

# 基于颜色和形状特征的彩色图像检索方法

孙君顶<sup>1)</sup> 崔江涛<sup>1)</sup> 毋小省<sup>2)</sup> 周利华<sup>1)</sup>

<sup>1)</sup>(西安电子科技大学多媒体研究所,西安 710071) <sup>2)</sup>(焦作大学计算机系,河南焦作 454000)

**摘要** 针对基于内容的彩色图像相似性检索问题,提出了一种新的相似性彩色图像检索方法,该方法既考虑了图像的颜色特征,也考虑了图像的空间信息特征,即在对 HSV 颜色模型进行特殊处理的基础上,将提取的色调不变量作为图像的颜色特征,同时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。在进行图像间的相似性测量时,为了结合不同的子特征进行全局的相似性检索,还采用 Gaussian 模型对不同的子特征间的距离进行了归一化处理。用不同类型的图像对这种方法进行的试验结果表明,它用于图像的相似性检索是很有效的,并具有较高的检索效率。

**关键词** 相似性检索 色调直方图 状态矩阵 基于内容的图像检索(CBIR) 转移概率矩阵

中图法分类号: TP391.3 文献标识码: A 文章编号: 1006-8961(2004)07-0820-08

## Color Image Retrieval Based on Color and Shape Features

SUN Jun-ding<sup>1)</sup>, CUI Jiang-tao<sup>1)</sup>, WU Xiao-sheng<sup>2)</sup>, ZHOU Li-hua<sup>1)</sup>

<sup>1)</sup>(Research Institute of Multimedia, Xidian University, Xi'an 710071)

<sup>2)</sup>(Department of Computer Science, Jiaozuo University, Jiaozuo Henan 454000)

**Abstract** The color histogram based image retrieval method is simple and invariant for translation and rotation of the images but losing the spatial information of the color. Recently many methods, such as accumulative histogram, color correlograms, local color histogram, etc, are introduced to improve the color histogram method. In this paper, a new content-based color image retrieval method is proposed, in which both the color content and the shape feature of the image have been taken into account. Firstly, based on the special disposal on the HSV color space, an improved accumulative histogram of the hue is calculated as the color feature. To attain the spatial information, H-, S-, and V-component of the image are firstly divided into  $n \times n$  blocks which are classified into 3 status, flatness, texture and edge status. Then each gray image is translated into a matrix composed of those 3 status values. After that the status matrix is transformed into 1-dimension status sequence, the transition probability matrix of the sequence is calculated as the image's spatial distribution information. In matching the similarity of the images, the Gaussian model is used to normalize the different sub-characters distance. Experiments with different kinds of images indicate that this method is great effective in image's retrieval.

**Keywords** similarity retrieval, hue histogram, state matrix, content-based image retrieval (CBIR), transition probability matrix

## 1 引言

随着多媒体和网络技术的迅速发展,图像数据的来源不断扩大,为了能从大量的图像数据中快速、准确地找到用户所需内容,基于内容的图像检索(content-based image retrieval, CBIR)技术正越来越

成为人们研究的焦点。近年来,人们对与图像内容处理有关的新方法和工具进行了相当深入的研究,相继推出了不少基于内容的图像检索系统。由于颜色作为图像最重要的一种低层特征,其对于区分图像十分有用,因此基于颜色的图像表示方法自然就成为主要的图像索引技术,并得到相当广泛和深入的研究。常用的颜色索引技术包括颜色直方图相交

法<sup>[1]</sup>,该方法虽具有特征提取和相似度计算简便的优点,但是在进行颜色量化时会丢失颜色信息,同时,全局颜色直方图只记录了全局的颜色统计信息,而丢失了颜色的空间分布信息;累加直方图方法<sup>[2,3]</sup>虽可通过累加方式来增加直方图的鲁棒性,但仍会丢失颜色的空间分布信息;类似纹理共生矩阵的 color correlograms<sup>[4]</sup>方法虽然能在一定程度上提取图像的颜色信息和空间信息,但存在计算复杂度大等问题。除了图像的颜色特征外,形状信息是另一种重要的图像低层特征,由于形状表示了图像中有意义的区域或相关对象,因此基于形状的图像索引技术允许用户查询具有相似对象的数据库图像。目前,基于空间局部化或区域匹配的图像检索技术已成为又一研究方向,并且已提出了一些有效的方法,如不需分割的图像直方图方法<sup>[5]</sup>、区域近似分割方法<sup>[6]</sup>以及图像的空间结构方法<sup>[7]</sup>等。

本文旨在研究一种有效的基于图像颜色和形状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形状信息的思路。这样就得到了一种新的基于图像颜色分布特征和形状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 模型对不同子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,例如:距离为 50 的(0,0,0)与(50,0,0)两种 R、G、B 颜色被认为是同一黑色;而距离为 50 的(200,150,0)和(200,200,0)则被认为是差别很大的两种颜色。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,所以,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视觉特性,文中首先将图像中每一像素的红、绿、蓝分量值  $r, g, b$  转化为亮度、饱和度、色调分量值  $h, s, v$  ( $h \in [0, 360], s \in [0, 1], v \in [0, 1]$ )(其中从 RGB 颜色空间到 HSV 颜色空间转换可参阅文献<sup>[8]</sup>),然后根据 HSV 颜色模型的特性做如下特殊处理<sup>[9]</sup>:

(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用  $h, s, v$  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<sup>[10,11]</sup>。基于此,本文利用彩色图像的色调分量  $h$  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<sup>[1]</sup>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。同颜色直方图类似,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<sup>[2,3]</sup>体现了两种颜色在颜色轴上的距离与相似性之间的关系,所以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献[2,3]的方法来求取其累加直方图, 并将该累加直方图作为图像的颜色特征。

### 3 马尔可夫链与图像的形状表示

众所周知, 图像的细节区域是图像的重要特征, 在图像检索中具有十分重要的作用。鉴于图像在转化到 HSV 颜色空间后, 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形状信息, 并提出一种有效的形状描述方法, 即首先将 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, 将图像的状态矩阵按照一定的顺序转化为一维状态序列; 最后利用马尔可夫链的特性分别求取状态序列的转移概率矩阵, 并将所求取的转移概率矩阵作为图像形状特征的描述。

#### 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  $\sigma$  表示子块内像素灰度值的标准差,  $\sigma$  定义为

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

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  $\alpha_1, \alpha_2$  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。

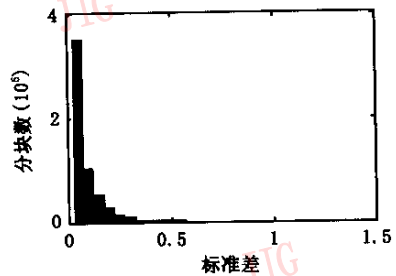


图 1 分块标准差分布

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 马尔可夫链

马尔可夫链最初是由 Markov 首先提出和研究的一类随机过程, 目前已广泛地应用于计算机、通信、自动控制等各个方面, 马尔可夫链及转移概率矩阵的详细描述过程可参阅文献[12]。设离散参数  $N_0 = \{0, 1, 2, \dots\}$ , 状态空间  $S = \{1, 2, \dots\}$ , 则有如下定义<sup>[12]</sup>:

定义 1 如果随机序列  $\{x_c, c \geq 0\}$  对于任意的  $a_0, a_1, \dots, a_c, a_{c+1} \in S, c \in N_0$  及概率  $p \{x_0 = a_0,$

$x_1 = a_1, \dots, x_c = a_c > 0$ , 有

$$p\{x_{c+1} = a_{c+1} | x_0 = a_0, x_1 = a_1, \dots, x_c = a_c\} \\ = p\{x_{c+1} = a_{c+1} | x_c = a_c\} \quad (2)$$

则称其为马尔可夫链,式(2)刻画了马尔可夫链的特性,称为马尔可夫性(或无后效性)。

定义 2  $\forall b_1, b_2 \in S$ , 称  $p\{x_{c+1} = b_2 | x_c = b_1\} \triangleq p_{b_1, b_2}(c)$  为  $c$  时刻的一步转移概率。记  $P = (p_{b_1, b_2})$ ,  $P$  称为  $\{x_c, c \geq 0\}$  的一步转移概率矩阵。

同时,由文献[12]的分析可知,由于一个马尔可夫链运动规律的概率特性取决于它的转移概率矩阵特性,因此,本文在对图像的状态矩阵进行处理时,即利用了马尔可夫链的这种特性。

### 3.3 基于马尔可夫链的图像特征抽取与表示

众所周知,若在二维图像空间中,像素的空间位置越近,则其相关性就越强,也就是说,对于图像的区域子块,空间位置越近,相应的相关性也越强。同样,对于图像的状态矩阵,也具有同样的特性,即空间位置越近的状态之间的相关性越强。为了从图像的状态矩阵中提取图像的形状信息和空间位置信息,可采用 Z 型扫描的方法将图像的状态矩阵转化为一维随机状态序列  $\{y_1, y_2, \dots, y_{n \times n}\}$  (其中,  $y_e$  表示第  $e$  个子块的状态,  $e = 1, 2, 3, \dots, n \times n$ ), 该序列对应的状态空间为  $U = \{0, 1, 2\}$ 。采用 Z 型扫描方法不仅可以尽可能将图像状态矩阵中相关性强的状态联系在一起,而且这样更有利于提取图像的形状特征。由于在图像的一维状态序列中,位置越靠近的两个状态的相关性越强,若某一状态为平坦态,则其相邻的两个状态在很大程度上可能为平坦态,但随着不同状态在序列中距离的增大,它们间的相关性会逐步减弱,本文仅考虑某一状态同其前一状态间的相关性,因此,做如下假设:在图像的状态序列中,任何一种状态的出现仅与该状态的前一状态有关,而与该状态更前的状态及后继状态无关,即当前子块状态是“平坦态”、“纹理态”还是“边缘态”仅仅与它前一子块的状态有关。此假设条件是为了使随机序列满足马尔可夫链的一步转移条件,在此条件下,该图像状态序列满足:(1) 序列的状态数目有限(状态数为 3);(2) 对于某一幅图像,其对应的一维状态序列可唯一确定,且该状态序列同时间的变化没有关系;(3) 进入某状态的概率仅与该状态之前的状态有关。

结合 3.2 节的内容,就可以采用一步转移概率矩阵来描述该图像随机序列的特性。设  $r$  及  $t$  是图像状态随机序列中任意两个邻接状态,对于转移概

率矩阵可采用式(3)来描述

$$P = \{p_{r,t}\} = \begin{bmatrix} p_{0,0} & p_{0,1} & p_{0,2} \\ p_{1,0} & p_{1,1} & p_{1,2} \\ p_{2,0} & p_{2,1} & p_{2,2} \end{bmatrix} \quad (3)$$

其中,  $p_{r,t} = p\{y_e = t | y_{e-1} = r\} = \frac{N_{r,t}}{\sum_{z=0}^2 N_{r,z}}$ ,  $e = 1, 2, 3, \dots, n \times n$ ,  $r, t \in U$ ,  $U = \{0, 1, 2\}$

若将上述的邻接状态  $r$  及  $t$  称为一状态对,记为  $(r, t)$ , 则在上式中  $N_{r,t}$  表示在图像随机状态序列中,状态对  $(r, t)$  出现的总次数。依据上述方法,就可以求出彩色图像 3 个分量图像的转移概率矩阵,并将所求出的 3 个转移概率矩阵作为图像的形状特征的描述。

对于由图像状态序列所求取的转移概率矩阵同数学意义上由随机序列所求取的转移概率矩阵稍有不同,假设随机状态序列的状态空间  $U = \{0, 1, 2\}$ , 则在数学意义上,转移概率矩阵  $\{p_{r,t}\}$  满足

$$\sum_{t=0}^2 p_{r,t} = 1, r \in U \quad (4)$$

但由图像状态矩阵转化来的图像状态序列却并不完全满足上式,由于对图像库中与具体某幅图像不同灰度分量对应的随机状态序列,其状态空间  $\hat{U} \subseteq U$ , 且其状态空间为图像库状态空间的子集,因此,如果与某图像对应的随机状态序列中不存在某种状态,那么采用式(3)来计算状态序列的转移概率矩阵时就会出现  $\sum_{t=0}^2 p_{r,t} = 0$  的情况,即分母为零的情况。为了防止这种情况,在实际处理过程中,可首先判断  $\sum_{t=0}^2 p_{r,t}$  的取值,如果  $\sum_{t=0}^2 p_{r,t} = 0$ , 则直接令  $p_{r,0} = p_{r,1} = p_{r,2} = 0$ , 以防止出现分母为零的情况。

## 4 距离度量与相似检索

### 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)<sup>[1]</sup>来进行直方图间的距离度量,直方图相交算子定义为

$$d_1(V, W) = 1 - \sum_{L=1}^{18} \min(v_L, w_L) \quad (5)$$

其中,  $L$  表示量化的色调数;  $V = \{v_1, v_2, \dots, v_{18}\}$ ,  $W = \{w_1, w_2, \dots, w_{18}\}$  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。

对于由式(3)计算出的图像子特征,本文提出了采用转移概率矩阵间的相关系数来进行距离度量的方法,矩阵间的相关系数可通过式(6)计算,设 $P_A$ 、 $P_B$ 分别表示两幅图像 $A$ 、 $B$ 对应分量的转移概率矩阵, $u$ 及 $o$ 表示转移概率矩阵的维数, $u=o=3$ 。

$$R(P_A, P_B) = \frac{\sum_u \sum_o ((P_A)_{u,o} - \mu_A)((P_B)_{u,o} - \mu_B)}{\sqrt{(\sum_u \sum_o ((P_A)_{u,o} - \mu_A)^2)(\sum_u \sum_o ((P_B)_{u,o} - \mu_B)^2)}} \quad (6)$$

其中

$$\mu_A = \frac{1}{u \times o} \sum_u \sum_o (P_A)_{u,o}$$

$$\mu_B = \frac{1}{u \times o} \sum_u \sum_o (P_B)_{u,o}$$

对于两幅彩色图像间距离度量定义如下:

$$d_2(P_A, P_B) = \sum_{K=1}^3 R_K(P_A, P_B) \quad (7)$$

其中, $K=1, 2, 3$ 分别表示彩色图像的3个灰度分量, $R_K$ 表示两幅图像对应灰度分量状态矩阵间的相关系数。

## 4.2 颜色与形状特征的组合检索

对于所提取的不同图像特征,需要定义不同的距离度量来计算图像间的相似程度,为了结合不同的子特征来进行图像检索,则需要将不同子特征所计算出的距离进行线性组合,即将两个图像之间的距离定义为子特征间距离的加权和。由于不同的子特征会采用不同的度量方式,因此,为了使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<sup>[13]</sup>。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 $\mu_M$ 和 $\sigma_M$ ,在 Gaussian 模型的假设下, $d_M$ 相应的归一化距离度量 $\hat{d}_M$ 定义为

$$\hat{d}_M(Z_1, Z_2) = \frac{d_M(Z_1, Z_2) - (\mu_M - 3\sigma_M)}{6\sigma_M} \quad (8)$$

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,若 $M=1$ ,则 $Z_1$ 和 $Z_2$ 表示上文提取图像的颜色特征矢量,若 $M=2$ ,则 $Z_1$ 和 $Z_2$ 表示上文提取的描述图像空间分布特征转移概率矩阵。据研究,上式不仅可保证 $\hat{d}_M$ 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(4)和式(6)计算出图像子特征之间的距离,利用式(7)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

$$d(Q, G) = \sum_{M=1}^2 W_M \hat{d}_M(q_M, g_M) \quad (9)$$

其中, $q_M$ 和 $g_M$ 分别为图像 $Q$ 和 $G$ 的第 $M$ 个子特征的特征向量, $\hat{d}_M(q_M, g_M)$ 表示子特征 $M$ 归一化后的距离, $W_M$ 为子特征 $M$ 的权重,可用其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 $W_1 + W_2 = 1, W_1, W_2 \in [0, 1]$ 。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,选用两个图像库分别进行了几组不同的试验,其中一个图像库包含球体、书本等在内共6类50幅在不同光照下形成的图像,该测试图像库可通过(<http://www.cs.sfu.ca/~colour/>)网站下载;另一图像库包含592幅光照情况基本相同的图像,它们包括飞机、动物、建筑、自然景物等。图像的尺寸取 $256 \times 192$ ,图像区域子块的尺寸取为 $8 \times 6$ ,在利用式(9)进行图像间的距离度量时,取 $W_1 = W_2 = 0.5$ ,这样就赋予颜色特征和形状特征同等重要性。

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)”<sup>[14]</sup>作为相似检索的评价准则。其中精确度 $I$ 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

$$I = Y/X \quad (10)$$

检索率 $J$ 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

$$J = Y/F \quad (11)$$

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,  $I \in [0, 1], J \in [0, 1]$ 。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。

对于试验 1, 本文是采用“检索率”来对检索结果进行评价。实验是从每类图像中分别抽取 2 幅图像作为检索实例图像进行了 12 次检索, 在 12 次检索结果中, 检索出的前 7 幅图像中检索率的平均值为 72.3%, 采用文献[9]方法的 12 次检索结果中, 在检索出的前 7 幅图像中检索率的平均值为 58.9%, 而采用颜色直方图方法[1]的 12 次检索结果中, 在检索出的前 7 幅图像中检索率的平均值仅为

45.2%。图 2 给出了本文方法、文献[9]方法和直方图方法的其中 1 次的检索结果, 其中每行最左边的一幅图像为示例图像兼检索结果, 随后为检索结果图像队列, 从左至右相似度是由大到小顺序排列的。

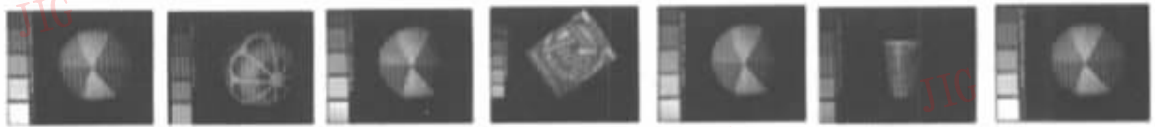
在光照不同的情况下, 由于本文方法是采用颜色的不变量(色调)作为图像的颜色特征, 因此该特征相对于直方图方法来说, 受光照的影响相对较小, 同时由于本文方法同时采用颜色信息和空间信息作为检索依据, 因此取得了较好的检索效果。文献[9]的方法虽对于有明显物体出现的图像有较好的检索效果, 但对于该类图像, 由于受光照变化的影响, 图像中物体的特征发生了较大的变化, 因而影响了该方法的检索性能, 而直方图方法则由于完全丢弃了图像的空间信息, 因此直方图方法的检索结果中包含了许多与示例图像完全不相关的图像。



(a) 本文方法检索结果



(b) 文献[9]方法检索结果



(c) 直方图方法检索结果

图 2 试验 1 示例图像和检索图像队列

对于试验 2, 本文利用精确度和检索率相结合的方法来检验本文方法的检索结果。试验是先在图像库中选取 5 类图像, 且在每类图像中抽取 5 幅图像共组成 25 次查询, 然后结合这 25 次查询, 计算出在不同检索率条件下系统的平均精确度, 同时计算出直方图方法和文献[9]方法在相应检索率条件下的平均精确度, 3 种方法的统计结果如图 3 所示。从图 3 还可以看出, 本文方法明显优于直方图方法, 对于文献[9]的方法, 由于采用区域生长的方法对图像进行分割来提取图像中的物体, 因此当图像中有明显物体出现时, 采用该方法可取得较好的检索效果, 从图 3 还可看出, 在较低的检索率条件下, 文献[9]

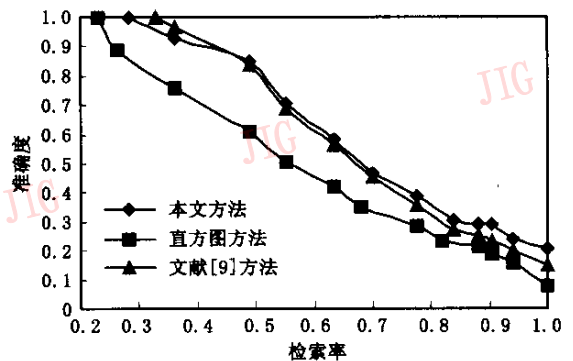


图 3 检索率与精确度变化曲线

的方法稍优于本文方法,但在其他条件下,本文方法明显占优,同时在提取图像的特征时,本文方法的计

算复杂度明显低于文献[9]的方法。图4给出了本文方法、文献[9]方法和直方图方法的一次检索结果。



(a) 本文方法检索结果



(b) 文献[9]方法检索结果



(c) 直方图方法检索结果

图4 试验2 示例图像和检索图像队列

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,但是由于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,可通过调整式(9)中参数 $W_1, W_2$ 的取值,以便获取更好的检索效果。

## 6 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信息特征用于彩色图像的相似性检索,即采用改进的颜色空间进一步提高了颜色模型的准确性,同时由于颜色的色调分量反映了图像颜色的不变量,且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,因此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研究内容。

## 参考文献

- Swain M, Ballard D. Color indexing[J]. International Journal of Computer Vision, 1991, 7(1):11~32.
- Stricker M, Orengo M. Similarity of color images [A]. In: Proceedings of SPIE Storage and Retrieval for Image and Video Databases[C], 1995, 2420:381~392.
- 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.
- Huang J. Image indexing using color correlograms [A]. In: Proceeding on the IEEE Computer Society Conference on Computer Vision and Pattern Recognition[C], San Juan, Puerto Rico, USA, 1997:762~768.
- Wan X, Kuo C J. A new approach to image retrieval with hierarchical color clustering[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1998, 8(5):628~643.
- Belongie S, Carson C, Greenspan H, et al. Color- and texture-based image segmentation using EM and its application to content-based image retrieval [A]. In: Proceedings of the 6th International Conference on Computer Vision (ICCV'98) [C], Bombay, India, 1998:1~7.
- Bimbo D, Vicario E. Using weighted spatial relationships in retrieval by visual contents [A]. In: Proceedings of IEEE Workshop on Image and Video Libraries [C], Santa Barbara, California, USA, 1998:75~79.
- Huang J. Color-spatial image indexing and applications [D]. Cornell University, New York, USA, 1998:20~32.

- 9 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.
- 10 Gevers T, Smeuder A W M. Evaluating color and shape invariant image indexing of consumer photograph [A]. In: Proceedings of the 1st International Conference on Visual Information Systems [C], Melbourne, Australia, 1996:254~261.
- 11 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.
- 12 林元烈. 应用随机过程[M]. 北京:清华大学出版社, 2002:78~130.
- 13 Nastar C, Mitschke M, Meihac C. Efficient query refinement for image retrieval [A]. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C], Santa Barbara, California, USA, 1998:547~552.
- 14 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.



崔江涛 1975 年生, 2001 年获西安电子科技大学计算机应用专业硕士学位, 现为西安电子科技大学多媒体技术研究所在职博士研究生。主要研究方向为多媒体数据库管理、图像处理、多媒体网络互联。

毋小省 1974 年生, 1998 年获焦作工学院计算机应用专业学士学位, 现为河南焦作大学计算机系讲师。主要研究方向为多媒体技术、图像处理。

周利华 1942 年生, 西安电子科技大学教授、博士生导师, 1965 年毕业于西安军事电讯工程学院自动控制专业。主要研究方向为多媒体网络互联技术、多媒体应用技术。



孙君顶 1975 年生, 2001 年获河南焦作工学院工业自动化专业硕士学位, 现为西安电子科技大学多媒体研究所博士研究生。主要研究方向为图像处理、基于内容的图像检索、多媒体技术。

E-mail:sunharry@mti.xidian.edu.cn