

DOI: 10.7652/xjtuxb201706014

一种三结构描述子的图像检索方法

李莉¹, 冯林^{1,2}, 吴俊², 刘胜蓝³

(1. 大连理工大学计算机科学与技术学院, 116024, 辽宁大连; 2. 大连理工大学创新创业学院, 116024, 辽宁大连; 3. 大连理工大学控制科学与工程学院, 116024, 辽宁大连)

摘要: 针对基元结构描述子在颜色空间的基础上再提取其他特征, 导致偏重对颜色信息的描述而降低了图像检索性能的问题, 提出一种应用在 HSV 颜色空间上的三结构描述子(TSD)的特征提取方法。该方法在 HSV 颜色空间中分别提取颜色和纹理信息, 考虑到了颜色和纹理特征的同等重要性, 同时避免了颜色信息的过多干扰; 在纹理特征提取中, TSD 利用像素间的信息变化来表示局部空间结构信息, 解决了传统的局部模式方法忽略对局部结构的空问关系描述的问题, 获得了更多的空间结构信息。实验结果表明, 该方法在 3 个图像库 Corel-1000、Corel-5000 和 Corel-10000 上的检索准确率分别达到 78.08%、38.12% 和 52.12%, 与以往基元结构方法相比, 检索准确率得到了提高。

关键词: HSV 颜色空间; 三结构描述子; 纹理描述子; 图像检索

中图分类号: TP391 **文献标志码:** A **文章编号:** 0253-987X(2017)06-0086-06

Image Retrieval Method Based on Three-Structure Descriptor

LI Li¹, FENG Lin^{1,2}, WU Jun², LIU Shenglan³

(1. School of Computer Science and Technology, Dalian University of Technology, Dalian, Liaoning 116024, China;
2. School of Innovation and Entrepreneurship, Dalian University of Technology, Dalian, Liaoning 116024, China;
3. School of Control Science and Engineering, Dalian University of Technology, Dalian, Liaoning 116024, China)

Abstract: A novel feature extraction method called three-structure descriptor (TSD) used in HSV color space is proposed to address the problem that the image description methods based on texon pay undue attention to the color description so reduce the performance of image retrieval due to directly extracting other features in color space. The proposed method extracts color and texture information in HSV color space and the same importance of both the color and texture is taken into account while the excessive interference of color information is avoided. TSD uses the information change among pixels to represent the local spatial structure information in texture feature extraction, which solves the problem that the traditional local pattern methods disregard the spatial correlation in local structure and captures more information regarding to spatial structure information. Experimental results show that the retrieval precisions of the proposed method reach 78.08%, 38.12% and 52.12% on the Corel-1000, Corel-5000 and Corel-10000 databases respectively, and its retrieval precisions are higher than those of existing methods based on texon.

Keywords: HSV color space; three-structure descriptor; texture descriptor; image retrieval

收稿日期: 2016-11-21。 作者简介: 李莉(1984—),女,博士生;冯林(通信作者),男,教授,博士生导师。 基金项目: 国家自然科学基金资助项目(61370200,61672130,61602082)。

网络出版时间: 2017-03-28 网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20170328.1133.002.html>

近年来,随着网络技术和多媒体技术的快速发展,图像资源越来越多,如何快速、高效地从海量的图像库中找到所需图像已成为模式识别领域的热点问题。传统的图像检索系统使用关键字、标题或文件名等文本特征标注图像,然而对于一个海量图像库,利用文本信息很难获得有效的图像内容信息,因此基于内容的图像检索系统得到广泛使用。

基于内容的图像检索系统^[1],首先提取图像的颜色、纹理、形状或其他多种底层视觉特征,之后选择合适的相似性度量方法找出待查询图像与数据库中相似的图像。图像底层特征提取是图像检索系统的关键步骤,直接影响了图像的检索结果。

颜色特征统计了图像中各像素点的颜色值分布情况,它是最基本且有效的视觉特征。颜色直方图^[2]描述了图像的全局特征,但丢失了图像的位置信息。针对此问题,Stricker等提出了颜色矩方法^[3],它描述了图像的颜色空间分布信息。纹理特征也是一种重要的视觉特征,其中灰度共生矩阵利用图像像素的共生关系描述图像的纹理特征^[4]。Siqueira等将灰度共生矩阵扩展为多尺度描述子^[5],并将得到的特征应用在纹理分类中;Ojala等提出局部二值模式(local binary pattern, LBP)描述子用于纹理分类^[6]。此后, LBP的变种相继出现应用在纹理分类、人脸识别等领域。形状特征在图像分析中也具有重要的作用,其中边缘方向直方图^[7]、傅里叶描述子^[8]等基于轮廓的描述方法,描述了图像的边缘信息。

颜色、纹理和形状特征都是图像的重要底层特征,但是利用单一的特征进行图像检索效果往往不够理想,因此研究者提出多特征融合的方法。Vipparthi等提出基于modelling的多连接直方图(multi-joint histogram based modelling, MJHM)^[9],利用结构元和主题模式联合统计表达图像信息;微结构描述子(micro-structure descriptor, MSD)在颜色特征的基础上结合边缘方向表示图像特征^[10];Liu等提出的多纹理直方图(multi-texton histogram, MTH)^[11]和颜色差分直方图(color difference histogram, CDH)^[12]利用了共生矩阵和直方图,同时描述了图像的颜色、纹理和形状特征。

以上方法存在以下不足:①在纹理特征提取方面,传统的局部模式方法只考虑了中心像素点和邻居像素点间的关系,忽略了图像像素间的结构区分性信息;②在颜色特征上统计结构基元的方法会偏重对颜色特征的描述,颜色信息会占主导地位,这样

会降低图像的检索效果。为了克服上述不足,本文提出一种新的特征表示方法提取图像的颜色和纹理特征。针对问题①本文提出一种新的三结构描述子(three structure descriptor, TSD)的纹理特征提取方法,该方法描述了局部结构中像素间的相关性信息;针对问题②本文提出在HSV颜色空间中分别提取图像的颜色和纹理信息,避免了颜色特征的过多干扰。实验结果表明本文方法比其他方法有更好的检索性能。

1 颜色和纹理描述子

1.1 HSV颜色空间

HSV颜色空间是一种常用的颜色空间模型,其中H、S和V分别代表图像的色调、饱和度和亮度,色调H由0~360°的不同角度表示不同的颜色。饱和度S的取值范围为0~1,随着S的增大,颜色越饱和。亮度V表示色彩的明亮程度,取值范围为0~1。

1.2 局部二值模式

Ojala等提出局部二值模式用于纹理分类^[6]。对于图像中给定的中心像素点,其LBP值是通过比较中心像素点与其周围邻居像素点的差值来计算的,即

$$L_{P,R} = \sum_{i=0}^{P-1} 2^i s(g_i - g_c) \quad (1)$$

式中:P和R分别为邻居像素点的个数和中心像素点与邻居像素点的半径; g_c 为中心像素点; g_i 为周围邻居像素点; $s(x)$ 为阈值函数,其定义为

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

2 图像特征提取

2.1 图像特征提取方法

在彩色图像检索中,颜色和纹理是图像的两种非常重要的特征。由Julesz基元理论^[13]可知,相似的图像会有相似的局部结构,根据对纹理基元的不同定义,研究者提出基于基元结构的方法。这类方法直接在颜色特征的基础上统计基元结构,但该方法过分重视图像的颜色特征,影响了图像的检索效果。针对此问题,本文提出一种新的特征提取方法,在HSV颜色空间中分别提取图像的颜色和纹理信息,以避免颜色特征占主导因素给图像检索带来干扰。

本文的图像特征提取方法分3步执行:①由于

HSV 颜色空间与人眼对视觉感知的理解非常接近,因此选择在 HSV 颜色空间中,将 H 分量量化为 18、36 和 72 维, S 分量量化为 10 维和 20 维,采用不同的量化方案,选出最好的颜色信息,并构建 H 和 S 分量的颜色直方图;②在 HSV 颜色空间中,由于 V 分量非常接近 RGB 图像的灰度转换,因此在 V 分量上提取不同结构的纹理信息;③合并得到的颜色和纹理特征向量来表示图像的特征。

本文的图像特征提取方法,考虑到了颜色和纹理特征的同等重要性,利用从 H 和 S 分量提取到的颜色直方图表示全局颜色信息,使用三结构描述子提取出的局部结构中像素间的相关性信息表示纹理特征。

2.2 三结构描述子

图像可看成是由规则的局部结构组成。对这些局部结构的描述可获得图像的纹理信息。根据对局部结构的不同定义,衍生出不同的纹理特征描述子,局部二值模式就是其中的一种。但是,局部二值模式只考虑了中心像素点和邻居像素点间的关系,忽略了局部结构像素间的相关性。针对局部二值模式的不足,本文提出了三结构描述子(TSD)方法,该方法由外结构 D_O 、内结构 D_I 和相等结构 D_E 组成,描述了局部结构像素间梯度方向上的空间结构分布信息,下面具体介绍 TSD 的计算过程。

假设半径为 1 的邻居像素点为 $g_{R,8 \times n,m}$ (半径 $R=1$;半径为 R 的邻居像素点的总数是 8 的倍数,将半径为 R 的邻居像素点分为 8 组,即取半径为 1 的每组像素点个数 n 为 1;半径为 1 的邻居像素点的个数为 $8 \times n; m=0,1,\dots,7$)。为了与半径为 1 的邻居像素点进行比较,并考虑到邻居像素点间比较的合理性,利用圆弧上相邻像素点间的均值将半径为 2 的邻居像素点作如下转换,得到新的半径为 2 的邻居像素点

$$G_{R+1,a,m} = \frac{1}{a} \sum_{k=0}^{a-1} g_{R+1,8 \times a, a \times m + k}, \quad m = 0, 1, \dots, 7 \quad (3)$$

式中:半径为 2 的每组像素点个数为 $a, a = n + 1 = 2$;半径为 2 的邻居像素点的个数为 $8 \times a$ 。

通过式(3)的转换,半径为 2 的周围邻居像素点的总个数为 8。将中心像素点 $g_{R,n,c}$ 与半径为 1 的周围邻居像素点 $g_{R,8 \times n,m}$ 作比较,可得到中心像素点与半径为 1 的周围邻居像素点间的局部差分

$$\tilde{I}(g_{R,1,m}) = g_{R,n,c} - g_{R,8 \times n,m}, \quad m = 0, 1, \dots, 7 \quad (4)$$

半径为 2 的邻居像素点 $G_{R+1,a,m}$ 与半径为 1 的周围邻居像素点 $g_{R,8 \times n,m}$ 的局部差分 $\tilde{I}(g_{R+1,2,m})$,可通过下式计算

$$\tilde{I}(g_{R+1,2,m}) = G_{R+1,a,m} - g_{R,8 \times n,m}, \quad m = 0, 1, \dots, 7 \quad (5)$$

三结构描述子的 D_O 结构编码可由下面的判断函数计算得到

$$f_1(\tilde{I}(g_{R,1,m}), \tilde{I}(g_{R+1,2,m})) = \begin{cases} 1, & \tilde{I}(g_{R,1,m}) > 0; \tilde{I}(g_{R+1,2,m}) > 0 \\ 0, & \text{其他} \end{cases} \quad (6)$$

三结构描述子的 D_I 结构编码可由下面的判断函数计算得到

$$f_2(\tilde{I}(g_{R,1,m}), \tilde{I}(g_{R+1,2,m})) = \begin{cases} 1, & \tilde{I}(g_{R,1,m}) < 0; \tilde{I}(g_{R+1,2,m}) < 0 \\ 0, & \text{其他} \end{cases} \quad (7)$$

三结构描述子的 D_E 结构编码可由下面的判断函数计算得到

$$f_3(\tilde{I}(g_{R,1,m}), \tilde{I}(g_{R+1,2,m})) = \begin{cases} 1, & \tilde{I}(g_{R,1,m}) = 0; \tilde{I}(g_{R+1,2,m}) = 0 \\ 0, & \text{其他} \end{cases} \quad (8)$$

与原始的局部二值模式(LBP)编码方案类似,取邻居像素点的个数为 P ,TSD 特征向量定义如下

$$T = \{D_O; D_I; D_E\} \quad (9)$$

其中

$$D_O = \sum_{m=0}^{P-1} 2^m f_1(\tilde{I}(g_{R,1,m}), \tilde{I}(g_{R+1,2,m})) \quad (10)$$

$$D_I = \sum_{m=0}^{P-1} 2^m f_2(\tilde{I}(g_{R,1,m}), \tilde{I}(g_{R+1,2,m})) \quad (11)$$

$$D_E = \sum_{m=0}^{P-1} 2^m f_3(\tilde{I}(g_{R,1,m}), \tilde{I}(g_{R+1,2,m})) \quad (12)$$

图 1 给出了 TSD 的计算示例,先利用式(3)将半径为 2 的周围邻居像素点作转换,得到 8 个新的半径为 2 的邻居像素点;之后将圆形邻域对应转换到 5×5 的块邻域;再通过式(4)和式(5)得到中心像素点与半径为 1 的邻居像素点及半径为 2 的邻居像素点与半径为 1 的邻居像素点的 5 种方向走向,箭尾到箭头的指向表示像素值由小到大变化,直线表示两个像素值相等;最后由式(6)~式(8)得到 D_O (即 2 个向外方向的箭头表示相应的二进制位为 1,其他方向编码为 0)、 D_I (即 2 个向内方向的箭头表示相应的二进制位为 1,其他方向编码为 0)和 D_E (即 2 条直线表示相应的二进制位为 1,其他方向编码为 0)的结构编码。下面通过具体实例说明 TSD 的编码过程。

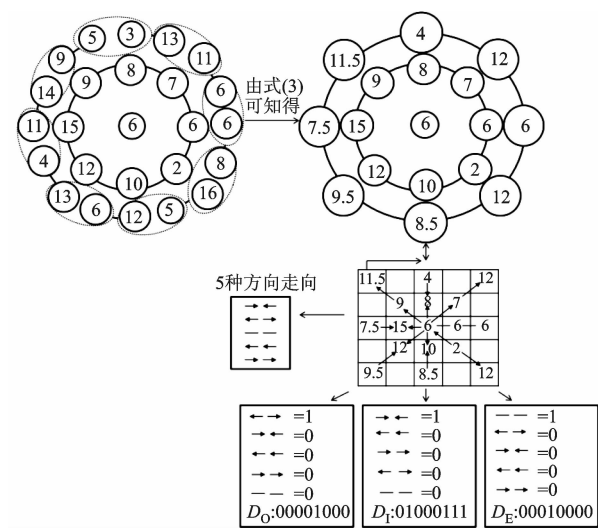


图 1 TSD 计算示例

如图 1 所示,对于中心像素点的值为 6,半径为 1 的第一个邻居像素点的值为 9,半径为 2 的第一个邻居像素点的值为 11.5。在 D_0 结构中,将中心像素点的值与半径为 1 的第一个邻居像素点的值作比较,值小于 0,而将半径为 2 的第一个邻居像素点与半径为 1 的第一个邻居像素点比较后值大于 0,不满足 D_0 中两值同时大于 0 的条件,因此 8 位编码中的第一位为 0。类似地,计算剩余的 7 位编码值,得到 D_0 最终的结构编码为 00001000。在 D_1 结构中,将中心像素点的值与半径为 1 的第一个邻居像素点的值作比较,值小于 0,而将半径为 2 的第一个邻居像素点与半径为 1 的第一个邻居像素点比较后值大于 0,不满足 D_1 中两值同时小于 0 的条件,因此 8 位编码中的第一位为 0。类似地,计算剩余的 7 位编码值,得到 D_1 最终的结构编码为 01000111。同理,在 D_E 结构中,针对 8 位编码中的第一位,通过比较计算,不满足 D_E 中两值相等的条件,因此 8 位编码中的第一位为 0。类似地,计算剩余的 7 位编码值,得到 D_E 最终的结构编码为 00010000。最

表 1 本文方法在 HSV 颜色空间上不同量化维数的查准率和查全率

(d_H, d_S, d_V)	查准率/%			查全率/%		
	Corel-1000	Corel-5000	Corel-1000	Corel-1000	Corel-5000	Corel-10000
(36,10,768)	77.02	36.73	51.94	29.92	9.94	15.32
(36,20,768)	77.71	37.45	52.12	30.41	10.30	15.42
(72,10,768)	77.32	36.79	51.22	30.18	10.08	15.13
(72,20,768)	78.05	37.89	51.84	30.79	10.49	15.37
(108,10,768)	77.46	36.97	50.79	30.27	10.10	15.02
(108,20,768)	78.08	38.12	51.49	30.86	10.57	15.30

后,通过合并这 3 个结构的特征向量得到最终的一个 $3 \times 256 = 768$ 维的特征向量。

2.3 本文方法与其他方法比较分析

本文提出的应用在 HSV 颜色空间上的三结构描述子特征提取方法,利用了像素间的信息变化表示局部空间结构信息。它与传统的局部二值模式方法不同,局部二值模式仅提取了中心像素点和邻居像素点间的关系。因此,三结构描述子与传统的局部二值模式方法相比,能够获得更多的空间结构信息。另外,局部二值模式还忽略了颜色特征的描述。

此外,在颜色特征上统计结构基元的方法有 MSD 方法、CDH 方法等。MSD 方法在颜色特征的基础上结合边缘方向表示图像特征;CDH 方法利用共生矩阵和直方图描述图像的颜色、纹理和形状特征。但是,这些方法会偏重对颜色特征的描述,从而降低图像的检索性能。本文在 HSV 颜色空间中分别提取颜色和纹理信息,考虑到颜色和纹理特征的同等重要性。

3 实验设置与结果分析

3.1 图像检索实验设置

实验中使用的图像库为 Corel-1000、Corel-5000、Corel-10000 和 CIFAR-10。Corel-1000 包含 10 类图像,共 1 000 幅;Corel-5000 包含 5 000 幅图像,共 50 类;Corel-10000 包含 10 000 幅图像,共 100 类;CIFAR-10 包含 10 类图像,每类包含 6 000 幅图像。

在检索时,选择数据库中所有的图像作为查询图像,使用查全率和查准率衡量特征提取框架的有效性,选择改进 Canberra 距离作为特征间的相似性度量方法。

3.2 图像检索实验结果分析

本文比较了不同量化维数对特征提取框架的影响。假设 H 分量的维数为 d_H , S 分量的维数为 d_S ,

V 分量的级数为 d_v 。表 1 统计了不同图像库的不同量化维数下图像检索的查准率(返回的图像数为 10)和查全率(返回的图像数为 50)。从表 1 可以看出, d_H 分别取 36、72 和 108, d_s 取 10 和 20 时, 不同的量化方案在不同的数据库中表现出不同的性能, 这取决于不同图像库上图像的颜色和纹理特征。在 Corel-1000 和 Corel-5000 图像库上, d_H 取 108、 d_s 取 20 是个比较好的选择; 在 Corel-10000 图像库上, d_H 取 36、 d_s 取 20 是个比较好的选择。

为了验证本文方法的有效性, 选择 4 种基于基元结构的方法 MTH、CDH、MJHM 和 MSD 进行比较, 图 2 为本文方法与其他 4 种方法在 Corel-1000 图像库上的整体检索效果比较。设置返回图像数起始值为 10, 之后每次增加 10 幅, 一直到 100 幅结束。从图 2 可以看出, 返回前 10 幅图像时, 本文方法的查准率达到 78.08%, 而 MTH、CDH、MJHM 和 MSD 方法分别只有 73.65%、71.59%、71.18% 和 73.71%。随着返回图像数的增加, 本文方法总是好于其他 4 种方法。如图 3 所示, 给出了检索前 20 幅图像时, 10 类图像的查准率和查全率。从图 3 可以看出, 本文方法在大多数类别中取得了好的检索效果。

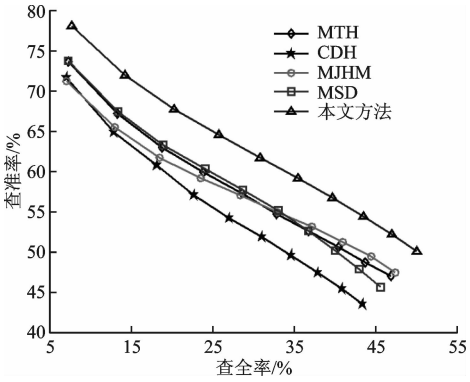
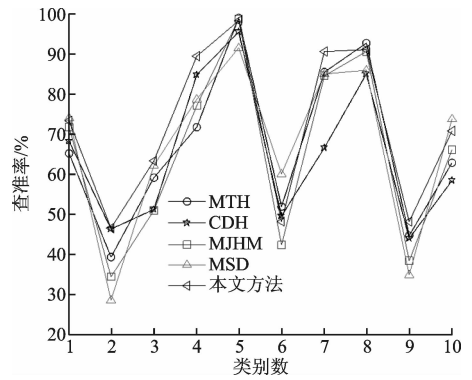


图 2 在 Corel-1000 图像库各方法的检索性能比较

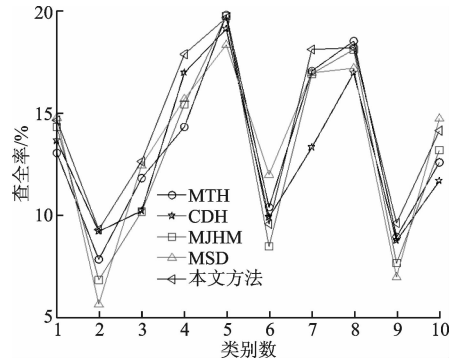
由图 4 可以看出, 在 Corel-5000 图像库中, 返回前 10 幅图像时, 本文方法的查准率达到 38.12%, 随着检索图像数的增加, 检索效果与 MSD 和 CDH 方法比较接近, 但仍分别比 MSD 和 CDH 方法高出 0.1% 和 0.4%, 同时查准率明显高于其他 2 种特征描述子, 验证了本文方法的有效性。

如图 5 所示, 在 Corel-10000 图像库中, 检索前 10 幅图像时, 本文方法查准率达到 52.12%, 而其他 4 种描述子都没有超过 50%, 验证了本文方法的优越性。

表 2 给出了本文方法与 LBP 在 CIFAR-10 图



(a) 查准率和类别数的关系



(b) 查全率和类别数的关系

图 3 在 Corel-1000 图像库中每类图像的检索结果

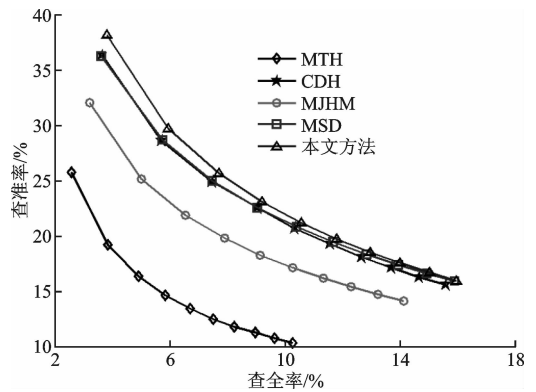


图 4 Corel-5000 图像库的检索性能比较

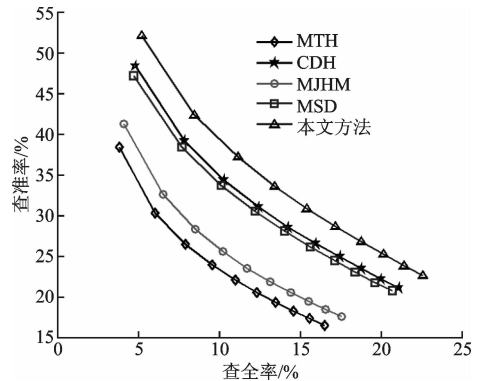


图 5 Corel-10000 图像库的检索性能比较

表 2 CIFAR-10 图像库中的检索性能比较

方法	查准率/%	查全率/%
LBP	32.69	0.05
本文	37.40	0.06

像库上的检索结果。在实验中,随机选择 5 000 幅作为查询图像,输出图像数设置为 10。由表 2 可以看出本文方法具有较优的性能。

表 3 比较了不同方法的特征向量维数和图像检索时间,在 Corel-10000 图像库上,本文方法的特征向量维数为 $d_H + d_S + d_V = 36 + 20 + 768 = 824$,在 Corel-1000 和 Corel-5000 图像库上,本文方法的特征向量维数为 $d_H + d_S + d_V = 108 + 20 + 768 = 896$ 。虽然本文方法的特征向量维数略高于其他方法,但本文方法的检索效果要优于其他方法。

表 3 不同方法的特征向量维数和图像检索时间

方法	特征向量维数	图像检索时间/s
MTH	82	0.021 2
CDH	108	0.026 1
MJHM	70	0.020 7
MSD	72	0.020 9
本文方法	824	0.039 3
本文方法	896	0.042 2

4 结 论

本文提出一种新的特征表示方法提取彩色图像的颜色和纹理特征。与在颜色特征的基础上统计基元结构的方法相比,本文在 HSV 空间中提取图像特征,同时考虑了图像的颜色和纹理信息,降低了颜色特征占主导因素的影响。另外,本文的基于局部结构的三结构描述子克服了局部二值模式只考虑图像中每个模式出现频率的问题,三结构描述子统计了局部结构像素间的相关性,保留了更多的区分信息。在图像库 Corel-1000、Corel-5000、Corel-10000 和 CIFAR-10 上的图像检索结果验证了本文方法的检索性能优于其他图像特征描述子。

参考文献:

[1] ELALAMI M E. A new matching strategy for content based image retrieval system [J]. Applied Soft Computing, 2014, 14(1): 407-418.

[2] SWAIN M J, BALLARD D H. Color indexing [J]. International Journal of Computer Vision, 1991, 7(1): 11-32.

[3] STRICKER M A, ORENGO M. Similarity of color images [C] // Proceedings of the SPIE, Storage and Retrieval for Image and Video Database III. Bellingham, WA, USA: SPIE, 1995: 381-392.

[4] HARALICK R M, SHANMUGAM K, DINSTEIN I H. Textural features for image classification [J]. IEEE Transactions on Systems, Man and Cybernetics, 1973, 3(6): 610-621.

[5] DE SIQUEIRA F R, SCHWARTZ W R, PEDRINI H. Multi-scale gray level co-occurrence matrices for texture description [J]. Neurocomputing, 2013, 120(10): 336-345.

[6] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987.

[7] JAIN A K, VAILAYA A. Image retrieval using color and shape [J]. Pattern Recognition, 1996, 29(8): 1233-1244.

[8] KAUPPINEN H, SEPPANEN T, PIETIKAINEN M. An experimental comparison of autoregressive and Fourier-based descriptors in 2D shape classification [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(2): 201-207.

[9] VIPPARTHI S K, NAGAR S K. Multi-joint histogram based modelling for image indexing and retrieval [J]. Computers and Electrical Engineering, 2014, 40(8): 163-173.

[10] LIU G H, LI Z Y, ZHANG L, et al. Image retrieval based on micro-structure descriptor [J]. Pattern Recognition, 2011, 44(9): 2123-2133.

[11] LIU G H, ZHANG L, HOU Y K, et al. Image retrieval based on multi-texton histogram [J]. Pattern Recognition, 2010, 43(7): 2380-2389.

[12] LIU G H, YANG J Y. Content-based image retrieval using color difference histogram [J]. Pattern Recognition, 2013, 46(1): 188-198.

[13] JULESZ B. Textons, the elements of texture perception, and their interactions [J]. Nature, 1981, 290(5802): 91-97.

(编辑 武红江)