

基于梯度基元聚合矢量的图像检索算法^{*}

张尤赛, 付 正, 朱志宇

(江苏科技大学 电子信息学院, 江苏 镇江 212003)

摘 要: 针对图像检索中多特征融合问题,提出了一种基于梯度基元聚合矢量的图像检索算法。该算法在改进的 HSV 颜色空间计算边缘梯度,通过定义的基元模板扫描梯度图像,生成梯度基元图像,将基元和非基元像素分别组合成聚合和非聚合像素集合;最后利用颜色自相关图算法对上述两个集合提取特征矢量,实现了融合颜色、形状、纹理和空间信息等多特征的图像检索。实验结果表明,该算法能够融合颜色、形状、纹理和空间信息,有效地提高了基于内容的图像检索的查准率和查全率。

关键词: 梯度基元;改进的 HSV 颜色空间;颜色自相关图;图像检索

中图分类号: TP391.3 **文献标志码:** A **文章编号:** 1001-3695(2012)03-1119-04

doi:10.3969/j.issn.1001-3695.2012.03.087

Image retrieval based on gradient texton coherence vector

ZHANG You-sai, FU Zheng, ZHU Zhi-yu

(School of Electronics & Information, Jiangsu University of Science & Technology, Zhenjiang Jiangsu 212003, China)

Abstract: This paper proposed a new image retrieval algorithm based on gradient texton coherence vector for fusion of multi-features. The algorithm computed the edge gradient first in the modified HSV color space, and then gained gradient texton map by scanning the gradient image through the special texton types. The texton pixels were combined into the coherence set, the other pixels were the non-coherence set. At last, it represented the feature vector of the image retrieval by color auto-correlogram in two sets, and realized the image retrieval based multi-features which fused color, shape, texture and spatial information. Experimental results show that the proposed algorithm can combine color, texture, shape and spatial characteristic effectively, and have valid precision and recall.

Key words: gradient texton; modified HSV color space; color auto-correlogram; image retrieval

0 引言

基于内容的图像检索是计算机视觉领域的一个热点问题。早期的研究者主要根据图像的颜色、纹理、形状以及空间关系等特征进行图像检索。但是实践表明,只通过一种特征进行图像检索,效果往往不能令人满意。因此,将两种或多种特征相融合来提高图像检索的性能,成为一种可行的解决方案。但是,如何有效地融合多种不同类型的特征^[1],以全面互补地表示图像信息,并能满足人对图像的主观感受,是一个值得深入研究的问题。

当前,基于内容的图像检索的多特征融合技术可以分为两个体系。第一个体系是加权求和的方法,即权重求和意义上的融合。在该体系中,根据加权的对象不同,又可分为特征矢量加权 and 单个特征的相似度加权两种方式,即通过多个单特征的加权处理,以使特征矢量能够融合图像的多种不同类型特征,其特点是简单易操作。但是,这种融合只是数学意义上的合并,并没有从图像的角度分析,而且权重的选择需要专业知识,对检索结果影响较大,所以在实际应用中难以取得最佳的效果。第二个体系是从图像内容的角度考虑真正意义上的深层特征的信息融合,即在特征提取的过程中,分别综合考虑不同

种类的特征信息,所提取的特征矢量中自然地包含了多种底层信息,从而能全面互补地表示图像信息。因此,该体系显示出较高的研究价值和发展潜力。近年来,该体系的研究工作受到了人们的关注,并取得了许多研究成果。

Vadivel 等人^[2]提出了一种综合颜色和亮度的改进共生矩阵,在计算共生矩阵的同时加入了不同饱和度和强度对视觉感知的影响,从而免去了选取颜色和亮度特征权重时人为因素的影响。但是该方法的特征维数较多,计算量大,且特征矢量不易构造。刘广海等人^[3]提出了一种四像素共生矩阵,充分利用了图像的彩色梯度信息和像素对之间的空间相关性,能够融合颜色、纹理和形状信息。尽管该方法对于粗糙表面属性的彩色图像检索取得了较好的效果,但并不能很好地适用于精细纹理图像检索。Liu 等人^[4]提出了一种多基元直方图检测纹理特征的方法,通过统计边缘方向角和颜色的联合概率分布,利用多基元直方图描述共生矩阵的特征,从而综合了纹理和颜色的空间相关性,融合了共生矩阵和直方图各自的优点,能有效提取图像的颜色和纹理特征。但是,由于该方法定义的基元类型不能包含所有可能的方式,所以不能完全表述自然图像的纹理多样性特征。

本文针对图像检索中多特征融合问题,提出了一种基于梯

收稿日期: 2011-08-05; **修回日期:** 2011-09-22 **基金项目:** 国家自然科学基金资助项目(61075028)

作者简介: 张尤赛(1959-),男,江苏镇江人,教授,博士,主要研究方向为图像处理、计算机三维可视化、模式识别(yszhang100@163.com);付正(1988-),男,硕士研究生,主要研究方向为图像检索;朱志宇(1971-),男,教授,博士,主要研究方向为模式识别与智能系统、控制理论与控制工程。

度基元聚合矢量的图像检索算法。该算法通过所定义的基元模板扫描梯度图像,得到梯度基元和非梯度基元图像;并利用颜色自相关算法提取两部分图像的颜色分布,作为图像检索的特征矢量。实验结果表明,该算法利用梯度基元表征图像的纹理和形状信息,并采用颜色自相关图来描述图像的颜色信息及其空间相关性,因此很好地融合了颜色、形状、纹理和空间信息,能有效地提高图像检索的查准率和查全率。

1 基于梯度基元聚合矢量的图像检索

1.1 特征提取基本流程

特征提取的基本流程如图 1 所示。

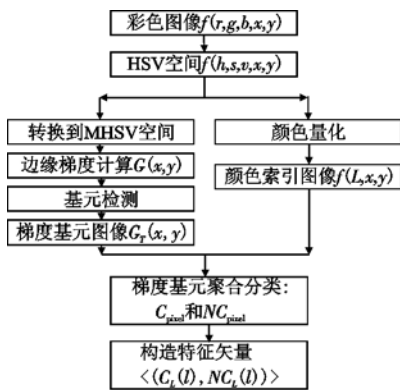


图1 特征提取基本流程

首先将彩色图像转换至 HSV 颜色空间,得到色调、饱和度和亮度三个分量 $H(x,y)$ 、 $S(x,y)$ 、 $V(x,y)$,并将三个分量转换到具有笛卡尔性质的改进 HSV(modified HSV, MHSV)颜色空间,分别为 $H'(x,y)$ 、 $S'(x,y)$ 、 $V'(x,y)$,计算出边缘梯度图像 $G(x,y)$,然后利用定义的基元模板扫描梯度图像,得到梯度基元图像。与之并行的另一个流程是将 HSV 的三个分量 $H(x,y)$ 、 $S(x,y)$ 、 $V(x,y)$ 进行颜色量化,计算颜色索引图像 $f(L,x,y)$ 。最终将颜色索引图像按照梯度基元图像分类,利用基于颜色自相关算法计算图像检索的特征矢量。

1.2 HSV 空间及颜色量化

为了满足人对图像的主观感受,选择了能够反映人眼色彩感知的 HSV 颜色模型,作为图像检索特征提取的颜色空间。为了便于计算特征和降低特征维数,对真彩色图像进行非均匀颜色量化,具体量化方法如下:

$$H = \begin{cases} 0, h \in (315^\circ, 20^\circ] \\ 1, h \in (20^\circ, 40^\circ] \\ 2, h \in (40^\circ, 75^\circ] \\ 3, h \in (75^\circ, 155^\circ] \\ 4, h \in (155^\circ, 190^\circ] \\ 5, h \in (190^\circ, 270^\circ] \\ 6, h \in (270^\circ, 295^\circ] \\ 7, h \in (295^\circ, 315^\circ] \end{cases} \quad (1)$$

$$S = \begin{cases} 0, s \in [0, 0.2] \\ 1, s \in (0.2, 0.7] \\ 2, s \in (0.7, 1] \end{cases} \quad (2)$$

$$V = \begin{cases} 0, v \in [0, 0.2] \\ 1, v \in (0.2, 0.7] \\ 2, v \in (0.7, 1] \end{cases} \quad (3)$$

1.3 颜色索引图

为了便于获得图像像素的颜色值,采用颜色索引图的一维颜色矢量表述原彩色图像的三维颜色矢量。其方法是将颜色量化后得到的三个分量 $H \in [0, 7]$ 、 $S \in [0, 2]$ 、 $V \in [0, 2]$ 在一维方向上展开,合成为一维矢量 L :

$$L = HQ_sQ_v + SQ_v + V \quad (4)$$

式中: Q_s 、 Q_v 分别为 S 、 V 分量的量化级,取 $Q_s = 3$ 、 $Q_v = 3$ 。因此,矢量 $L \in [0, 71]$,共 72 量化级(72 维)。利用 L 代替原彩色图像各像素的颜色值,即得到颜色索引图 $f(L,x,y)$ 。

1.4 MHSV 空间的边缘梯度计算

1.4.1 MHSV 空间

由于 HSV 颜色空间是一种圆锥坐标系,为了能利用笛卡尔坐标系中矢量点积和矢量夹角式来求取边缘梯度。这里采用一种 MHSV 颜色模型^[5],将 HSV 颜色空间从圆锥坐标系转换到笛卡尔坐标系,其转换公式为

$$H' = S \cos(H) \quad (5)$$

$$S' = S \sin(H) \quad (6)$$

$$V' = V \quad (7)$$

该转换式将 HSV 空间的三个分量 $H \in [0^\circ, 360^\circ]$ 、 $S \in [0, 1]$ 、 $V \in [0, 1]$ 转换为 MHSV 空间的三个分量 $H' \in [-1, 1]$ 、 $S' \in [-1, 1]$ 、 $V' \in [0, 1]$ 。由此可见,MHSV 满足笛卡尔坐标系的条件。

1.4.2 边缘梯度计算

传统意义上定义的边缘检测算子(如 Sobel、Robert、LOG 等)只能应用于灰度图像的边缘检测。因此,为了能够同时提取图像的梯度和颜色信息,这里采用了以下的梯度计算方法^[4]:

a) 对 MHSV 的三个分量分别求 x 和 y 方向的梯度,并构造矢量 P 和 Q :

$$P(x,y) = \left[\frac{\partial H'(x,y)}{\partial x}, \frac{\partial S'(x,y)}{\partial x}, \frac{\partial V'(x,y)}{\partial x} \right]^T = [H'_x(x,y), S'_x(x,y), V'_x(x,y)]^T \quad (8)$$

$$Q(x,y) = \left[\frac{\partial H'(x,y)}{\partial y}, \frac{\partial S'(x,y)}{\partial y}, \frac{\partial V'(x,y)}{\partial y} \right]^T = [H'_y(x,y), S'_y(x,y), V'_y(x,y)]^T \quad (9)$$

b) 利用笛卡尔坐标系中矢量点积和矢量夹角式,计算图像的方向角 $\theta(x,y)$:

$$\theta(x,y) = \arccos \frac{P(x,y) \cdot Q(x,y)}{\|P(x,y)\| \times \|Q(x,y)\|} \quad (10)$$

c) 由下列公式计算图像的彩色边缘梯度 $G(x,y)$ 幅值:

$$A(x,y) = H_x^2 + V_x^2 + S_x^2 + H_y^2 + V_y^2 + S_y^2 \quad (11)$$

$$B(x,y) = H_x^2 + V_x^2 + S_x^2 - H_y^2 - V_y^2 - S_y^2 \quad (12)$$

$$C(x,y) = H'_x H'_y + V'_x V'_y + S'_x S'_y \quad (13)$$

$$G(x,y) = (0.5 \times [A(x,y) + B(x,y) \times \cos 2\theta(x,y) + 2 \cdot C(x,y) \times \sin 2\theta(x,y)])^2 \quad (14)$$

d) 梯度幅值均匀量化:

$$G_{\text{index}}(x,y) = \left[\frac{G(x,y)}{\max(G(x,y))} \times K + 0.5 \right] \quad (15)$$

式中: K 表示梯度量化级。

1.5 基元检测及梯度基元图像

为了体现相同的边缘梯度值的颜色分布,这里将图像的边缘梯度作为基元分布,通过构造基元规则进行梯度基元检测。

基元的构成规则为:a) 选取大小为 3×3 个元素构成的模

板, w_0 为中心元素; b) 以梯度图像 $G_{\text{index}}(x, y)$ 作为基元的属性; c) 采用重叠式的模板扫描; d) 模板内与 w_0 梯度值相等的元素构成一个基元; 模板内 8 邻域元素与 w_0 梯度值均不相等, 则该模板不构成基元。

图 2 给出了 3×3 模板及几种典型基元结构的示例, 其中黑色元素梯度值相同, 构成一个基元; 白色元素梯度值与 w_0 不同, 不属于基元元素。

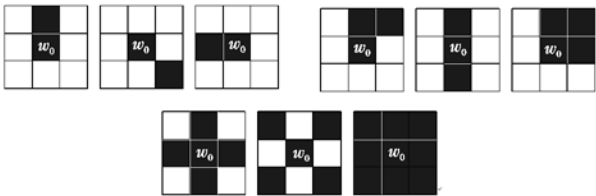


图2 3×3模板及几种典型基元结构示例

- 基元检测步骤^[6]为:
- a) 从图像的第一个像素开始, 以 3×3 模板将图像分为互不相交的各个子块。
 - b) 在每个 3×3 的子块中, 以 w_0 为参照构成基元, 所有基元元素保留原梯度值, 所有非基元元素的值置为 -1 。
 - c) 在 3×3 子块中, 8 邻域中所有元素均与 w_0 不相等, 则该子块不构成基元, 所有元素值置为 -1 。
 - d) 据此得到子基元图像 $G_{(0,0)}(x, y)$ 。
 - e) 分别从图像的第 1 行第 2 列、第 2 行第 1 列和第 2 行第 2 列的像素开始, 重复步骤 a) c), 分别得到子基元图像 $G_{(0,1)}(x, y)$ 、 $G_{(1,0)}(x, y)$ 和 $G_{(1,1)}(x, y)$ 。
 - f) 融合上述四个子基元图像, 对应元素取四个基元图像中的最大值, 最后得到梯度基元图像:
- $$G_T(x, y) = \max \{ G_{(i,j)}(x, y) \mid i=0,1; j=0,1 \} \quad (16)$$

1.6 梯度基元聚合分类与图像检索特征矢量

梯度基元图像 $G_T(x, y)$ 表示纹理属性、排列以及重复结构, 并且包含了一定的形状特征。根据梯度基元图像将像素分为两类: 一类元素值 $G_T(x, y) \geq 0$, 称为聚合像素; 另一类元素值 $G_T(x, y) < 0$, 称为非聚合像素。

假设聚合像素集合 C_{pixel} 和非聚合像素集合 NC_{pixel} 为

$$C_{\text{pixel}} = \{ (x, y) \mid G_T(x, y) \geq 0 \} \quad (17)$$
$$NC_{\text{pixel}} = \{ (x, y) \mid G_T(x, y) < 0 \} \quad (18)$$

颜色索引图 $f(L, x, y)$ 根据聚合像素集合 C_{pixel} 和非聚合像素集合 NC_{pixel} 分为 $f_c(L, x, y)$ 和 $f_{nc}(L, x, y)$ 两部分。

$$f_c(L, x, y) = f(L, x, y) \quad (x, y) \in C_{\text{pixel}} \quad (19)$$
$$f_{nc}(L, x, y) = f(L, x, y) \quad (x, y) \in NC_{\text{pixel}} \quad (20)$$

对 $f_c(L, x, y)$ 和 $f_{nc}(L, x, y)$ 分别计算相同颜色索引值同时出现的条件概率, 利用颜色自相关图算法进行特征提取。定义 $C_L(l)$ 表示集合 $f_c(L, x, y)$ 中颜色值为 l 的特征值, $NC_L(l)$ 表示集合 $f_{nc}(L, x, y)$ 中颜色值为 l 的特征值, $l \in L$ 。Pr 表示颜色值为 l 的像素对 (x_1, y_1) 和 (x_2, y_2) 同时出现的概率。D 表示两个像素点之间的距离。

$$C_L(l) = Pr \{ f(L, x_1, y_1) = f(L, x_2, y_2) = l \mid (x_1, y_1) - (x_2, y_2) \leq D, f(L, x_1, y_1), f(L, x_2, y_2) \in f_c(L, x, y) \} \quad (21)$$
$$NC_L(l) = Pr \{ f(L, x_1, y_1) = f(L, x_2, y_2) = l \mid (x_1, y_1) - (x_2, y_2) \leq D, f(L, x_1, y_1), f(L, x_2, y_2) \in f_{nc}(L, x, y) \} \quad (22)$$

据此构造图像检索的特征矢量:

$$M(l) = \langle (C_L(l), NC_L(l)) \rangle \mid l \in L \quad (23)$$

在进行相似度量时, 检索图像 S 和图像库中的图像 T 之间的距离函数为

$$D_{ST} = \sum_{f=0}^{71} (|C_L(f)_S - C_L(f)_T| + |NC_L(f)_S - NC_L(f)_T|) \quad (24)$$

2 实验结果分析

实验采用了 Corel-1000 图像库^[7], 共有非洲土著 (Africa)、海滩 (beach)、建筑 (building)、公共汽车 (buses)、恐龙 (dinosaurs)、大象 (elephants)、花卉 (flowers)、马匹 (horses)、山川和极地 (mountains glaciers) 以及食物 (food) 10 类图像, 每类 100 幅, 共计 1 000 幅图像。图 3 给出了一些不同类型的典型图像示例。

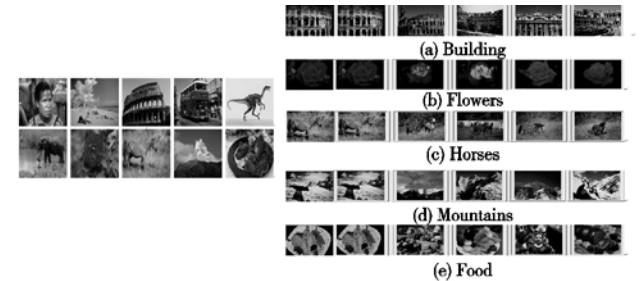


图3 10种不同类型的典型图像示例 图4 本文算法的图像检索示例

实验以 Corel-1000 图像库 (1 000 幅图像) 作为检索集, 从每类 100 幅图像中随意选取 20 幅不同的图像作为检索图像, 进行 20 次图像检索实验, 10 类图像共计 200 次实验。每次实验将相似度最大的前 20 幅图像作为检索返回结果, 计算出每次实验的查全率和查准率, 最后得到每类图像 20 次实验的平均查全率和平均查准率, 作为算法性能的平均指标。查全率和查准率的定义为

$$\text{查准率} = \frac{\text{正确检索到的图像数}}{\text{设定的返回图像数}} \times 100\% \quad (25)$$

$$\text{查全率} = \frac{\text{正确检索到的图像数}}{\text{图像库中相关图像的总数目}} \times 100\% \quad (26)$$

图 4 是本文算法针对 building、flowers、horses、mountains 和 food 五种类型图像的检索实例。其中, 第一幅为检索图像, 后面的 5 幅为检索返回的查准率最高的前 5 幅图像。

为了验证算法的性能, 实验复现了基于 ICICM 算法的图像检索^[2]和基于 MTH 算法的图像检索^[4], 并与本文算法进行了性能比较。根据实验检索效果和特征提取时间综合考虑, 本文算法 HSV 颜色量化为 72 维, 梯度量化值 $K = 30$, 距离参数 $D = 1$ 。ICICM 算法^[2] HSV 量化为 72 维, r_1 选择 0.1, r_2 选择 0.85。MTH 算法^[4] RGB 颜色量化为 64 维, 方向角量化为 18。

表 1、2 分别给出了上述三种算法针对 10 种类型图像, 每类 20 次实验的平均查准率和平均查全率。

表 1 不同种类图像平均查准率比较 %

编号	图像类别	MTH	ICICM	本文算法
1	africa	33.0	33.5	43
2	beach	27.0	20.8	27.5
3	building	44.5	43.5	65.8
4	buses	48.0	67.0	72.0
5	dinosaurs	92.8	91.0	99.0

续表 1

编号	图像类别	MTH	ICICM	本文算法
6	elephants	35.3	34.0	39.3

编号	图像类别	MTH	ICICM	本文算法
7	flowers	45.0	67.3	87.0
8	horses	73.5	71.0	92.8
9	mountains	22.5	26.5	37.5
10	food	30.5	38.0	50.5

表 2 不同种类图像平均查全率比较 %

编号	图像类别	MTH	ICICM	本文算法
1	africa	6.6	6.7	8.6
2	beach	5.4	4.2	5.5
3	building	8.9	8.7	13.1
4	buses	9.6	13.4	14.4
5	dinosaurs	18.6	18.2	19.8
6	elephants	7.1	6.8	7.9
7	flowers	9.0	13.5	17.4
8	horses	14.7	14.2	18.6
9	mountains	4.5	5.3	7.5
10	food	6.1	7.6	10.1

图 5、6 是表 1、2 中三种算法的平均查准率和平均查全率的对比图,图 7 是三种算法的平均查全率—查准率曲线。

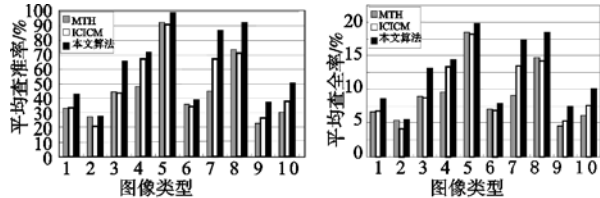


图5 三种算法的平均查准率对比

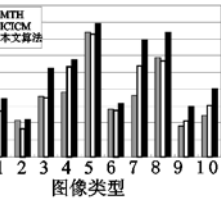


图6 三种算法的平均查全率对比

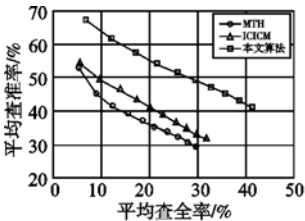


图7 三种算法的平均查全率-查准率曲线

表 3 给出了上述三种算法对一幅图像的特征提取时间和检索时间。实验的硬件环境为 Pentium® Dual-Core CPU E5400 2.70 GHz,内存 2 GB,软件平台为 Visual C++ 6.0。

表 3 三种算法的特征提取时间与图像检索时间比较

比较项	MTH	ICICM	本文算法
特征提取时间	65 ms	1 562 ms	175 ms
图像检索时间	6.5 s	6.2 s	6.4 s

图像检索分为特征提取和检索查询两个步骤。因为特征提取是在图像进入数据库时一次完成,在以后的查询中即可以重复使用;而对待检索的图像也是一次提取,所以特征提取时间对图像检索时间的影响较小。因此图像检索效率主要体现在检索查询上,检索查询时间主要与提取的特征维数有关。从表 3 可以看出,三种算法的检索查询时间在同一数量级上,差别不大。

实验表明,本文提出的基于梯度基元聚合矢量的图像检索算法具有以下优点:

a) 融合了颜色、纹理、形状和空间多种单特征,提高了图像检索的查准率和查全率。由表 1、2 和图 5、6 可以看出,本文算法针对每种类型图像检索的查全率和查准率均高于 ICICM 和 MTH 算法,特别是针对 building、flowers、horses、mountains 和 food 五种类型图像的检索性能得到了显著提高。其原因是本文算法通过检测梯度图像的基元结构表达了纹理和形状特征,

提取基元和非基元像素的颜色分布作为特征矢量,融合了颜色和空间信息。而 ICICM 算法在考虑纹理特征的同时,主要侧重颜色和亮度的融合,没有考虑形状特征,因此对 beach、building、dinosaurs、elephants 和 horses 五类富有形状特征的图像检索效果较差。MTH 算法融合了纹理、颜色和形状的特征,忽略了空间关系,所以在 buses、flowers、mountains 和 food 四类富有空间关系的图像检索效果较差。图 7 给出的平均查全率—查准率曲线表明,本文算法的曲线远高于其他两种算法,具有优越的检索性能。

b) 特征提取与融合反映了人类视觉特性。采用梯度基元图像和颜色索引图像相融合的方法,在提取特征的过程中,从图像内容的角度自然地考虑了颜色、纹理、形状和空间等多特征,避免了传统方法中简单的数学加权的多特征综合所造成的视觉信息失真,符合人类视觉特性。

c) 对不同类型图像的适应性强。通过定义各种基元结构,利用重叠式扫描方式检测基元,可以涵盖所有的基元类型,无论对于颜色、纹理、形状和空间关系复杂多样的自然图像,还是对于相对单一的纹理图像,均有很好的适应性和检索效果。

3 结束语

本文提出的基于梯度基元聚合矢量的图像检索算法,通过检测梯度图像的基元结构,提取基元和非基元像素的颜色分布特征,自然地融合了颜色、纹理、形状和空间等特征。实验结果表明,本文算法适应性好,具有较理想的图像检索效果,明显提高了图像检索的查准率和查全率。

在基于内容的图像检索中,视觉特征和语义特征是两种重要的图像特征。视觉特征反映了人对图像的直观视觉感受,属于底层特征;语义特征则体现了人对图像内容的观察理解,属于高层特征,在它们之间有时存在着巨大的语义鸿沟,如何克服这种鸿沟是一个尚未很好解决的问题。因此,深入研究视觉特征与语义特征之间的映射关系,并将语义特征应用于基于内容的图像检索是今后重要的研究方向。

参考文献:

[1] 吴介,裴正定. 底层内容特征的融合在图像检索中的研究进展[J]. 中国图象图形学报, 2008,13(2):189-197.

[2] VADIVEL A, SHAMILK S, MAJUMDAR A K. An integrated color and intensity co-occurrence matrix[J]. Pattern Recognition, 2007, 28(8):974-983.

[3] 刘广海,杨静宇. 基于四像素共生矩阵的图像检索[J]. 中国图象图形学报, 2008,13(9):1783-1788.

[4] LIU Guang-hai, ZHANG Lei, HOU Ying-kun, et al. Image retrieval based on multi-texton histogram[J]. Pattern Recognition, 2010, 43(7):2380-2389.

[5] RUI Min, CHENG H D. Effective image retrieval using dominant color descriptor and fuzzy support vector machine[J]. Pattern Recognition, 2009, 42(1):147-157.

[6] LIU Guang-hai, LI Zuo-yong, ZHANG Lei, et al. Image retrieval based on micro-structure descriptor[J]. Pattern Recognition, 2011, 44(9):2123-2133.

[7] WANG J Z. Test database used in simplicity paper [DB/OL]. (2001-07). <http://wang.ist.psu.edu/docs/related>.