

# 一种基于颜色和纹理的图像检索方法

徐慧英 袁 杰 赵建民 朱信忠

(浙江师范大学数理与信息工程学院 金华 321004)

**摘要** 提出了一种基于颜色和纹理的新的图像检索方法。该方法在 HSV 颜色空间量化过程中,采用一种基于线性隶属度的模糊量化方法,使量化结果更接近于人类感知;在提取分块主色的基础上引入一种颜色相邻矩阵,用于描述颜色的相对空间分布情况,以提高对图像变形的鲁棒性;还加入图像的纹理信息,最后结合这些颜色和纹理特征进行复合图像检索。实验表明,该方法能有效提高图像检索精度。

**关键词** 基于内容图像检索,颜色空间,颜色量化,颜色直方图,纹理

## Image Retrieval Based on Color and Texture

XU Hui-ying YUAN Jie ZHAO Jian-min ZHU Xin-zhong

(College of Mathematics, Physics and Information Engineering, Zhejiang Normal University, Jinhua 321004, China)

**Abstract** A new picture retrieval algorithm based on color and texture was presented. In the process of scaling color in HSV color space, the new algorithm used a scaling method based on fuzzy scaling to better meet human's sensation. The algorithm also introduced a color adjacent matrix after extracting every picture block's main color to describe colors' relative spacious distribution and enhanced its robustness when the target picture had changed. It added pictures' texture information too, at last, it combines all those above to search involved pictures. Experiments indicate that this algorithm can improve the searching precision effectively.

**Keywords** CBIR, Color space, Color scaling, Color histogram, Texture

## 1 引言

基于内容的图像检索方法(CBIR)是最近十几年来发展的一种新的图像检索方法,它摒弃了以前那种以文找图的基于文本的图像检索方法(TBIR),而采用以图找图、按图索骥的方法。基于内容的图像检索主要是指将图像的可视特征(如颜色、纹理、形状、位置关系)作为图像的内容进行匹配、查找,利用某些算法,进行特征提取和匹配。基于内容的图像检索技术融合了低级的图像理解技术,从而可以提供更加有效的检索途径,而且可以增加反馈环节,进一步修正和优化检索结果,具有很好的用户体验。

颜色是最直接地描述图像内容的视觉特征。基于颜色的图像索引技术也是目前为止较成熟、应用最广泛的一种 CBIR 技术。Swain<sup>[1]</sup>首先提出了使用颜色直方图作为图像内容的索引特征,图像之间的相似性测度则采用直方图的交;类似地, Hafner<sup>[2]</sup>在 QBIC 中则采用了直方图的二次距离作为图像的相似性测度;另外, Stricker 等<sup>[3]</sup>采用了颜色的前三阶矩来表征图像的内容。近年来又提出了基于模糊 C 均值聚类的模糊颜色直方图方法等。纹理特征是一种不依赖于颜色或亮度来反映图像中同质现象的视觉特征,它包含了物体表面结构组织排列的重要信息,以及它们与周围环境的联系。它

是不同物质共有的本质特性,例如沙漠、树木、织物等都有各自的纹理特征。纹理的描述方法主要有灰度共生矩阵等。

物体具有各方面的特性,每个单方面的特性都不足以恰当地描述物体,所以在检索中也要结合物体的多种特性。本文结合颜色和纹理特征对图像进行检索。在颜色特征提取中,颜色矩对图像刻画过于粗糙,一般不单独使用,常作为多级过滤查询中的第一级;颜色直方图可以分为全局颜色直方图和分块颜色直方图。全局颜色直方图对整幅图像的宏观颜色统计信息能很好刻画,从而对图像的旋转、平移、缩放等有很好的鲁棒性,但没有描述颜色的空间位置分布,会检索出许多不相关的图像;分块颜色直方图由于引入了位置信息,对图像的描述更加精细,但其只能对对应位置的颜色信息进行比较,对图像的旋转、平移、缩放等比较敏感,鲁棒性不好,检索精度也不能得到有效的提高。

本文旨在研究一种有效的基于颜色和纹理特征的图像表示方法及相应的检索技术。颜色特征方面,首先在颜色量化时对 HSV 颜色空间的量化方法做了一定的改进,引入模糊量化方法,使量化结果更接近人类感知;然后采用基于 HSV 颜色空间的颜色分块直方图描述方法,同时引入一种刻画颜色相对位置分布的颜色相邻矩阵,这样很好地提高了分块直方图的鲁棒性。纹理特征方面,由于一般图像的纹理特征并

到稿日期:2008-08-25 本文受国家自然科学基金(60773197),浙江省自然科学基金(Y107750)资助。

徐慧英(1977-),女,讲师,主要研究方向为网上多媒体检索、图像处理与 WEBGIS 应用等, E-mail: xhy@zjnu.cn;袁 杰(1981-),男,硕士生,主要研究方向为互联网信息挖掘和智能信息检索;赵建民(1951-),男,教授,主要研究方向为模式识别、图像处理与数据挖掘;朱信忠(1975-),男,副教授,主要研究方向为模式识别、图像处理与多媒体检索等。

不是很显著,所以采用了一种较简单的类似于灰度共生矩阵的表示方法。最后将根据这两种特征算出的相似度进行加权综合,得出总的相似度。在相应的检索实验中,该方法相对一般的颜色直方图检索精度有了一定的提高,得到了很好的效果。

## 2 HSV 颜色空间及量化

图像中广泛使用的 RGB 彩色模型是一般的显示设备(如电视机、显示器等)使用的颜色空间,它容易用硬件实现,但它与人类的感官差别较大,RGB 颜色空间中相同欧氏距离的颜色人眼看起来差别不一样<sup>[4]</sup>。由于在 RGB 颜色空间中计算出的图像间的距离对图像的区分能力并不好,因此本文选用另一种颜色空间——HSV 颜色空间。

HSV 颜色空间采用色度  $H$ 、饱和度  $S$  和亮度  $V$  分离的方式实现颜色的定量描述。三维分量都存在明确的语义对应关系,相对于其它颜色模型,它更为准确地反映了人类视觉系统对色彩的理解方式,且不同颜色在该颜色空间中的欧式距离与视觉距离具有大体一致性,对图像区分度较好。

RGB 颜色空间到 HSV 颜色空间的转化是一个非线性过程,也有快速计算方法。将 RGB 颜色空间中的颜色转化到 HSV 颜色空间后,HSV 各分量的范围为: $h \in [0, 360]$ ,  $s \in [0, 1]$ ,  $v \in [0, 1]$ 。为了不至于混淆,规定  $h, s, v$  为 RGB 空间颜色转换到 HSV 空间的连续值,而  $H, S, V$  为  $h, s, v$  3 个量化后的离散值。

为了减少信息量并简化处理过程,需要对转换后的  $hsv$  值进行进一步量化,并将量化后的值统一成一个一维颜色矢量。本文采用的量化规则如下:

$$H = \begin{cases} 0 & h \in [0, 20] \cup [315, 360] & \text{红色} \\ 1 & h \in [20, 50] & \text{橙色} \\ 2 & h \in [50, 75] & \text{黄色} \\ 3 & h \in [75, 155] & \text{绿色} \\ 4 & h \in [155, 195] & \text{青色} \\ 5 & h \in [195, 275] & \text{蓝色} \\ 6 & h \in [275, 315] & \text{紫色} \end{cases} \quad (1)$$

$$S, V = \begin{cases} 0, s, v < 0.2 \\ 1, 0.2 \leq s, v < 0.7 \\ 2, s, v \geq 0.7 \end{cases} \quad (2)$$

这样  $H$  量化为 7 个值,  $S, V$  各量化为 3 个值,量化后合成一维的颜色矢量  $C$ 。合成方法如下:

① 对于  $v < 0.2$  的颜色认为是黑色,  $C=0$ ;

② 对于  $s < 0.1$ , 且  $v > 0.2$  的颜色按亮度  $v$  划分为 3 种灰度,分别为  $v \in (0.2, 0.5]$  深灰,  $v \in (0.5, 0.8]$  浅灰,  $v \in (0.8, 1]$  白色,  $C$  的相应值分别为  $C=1, 2, 3$ 。

③ 其它情况先按上述量化方法量化,然后算出  $C$ :

$$C = 4 + 9H + 3S + V \quad (3)$$

这种量化方法也存在一定问题。当要量化的值恰好在量化边界上或在量化边界附近时,由于这种截断量化天然的量化方式,该值只能归入一种颜色中,而对另一种颜色一点贡献也没有,这就割裂了相近颜色的相关性和连续性,引入了一定的量化误差。比如当  $h=75, s, v > 0.2$  时,此时它处于黄绿两种颜色交界处,按式(1)量化方法,  $H$  只能量化为 3, 即绿色; 而当  $h=74, s, v$  不变时它又量化为黄色。事实上,  $h=74$  与

$h=75$  时视觉差别非常小,人眼根本分辨不出,但量化后却成了截然不同的两种颜色。

为了尽量减少量化中引入的误差,本文借鉴模糊数学的思想,引入一种模糊量化方法。一般而言,空间中的色调  $H$  是最为重要的分辨性视觉特征,颜色之间的差异主要由颜色色调的差异描述。鉴于此,本文拟对  $h$  分量进行更精细的量化,  $s$  和  $v$  的量化方法不变。大体思路是,对上述红橙黄绿青蓝紫 7 种颜色,每种颜色都规定一个核,当要映射的像素的  $h$  分量落在相应颜色的核对应的  $h$  值的范围中时,则  $H$  值直接映射为上述量化图中相应的数值; 而当  $h$  分量没有落在核中,即  $h$  分量落在上述量化图的边界点附近时,使用某种映射函数将其映射为两个量化值中间的某个量,这样边界两边的量化值都对当前量化值产生了影响。以下用图 1 来说明。

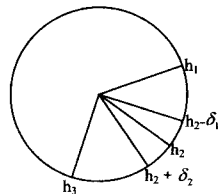


图1 模糊量化示意

在图 1 中,设  $h_1, h_2, h_3$  分别是以上量化方法中的量化边界,也即  $h \in [h_1, h_2]$  时  $H(h)=k$ ;  $h \in [h_2, h_3]$  时,  $H(h)=k+1$ 。取  $\delta_1 = k_1(h_2 - h_1)$ , 同理  $\delta_2 = k_1(h_3 - h_2)$ ,  $k_1$  为一常数,比如  $\frac{1}{4}$ 。由于  $h$  是非等间距量化,所以一般情况  $\delta_1$  与  $\delta_2$  不相等。

当  $h \in [h_1 + \delta_1, h_2 - \delta_1]$  时,  $h$  处于某种颜色对应的  $h$  值的核区,量化不变,  $H(h)=k$ 。

当处于两种颜色交界点附近时,即  $h \in (h_2 - \delta_1, h_2 + \delta_2)$  时,  $H_{new}(h)$  由计算得出:

$$H_{new}(h) = \omega_1 \cdot k + \omega_2 \cdot (k+1) \quad (4)$$

其中,  $H_{new}(h)$  表示按新规则量化后的  $H$  的值,  $\omega_1, \omega_2$  是权值,本文中取  $\omega_1 = \omega_2 = 0.5$ 。经过这种量化后  $H$  就可能取小数了。此时  $C$  的计算公式改为:

$$C = \text{Round}(4 + 9H + 3S + V) \quad (5)$$

其中  $\text{Round}$  是四舍五入函数。

经这种方法量化后,感官上相近的两种颜色其  $h$  分量量化得到的  $H$  值也很相近,最后得到的直方图对图像的颜色分布的描述更准确,也在一定程度上解决了量化边界不连续的问题,使得量化后的直方图更接近人的主观视觉感知,从而提高了检索的效率。

## 3 特征提取

### 3.1 分块颜色直方图

将图像等分成  $4 \times 4$  共 16 块,提取每块的颜色直方图作为特征矢量。由于在图像分割时,已经将空间信息嵌入到每一块中,又用颜色直方图表示图像块的颜色信息,所以形成的特征矢量既包含颜色信息又包含空间信息。

每一个块的特征包含了颜色的分块分布信息,而所有块的空间分布又代表了图像内容的颜色模式结构。由于上面已将颜色量化成 67 柄,故系统维护 16 个 67 柄的颜色直方图。

### 3.2 颜色相邻矩阵

分块颜色直方图由于引入了绝对位置,从而当图像变形时鲁棒性大大降低。怎样既引入颜色的空间分布又不至于显

著降低鲁棒性呢?

本文提出一种建立在 HSV 量化颜色空间之上的刻画颜色相对分布位置的方法,称之为颜色相邻矩阵法。当然,也可以推广到其它颜色空间。由于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,基于此,本文只取  $h$  的量化值  $H$  构成主色  $C'$ 。设  $C'$  共有  $k$  级,即  $C'=0,1,\dots,k-1$ 。本文实验中取  $k=10$ ,则  $C'$  的范围为 0 到 9。具体取值如:

$$C' = \begin{cases} H, C \neq 0, 1, 2, 3 \\ 7, C = 0 \\ 8, C = 1, 2 \\ 9, C = 3 \end{cases} \quad (6)$$

$C$  是量化后的  $H, S, V$  线性组合得到的一维颜色矢量。

得到一幅图像的颜色相邻矩阵步骤如下:

(1) 将一幅图像分成  $m \times n$  个小块,每一小块的大小都为  $s \times t$  个像素。

(2) 得出每一小块的主  $C'$  值,也就是该小块中  $C'$  取得次数最多的那个值。这样就形成一个二维主色矩阵,大小为  $m \times n$ ,表示为  $A = \{a_{i,j}\}, i=1,2,\dots,m, j=1,2,\dots,n$ 。

(3) 建立一个  $k \times k$  的矩阵  $P$ ,各元素初始值为 0。将上步得到的矩阵  $A$  按某种顺序进行扫描,设  $H_{i,j}$  与  $H_{p,q}$  是扫描序列中的一对相继出现的颜色( $H_{i,j}$  在  $H_{p,q}$  的前面),则  $P$  中相应元素  $P[H_{i,j}, H_{p,q}]$  自增 1,并将  $P[H_{i,j}, H_{p,q}]$  的值赋给  $P[H_{p,q}, H_{i,j}]$ 。如此反复,直到扫描完成。可以看出,矩阵  $P$  为一个对称阵。

(4) 建立  $k \times k$  的矩阵  $D$ ,  $D$  中元素的计算公式如下:

$$D_{i,j} = P_{i,j} / \sum_{n=1}^k P_{i,n} \quad (7)$$

则矩阵  $D$  就是该图像的颜色相邻矩阵。

上述第(3)步中矩阵的扫描顺序并没有规定,可以按具体情况选择一种扫描顺序。本文根据实际情况按图 2 所示的向右、右下、向下、左下 4 种方向扫描主色矩阵,从而得到 4 个颜色相邻矩阵。

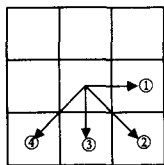


图 2 扫描方向

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \end{pmatrix}$$

图 3 示例矩阵

从上面的步骤可以看出,一方面矩阵  $D$  中的每个元素都描述了图像中各颜色的相邻关系,从而矩阵  $D$  能描述颜色的空间分布情况;另一方面,这种方法只是描述颜色间的相对相邻关系,它又不像分块颜色直方图那样引入绝对位置分布,因此对平移、旋转等变形具有鲁棒性。

图 3 所示的矩阵得到的 4 个颜色相邻矩阵分别为:

$$D_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, D_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

$$D_3 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}, D_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

### 3.3 纹理特征提取

纹理特征是图像重要特征之一,其本质是刻画像素的邻域灰度空间分布规律。任何物体表面经放大都会呈现出一定的纹理。图像纹理分析是图像分析处理研究中的一个重要组成部分。纹理不同于诸如灰度和颜色的特征,它所包含的信息不仅仅来源于单个像素,还和该像素周围空间邻域的灰度分布情况有密切的联系。

在频域分析和小波分析中,纹理一般处于频谱中的高频部分。而在空域分析中,20 世纪 70 年代 Haralick 等人提出了基于二阶灰度统计特征的共生矩阵方法<sup>[5]</sup>;基于人类对纹理视觉感知的研究, Tamura 等人提出了纹理特征的表达形式。Tamura 纹理特征的 6 个分量对应于心理学角度上纹理特征的 6 种属性,分别是粗糙度(Coarseness)、对比度(Contrast)、方向度(Directionality)、线性度(Linearity)、规整度(Regularity)和粗略度(Roughness)<sup>[6]</sup>。小波分析在纹理分析中也发挥着越来越重要的作用<sup>[7]</sup>。

本文采用一种较简单的表示法描述图像的纹理信息,该方法在文献[8]的基础上做了一些改进。首先将图像转换成灰度图,在灰度图的基础上进行处理。一般的图像大致可以分为前景部分和背景部分,前景部分也即目标物体。前景部分和背景部分各自内部,灰度值变化比较平缓,而前景部分和背景部分交界处灰度值变化一般比较尖锐。在纹理图像中,根据纹理基元的大小可以将纹理分为粗纹理和细纹理。对具有不同结构的纹理模式来说,基元尺寸越大或者基元重复次数越少,则给人的感觉越粗糙。综上所述,可以把一幅图像(包括纹理图像和非纹理图像)分为 4 个区:平滑区、粗纹理区、细纹理区和边界区。具体特征提取方法如下:

将图像转换成灰度图,并将大小规格化为  $128 \times 128$ 。规格化后的图像分成  $m \times n$  个小块,每一小块的大小都为  $s \times t$  个像素,可使本次分块大小与上面提取颜色相邻矩阵的分块大小保持一致。

按下式算出每一块颜色的均值和标准差:

$$\mu_{i,j} = \frac{1}{s \times t} \sum_{p=1}^s \sum_{q=1}^t G_{p,q} \quad i=1,2,\dots,m \quad j=1,2,\dots,n \quad (8)$$

$$\delta_{i,j} = \left( \frac{1}{s \times t} \left( \sum_{p=1}^s \sum_{q=1}^t (G_{p,q} - \mu_{i,j})^2 \right) \right)^{\frac{1}{2}} \quad (9)$$

$G_{p,q}$  表示  $block_{i,j}$  中相对位置为  $(p,q)$  的点的灰度值。

(3) 建立一个  $m \times n$  大小的矩阵  $T$  表示纹理。取 3 个阈值  $\delta_1 < \delta_2 < \delta_3$ , 比较  $\delta_{i,j}$  与上述 3 个阈值的值。当  $\delta_{i,j} < \delta_1$  时,认为此小块变化比较平缓,令  $T_{i,j} = 1$ ; 当  $\delta_{i,j} \in [\delta_1, \delta_2)$  时,认为是细纹理,令  $T_{i,j} = 2$ ;  $\delta_{i,j} \in [\delta_2, \delta_3)$  时,认为是粗纹理,令  $T_{i,j} = 3$ ; 当  $\delta_{i,j} > \delta_3$  时,认为是边缘,令  $T_{i,j} = 4$ 。

这样就将提取出的纹理特征放在矩阵  $T$  中了。而对于  $\delta_1, \delta_2, \delta_3$  的取值,经过多次实验得知,  $\delta_1 = 0.05, \delta_2 = 0.2, \delta_3 = 0.35$  时分辨效果较好。

### 4 图像相似性度量

相似性度量是图像检索技术中的一个关键问题,同时是

其中的一个难点。在基于文本的检索方法中,一般采用基于精确的匹配,而在基于内容的图像检索中则要通过计算用户例图和数据库中图像之间在视觉特征上的相似度进行,它只能是图像的近似表示,因此选择一个合适的相似度量方法对检索的效果影响较大。

由于在基于内容的图像检索中图像的特征多是用向量表示的,所以常用的相似性度量方法都是建立在向量空间模型中。相似性度量方法就是利用相关距离计算函数,计算被检索对象和目标对象的特征序列之间的结果,用该计算值来判别对象之间的相似性程度。

本文提取了颜色和纹理两个方面的特征,其中颜色特征又分为分块颜色直方图和颜色相邻矩阵。下面分别介绍各种分特征的相似性度量方法。

#### 4.1 分块颜色直方图相似性计算

计算两个块  $b_1, b_2$  间的颜色直方图距离时,本文采用经典的直方图相交距离,其数学描述为

$$d(b_1, b_2) = \sum_{i=1}^n \min(b_{1,i}, b_{2,i}) \quad (10)$$

$n$  为直方图的柄数,即量化颜色数。 $\min(b_{1,i}, b_{2,i})$  表示块  $b_1$  和  $b_2$  中颜色值为  $i$  的点数的极小值。

而两个块间相似性度量为:

$$\text{Similarity}(b_1, b_2) = \frac{\sum_{i=1}^n \min(b_{1,i}, b_{2,i})}{\sum_{i=1}^n b_{1,i}} \quad (11)$$

比较时,两幅图像对应位置的块进行比较。这样算出 16 个对应块的相似度,然后由式(12)得出两幅图像  $A$  和  $B$  的分块颜色直方图法相似度:

$$\text{similarity}_1(A, B) = \frac{1}{16} \sum_{i=1}^{16} \text{similarity}(b_i(A), b_i(B)) \quad (12)$$

#### 4.2 颜色相邻矩阵相似性计算

两幅图像,当提取出它们各自的颜色相邻矩阵后,算出这两个矩阵的相关系数作为这两幅图像的另一种相似度。当只有一个颜色相邻矩阵时,设  $DA, DB$  分别表示两幅图像  $A, B$  的颜色相邻矩阵,大小皆为  $k \times k$  (本文中  $k=10$ ),则两幅图像间相似度定义如下:

$$\text{Similarity}_2(A, B) = \frac{\sum_{i=1}^k \sum_{j=1}^k (D_{A_{i,j}} - \mu_A)(D_{B_{i,j}} - \mu_B)}{\sqrt{(\sum_{i=1}^k \sum_{j=1}^k (D_{A_{i,j}} - \mu_A)^2)(\sum_{i=1}^k \sum_{j=1}^k (D_{B_{i,j}} - \mu_B)^2)}} \quad (13)$$

其中,

$$\mu_A = \frac{1}{k^2} \sum_{i=1}^k \sum_{j=1}^k D_{A_{i,j}} \quad (14)$$

$$\mu_B = \frac{1}{k^2} \sum_{i=1}^k \sum_{j=1}^k D_{B_{i,j}} \quad (15)$$

由于上文提取出了 4 个方向的颜色相邻矩阵,所以按上式算出两幅图像对应方向的颜色相邻矩阵后,按式(16)进行综合:

$$\text{similarity}_2(A, B) = \frac{1}{t} \sum_{i=1}^t \text{similarity}_{2,i}(A, B) \quad (16)$$

本文中  $t$  取 4。

#### 4.3 纹理相似性计算

设两幅图像  $A, B$  各自的纹理特征矩阵分别为  $T_A, T_B$ ,则两幅图像的纹理相似性如下计算:

$$\text{similarity}_3(A, B) = 1 - \frac{1}{2} \left( \frac{|\mu_B - \mu_A|}{\max_{i=1,2,\dots,n} |\mu_i - \mu_A|} + \right.$$

$$\left. \frac{|\delta_B - \delta_A|}{\max_{i=1,2,\dots,n} |\delta_i - \delta_A|} \right) \quad (17)$$

其中  $\mu_A, \mu_B$  分别是  $T_A$  和  $T_B$  的均值,而  $\delta_A, \delta_B$  分别是  $T_A$  和  $T_B$  的标准差,它们的计算公式与颜色相邻矩阵中均值和标准差计算公式一样。

#### 4.4 综合颜色特征和纹理特征的图像相似性计算

上面已经得到了 3 个相似度,它们从不同的角度描述了两幅图像的相似情况。本文将这三者结合起来进行合成检索。因每个相似度的取值范围已经是介于 0 与 1 之间,所以不需再对其进行归一化操作。合成相似度如下:

$$\text{Similarity}_{\text{total}}(A, B) = \sum_{i=1}^3 W_i \text{Similarity}_i(A, B), (W_1, W_2, W_3 \in [0, 1], W_1 + W_2 + W_3 = 1) \quad (18)$$

其中  $W_1, W_2$  和  $W_3$  的值可按实际情况灵活设置,本文中设置  $W_1=0.5, W_2=0.3, W_3=0.2$ 。

#### 4.5 综合检索步骤

本文采用基于用户例图查询的工作方式,由用户提供例图,系统返回库中与例图计算相似度最大的前 30 幅图形。为简便起见,图像统一放在某一特定目录下而并没有放在数据库中,图像特征提取也是实时进行的。具体流程如下:

(1)将例图和当前检索图像大小规格化为  $128 \times 128$  大小。

(2)将规格化后的图像颜色空间转换到 HSV 空间,同时进行模糊量化,量化后共 67 级。

(3)将图像分成  $64 \times 64$  大小的块共 16 块,得出每块的颜色直方图,同时将这些块继续分成  $8 \times 8$  大小的小块,提取各小块的主色  $C'$ ,得到整幅图形的主色矩阵。

(4)按上面介绍的分块直方图相交法算出两幅图像的相似度  $\text{Similarity}_1$ 。

(5)扫描主色矩阵,得出图像的 4 个颜色相邻矩阵,并算出两幅图像的相邻矩阵相关系数,也即  $\text{Similarity}_2$ 。

(6)将图像转换成灰度图,得到该图像的纹理特征矩阵,计算出两幅图像的纹理特征相似度  $\text{Similarity}_3$ 。

(7)合成 3 个相似度,得到两幅图像总的相似度。

(8)对图像库中的每幅图像重复以上步骤,算出每幅图像与例图的相似度。

(9)将图像按相似度的逆序排列,最前面的 30 幅返回给用户。

## 5 试验结果及评介

本文采用检索准确率和自定义的一个参数对检索结果进行评介。检索准确率  $P$  定义为某次检索中检索出的相关图像数与检索出的总图像数的比率。自定义参数描述如下:设在某次查询  $Q$  中,检索到的前  $n$  幅图像中有  $n_1$  幅是相关图像,分别为  $PR_1, PR_2, \dots, PR_{n_1}$ ,在检索结果中相应的顺序号分别为  $NO_{PR_1}, NO_{PR_2}, \dots, NO_{PR_{n_1}}$ 。另前  $n$  幅图像中有  $n_2$  幅是不相关的,分别为  $PN_1, PN_2, \dots, PN_{n_2}$ ,相应的序号分别为  $NO_{PN_1}, NO_{PN_2}, \dots, NO_{PN_{n_2}}$ ,  $n_1 + n_2 = n$ 。定义

$$\text{AvgP} = \frac{1}{2} + \frac{1}{n(n+1)} \left( \sum_{i=1}^{n_1} (n+1 - NO_{PR_i}) - \sum_{i=1}^{n_2} (n+1 - NO_{PN_i}) \right) \quad (19)$$

当前  $n$  幅图像全部相关时,  $\text{AvgP}$  取得最大值 1;当前  $n$

幅图像都不相关时, AvgP 取得最小值 0; 其它情况介于 0 与 1 之间。越接近 1 说明效果越好, 这样 AvgP 就能很好地刻画检索效果。本测试中  $n$  取 30, 即只显示前 30 幅图像, 且以下各种参数计算也都只针对前 30 幅来进行。

用 MATLAB 实现本文提出的算法并在 SIMPLcity<sup>[9]</sup> 集的一个子集上进行了实验。SIMPLcity 测试集是从 Corel 图像库中选取的, 共包含 10 个语义类。10 个语义类分别为: 非洲人、海滩、建筑物、公共汽车、恐龙、大象、花、马、高山和食物。本文实验中在每个语义类中取前 30 幅图片, 这样测试库由 300 幅图像组成。本文比较了全局颜色直方图、分块颜色直方图和本文提出的方法共 3 种方法的试验结果。具体测试时, 每种方法每个类别中取 6 幅图像作为例图检索, 这样每种方法共进行 60 次检索, 待比较的 3 种方法共进行 180 次检索。对于每类图像而言, 计算其 6 次查询结果的检索准确率平均值和 AvgP 平均值, 并将其作为该类图像的平均准确率和平均 AvgP 值。

表 1 显示了在用 3 种方法对所有类别图像进行检索时的评测值。

表 1 试验结果评测

图像类别	方法		
	全局直方图	分块直方图	本文方法
黑人	0.70	0.71	0.78
海滩	0.60	0.58	0.65
建筑	0.51	0.56	0.61
公汽	0.81	0.86	0.91
恐龙	0.91	0.91	1.00
大象	0.56	0.55	0.64
花	0.69	0.77	0.89
马	0.81	0.86	0.91
高山	0.35	0.33	0.42
食物	0.40	0.43	0.46

由于 AvgP 既考虑到相关图像在结果序列中的序号, 又考虑到不相关图像的序号, 所以它对检索结果的刻画相比于精度而言更全面, 比一般的相关图像平均排序指标也更全面。各种方法的准确率如图 4 所示。

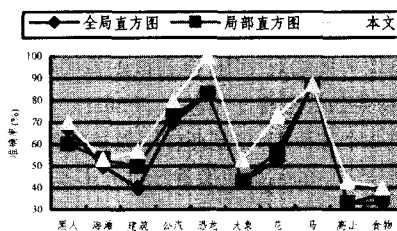
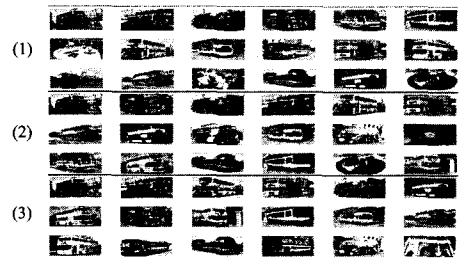


图 4 试验准确率

应该注意到, 检索准确率只注意检出的相关图像与全部检出图像的比率, 与图像排序位置无关。所以, 准确率与参数 AvgP 存在不一致性, 即有时某种方法准确率比其它方法高但 AvgP 反而低, 反之亦然。

总体来说, 恐龙图像比较简单, 整幅图像可以很清楚地分为前景和背景两部分, 从而各种方法都有较高的检索准确率; 而建筑、高山、食物等主题的图像构成较复杂, 所含对象也较多, 提取的特征并不能很好地表征图像, 所以各方法准确率也较低; 而其它语义类的图像复杂度介于二者之间, 准确率也介于它们之间。按方法来说, 由于并没有对这些图像进行旋转、平移等变形, 全局直方图相对于分块直方图鲁棒性好的优点没表现出来, 故分块直方图法的效果要好于全局直方图法, 而本文方法的检索效果又普遍好于以上两种方法。在检索恐龙时, 本文方法的准确率甚至达到了 100%, 比其它方法的精度提高差不多 20%。

图 5 是检索例图为红色公汽时各种方法的检索结果(由于篇幅限制, 每种方法只列出前 18 幅图像的检索结果)。



(1)为全局直方图法; (2)为分块直方图法; (3)为本文方法。

图 5 检索结果

在图 4 的检索结果中, 全局直方图法由于对图像刻画得不够细致, 导致误检率较高, 前 18 幅图像中有 4 幅不相关; 局部直方图由于结合了位置元素, 对图像描述更加充分, 从而效果有一定的提高, 前 18 幅只有 2 幅不相关; 而在本文方法中, 既引入了绝对位置, 又引入了相对位置, 颜色量化也更科学, 同时结合了图像的粗略纹理特征, 对图像的刻画也更精细和全面, 从而效果也进一步提高, 前 18 幅图像只有 1 幅不相关, 位置也很靠后, 结果较令人满意。

**结束语** 本文提出了一种基于颜色和纹理的图像检索方法, 该方法先将 RGB 颜色空间转换到 HSV 颜色空间并进行模糊量化, 再提取分块颜色直方图, 颜色相邻矩阵和纹理特征矩阵作为图像特征向量。实验证明, 其检索的效果比单用全局颜色直方图和分块颜色直方图效果都好, 这主要是它在引入颜色空间位置分布信息的同时又兼顾了图像变形的鲁棒性, 在二者之间取得了很好的平衡。

但该方法也有它的一些不足之处, 如它对图像的纹理描述比较粗糙, 对复杂图像的描述还是不够准确, 从而导致对复杂图像的检索准确率较低等。本项研究的下一步工作将针对这些缺点, 提高特征向量对图像内容描述的准确性, 并通过引入相关反馈技术把用户的习惯、爱好、情感等主观或感性化的因素考虑进来, 以进一步提高检索精度。

## 参考文献

- [1] Swain M J, Ballard D H. Color indexing [J]. International Journal of Computer Vision, 1991, 7(1): 11-32
- [2] Hafner J, Sawhney H S, et al. Efficient color histogram indexing for quadratic form distance functions [J]. IEEE Trans. On Pattern Analysis and Machine Intelligence, 1995, 17(7): 729-736
- [3] Stricker M, Orengo M. Similarity of color images [C] // IS&T/SPIE Conf. on Storage and Retrieval for Image and Video Database 3, Vol 2420, San Jose, CA, Feb. 1995: 381-392
- [4] 安志勇, 赵珊, 周利华. 基于形状和纹理的图像检索 [J]. 计算机科学, 2006, 33(11): 225-228
- [5] Haralick R, Shanmugan K, Dinstein I. Textural features for image classification [C] // IEEE Trans System Man Cybernetics, 1973, 3: 610-621
- [6] Tamura H, Mori S, Yamawaki T. Textural features corresponding to visual perception [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1978, 8(6): 460-472
- [7] 王涛, 胡事民, 孙家广. 基于颜色-空间特征的图像检索 [J]. 软件学报, 13(10): 2031-2036
- [8] 孙君顶, 张喜民, 崔江涛, 等. 一种新的基于颜色和空间特征的图像检索方法 [J]. 计算机科学, 2005, 32(6): 158-160
- [9] <http://wang.ist.psu.edu/docs/related.shtml>