

# 一种多特征融合的图像检索新方法

金 铭 汪友生 边 航

(北京工业大学 电子信息与控制工程学院 北京 100124)

**摘要:** 如何提高基于内容的图像检索的准确率是图像检索领域中的一大难题,其关键在于特征提取方法。针对单一特征检索效果的不足,从图像的多特征入手,先使用颜色和边缘的方向性描述符提取图像的颜色和纹理特征,再使用分层梯度方向直方图提取图像的形状特征,最后采用图像底层特征串行关联的方法将3种特征结合进行检索。使用Corel标准图像集进行测试,结果表明本文方法的查准率和 $AvqP$ 值比两种原算法均有提高,是一种有效的多特征融合的图像检索方法。

**关键词:** 图像检索;多特征融合;颜色和边缘方向性描述符;分层梯度方向直方图

**中图分类号:** TP391.9 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## New image retrieval method based on multi-feature integration

Jin Ming Wang Yousheng Bian Hang

(College of Electronic and Control Engineering, Beijing University of Technology, Beijing 100124, China)

**Abstract:** How to improve the accuracy of content-based image retrieval is a difficult problem, while its key point is feature extraction method. Aiming at the deficiency of single feature image retrieval, we propose an effective image retrieval method based on multi-features of images. We firstly use color and edge directivity descriptor (CEDD) to extract color and texture features of the images, then use pyramid histogram of oriented gradients (PHOG) to extract shape features, finally we adopt the feature serial correlation method to integrate the three image features into retrieval. The experiment using Corel image sets shows that the precision and the  $AvqP$  value of our method both perform better than the two original algorithms.

**Keywords:** image retrieval; multi-feature integration; CEDD; PHOG

## 1 引言

基于内容的图像检索(content-based image retrieval, CBIR),是使用图像的可视特征对图像进行相似性匹配的检索方式。这种检索方式可以自动提取图像库中所有图片的底层视觉特征,以高维向量的形式存储并建立特征库,检索时以计算向量间距离的方式获得相似性最高的若干张图像返回给用户<sup>[1-3]</sup>。与基于文本的图像检索方式相比,CBIR系统工作时更高效,但其检索准确性依赖于特征提取算法。

图像检索中常用的图像特征主要有颜色特征<sup>[4]</sup>、纹理特征<sup>[5]</sup>和形状特征<sup>[6]</sup>等。颜色特征是描述图像最直观的底层视觉特征,也是在图像检索中应用最广泛的视觉特征。颜色特征通常和图像中所包含的对象或图像的场景非常相关,能最直接地描述人类视觉系统对图像的感知。此外,与其他底层视觉特征相比,颜色特征对图像自身的尺寸、视

角、方向等敏感度不高,从而在检索时具有较高的鲁棒性。纹理特征是描述图像的另一种重要的底层特征,通常认为纹理是图像上表现出的灰度或颜色分布的某种规律性,这种规律性在不同类别的纹理中有着不同的特点,能够刻画出像素邻域的灰度空间分布规律,因此可以用来识别出不同的图像。形状特征是描述物体最本质的特征,也是描述难度非常大的图像特征,是进行物体识别的关键信息。其不因周围环境而改变,鲁棒性非常高,对形状的理解往往决定了对整幅图像涵义的理解。目前,已经有很多文献基于以上单一特征进行图像检索,而单一特征只有在对相应特征比较敏感的特定图像集内才能取得较好的检索效果,对于那些特征综合性较强、内容比较复杂的图像,检索效率则比较低,所以迫切地需要研究融合多种图像特征的检索方法来提高CBIR的工作效率<sup>[7-9]</sup>。

目前有一些特征提取算法结合了图像的颜色和纹理特征<sup>[10-11]</sup>,但是如何将这两种底层特征紧凑地表示成向量,是

一个很有难度的问题。SA Chatzichristofis 等人提出的颜色和边缘方向性描述符 (color and edge directivity descriptor, CEDD) 算法, 结合了图像颜色和纹理两方面信息, 将这两种信息紧凑表达为特征描述符<sup>[12]</sup>。该描述符占用空间小, 实现简单, 但没有考虑到图像的形状特征, 导致检索准确率不高。分层梯度方向直方图 (pyramid histogram of oriented gradients, PHOG) 最早由 Bosch 等提出, 是一种描述空间形状能力较强的特征向量, 但由于它仅仅包含了图像的形状特征, 故对颜色和纹理特征显著的自然背景图像检索效果不佳<sup>[13]</sup>。考虑到 CEDD 和 PHOG 的缺点, 将两者结合能提取图像的 3 种基本特征, 且 2 种算法能起到互补的作用。本文首先使用颜色和边缘的方向性描述符提取图像的颜色和纹理特征, 再采用分层梯度方向直方图来描述图像的空间形状特征, 并通过图像底层特征串行关联方法将 3 种图像特征相结合再进行图像检索, 使用 Corel 标准图像集进行测试, 结果表明本文方法的查准率和  $AvgP$  值比两种原算法均有提高。

## 2 颜色和边缘方向性描述符

首先提取图像的颜色和纹理特征。将任意大小的图像都分成 1 600 个图像块, 之后计算每个图像块的颜色和纹理特征向量, 计算完成后将所有图像块的特征向量进行合并和量化, 其基本框图如图 1 所示。

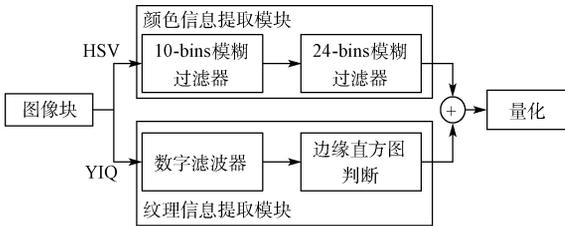
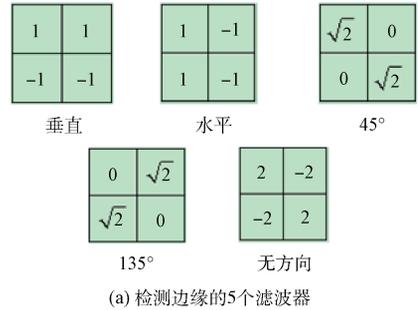


图 1 颜色和纹理特征提取步骤

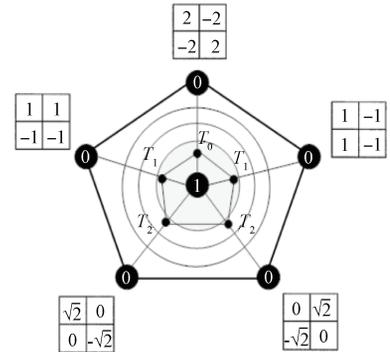
在颜色信息提取模块, 首先将图像块的颜色空间从 RGB 空间转到 HSV 空间, 再经过一个两级模糊过滤器, 由 10-bins 模糊过滤器和 24-bins 模糊过滤器串联组成。第一级模糊过滤器以 HSV 值作为 3 个信道的输入, 过滤后产生 10 维的模糊直方图。第二级模糊过滤器以 10 维模糊直方图和 H, V 值为输入, 这一级过滤器的目的主要是为了给 7 维彩色信息中加上亮度值, 所以在滤波处理中, 10 维直方图中的前 3 维并不参与计算。经过第二级模糊过滤后, 可以将 10-bins 模糊过滤器输出的每种色区再分为 3 个 H 值区域, 输出 1 个 24 维向量, 构成图像块的颜色信息。

在纹理信息提取模块, 先将图像块的颜色空间转到 YIQ 空间, 提取出图像的亮度值, 然后采用 MPEG-7 标准中边缘直方图描述符提取纹理特征<sup>[14]</sup>。每个图像块输入后, 都被分成 4 个子图像块, 每个子图像块的边缘分为 5

类, 包括 4 个有方向边缘 (分别是垂直方向、水平方向、45° 方向、135° 方向) 和 1 个无方向边缘, 由图 2(A) 所示的 5 个数字滤波器计算。提取出的纹理信息是 1 个 6 维的直方图, 各维的含义分别是: (0) 无边缘信息; (1) 无方向的边缘信息; (2) 水平方向的边缘信息; (3) 垂直方向的边缘信息; (4) 45° 方向的边缘信息; (5) 135° 方向的边缘信息。设定阈值  $T_{Edge} = 14$ ,  $T_0 = 0.68$ ,  $T_1 = T_2 = 0.68$ , 判断每个子像素块纹理信息所属直方图区域的框图如图 2(b) 所示。最终经过纹理信息提取单元的处理, 得到 1 个 6 维向量, 构成图像块的纹理特征。



(a) 检测边缘的 5 个滤波器



(b) 边缘方向阈值判断  
图 2 纹理信息计算

通过颜色信息提取模块和纹理信息提取模块的处理, 每个图像块上提取得到的所有信息被分为 6 个代表不同纹理的区域, 其中每个区域中都有 24 维的颜色信息, 也就是得到了 144 维的直方图, 如图 3 所示。1 600 个图像块中的每个图像块均有 144 维直方图特征, 构成整个图像的颜色和纹理特征。为了限制特征描述符的长度, 参考文献<sup>[12]</sup>对特征向量进行了 3 bit 的量化, 经过量化可以将特征描述符的长度控制在  $144 \times 3 = 342$  bit。

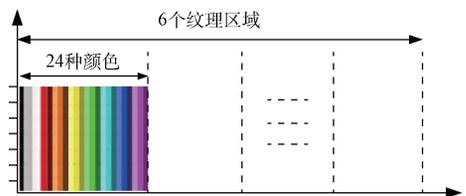


图 3 颜色和纹理特征结构

### 3 分层梯度方向直方图

接着提取图像的形状特征。先提取图像的边缘轮廓,再对图像轮廓进行逐层分割,在每一层上统计边缘图像的梯度方向直方图分布,最后按权值综合各个层中每个分块的梯度方向直方图,组成 1 个大的梯度方向直方图金字塔,代表图像的空间形状特征。本文对图像的形状特征提取步骤是:先将彩色图像转换成灰度图像,使用 Canny 算子对灰度图像进行边缘检测;对检测后的边缘图像进行金字塔分割,设图像共被分割成  $L$  层( $L$  从 0 开始,本文取  $L = 2$ ),则在第  $l$  层,图像被分成  $4^l$  块,在每一个图像块上计算轮廓点处的梯度方向和梯度模,将梯度方向单位以度来表示时,范围为  $[0^\circ, 180^\circ]$  或  $[0^\circ, 360^\circ]$ ,本文取  $[0^\circ, 180^\circ]$ ;将梯度角度分成  $K$  个区间,在各个区间上累加梯度模的值作为该区间的权值,参考文献[6]中对  $K$  不同取值时算法性能的测试,范围为  $[0^\circ, 180^\circ]$  时取  $K = 20$  算法性能较好,故本文选取  $K = 20$ ,即 20 个区间中每个区间  $9^\circ$ ;依照权值得到每个区间的梯度方向直方图,将梯度直方图合并成为形状特征,合并时需要归一化,即将第  $l$  层梯度方向直方图合并时需乘以  $\frac{1}{L4^l}$ 。最终得到的形状特征维数为  $K \sum_{l \in L} 4^l$  维,本文中为 420 维。

### 4 融合颜色、纹理和形状特征的图像检索

本文尝试融合图像的 3 种主要特征进行图像检索,先使用颜色和边缘方向性描述符提取图像的颜色和纹理特征,针对颜色和边缘方向性描述符没有考虑图像形状特征的缺点,使用分层梯度方向直方图提取图像的形状特征,并采用图像底层特征串行关联的方法融合颜色、纹理、形状特征进行检索。具体的实现步骤如下:

Step1:对图像库中所有图像  $I$ ,提取颜色和边缘方向性描述符特征和分层梯度方向直方图特征,分别存入 CEDD 特征库和 PHOG 特征库,并建立倒排索引;

Step2:输入 1 张检索图片  $I_r$ ,提取  $I_r$  的颜色和纹理特征向量,记为  $X^c$ ;

Step3:计算  $s$  与 CEDD 特征库中各特征的欧式距离,对距离进行排序,取前  $P$  个数据,即取前  $P$  幅相似图像,记相似图像集合为  $S_p = \{I_{1P}, I_{2P}, \dots, I_{PP}\}$ ;

Step4:提取  $I_r$  的分层梯度方向直方图特征向量,记为  $X^s$ ;

Step5:计算  $X^s$  与图像集合  $S_p$  中各图像的 PHOG 特征向量的距离,本文在这里使用  $\chi^2$  距离。对计算后的  $\chi^2$  距离进行排序,取  $P$  个距离中的前  $Q$  个,即从  $P$  幅图像中使用 PHOG 特征过滤,取前  $Q$  幅图像,作为检索的结果。

在本文的检索系统中,CEDD 特征和 PHOG 特征占用的空间均较小,且使用特征串行关联的方法将图像的 3 种主要特征进行融合,弥补了 CEDD 没有考虑图像形状信息

的缺点,并且计算量小,灵活性大,实现了对检索结果的过滤功能。从计算步骤上来看,是一种有效且计算相对简单的检索方法。

## 5 实验分析

### 5.1 实验数据

为验证本文检索方法的可行性,选取 Corel 图像库中的 10 类图像,每类 100 幅,共 1 000 幅图像进行测试,每幅图像大小均为  $256 \times 384$  或  $386 \times 256$ ,选取的 10 个类别如图 4 所示。



图 4 测试算法所用 10 类图像示例

### 5.2 评价参数

本文采用两种检索结果评价参数,第一种是查准率,其计算方式是:查准率 = 检索结果中与  $I_r$  相关的图片数量/返回的检索结果数量;第二种参数是文献[15]中提出的准确率评价参数  $AvgP$ 。假设在某次检索过程中,检索到的前  $n$  幅图像中,有  $m$  幅图像是相关的,分别为  $I_{1m}, I_{2m}, \dots, I_{mm}$ ,在结果中排列的序号分别为  $No_{IR_1}, No_{IR_2}, \dots, No_{IR_m}$ ,有  $z$  幅图像是不相关的,分别为  $I_{1z}, I_{2z}, \dots, I_{zz}$ ,在结果中排列的序号为  $No_{IN_1}, No_{IN_2}, \dots, No_{IN_z}$ ,其中  $n = m + z$ ,定义参数:

$$AvgP = \frac{1}{2} + \frac{1}{n(n+1)} \left[ \sum_{i=1}^m (n+1 - No_{IR_i}) - \sum_{i=1}^z (n+1 - No_{IN_i}) \right] \quad (1)$$

当结果中的前  $n$  幅图像全部相关时,  $AvgP$  取得最大值 1;相反  $AvgP$  取得最大值 0。根据  $AvgP$  的计算方式可知,其不仅考虑了检索过程的准确性,并考虑了检索结果中图片位置的因素,位置越前,其量化值越高,从而更全面地评价检索算法。

### 5.3 实验结果

实验选取的 10 类图像中,每类图像有各自的特点,使

用同样的算法检索不同类别图片时准确率差距大,故下面的检索准确率统计按类进行。为了保证实验的公平性,在每类图像中随机选取 1 幅图像进行检索,并在每类图像中将实验重复 100 次,实验结果评价参数取 100 次实验结果的平均值。所有检索过程均在 WIN7 平台下进行,硬件配置是:3.2 GHz 双核 CPU,4 GB 内存。

$P = 40, Q = 21$  时,使用颜色和方向性描述符的检索结果和使用本文方法融合形状特征对返回图片进行过滤后的检索结果如图 5 所示。其中前者的查准率为 87.50%,  $AvgP$  值为 0.919 5,而过滤后查准率为 95.24%,  $AvgP$  值为 0.995 7。可以看出过滤后的图像明显与原图形状更相似,准确率和  $AvgP$  值也有明显提高。

根据第三部分的介绍,实验中, $P, Q$  为返回图片的数量,当返回数量减少时,查准率和  $AvgP$  值会相应提高,所以为了公平,在下面评价 CEDD、PHOG、本文算法这 3 种算法的检索结果时,前两种算法的图片返回数量取  $Q$ ,本文算法过滤前和过滤后返回数量分别取  $P$  和  $Q$ 。本文选取了 3 组  $P, Q$  值,3 种算法的平均检索指标对比如表 1、表 2 所示。

从表 1 表 2 中可以看出,在一些形状特征比较明显的图像类别,如花朵、恐龙、公交车等,PHOG 的查准率和  $AvgP$  值要优于 CEDD,但在一些形状特征不明显的图像类别,如海滩、食物等,这些图像的颜色特征和纹理特征比较显著,所以使用 CEDD 检索能得到较好的结果。在 3 组不同的  $P, Q$  参数下,在非洲、海滩、食物 3 个形状特征不明显的图像类别中,本文算法的查准率和  $AvgP$  值与 CEDD



(a) 颜色和方向性描述符检索结果



(b) 融合形状特征过滤后结果

图 5 图像检索结果

表 1 不同参数时 3 种算法查准率对比 (%)

类别	$P=60, Q=40$			$P=50, Q=35$			$P=40, Q=21$		
	CEDD	PHOG	本文	CEDD	PHOG	本文	CEDD	PHOG	本文
花朵	59.58	61.42	74.37	60.16	65.25	76.42	63.79	69.58	79.98
非洲	57.91	52.92	57.59	60.45	56.92	61.27	58.38	58.92	59.13
沙滩	45.52	41.53	45.37	46.61	43.54	46.12	47.75	45.50	48.51
恐龙	93.41	94.42	94.90	95.55	95.28	95.73	96.48	96.42	97.29
食物	42.91	34.23	42.45	44.42	39.25	44.34	45.02	41.92	45.46
马	81.65	84.27	85.25	84.65	86.22	87.26	88.65	87.68	91.54
山峰	44.52	43.92	44.99	47.02	46.52	47.25	49.76	47.52	52.23
建筑	53.65	51.98	54.26	57.23	54.65	58.55	58.60	56.65	59.53
大象	39.67	41.54	43.51	41.47	43.79	49.76	43.67	45.67	51.96
公交车	72.32	83.48	85.25	80.32	85.32	87.24	84.38	89.05	92.33

表 2 不同参数时 3 种算法的  $AvgP$  值对比

类别	$P=60, Q=40$			$P=50, Q=35$			$P=40, Q=21$		
	CEDD	PHOG	本文	CEDD	PHOG	本文	CEDD	PHOG	本文
花朵	0.56	0.62	0.73	0.60	0.66	0.77	0.64	0.72	0.80
非洲	0.51	0.51	0.55	0.54	0.52	0.57	0.54	0.53	0.58
沙滩	0.46	0.41	0.46	0.43	0.44	0.48	0.48	0.44	0.51

续上表

类别	$P=60, Q=40$			$P=50, Q=35$			$P=40, Q=21$		
	CEDD	PHOG	本文	CEDD	PHOG	本文	CEDD	PHOG	本文
恐龙	0.94	0.95	0.95	0.95	0.95	0.96	0.96	0.97	0.97
食物	0.41	0.31	0.40	0.43	0.32	0.41	0.44	0.38	0.50
马	0.81	0.85	0.82	0.83	0.85	0.87	0.84	0.87	0.90
山峰	0.42	0.41	0.43	0.44	0.42	0.46	0.47	0.45	0.48
建筑	0.51	0.59	0.63	0.53	0.62	0.65	0.58	0.65	0.66
大象	0.41	0.44	0.44	0.42	0.47	0.47	0.44	0.49	0.52
公交车	0.73	0.85	0.85	0.76	0.86	0.88	0.79	0.90	0.93

不相上下,除这 3 类图像以外本文的算法的查准率和  $AvgP$  值比 CEDD 和 PHOG 的检索结果均有提高,证明本文算法是一种可行、有效、准确率较好的检索算法。

## 6 结 论

本文提出了 1 种颜色、纹理、形状 3 种特征结合的图像检索方法,先使用颜色和边缘方向性描述符提取图像的颜色和纹理特征,再使用分层梯度方向直方图提取图像的形状特征,并使用底层特征串行关联的方法将 3 种特征融合进行检索。使用查准率和  $AvgP$  值这两种评价参数对两种原算法和本文算法的检索结果进行比较,证明本文方法的查准率和  $AvgP$  值在大部分图像类别中均优于两种原算法,是一种可行有效的检索方法。

## 参考文献

- [1] DAHANE G M, VISHWAKARMA S. Content based image retrieval system[J]. IJEIT, 2012, 1(5): 92-96.
- [2] 周美丽, 白宗文. 基于形状特征的图像检索系统的设计[J]. 国外电子测量技术, 2015(6): 82-84.
- [3] 施智平, 李清勇, 赵晓东, 等. 基于内容图像检索中的优化鉴别特征[J]. 计算机辅助设计与图形学学报, 2012, 24(12): 1592-1598.
- [4] 顾晓东, 杨诚. 新的颜色相似度衡量方法在图像检索中的应用[J]. 仪器仪表学报, 2014, 35(10): 2286-2292.
- [5] JACOB I J, SRINIVASAGAN K G, JAYAPRIYA K. Local opponent color texture pattern for image retrieval system [J]. Pattern Recognition Letters, 2014, 42(6):72-78.
- [6] 张楚金, 王耀南, 卢笑, 等. 基于假设验证和改进 HOG 特征的前车检测算法[J]. 电子测量与仪器学报, 2015(2): 165-171.
- [7] 张永库, 李云峰, 孙劲光. 基于多特征融合的图像检索[J]. 计算机应用, 2015, 35(2): 495-498.
- [8] 傅启明, 刘全, 王晓燕, 等. 遗传反馈的多特征图像检索 [J]. 中国图象图形学报, 2011, 16(10): 1858-1865.

- [9] 胡敏, 罗珣, 马韵洁. 基于语义矩阵反馈的多特征融合三维模型检索方法[J]. 电子测量与仪器学报, 2012, 26(4):325-330.
- [10] JAYANTHI K, KARTHIKEYAN M. Efficient fuzzy color and texture feature extraction technique for content based image retrieval system[C]. 2014 IEEE International Conference on Computational Intelligence and Computing Research (ICIC), IEEE, 2014.
- [11] LAI C C, CHEN Y C. A user-oriented image retrieval system based on interactive genetic algorithm [J]. IEEE Transactions on Instrumentation and Measurement, 2011, 60(10): 3318-3325.
- [12] BAMPIS L, IAKOVIDOU C, CHATZICHRISTOFIS S A, et al. Real-time indexing for large image databases: color and edge directivity descriptor on GPU[J]. Journal of Supercomputing, 2015, 71(3): 1-29.
- [13] 袁杰, 魏宝刚, 王李冬. 一种综合 PHOG 形状和小波金字塔能量分布特征的图像检索方法[J]. 电子学报, 2011, 39(9):2114-2119.
- [14] IAKOVIDOU C, ANAGNOSTOPOULOS N, KAPOUTSIS A C, et al. Searching images with MPEG-7 (& MPEG-7-like) powered localized descriptors: The simple answer to effective content based image retrieval[C]. 2014 12th International Workshop on Content-Based Multimedia Indexing (CBMI), IEEE, 2014:1-6.
- [15] 徐慧英, 袁杰, 赵建民, 等. 一种基于颜色和纹理的图像检索方法[J]. 计算机科学, 2009, 36(5): 282-286.

## 作者简介

金铭, 1992 年出生, 在读硕士研究生, 主要研究方向为图像处理与分析。

汪友生, 1965 年出生, 博士, 副教授, 主要研究方向为图像处理与分析。

E-mail: wangyousheng@bjut.edu.cn