

# 基于局部颜色空间特征的图像检索

汪慧兰<sup>1,2</sup>, 赵海峰<sup>1</sup>, 罗 斌<sup>1</sup>

(1. 安徽大学 计算机科学与技术学院, 安徽 合肥 230039;

2. 安徽师范大学, 安徽 芜湖 241000)

**摘 要:** 基于图像中主要物体的尺寸动态地对图像分块来获得颜色的空间分布特征, 并选取了符合人眼感知特性的 HSV 颜色空间进行非均匀量化, 提取各个分块的主色及其所占的百分比作为颜色综合特征矢量, 根据所提出的综合特征矢量给出相应的相似度计算方法, 同时构造出相似矩阵进行基于颜色内容的图像检索, 能通过调整各个分块的权值来进一步提高检索的精度, 最后将文中提出的方法与传统的全局直方图方法相比较, 可知文中的方法因为利用了动态分块的方法, 考虑了颜色的空间信息以及权值的调整, 从而取得了比较好的检索结果。

**关键词:** 主色; 空间关系; 相似矩阵; 图像检索

**中图分类号:** TP391.41; G354.4

**文献标识码:** A

**文章编号:** 1005-3751(2006)01-0076-04

## Image Retrieval Based on Local Color and Spatial Features

WANG Hui-lan<sup>1,2</sup>, ZHAO Hai-feng<sup>1</sup>, LUO Bin<sup>1</sup>

(1. School of Computer Science and Technology, Anhui University, Hefei 230039, China;

2. Anhui Normal University, Wuhu 241000, China)

**Abstract:** In this paper, image color distribution is obtained based on dynamic sub-block splitting. The image is split into sub-blocks according to the size of object in the query image. Adopt non-uniform quantization of HSV color space, which is in accordance with human visual perception. Combined color feature vector consists of dominant colors and their percentage in each sub-block. A similarity measure is defined for the combined color feature vector. Content-based image retrieval is achieved by using a constructed similarity matrix. It can raise the efficiency of retrieval by adjusting sub-block weight value. Compare the proposed method with traditional global histogram method. Based on the dynamic splitting, spatial information and weight value, better retrieval results are obtained.

**Key words:** dominant color; spatial relationship; similarity matrix; image retrieval

### 0 引 言

随着多媒体和网络技术的迅猛发展, 每天都会有成千上万的新图像产生。为了能够从大量图像数据中快速、准确地找到用户自己所需的图像, 图像检索的研究自从 1970 年以来就成了人们研究的热点。而传统的图像检索大多是基于关键字、文本的方法来管理和检索图像的, 当图像的数量非常庞大时, 对图像进行文字描述的工作量也很大, 而且文字描述的主观性和不准确性则导致不准确和不完全的检索结果, 为了克服这些存在的问题, 基于内容的图像检索(CBIR)技术正越来越成为人们研究的焦点。目前, 比较著名的 CBIR 系统有: QBIC, VisualSeek, PhotoBook, Virage, Retrieval ware, MARS, Netra 等等, 这些系统主要使用图像的颜色、纹理、形状和空间关系等特征进行

检索。

因为颜色具有与生俱来的旋转不变性和尺度不变性, 因此在基于内容的图像检索中, 颜色是使用最广泛的特征之一。从 1991 年 Swain 和 Ballard<sup>[1]</sup>提出了使用颜色直方图描述彩色图像和颜色直方图相交方法进行相似度量以来, 很多研究方法都将颜色直方图作为一种常用的特征矢量用于 CBIR 的研究。但是两个内容很不相同的图像可能具有相同的颜色直方图, 这主要是由于仅用颜色直方图时不能反映颜色的空间分布导致的。为了克服全局颜色直方图的这种缺陷, 改进颜色直方图特征的方法包括: 累加直方图方法<sup>[2]</sup>, 这种方法通过累加方式增加了直方图的鲁棒性, 但仍丢失了颜色的空间分布信息; 后来为了增加颜色特征的空间信息, 提出了简单的局部直方图方法, 将图像进行固定分块, 计算每一个子块的颜色直方图, 然后比较两幅图像相应子块间的相似距离并进行求和作为整幅图像的相似距离值, 但子块间的空间约束关系太强, 丢失了全局颜色直方图所具有的旋转、平移不变性的优点; Smith and Chang<sup>[3]</sup>提出了二进制颜色集(binary color sets), Stricker<sup>[4]</sup>提出了颜色矩(color moments)等等, 另外

收稿日期: 2005-04-12

基金项目: 教育部“优秀青年教师资助计划”项目(教人司[2003]355号); 安徽大学 211 工程学术创新团队资助

作者简介: 汪慧兰(1978—), 女, 安徽怀宁人, 硕士研究生, 研究方向为基于内容的视觉信息检索、模式识别等。

还有颜色相关矢量 CCV(color coherence vector)<sup>[5]</sup>和颜色关联图 Color Correlograms<sup>[6]</sup>方法等等,但这些方法仍没有准确地描述图像中物体的颜色、位置等信息。另外还有方法是利用颜色—空间特征进行检索<sup>[7]</sup>,这种方法首先通过图像分割得到图像中的主要物体,然后提取每个物体的颜色和位置特征,最后进行图像间相似度的计算,但由于图像分割对任意图像很难准确地找到感兴趣物体的范围,因此实验的结果往往不够理想。

一般来说,有 3 种方式划分图像<sup>[8]</sup>:固定方式,半动态方式和动态方式。因为图像中检索物体的尺寸是决定划分图像的重要因素,因此文中采用了依据图像中主要物体的尺寸动态地对图像分块的方法来获取颜色的空间分布信息,并选取符合人眼感知特性的 HSV 颜色空间进行非等间隔量化,然后提取各个分块的主色及其所占的百分比作为颜色综合特征矢量,以提高检索的效率,再根据综合特征矢量进行相应的相似度计算,构造相似矩阵,进行基于内容的图像检索,并能通过调整各个分块的权值来进一步提高检索的精度。

## 1 基于综合特征的图像检索

基于内容的图像检索过程主要有 3 个方面:

- (1) 选取恰当的图像特征,如:颜色、纹理、形状等等;
- (2) 有效的特征提取方法;
- (3) 准确的特征匹配算法(相似度量)。

文中选取了基于颜色的综合特征进行图像检索,其具体的检索步骤如下。

### 1.1 颜色空间的选取与量化

RGB 颜色空间是一种常见的颜色表示法,但是 RGB 颜色空间与人眼的感知差异很大。例如,距离为 50 的(0, 0, 0)与(50, 0, 0)两种 RGB 颜色认为是同一黑色,而距离为 50 的(200, 150, 0)和(200, 200, 0)则是差别很大的两种颜色(黄色和绿色),也就是说其空间的相似不代表实际颜色的相似。因此为了使得算法对于颜色距离的定义符合人的视觉特征,采用了较能符合人眼感知特性的 HSV 颜色空间(色调 H、饱和度 S 和亮度 V)。由于 HSV 颜色空间几乎是均匀的,因此,两个 HSV 颜色的相似度可以表达为它们在 HSV 空间的相似性。又因为一幅 24bit 的 RGB 真彩色图像可能包括 16 777 216 种颜色,由于人眼不可能分辨出如此多种的颜色,同时为了提高检索的效率,笔者在 HSV 空间中对图像进行了量化。

本过程的具体步骤:首先将图像中每一像素的 r, g, b 值转换为 h, s, v 值( $h \in [0, 2\pi]$ ,  $s \in [0, 1]$ ,  $v \in [0, 1]$ ),然后对 HSV 颜色空间进行非均匀量化得到 256 种代表颜色。

具体的颜色量化和编码方法如下<sup>[9]</sup>:

$$H = \begin{cases} 0, & \text{if } h \in [345, 15] \\ 1, & \text{if } h \in [15, 25] \\ 2, & \text{if } h \in [25, 45] \\ 3, & \text{if } h \in [45, 55] \\ 4, & \text{if } h \in [55, 80] \\ 5, & \text{if } h \in [80, 108] \\ 6, & \text{if } h \in [108, 140] \\ 7, & \text{if } h \in [140, 165] \\ 8, & \text{if } h \in [165, 190] \\ 9, & \text{if } h \in [190, 220] \\ 10, & \text{if } h \in [220, 155] \\ 11, & \text{if } h \in [255, 275] \\ 12, & \text{if } h \in [275, 290] \\ 13, & \text{if } h \in [290, 316] \\ 14, & \text{if } h \in [316, 330] \\ 15, & \text{if } h \in [330, 345] \end{cases} \quad S = \begin{cases} 0, & \text{if } s \in [0, 0.15] \\ 1, & \text{if } s \in [0.15, 0.4] \\ 2, & \text{if } s \in [0.4, 0.75] \\ 3, & \text{if } s \in [0.75, 1] \end{cases} \quad V = \begin{cases} 0, & \text{if } v \in [0, 0.15] \\ 1, & \text{if } v \in [0.15, 0.4] \\ 2, & \text{if } v \in [0.4, 0.75] \\ 3, & \text{if } v \in [0.75, 1] \end{cases}$$

在 HSV 颜色空间中按照上面的量化方法得到了 256 种代表颜色,每种颜色用相应的平均颜色来表示。为了减小量化噪声和局部纹理特征对检索结果的影响,采用了中值滤波对量化后的图像进行平滑。

### 1.2 提取局部颜色空间特征

(1) 对 HSV 空间进行量化后,将 HSV 合成一维特征矢量:

$$L = 16H + 4S + V \quad (1)$$

这样 H, S, V 三个分量就在一维矢量上分布开来。在式(1)中,对色调 H 取的权重为 16,对饱和度 S 取的权重为 4,对亮度 V 取的权重为 1,这是因为色调包含了绝大多数的信息,同时减轻了图像亮度和饱和度对检索结果的影响,于是对颜色分布不同的图像能很好地检索出来,故能充分利用图像的颜色信息特征进行检索。

(2) 根据检索图像中物体的尺寸动态地对图像进行分块,然后提取各个分块的主要颜色及其对应的百分比,则第 j 分块的特征矢量为:

$$f_j = \{L_i, P_i\}, i = 1, 2, \dots, N \quad (2)$$

其中 N 代表各个分块提取的颜色总数,  $L_i$  是一维颜色特征向量,通过颜色查找表可得到相应的 HSV 空间中的三维颜色矢量( $h_i, s_i, v_i$ ),  $P_i$  是  $L_i$  占相应分块的百分比,并且  $\sum_i P_i = 1$ , N 对于各个分块其值是不同的。

对于每一幅图像能得到综合特征向量  $F = (f_1, f_2, \dots, f_m)$ , 其中 m 为图像的动态分块数。

### 1.3 相似度计算

根据式(2),设  $f_{q1} = \{L_i, P_i\}, i = 1, 2, \dots, N_1\}$ ,  $f_{d1} = \{L_j, P_j\}, j = 1, 2, \dots, N_2\}$  是两个图像块的特征矢量,则  $f_{q1}$  与  $f_{d1}$  之间的相似距离度量可定义为<sup>[10]</sup>:

$$D^2(f_{q1}, f_{d1}) = \sum_{i=1}^{N_1} P_i^2 + \sum_{j=1}^{N_2} P_j^2 - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2A_{ij}P_iP_j \quad (3)$$

$$\text{式中 } A_{ij} = \begin{cases} 1 - d_{ij}/d_{\max}, & d_{ij} \leq T_d \\ 0, & d_{ij} > T_d \end{cases} \quad (4)$$

$d_{\max} = \alpha T_d$  ( $T_d$  为两颜色相似的最大距离,  $\alpha$  通过实

验设定)

其中在 HSV 颜色空间中<sup>[11]</sup>:

$$d_{ij} = [(V_i - V_j)^2 + (S_i \cos H_i - S_j \cos H_j)^2 + (S_i \sin H_i - S_j \sin H_j)^2]^{\frac{1}{2}}$$

$$d_{\max} = \sqrt{5} \quad (5)$$

总相似度  $S$  计算的具体步骤:

① 设待检索图像与数据库图像的综合特征向量分别为  $F_q = (q_1, q_2, \dots, q_m)$ ,  $F_p = (p_1, p_2, \dots, p_n)$ , 计算相似度矩阵  $S = [S_{i,j}]_{m \times n}$ , 其中  $S_{i,j}$  由式(3)定义。即:

$$S = \begin{bmatrix} S_{1,1} & \dots & S_{1,n} \\ \vdots & & \vdots \\ S_{m,1} & \dots & S_{m,n} \end{bmatrix} \quad (6)$$

② 对于  $S_{i,j}$  小于某一阈值的置 0, 以减少完全不相似子图对检索结果的影响。

③ 查找检索图像各块与数据库图像中最为相似的分块, 即计算相似矩阵  $S$  各行的最大值  $S_{\max}(i)$ , 并记录最大值  $S_{\max}$  所对应的数据库图像的子图像的序号, 然后将相似矩阵  $S$  相应的列置 0, 这是为了排除数据库图像中已经比较过的子图像的干扰。对检索图像中所有的子块重复此步骤。

④ 计算两图像之间的总的相似度  $S$ :

$$S(p, q) = \sum_{i=1}^m W(i) * S_{\max}(i)$$

$$(\sum_i W(i) = 1)$$

其中  $W$  为权重矩阵, 文中根据各种类型图像的特点设定了一系列的权重矩阵  $W$  并进行了实验比较。

⑤ 按  $S$  由大到小的顺序排列从数据库中检索到的图像,  $S$  越大, 说明两图像越相似。

## 2 实验结果分析

使用文中所提出的检索方法, 笔者进行了基于内容的图像检索, 图像内容包括: 自然风光、花朵、动物、汽车等 300 幅(分辨率为  $600 \times 400$ )。按上述步骤提取检索图像的特征, 在实验中, 根据图像检索内容的需要, 进行动态分块, 如对于花分成 25 块, sunset 分成 49 块,  $T_d$  设为 0.1, 而对于分块中百分比小于 10% 的颜色, 按式(5)计算使其归入最相似的颜色, 特别地当分块中全部百分比都小于 10% 的颜色, 取其平均颜色。一般情况下每个分块选取 3、4 个颜色就能表达其特征, 按式(6)构造出相似矩阵, 并将其与预先计算好并保存在图像数据库中的特征矢量进行相似匹配, 得到检索结果。

下面的实验是取两例图像按其相似程度排列的前 10 幅图像检索的结果, 其中第一幅为待检索图像, sim 是各图像与待检索图像的相似度值。

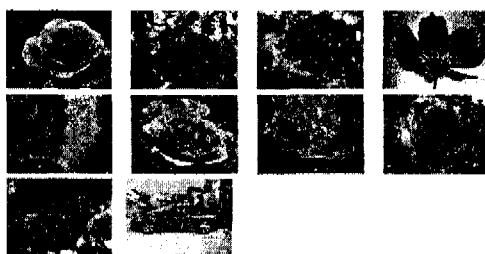


sim:	1.000	0.9523	0.8762	0.8634
	0.8562	0.9392	0.7943	0.7856
	0.7683	0.7216		

图 1 传统的基于颜色直方图特征的检索结果

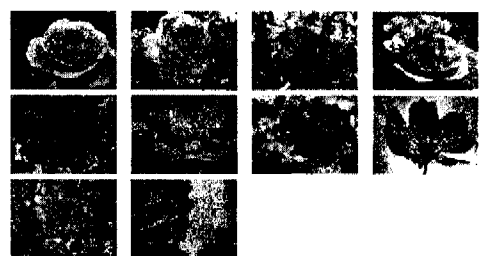


sim:	1.000	0.8622	0.8492	0.8368
	0.8292	0.8289	0.8224	0.7970
	0.7953	0.6042		



sim:	1.0000	0.7572	0.7291	0.6722
	0.6695	0.6684	0.6204	0.5818
	0.5477	0.5436		

图 2 基于文中的方法(各块权值相同时)的检索结果



sim:	1.000	0.9628	0.9462	0.8934
	0.8762	0.8526	0.7834	0.7556
	0.7283	0.6916		

图 3 将各分块调整为不同权值时的检索结果

为了对于检索结果作一个较为客观的评价, 对正确率进行了比较, 其定义如下<sup>[12]</sup>:

$$\text{正确率} = \frac{\text{有关联的正确检索结果}}{\text{所有检索到的结果}}$$

图 1 是传统的基于颜色直方图特征的对于上面是蓝天、中间为房子和树的图像的检索结果, 因为其没有考虑空间信息, 前 10 幅的正确率为 50%。图 2 进行了两组数

据的比较,分别为花和房子,它是文中的方法在各分块权值相同时的检索结果,前 10 幅的正确率为 70%,从而取得在颜色空间分布上更为相似的图像,比颜色直方图方法提高了检索精度。但有时当想要检索图像库中是否有与示例图像中的某一物体颜色相近的图像时,则文中提出的方法就可以根据需要调整权值。从图 3 可知当需要检索图像库中是否有红花的图像时,此方法明显取得了更符合人感知的检索结果。而且由于相似矩阵的引入,此方法还有对平移、旋转不敏感的优点。

### 3 结 论

基于局部颜色空间关系的图像检索方法,是利用动态分块的方法得到关于图像颜色的综合特征矢量,再通过相似度矩阵的计算进行检索,并通过调整分块的权值获得与人的感知相近的检索结果,从而提高了检索精度。实验表明,文中提出的算法能解决分块的局部颜色特征对平移、旋转敏感的缺陷,当然这仅仅是从颜色特征来进行图像检索的,且权值的调整不能进行人机交互,这有待于改进。以后的发展趋势应当是使用图像的多个特征如颜色、纹理、形状、拓扑和语义等相结合的方法和基于人机交互的相关反馈技术来进一步提高基于内容图像检索的精度。

#### 参考文献:

- [1] Swain M, Ballard D. Color indexing[J]. International Journal of Computer Vision, 1991, 7(1):11-32.
- [2] Stricker M, Orengo M. Similarity of color images[A]. In: Proceedings of SPIE Storage and Retrieval for Image and Video Databases III[C]. San Jose, CA: [s. n.], 1995, 2420: 381-392.
- [3] Smith J R, Chang S F. Tools and techniques for color image retrieval[A]. In: Proc SPIE of the Storage & Retrieval for

Image and Video Databases IV[C]. San Jose, CA: [s. n.], 1996, 2670: 426-437.

- [4] Stricker M, Dimai A. Color indexing with weak spatial constraints[A]. In: Proc SPIE Storage Retrieval Still Image Video Databases IV[C]. San Jose, CA: [s. n.], 1996, 2670: 29-40.
- [5] Pass G, Zabih R. Histogram refinement for content based image retrieval[A]. In: Proc IEEE Workshop on Applications of Computer Vision[C]. Washington, DC, USA: IEEE Computer Society, 1996. 96-102.
- [6] Huang J, Kumar S R. Image indexing using color correlograms [A]. In: Proc IEEE Conf on Computer Vision and Pattern Recognition[C]. San Juan, Puerto Rico: IEEE Computer Society, 1997. 762-768.
- [7] Messer K, Kittler J. A region based image database system using color and texture[J]. Pattern Recognition Letters, 1999 (20):1323-1330.
- [8] Li Xue long. Image retrieval based on perceptive weighted color blocks[J]. Pattern Recognition Letters, 2003, 24: 1935-1941.
- [9] 田玉敏,林高全. 基于颜色特征的彩色图像检索方法[J]. 西安电子科技大学学报(自然科学版), 2002, 29(1):43-46.
- [10] Yining D, Manjunath B S. An efficient color representation for image retrieval[J]. IEEE transaction on image processing, 2001, 10(1):140-147.
- [11] Smith J R, Chang S F. Visualseek: a fully automated content-based image query system[A]. In: Proc ACM Multimedia 96[C]. Boston, MA: [s. n.], 1996. 147-156.
- [12] Patel N V, Sethi I S. Video shot detection and characterization for video database[J]. Pattern Recognition, 1997(30):583-592.

(上接第 75 页)

Agent 技术提出了系统的实现模型,通过多 Agent 的合作与协商,最终产生符合用户要求的学习对象列表。

随着 SCORM 规范被更为广泛的认可,符合 SCORM 规范的学习资源以及学习内容管理系统将越来越多,如何实现 SCORM 规范的目标,实现学习资源最大限度的共享,在众多学习资源和学习者之间建立沟通的桥梁,将会成为越来越突出的问题。本模型对于这个问题的解决作出了有益的尝试。

以后的工作可以将该模型拓展到在多个不同的符合 SCORM 规范的学习内容管理系统 LCMS(Learning Content Management System)的学习对象中间进行检索和匹配。匹配代理将试图在多个代表不同 LCMS 的学习资源代理中寻求匹配的学习内容。

#### 参考文献:

- [1] ADL Technical Team. Sharable Content Object Reference Model (SCORM) 2004 2nd Edition[EB/OL]. <http://www.adlnet.org>, 2004.
- [2] IMS Global Learning Consortium, Inc. IMS Content Packaging Information Model [EB/OL]. <http://www.imsglobal.org/content/packaging/cpv1p1p4/imscp-infov1p1p4.html>. 2004-10-04
- [3] 马利霞,曹树金. 描述教育信息资源的元数据标准[J]. 大学图书馆学报, 2004(2):5-9.
- [4] Wooldridge M. An Introduction to MultiAgent Systems[M]. New York: John Wiley & Sons, Inc, 2002.
- [5] Zambonelli F, Jennings N R, Wooldridge M J. Developing Multiagent Systems: the Gaia Methodology[J]. ACM Transactions on Software Engineering and Methodology, 2003, 12(3):317-370.