

一类基于局部信息的特征描述子的图像检索

黄朝兵 余胜生 周敬利

(华中科技大学计算机科学技术学院信息存储系统国家实验室 武汉430074)

摘要 基于内容的图像检索准确性大大依赖于低层可视特征的描述。本文提出一类创新的彩色图像空间描述子、纹理描述子、边缘描述子和颜色描述子,空间描述子由局部均值直方图表示,纹理描述子由局部方向差单元直方图表示,边缘描述子由局部极大-极小差直方图表示,颜色描述子由量化 HSV 模型颜色直方图表示。这四种描述子被用作特征索引,它们对彩色图像,尤其是对具有相对规则的结构或纹理特征的图像具有很强的描述力。实验结果表明,用这种特征索引来检索图像,可以得到比其它基于颜色-空间方法高得多的精确度。

关键词 基于内容的图像检索,颜色直方图,局部均值直方图,局部方向差单元直方图,局部极大-极小差直方图

A Novel Feature Descriptor Based Local Information for Image Retrieval

HUANG Chao-Bing YU Sheng-Sheng ZHOU Jing-Li

(National Storage System Laboratory, College of Computer Science & Technology, Huazhong University of Science & Technology, Wuhan 430074)

Abstract The accuracy of content-based image retrieval depends greatly on the description of low-level visual features. In this paper, a novel spatial, texture, edge, and color descriptor for color image are proposed. Spatial descriptor is represented by local mean histogram, texture descriptor is represented by local directional difference unit histogram, edge descriptor is represented by local max-min difference histogram, color descriptor is represented by color histogram under the quantized HSV model. These four descriptors are used as feature indexes, which have powerful descriptive power for the color images, especially for images with relatively regular structure characteristic or texture characteristic. Experimental results show that the retrieval method using the feature indexes can achieve much higher precision than other color-spatial based methods.

Keywords Content based image retrieval (CBIR), Color histogram, Local mean histogram, Local directional difference unit histogram, Local max-min difference histogram

1 引言

随着多媒体、通信技术和 Internet 的飞速发展,多媒体数据规模急剧增长,为了帮助用户快速准确地找到感兴趣的图像或视频,基于内容的图像检索(Content-Based Image Retrieval, CBIR)技术成为人们研究的焦点。经过十年来的努力, CBIR 技术的研究和实现已经取得了很多成果^[1~3],它们主要利用图像的颜色、纹理、形状、空间等低层特征及其任意组合进行检索,检索准确性大大依赖于低层可视特征的描述,利用相关反馈技术和高层语义特征进行检索是 CBIR 技术的一个发展方向。高层语义特征由低层可视特征导出,利用高层语义特征进行图像检索的精度仍然依赖于低层可视特征的描述和表达。

基于颜色直方图方法^[4]计算简单,对图像的平移和旋转具有不变性,但是颜色直方图反映的是颜色的全局统计信息,丢失了颜色的空间分布信息,具有相似直方图的图像可能外形完全不同。累加直方图方法由文[5]提出,它对量化参数更鲁棒,但是它也没有反映颜色的空间分布信息。为了较全面地描述图像的颜色、空间和纹理等方面的性质,人们提出了很多

方法。在颜色直方图中增加空间信息,即基于颜色-空间特征的方法中,较有代表性的如文[6~10]。文[6]提出颜色相关向量方法(Color Coherence Vector-CCV),将像素被分类成相关或不相关,以相关和不相关的像素颜色直方图作为特征索引。文[7]提出颜色相关图法(Color Correlogram),以同一颜色或不同颜色间按照不同距离在图像中的分布为特征索引。文[8]提出环形直方图方法(Annular Color Histogram-ACH),对每一个颜色值,计算其质心,再以质心为圆心,按不同的半径作圆,将圆分成圆环,以各个圆环内的像素的颜色直方图为特征索引。文[9]提出空间彩色直方图方法(Spatial-Chromatic Histogram-SCH),对任一颜色值 i ,计算与其质心的标准偏差 $\sigma(i)$,以所有 $\sigma(i)$ 为特征索引。文[10]提出地理统计方法(Geographical Statistics-Geostat),对任一颜色值 i ,计算与其质心的松散度 $L(i)$,以所有 $L(i)$ 为特征索引。文[7~10]比文[6]更能充分反映图像的颜色-空间特征,但它们对图像旋转、平移敏感。常用于表现图像纹理特征的方法有统计方法、结构方法和频谱分析方法。文[12]发现 Gabor 小波变换符合人类的视觉特征,但是提取图像的 Gabor 纹理计算量很大,耗时。文[13]利用图像纹理谱描述方法来刻画图像像素点在邻域内

黄朝兵 博士研究生,高级工程师,主要研究方向为基于内容的图像检索,多媒体信息处理与通信。余胜生 教授,博士生导师,主要研究方向为多媒体信息处理与通信,附网存储和存储区域网络技术。周敬利 教授,博士生导师,主要研究方向为多媒体信息处理与通信,附网存储和存储区域网络技术。

灰度的变化情况,计算复杂度减少,但特征索引维数太高,而且仅仅用到灰度信息,没有利用颜色信息。

为了全面反映图像的特征,本文提出一种创新的颜色、空间、纹理和边缘描述子。第2节给出这些新的描述子。第3节用这些描述子作为特征索引检索图像。第4节给出实验结果。最后得出结论。

2 创新特征描述子

我们知道仅用一种低层可视特征不足以区分图像。直方图形式的特征与图像的位置无关,从而对图像的平移和旋转不变,归一化后对图像的尺度不变;而且直方图没有度量。为此,本文提出创新的直方图形式的空间、纹理和边缘描述子。

2.1 颜色描述子

一般认为 RGB 颜色空间与人眼的感知差异较大,另外,人很难区分两个相近 RGB 值,描述图像内容通常使用术语,比如红色、深黄等,而不用 RGB 值;而且我们倾向于集中记住高亮度饱和区域和大的颜色区域。本文采用较好地符合人眼感知特性的 HSV(hue, saturation, value)颜色模型,并对其进行符合人的视觉的非均匀量化处理。

RGB 颜色空间到 HSV 颜色模型的转换参考文[11],转换得到的 h, s, v 值,其取值范围为: $h \in [0, 2\pi], s \in [0, 1], v \in [0, 1]$ 。

对 HSV 空间进行非均匀量化得到166种代表颜色,具体颜色量化和编码如下:

(1)对于 $v < 25\%$ 的颜色均归入黑色,其编码为 $q=0$ 。

(2)对于 $s < 20\%$ 且 $v \geq 25\%$ 的颜色按亮度 v 值均匀量化成三个值,分别为深灰、浅灰、白色,其对应的编码分别为: $q=1, 2, 3$ 。

(3)其它的颜色($s \geq 0.2$ 且 $v \geq 0.25$)认为是彩色,将 $h \in [0, 2\pi]$ 均匀量化成18个值,对应取值 $H=0, 1, \dots, 17, s \in [0.2, 1]$ 均匀量化成3个值,对应取值 $S=0, 1, 2, v \in [0.25, 1]$ 均匀量化成3个值,对应取值 $V=0, 1, 2$, 这样共得到162种彩色,其编码为: $q=9 \times H + 3 \times S + V + 4$ 。

这样图像中的每个像素值 q 都可用0到165中的整数码值来表示,记 $Q=\{0, 1, 2, \dots, 165\}$ 。则有: $q \in Q$ 。

给定一幅大小为 $m \times n$ 的图像 f , 设它在坐标为 (i, j) 处的像素点的 HSV 量化编码值为 q , 即 $q=f(i, j) \in Q$, 那么计算量化颜色值在 Q 中的分布,得到归一化的量化颜色直方图为:

$$H_c(q) = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \delta(f(i, j) - q), \forall q \in Q$$

其中 $\delta(x)$ 是单位冲击函数。

2.2 空间描述子

我们用邻域的颜色均值来反映图像的空间信息。本文以像素点的 3×3 邻域为单位, 计算每个点的邻域均值, 以邻域均值的直方图作为空间描述子。

给定一幅大小为 $m \times n$ 的图像 f , 设它在坐标为 (i, j) 处像素点的 HSV 量化编码值为 $f(i, j)$, 则其 3×3 邻域的均值为:

$$e(i, j) = \frac{1}{9} \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{j+1} f(k, l)$$

对于 $\forall f(i, j) \in Q$, 有 $0 \leq f(i, j) \leq 165$, 则由上述定义, 我们可以得到 $0 \leq e(i, j) \leq 165$, 对所有 $i=0, 1, \dots, m, j=0, 1, \dots, n$, 都成立。再对 $e(i, j)$ 执行四舍五入取整, 那么它

们仍然属于 Q 。

$e(i, j)$ 是图像在点 (i, j) 处的 3×3 邻域的颜色均值($i=0, 1, \dots, m, j=0, 1, \dots, n$), 它反映了图像在该点的局部空间信息, 那么, 对整个图像的局部均值进行统计, 将揭示图像的全局空间信息。类似于2.1节的定义, 归一化的像素点 3×3 邻域的局部均值直方图为:

$$H_{ne}(e) = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \delta(e(i, j) - e), \forall e \in Q$$

2.3 纹理描述子

图像的纹理表示图像的颜色在某个邻域内的一种变化, 表现为图像的不同方向上的变化度量。文[13]提出纹理谱的概念, 它认为, 纹理图像可以看作是纹理单元的集合, 纹理单元标识给定像素及其邻域的局部纹理性质。对整个图像的纹理单元进行统计, 将揭示图像的全局纹理性质。借鉴这一思想, 我们提出一种新的纹理描述子——局部方向差单元直方图。

考虑像素点 (i, j) 的 3×3 邻域, 如图1所示, 其中 $y(i, j)$ 表示图像在该像素点的灰度(24-bit 真彩色图像的 RGB 值到灰度的转换为: $y=0.299R + 0.587G + 0.114B$), $h(i, j)$ 表示图像在该像素点的未经量化的色度值(hue), 可由2.1节得到。

$y(i-1, j-1),$ $h(i-1, j-1)$	$y(i, j-1), h(i, j-1)$	$y(i+1, j-1),$ $h(i+1, j-1)$
$y(i-1, j),$ $h(i-1, j)$	$y(i, j), h(i, j)$	$y(i+1, j),$ $h(i+1, j)$
$y(i-1, j+1),$ $h(i-1, j+1)$	$y(i, j+1), h(i, j+1)$	$y(i+1, j+1),$ $h(i+1, j+1)$

图1 图像像素的 3×3 邻域表示

我们用二进制序列 $R_k(k=0, 1, 2, \dots, 7)$ 来刻画该点邻域像素值在八个方向上的显著变化, 它们定义如下:

(1)水平方向1: if $|y(i+1, j-1) - y(i-1, j-1)| > BTH$ or $|h(i+1, j-1) - h(i-1, j-1)| > HTH, R_0=1$; otherwise $R_0=0$ 。

(2)水平方向2: if $|y(i+1, j) - y(i-1, j)| > BTH$ or $|h(i+1, j) - h(i-1, j)| > HTH, R_1=1$; otherwise $R_1=0$ 。

(3)水平方向3: if $|y(i+1, j+1) - y(i-1, j+1)| > BTH$ or $|h(i+1, j+1) - h(i-1, j+1)| > HTH, R_2=1$; otherwise $R_2=0$ 。

(4)垂直方向1: if $|y(i-1, j-1) - y(i-1, j+1)| > BTH$ or $|h(i-1, j-1) - h(i-1, j+1)| > HTH, R_3=1$; otherwise $R_3=0$ 。

(5)垂直方向2: if $|y(i, j-1) - y(i, j+1)| > BTH$ or $|h(i, j-1) - h(i, j+1)| > HTH, R_4=1$; otherwise $R_4=0$ 。

(6)垂直方向3: if $|y(i+1, j-1) - y(i+1, j+1)| > BTH$ or $|h(i+1, j-1) - h(i+1, j+1)| > HTH, R_5=1$; otherwise $R_5=0$ 。

(7)对角方向1: if $|y(i+1, j-1) - y(i-1, j+1)| > BTH$ or $|h(i+1, j-1) - h(i-1, j+1)| > HTH, R_6=1$; otherwise $R_6=0$ 。

(8)对角方向2: if $|y(i-1, j-1) - y(i+1, j+1)| > BTH$ or $|h(i-1, j-1) - h(i+1, j+1)| > HTH, R_7=1$; otherwise $R_7=0$ 。

其中, BTH 、 HTH 分别为灰度阈值、色度阈值, 都为事先给定的正常数。由于每个 $R_k(k=0, 1, 2, \dots, 7)$ 都有两种可能的值, 那么这八个的所有组合将得到总共 $2^8=256$ 种可能的

值,我们用这256种可能的值来定义像素点 (i,j) 处的局部方向差单元如下:

$$r(i,j) = \sum_{k=0}^3 2^k \times R_k$$

对于每个像素点 (i,j) ,我们都可以得到它的局部方向差单元 $r(i,j)$,其取值范围是: $R=\{0,1,2,\dots,255\}$,统计 $r(i,j)$ 在 R 上的分布 $(i=0,1,\dots,m,j=0,1,\dots,n)$,得到归一化的局部方向差单元直方图为:

$$H_{ar}(r) = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \delta(r, (i,j) - r), \forall r \in R$$

2.4 边缘描述子

边缘是图像的一个基本特征,文[14]采用某个像素集合内的像素值的极大值与极小值的差值来度量其边缘程度。本文给出一个新的边缘描述子。以 $N(i,j)$ 表示像素点 (i,j) 的 3×3 邻域,计算该邻域内量化颜色值的极大值和极小值的差值,即:

$$d(i,j) = \max_{\substack{i-1 \leq k \leq i+1 \\ j-1 \leq l \leq j+1}} f(k,l) - \min_{\substack{i-1 \leq k \leq i+1 \\ j-1 \leq l \leq j+1}} f(k,l)$$

如果像素点邻域内没有边缘的话,则它们的颜色值会相互接近,从而 $d(i,j)$ 接近于0;否则, $d(i,j)$ 会远离0,为一个正整数,其值若大,则表示强边缘,其值若小,则表示弱边缘。故 $d(i,j)$ 可看作是 $N(i,j)$ 边缘程度的一个度量。对整个图像的边缘程度进行统计,得到图像的全局边缘信息。从上述定义可知, $0 \leq d(i,j) \leq 165$,且 $d(i,j) \in Q$ 。对于每个像素点 (i,j) ,我们都可以得到它的极大-极小差 $d(i,j)$,其取值范围是 Q ,统计 $d(i,j)$ 在 Q 上的分布 $(i=0,1,\dots,m,j=0,1,\dots,n)$,得到归一化的局部极大-极小差直方图为:

$$H_{ad}(d) = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \delta(d(i,j) - d), \forall d \in Q$$

3 图像检索

图像检索问题可以简述如下:设 D 是一个图像数据库, M 是查询图像,提取 M 和 D 中图像的特征并存储,对 D 中的每一个图像 I ,采用某种距离度量,根据特征计算得到 I 与 M 的距离,然后按照距离值从小到大对 D 中的图像进行排列,显示给用户。

通常使用的距离度量有 L_1 和 L_2 -Euclidean距离,文[7]认为对于基于直方图的方法,采用 L_1 -相对距离简单、鲁棒,用来比较图像更准确,我们的实验结果也印证了这一点,本文也采用这一距离度量。设 M 为查询图像, I 为图像库中的任意一图像, M 与 I 的量化颜色直方图间的距离为:

$$d_1(M,I) = \sum_{q \in Q} \frac{|H_c^M(q) - H_c^I(q)|}{1 + H_c^M(q) + H_c^I(q)}$$

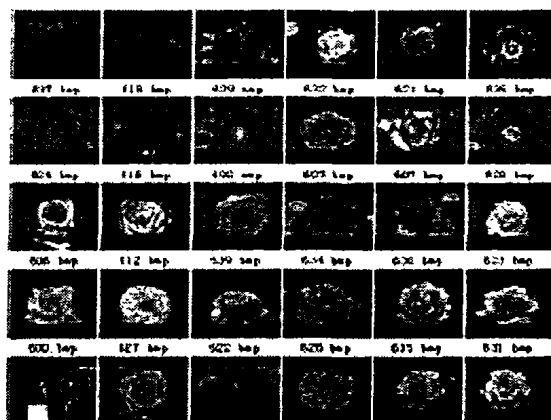
类似地可以定义,它们的局部均值直方图间的距离 $d_2(M,I)$,它们的局部方向差单元直方图间的距离 $d_3(M,I)$,它们的局部极大-极小差直方图间的距离 $d_4(M,I)$ 。那么图像 M

与 I 的距离定义为: $d(M,I) = \sum_{i=1}^4 w_i d_i(M,I)$,其中 $w_i(i=1,2,3,4)$ 为权重。最终,按此距离对图像的特征进行计算,返回基于内容图像检索的查询结果给用户。

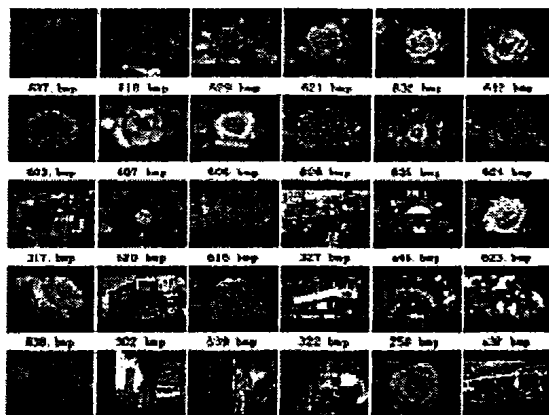
4 实验结果

本文实验所用图像库为1000幅24-bit真彩色JPEG图像,从互联网上下载于文[3],图像大小为 384×256 或 256×384 ,它们被分为10类,每类100幅,这十类图像包括公共汽车、花卉、马、象、海滨、建筑物、人等。我们分别采用三种方法进行

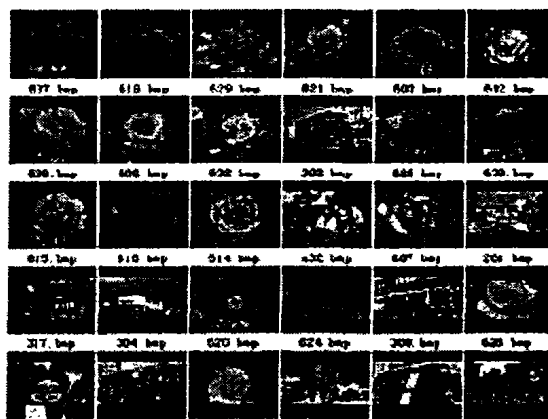
示例查询以便比较性能,方法1:采用文[9]中的方法SCH,方法2:采用文[10]中的方法Geostat,方法3:采用本文提出基于局部特征方法(以下简称LF),实验中将四个权重都选择为1,即 $w_i=1(i=1,2,3,4)$,从实验中发现,两个阈值分别取 $BTH=5$ 、 $HTH=0.198$ 时,其检索性能最优。



(a) 以“花卉”为查询图像用LF方法的查询结果



(b) 以“花卉”为查询图像用SCH方法的查询结果



(c) 以“花卉”为查询图像用Geostat方法的查询结果

图2 以“花卉”为查询图像用三种方法的查询结果

设实验所用1000幅图像组成的图像库为 D ,从 D 中选取100幅同类图像构成查询集合 G ,检索性能的评价采用基于“检索精度-查全率”(precision-recall)的相似检索评价准则,通过执行查询集合 G 中的各个查询,就可以计算出这些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。设 L 为查询返回的图像个数, K 是查询结果中与查询图像相关的图像个数, T 表示测试集合 D 中与查询图像相关的图像总数,这里与某图像相关是指与该图像属于同一类。检索

(下转封四)

(上接第 218 页)

精度和查全率定义如下:

检索精度 $p=K/L$, 它定义为查询结果中, 与检索图像相关的图像所占的百分比, $p=100\%$ 表示查询出的图像均为相关图像。查全率 $r=K/T$, 它定义为查询结果中与检索图像相关的图像占全部相关图像的百分比, $r=100\%$ 表示全部相关图像均被检索出来。检索性能的评价采用“查准率-查全率”曲线 (precision-recall curve), 即查准率作为查全率的函数, 可以这样评价 CBIR 系统的高效性: 在相同查全率条件下, 若系统的查准率越高, 则此系统越高效。

对 10 类中的每一类图像进行测试, 我们发现, 在三种方法中 LF 的性能最好, 同时我们还发现, 相对于其它类型图像而言, LF 对具有相对规则的结构或纹理特征的图像 (比如, “公共汽车”、“花卉”) 的检索效率更高, 可见, LF 更适合于这类图像的检索。

图 2 显示了以“花卉”为查询图像时, 采用三种方法的前 30 幅查询结果图像 (左上角为查询图像), 按从左至右、从上至下的顺序排序。

从图 2 可以看出, 在前 30 幅检索出的图像中, 采用 LF 方法属于“花卉”的为 29 幅, 仅有 1 幅例外, 而用 SCH 方法和 Geostat 方法属于“花卉”的都为 19 幅, 都有 11 幅例外。

图 3 显示了以“花卉”类为查询图像采用三种方法的“precision-recall”曲线, 它是通过分别对“花卉”类 100 幅图像中的各个图像进行查询, 在不同的查全率条件下, 计算出查准率, 再取平均后的值绘制而成的。

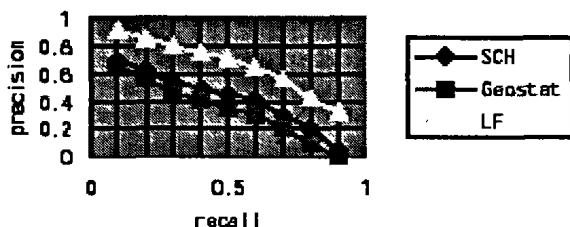


图 3 以“花卉”类为查询图像的平均“precision-recall”曲线

从图 3 可以看出, 在所有任一相同的查全率下, 三种方法中 LF 方法的检索精度最高。对所有查全率, LF 的平均检索精度高出 SCH 方法 21%~27%, 高出 Geostat 方法 25%~34%。另外, 我们可以根据特定的应用在距离度量中选择适当的权重来取得更好的检索性能。

结论 本文给出了一种新的基于局部特征的图像检索方

法, 在此方法中, 我们提出了新的低层可视特征描述子, 这些描述子能够全面地描述图像的特征, 对彩色图像, 尤其是对具有相对规则的结构或纹理特性的图像具有很强的描述力, 同时所提出的颜色描述子、空间描述子、边缘描述子对图像的平移、旋转和尺度不变, 纹理描述子对图像的平移和尺度不变。实验结果显示, 用本文提出的方法进行图像检索时, 其效率比其它基于颜色-空间的方法高得多。今后的工作将着眼于新的高效低层可视特征描述子的表示并从中导出高层语义特征, 进行智能图像检索的研究。

参考文献

1. Rui Y, Huang T, Chang S F. Image retrieval: current technique, promising directions and open issues, Journal of Visual Communication and Image Representation, 1999, 10: 39~62
2. Licker M F, et al. Query by image and video content: The QBIC system. IEEE Computer, 1995, 28(9): 23~32
3. Wang J, Wiederhold G. SIMPLicity: semantics sensitive integrated matching for picture libraries. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(8): 1~17
4. Swain M, Ballard M. Color indexing, International Journal of Computer Vision, 1991, 7(1): 11~32
5. Strick M, Oren M. Similarity of color images, In: Proc. of the SPIE2420, Storage and Retrieval for Image and Video Database II, San Jose CA: SPIE, 1995. 381~392
6. Pass G, Zabih R. Histogram refinement for content-based image retrieval. In: Proc. of the 3rd IEEE Workshop on Applications of Computer Vision, Sarasota: IEEE Computer Society, 1996. 96~102
7. Huang J. Image indexing using color correlograms. In: Proc. of the IEEE Computer society Conf. on Computer Vision and Pattern Recognition, San Juan: IEEE Computer Society, 1997. 762~768
8. Rao A, Srihari R, Zhang Z. Spatial color histograms for content-based image retrieval. In: 11th IEEE Intl. Conf. on Tools With Artificial Intelligence, 1999
9. Cinque L, Levialdi S, Olsen K, Pellicano A. Color-based image retrieval using spatial-chromatic histograms. In: IEEE Intl. Conf. on Multimedia Computing and System, 1999, 2: 969~973
10. Lim S, Lu G. Spatial statistics for content based image retrieval. In: IEEE Intl. Conf. on Information Technology: Computers and Communications, 2003
11. Androustos D, et al. A novel vector-based approach to color image retrieval using a vector angular-based distance measure, Computer Vision and Image Understanding, 1999, 75(1/2): 46~58
12. Manjunath B S, Ma W Y. Texture features for browsing and retrieval of image data. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(8): 837~841
13. He Dong-Chen, Wang Li. Texture features based on texture spectrum, Pattern Recognition, 1991, 24(5): 391~399
14. Pal Sankar K, Majumder Dwijesh k Dutta. Fuzzy Mathematical Approach to Pattern Recognition. Wiley Eastern Limited, New Delhi, India, 1986

计算机科学

(1974 年 1 月创刊)

第 32 卷第 2 期 (月刊)

2005 年 2 月 25 日出版

ISSN 1002-137X
CN50-1075/TP

定价: 25.00 元 国外定价: 5 美元

邮发代号: 78-68

发行范围: 国内外公开

主管单位: 国家科学技术部

主办单位: 国家科技部西南信息中心

编辑出版: 《计算机科学》杂志社

重庆市渝中区胜利路 132 号 邮政编码: 400013

电话: (023) 63500828 E-mail: jsjcx@swic.ac.cn

网址: www.jsjcx.com

社长: 牟炳林

主编: 彭丹

副主编: 朱宗元

主编助理: 徐书令

印刷者: 重庆科情印务有限公司

总发行处: 重庆市邮政局

订购处: 全国各地邮政局

国外总发行: 中国国际图书贸易总公司 (北京 399 信箱)

国外代号: 6210-MO