

图像自相关特征识别的算法改进

陈杨华,张少林

(浙江科技学院 理学院,杭州 310023)

摘要: 将理论分析与具体操作相结合,简析相似图像识别的流程,包括关键点信息提取,如何利用关键点信息对颜色特征、纹理特征建立图内自相关模型。通过借鉴分词系统的思想,提出一种能够对特征数据进行结构化处理的方法。仿真结果表明,该方法能较准确地分配索引,并且能较快地匹配到相似图像。

关键词: 图像识别;自相关特征;数据索引;算法改进

中图分类号: TP391.413 **文献标志码:** A **文章编号:** 1671-8798(2017)01-0024-07

Improved algorithm of image auto-correlative feature recognition

CEHN Yanghua, ZHANG Shaolin

(School of Sciences, Zhejiang University of Science and Technology, Hangzhou 310023, Zhejiang, China)

Abstract: The paper combines theoretical analysis with specific operation to briefly analyze the recognition process of similar images, on how to extract information of key points and how to set up auto-correlative models within images for color features and texture features by using information of key points. By referring to the Lucene word segmentation system, a structure approach is proposed to process feature data. The simulation results show that this approach can accurately allocate index and match similar images at a fast speed.

Keywords: image identification; auto-correlative feature; data index; improved algorithm

图像识别与处理作为近几年国内外研究的热点,目前已有了深度学习、卷积神经网络等识别率较高的算法^[1-2],但这两种算法在具体图像识别应用中,如果学习率大就很有可能陷入局部极小值的问题。为了维持模型的稳定性,将学习率设置得很小,用大量的数据和训练时间来弥补缺陷,因此成本较大^[3]。这两种算法内部属于黑盒,存在大量的阈值和权值^[4],无法将局部特征单独提取成对应属性。针对上述不

收稿日期: 2017-01-17

基金项目: 国家级大学生创新创业训练计划项目(201511057015)

通信作者: 张少林(1961—),男,湖北省监利人,副教授,硕士,主要从事计算机应用研究。E-mail:1394009585@qq.com。

足,在中小平台没有巨大的数据支持情况下,使用笔者研究的方法具有以下优点:成本低,性价比高,采用距离匹配原则,不需要大量的数据和训练时间;使用高斯滤波进行特征提取的优化,因其对称及模糊化的特性,保证了特征匹配的精确性和稳定性,由此减轻了高维度下产生的过拟合影响;相比于其他特征提取方法,关键点特征提取在图像关键信息保留的情况下,有效减少了计算机运算时间和空间存储量,也减少了非关键像素产生的差样本过拟合现象的发生;在用图内自相关特征保证旋转不变性和光照不变性^[5]的同时,区分关键点权重提高准确性,而机器学习和神经网络算法它们本身是没有这种特性的;建立倒排表^[6]和特征词典的数据索引^[7],加快了检索的效率;过程透明,词典特征可以用来匹配对应文本属性进行检索。在本研究中,图像自相关特征识别的流程主要分为图像预处理、特征提取和索引建立三部分,其中,图像预处理部分使用关键点提取的方式,特征提取部分使用颜色自相关信息提取和纹理自相关信息提取方式,分别对应匹配图像的颜色和空间结构,索引建立部分使用特征词典的方法。

1 图像预处理——关键点信息提取

1.1 关键点信息提取的优点

相比其他预处理方法,关键点信息提取^[8]有以下几个优点:过滤了非关键信息,仅让关键信息进入纹理特征的提取,从而减少了差样本拟合现象的发生;增加关键点在图像特征中的比重,以减少噪声和背景的干扰,提高匹配的精确性;筛选出能近似表示图像信息的特殊点集合,确保在后续计算过程中保留较完整图像信息的同时,减少程序处理的数据量,加快检索速度。此外,因为要做特征词典链,如果非关键点加入提取,就会产生信息冗余,而单图像提取的特征尽量要简洁且具有代表性,这样才能高效地检索到目标图像,而关键点信息提取恰好有这样的功能。

1.2 关键点信息提取的相关步骤

1)对计算机中图像对应存储的 ARGB 颜色模型(透明度、红色通道、绿色通道、蓝色通道)进行灰度化,得到单通道灰度集 Y ,并将其中的物理存值简化到 8 位二进制保存。编程实现过程中可以使用位运算或字节转换等方式将 ARGB 模型通道分离。该过程能够大大减少纹理特征提取所需的时间。单像素灰度化过程如图 1 所示,其中 Y 的加权公式由人眼对光的敏感程度所得^[9]。

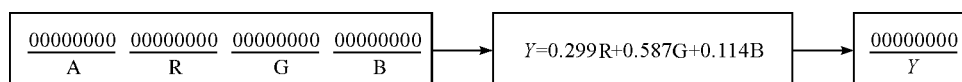


图 1 灰度化过程

Fig. 1 Gray processing

2)用不同标准差的二维高斯核 G 去对图像做卷积,在减轻噪声的同时能将原图构建出不同模糊程度和不同远近的图像尺度空间^[5]。高模糊度用来匹配轮廓,低模糊度用来匹配细节,其主要公式为

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}, \quad (1)$$

$$Y_{\text{center}} = \sum_{i=0}^N \sum_{j=0}^N G(i, j, \sigma) \times Y(i, j). \quad (2)$$

式 2 中: $Y(i, j)$ 与 Y_{center} 分别为卷积区域在 (i, j) 位置的灰度及卷积中心灰度; N 为高斯卷积窗口大小。根据二维高斯滤波函数的特性可知,在 $N=6\sigma+1$ 时能保留 99%以上的灰度能量,此时很难造成图像变暗。

3)对相邻尺度空间进行高斯差分,就是相邻两幅图对应位置相减,得到差分点集 D ,因为高斯滤波在不同的标准差下会形成不同的模糊程度,而关键点有边缘响应的特性,因此,通过相减的方法能够快速定位潜在关键点位置。

4)将点集 D 的灰度差看做第三维变量,用领域内 8 个点的灰度拟合三维二次函数。用 Hessian 矩阵、领域极值点、区域方差对点集做曲率筛选^[10]。Hessian 矩阵为

$$\mathbf{H}(x, y) = \begin{pmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{pmatrix}.$$

5) 合并图像在不同高斯核标准差下的公共关键点, 最终得到关键点位置 (i, j) 、主曲率及标准差 σ 。

2 特征信息提取

2.1 颜色自相关信息提取

颜色自相关特征提取的意义在于单一颜色直方图往往不能反映出图像结构, 为此, 在颜色直方图的基础上, 建立图内像素空间上的关联及加大关键点对颜色空间贡献度^[11]。其步骤如下:

1) 先将图像的 ARGB 颜色空间转为 HSV (Hue 色调, Saturation 饱和度, Value 明度) 颜色空间^[7], 将 HSV 颜色空间划分为 n 个区域, 获得区域集合 $C\{c_1, c_2, c_3, c_4, \dots, c_n\}$ 。其中每个 c 表示图内满足该区域颜色的像素点信息集合, n 可以视具体使用场景定义, 若图像库颜色类别多则细分; 反之亦然。

2) 由小到大建立距离集 $D\{d_1, d_2, d_3, d_4, \dots, d_m\}$, $d_1 = 0$, $d_m =$ 图像长和宽之间的最大值, 正常情况下 $d_m = m$, 每次距离集增量为 1, 当然, 这可以根据使用场景给定更好的值。

3) 对关键点位置像素赋予高权重, 提高相似匹配的精确性, 但要遵守一定的规则, 若关键点的数量占图像所有像素点数量的比例较大, 那么分给每个关键点的权值也就越小, 但关键点的权值仍大于等于非关键点的权值。

4) 依次在不同距离下 (D), 统计各区域 (C) 像素之间满足距离约束的像素总数, 形成 m 个空间像素量集合 S 。在距离为 d_k ($1 < k < m$) 下时, 单个 S_k 为

$$S_k = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ & s_{22} & \cdots & s_{2n} \\ & & \ddots & \vdots \\ & & & s_{mn} \end{pmatrix}.$$

5) 以长宽为 500 像素的图像和 HSV 模型划分为 64 个区域为例, 在上述过程后, 得到单图向量维度为 $500 \times 64 \times 64$, 单特征的维度为 64×64 需要进行低维化, 考虑到 k 从 1 到 500 的距离递增过程中, 距离集合中每个空间像素量集合是同等地位的^[12], 因此, 可以将所有空间颜色集合对应位置相加合并, 这样, 单图向量维度等于单特征向量维度仅为 64×64 。

6) 进行归一化, 去除缩放影响, 以方便颜色相似度测量^[12]。

2.2 纹理自相关信息提取

根据人眼识别特性进行图像灰度化后, 统计关键点领域梯度信息作为描述子, 将所有描述子的集合作为图像的纹理特征^[6]。步骤如下:

1) 获取大小为 $N \times N$ 的特征点窗口内像素梯度方向, 以相同窗口大小的二维高斯核平滑后, 做向量叠加, 获得主梯度方向。

2) 旋转 x 轴正方向至主梯度方向, 这样能够保证图像的旋转不变性。统计大小为 $N \times N$ 窗口内像素梯度方向 (经坐标旋转后) 与主方向的自相关角度差 β ($-\pi \leq \beta < \pi$), 其中 π 代表弧度制中的 180° 。以大小为 16 的高窗口 1 张图含 100 个关键点为例, 那么单图向量维度为 25 600, 单特征向量维度为 256, 可以通过固定区域的形式把高维向量映射到低维度区域上来减少维灾的发生。

3) 低维化, 构造区域直方图^[10], 将角度差划分为 n 个区域, 每个区域大小为 $2\pi/n$, 当 β 落在某个区域中时, 将其梯度模长加入到当前区域直方图 Histogram。如果将角度划分为 10 个区域, 那么 1 张图的向量维度为 1 000, 单特征向量度仅为 10。

4) 进行归一化, 即

$$\text{Histogram}_i = \text{Histogram}_i / \sum_j^n \text{Histogram}_j,$$

让周围的点分到的是相对于中心关键点的权重,这样做能够减轻光照的影响。

3 数据结构化

3.1 结构化方法——Lucene 分词系统

Lucene 分词系统^[13]以文章中的关键词作为文章特征,通过关键词在所有文章中的比例,以及在对应文章中出现的频次来分析关键词和文章的对应关系,建立关键词索引,合并相同词汇后达到树搜索的效果,从而大大加快了搜索速度。延伸到图像识别过程中,通过分析图像的关键点处的特征信息,建立图像特征词典用来模拟 Lucene 中的关键词索引,以达到树搜索的目的,并且加以改进,让搜索路线按照可能最大相似度为目的进行动态搜索,通过频率和编号的调整来达到优化存储的效果。

3.2 结构化的原因

1)完整图像信息所占内存非常大,特别是在生成自相关特征描述子阶段,如果不进行优化,那么处理的数据量增大,内存中的保存图像数量也会变少,导致文件读取的 IO 操作变多,而这样非常耗时。因此,尽量把信息合并和压缩后放到内存中去,才能够大大加快检索效率。

2)顺序匹配的效率非常低下,会造成很多资源的浪费,比如选取 TOP5(相似度最高的 5 幅图像所组成的集合)时,需要匹配所有特征后才能得出最接近的 TOP5 图像集。但是,如果用词典序检索,因为特征的合并,能够一次匹配多个特征,并且能计算出对应图像剩余可能相似度,当其不足以改变 TOP5 的排名时,不进行匹配。

3)从代码层面来看,检索功能和特征提取功能也可以分别建立模块,这有利于减少功能间的相互影响,更有利于系统的扩展和代码的复用。模块图如图 2 所示。

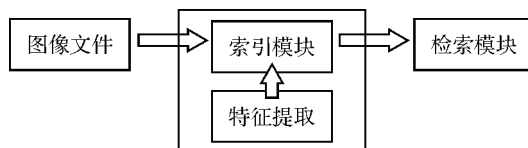


图 2 检索功能和特征提取模块图

Fig. 2 Retrieval function and feature extraction module

3.3 建立图像词典

建立词典对数据进行压缩,对索引进行结构化。自相关特征主要遵循两点规则:横向比较,若该特征在所有的特征中出现比例低,则越能体现该特征对于索引图像的重要性;纵向比较,若该特征相似合并后占图内特征数量比大,则也越能体现该特征对于索引图像的重要性。相关过程如下:

1)提取图像的关键点、颜色特征、纹理特征的文本信息。

2)用欧氏距离对图像集内的每个特征进行距离及相似度的计算,合并相似特征,其中占比较大的称为高频特征,提高其词典所占权重。以圆为例其高频特征就是在高斯窗口内的那段弧,如图 3 所示。

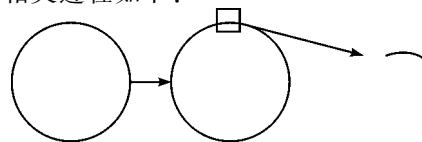


图 3 高斯窗口内的弧

Fig. 3 Arc within the Gauss window

3)将特征作为词典,以图像链表的方式进行存储,称之为倒排表^[13]。确定匹配模型的图像与特征的关系,建立特征匹配的索引结构。建立数据索引倒排表,结构如图 4 所示。

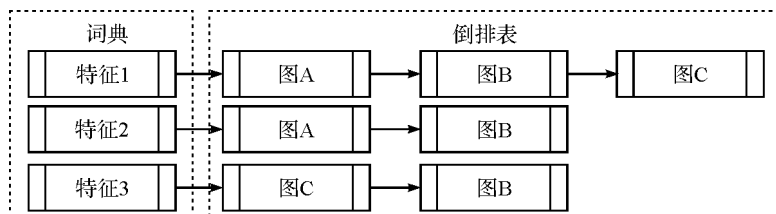


图 4 特征匹配的索引结构

Fig. 4 Index structure of feature matching

4)将词典和倒排表转换成文本信息。

3.4 压缩倒排表

由于倒排索引文件往往占用巨大的磁盘空间,引进压缩算法使得磁盘占用减少,从而加快检索效率。压缩数据处理倒排表包括索引编号、出现频次和位置序列三部分。描述一个压缩算法的三大指标为压缩率、压缩速度和解压缩速度,压缩过程如下:

1)依次比较属性规定的特征与其他自相关特征的相似度。如果极为相似,则记录下属性编号,匹配特征位置序列及统计该特征出现频次,得到如图 5 所示的词典链表。

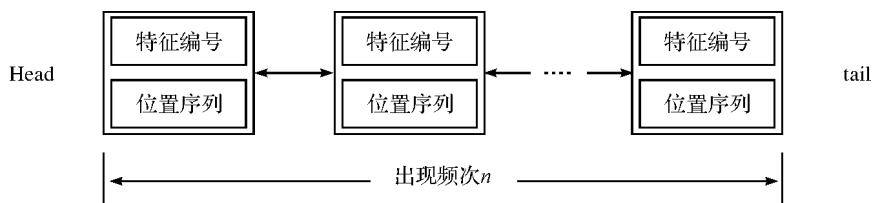
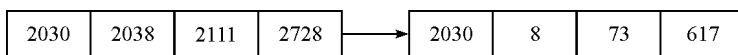


图 5 记录特征出现频次的词典链

Fig. 5 Dictionary chain for recording feature occurrence frequency

2)对所有词典链表按照出现情况频次进行递增排序。根据 zipf 定律^[14],在相对较大图像库中频次排名前 20% 的高频特征占有所有图像的 80%,而其余 80% 的特征仅占 20%。因此,重新分配编号,对频次较高的属性给出占用内存小的编号。同理,将剩下占用内存较大的编号分配给频次低的属性特征。

3)将编号存储改用编号差值存储,对检索大量图片能够有效降低存储量^[15],例如:



4 JAVA 下的测试实例

4.1 实验环境

正常情况下 JAVA 处理的速度是 C 处理速度的 1/20~1/10,因此图像检索速度比较慢,但不影响算法对比。运行环境配置:win10、64 位操作系统、虚拟内存 4 G、CPU 功率 2.0 GHZ、JDK1.8。实验图片信息:分辨率 500×500,位深 24 位,占用内存 732 kB,格式为 PNG。

4.2 单图提取特征时间

1)图像颜色数据灰度化,灰度 Y 集合用数组数据保存,平均花费时间约 30 ms。程序默认参数设置最小图像长宽大小为 64 像素,因此,默认尺度为分别 1、0.5、0.25,初始高斯核标准差为 1.6。高斯核过滤及完成高斯差分 DOG 运行大约时间分别为 168、31、8 ms。由所处位置能轻易获取到点的主曲率、所在尺度及区域方差等信息。原图及过滤后关键点位置如图 6 所示。



图 6 关键点提取

Fig. 6 Extract key points

2)原 RGB 模型转 HSV 模型所需时间大约 97 ms。颜色空间集合转换时间约 320 ms。默认颜色空间距离集合从 1 到最大距离 500,颜色划分为 16 个区域,共 256 维。

3)描述子生成时间约 798 ms。

4)索引建立时间约 401 ms。

在上述配置中,完成 1 次自相关特征提取并建立索引所需时间约为 1 853 ms,索引建立时间约占总提取时间的 21.64%。

4.3 单图特征数据量

1)归一化之后数据存储为 double 类型,一个描述子 128 维,占数据量大,纹理自相关特征数据量约为 202 kB。

2)相比较纹理特征,颜色自相关特征数据量就比较少了,约为 1 kB。

3)词典建立后,由于信息量的增加,数据量约为 225 kB,压缩后数据量约为 56 kB。

4)实验将原图复制 1 份并顺旋转 90°与原图构成新的图像库。压缩后数据量约为 70 kB。

500 个空间颜色集合合并后仅存储 1 kB。纹理特征在结构化时,因单幅图像的特征额外附带描述子大小、所在尺度、位置序列及出现频次等结构信息,因此所占内存反而增大。而压缩词典过程中,相似的描述子不再记录 128 维的 double 类型数据,仅仅存放结构信息,所以数据量减少。

4.4 检索结果仿真

由于是建立在中小数据量的前提下,因此,用本研究检索算法只与常用特征提取的算法进行比较,没有必要与深度学习和神经网络等需要海量数据的算法进行比较。笔者随机选取 100 张经过有损压缩后的 JPG 图像放入图像库,分为 T 恤、毛衣、裙子、裤子、鞋子 5 类,每类 20 张,如图 7 所示。所有图像经过旋转和调节明暗程度后放入等待检索库,测试每张图检索结果的 TOP3 准确率与类目 TOP5 召回率。每个图像大小接近 10 kB,图像库文件夹大小为 987 kB,图像索引总数据量为 142 kB。同一内存下,索引能表达的图像数量为初始文件所能表达图像数量的 6.95 倍左右。仿真结果见表 1。



图 7 检索图像

Fig. 7 Retrieval image

表 1 仿真结果

Table 1 Simulation results

采用算法	TOP3 查找准确率/%	TOP5 类目召回率/%	平均所用时间/ms
颜色空间法	73.00(73)	64.40(322)	234.64
CEDD 算法	71.00(71)	66.80(334)	316.58
LBP 旋转不变模式	73.00(73)	69.40(347)	670.35
Sift 算法	76.00(76)	68.00(340)	819.04
Surf 算法	78.00(78)	71.10(356)	454.29
本研究的检索算法	82.00(82)	74.20(371)	377.12

注:括号内为匹配到的图像数量。

5 结 语

本研究对图像自相关特征检索流程进行分析,与常用特征提取方法进行相比较,其准确率和用时有

所提升。实际图像特征提取的编程实现过程非常复杂,可以借助一些已有的框架进行改良。本研究的特征提取算法为改良后的 sift 自相关算法与颜色空间自相关算法的结合。有了初步的基础后,可以进一步研究如何在建立特征词典与对应属性的联系时,用文本属性去描述图像特征,整合文本检索。此外,索引的文档目录树还需进一步拆分,以提升识别速度。

参考文献:

- [1] SUN Y, WANG X G, TANG X O. Deep learning face representation from predicting 10,000 classes[C]//Computer Vision and Pattern Recognition(CVPR). Columbus: IEEE,2014:1891.
- [2] ANDREAS J, ROHRBACH M, DARRELL T, et al. Deep compositional question answering with neural module networks[C]//Computer Vision and Pattern Recognition(CVPR). Las Vegas: IEEE,2016.
- [3] GRAVES A, MOHAMED A R, HINTON G. Speech recognition with deep recurrent neural networks[C]//IEEE International Conference on Acoustics. Vancouver: IEEE,2013:6645.
- [4] HERMANN K M, KOCISKY T, GREFFENSTETTE E, et al. Teaching machines to read and comprehend[C]//Advances in Neural Information Processing Systems(NIPS). Quebec: NIPS,2015:1684.
- [5] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004,60(2):90.
- [6] 杨安生. 基于倒排表的中文全文检索研究[J]. 情报探索,2009(7):78.
- [7] SUN H A, ZHANG L. The design and implementation of the medical education vertical retrieval based on lucene. net [J]. Open Education Research,2011,58(2):295.
- [8] 窦育民,黄仁杰,李涛,等. 基于关键点的轮廓描述方法[J]. 计算机应用研究,2015,32(5):1525.
- [9] 刘庆祥,蒋天发. 彩色与灰度图像间转换算法的研究[J]. 武汉理工大学学报(交通科学与工程版),2003,27(3):345.
- [10] 彭兴璇,唐雪娇,董星. 基于改进 SIFT 算法在图像匹配中的研究[J]. 微型机与应用,2015(20):37.
- [11] 王保平,赵静,苏建康,等. 基于局部颜色-空间特征的图像检索方法研究[J]. 计算机应用与软件,2012,29(5):28.
- [12] 汤慧,唐朝晖,彭铁光. 基于 HSV 空间颜色特征的图像内容检索及应用[J]. 中国科技信息,2011(12):90.
- [13] 郑榕增,林世平. 基于 Lucene 的中文倒排索引技术的研究[J]. 计算机技术与发展,2010,20(3):80.
- [14] 刘胜久,李天瑞,珠杰. Zipf 定律与网络信息计量学[J]. 中文信息学报,2015,29(4):90.
- [15] 林俊鸿,姜琨,杨岳湘. 倒排索引查询处理技巧[J]. 计算机工程与设计,2015,36(3):573.