

基于色彩量化及索引的图像检索^{* *}汪华章¹, 何小海^{1*}, 宰文姣², 王 炜¹

(1. 四川大学电子信息学院, 四川 成都, 610064; 2. 四川师范大学工程系, 四川 成都, 610066)

摘要: 提出了一种基于色彩量化及索引的图像检索新方法。结合人眼视觉感知特性, 首先将图像划分成 4×4 大小的非重叠子块, 通过块梯度的大小, 在亮度空间将图像子块划分为视觉均衡块和非均衡块。若为均衡块, 则用该块 RGB 空间的颜色均值作为其代表颜色值, 然后转换到 HSV 空间, 并量化成 32 个等级, 形成 32 维索引直方图(S_HIST); 若非均衡块, 根据保持颜色矩不变技术将该图像子块在 RGB 空间量化两种颜色, 然后转化到 HSV 空间, 并将每种颜色量化成 32 个等级, 形成 496 维的索引直方图(D_HIST)。最后, 综合索引特征, 进行图像检索。实验结果表明: 该算法是非常有效的。

关键词: 色彩量化; 块梯度; 均衡块; 保持颜色矩不变

中图分类号: TP391 **文献标识码:** A **文章编号:** 1005-0086(2008)02-0253-05

Image retrieval based on color quantization and indexing

WANG Hua-zhang¹, HE Xiao-hai^{1*}, ZAI Wen-jiao², WANG Wei¹

(1. Institute of Electronics and Information Engineering, Sichuan University, Chengdu 610064, China; 2. Department of Engineer, Sichuan Normal University, Chengdu 610066, China)

Abstract: A novel image retrieval algorithm based on color quantization and indexing is proposed, which is integrated with the human vision. Firstly, the image is divided into 4×4 non-overlapping blocks. Each block, based on its block gradient magnitude, is classified into visual uniform or visual non-uniform block. If the block belongs to a uniform block, the average of each color component for the pixels of it in RGB space is used as its representative color, then it is converted from RGB space into HSV space, and quantized into 32 levels, so an indexing histogram (S_HIST) of 32 dimensions is constructed. If the block belongs to a non-uniform block, then it is quantized into two representative colors by color moment-preserving. The two colors are transformed to HSV space, and each color is quantized into 32 levels, so an indexing histogram (D_HIST) of 496 dimensions is constructed. Finally, a novel image retrieval algorithm is formed by integrating all indexing features. The result of retrieval shows that the algorithm is effective.

Key words: color quantization; block gradient; uniform block; color moment-preserving

1 引 言

多媒体和网络技术的飞速发展以及大容量存储器和数字化设备的广泛应用, 使图像数据每天以惊人的速度增长, 如何有效组织、快速检索相关图像变得越来越重要。基于内容的图像检索(CBIR)技术受到青睐, 目前已有一些著名的原型系统, 如 QBIC、Virage、RetrievalWare、Photobook、MARS、VisualSEEK 和 Netra 等, 主要使用图像颜色、纹理、形状和空间关系以及以上多个特征的组合进行检索。

现阶段, 对图像理解尤其是高层语义理解不是很透彻的情况下, 决定图像检索性能主要存在两个问题: 1) 如何利用低层的视觉特征(如颜色、纹理和形状等)来表征图像内容; 2) 如何提取相似的特征以及如何进行有效的度量。针对这两个问题, 涌现了许多新的算法, 而且其检索的性能也在不断地提高。本文在深入理解图像内容表达方式的同时, 结合人眼的视觉特

性, 采用图像块编码的思想, 对图像进行量化, 并形成索引直方图进行检索。

2 算法介绍

2.1 特征的选择

颜色描述符是图像分析和检索中最重要的特征。由于颜色特征对平移、旋转和尺度等具有不变性的特点, 因此在许多检索算法中, 基于颜色的图像表示方法成为主要的图像索引术之一。颜色特征表示方法主要有颜色直方图(color histogram)^[1]、颜色集(color set)^[2]、颜色矩(color moments)^[3]、颜色聚合向量(color coherence vector)^[4]和颜色分布特征^[5]等, 其中直方图的应用尤为广泛。其计算简单、运算效率高, 同时归一化后对图像的尺寸变化不受影响, 但不足之处是无法反映颜色的空间分布信息。本文在充分利用颜色直方图的优点的同时, 对其不足之处进行改进, 提出一种新颖的基于色彩量化及索引

* 收稿日期: 2007-03-27 修订日期: 2007-05-28

* E-mail: hxh@scu.edu.cn

的图像检索算法,该算法不仅突出了目标的局部信息,同时兼顾了图像的全局信息,和同类算法相比,大大地提高了检索性能。

2.2 特征提取

基于色彩的量化,已经提出了大量的算法,主要分成2类:一类称为分块算法,即将图像划分为互不相连的区域,然后对其进行中值划分^[6]、均值划分^[7]或是方差划分^[8]等,该类算法应预先设定量化的颜色数目,若数目设定后,量化期间不能改变;另一类称为聚类算法,即对色彩进行聚类,如G-均值聚类算法^[9]、模糊G-均值算法^[10],这些算法有效地代表了图像的内容,同时通过合理的索引机制大大地节省了计算量。本文提出一种新颖的色彩量化方法,将图像分成互不重叠的 4×4 的图像子块^[11,12],由于每个子块内各像素间具有颜色近似的相关性,因此按照每个子块的统计特性进行局部量化,找出最能代表该小块的颜色值进行近似,以达到图像信息的最大保留。从视觉上讲,任何一幅图像可以看成是由“色彩平滑区”和“色彩非平滑区”构成,那么对于每一个 4×4 的子块,可以划分为“色彩均衡块”和“色彩非均衡块”。如何判断图像中子块属于哪种类型,又如何在最大限度不失真的情况下表征该图像子块内容,我们引入“块梯度”的概念。

假设大小为 $M \times N$ 的图像 IM ,可以将其划分为 $M \times N / (4 \times 4)$ 个图像子块,对其中任何一个图像子块, $I = \{I_{ij}; 0 \leq i < 4; 0 \leq j < 4\}$,如图1所示。

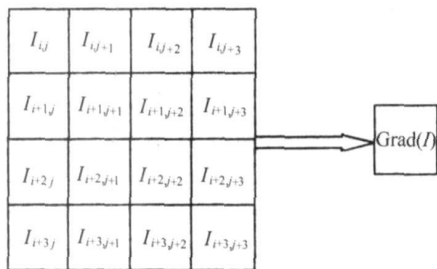


图1 任意 4×4 非重叠亮度块及块梯度

Fig. 1 Any 4×4 non-overlapping intensity block and block gradient

上述图像块可以表示成

$$I = \begin{bmatrix} I_{i,j} & I_{i,j+1} & I_{i,j+2} & I_{i,j+3} \\ I_{i+1,j} & I_{i+1,j+1} & I_{i+1,j+2} & I_{i+1,j+3} \\ I_{i+2,j} & I_{i+2,j+1} & I_{i+2,j+2} & I_{i+2,j+3} \\ I_{i+3,j} & I_{i+3,j+1} & I_{i+3,j+2} & I_{i+3,j+3} \end{bmatrix} \quad (1)$$

令

$$G_x = \frac{1}{8} \left[\sum_{n=0}^3 I_{i+n,j+2} + \sum_{n=0}^3 I_{i+n,j+3} - \sum_{n=0}^3 I_{i+n,j+1} - \sum_{n=0}^3 I_{i+n,j} \right]$$

$$G_y = \frac{1}{8} \left[\sum_{n=0}^3 I_{i+2,j+n} + \sum_{n=0}^3 I_{i+3,j+n} - \sum_{n=0}^3 I_{i+1,j+n} - \sum_{n=0}^3 I_{i,j+n} \right] \quad (2)$$

那么对整个图像块的梯度(即块梯度),其定义为 $\text{Grad}(I) = \sqrt{G_x^2 + G_y^2}$,它代表该块亮度的变化程度。其图像块亮度越均衡, $\text{Grad}(I)$ 值越小,说明该子块可能对应原图像中的背

景或目标中平滑变化的区域。换句话说, $\text{Grad}(I)$ 值越大反映在图像上视觉变化越明显,则该子块可能对应原图像中目标的边缘。如图2所示,假设对一幅完整的原始图像,将其中所有互不相连的 4×4 亮度块索引成1个像素点,并用它的“块梯度”值代表该点的像素值,则可以清楚地发现,其索引后的图像长、宽均为原图的 $1/4$,原图中目标及边缘的轮廓清楚地凸现在索引图中。对于原始图像中平滑的区域表现在梯度索引图中更平滑(为了便于观察,梯度索引图进行了取反运算),对视觉的影响非常小,说明“亮度非均衡块”其视觉贡献远远大于“亮度均衡块”。因此,根据人类视觉临界区分(JND)能力,可以选取一个固定的阈值将每一个图像子块划分为“亮度均衡块”和“亮度非均衡块”。研究表明^[13,14]: $\text{Grad}(I)$ 为13时为临界值,即 $\text{Grad}(I)$ 小于13图像的视觉变化不明显,若 $\text{Grad}(I)$ 大于或等于13说明图像子块极可能是目标轮廓或边缘块。“亮度均衡块”由于色彩变化不是很明显,因此可以利用该块中所有像素RGB 3分量的均值作为该块的代表色。而对于“亮度非均衡块”用同样的方法进行处理显然不合理,因此本文提出一种基于颜色矩保持不变^[15,16]的方法进行量化,即最大限度保持颜色内容不变的情况下,将每个“亮度非均衡块”量化成2种颜色,不妨假设量化后的2种代表色分别为 (R_1, G_1, B_1) , (R_2, G_2, B_2) ,假设 P_1, P_2 分别代表上面所述2种颜色在该块中所占的比例,那么通过

$$\begin{aligned} P_1 R_1 + P_2 R_2 &= m_r, P_1 G_1 + P_2 G_2 = m_g, \\ P_1 B_1 + P_2 B_2 &= m_b, P_1 R_1^2 + P_2 R_2^2 = m_{r2}, \\ P_1 G_1^2 + P_2 G_2^2 &= m_{g2}, P_1 B_1^2 + P_2 B_2^2 = m_{b2}, \\ P_1 R_1 G_1 B_1 + P_2 R_2 G_2 B_2 &= m_{rgb}, P_1 + P_2 = 1 \end{aligned} \quad (3)$$

可以求出2种代表色的值。其中: $m_r = \frac{1}{16} \sum_{i=0}^{15} r_i$, $m_g = \frac{1}{16} \sum_{i=0}^{15} g_i$, $m_b = \frac{1}{16} \sum_{i=0}^{15} b_i$,分别为该“亮度非均衡块”对应原图中R、G、B 3分量所有像素的均值; $m_{r2} = \frac{1}{16} \sum_{i=0}^{15} r_i^2$, $m_{g2} = \frac{1}{16} \sum_{i=0}^{15} g_i^2$, $m_{b2} = \frac{1}{16} \sum_{i=0}^{15} b_i^2$,则相应为3分量方差的均值; $m_{rgb} = \frac{1}{16} \sum_{i=0}^{15} r_i g_i b_i$,为R、G、B 3分量间的相关性。

2.3 HSV空间量化及索引直方图的建立

在图像检索过程中通常采用欧氏距离来度量2幅图像间的差异,最常用的方法是在RGB空间直接提取图像特征进行度量,但RGB颜色空间与人眼的视觉感知相差很大,同样差异的2种颜色视觉感知可能完全不一样,因此本文采用HSV颜色模型,因为该模型中H分量对彩色的描述能力来说与人的视觉最接近,区分力也比较强。同时,H、S、V分量相对独立,对视觉贡献不同,很容易通过设定不同权值将三维分量转换到一维上进行度量。实验表明:颜色直方图的维数越多,图像信息越丰富,对全面描述颜色内容比较有利;但维数越多,用来存储特征矢量的空间越大,同时计算量会大大增加,因此,综合多方面的因素,在降低直方图的维数并保留足够的颜色信息前提下,对HSV空间进行非均匀量化,量化得到32种代表颜色,由RGB空间到HSV空间的转换,转换后的h、s、v值($h \in [0,$

360), $s \in ([0, 1], v \in [0, 1])$, 则其量化及编码规则如下: 1) 若 $V < 0.1$, 则认为该颜色为黑色, $\text{Code} = 0$; 2) 若 $S < 0.1$ 且 $V > 0.1$, 则将颜色按亮度 V 划分为 3 个灰度, 分别为 $(0.1, 0.4]$ 、

$(0.4, 0.7]$ 和 $(0.7, 1]$ 3 个区间, 分别编码为 $\text{code} = 1, 2, 3$; 3) 将其它的颜色按彩色进行量化, 将其划分为 28 种颜色, $\text{code} = 4 \times h' + 2 \times s' + v' + 4$, 其中

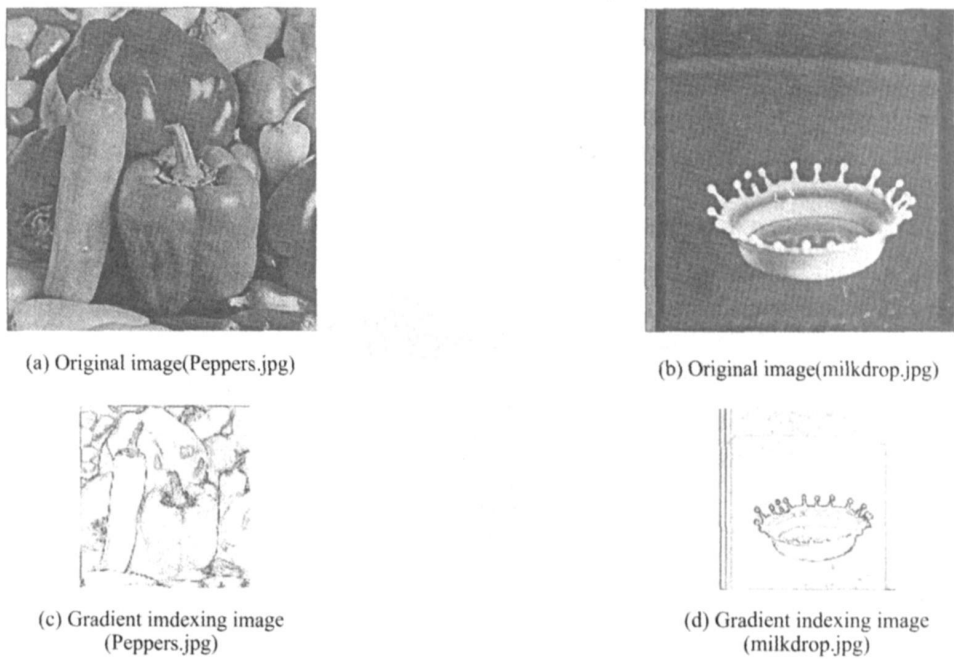


图 2 原始图像 (a)、(b) 及相应的梯度索引图 (c)、(d)
Fig. 2 Original images (a), (b) and corresponding gradient indexing image (c), (d)

$$h' = \begin{cases} 0 & h \in (0, 20] \cup (330, 360] \\ 1 & h \in (20, 45] \\ 2 & h \in (45, 75] \\ 3 & h \in (75, 155] \\ 4 & h \in (155, 210] \\ 5 & h \in (210, 270] \\ 6 & h \in (270, 330] \end{cases}$$
$$s' = \begin{cases} 0 & s \in (0.1, 0.5] \\ 1 & s \in (0.5, 1.0] \end{cases} \quad v' = \begin{cases} 0 & v \in (0.1, 0.5] \\ 1 & v \in (0.5, 1.0] \end{cases} \quad (4)$$

显然, 经过 RGB 空间的转换以及 HSV 空间的量化, 每一个“单色均衡块”则被量化成 32 种颜色, 因此, 作为所有均衡块的索引, 构建量化等级数为 32 的直方图, 然后对其进行标准化, 即得到“单色均衡块直方图”, 记为 S_HIST 。用同样的方法, 将“双色非均衡块”进行量化, 2 种颜色分别被量化成 32 种颜色。那么对于整个双色块, 则直方图的维数应该为 $C_{32}^2 = 32 \times 31/2 \times 1 = 496$ 。同理, 作为所有非均衡块的索引, 构建量化等级数为 496 的直方图, 同样对其进行标准化, 即得到“双色非均衡块直方图”, 记为 D_HIST 。

3 图像内容相似性计算

表现图像内容的特征被提取出来以后, 形成特征向量。则图像间的相似性可以转变为特征向量是否相似来计算。特征向量间的距离越大, 说明图像间的差别越大。针对不同的向量模型, 可以采用不同的相似度匹配算法, 可以选择欧拉距离、二次

式距离和马氏距离等。对于同维度的各分量, 其重要程度若相同, 则可以采用欧拉距离 (L1 或 L2 距离) 进行度量, 其中 L1 距离可以表示为

$$D1 = \sum_{i=1}^N |A_i - B_i|$$

式中, N 为特征向量的维数。类似地, L2 距离可以表示为

$$D2 = \sum_{i=1}^N |A_i - B_i|^2$$

本文采用 L2 距离进行度量。设 q 代表查询图像, 相应地 r 代表图像库中被检索的图像。很显然, 单色块和双色块在图像检索过程中其重要性相同, 因此在计算相似性的过程中, 可以通过设置不同的权值来突出其重要性, 假设 ω_1 代表单色块的权重, ω_2 代表双色块的权重, 那么图像间的相似性可以表示为

$$\text{DISTANCE}(q, r) = \omega_1 \sqrt{\sum_{i=1}^{32} |S_HIST_q(i) - S_HIST_r(i)|^2} + \omega_2 \sqrt{\sum_{i=1}^{496} |D_HIST_q(i) - D_HIST_r(i)|^2} \quad (5)$$

其中, $\omega_1 + \omega_2 = 1$ 。通过设置不同权值, ω_1 、 ω_2 初始值分别设为 0.4、0.6, 在实验过程中可以根据图像的特点调节权值大小。

4 评价准则及测试方案

4.1 评价准则

采用“基于例子查询”的方法, 给出例子图像, 找出图像数据库中与之相似的所有图像。算法采用查全率 (recall) 和查准率 (precision)^[17] 作为评价标准, 假设 R 代表正确地检索到相

关图像的数目, N 代表检索结果返回的图像总数, Q 代表图像库中与查询图像相关的目标图像总数。其中, 查全率定义为检索队列中检索到的相关图像数与数据库中所有与例子图像相关的图像数目; 查准率定义为检索结果队列中检索的目标图像数与队列中的图像数目之比。对不同图像检索算法, 在相同的查全率条件下, 查准率越高, 表明该算法的检索效果越好。上述指标可以表示为

$$\text{precision} = R/N, \text{recall} = R/Q \quad (6)$$

然后作出查全率和查准率曲线图, 则可以很直观的判断检索性能的好坏。

4.2 测试方案

以 1000 幅图像作为测试库, 图像库共分 10 组, 均来自于 <http://wang.ist.psu.edu/jwang/test1.tar>, 其中包含花、马、恐龙、车、海滩、食物、建筑和大象等类别。检索过程中, 相似图像的判别原则主要有两点: 1) 基本上认为同一组的图像属于相似的图像; 2) 主观判断 2 幅图像的相似性, 包括局部相似和整体相似。检索过程中, 对同一类图像取其中任何一幅作为例子图像进行检索, 为了验证算法的有效性和鲁棒性, 通常对同类图像进行多次检索求其平均查全率, 整个算法的指标是对各组图像进行检索后的综合平均指标。

5 实验结果及分析

为了直观地表示检索性能的变化情况以及同类算法相比的优缺点, 图 3 显示了该算法测试图像库中不同组图像获得的查全率和查准率曲线, 图 4 显示了不同算法所获得查全率与查准率曲线。

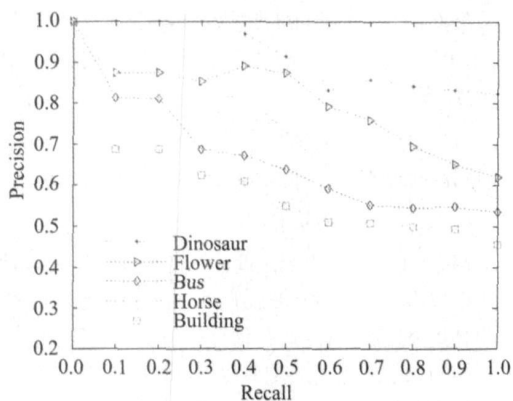


图 3 5 类不同图像查全率查准率曲线图

Fig.3 Recall and precision graphs of five different classes images

图 3 可以清楚的发现: 对于目标和背景比较单一的图像, 该算法其检索率较高, 例如“恐龙”及“花朵”, 对于背景比较复杂, 目标比较分散或目标和背景难以区分的图像其性能明显降低, 例如建筑。值得注意的是: “马”从理论上分析其效果应该比较好, 其目标和背景比较明显, 但是该图像库中有类似的“大象”, 因此导致其效果有所下降。图 4 表明, 整体上该算法的效果要比同类算法的检索效果好。与块梯度算法及全局直方图算法相比, 该算法不仅考虑到目标的边缘信息, 同时考虑到了

图像的全局信息, 并将块内和块间的相互关系有效地结合起来, 充分利用了图像压缩编码的思想, 使算法检索性能得到了较大的提高。

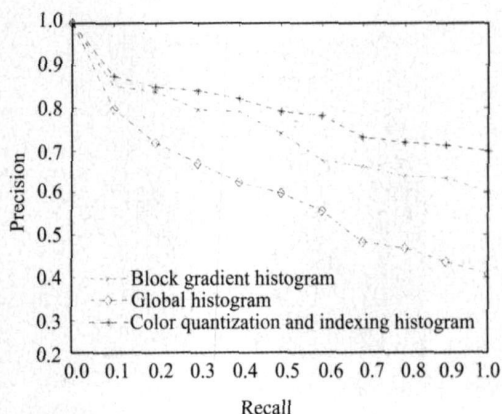


图 4 几种算法性能比较

Fig.4 Performance contrast of several algorithms

6 结论

结合人眼视觉感知特性, 对图像进行分块量化得到视觉均衡块和视觉非均衡块, 并根据不同块对视觉贡献的大小分配权重, 不仅考虑图像的全局信息, 同时注重图像的局部信息, 和同类算法相比, 性能大大提高, 尤其适合目标和背景比较单一的图像的检索。

参考文献:

- [1] Swain M J, Ballard D H. Color indexing[J]. International Journal of Computer Vision, 1991, 7(1): 11-32.
- [2] Misra P A, Rodriguez J J. Efficient indexing of multi-color sets for content-based image retrieval [A]. 4th IEEE Southwest Symposium 2000, 116-120.
- [3] Pass G, Zabih R. Histogram refinement for content-based image retrieval [A]. IEEE Workshop on Applications of Computer Vision [C]. 1996, 96-102.
- [4] SUN Jin-ding, WU Xie-sheng. Image retrieval based on color distribution features [J]. Journal of Optoelectronics · Laser (光电子·激光), 2006, 17(8): 1009-1013. (in Chinese)
- [5] Heckbert P. Color image quantization for frame buffer display [J]. Comput. Graphics, 1982, 16: 297-307.
- [6] Wan S J, Prusinkiewicz P, Wong S K M. Variance-based color image quantization for frame buffer display [J]. Color Res, 1990, 15(1): 52-58.
- [7] Jain A K, Murty M N, Flynn P J. Data clustering: a review [J]. ACM Comput. Surv, 1999, 31(3): 264-323.
- [8] Karayianis N B, Bezdek J C. An integrated approach to fuzzy learning vector quantization and fuzzy c-means clustering [J]. IEEE Trans on Fuzzy Systems, 1997, 5(4): 622-628.
- [9] Chan K W, Chan K L. An adaptive pattern based image coding technique using visual pattern and texture pattern [A]. 1994 International Symposium on Speech, Image Processing and Neural Networks [C]. 1994.

- [10] CHa-Hung Yeh, Chung J Kuo. Content-based image retrieval through compressed indices based on vector quantized images[J]. Optical Engineering 2006, 45(1): 017001.
- [11] ZHAO Shan, SUN Jun ding, ZHOU Li-hua. Image texture extraction and retrieval based on block truncation coding[J]. Journal of Optoelectronics • Laser (光电子 • 激光), 2006, 17(8): 1014-1017. (in Chinese)
- [12] Chen D, Bpvol A C. Visual Pattern Image Coding[J]. IEEE Trans Communi 1990, 38(12): 2137-2146.
- [13] Nezamabadi-pour H, Kabir E. Image Retrieval Using Histograms of uni-color and Bi-color Blocks and directional changes in intensity gradient[J]. Pattern Recognition Letters, 2004, PRL(25): 1547-1557.
- [14] Chen L H, Tsai W H. Moment-preserving curve detection[J]. IEEE trans Syst Man Cybern, 1988, 18(1): 148-158.
- [15] Yang Chen Kuei, Lin Ja-Chen, Wen Hsiang, Tsai. Color Image Compression by Moment-Preserving and Block Truncation Coding Techniques[J]. IEEE Transactions on Communications, 1997, 45(12): 1513-1516.
- [16] Julia Vogel, Bernt Schiele. Performance evaluation and optimization for content-based image retrieval[J]. Pattern Recognition. 2006, 897-909.

作者简介:

汪华章 (1976-), 男, 博士生, 湖北人, 研究方向为多媒体通信、图像检索和模式识别。