

# 基于字典统计耦合归一化多重距离的图像检索算法

王荣芝<sup>1</sup>, 王莉<sup>2</sup>, 李佑婷<sup>3</sup>

(1.呼伦贝尔学院, 呼伦贝尔 021008; 2.北京遥感设备研究所, 北京 100195;  
3.内蒙古大学, 呼和浩特 010021)

**摘要:** 目的 针对当前彩色图像检索技术容易受到色彩的干扰, 使其鲁棒性不强等不足, 提出一种字典统计耦合归一化多重距离的彩色图像检索算法。**方法** 首先将图像量化并转换成一维信号, 然后引入字典统计, 将一维信号进行字典编码, 并计算编码后的图像多样值, 在归一化字典距离的基础上嵌入字典编码图像的多样值, 从而定义归一化多重距离(NMD)的相似度量准则, 利用NMD对查询图像与数据库图像的多样值进行比较与识别, 搜索出与查询图像具有相同特征的最相似图像, 完成目标检索。**结果** 在COREL数据库的实验结果表明, 相对于当前常用的检索技术, 该检索算法具有更高的查准率和查全率, 可对彩色图像完成精确检索, 有效减低了色彩对检索性能的干扰。**结论** 文中算法具有较好的检索精度, 能够较好地用于医疗、商标等领域的目标检索。

**关键词:** 图像检索; 字典统计; 归一化多重距离; 字典编码; 相似度量准则

**中图分类号:** TP391.3   **文献标识码:** A   **文章编号:** 1001-3563(2017)05-0228-06

## The Image Retrieval Algorithm Based on Dictionary Statistics and Normalized Multiple Distance

WANG Rong-zhi<sup>1</sup>, WANG Li<sup>2</sup>, LI You-ting<sup>2</sup>

(1.Hulunbuir University, Hulunbuir 021008, China; 2.Beijing Remote Sensing Equipment Research Institute, Beijing 100195, China; 3.Inner Mongolia University, Hohhot 010021, China)

**ABSTRACT:** The work aims to propose an image retrieval algorithm based on dictionary statistics and normalized multiple distance with respect to the disturbance easily caused to the current color image retrieval technology, thus leading to its inadequate robustness and other deficiencies. Firstly, the image was quantized and converted into a one-dimensional signal. Then the dictionary statistics were introduced to code the one-dimensional signals. The diversity value of the image were calculated. At last, on the basis of normalized dictionary distance, the diversity values of image obtained through dictionary coding were inserted and thus the similarity metric criterion of normalized multiple distance (NMD) was defined. NMD was adopted to compare and identify the diversity values of the query image and database image, and search out the image most similar to the query image with the same characteristics, so as to complete the retrieval task. Experimental results from the COREL database showed that: relative to the current commonly used retrieval techniques, the proposed retrieval algorithm had better precision ratio and recall ratio and could accurately retrieve color images, which effectively reduced the disturbance caused to retrieval performance by the color. In conclusion, the proposed algorithm has better retrieval precision and can be used for retrieval tasks in the fields of medical treatment and trademark, etc. in a better manner.

**KEY WORDS:** image retrieval; dictionary statistics; normalized multiple distance; dictionary coding; similarity metric criterion

随着近些年来摄像设备和互联网的快速发展, 数字图像数量呈现爆发性增长, 图像在生活中扮演的角

色越来越重要, 人们对图像信息的需求也日益增加, 获取感兴趣的图像信息难度升级, 图像检索的重要性

不断加大。传统的基于文本的图像检索算法存在工作量大, 具有主观性和不确定性, 无法满足现代检索功能, 因此, 基于图像内容的检索(CBIR)技术越来越受到关注, 已经成为当前图像检索的热点<sup>[1-3]</sup>。如, A.Hari等提出了一种基于局部二值模式(LBP)的图像检索算法<sup>[4]</sup>, 其通过计算中心像素与相邻像素的灰度值差异得到一个局部二进制模式, LBP对纹理区的每个像素都可以得到一个LBP编码。该技术具有计算简单、可适用性强, 广泛应用于模式识别、人脸识别和图像检索等领域, 但是LBP存在抗噪声能力低、光照变化不具备鲁棒性, 多尺度和维数相矛盾等缺点。陈文兵<sup>[5]</sup>等提出了一种基于轮廓特征点的图像检索算法, 该方法利用平均偏移提取兴趣特征, 以对象曲率的局部极值点作为信息点, 然后将信息点表示为特征向量, 利用查询图像与数据库中图像的特征向量间的距离匹配机制, 完成对象的检索任务。但是该算法对彩色图像不敏感, 对色彩丰富的彩色图像的检索率较低, 另外对于具有几何形变的图像以及特征不变性的图像检索性能很不稳定, 需要进一步的研究。吕彬<sup>[6]</sup>提出了一种基于全局颜色直方图的图像检索技术, 该技术提取颜色特征, 对特征点在图像中的位置关系分布进行相似性匹配, 能够对图像的旋转、平移和伸缩具有良好的检索性能。但是其无法反应颜色空间的位置关系, 如果检索图像具有相同颜色直方图那么可能会造成错误检索。

为解决彩色图像检索技术中查准率和查全率不高, 算法鲁棒性不强, 容易受到色彩的干扰和影响, 导致彩色图像检索算法性能不强等问题, 对此, 基于字典理论基础上, 提出了一种字典统计与归一化多重距离的彩色图像检索算法。将图像转换为一维信号, 利用字典统计特性进行编码并计算编码后图像多样值, 然后定义了NMD的相似度量准则, 根据字典的内部统计特性, 利用NMD对查询图像与数据库图像的多样值完成检索任务。在Corel数据库中测试了文中算法和对比算法的检索性能。

## 1 字典统计理论

字典理论是利用字典作为一个不同的单词的集, 每一个都定义为独特的字符序列<sup>[7]</sup>。2个字典之间的相似度根据它们之间相同单词的数量进行衡量。2个字典仅当它们具有不同单词集一样时是相等的。在文中, 利用内部字典结构分配给每个提取的字一个附加的参数, 其具有多样性。这个新的参数表示单词的重要性并且它有助于(在压缩中)字符串的编码过程。为了定义新的参数从而进一步分析字典生成过程。对于一个简单字典, 见图1a, 由通过3个字符组成{a, b, c}, 并且使得它们的基数等于8(包含8个不同的单

词)。图1b中的树状结构显示了其可视化过程。第一排的结构包含字典集的输入字符。这些构成了源词的初始集合和它们的数目被确定, 等于系统的量化等级。每一个单词度代表一个节点, 然而来源和产品之间的关联是通过一个边缘来表示的。根据这些初始集合, 每一个单词至少拥有一个字符, 同时产品单词可以被用于进一步产生后面的任何单词。如图1b中, 根据源单词{a}, 可产生2个字符的单词{aa}, 并且进一步用于生成含有3个字符的单词{aab}和{aac}, 因此, 在字典中每个单词都具有多样值, 并且与生成该单词的边缘联结的数量匹配。

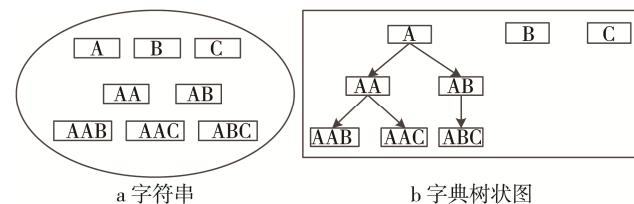


图1 字典结构

Fig.1 Dictionary structure

对于一个单词, 如果其多样值越大, 说明其包含的信息内容越丰富, 能够为更复杂单词的产生建立基础。当对2个字典比较时, 如果共有的单词具有巨大的多样值时, 那么该单词特性需特别注意。为了完成这套理论框架, 引入多样集。多样集是一个包含其自身元素多样性额外信息的集合。在文中, 利用2维数组表示字典, 其可以同时适应单词的集合和多样值。设x为一个字符串, 并且  $D(x)=\{w_1, w_2 \dots w_n\}$  为n个单词的无序集合, 其中:  $w_i$  为第i次提取的不同单词。多样集定义为:

$$MS(x)=\{D(x), m_x\}=\{(w_1, m_x(w_1)), (w_2, m_x(w_2)) \dots (w_n, m_x(w_n))\} \quad (1)$$

式中:  $m_x(w_i)$  为第i个单词的多样性, 表示第i个二维数组的( $w_i, m_x(w_i)$ )。在多样集中, 对多样值为0的源单词更新, 将多样值增加为1, 见图1b。多样集的基数为所有多样集的和, 其计算见式(2)。

$$\sum_{i=1}^n m_x(w_i)=|MS(x)| \quad (2)$$

根据多样字典理论, 设2个多样集  $MS(x)=\{D(x), m_x\}$  和  $MS(y)=\{D(y), m_y\}$ , 并定义为定义1和定义2。

定义1。并联多样集  $MS(z)=\{D(z), m_z\}$ , 式中:  $D(z)=D(x) \cup D(y); m_z=\max(m_x(w_i), m_y(w_i))$ 。

定义2。交叉多样集  $MS(z)=\{D(z), m_z\}$ , 式中:  $D(z)=D(x) \cap D(y); m_z=\min(m_x(w_i), m_y(w_i))$ 。

根据定义1和2, 并联多样集和交叉多样集是相互独立的操作, 例如:  $MS(x) \cup MS(y)=MS(y) \cup MS(x)$  和  $MS(x) \cap MS(y)=MS(y) \cap MS(x)$ 。如果2个多样集是完全相同的, 那么产生的结果与该多样集相同, 例如:  $MS(x) \cap MS(x)=MS(x)$  和  $MS(x) \cup MS(x)=MS(x)$ 。此外,

对于定义 1 和 2 来说，多样集满足单调性和分配性。

单调性，并联多样集  $MS(z)=\{D(z), m_z\}$  比单一的多样集更大，表示为  $|MS(z)|=|MS(x) \cup MS(y)| \geq |MS(x)|$ 。其中， $|MS(z)| \neq \emptyset$ ， $D(z)=D(x) \cup D(y) \leq D(x)+D(y) \leq D(x)$ ，其中基数通常是一个正数，如  $|MS(z)| \geq 0$ 。

分配性，设  $MS(x)=\{D(x), m_x\}$ ， $MS(y)=\{D(y), m_y\}$  和  $MS(z)=\{D(z), m_z\}$  为 3 个多样集，2 个多样集的并联集加上第 3 个多样集，总是小于或等于 2 个并联多样集之和，即  $|MS(x) \cup MS(y)|+|MS(z)| \leq |MS(x) \cup MS(z)|+|MS(y) \cup MS(z)|$ 。

多样集原理对于表示一个字典来说，是一种强有力得工具，它结合了已提取的压缩信息和每一个单词的统计信息，并且可运用到其他类似于以字典为基础的相似性度量。由此，可通过某些方法构建一个集，在该集中每一个单词都可以通过一个对应其多样集的确定值来直接加权。

## 2 归一化多重距离

根据归一化字典距离概念<sup>[8-9]</sup>，在文中设了一种归一化多重距离(NMD)的作为一种新的异于以字典为基础的度量方法。假设  $w_i$  为字典中的第  $i$  个单词，其多样值为  $m(w_i)$ ，并提取单词的总数  $\sum_{i=1}^n m(w_i) = |MS(x)|$ 。多样值表示的是从压缩过程中提取其关联词的出现次数。归一化的多重距离(NMD)定义见式(3)。

$$NMD(x, y) = \frac{|MS(x) \cup MS(y)| - \min\{|MS(x)|, |MS(y)|\}}{\max\{|MS(x)|, |MS(y)|\}} \quad (3)$$

式中： $|MS(x)|$  和  $|MS(y)|$  分别为压缩字符串  $x$  和  $y$  的基数， $|MS(x) \cup MS(y)|$  是 2 个字符串的并联多样集基数。与 NDD 类似，NMD 也是一种相似度量，其值属于  $[0, 1]$ ，并且满足以下特性：非负性， $NMD(x, y) \geq 0$ ；幂等性， $NMD(x, y) = 0$ ；对称性，通过式(3)的多重集互换，NMD 保持不变，例如  $NMD(x, y) = NMD(y, x)$ ，因为这 2 个多重集的并联集是相互独立的；三角不等式，假设 3 个多重集  $MS(x)$ ， $MS(y)$ ， $MS(z)$  2 个多重集之间的差异值总是小于或者等于其他多重集的差异值之和，例如  $NMD(x, y) \leq NMD(x, z) + NMD(y, z)$ 。

不同于归一化字典距离，为了改变 2 个相似字典的相似程度，归一化多重距离通过提取单词的多样值表示字典的内部统计特性。如图 2 所示，2 个字典用树状图来表示，并且由 3 个相同的字符组成，这 3 个字符为(A, B, C)，如图 2a 中，一个源词 {A} 可生成 2 个字符的单词 {AA}，并且进一步用于生成含有 3 个字符的单词 {AAB} 和 {AAC}，形成了包含 8 个单词。图 2b 中，同样的源词 {A} 可生成 2 个字符的单词 {AA}，并且进一步用于生成含有 3 个字符的单词 {AAA} 和 {AAC}。由此可见，这 2 个字典中总共包括 7 个相同

的单词和一个不同的单词，具有相同的重复单词，2 个字典形成了多样集，为了满足基于字典技术的多样集的限制，将实时更新每个字典的多样值。

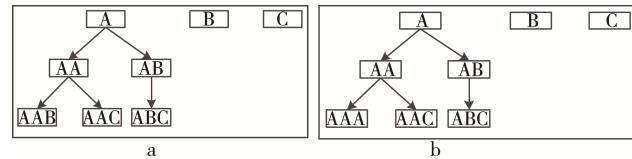


图 2 归一化多重距离字典结构  
Fig.2 Normalized multiple distance dictionary structure

## 3 文中检索算法

根据在前面 2 部分分析，文中提出了一种基于多信息提取的彩色图像检索技术。算法流程见图 3。主要由查询图像  $Q$  的特征提取和 NMD 相似度量 2 部分组成。

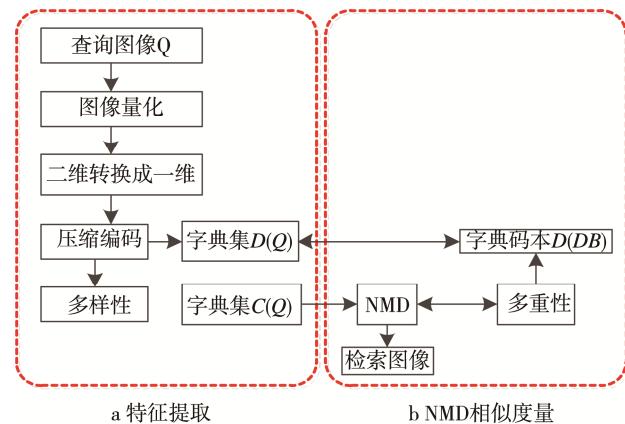


图 3 文中算法流程  
Fig.3 Flow chart of this algorithm

为了避免表示完整的色彩空间，对图像进行量化。通过利用单一字符替代每个像素来量化初始三维 RGB 色彩空间。为简化操作，在 RGB 色彩空间进行量化。被量化的图像转换为一维字符串  $x$ 。通过压缩编码对串行数据进行操作，利用光栅扫描方法，即以一排或一列为基础来扫描一个图像，通过图像平面中的像素扫描整个图像，最典型的技术为 Hilbert-Peano 曲线和 SCPO 2 种<sup>[10-11]</sup>。2 种光栅原理分别为行主导和列主导。行主导是将像素的强度给一字符串，一行接一行，从左上角开始，然后继续经过行，再继续下一个。以同样的方式，列主导开始于左下角像素并且继续上升到第 1 列，在移到右边之前，一列接一列的运行，并且运用于图像中，得到图像的一维序列。通过得到的一维序列，利用 LZW 压缩编码器提取相对应的多样集字典。

$$MS(Q)=\{D(Q), m_Q\} \quad (4)$$

式中： $D(Q)$  为字典集； $m$  为多样值。则相同的操作可在脱机状态下从数据库图像中重复提取多样集

$MS(Q)$ 。

在第2阶段, 根据引入的NMD相似度量, 比较查询图像 $MS(Q)$ 的多样集与图像数据库多样集关系, 根据NMD获得的相似度量可表示为:

$$D(Q,P)=\omega_1 D(Q,P)+\omega_2 D(Q,P) \quad (5)$$

式中:  $\omega_1$ 和 $\omega_2$ 分别为色彩调整特征度量距离和纹理特征度量距离的权重, 通过调整 $\omega_1$ 和 $\omega_2$ , 可突出重要特征和显著性。

## 4 实验与讨论

为验证提出的图像检索算法的有效性和准确性, 选用最常用的COREL数据库作为训练样本<sup>[12]</sup>。从COREL数据库中选择8种类别图像, 每种类别各含有50幅图像组成的400幅图像数据库DB, 8种类别图像分别为马、大象、海滩、花朵、建筑、恐龙、非洲和公交车, 见图4。为增加实验的说服力, 设置当前图像常用的图像检测算法(文献[4]、文献[5]和文献[6])作为对照。

### 4.1 性能评价指标

为了定量评价图像检索性能, 文中引入当前图像检索中广泛应用的评价准则: 查准率、平均查准率、查全率以及平均查全率<sup>[13—14]</sup>。假设查询数据库中的第*i*个图像 $Q_i$ , 查准率 $P_k(Q_i)$ 计算公式为:

$$P_k(Q_i)=\frac{N_r}{N} \quad (6)$$

式中:  $N_r$ 为正确检索图像的数量;  $N$ 为总图像数量。对于一个数据库 $N_q$ 的查询, 其平均查准率 $\bar{P}_k$ 定义见式(7)。

$$\bar{P}_k=\frac{1}{N_q} \sum_{i=1}^{N_q} P_k(Q_i) \quad (7)$$

查全率 $R_k(Q_i)$ 是正确检索图像与所有检索图像的比值, 其计算见式(8)。

$$R_k(Q_i)=\frac{N_r}{N_c} \quad (8)$$

式中:  $N_c$ 为所有检索得到图像的总数。对于数据集 $N_q$ 的平均查全率 $\bar{R}_k$ 见式(9)。

$$\bar{R}_k=\frac{1}{N_q} \sum_{i=1}^{N_q} R_k(Q_i) \quad (9)$$

### 4.2 实验结果与分析

3种检索算法对查询图像大象的检索结果显示, 文中算法的检索精度最高, 误检率较低, 从数据库图像中检索出的相似图像完整, 检索图像与查询图像的相似性高。而对照组3种技术的结果都不是很好, 检索图像与查询图像出现一些差别。

为了量化这些算法的检索精度, 文中引用P-R曲



图4 COREL数据库图像

Fig.4 COREL database image

线<sup>[15]</sup>来反应图像检索性能, 在控制查全率的情况下, 可获取P-R曲线对应的查准率,  $P$ 值越高, 说明在相同的查全率下检索的准确率越高, 检索性能越好。文中算法与文献[4]、文献[5]和文献[6]的P-R曲线见图5。从图5中可知, 文中检索算法P-R曲线表现最好, 说明文中的检索性能相对其他算法优异。

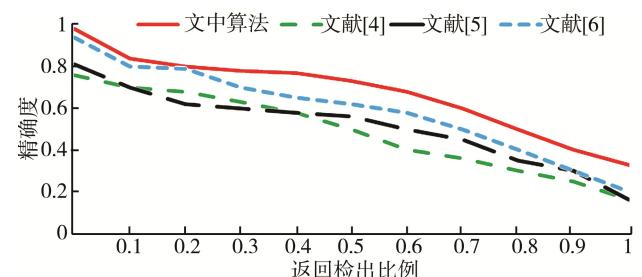


图5 4种算法的P-R曲线

Fig.5 P-R curves of the four algorithms

为了进一步证明在彩色图像检索任务的有效

性,根据检索返回图像数量,在COREL数据库进行了多个实验,分别测得不同返回图像的平均查准率和平均查全率,见图6。表1为不同类型图像的返回图像数  $k=10$  个最相似图像的查准率,实验结果见表2。表2为不同类型图像的返回图像数  $k=50$  个最相似图像的平均查准率,实验结果见表2。根据表1和表2得出,文中算法在检索数量  $k=10$  和  $k=50$  都具有明显的优势,算法的查准率相对其他3种算法在相同类别图像中较好。对于花朵、公交、恐龙和大象等类别,其检索的准确率更高,因为这些类别含有特定的形状和纹理信息,特性能够更容易被算法提取。

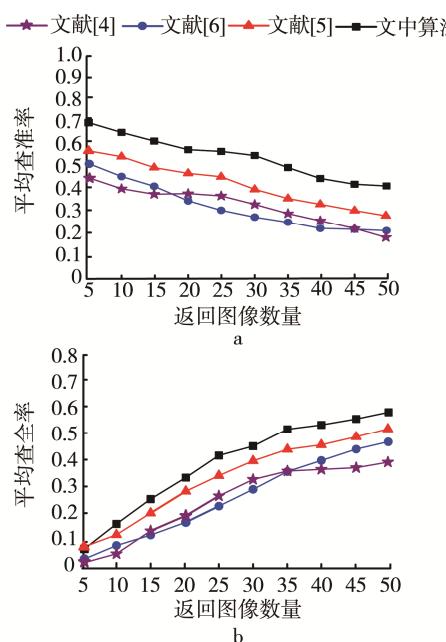


图6 各算法的平均查准率与平均查全率测试结果  
Fig.6 The average precision of each algorithm and the average recall test results

表1  $k=10$  条件下的3种算法的查准率结果  
Tab.1 The three algorithms under the condition of  $k=10$  precision results

分类	查准率/%			
	文中算法	文献[4]	文献[5]	文献[6]
马	69.56	45.64	51.23	44.35
大象	82.42	65.82	71.63	72.63
海滩	62.69	44.33	46.21	45.65
花朵	92.51	72.14	76.38	78.22
建筑	63.47	40.27	43.66	55.89
恐龙	84.20	61.44	65.25	70.19
非洲	52.98	38.56	42.85	48.41
公交车	85.61	70.39	75.45	76.52

表2  $k=50$  时查准率结果Tab.2 The three algorithms under the condition of  $k=50$  precision results

类别	查准率/%			
	文中算法	文献[4]	文献[5]	文献[6]
马	61.43	42.75	49.46	42.89
大象	79.44	62.40	68.12	69.31
海滩	55.21	42.19	43.87	44.85
花朵	89.15	68.22	74.54	75.84
建筑	59.12	38.16	42.27	50.59
恐龙	79.56	58.85	62.77	67.40
非洲	45.36	35.75	36.63	42.52
公交车	80.58	65.60	71.45	73.18

综合上述实验结果和定量评价指标结果可知,与对照组算法相比,所提算法具有较高的检索精度,具有更好的P-R特性曲线。在检索图像数目  $k=10$  和  $k=50$  的实验中,文中算法得到的平均查准率更高。其主要原因是文中算法对图像进行多特征提取,将图像量化,转换为一维信号字符串,利用基于字典统计特性进行编码,计算编码后图像特征值;然后定义了NMD的相似度量准则,NMD通过提取字典的内部统计特性,对查询图像与数据库图像的特征值进行比较,很好地完成检索任务。

## 5 结语

基于字典统计理论,设计了一种字典统计耦合归一化多重距离的图像检索算法。利用字典统计特性进行图像编码并计算编码后图像多样值,根据多样值作为彩色图像检索特征向量,设计了归一化多重距离相似度量,从图像库中检索出与查询图像具有相同特征值的相似图像。在COREL数据库中进行测试,设置当前常用的检索算法作为对照组实验,借助最常用的查准率、查全率、平均查准率和平均查全率作为算法性能评价指标。实验结果表明了所提出技术的评价指标要优于其他文献方法,具有更高的检索精度,能够有效地对彩色图像完成检索。

## 参考文献:

- [1] JOUILI S, TABBONE S. Hypergraph-based Image Retrieval for Graph-Based Representation[J]. Pattern Recognition, 2012, 45(11): 4054—4068.
- [2] 田靓. 基于颜色内容的图像检索方法的比较[J]. 包装工程, 2009, 30(4): 84—86.
- [3] TIAN Liang. Comparison of Image Retrieval Methods Based on Color Content[J]. Packaging Engineering, 2009, 30(4): 84—86.
- [4] 毋小省, 孙君顶. 基于感兴趣区域的图像检索方法

- [J]. 激光与红外, 2013, 43(2): 221—224.
- WU Xiao-sheng, SUN Jun-ding. Image Retrieval Based on Region of Interest[J]. Laser & Infrared, 2013, 43(2): 221—224.
- [4] REDDY A H, CHANDRA N S. Local Oppugnant Color Space Extrema Patterns for Content Based Natural and Texture Image Retrieval[J]. AEU-international Journal of Electronics and Communications, 2014, 69(1): 290—298.
- [5] 陈文兵, 成海燕, 陈允杰. 一种基于轮廓特征点的图像检索方法[J]. 计算机工程, 2012, 38(12): 197—200.
- CHEN Wen-bing, CHENG Hai-yan, CHEN Yun-jie. Image Retrieval Method Based on Contour Feature Point[J]. Computer Engineering, 2012, 38(12): 197—200.
- [6] 吕彬. 基于全局直方图及其空间关系的图像检索技术[D]. 太原: 太原理工大学, 2014.
- LYU Bin. Based on Global Histogram and Spatial Relationships Image Retrieval Technology[D]. Taiyuan: Taiyuan University of Technology, 2014: 23—25.
- [7] 周同驰, 程旭, 吴镇扬. 分层树结构字典编码的行为识别[J]. 中国图象图形学报, 2014, 19(7): 1054—1061.
- ZHOU Tong-chi, CHENG Xu, WU Zhen-yang. Action Recognition Using Hierarchically Tree-Structured Dictionary Encoding[J]. Journal of Image and Graphics, 2014, 19(7): 1054—1061.
- [8] COHEN A R, VITANYI P M. Normalized Compression Distance of Multisets with Applications[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(8): 1602—1614.
- [9] COHEN A R, VITANYI P M B. Normalized Compression Distance of Multiples[J]. Physics, 2012, 41(13): 121—130.
- [10] COSTA P, BARROSO J, FERNANDES H. Using Peano-Hilbert Space Filling Curves for Fast Bidimensional Ensemble EMD Realization[J]. EURASIP Journal on Advances in Signal Processing, 2012(1): 113—125.
- [11] YANG P, XIE L, LIU J. Simultaneous Celestial Positioning and Orientation for the Lunar Rover[J]. Aerospace Science & Technology, 2014, 34(4): 45—54.
- [12] BESIRIS D, ZIGOURIS E. Dictionary-based Color Image Retrieval Using Multiset Theory[J]. Journal of Visual Communication and Image Representation, 2013, 24(7): 1155—1167.
- [13] RAO L K, RAO D V. Local Quantized Extrema Patterns for Content Based Natural and Texture Image Retrieval [J]. Human-centric Computing and Information Sciences, 2015, 26(5): 1—24.
- [14] 邹彬, 潘志斌, 胡森. 基于局部投影与块 LBP 特征的图像检索[J]. 中国图象图形学报, 2012, 17(6): 671—677.
- ZOU Bin, PAN Zhi-bin, HU Sen. Image Retrieval Method Based on Local Projection and Block LBP Feature[J]. Journal of Image and Graphics, 2012, 17(6): 671—677.
- [15] 杨叶芬, 曾东海, 刘海. 逆向学习耦合多属性查询的图像排序/检索优化算法研究[J]. 包装工程, 2015, 36(7): 84—90.
- YANG Ye-fen, ZENG Dong-hai, LIU hai. Research on Image Ordering/Retrieval Optimization Algorithm for Inverse Learning Coupled Multi Attribute Query[J]. Packaging Engineering, 2015, 36(7): 84—90.