特征与步骤二中得到的图像集合中的每一幅图像的局部旋转不变特征按照相似性度量方法3计算相似度,并选择出部分相似度较高的图像集合。

[0063] 也就是说,每一步的判别都仅基于一种类型的视觉特征进行,并且每一步都能过滤掉一些图像,最终,步骤三得到的图像集合,就是在各种类型的视觉特征上都与查询图相似的图像,这就是所谓的级联式的判断。其中,各步骤中的相似性度量方法可以是相同的,也可以是不同的。也就是说,不同类型的视觉特征可以有不同的相似性度量方法。在申请实施例总,可以采用向量间的距离作为相似性度量方法。需要说明的是,本申请实施例中所采用的全局颜色分布特征、全局边缘特征以及旋转不变特征的比较顺序是可变的。

[0064] 这种级联式的判别方法相比训练分类器的方法,至少有以下几点优点:

[0065] 第一,该方法不需要任何训练样本,也不需要进行传统的分类器训练过程,可以节省大量的系统资源和分类器重新训练过程时所需资源;

[0066] 第二,该方法分层次进行相似度判别,每一层均采用不同类型的图像特征得到在某种视觉特征上与查询图像最相似的图像集合,并将其作为下一级的输入进行进一步筛选。

[0067] 第三,该方法只需要在离线阶段计算单一图像特征即可,后续特征计算均是实时计算。相比一次性计算图像的不同特征并进行融合的技术而言,这种方法对于系统的存储压力和计算资源的消耗都要更小。

[0068] 第四,该方法由于不需要传统机器学习的分类器训练过程,因此可拓展性强,能够扩展到更多的类目进行搜索。

[0069] 在得到与查询图在图像视觉上高度相似的图像组成的集合之后,就可以根据图像数据库中保存各个图像所属的类目,确定出当前查询图所属的类目。具体的,根据图像数据库中保存的各个图像的类目信息,确定相似度符合预置条件的各个图像分别对应的类目,

然后将出现次数最多的类目确定为查询图所属的类目信息。c例如,最终确定出与查询图最相似的图像一共有10幅,其中5幅都属于类目A,另外有2幅属于类目B,2幅属于类目C,1幅属于类目D,则可以确定出当前的查询图属于类目A。当然,具体实现时,还可以采用其他的决策方法来实现,这里不再一一介绍。

[0070] 进而,在确定出查询图所属的类目之后,还可以确定出该查询图的描述信息。具体的,可以提取相似度符合预置条件的图像中出现次数最多的类目对应的图像的描述信息,通过分析这些图像的描述信息获取当前查询图的描述信息。例如,在前述例子中,确定出查询图属于类目A之后,就可以取出该类目A对应的5幅图像,然后根据这5幅图像的标题等文字描述信息进行分词后,并进行分析,最终取出一些关键词作为该查询图的描述信息。

[0071] 当然,在实际应用中,也可以通过其他方式来预测查询图的描述信息,不一定是在预测出查询图的类目后再预测描述信息。另外,也可以仅预测查询图的类目信息或者描述信息,后续基于其中一方面的信息为用户提供搜索结果。当然,如果两方面的信息都获取到,则搜索结果的质量将会得到提高。

[0072] S104:根据所述查询图以及确定出的查询图所属的类目信息和/或描述信息进行搜索,返回搜索结果。

[0073] 在预测出查询图所属的类目信息和/或描述信息之后,就可以基于这些信息,从图像数据库中获取相关的搜索结果。具体的搜索过程,可以与用户提交查询图并提交了类目信息以及描述信息的情况相同,例如,可以首先在图像数据库中搜索该查询图所属类目下的所有业务对象,然后将该查询图的描述信息与各个业务对象的标题进行相似度判别,相似度满足条件的业务对象的图像再与查询图的图像特征进行比对,最终得到搜索结果返回给用户,等等,这里不再进行详述。

[0074] 总之,在本申请实施例中,如果用户需要进行图像搜索,则可以仅提交查询图,不需要同时提交查询图所属的类目、描述信息等其他信息,并且提交的查询图可以是图像数据库之外的任意图像;在接收到查询图之后,系统可以首先根据查询图的特征预测出该查询图所属的类目和/或描述信息,进而可以结合查询图以及类目和/或描述信息,共同得出与查询图属于同种类目和/或在款式、色彩等视觉特征上高度相似的图像集合作为搜索结果。这样,可以使得在不需要用户提供类目或描述信息的情况下,就能为用户提供符合用户需求的搜索结果,并且,通过查询图特征的对比预测得到的类目、描述信息等更为客观、准确,可以消除对用户输入信息的依赖。

[0075] 以上对本申请实施例提供的图像搜索的方法进行了详细地介绍,可以看出,在本申请实施例解决图像搜索过程中的问题的方案中,如何对一幅图像预测出所属的类目和/或描述信息,是解决问题的关键。而在实际应用中,如果能够针对一幅图像预测出所属的类目和/或描述信息,则在其他的应用场景下也能解决一些技术问题。例如,在电子商务交易平台中,卖方用户在上传业务对象的时候,需要为其选择相应的类目。然而,系统中的类目关系繁杂,卖方用户容易选错;同时有部分卖方用户可能会故意对类目进行错放,以达到搜索作弊等目的。而如果能够针对用户上传的业务对象的图像预测出其所所属的类目,则不需要卖方用户人工选取类目,减轻卖方用户选择类目的过程,提高用户体验;如果卖方用户选择了类目,也可以进行类目预测,如果预测出的类目与用户选择的类目完全不相关,则可以向系统管理员等发出提示信息,或者拒绝用户的提交,等等。这样,可以避免卖方用户通过

文本进行作弊的途径。因此,在本申请实施例中,将系统自动确定一幅图像的相关文本信息的方法单独进行保护。参见图2,该获取图像文本信息的方法可以包括以下步骤:

[0076] S201:获取待确定类目信息的目标图像,对所述目标图像进行视觉特征提取;

[0077] 这里的目标图像,就可以是指前文所述的用户在需要进行图像搜索时提交的查询图,或者卖方用户提交的业务对象的图像等等。具体在进行特征提取时,同样可以如前文所述,提取目标图像的全局和/或局部特征。

[0078] S202:将所述目标图像的视觉特征与图像数据库中各图像的视觉特征进行相似度判别;

[0079] 这里的图像数据库与前文所述的图像数据库类似,同样可以是在线下对数据库中已知类目以及描述信息的图像进行特征提取,并保存在数据库中。同样的,数据库中的同一幅图像可以提取多种不同类型的特征。因此,得到目标图像的特征之后,就可以与图像数据库中各图像的特征进行相似度判别。类似的,如果一幅图像对应多种不同类型的特征,则可以按照前文所述的级联式判决方法。

[0080] S203:根据与所述目标图像的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,获取所述目标图像所属的类目信息和/或描述信息。

[0081] 从图像数据库中找出与当前目标图像最相似的若干个图像之后,就可以根据这些图像的类目,预测出目标图像所属的类目信息。另外,还可以预测出目标图像的描述信息。

[0082] 需要说明的是,在图2所述的获取图像文本信息的方法,与图1所述的图像搜索方法中,获取查询图所属的类目信息和/或描述信息的过程实际是相同的,因此,相关的实现细节可以参见前文的介绍,这里不再赘述。

[0083] 总之,通过本申请实施例提供的上述获取图像文本信息的方法,可以根据用户提交的目标图像,自动根据目标图像的视觉特征以及图像数据库中图像的视觉特征,预测出目标图像所属的类目和/或描述信息。这样,在需要使用到目标图像的文本信息的应用中,就不再需要用户手动进行文本信息的输入,即使用户输入了文本信息,还可以根据预测出的信息对用户输入的信息进行验证,避免出现通过文本进行作弊等现象。

[0084] 与本申请实施例提供的图像搜索方法相对应,本申请实施例还提供了一种图像搜索装置,参见图3,该装置可以包括:

[0085] 特征提取单元301,用于接收到输入的查询图后,对所述查询图进行视觉特征提取;

[0086] 相似度判别单元302,用于将所述查询图的视觉特征与图像数据库中各图像的视觉特征进行相似度判别;

[0087] 预测单元303,用于根据与所述查询图的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,确定所述查询图所属的类目信息和/或描述信息;

[0088] 搜索结果返回单元304,用于根据所述查询图以及确定出的查询图所属的类目信息和/或描述信息进行搜索,返回搜索结果。

[0089] 具体实现时,为了避免图像背景等在进行相似度判别时的干扰,特征提取单元301具体可以包括:

[0090] 主体内容区域提取子单元,用于提取所述查询图的主题内容区域;

[0091] 特征提取子单元,用于对所述主体内容区域进行特征提取。

[0092] 其中,如果所述查询图的主要内容是服饰类内容,则特征提取单元301具体可以包括:

[0093] 人脸检测子单元,用于利用人脸检测技术对所述查询图进行人脸检测,检测出人脸区域的位置及面积;

[0094] 躯干区域确定子单元,用于根据所述人脸区域的位置、面积以及预置的人脸区域与躯干区域的比例关系,确定躯干区域的位置及面积;

[0095] 主体内容区域确定子单元,用于根据所述躯干区域的位置及面积,提取所述查询图的主体内容区域。

[0096] 具体在进行视觉特征提取时,特征提取单元301具体可以用于:

[0097] 对所述查询图进行全局特征和/或局部特征的提取。

[0098] 其中,所述全局特征包括全局边缘视觉特征和/或全局颜色分布特征,所述局部特征包括局部旋转不变特征。

[0099] 当提取出的特征为至少两种时,所述相似度判别单元302具体可以用于:

[0100] 按照预先为各种特征设置的先后顺序,进行级联式的分层计算;其中,在每层的计算中,仅基于其中一种特征进行相似度判别,并将符合层内预置条件的图像集合输入到下一层、以便基于下一种特征进行相似度判别。

[0101] 其中,预测单元303具体可以用于:

[0102] 根据图像数据库中保存的各个图像的分类信息,确定所述相似度符合预置条件的各个图像分别对应的类目;

[0103] 将出现次数最多的类目确定为所述查询图所属的类目信息。

[0104] 另外,特征提取单元301还可以用于:

[0105] 提取所述相似度符合预置条件的图像中所述出现次数最多的类目对应的图像的描述信息,通过分析该描述信息获取所述查询图的描述信息。

[0106] 总之,在本申请实施例中,如果用户需要进行图像搜索,则可以仅提交查询图,不需要同时提交查询图所属的类目、描述信息等其他信息,并且提交的查询图可以是图像数据库之外的任意图像;在接收到查询图之后,系统可以首先根据查询图的特征预测出该查询图所属的类目和/或描述信息,进而可以结合查询图以及类目和/或描述信息,共同得出与查询图属于同种类目和/或在款式、色彩等视觉特征上高度相似的图像集合作为搜索结果。这样,可以使得在不需要用户提供类目或描述信息的情况下,就能为用户提供符合用户需求的搜索结果,并且,通过查询图特征的对比预测得到的类目、描述信息等更为客观、准确,可以消除对用户输入信息的依赖。

[0107] 与本申请实施例提供的获取图像文本信息的方法相对应,本申请实施例还提供了一种获取图像文本信息的装置,参见图4,该装置还可以包括:

[0108] 特征提取单元401,用于获取待确定类目信息的目标图像,对所述目标图像进行视觉特征提取;

[0109] 相似度判别单元402,用于将所述目标图像的特征与图像数据库中各图像的视觉特征进行相似度判别;

[0110] 预测单元403,用于根据与所述目标图像的相似度符合预置条件的图像对应业务对象的类目信息和/或描述信息,获取所述目标图像所属的类目信息和/或描述信息。

[0111] 其中,具体实现是,预测单元403具体可以用于:

[0112] 根据图像数据库中保存的各个图像类目信息,确定所述相似度符合预置条件的各个图像分别对应的类目;

[0113] 将出现次数最多的类目确定为所述查询图所属的类目信息。

[0114] 通过本申请实施例提供的上述获取图像文本信息的装置,可以根据用户提交的目标图像,自动根据目标图像的视觉特征以及图像数据库中图像的视觉特征,预测出目标图像所属的类目和/或描述信息。这样,在需要使用到目标图像的文本信息的应用中,就不再需要用户手动进行文本信息的输入,即使用户输入了文本信息,还可以根据预测出的信息对用户输入的信息进行验证,避免出现通过文本进行作弊等现象。

[0115] 通过以上的实施方式的描述可知,本领域的技术人员可以清楚地了解到本申请可借助软件加必需的通用硬件平台的方式来实现。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可以以软件产品的形式体现出来,该计算机软件产品可以存储在存储介质中,如ROM/RAM、磁碟、光盘等,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或者实施例的某些部分所述的方法。

[0116] 本说明书中的各个实施例均采用递进的方式描述,各个实施例之间相同相似的部分互相参见即可,每个实施例重点说明的都是与其他实施例的不同之处。尤其,对于系统或系统实施例而言,由于其基本相似于方法实施例,所以描述得比较简单,相关之处参见方法实施例的部分说明即可。以上所描述的系统及系统实施例仅仅是示意性的,其中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模块来实现本实施例方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0117] 以上对本申请所提供的图像搜索、获取图像文本信息的方法及装置,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,对于本领域的一般技术人员,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。综上所述,本说明书内容不应理解为对本申请的限制。

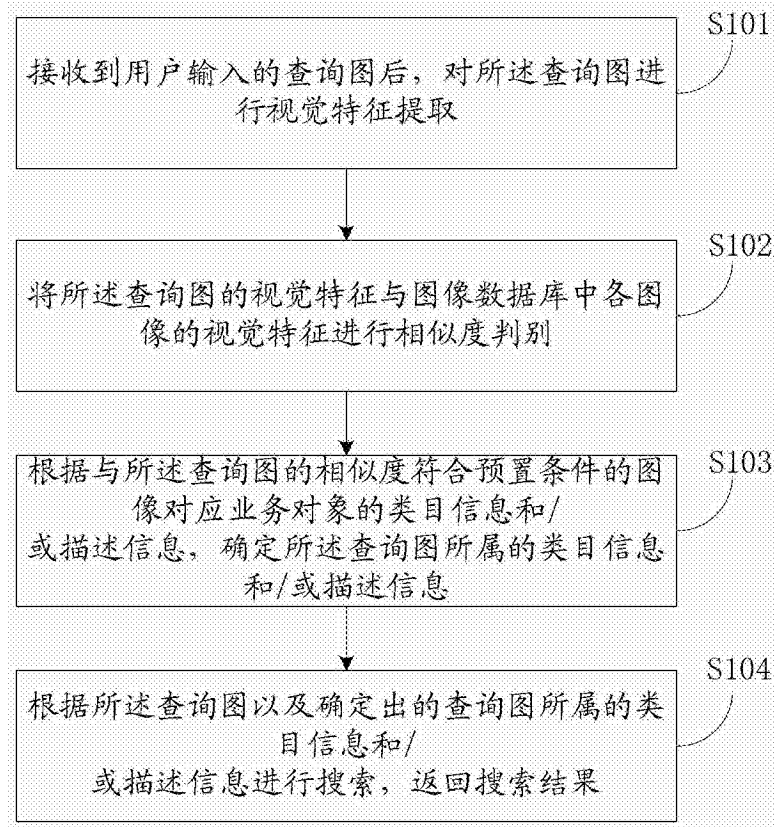


图1

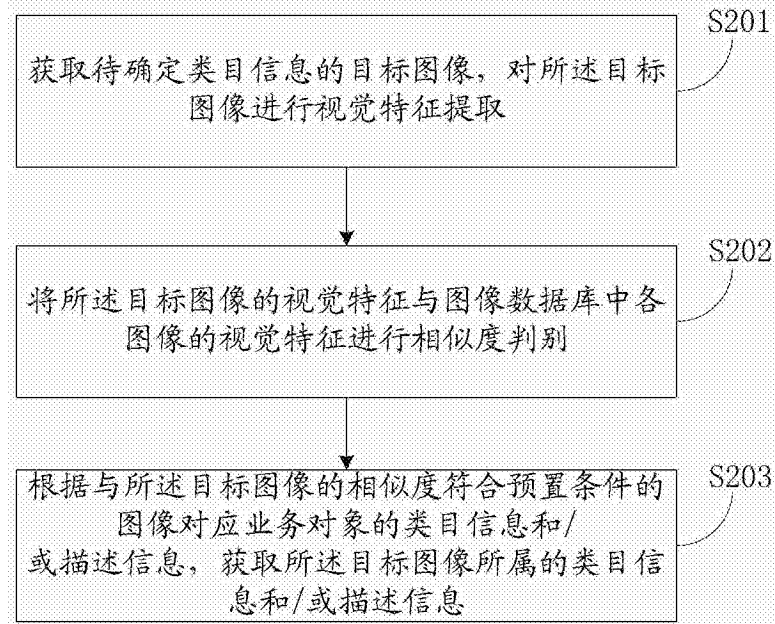


图2

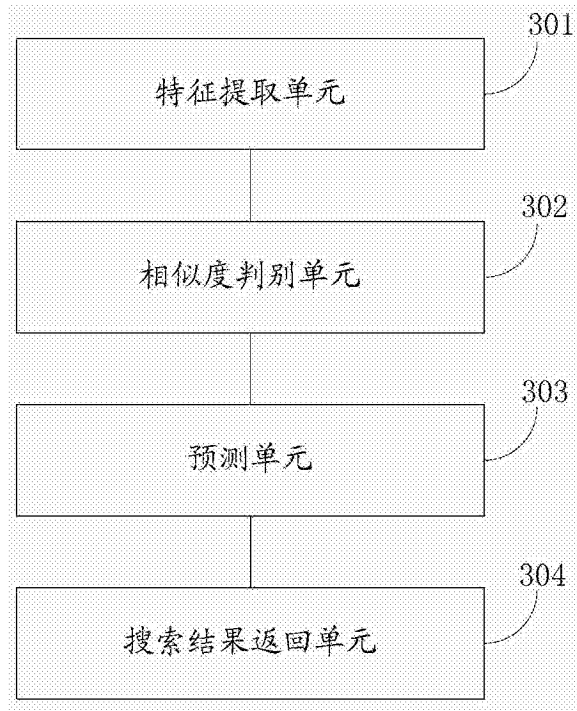


图3

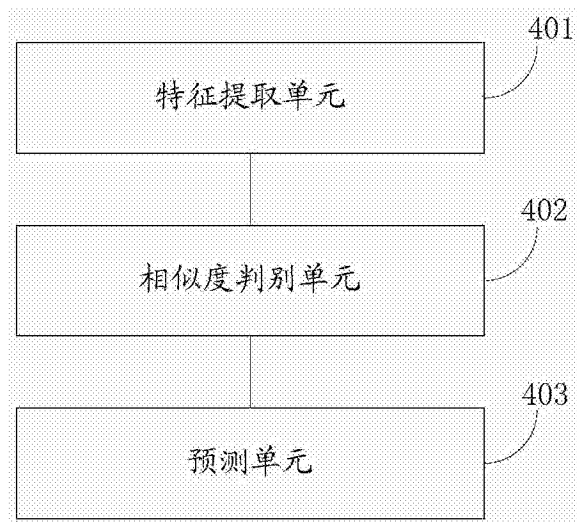


图4