

# 基于颜色体积直方图的图像检索

刘广海 吴璟莉

(广西师范大学计算机科学与信息工程学院 桂林 541004)

**摘要** HSV 颜色空间能够较好地模拟人类视觉系统感知颜色的模式。为了更好地利用 HSV 颜色空间的优点,提出了一个新型的特征描述子(称之为颜色体积直方图),并且将其应用于图像检索。它首先把彩色图像从 RGB 空间转换为 HSV 颜色空间,然后把彩色图像均匀量化为 72 种颜色,最后利用颜色体积来描述图像特征。该算法在具有 15000 张彩色图像的 Corel 图库上进行了实验。结果表明,提出的彩色体积直方图能够结合颜色和空间特征,检索性能明显高于颜色直方图和局部二值模式直方图。

**关键词** 图像检索, HSV 颜色空间, 颜色量化, 颜色体积直方图

**中图分类号** TP391 **文献标识码** A

## Image Retrieval Based on Color Volume Histogram

LIU Guang-hai WU Jing-li

(The College of Computer Science and Information Technology, Guangxi Normal University, Guilin 541004, China)

**Abstract** HSV color space could mimic human color perception well. In order to take the advantage of HSV color space, a new feature descriptor, namely color volume histogram, was proposed in this paper for image representation and image retrieval. First, it converts color image from RGB color space to HSV color space, and then uniformly quantizes the color image into 72 colors. Finally, color volumes are used to represent image features. The proposed algorithm is extensively tested on Corel datasets with 15000 natural images. Image retrieval experimental results show that the color volume histogram has the discrimination power of color and spatial features, and the performance is better than that of the local binary pattern histogram and color histogram significantly.

**Keywords** Image retrieval, HSV color space, Color quantization, Color volume histogram

## 1 前言

颜色是一种很重要的视觉属性,它能够提供更丰富的信息来表达图像内容,并且广泛应用于模式识别、基于内容的图像检索、计算机视觉和彩色图像压缩等领域。颜色特征对图像的背景、尺度变化、方向等具有很强的鲁棒性<sup>[1,2]</sup>。常见的颜色特征描述方法主要有颜色直方图、颜色矩、颜色集、颜色聚合向量和颜色相关图<sup>[2-6]</sup>。MPEG-7 标准推荐的颜色描述子有主颜色、颜色结构、可伸缩颜色和颜色布局<sup>[6]</sup>。颜色直方图能够描述不同颜色在整幅图像中所出现的次数和频率。它因为具有提取简单的优点,所以被广泛地应用于许多图像检索系统中。颜色矩是一个非常简单而有效的颜色特征,并且图像中任何颜色分布均可以用它的矩来表示。由于颜色分布信息主要集中在低阶矩中,因此只采用一阶矩、二阶矩和三阶矩就可以表达图像颜色分布<sup>[2,3]</sup>。颜色集近似于颜色直方图方法,它可以表达为二进制的特征向量,可以构造二叉树来加快检索速度,非常适合对大规模图像库进行图像检索<sup>[2,3]</sup>。颜色聚合向量是颜色直方图的一种演变,它包含了颜色分布的空间信息。Huang 等人提出的颜色相关图能够表达不同颜

色分布的空间相关性。实验证明,它比颜色直方图和颜色聚合向量具有更高的检索效率<sup>[5]</sup>。在图像检索系统中,颜色特征经常和其它低级视觉特征相结合,以提高检索性能。例如,颜色特征可以与纹理、边缘特征以及空间属性相结合<sup>[8-10]</sup>。

颜色特征与人类视觉感知有着密切关系。在均匀颜色空间中,任意两点之间的颜色差别与人类视觉感知具有相似性,但如何利用均匀颜色属性来描述图像特征仍然是一个难题。为此,本文提出了一个新型的特征描述子(称之为颜色体积直方图),并且将其应用于图像检索。实验结果证明,本文提出的颜色体积直方图不仅能够描述颜色特征,还能够表达颜色的空间属性。

## 2 颜色体积直方图

人类眼睛对颜色的敏感度非常高,所以颜色特征非常适合于描述彩色图像内容。图像颜色特征的表达涉及到颜色空间和颜色量化的问题<sup>[1]</sup>。颜色特征的提取与选择颜色空间有着密切的关系。颜色空间与人类视觉系统密切相关,常见的颜色空间有 RGB, Lab, Luv, YUV, HSV, YIQ, YCbCr 等等<sup>[1,2]</sup>。RGB 颜色空间是一个不均匀的空间,它与人类视觉

到稿日期:2011-08-26 返修日期:2011-11-03 本文受广西自然科学基金项目(2011GXNSFB018070,2011GXNSFB018068)资助。

刘广海(1977-),男,博士,副教授,主要研究方向为模式识别和图像分析等,E-mail:liuguanghai009@163.com;吴璟莉(1978-),女,博士,副教授,主要研究方向为生物信息学等。

系统没有直接联系,不符合人类视觉感知特性<sup>[1]</sup>。HSV 颜色空间是一种基于色调、亮度和饱和度的颜色模型,并且 HSV 颜色空间色彩与人类的视觉感知基本一致,它与 RGB 颜色空间的转换是线性的,其简单快速,所以广泛应用于基于内容的图像检索<sup>[2]</sup>。

本文根据 HSV 颜色空间的属性提出了一种新型的颜色特征描述方法,它采用颜色体积信息来描述图像内容,能够体现颜色特征和空间特征。

## 2.1 HSV 颜色空间

HSV 颜色空间可以用圆柱体模型来解释。色调( $H$ )由颜色名称来辨别,对应于颜色轮上的角度( $0^\circ \sim 360^\circ$ )。饱和度( $S$ )是指颜色的深浅,用半径表示。亮度( $V$ )是指颜色的明暗程度,可用连接黑色点和白色点的中轴线上的某段距离来表示<sup>[1,2]</sup>。

从 RGB 空间转换到 HSV 空间的变换是线性的。为了获得颜色体积直方图,本文首先对彩色图像进行颜色空间变换。RGB 空间转换为 HSV 空间的过程如下所示<sup>[1]</sup>:

$$S = \begin{cases} \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)}, & \max(R,G,B) > 0 \\ 0, & \text{其他} \end{cases} \quad (1)$$

$$V = \frac{\max(R,G,B)}{255.0} \quad (2)$$

当  $R=G=B$  时,是非彩色像素,饱和度  $S=0$ ,色调无定义。为了方便计算色调  $H$ ,首先将 RGB 3 个分量进行归一化:

$$R' = \frac{\max(R,G,B) - R}{\max(R,G,B) - \min(R,G,B)} \quad (3)$$

$$G' = \frac{\max(R,G,B) - G}{\max(R,G,B) - \min(R,G,B)} \quad (4)$$

$$B' = \frac{\max(R,G,B) - B}{\max(R,G,B) - \min(R,G,B)} \quad (5)$$

则可以得到初始的色调值  $H'$ :

$$H' = \begin{cases} B' - G', & R = \max(R,G,B) \\ R' - B' + 2, & G = \max(R,G,B) \\ G' - R' + 4, & B = \max(R,G,B) \end{cases} \quad (6)$$

因为  $H' \in [-1,5]$ ,将其归一化到区间  $[0,1]$ ,则最终的色调值为:

$$H = 60 \cdot \begin{cases} H' + 6, & H' < 0 \\ H', & \text{其他} \end{cases} \quad (7)$$

最后,HSV 3 个分量的取值范围是  $H \in [0,360]$ ,  $S \in [0,1]$  和  $V \in [0,1]$ 。

## 2.2 颜色量化

颜色是一种非常重要的视觉属性,它能够提供更丰富的信息来表达图像内容。颜色特征提取与颜色空间有着密切关系。常见的颜色空间有 RGB, Lab, Luv, HSV, YIQ 和 YCbCr 等<sup>[1,4]</sup>。颜色信息是组成图像内容的基本要素,在模式识别和图像检索中,甚至在没有形状信息的情况下,仍然可以取得较好的效果<sup>[1-3]</sup>。人眼对颜色的敏感度非常高,为了描述颜色特征,以及简化操作,本文选择 HSV 颜色空间来进行颜色量化。在 HSV 颜色空间,一幅大小为  $M \times N$  的彩色图像,设  $\text{bin}(H)$ ,  $\text{bin}(S)$  和  $\text{bin}(V)$  分别表示  $H$ ,  $S$  和  $V$  分量的颜色量化数。因为人眼对颜色的辨析能力有限,按照人的视觉感知属性,本文将彩色图像量化为  $8 \times 3 \times 3 = 72$  种颜色,即

$\text{bin}(H)=8$ ,  $\text{bin}(S)=3$  和  $\text{bin}(V)=3$ ,则 72 种颜色的索引图像可表示为  $C(x,y)$ ,其中  $x \in [0,1,\dots,M-1]$ ,  $y \in [0,1,\dots,N-1]$ 。

## 2.3 特征描述

直方图是描述颜色特征最常用的方法之一,它具有高效的索引性能,并且计算简单。如果直接利用直方图来描述特征,则无法表达图像对象的空间关系。为了克服颜色直方图无法描述空间关系的缺点,同时发挥其优点,本文提出通过颜色体积直方图(CVH)来描述图像内容,将颜色和空间信息考虑在内,并且将它们结合为一个整体。

假设颜色索引图像  $C(x,y)$  的值  $C(p)=w$ ,  $w \in [0,1,\dots,W-1]$ ,  $P=(x,y)$  为像素点位置,  $W$  为颜色量化总数目。假设从图像左上角开始,取一个  $3 \times 3$  的窗口,它的中心点位置为  $P_0=(x_0,y_0)$ ,周围八邻域像素的位置为  $P_i=(x_i,y_i)$ ,  $i \in [1,2,\dots,8]$ 。假设中心点像素的颜色值为  $(H_0,S_0,V_0)$ ,周围八邻域像素的颜色值为  $(H_i,S_i,V_i)$ ,  $i \in [1,2,\dots,8]$ 。如果从左到右,从上到下,不断移动  $3 \times 3$  窗口,移动步长为 1 个像素,则颜色体积直方图(CVH)可以定义为:

$$H(C(x_0,y_0)) = \sum_{x=1}^{M-1} \sum_{y=1}^{N-1} \left\{ \sum_{i=0}^8 \left( 1 - \frac{CV_i}{9MN} \right) \right\} \quad (8)$$

式中,  $CV_i$  为像素点  $P_i=(x_i,y_i)$  在 HSV 圆柱体模型中所占的体积,  $CV_i = |\pi \cdot S_i^2 \cdot V_i \cdot H_i / 360|$ 。图 1 所示为房子和邮票图像的颜色体积直方图示例。

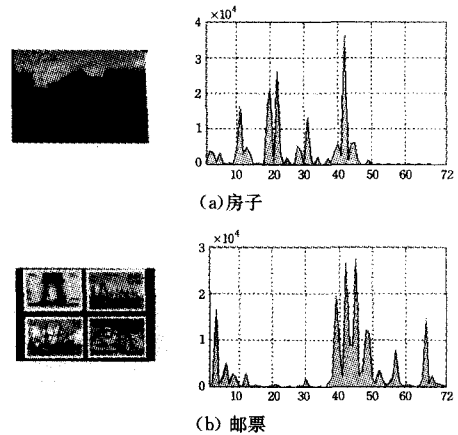


图 1 颜色体积直方图(CVH)的示例

HSV 圆柱体模型中的任意两个颜色点,它们之间的颜色差别可以采用这样的方式来度量,即首先将圆柱体模型转换为笛卡尔坐标系,然后采用欧几里德距离来度量,但是欧几里德距离无法说明两个颜色点分别属于什么颜色。

采用颜色体积来表达颜色特征,在很大程度上,可以区分两个颜色点的归属。因为不同的颜色,它们具有不同颜色体积的可能性比较大。此外,在颜色量化环节中,已经对颜色进行了一次划分,相当于进行了一次颜色归属的区分。在图像处理过程中,任何颜色量化算法都会出现颜色误差。如果在颜色直方图构建中,采用  $3 \times 3$  图像区域中的某种颜色体积没有出现的比例来代替颜色出现的次数或者频率,可以大幅度降低颜色误差。因此,颜色体积直方图根据 HSV 圆柱体模型的体积信息来描述图像内容,不仅包含了颜色特征,还具有空间特征描述能力。

## 3 实验与性能分析

为了验证本文算法的检索效果,分别采用颜色体积直方

图(CVH)、局部二值模式直方图(LBPH)<sup>[7]</sup>和颜色直方图(CH)来进行对比实验。三者均属于采用直方图来描述图像特征的算法,其中LBPH的特征维数为256维,在对比实验中,它分别提取R,G和B3个分量的特征,然后采用它们的平均值作为最终特征向量。CH采用RGB颜色空间进行颜色量化,它的特征维数为64维,CVH的特征维数为72维。在两个实验中,每个图像类别分别选择50张图像作为例子图像,通过计算平均检索精确度和查全率作为实验结果。

本文实验系统采用两个图像库,它们由15000张Corel图像组成。第一个图像库是Corel-5000图像库。它包含5000张图像,主要有热气球、大象、花卉、土著人、海滩、建筑物、公交车、马和海浪等50类图像。每类图像的数目为100张,大小为192×128或者128×192,图像格式为jpg。第二个图像库是Corel-10000图像库。它包含10000张图像,主要有飞机、花卉、埃及风光、海滩、建筑物、落日、桥梁、马、食物和油画等100类图像。每类图像的数目为100张,大小为192×128或者128×192,图像格式为jpg。

采用彩色体积直方图来提取特征,每一幅图像可以得到72维特征向量 $T=[T_1, T_2, \dots, T_{72}]$ 应用于基于内容的图像检索。假设例子图像的特征为 $Q=[Q_1, Q_2, \dots, Q_{72}]$ ,图像之间的相似程度可采用L1距离来计算。因为L1距离计算简单,并且计算量和复杂度都非常小,非常适合于大规模图像检索。L1距离可以表示为:

$$D(i, j) = \sum_{i=1}^{72} |T_i - Q_i| \quad (9)$$

本文采用精确度和返回率来评价图像检索性能。精确度就是相似图像数目与返回图像数目的比率。查全率就是相似图像数目与图像库中相似图像总数的比率<sup>[8-10]</sup>。它们的定义如下所示:

$$P(N) = I_N / N \quad (10)$$

$$R(N) = I_N / M \quad (11)$$

式中, $I_N$ 为界面上相似图像的数目, $M$ 为图像库中相似图像总数, $N$ 是界面返回的图像数目。在本文图像检索系统中,界面返回的图像数目 $N=12$ ,图像库中相似图像总数 $M=100$ 。

在图像检索实验中,CH采用直方图相交匹配方法,LBPH采用L1距离。从图2的检索性能曲线图可以看出,本文所提出的CVH的检索性能明显地高于CH和LBPH方法。

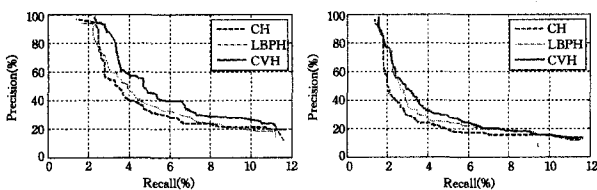


图2 检索性能曲线

表1的实验数据显示了CVH在颜色量化数目为54,72,108和192条件下的检索性能。根据表1的数据来看,随着颜色量化数目的增加,检索性能也得到相应提高,但计算复杂度基本保持不变。在图像检索实验中,检索性能不仅仅与精确度密切相关,而且与特征向量维数相关。特征维数过高,势必耗费更多的检索时间和计算量。为了平衡检索性能、检索时间和计算量,本文所提出的CVH采用72维特征向量来进行图像检索。

表1 CVH在不同颜色数量下的检索性能(Corel-5000库)

颜色总数量	检索性能	
	精确度	返回率
54	48.11%	5.77%
72	49.63%	5.96%
108	50.90%	6.11%
192	53.07%	6.37%

根据表2的检索实验数据来看,在Corel-5000图像库中,CVH方法的检索精确度分别比CH和LBPH方法提高了6.28%和5.81%。在Corel-10000图像库中,CVH方法的检索精确度分别比CH和LBPH方法提高了5.49%和2.25%。LBPH主要表达了局部结构信息的空间分布,CH方法则表达了颜色信息出现的次数或者频率。CVH表达了感知颜色的空间分布,即表达了空间信息,同时也表达了颜色特征,所以CVH的检索精确度均高于LBPH和CH方法。

表2 3种方法的检索性能数据

图像库	性能	方法		
		CH	LBPH	CVH
Corel-5000	精确度	43.35%	43.82%	49.63%
	返回率	5.21%	5.26%	5.96%
Corel-10000	精确度	33.99%	37.23%	39.48%
	返回率	4.08%	4.47%	4.74%

根据Corel-5000图像库和Corel-10000图像库的实验数据来看,LBPH作为著名的纹理特征描述子,同样是针对彩色图像,CVH的检索精确度却比LBPH分别提高了5.81%和2.25%,说明CVH具有描述彩色纹理特征的能力,同时也反映出颜色特征在自然图像检索中能够体现出良好性能。一般情况下,基于颜色特征的图像检索性能要优越于纹理特征,甚至优于基于灰度模式扩展为彩色模式的算法。例如,LBPH就是一种灰度模式扩展为彩色模式的算法,它分别在RGB3个分量上提取特征,最后用它们的平均值作为最终特征向量。CVH和LBPH最为显著的区别是:LBPH的特征维数为256维,而CVH仅仅只有72维特征向量。

图3和图4分别显示了CVH方法在Corel-10000库中的两个检索示例:(a)例子图像为一幅房子图像。图3所返回的绝大部分图像均为相似图像。CVH方法检索出来的房子图像基本上具有相似的颜色分布属性和结构相似性。

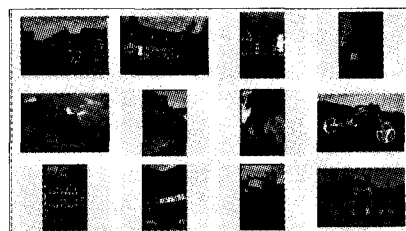


图3 房子图像的检索示例(Corel-5000图像库)

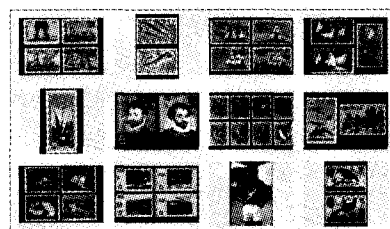


图4 邮票图像的检索示例(Corel-10000图像库)

间甚少,对整个算法的运行时间增加甚少。从以上分析可以看出,在本文提出的快速算法中,目标函数运算次数、加法、乘法和取对数运算次数都大量减少,所以分割时间少。总之,本文提出的快速算法是有效的。

**结束语** 1) 矩不变法无须迭代、运行速度快,且获取的阈值反映了图像信息的整体情况,对其调整二维 Shannon 熵斜分法的阈值是可行的;2) 为了克服单独使用一种二维 Shannon 熵斜分法的不足,提出了一种使用矩不变法来调整二维直方图斜分 Shannon 熵的阈值法,该方法是有用的,提高了二维 Shannon 熵斜分法的分割性能;3) 本文成功构造出二维 Shannon 熵斜分法的一般递推算法,并将二维直方图的特性与这种算法有机结合,得到一种新型快速递推算法,更进一步提高了分割速度,这种算法具有一定的推广价值。

## 参考文献

- [1] 张新明,郑延斌,张慧云. 应用混沌多目标规划理论融合的图像分割[J]. 小型微型计算机系统, 2010, 31(7): 1416-1420
- [2] Kapur J N, Sahoo P K, Wong A K C. A new method for grey-level picture thresholding using the entropy of the histogram [J]. Computer Vision, Graphics and Image Processing, 1985, 29(3): 273-285
- [3] Sahoo P K, Slaaf D W, Albert T A. Threshold selection using a

minimal histogram entropy difference [J]. Optical, Engineering, 1997, 36(7): 1976-1981

- [4] Abutaleb A S. Automatic thresholding of gray-level pictures using two-dimensional entropies [J]. Pattern Recognition, 1989, 47(1): 22-32
- [5] Brink A D. Thresholding of digital images using two-dimensional entropies [J]. Pattern Recognition, 1992, 25(8): 803-808
- [6] Chen W T, Wen C H, Yang C W. A fast two-dimensional entropic thresholding algorithm [J]. Pattern Recognition, 1994, 27(7): 885-893
- [7] Gong J, Li L Y, Chen W N. Fast recursive algorithm for two-dimensional thresholding [J]. Pattern Recognition, 1998, 31(3): 295-300
- [8] 吴一全,潘喆,吴文怡. 二维直方图区域斜分的最大熵阈值分割算法[J]. 模式识别与人工智能, 2009, 22(1): 162-168
- [9] Tsai W H. Moment-preserving thresholding: A new approach [J]. Computer, Graphics and Image Processing, 1985, 29(3): 377-393
- [10] 张新明,李双,郑延斌,等. 傅里叶谱和矩不变法结合的图像阈值分割[J]. 计算机应用, 2010, 30(8): 2094-2097
- [11] 雷博,范九伦. 加权调和平均型最大熵图像阈值选取法[J]. 模式识别与人工智能, 2009, 22(6): 884-890

(上接第 275 页)

(b) 例子图像为一幅邮票图像。CVH 方法检索出来的邮票图像基本上具有相似的颜色成分以及颜色结构分布属性。图 4 返回的 12 幅图像中,有 11 幅图像为相似图像,在返回的 11 幅相似图像中,大部分邮票的颜色基本具有类似的若干种颜色,这说明 CVH 能够表达图像的颜色特征。

CVH 吸收了颜色直方图所具备的优点,同时考虑了周围像素的空间关系,取得了较好的检索性能。尽管 CVH 包含了颜色和空间特征,但它无法充分表达彩色图像的局部信息。在表达图像局部信息方面,LBPH 要优越于 CVH,所以 CVH 在部分图像类别的检索精确度并不理想。从实验对比结果以及分析中可以看出,能够体现局部结构和颜色特征的算法将会大幅度提高检索性能,这也将成为本文在下一阶段继续研究的问题。

本文实验的硬件配置为: Intel Centrino 2. 83GHZ 的 4 核 CPU, 4G 内存和 500G 硬盘,操作系统为 Windows 7 的台式机,检索系统采用 Visual C# 2010 开发。CH, LBPH 以及本文提出的 CVH 对同一幅  $192 \times 128$  的彩色图像提取特征,它们所耗费的时间分别为 34. 36/ms, 49. 24/ms 和 36. 70/ms。CVH 所耗费的时间低于 LBPH, 而与 CH 方法耗费的时间基本相似。

综合检索精确度、特征提取耗时间以及特征维数等因素来看,本文提出的 CVH 要优越于 LBPH 以及 CH。

**结束语** HSV 颜色空间能够较好地模拟人类视觉系统感知颜色的模式。本文根据 HSV 颜色空间的属性提出了一个新型颜色特征描述子(称之为颜色体积直方图),并且将其应用于图像检索。传统的颜色直方图统计颜色出现的频率或者比例,具有高效的索引性能,并且计算非常简单,但它无法描述空间关系。本文提出的颜色体积直方图吸收了传统颜色直方图的优点,并且克服了它的缺点。实验结果表明: 本文提

出的彩色体积直方图能够结合颜色特征和空间特征,检索性能优越于传统颜色直方图和局部二值模式直方图。本文将在下一步继续研究局部颜色结构的特征提取以及图像检索等方面的问题。

## 参考文献

- [1] Burger W, Burge M J. 数字图像处理 Java 语言算法描述[M]. 黄华,等译. 北京:清华大学出版社, 2009
- [2] 陆建江,张亚非,等. 智能检索技术[M]. 北京:科学出版社, 2009
- [3] 沈兰荪,张青,李晓光. 图像检索与压缩域处理技术的研究[M]. 北京:人民邮电出版社, 2008
- [4] Gonzalez R C, Woods R E. Digital Image Processing (3rd edition)[M]. Prentice Hall, 2007
- [5] Huang J, Kumar S R, Mitra M, et al. Image indexing using color correlograms [A]// IEEE Conference on Computer Vision and Pattern Recognition [C]. 1997
- [6] Manjunath B S, Ohm J-R, Vasudevan V V, et al. Color and Texture Descriptors[J]. IEEE Transactions on Circuit and Systems for Video Technology, 2001, 11(6): 703-715
- [7] Ojala T, Pietikainen M, Maenpaa T. Multi-resolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987
- [8] Liu Guang-hai, Yang Jing-yu. Image retrieval based on the tex-ton co-occurrence matrix [J]. Pattern Recognition, 2008, 41(12): 3521-3527
- [9] Liu Guang-hai, Zhang Lei, et al. Image Retrieval Based on Multi-Texton Histogram[J]. Pattern Recognition, 2010, 43(7): 2380-2389
- [10] Liu Guang-hai, Li Zuo-yong, Zhang Lei, et al. , Image retrieval based on micro-structure descriptor [J]. Pattern Recognition, 2011, 44(9): 2123-213