

一种新的基于颜色和空间特征的图像检索方法^{*}

孙君顶 张喜民 崔江涛 周利华

(西安电子科技大学 多媒体研究所 西安710071)

摘 要 提出了一种新的基于颜色和空间特征的图像检索方法。首先提取色调不变量作为图像的颜色特征,然后设计了图像状态矩阵来描述图像的形状信息和空间位置信息,在进行相似检索时,采用 Gaussian 模型对不同于特征间的距离进行了归一化处理。试验表明利用该算法提取的图像的颜色和空间特征在进行图像检索时效果显著。

关键词 基于内容的图像检索(CBIR),色调直方图,状态矩阵,状态关联图

A New Approach for Image Retrieval Based on Color and Spatial Features

SUN Jun-Ding ZHANG Xi-Min CUI Jiang-Tao ZHOU Li-Hua

(Research Inst. of Multimedia, Xidian Univ., Xi'an 710071)

Abstract A new content-based color image retrieval method is proposed in this paper, in which both the color content and the spatial features have been taken into account. Based on the special disposal on the HSV color space, the image invariable, hue is selected as the color character, and then an improved accumulative histogram of the hue is calculated as the color feature. To attain the image's shape and spatial information, the image's H-, S-, and V- component are first transformed into status matrix. Then the state correlogram of the status matrix is calculated as the image's spatial distribution information. In matching the similarity of images, the Gaussian model is used to normalize the different sub-characters distance. Experiments indicate that this method is great effective in image's similarity retrieval, and has a higher retrieval-rate.

Keywords Content-based image retrieval, Hue histogram, State matrix, State correlogram

1 引言

基于内容的图像检索(content-based image retrieval, CBIR)技术是当前信息世界研究的热点之一。随着多媒体数据库尤其是大量图像库的飞速发展,对图像检索效率的要求越来越高,因此采用合适的检索方法至关重要。根据图像特征的不同,常用的图像检索特征有颜色、纹理、形状和空间关系等。但基于图像单一特征的检索方法只能表达图像的部分属性,缺少足够的全局信息。因此,在进行图像检索中往往需要结合多种图像特征进行检索,以提高检索的效率及准确度^[1,2,11]。

本文旨在研究一种有效的基于图像颜色和空间特征的图像表示方法及相应的检索技术,即在对 HSV 颜色模型进行特殊处理的基础上,将提取的色调(Hue)作为颜色不变量,并采用改进的累加直方图方法来描述图像的颜色特征。同时,针对图像在 HSV 颜色空间的3个灰度分量,提出采用图像状态矩阵来描述图像的空间信息,并以此为基础来提取图像的空间分布信息的思路。这样就得到了一种新的基于图像颜色特征和空间分布特征的图像表示方法。在进行图像间相似性度量时,为了能结合不同的子特征进行全局的相似性检索,文中采用 Gaussian 模型对不同于特征间的距离进行了归一化处理。

2 色调直方图(hue histogram)

2.1 颜色空间与图像的颜色不变量

大家知道,对彩色图像的研究,必须在彩色图像的特定颜

色空间中进行,实际应用中常用的颜色空间有 RGB、YIQ、La*b*、HSV 等。一般认为,RGB 颜色空间与人眼的感知差别很大,而 HSV 颜色空间直接对应于人眼色彩视觉特征的3要素,在色调(Hue)、饱和度(Saturation)和亮度(Value)3个分量中,色调尤其影响着人类的视觉判断,由于 HSV 空间各轴在视觉上彼此无关,空间距离又符合人眼视觉特征,并且从 RGB 到 HSV 的转换是一个简单且快速的非线性变换,因此,本文使用 HSV 颜色空间作为彩色图像的色彩空间。

为了使 HSV 颜色模型能更好地符合人眼的特性,文中首先将图像中每一像素的 r, g, b 值转化为 h, s, v 值($h \in [0, 360], s \in [0, 1], v \in [0, 1]$),然后根据 HSV 颜色模型的特性作如下特殊处理^[1]:

(1)黑色区域:所有 $v < 0.15$ 的颜色均归入黑色,令 $h=0, s=0, v=0$;

(2)白色区域:所有 $s < 0.1$, 且 $v > 0.8$ 的颜色归入白色,令 $h=0, s=0, v=1$;

(3)彩色区域:位于黑色区域和白色区域以外的颜色,其 h, s, v 值保持不变。

这样经过图像颜色空间的变换,图像中每个像素的颜色就可由 h, s, v 值表示,同时由于将与黑色、白色相近的颜色分别作为同一种颜色对待,从而进一步提高了颜色模型的准确性。

为了获取一个独立于观察点的图像颜色描述符,需要一种颜色特征,在事先考虑到阴影、遮蔽以及亮度变化等因素的影响下,它仍能独立于对象表面的形状和观察的角度。人们已

^{*}基金项目:十五国防科技(电子)预研项目(413160501)。

经发现,人眼对于颜色的色调尤为敏感,并且颜色的色调(Hue)是一个独立于观察点的颜色特征^[3,4]。基于此,本文利用彩色图像的色调分量(Hue)作为图像的颜色不变量。

2.2 色调直方图

在颜色特征表示方面,由于颜色直方图^[5]描述了图像颜色的统计分布特征且具有平移、尺度、旋转不变性,因此在颜色检索中被广泛采用。与颜色直方图类似,如果直接将色调作为图像的颜色特征,那么就可以得到相应的色调直方图方法。由于累加直方图^[6,7]体现了两种颜色在颜色轴上的距离与相似性之间的关系,因此累加直方图法在检索效率上优于一般直方图法。为此,本文对描述图像颜色特征的色调直方图也采用了累加的方法进行处理,为了进一步提高所提取的色调直方图的准确性,还对累加直方图的方法进行了适当的改进。

具体检索时,先将色调均匀量化成18个不重叠的局部区间 $[20k, 20(k+1)]$, $k=0, 1, \dots, 17$,然后计算出此时的色调直方图;由于色调在色调轴上的分布实际上是连续过渡的,且各色调区间并不存在截然的界限,因此可再改变区间的划分为 $[20k+10, (20(k+1)+10) \bmod 360]$, $k=0, 1, \dots, 17$,并计算出此时的色调直方图,然后将这两个直方图逐项相加取平均,最后将所求取的平均色调直方图采用文[6]的方法求取其累加直方图,并将该累加直方图作为图像的颜色特征。

3 空间特征提取

由于彩色图像HSV颜色空间的3个分量图像是相互独立的灰度图像,为此本文提出通过综合这3个分量灰度图像来提取图像的空间信息,并提出一种有效的空间特征描述方法。首先将3个灰度图像分别转化为一种包含图像空间信息的状态矩阵,然后参考颜色关联图(Color Correlogram)^[8]的方法,我们提出采用状态关联图(State Correlogram)来反映状态空间中不同的状态点按照空间距离的分布特征。

3.1 图像状态矩阵的定义

为了描述图像的空间位置信息,可将图像的色调、饱和度和亮度3个灰度图像分别划分为 $n \times n$ 个小区块子块。对于某一区域子块,如果该区域内像素间的灰度值差别较小,那么在人眼看来,该区域表现得就比较平坦;如果该区域内像素间的灰度值差别较大,则该区域表现得就比较起伏,依据这种特性,就可以将图像的区域子块划分为不同的类型,而区域内像素灰度值的变化则可通过该区域内像素灰度值的标准差来反映,对于某一子块,如果其内部像素灰度值的标准差很小,则表明该子块内像素的灰度值比较接近,因此该子块对应图像中灰度变化平坦的区域;如果其内部像素灰度值的标准差很大,则表明该子块内像素的灰度值变化较大,而只有包含图像边缘细节的区域子块才具有这种特性;而介于二者之间的子块,像素间的灰度值相差则不是很大,这种子块在一定程度上反映了图像纹理特征。依据子块内像素灰度值的标准差,可将图像区域子块划分为平坦态、纹理态和边缘态3种状态。对于这3种状态定义如下:设 σ 表示子块内像素灰度值的标准差, σ 定义如下:

$$\sigma = \left(\frac{1}{l \times m} \sum_{i=1}^l \sum_{j=1}^m (f(i, j) - \mu)^2 \right)^{\frac{1}{2}} \quad (1)$$

其中, $\mu = \frac{1}{l \times m} \sum_{i=1}^l \sum_{j=1}^m f(i, j)$, $f(i, j)$ 表示坐标为 (i, j) 的像素灰度值, l, m 表示所划分子块的大小。

对于给定的阈值 $\alpha_1, \alpha_2 (\alpha_1 < \alpha_2)$,若 $\sigma < \alpha_1$,则定义该子块的状态为“平坦态”,并用“0”表示该种状态;若 $\sigma > \alpha_2$,则定义

该子块的状态为“边缘态”,并用“2”表示该种状态;若 $\alpha_1 < \sigma < \alpha_2$,则定义该子块的状态为“纹理态”,并用“1”表示该种状态。为确定参数 α_1, α_2 的取值,可采用统计的方法来进行估计,即从图像库中任意选取100幅图像,再对子块内像素灰度值标准差的分布进行统计,统计结果如图1所示,其中横坐标表示标准差的变化情况,纵坐标表示在100幅图中不同标准差取值所对应的分块总数。

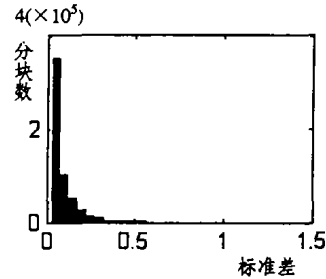


图1 分块标准差分布

由图1可以看出图像子块内像素灰度值的标准差主要分布在区间 $[0, 0.6]$,且分布在区间 $[0, 0.05]$ 的子块占了大多数,而分布在区间 $[0.6, 1]$ 的子块几乎没有,根据这种统计结果及基于图像整体特性的考虑,在试验中,取 $\alpha_1 = 0.05, \alpha_2 = 0.4$ 。

根据上述对子块类型的定义,可以将图像中的每一区域子块分别采用其状态值 $\{0, 1, 2\}$ 来表示,这种采用分块状态来表示的图像矩阵称为图像的状态矩阵。这样,与彩色图像对应的每个灰度图像均可转化为由3种状态 $\{0, 1, 2\}$ 构成的二维状态矩阵,而在状态矩阵中就包含了图像的空间信息。

3.2 状态关联图(State Correlogram)

为了提取图像的空间信息,针对图像的状态矩阵,我们采用类似Color Correlogram^[8]的方法,通过定义条件概率来描述图像状态矩阵中各个状态之间的空间关系特征,即采用状态关联图来描述图像的空间特征。

对某图像状态矩阵 T ,设其状态数记为 m, b_1, b_2, L, b_m 表示不同的状态值。对于状态矩阵中的一个状态点 $f(x, y) \in T$, $T(f)$ 表示该点的状态值。令 $T_b = \{f \in T | T(f) = b\}$,则 $f \in T_b$ 等价于 $f \in T, T(f) = b$ 。对于任意两个像素, $f_1 = (x_1, y_1), f_2 = (x_2, y_2)$,定义两状态间的空间距离为:

$$k = |f_1 - f_2| @ \max\{|x_1 - x_2|, |y_1 - y_2|\} \quad (2)$$

设 $H_{b_i}(T)$ 代表状态矩阵 T 中状态为 b_i 的状态点的总个数, $1 \leq i \leq m$,则状态矩阵 T 中状态 b_i 出现的概率可表示为:

$$h_{b_i}(T) = \Pr[f \in T_{b_i}] = \frac{H_{b_i}(T)}{n^2} \quad (3)$$

设 k 表示状态矩阵中不同状态点间的距离, $1 \leq i, j \leq m$,则状态矩阵 T 的状态关联图定义为:

$$\gamma_{b_i, b_j}^{(k)}(T) @ \Pr[f_2 \in T_{b_j}, |f_1 - f_2| = k | f_1 \in T_{b_i}] \quad (4)$$

如果仅仅考虑图像状态矩阵中相同状态间的空间关系,此时的状态关联图称为状态自关联图,其定义如下:

$$\beta_{b_i}^{(k)}(T) @ \gamma_{b_i, b_i}^{(k)}(T) \quad (5)$$

3.3 状态关联图计算

$$\Gamma_{b_i, b_j}^{(k)}(T) @ |\{f_1 \in T_{b_i}, f_2 \in T_{b_j} | |f_1 - f_2| = k\}| \quad (6)$$

上式统计了距离为 k 的状态分别为 b_i, b_j 的状态对的个数。

如果忽略边界条件,设 b_i, b_j 分别代表状态矩阵中任意两点 f_1, f_2 的状态,则:

$$\Pr(b_1=b_i, b_2=b_j, |f_1-f_2|=k) = \frac{\Gamma_{b_i, b_j}^{(k)}(T)}{n^2 \times 8k} \quad (7)$$

上式中,分母表示状态矩阵 T 中距离为 k 的状态点的总个数,分子即为式(6)的统计值。对于每一个状态点,和它距离为 k 的状态点的个数为 $8k$,由此得到条件概率:

$$\begin{aligned} \Pr(b_2=b_j, |f_1-f_2|=k | b_1=b_i) \\ = \frac{\Pr(b_1=b_i, b_2=b_j, |f_1-f_2|=k)}{\Pr(b_1=b_i)} \\ = \frac{n^2}{H_{b_i}(T)} \times \frac{\Gamma_{b_i, b_j}^{(k)}(T)}{n^2 \times 8k} = \frac{\Gamma_{b_i, b_j}^{(k)}(T)}{H_{b_i}(T) \times 8k} \end{aligned} \quad (8)$$

上式即为所要计算的 $\gamma_{b_i, b_j}^{(k)}$ 的值。

4 相似性度量

4.1 子特征的距离度量

对于计算色调直方图间的距离,可采用经典的直方图相交算子(histogram intersection)^[5]来进行距离间的度量,直方图相交算子定义为:

$$d_1(V, W) = 1 - \frac{\sum_{L=1}^{18} \min(v_L, w_L)}{\sum_{L=1}^{18} w_L} \quad (9)$$

其中, $V = \{v_1, v_2, L, v_{18}\}$ 和 $W = \{w_1, w_2, L, w_{18}\}$ 分别表示两幅图像的色调直方图,若两幅图像的色调分布相同,则它们之间的距离为0(相似度为1);否则,为0和1之间的一个数。

在采用上述方法提取图像状态矩阵的特征后,我们采用 Jeffrey 距离进行空间特征间的相似性度量:

$$\begin{aligned} d_J(T, T') = \\ \sum_{1 \leq i, j \leq n, k \in D} \left\{ \gamma_{b_i, b_j}^{(k)}(T) \log \left[\frac{2\gamma_{b_i, b_j}^{(k)}(T)}{\gamma_{b_i, b_j}^{(k)}(T) + \gamma_{b_i, b_j}^{(k)}(T')} \right] \right. \\ \left. + \gamma_{b_i, b_j}^{(k)}(T') \log \left[\frac{2\gamma_{b_i, b_j}^{(k)}(T')}{\gamma_{b_i, b_j}^{(k)}(T) + \gamma_{b_i, b_j}^{(k)}(T')} \right] \right\} \end{aligned} \quad (10)$$

这里选用3个像素间距, $D = \{1, 3, 5\}$, $k \in D$ 。

4.2 颜色与空间特征的组合检索

由于不同的子特征会采用不同的度量方式,因此为了不同的子特征具有可比性,需要对子特征距离进行归一化处理。本文采用 Gaussian 模型对所计算出的距离进行归一化处理^[9]。为了确定 Gaussian 模型的参数,本文提出采用训练集的方法来确定这些参数的取值,即随机地从图像库中抽取若干对图像组成训练集,并假定选取的训练集足够大,且具有多样性,还有训练集中图像的分布特征足以能描述整个图像集的分布特征,这样,当向图像集中加入其他新图像时,就不会对整个图像库的分布特征产生较大的影响;然后对于每个子特征分别计算训练集中所有图像对间的距离,在此基础上再计算这些距离的均值和标准差。

设子特征 M 的距离度量为 d_M ,相应的距离均值和标准差的训练值分别为 μ_M 和 σ_M ,在 Gaussian 模型的假设下, d_M 相应的归一化距离度量 S_{d_M} 定义为

$$S_{d_M}(Z_1, Z_2) = \frac{d_M(Z_1, Z_2) - (\mu_M - 3\sigma_M)}{6\sigma_M} \quad (11)$$

其中, Z_1 和 Z_2 是任意两幅图像对应子特征 M 的特征表示。上式可保证 S_{d_M} 中99%的取值落入 $[0, 1]$ 区间,并且可以同其他的归一化距离度量进行线性组合。

假设例子图像为 Q (query image), G 为图像库中任意一目标图像,在实际的检索处理中,可先分别利用式(9)和式(10)计算出图像子特征之间的距离,利用式(11)对所计算的

距离进行归一化处理,然后,就可计算例子图像与数据库图像之间的全局相似度。图像间的全局距离度量定义为:

$$d(Q, G) = \sum_{M=1}^2 \omega_M S_{d_M}(q_M, g_M) \quad (12)$$

其中, q_M 和 g_M 分别为图像 Q 和 G 的第 M 个子特征的特征表示, $S_{d_M}(q_M, g_M)$ 表示子特征 M 归一化后的距离, ω_M 为子特征 M 的权重,用来控制图像颜色特征和细节特征在检索中的相对重要性,并且满足 $\omega_1 + \omega_2 = 1$, $\omega_1, \omega_2 \in [0, 1]$ 。权值可以通过试验确定,也可由用户指定。

5 试验结果

为了检验本文方法的性能,采用的图像库包含2200幅图像,它们包括动物、建筑、自然景物等。实验中,图像区域子块尺寸取为 8×6 ,在利用图像的状态矩阵提取其空间特征时,采用状态自关联图(式5)进行计算,即只考虑相同状态的空间关系,在利用式(12)进行图像间的距离度量时,取 $\omega_1 = \omega_2 = 0.5$,这样就赋予颜色特征和空间特征同等重要性。

为了评价本文提出的方法的效果,采用“精确度(precision)”和“检索率(recall)”^[10]作为相似检索的评价准则。精确度 p 定义为检索结果队列中检索到的目标图像数与检索结果队列中所有的图像数之比,即

$$p = Y/X \quad (13)$$

检索率 r 定义为检索结果队列中检索到的目标图像数与数据库中全部的目标图像数之比,即

$$r = Y/F \quad (14)$$

这里 X 代表检索结果返回的图像总数, Y 代表查询结果中与例子图像相关的目标图像数, F 表示图像库中与例子图像相关的目标图像总数。由精确度和检索率的定义可知, $p \in [0, 1]$, $r \in [0, 1]$ 。通过执行例子集合中的各个查询来计算出查询的平均检索精度和查全率,据此就能给出系统的检索性能评价。

为了检验本文方法的性能,我们将本文方法同直方图方法^[6]和基于颜色和形状的方法(CSH)^[11]进行比较。试验时先在图像库中选取5类图像,且在每类图像中抽取4幅图像共组成20次查询,然后结合这20次查询,计算出在不同检索率条件下系统的平均精确度,三种方法的检索率与精确度变化曲线如图2所示,由该图可以看出本文的方法明显优于其它两种方法。图3给出了其中一次采用直方图方法、CSH方法和本文方法的检索结果。

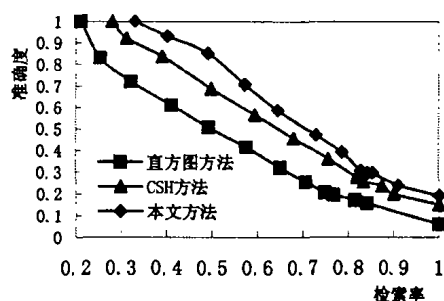


图2 检索率与精确度变化曲线

为了检验本文方法对旋转和平移等形变的敏感程度,还用多幅示例图像进行了不同角度和各种平移等形变处理的检索试验,发现本文方法仍具有较好的检索效果。另外在试验过程中,还对图像区域子块的尺寸进行了调整试验,结果发现对

(下转第184页)

计模式的实例。

参考文献

- 1 Gamma E, Helm R, Johnson R, Vlissides J. Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley Pub Co, 1995
- 2 Ferenc R, Beszedes A, Tarkainen M, Gyimothy T. Colubus - Reverse Engineering Tool and Schema for C++. In: Proc. of the 6th Intl. Conf. on Software Maintenance, IEEE Computer Society, Oct. 2002. 172~181
- 3 Kramer C, Prechelt L. Design Recovery by Automated Search for Structural Design Pattern in Object-Oriented Software. In: Proc. of the Third Working Conf. on Reverse Engineering, Nov. 1996. 208~215
- 4 Houston T X. Paradigm Plus 2.01 Reference Manual, ProtoSoft Inc., 1994
- 5 Brondby D. Visual Prolog 4.0, Prolog Development Center, 1996
- 6 Rumbaugh J, Blaha M, et al. Object-Oriented Modeling and Design. Prentice-Hall, Englewood Cliffs, NJ, 1991
- 7 Antoniol R F G, Cristoforetti L. Using Metrics to Identify Design Patterns in Object-Oriented Software. In: Proc. of the Fifth Intl. Symposium on Software Metrics, Nov. 1998. 23~34
- 8 Petroni A. Rappresentazione di design orientati agli oggetti ed analisi basata su metriche. Tesi di Diploma, University of Trento, Trento, Italy, March 1997
- 9 Balanyi Z, Ferenc R. Mining Design Patterns from C++ Source Code. In: Proc. Intl. Conf. on Software Maintenance, Sept. 2003. 305~314

(上接第160页)

同一类图像,区域子块的尺寸越小,相似性测量的效果越好,而且对图像中的小尺寸物体也有更好的识别能力,但是减小子块的尺寸以后,检索速度却受到了较大影响。因此,在实际

应用中可以折中考虑两种因素,选取合适的分块尺寸。同时,对不同类别的图像,通过调整式(12)中参数 ω_1, ω_2 的取值,可获取更好的检索效果。

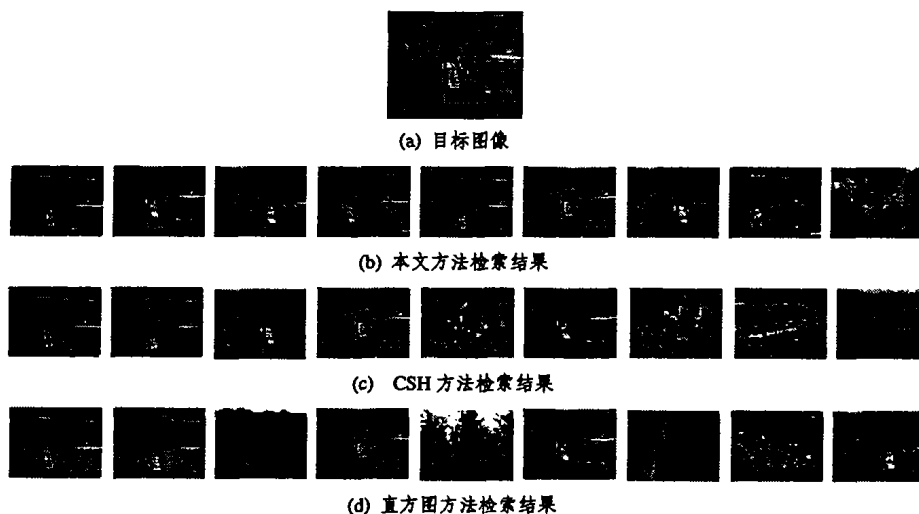


图3 实验结果

结论 本文提出了一种新方法,将彩色图像的颜色特征和空间特征用于彩色图像的相似性检索。色调分量反映了图像颜色的不变量,图像的状态矩阵不仅反映了图像的空间位置关系,也反映了图像的形状信息,采用两种特征相结合的方法进一步提高了图像检索的准确性。本文方法推导简单,且对图像没有特殊的限制和预处理要求。试验结果表明,本文方法不但有效,而且具有较好的检索性能,其检索结果能较好地接近人的视觉感知效果。但如何在检索过程中引入学习机制(如相关反馈技术),是本文方法下一步的工作。

参考文献

- 1 王涛,胡事民,孙家广. 基于颜色-空间的图像检索[J]. 软件学报, 2002, 13(10): 2031~2036
- 2 孙君顶,武学东,周利华. 基于颜色和形状的图像检索[J]. 计算机科学, 2004, 31(5): 180~183
- 3 Gevers T, Smeuder A W M. Evaluating Color and Shape Invariant Image Indexing of Consumer Photograph[A]. In: Proc. of the 1st Intl. Conf. on Visual Information Systems[C], Melbourne, Australia, 1996. 254~261
- 4 Gevers T, Smeuder A W M. Content-Based Image Retrieval by Viewpoint-Invariant Image Indexing[J]. Image and Vision Computing, 1999, 17(7): 475~488
- 5 Swain M J, Ballard D H. Color Indexing. Intern Journal of Computer Vision, 1991, 7(1): 11~32
- 6 Stricker M, Orengo M. Similarity of Color Images[J]. In: Proc. of SPIE Storage and Retrieval for Image and Video Databases, 1995, 2420: 381~392
- 7 刘忠伟. 利用局部累加直方图进行彩色图像检索[J]. 中国图象图形学报, 1998, 3(7): 533~537
- 8 Huang J. Image Indexing Using Color Correlograms[A]. In: Proc. on the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition[C], San Juan, Puerto Rico, USA, 1997. 762~768
- 9 Nastar C, Mitschke M, Meihac C. Efficient Query Refinement for Image Retrieval[A]. In: Proc. of IEEE Computer Society Conf. on Computer Vision and Pattern Recognition[C], Santa Barbara, California, USA, 1998. 547~552
- 10 Smeulders A W M, Santini S, Worring M, et al. Content Based Image Retrieval at the End of the Early Years[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(12): 1349~1380
- 11 Stehling R O, Nascimento M A, Falcao A X. On 'shapes' of colors for content-based image retrieval[A]. In: Proc. of the 2000 ACM Workshops on Multimedia, California, United States 2000. 171~174