

文章编号: 1004-4353(2017)02-0167-06

基于多样字典理论与多尺度距离度量的 彩色图像检索

梁 晨

(齐鲁师范学院 信息科学与工程学院, 山东 济南 250200)

摘要: 针对在图像检索中因色彩因素导致的相关算法正确率低、稳定性差等问题, 提出了一种多样字典理论与多尺度距离度量的彩色图像检索算法. 首先, 对输入图像进行量化, 将其转换为一维字符串形式; 其次, 采用多样字典统计对图像视觉模式编码, 并计算编码后的图像特征值; 最后, 给出多尺度距离的相似度量准则, 并根据该准则对查询图像与数据库图像的特征值进行处理, 寻找与其匹配的特征图像. 实验结果表明: 本文所提出的算法在查准率与查全率上要优于当前流行的检索方法, 其对彩色图像的检索精度和稳定性也有了明显提高, 因此具有较好的应用价值.

关键词: 图像检索; 多样字典理论; 多尺度距离度量; 色彩干扰

中图分类号: TP391.4

文献标识码: A

The color image retrieval algorithm based on multiple dictionary theory and multi-scale distance measure

LIANG Chen

(Institute of Information Science and Engineering, Qilu Normal University, Jinan 250200, China)

Abstract: A new color image retrieval algorithm based on the theory of multiple dictionaries and multi-scale distance measure was proposed to improve the low accuracy and instability which are the problems caused by color factor. Firstly, quantifying the input image and converting it to a string form; secondly, encoding the image visual pattern by calculating the multi-dictionaries, and then evaluating encoded image feature value. Finally, this paper presents a similarity metric for multi-scale distance. According to the criterion, the characteristic values of the query image and the database image are processed, and then a matching image is found to complete the search task. Experimental results shows that this algorithm proposed in this paper was superior to the current popular methods in precision and recall level and the accuracy and stability of the color image retrieval has been significantly improved. Therefore, the algorithm has good application value.

Keywords: image retrieval; multiple dictionary theory; multiple distance measure; color interference

0 引言

随着传感器技术的迅速发展, 数字图像数量呈爆发式增长, 面对海量的图像, 人们获取所需的图像的难度不断加大, 因而研究图像自动检索具有重要意义. 由于基于文本的图像检索技术具有

一定的主观性与不确定性, 目前已经无法满足现有的检索要求, 因此, 基于图像内容的检索方法 (CBIR) 越来越受到人们的重视^[1-3]. 近年来, 许多学者对图像检索进行了研究, 并取得了一些成果. 例如: 文献[4]提出了一种基于 LBP 的图像检索

方法,该方法通过计算中心像素与邻像素的灰度值差异得到一个局部二进制, LBP 对区域中的每个像素都可以获得一个 LBP 编码;该方法计算简单、适用性强,但缺点是 LBP 对噪声敏感,受光照变化影响较大,并且多尺度与维数相矛盾. 文献[5]提出了一种基于轮廓特征点的图像检索技术,其通过平均偏移提取兴趣特征,将对象曲率的局部极值点作为信息点,并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实现目标检索;该方法的缺点是对彩色图像不敏感,而且对几何形变与特征不变性图像的检索性能不稳定. 文献[6]提出了一种全局颜色直方图的图像检索方法,其通过提取颜色特征,利用特征点在图像中的位置关系进行相似性匹配;该方法对图像的旋转、平移和伸缩具有良好的检索率,但是该方法无法准确表示颜色空间关系,无法有效检索具有相似颜色的图像. 为了进一步解决彩色图像检索中检索率不高、容易受到色彩的干扰等问题,本文提出了一种多样字典理论与多尺度距离度量的彩色图像检索算法,并通过实验验证了本文方法的检索性能.

1 多样字典理论

对于两个字典,其相似度一般根据它们之间相同单词的数量进行度量^[7]. 在本文中,通过内部字典结构对每个提取的字分配一个附加参数(参数具有多样性),该参数表示单词的重要性,并且它有助于字符串的编码过程. 下面,根据定义的参数,分析多样字典形成过程.

如图 1(a) 所示,对于一个简单的字典,其可通过 $[a,b,c]$ 3 个字符组成,使其基数等于 8 即可. 图 1(b) 中的树状结构表示多样字典的可视化过程,其第一排的结构为字典集的字符,它构成了源词的初始集合. 每个单词代表一个节点,字符与单词之间的关系通过一个边缘来表达. 图 1 中,每个单词至少包含了一个字符,同时生成的单词也可以被当作一个字符,从而进一步生成后续的其他单词. 如图 1(b) 中,根据源单词 $[a]$,可生成含有 2 个字符的单词 $[aa]$,进一步还可以得到包含 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,并且与形成该单词的字符数量有关.

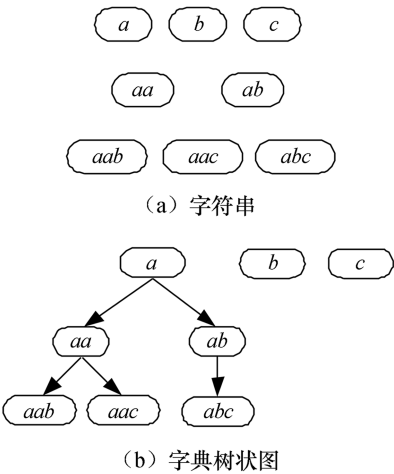


图 1 字典结构

对于一个单词,如果多样值越大,那么其包含的信息越丰富;因此,当比较两个字典时,如果共有的单词具有较大的多样值时,需特别注意该单词. 为了完善多样字典理论,本文引入了多样集.

多样集是一个含有其自身单词多样性额外信息的集合. 对于一个字典来说,多样集是一个非常强有力的工具,它结合了已提取的压缩信息和每个单词的统计信息,并且可以运用到其他类似于以字典为基础的相似性度量. 本文中,通过二维数组表示字典,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. 因此,多样集定义如下:

$$MS(x) = \{D(x), m_x\} = \{(\omega_1, m_x(\omega_1)), (\omega_2, m_x(\omega_2)), \dots, (\omega_n, m_x(\omega_n))\}, \quad (1)$$

式中 $m_x(\omega_i)$ 表示第 i 个单词的多样值, $(\omega_i, m_x(\omega_i))$ 表示第 i 个二维数组的多样集. 在多样集中,对多样值为 0 的源单词进行更新,将多样值增加为 1,如图 1(b) 所示. 多样集的基数为所有多样集的和,其计算公式如下:

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

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:

1) 交叉多样集 $MS(z) = \{D(z), m_z\}$, 其中:
 $D(z) = D(x) \cap D(y)$, (3)

$$m_z(z) = \min(m_x(w_i), m_y(w_i)). \quad (4)$$

2) 并联多样集 $MS(z) = \{D(z), m_z\}$, 其中:

$$D(z) = D(x) \cup D(y), \quad (5)$$

$$m_z(z) = \max(m_x(w_i), m_y(w_i)). \quad (6)$$

根据式(1)和(2)知, 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 $MS(x) \cap MS(y) = MS(y) \cap MS(x)$. 如果两个多样集完全相同, 那么产生的结果与该多样集相同, 例如: $MS(x) \cap MS(x) = MS(x)$, $MS(x) \cup MS(x) = MS(x)$. 此外, 多样集满足单调性和分配性.

1) 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, 两个多样集的并联集加上第 3 个多样集之和总是小于或等于 2 个并联多样集之和, 表示为:

$$|MS(x) \cup MS(y)| + |MS(z)| \leq |MS(x) \cup MS(z)| + |MS(y) \cup MS(z)|. \quad (7)$$

2) 单调性. 并联多样集 $MS(z) = \{D(z), m_z\}$ 比单一多样集大, 表示为

$$|MS(z)| = |MS(x) \cup MS(y)| \geq |MS(x)|, \quad (8)$$

其中 $|MS(z)| \neq \emptyset$, $D(z) = D(x) \cup D(y) \leq D(x) + D(y) \leq D(x)$, 基数通常是一个正数, 如 $|MS(z)| \geq 0$.

2 多重距离度量

根据多样字典理论^[8-9], 本文定义了一种基于多样字典理论的多重距离测量(NMD).

设 w_i 为字典中的第 i 个单词, 其多样值为 $m(w_i)$, 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)定义如下:

$$NMD(x, y) = \frac{|MS(x) \cup MS(y)| - \min\{|MS(x)|, |MS(y)|\}}{\max\{|MS(x)|, |MS(y)|\}}, \quad (9)$$

其中 $|MS(x)|$ 和 $|MS(y)|$ 分别为压缩字符串 x 和 y 的基数, $|MS(x) \cup MS(y)|$ 是两个字符串的并联多样集基数. 与归一化字典类似, NMD 也是一种相似度量, 其值属于 $[0, 1]$, 并且具有下列特点:

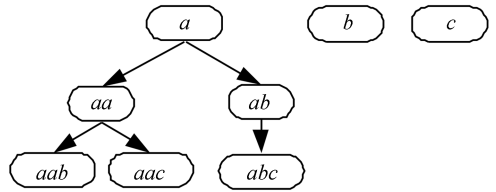
1) 对称性. 通过公式(9)的多重集互换, NMD 保持不变, 例如: $NMD(x, y) = NMD(y, x)$, 因为这两个多重集的并联集是相互独立的.

2) 非负性. $NMD(x, y) \geq 0$.

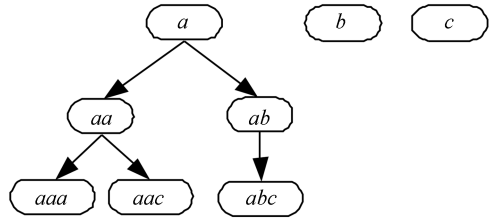
3) 三角不等式性. 假设 3 个多重集($MS(x)$, $MS(y)$ 和 $MS(z)$) 中两个多重集之间的差异值总是小于或者等于其他多重集的差异值之和, 表示为 $NMD(x, y) \leq NMD(x, z) + NMD(y, z)$.

4) 幂等性. $NMD(x, y) = 0$.

为了改变两个相似字典的相似程度, NMD 通过提取单词的多样值表示字典的内部统计特性. 将用 3 个相同的字符(a, b, c) 组成的两个字典用树状图来表示, 如图 2 所示, 可知这两个字典中总共包括 7 个相同的单词和一个不同的单词, 但它们的基数仍等于单词的总数.



(a) 字典树状图



(b) 字典树状图

图 2 NMD 字典结构

特征提取后, 得到查询图像 Q 的特征向量 $f_Q = (f_{Q1}, f_{Q2}, \dots, f_{QL_g})$. 同样, 数据库中的每个图像的特征向量都可表示为 $f_{DB_k} = (f_{DB_{k1}}, f_{DB_{k2}}, \dots, f_{DB_{kL_g}})$, $k = 1, 2, \dots, DB$. 通过数据库中的图像特征的匹配, 选择与检索图像最相似的 n 个图像. 为了匹配图像, 引入相似距离进行度量 $D(Q, DB)$, 其定义如下:

$$D(Q, DB) = \sum_{i=1}^{L_g} \left| \frac{f_{DB_{ki}} - f_{Q_i}}{1 + f_{DB_{ki}} + f_{Q_i}} \right|, \quad (10)$$

式中 Q 和 DB 分别表示查询图像和数据库图像, $f_{DB_{ki}}$ 表示数据库 DB 中第 k 个图像的第 i 个特征, f_{Q_i} 为查询图像 Q 的第 i 个特征.

3 本文算法机制

本文提出的多样字典理论与多尺度距离度量的彩色图像检索算法流程如图 3 所示. 算法的主要步骤是: 首先, 对输入图像进行量化, 将其转换为一维字符串形式; 然后, 利用字典统计对图像视觉模式编码, 并计算编码后的图像特征值; 最后, 设计多尺度距离的相似度量准则, 根据设计的多尺度距离测量对查询图像与数据库图像的特征值进行匹配, 寻找与查询图像具有相同特征图像, 完成检索任务.

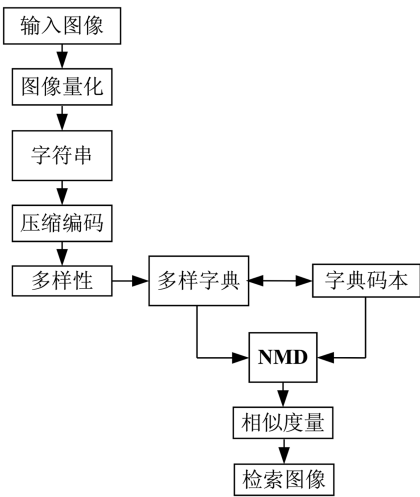


图 3 本文算法流程图

为了有效地表示完整的色彩空间, 需要对图像进行量化. 对于三维 RGB 色彩空间, 通过单一字符替代每个像素, RGB 被转换为一维字符串 x . 通过压缩编码对串行数据进行操作, 利用 Hilbert-Peano 曲线和 SCPO 光栅中的扫描方法^[10-11]对像素扫描, Hilbert-Peano 与 SCPO 两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值一行接着一行扫描, 从左上角开始, 然后继续经过下一行. 列主导开始于左下角像素并且继续上升到第一列, 在移到右边之前, 一列接一列地运行, 并且运用于图像中, 得到图像的一维序列.

得到一维序列后, 利用 LZW 压缩编码器提取相对应的多样集字典:

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

式中 $D(Q)$ 表示字典和, m 为多样值. 由此可知, 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 $M(DB)$.

提取多样集后, 再根据引入的 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的关系. 根据 NMD 计算的相似度定义如下:

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

其中 ω_1 和 ω_2 分别表示色彩调整特征度量距离和纹理特征度量距离的权重, 通过调节 ω_1 和 ω_2 可以突出色彩和细节等重要特征.

4 实验与分析

利用常用的 COREL 数据库^[12] 验证本文算法的性能, 并分别与文献[4](算法 I)、文献[5](算法 II) 和文献[6](算法 III) 的方法进行对比. 在 COREL 数据库中选择 8 种类别彩色图像, 每种类别各含有 40 幅图像, 由此组成 320 幅图像数据库 DB. 8 种类别图像分别为: 大象、海滩、飞机、花朵、建筑、恐龙、非洲人和公交车, 如图 4 所示.

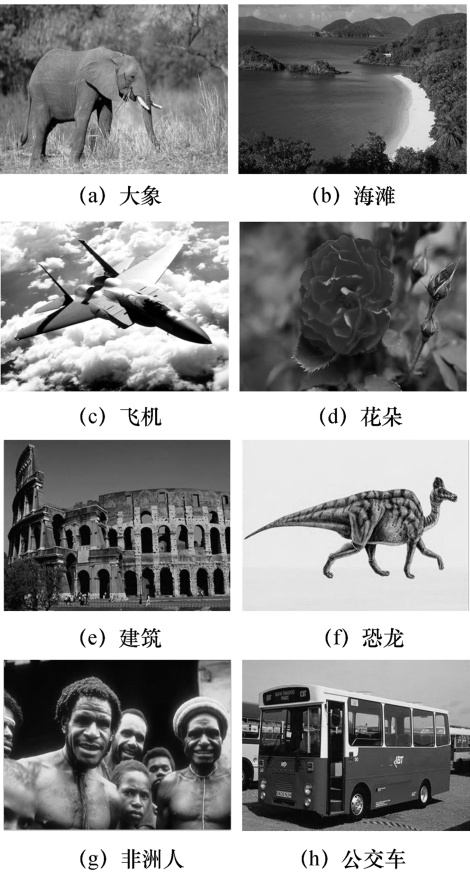


图 4 8 种类别图像

为了客观地评价各方法的检索性能, 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.

设查询数据库中的第 i 个图像为 Q_i , 其查准率 $P_k(Q_i)$ 的计算公式^[8] 为

$$P_k(Q_i) = \frac{N_r}{N}, \tag{11}$$

其中 N_r 是正确检索图像的数量, N 为总图像数量. 定义数据库 N_q 的平均查准率 \bar{P}_k ^[13] 为

$$\bar{P}_k = \frac{1}{N_q} \sum_{i=1}^{N_q} P_k(Q_i). \tag{12}$$

查全率 $R_k(Q_i)$ 的计算公式为

$$R_k(Q_i) = \frac{N_r}{N_c}, \tag{13}$$

其中 N_c 是所有检索得到的图像总数. 数据集 N_q 的平均查全率 \bar{R}_k 表示为

$$\bar{R}_k = \frac{1}{N_q} \sum_{i=1}^{N_q} R_k(Q_i). \tag{14}$$

图 5、图 6、图 7 和图 8 分别为本文算法、算法 I、算法 II 和算法 III 对查询图像飞机的检索结果. 由图 5 可以看出, 本文算法检索出的相似图像完整, 查询图像特征与所检索图像的特征点有较好的匹配度, 相似性高, 如图中黑色框区域所示. 从图 6、图 7 和图 8 中可以看出, 其检索出的匹配特征点较少, 如图中的机翼、机头, 检索图像与查询图像出现差别, 相似度不高, 说明这 3 种对照组检索算法检索的结果并不是很理想.



图 5 本文算法的检索结果

查准-查全曲线能综合体现图像的检索性能. 在控制查全率的情况下, 曲线可得到相应的查准率, 曲线越高, 说明在相同的查全率下检索的准确率越高, 检索性能越好. 图 9 为本文算法与算法 I、算法 II 和算法 III 的 P - R 曲线图. 由图 9 可以看出, 本文检索算法的 P - R 曲线表现最好, 说明本文的检索性能优于其他算法的性能.



图 6 文献[4]的检索结果



图 7 文献[5]的检索结果



图 8 文献[6]的检索结果

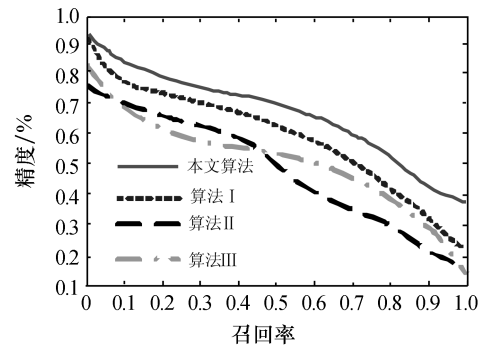


图 9 4 种算法的 $P-R$ 曲线图

为了进一步验证本文算法对彩色图像检索的性能,在 COREL 数据集上进行 2 组实验:在实验 1 中,检索 $k=10$ 个最相似图像的平均查准率,结果见表 1 所示;在实验 2 中,检索 $k=100$ 个最相似图像的平均查准率,结果见表 2 所示。

表 1 $k=10$ 的查准率结果

分类	查准率($k=10$)/%			
	本文算法	算法 I	算法 II	算法 III
飞机	70.12	43.37	50.18	45.16
大象	81.20	64.68	70.63	71.52
海滩	61.86	45.10	45.38	46.14
花朵	91.36	71.88	75.71	77.45
建筑	64.35	41.32	44.85	54.98
恐龙	85.14	60.33	66.16	71.35
非洲人	53.11	39.76	43.75	49.15
公交车	84.74	69.85	74.18	75.71

表 2 $k=100$ 的查准率结果

分类	查准率($k=100$)/%			
	本文算法	算法 I	算法 II	算法 III
飞机	62.55	41.87	48.51	43.40
大象	78.85	61.78	69.02	69.54
海滩	56.73	43.30	42.56	45.32
花朵	88.45	69.24	73.32	74.14
建筑	60.08	39.43	43.16	51.28
恐龙	80.44	59.56	63.74	68.36
非洲人	46.49	36.23	37.16	43.52
公交车	79.81	66.48	70.63	72.34

由表 1 和表 2 可知,本文算法在 $k=10$ 和 $k=100$ 时,查准率相对其他 3 种算法在相同类别图像中表现较好.此外,从表 1 和表 2 还可以看出,对于花朵、公交车、恐龙和大象等类别的检索,本文算法检索的查准率相对更高,其主要原因是这些类别图像的形状和纹理信息更加明显,特征容易被提取。

5 结论

本文提出了一种多样字典理论与多尺度距离度量相结合的彩色图像检索算法.实验结果表明,本文提出的检索算法的查准率与查全率均优于文献[4]、[5]和[6]中的方法,因此本文算法在彩色图像检索中具有较好的应用价值.本文方法在大量的图像检索任务中花费的时间成本相对较大,因此还需要进一步提高本文算法的效率,以完善本文方法。

参考文献:

[1] Avi Arampatzis, Konstantinos Zagoris. Dynamic two-stage image retrieval from large multimedia databases[J]. Information Processing and Management, 2013,49(6):274-285.

[2] 杨晓慧,姚雪彦.结合 LBP 和 Brushlet 的自适应图像检索[J].计算机工程,2013,39(2):233-236.

[3] 吴晓雨,何彦,杨磊.基于改进形状上下文特征的二值图像检索[J].光学精密工程,2015,23(1):303-310.

[4] Reddy A Hariprasad, Chandra N Subhash. Local oppugnant color space extrema patterns for content based natural and texture image retrieval[J]. International Journal of Electronics and Communications, 2015,69(1):290-298.

- [2] 陆小川,伊兵哲,平西建,等. 含噪文本图像的中英文文种识别研究[J]. 计算机工程与设计, 2007, 28(21): 5150-5152.
- [3] Busch A, Boles W W, Sridharan S. Texture for script identification[J]. IEEE Trans Pattern Anal Mach Intell, 2005, 27(11): 1720-1751.
- [4] 金璟璇,崔荣一,崔旭. 基于小波统计特征的行块级朝汉文种辨识[J]. 延边大学学报(自然科学版), 2013, 39(4): 277-280.
- [5] Rafael C Gonzalez, Richard E Woods. 阮秋琦译. 数字图像处理[M]. 2 版. 北京: 电子工业出版社, 2007.
- [6] 郭龙,平西建,周林,等. 基本图像特征用于文本图像文种识别[J]. 应用科学学报, 2011, 29(1): 56-60.
- [7] Griffin L D. The second order local-image-structure solid[J]. IEEE Trans Pattern Anal Mach Intell, 2007, 29(8): 1355-1366.
- [8] Griffin L D, Lillholm M. Symmetry sensitivities of derivative-of-gaussian filters[J]. IEEE Trans Pattern Anal Mach Intell, 2010, 32(6): 1072-1082.
- [9] Crosier M, Griffin L D. Texture classification with a dictionary of basic image features[C]//2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008). Anchorage, Alaska, USA: 2008.
- [10] Bello M G. Enhanced training algorithms, and integrated training/architecture selection for multilayer perceptron networks[J]. IEEE Trans Neural Networks, 1992, 3(6): 864-875.
- [11] 朴明姬. 自然图像中文字语种辨识方法的研究[D]. 延吉: 延边大学, 2014: 25-29.
- ~~~~~

(上接第 172 页)

- [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.
- [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: 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.
- [8] Rao L Koteswara, Rao D Venkata. Local quantized extrema patterns for content-based natural and texture image retrieval[J]. Human-Centric Computing and Information Sciences, 2015, 26(5): 1-24.
- [9] Cohen Andrew R, Vitanyi Paul M B. Normalized compression distance of multiples[J]. Physics, 2012, 41(13): 121-130.
- [10] Costa Paulo, Barroso João, Fernandes Hugo, et al. Using Peano-Hilbert space filling curves for fast bidimensional ensemble EMD realization[J]. EURASIP Journal on Advances in Signal Processing, 2012, 2012(1): 113-125.
- [11] Yang Peng, Xie Li, Liu Jilin. Simultaneous celestial positioning and orientation for the lunar rover[J]. Aerospace Science and Technology, 2014, 34(3): 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] 谭文明,李斌,张文聪. 基于中心对称局部二值模式的背景建模方法研究[J]. 中国科学技术大学学报, 2010, 40(11): 1112-1117.