

基于 R 树和特征提取的相似图片搜索

彭友
软件 22

2012013312
18811369567

a623917@126.com

邱泓钧
软件 22

2011013259
18801375020

children19930928@gmail.com

赵馨逸
软件 21

2012013277
13426413257

stefanie_xin@163.com

简介

在本文中介绍了使用 R 树进行相似图片的检索实现，以及结果的对比分析。基于 R 树，我们提供了包括 RGB，HSV 以及灰度共生矩阵三大类不同维度的 16 种特征测试。

关键词

R 树;图片特征提取;

1. 引言

本次作业内容为图像检索，其中多维数据的处理是一个关键问题，图像检索的重要部分是对图像的特征向量索引和图片的特征向量提取两个方面。本文将从这两个方面分别介绍我们的方法，探究 RGB，HSV，灰度等特征向量在 R 树索引组织中的 K 近邻运行效果。

2. 特征提取方法

2.1 颜色特征

2.1.1 简介

获取每张图片的各像素的 RGB 或 HSV 值，并根据颜色空间的三维划分方式将像素点对应到不同区块，最后根据各个区块的像素比例作为图像的特征。这种方法可以描述不同色赛在整幅图像所占的比例，但是忽略了每种色彩的空间位置。

2.1.2 统计方法

对三维的颜色空间，每个维度可以自定义划分成 dimI, dimJ, dimK 份，从而形成 $\text{Dim} = \text{dimI} * \text{dimJ} * \text{dimK}$ 个不同的区块，统计每个区块包含有多少像素点。另外，考虑到每张图片的尺寸不同，进行归一化处理，使得每个直方图的值表示像素点落在该颜色区间的概率。

2.1.3 RGB 颜色空间

RGB 色彩模式是工业界的一种颜色标准，RGB 代表红、绿、蓝三个通道的颜色，范围都是 0~255。

主要划分成 8 维(2*2*2)、12 维(2*2*3, 2*3*2, 3*2*2)、16 维(2*2*4, 2*4*2, 4*2*2)、18 维(2*3*3, 3*3*2, 3*2*3)颜色空间作为向量，用于 R 数索引分析。

2.1.4 HSV 颜色空间

HSV 另一种颜色的表示方法，而且更接近人对颜色的感觉。这个模型中颜色的参数分别是：色调 (H)，饱和度 (S)，亮

度 (V)。其中 H,S,V 的范围分别为 0~360,0~1,0~1。

通过 RGB 可以计算得到 HSV。

主要划分成 8 维(2*2*2)、12 维(3*2*2)、16 维(4*2*2)、20 维(5*2*2)、24 维(6*2*2) 颜色空间作为向量，用于 R 树索引分析。

2.2 纹理特征

2.2.1 简介

由于纹理是由灰度分布在空间位置上反复出现而形成的，因而在图像空间中相隔某距离的两个像素之间存在灰度关系。这种方法可以描述出图片一定的纹理相关特性。

2.2.2 灰度共生矩阵

首先通过公式： $Y=0.229*R+0.587*G+0.114*B$ 将原始的彩色图像转换成灰度图像，并将 0~255 的灰度划分成 16 个级别，通过公式计算出 16*16 的灰度共生矩阵。

在这里我们计算出 0° ， 45° ， 90° ， 135° 四个方向的灰度共生矩阵，并对这四个方向分别计算它的能量、对比度、熵、IDM 相关度以及自相关这 5 种特征，接对所有图像得到的结果在[-1,1]区间进行归一化，得到一个 20 维向量用于 R 树索引。

3. R 树索引机制

3.1 实现方式

在试验中，我们采用了用 R 树来实现图片的特征向量存储与查询，接下来我们来简要介绍部分 R 树的实现方式。

R 树我实现了最小包围矩形的形式实现了区域的划分。本实验中 R 树的插入和老师上课讲的一样，是通过递归的思路使得插入。分裂节点采取的是递归分裂，使得最终插入的节点面积最小。而 k 近邻查询是通过维护一个最近邻数组，在查询的不断的修改该数组，使得在该数组始终是按照到 data 点的紧邻的远近而排序，故在便利节点中始终只需要比较数组的最后一个节点到 data 的距离和现在查询节点到 data 的距离，而这样同时也可以进行剪枝，根据比较点到 MBR 的最短距离与当前序列中最长的距离来进行剪枝。对于 r 树的插入采用的方法是首先得到节点到区域的距离的到区域的索引，通过然后通过判断是否是叶子节点和非叶子节点，如果是叶子节点则插入到节点中，如果需要分裂则分裂两个节点，将一个节点保存，然后向上构建非叶子节点来保存这两个节点，然后就通过非叶子向下递归分裂节点，然后改变非叶子节点的区域，调整非叶子节点，如果不分裂则调整包含区域，如果分裂则将分裂出的节点区域插入。

3.2 具体改进[5]

使用了分支与界法的思路，核心思路是比较每个块到查询节点的最短距离和到节点到当前所有块最远距离中最小值进行比较，不断对查询的块进行剪枝（如果该块到待查询节点的最小距离大于当前所有块到节点的最远距离中的最小值）来实现 k 近邻查询。

3.3 查询算法[6]

我们实现了范围查询（r 树的头文件中自带的），欧几里得距离的 KNN 近邻查找、曼哈顿距离的 KNN 近邻查询和切比雪夫距离的 KNN 查询。

4. 结果比较与评估

这部分比较不同特征向量提取情况下的 R 树磁盘访问次数以及搜索效果

4.1 实验环境

系统：window 8
开发工具：VS2012
语言：C++
运行要求：不能配OpenCV
调试模式：Release

4.2 测试方法

本实验采用了两种测试方法进行测试，分别是随机生成数据和利用自己提取的老师给的图片集的特征文本进行测试。

测试集的数据量分别为 1000,2000,3000,4000,5000。图片集使用老师所给的图片集，图片个数为 5613。

4.3 R 树性能

R 树的性能分析如图所示：

4.3.1 10 近邻，分裂结点的上下限分别是 32-16

维度 数据量	HSV8 (222)	HSV12 (322)	HSV16 (422)	HSV24 (622)
1000	45.972	55.638	45.967	42.902
2000	93.692	132.662	105.197	92.902
3000	144.418	190.748	167.81	137.902
4000	193.418	287.34	232.81	184.902
5000	242.13	339.802	296.81	224.902

特征 数据量	RGB (222)	RGB (233)	RGB (322)	RGB (422)
1000	46	61	57.156	62
2000	93.998	122	139.326	128
3000	145.908	186	198.148	192
4000	194.786	250	302.838	256
5000	244.872	313	355.824	319

数据趋势：R 树的平均查找次数随维度上升先是有巨大的提升，然后硬盘访问次数会先增大后减小，我们生成的 HSV 的 10 近邻的硬盘访问量大概会在 12 维左右的时候达到最大硬盘访问量。

从列表中我们可以看出，在相同的总维度下，当蓝色的划分比较高的时候，匹配的准确率明显高于偏重划分红色和绿色的特征。

特征 数据量	HSV(16 维)	HSV(20 维)	HSV(16 维)	HSV(20 维)
1000	0.300533	0.2927	0.287025	0.27635
2000	0.326878	0.327989	0.312492	0.312542
3000	0.310278	0.314333	0.295563	0.299542
4000	0.30645	0.311847	0.29204	0.29763
5000	0.300038	0.30628	0.286077	0.292608

HSV 由于比较符合人眼的认知，所以最后的正确率在 30、40 近邻的情况下还达到了 30%左右，相对是比较高的。

数据分析：当 R 树的维度比较高的时候，各个节点相等于被稀释了，然后剪枝的效率比较高，所以到比较高维的时候磁盘访问量就会急剧下降。

4.3.2 随机生成的数据测试：

结点上下限分别是8和4，随机数的范围是0-10

维度 节点 个数	4	8	12	16	20
1000	272.12	210	259	185	173
2000	587.33	489.994	460	410	368
3000	877.194	874.999	720	611	558
4000	1244.77	1236	996	833	759
5000	1611.19	1456.99	1221	1006	940

数据分析：由于随机生成的数十分的平均，所以随着维度的升高，剪枝的效率会变大，导致硬盘访问量会下降。

4.4 搜索精度

测试都是基于测量 10,20,30,40-近邻查询的平均数据，并将其置为表格。（其中准确率是 0.01 表示 1%）

这

4.4.1 颜色特征

颜色特征方面我们测了 15 种不同的特征，其中 RGB 有 8、12、16、18 四个维度的测试，每个维度进行了更细化的维度分配。

限于篇幅，我们贴出了部分数据，全部数据可以参见附录（其中准确率是 0.01 表示 1%）

30 近邻与 40 近邻

特征 数据量	RGB (233)	RGB (332)	RGB (224)	RGB (422)
1000	0.265533	0.2392	0.269533	0.233767
2000	0.302133	0.284217	0.286217	0.27065
3000	0.276522	0.264778	0.264556	0.249211
4000	0.28892	0.28455	0.276067	0.264467
5000	0.278607	0.27846	0.269833	0.261673

4.4.2 HSV

部分数据如下，详情参见附录：（其中准确率是 0.01 表示 1%）

4.4.3 灰度共生矩阵

其中准确率是 0.01 表示 1%，灰度共生矩阵是 20 维的。

近邻数 数据量	10	20	30	40
1000	0.3069	0.19065	0.175867	0.167475
2000	0.313133	0.221483	0.204456	0.195
3000	0.299733	0.220008	0.201644	0.191642

4000	0.30516	0.232545	0.212923	0.20178 2
5000	0.31002	0.242273	0.221631	0.20988 7

分析可知，灰度共生矩阵在 10 近邻的时候正确率比较高，但是随近邻数下降的很快。

4.4.4 距离测量方式

我们实现了三种距离测量，欧几里得距离，曼哈顿距离和切比雪夫距离，来实现 k 近邻搜索。

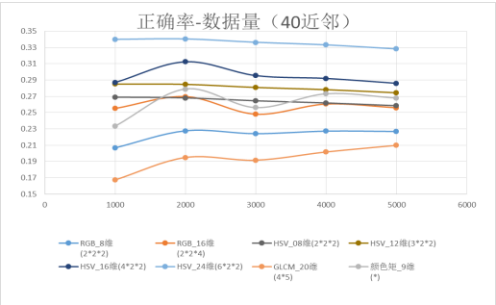
欧几里得距离计算方式是：每个维度的差值的平方和（本实验中没有开方），曼哈顿距离是每个维度的差的绝对值的和，切比雪夫距离是所有维度差值的绝对值中的最大值。

4.4.5 已给样例特征

近邻 数据量	10	20	30	40
1000	0.3124	0.2642 5	0.2430 3	0.23357
2000	0.36265	0.3118 7	0.2896 1	0.27913
3000	0.3462	0.2912 5	0.2684 7	0.25623
4000	0.3619	0.3078 5	0.2856 7	0.27319
5000	0.35362	0.3016 9	0.2804 0	0.26798

我们可以看出，样例的特征在 9 维的低维度的情况下达到大概 30% 的正确率，相比之下，效果还是不错的。

4.4.6 总结



从上面的数据我们可以看出，

各特征随数据量增大正确率的变化曲线特征：

RGB：曲线呈现“M”型，在 2000 和 4000 数据量达到极大值，在 1000,3000 和 5000 时正确率较小。

HSV：正确率普遍较高，其中 24、12、8、4 维的曲线是下降的近直线，16 和 20 维曲线在 2000 数据最大并向两侧递减。

总体来说，RGB 的准确率没有 HSV 的高，样例中的颜色矩能够达到一个很高的正确度。

5. 改进策略（加分项）

我们改进了相似度匹配的方法，尝试了欧几里得距离、曼哈顿距离和切比雪夫距离。然后还通过更改各个维度的权重实现更精确的匹

配，以 RGB 直方图（见 4.4.1 对 RGB 划分的分析）为例，我们在相同维度下通过改变 RGB 的划分数目来改变各个颜色的权重。如下图所示，图中是曼哈顿距离和欧式距离的对比，以及切比雪夫距离和欧式距离的对比。

RGB_8 维
(2*2*2)

特征 数据量	欧几里得	曼哈顿距离	切比雪夫	曼/欧	切/欧
1000	0.2982	0.3008	0.3008	100.87%	100.87%
2000	0.319233	0.320833	0.320833	100.50%	100.50%
3000	0.315683	0.317717	0.317717	100.64%	100.64%
4000	0.3184	0.32119	0.3212	100.88%	100.88%
5000	0.3164	0.31953	0.31928	100.99%	100.91%

RGB_16 维
(2*2*4)

特征 数据量	欧几里得	曼哈顿距离	切比雪夫	曼/欧	切/欧
1000	0.3441	0.3566	0.3567	103.63%	103.66%
2000	0.36225	0.363667	0.363733	100.39%	100.41%
3000	0.341267	0.355917	0.355967	104.29%	104.31%
4000	0.352275	0.35878	0.35882	101.85%	101.86%
5000	0.34252	0.356913	0.356947	104.20%	104.21%

HSV_08 维 (2*2*2)

特征 数据量	欧几里得	曼哈顿距离	切比雪夫	曼/欧	切/欧
1000	0.3207	0.3235	0.3235	100.87%	100.00%
2000	0.331167	0.334233	0.334233	100.93%	100.00%
3000	0.32365	0.327567	0.327567	101.21%	100.00%
4000	0.32343	0.32737	0.32737	101.22%	100.00%
5000	0.320367	0.324247	0.32424	101.21%	100.00%

实际的数据显示，曼哈顿距离好于欧式距离和切比雪夫距离，这是由于不同图片的维度是相互独立的，这样的情况下曼哈顿距离更好的拟合了这种方式。

6. 参考文献

- [1]基于 R-树的空间最近邻查询研究 韩冬柏
- [2]刘鹏宇，基于内容的图像特征提取算法的研究，2002.09-2004.06<http://www.docin.com/p-344263491.html>
- [3]彭辉，基于纹理特征的图像分类识别，计算机与信息技术，2007