

综合 PHOG 和 LWT 的图像检索算法

弓云峰, 崔得龙

(广东石油化工学院, 茂名 525000)

摘要: **目的** 研究物体的形状特征在图像描述及图像检索中的区分度和检索性能。**方法** 设计一种综合 PHOG 形状和提升小波变换的图像检索算法。算法首先对原始图像进行极坐标系方向归一化, 提取图像旋转不变特征; 其次提取分层图像的 PHOG 形状特征; 然后提取分层图像低频变换系数均值和方差作为提升小波变换特征; 最后将各种特征进行融合并用于图像检索, 并定义距离衡量公式。**结果** 通过文中设计算法提取的图像形状特征可使各标准测试图像间距离均值为 0.2352。**结论** 在 Corel 图像集上的检索实验结果优于 RIM 算法和 FWTH 算法, 表明文中算法图像检索领域具有一定的应用前景。

关键词: PHOG; 形状特征; 提升小波; 图像检索

中图分类号: TP391 **文献标识码:** A **文章编号:** 1001-3563(2017)15-0202-05

An Image Retrieval Method Based on PHOG and LWT

GONG Yun-feng, CUI De-long

(Guangdong University of Petrochemical Technology, Maoming 525000, China)

ABSTRACT: The work aims to study discriminative power and retrieval performance of shape feature of the objects in image description and image retrieval. An image retrieval method synthesizing PHOG shape and lifting-based wavelet transform was designed. Firstly, the orientation normalized polar coordinate system of the original image was achieved to extract the image rotational invariant features. Secondly, PHOG shape feature of layered image was extracted. Then, the mean value and variance of low-frequency transform coefficient of the layered image were extracted as the transform feature of the lifting-based wavelet. Finally, all kinds of features were synthesized and used for image retrieval, and the distance measurement formula was defined. The image shape features extracted by the designed algorithm proposed herein could enable the mean value of distances between all standard test images to be 0.2352. The retrieval experiment results on Corel image set are superior to RIM algorithm and FWTH algorithm, which indicates that the proposed algorithm has certain application prospects in the field of image retrieval.

KEY WORDS: PHOG; shape feature; lifting-based wavelet; image retrieval

图像检索作为人们获取信息的重要方式之一, 从 20 世纪 70 年代发展至今经历了基于文本、基于内容、基于相关反馈和基于压缩域等 4 个主要发展阶段。图像的底层特征作为图像自身重要特征在图像检索技术发展过程中起着重要的作用^[1-4], 如颜色、纹理、形状以及局部描述子等。常用的颜色特征有颜色直方图、颜色矩、颜色相关图、颜色聚合向量等; 纹理描述有灰度共生矩阵和小波变换等; 局部描述子有

Harris、形殊点、SIFT 等。

形状特征在人类视觉系统的目标感知和识别中起着非常重要的作用, 因此, 检索对象形状的提取和描述一直是基于内容图像检索研究中的重点和难点。文献中形状描述方法可主要分为基于区域的形状描述和基于轮廓的形状描述两大类方法。基于区域的形状描述对整幅图像进行特征提取, 主要包括各种不变矩, 如 Hu 矩、Zernike 矩、Fourier-mellin 矩等, 该

收稿日期: 2016-09-10

基金项目: 国家自然科学基金(61473331); 广东省云机器人工程中心项目(660010); 茂名市石油化工自动化工程技术研究中心项目(517013)

作者简介: 弓云峰(1977—), 男, 硕士, 广东石油化工学院实验师, 主要研究方向为图像处理、智能信息处理。

类描述方法一般对噪声具有较强的鲁棒性,但存在计算量大和计算结果精确度较低等不足。基于轮廓的形状描述对图像轮廓进行特征提取,主要包括 Fourier 法、曲率尺度空间法和 Freeman 码法等。该类描述方法在图像总体形状特征-轮廓的基础上增加了对形状细节的描述,更加贴近人类视觉系统,但检索结果依赖于图像轮廓提取的质量,而图像轮廓的提取通常对噪声较敏感。

图像形状梯度方向直方图金字塔(PHOG)是近年来图像视觉描述领域取得的重要研究成果,已在图像分类^[5-6]、图像检索^[7-8]、图像识别^[9-10]等领域得到了广泛的应用。Li 等^[11]设计了一种基于颜色分割和形状匹配的鲁棒交通信号检索算法。算法中应用 PHOG 进行感兴趣区域形状提取和分类。Atoany 等^[12]设计了一种基于内容的图像检索系统,系统使用颜色相关直方图抽取颜色特征,使用 PHOG 提取形状特征。Asha 等^[13]设计了基于内容安全的图像检索系统,系统采用 PHOG 特征描述原始图像,并采用双密钥 RSA 算法进行加密。袁杰等^[14]设计了一种综合 PHOG 形状和小波金字塔能量分布特征的图像检索方法。算法基于 SVM 分类器框架,动态结合图像 PHOG 特征和小波金字塔能量分布特征,并在 Corel 图像集上验证了算法的有效性。

PHOG 作为一种多尺度形状特征描述方法,虽然不仅可表征图像的整体形状,而且还表征图像的局部形状以及它们的空间关系,但仍然存在检索目标准确分割困难、形状间相似性度量不便、以及依赖于图像轮廓提取质量等不足。文中针对 PHOG 形状特征在图像检索应用中存在的不足,提出一种综合 PHOG 形状特征和提升小波变换的图像检索算法。算法首先对图像的灰度图进行方向归一化,然后分别提取 PHOG 形状特征和提升小波变换特征,最后进行特征的动态融合进行图像检索。

1 旋转不变特性提取

文中采用基于极坐标方向能量归一化^[15]的方法获取原始图像的旋转不变特征,具体步骤如下所述。

1) 计算原始图像 $I(M \times N)$ 的质心坐标 $(\bar{x}, \bar{y})$, 并将其作为图像的原点。若原始图像大小为 $M \times N$, 则其质心坐标 $(\bar{x}, \bar{y})$ 定义为:

$$\begin{cases} \bar{x} = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N x f(x, y) \\ \bar{y} = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N y f(x, y) \end{cases} \quad (1)$$

2) 坐标系转换。将式(1)从平面直角坐标系转换到极坐标系,可得原始图像的极坐标表示为 $f(r, \theta)$, 各变量的说明见式(2)。

$$\begin{cases} r = \sqrt{(x - \bar{x})^2 + (y - \bar{y})^2} \\ \theta = \arctan \frac{y - \bar{y}}{x - \bar{x}} \\ f(r \cos \theta, r \sin \theta) = f(r, \theta) \end{cases} \quad (2)$$

3) 确定方向能量分布序列。原始图像极坐标系下任一方向 $\theta_j (0 \leq \theta_j \leq 2\pi)$ 的能量分布函数 $\rho(\theta_j)$ 定义为:

$$\rho(\theta_j) = \frac{1}{n_j} \sum_{i=0}^{N-1} f(r_i, \theta_j), \quad (0 \leq \theta_j < M) \quad (3)$$

式中: n_j 为方向 θ_j 上所占区域内像素的个数; N 为径向变量 r 的量化级数; M 为方向变量 θ 的量化级数。则原始图像 $I(M \times N)$ 的方向能量分布序列可表示为: $\rho(\theta) = \{\rho(\theta_0), \rho(\theta_1), \dots, \rho(\theta_{M-1})\}$ 。

4) 提取旋转不变特征。利用循环左移操作提取旋转不变特征。循环左移算子规则为: 将原始序列的最大值元素放在旋转不变特征中的首位。若最大值元素不唯一时,则比较序列中最大值右边第 1 个值,选择较大值的序列为循环左移后的序列。若右边第 1 个值相等,则比较第 2 个值,依次类推,直至选择出足够数量的旋转不变特征。

2 PHOG 形状特征提取

PHOG^[15]作为一种形状的空间描述,在通用目标分类等领域得到较好应用。通常 PHOG 特征提取主要步骤如下所述。

1) 检测图像边缘。利用边界检测算子 Canny 提取图像的边缘。

2) 图像的层次化表征。图像分层示例见图 1。其中图 1a 为原始图像;图 1b 为第 1 层,即原始图像经 Canny 算子提取的轮廓;图 1c 为第 2 层,即图 1b 的均分;图 1d 为第 3 层,即图 1c 的进一步均分。

3) 提取梯度直方图特征。对每层次的各子区域,分别提取梯度直方图特征。

4) 构造 PHOG 特征。级联各子区域的梯度直方图特征,得到原始图像的 PHOG 特征。

从图 1 中可以看出,随着分割层数的增加,梯度方向直方图对图像形状的刻画也越来越局部化和精细化。提取出的 PHOG 特征还要进行归一化。第 1 层梯度方向直方图合并时乘以 $\frac{1}{L^4}$ 就能得到该层的归一化值。文中采用式(4)所示的 χ^2 距离计算 2 个 PHOG 特征间的距离。

$$distance(P_i, P_j) = \frac{1}{2} \sum_{k,l} \frac{1}{L^4} |p_{ik} - p_{jk}| \quad (4)$$

3 LWT 特征提取

LWT 即二代小波变换是基于 Sweldens 设计的提升

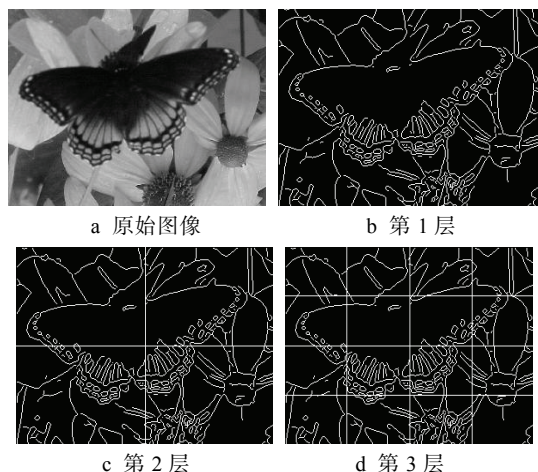


图1 3层 PHOG 特征提取示例

Fig.1 Example of 3-layer PHOG feature extraction

方案,该方案主要包括分裂、预测和更新等3个步骤^[16-17],各步骤主要实现的功能如下所述。

1) 分裂。将原始信号 c_{j-1} 分割成2个不相交的样本子集。即:

$$(o_{j-1}, e_{j-1}) = \text{Split}(c_{j-1}) \quad (5)$$

分解后的信息通常分别称为原始信号的奇偶子集。

2) 预测。利用相邻信号间的相似性进行子集预测,以 P 表示预测算子,则用偶子集来预测奇子集可表示为:

$$d_j = o_{j-1} - P(e_{j-1}) \quad (6)$$

奇偶信号间的差值 d_j 包含原始信号 c_{j-1} 的细节信息,体现出原始信号 c_{j-1} 的瞬变特性。

3) 更新。更新细节信息 d_j ,以 U 表示更新算子,并与偶子集进行求和以得到低频信息:

$$c_{j-1} = e_{j-1} + U(d_j) \quad (7)$$

二代提升小波的逆变换实现起来非常简单,只需将前向计算公式中的加号变为减号,二代提升小波的正反变换过程见图2。

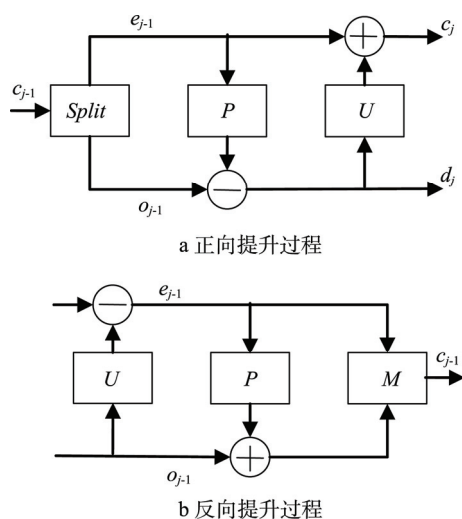


图2 二代小波的正反提升过程

Fig.2 Forward and backward lifting process of second generation wavelet

为便于稳健特征的构造,文中采用低频逼近小波系数均值和方差作为统计特征,并采用归一化汉明距离(NHD)进行度量,定义如下:

$$NHD(h_1, h_2) = \frac{1}{L} \sum_{k=1}^L |h_1(k) - h_2(k)| \quad (8)$$

式中: h_1 和 h_2 分别为2幅图像提取的特征序列; L 为序列的长度;由归一化汉明距离的定义式可知,若2幅图像越相似,则NHD值越接近1,否则越接近0。

4 图像检索算法

文中设计的图像特征提取和图像检索步骤如下所述。

1) 将原始图像进行极坐标方向能量归一化,确定图像的主轴。

2) 对规范化之后的图像分别提取 PHOG 特征和提升小波特征。

3) 综合 PHOG 特征和提升小波特征,根据式(8)进行距离衡量。

$$D(I_i, I_j) = \sum_k \omega_k D_k(I_i, I_j), (k=1,2) \quad (9)$$

式中: $D(I_i, I_j)$ 为图像间总的距离; $D_k(I_i, I_j)$ 为分别用 PHOG 特征和提升小波特征测算出的图像 I_i 和 I_j 的距离,分别见式(4)和式(8); ω_k 为相应的权值,所有权值之和为1。文中提取的 PHOG 特征和提升小波特征在进行距离度量时同等重要,因此 ω_1 和 ω_2 设置为0.5。

4) 返回与原始图像距离最小的若干幅图像作为图像检索结果。文中算法选取距离最近的15幅图像作为检索结果。

5 实验分析

5.1 区分度分析

选用 Lena, Baboon, Plane, Peppers, Boats, Lake, Lax, Bridge 等标准测试图像进行图像检索区分度分析,部分测试图像见图3。

文中算法下 PHOG 形状特征提取中角度范围为 $[0, 180]$, 角度区间个数为60, 级数为3级,从而每幅图像的 PHOG 特征有1260维。提升小波变换中分块图像大小为 32×32 , 变换级数为2级;综合检索中 ω_k 取值各为0.5。

在标准测试图像下进行的区分度分析结果见表1。由表1可知,各标准测试图像间最小距离为0.1359, 最大距离为0.4063, 均值为0.2352。以0.5作为阈值则可完全区分各标准测试图像,实验结果表明文中算法对内容不同的标准测试图像具有很好的区分度。

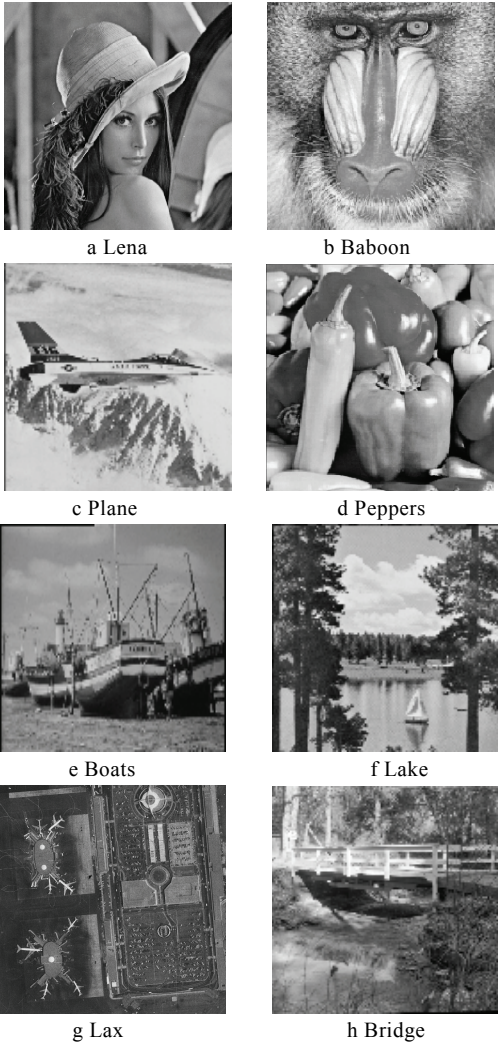


图 3 部分标准测试图像
Fig.3 Part of benchmark test images

5.2 检索性能分析

选取 Corle 的一个子集作为测试图像，该集合包括飞机、雕塑、手机、轮椅、犀牛、人脸、游艇、纸币、手表、风笛等共 10 类图像，每类各 30 幅图像，

共 300 幅，图像大小为 256×384 或 384×256 像素。文中采用 A 指标对检索效果进行衡量，其定义如下：若某次检索获得的 n 幅图像中，其中 m 幅与被检索图像相关，在检索结果中的排序为， $N_{IR_1}, N_{IR_2} \dots N_{IR_m}$ ，剩余 $z=n-m$ 幅与被检索图像无关，在检索结果中的排序为 $N_{IN_1}, N_{IN_2} \dots N_{IN_z}$ ，则

$$A = \frac{1}{2} + \frac{1}{n(n+1)} \left[\sum_{i=1}^m (n+1 - N_{IR_i}) - \sum_{i=1}^z (n+1 - N_{IN_i}) \right] \quad (10)$$

由式(10)可知，当检索得到的所有结果图像 n 中与被检索图像相关的图像数量越多时， A 的值越接近 1，否则越接近 0。 A 值不仅考虑了检索结果的准确性，而且考虑了检索结果图像的排序，因此该指标能够较全面地衡量图像检索算法的综合检索能力。文中算法在不同类别上的平均 A 值见图 4，为了验证文中算法的有效性，图 4 中同时列出了 PHOG 算法和 LWT 算法下的 A 值。

从图 4 的实验结果可知，文中算法结合了 PHOG 提取的图像形状特征和 LWT 提取的图像细节特征，在测试子集下的 A 值均优于对比算法，同时，对形状特征区别较大的测试子集，例如人脸、游艇、手表等，优于形状特征区别较小的测试子集，例如手机、犀牛等。

5.3 与类似算法的比较

实验中随机在图像库中选取 10 类图像，且在每类图像中抽取 4 幅图像共组成 100 次检索，然后结合 40 次检索，计算在不同检索条件下的算法的平均精确度，实验结果见图 5。图 5 中横坐标为检索率，定义为检索结果队列中检索到的目标图像与数据库中全部的目标图像数量的比值，纵坐标为精确度，定义为检索结果队列中检索到的目标图像与检索结果队列中所有图像数量的比值。为了与类似算法比较，图 5 同时列出了 RIM 算法^[18]、FWHT 算法^[19]和 DCCD 算法^[20]的平均精确度。

表 1 部分标准测试图像间的区分度
Tab.1 Discrimination among parts of benchmark test images

图像	Lena	Baboon	Plane	Peppers	Boats	Lake	Lax	Bridge
Lena	0	0.2067	0.2282	0.1637	0.1947	0.2281	0.2400	0.3026
Baboon	0.2067	0	0.1696	0.1492	0.2389	0.2389	0.2200	0.2586
Plane	0.2282	0.1696	0	0.1923	0.1359	0.2295	0.2399	0.1776
Peppers	0.1637	0.1492	0.1923	0	0.2107	0.1888	0.2353	0.3325
Boats	0.1947	0.2389	0.1359	0.2107	0	0.2759	0.2807	0.3870
Lake	0.2281	0.2389	0.2295	0.1888	0.2759	0	0.2307	0.2237
Lax	0.2400	0.2200	0.2399	0.2353	0.2807	0.2307	0	0.4063
Bridge	0.3026	0.2586	0.1776	0.3325	0.3870	0.2237	0.4063	0

由于文中算法充分利用了图像极坐标方向能量归一化以及综合 PHOG 和提升小波进行特征提取，因此文中算法的检索性能有一定程度的提高，优于 RIM 算法和 FWTH 算法，与 DCCD 算法相比，性能

也有一定的提升,尤其是在检所率大于 0.3 之后。

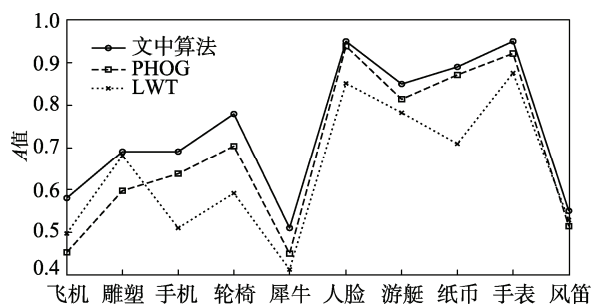


图 4 文中算法在不同类别上的平均 A 值

Fig.4 Average A value of different algorithms among various categories

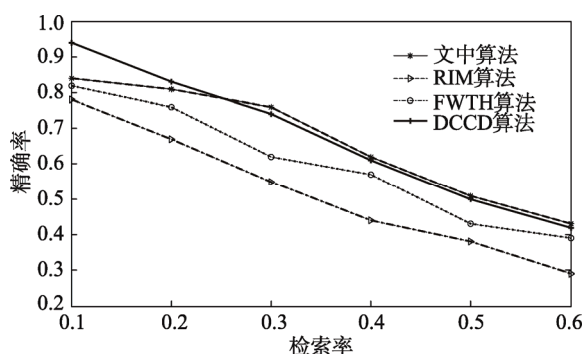


图 5 各种算法的检索性能分析

Fig.5 Analysis of retrieval performance of various algorithms

6 结语

物体的形状特征在图像描述及图像检索算法中占据着非常重要的地位。文中使用图像极坐标方向能量归一化,得到了原始图像的旋转不变特性描述,同时结合 PHOG 形状描述和提升小波变换进行特征提取,实验结果证明了文中算法的有效性。

参考文献:

- [1] KRYSTIAN M, CORDELIA S. Scale & Affine Invariant Interest Point Detectors[J]. International Journal of Computer Vision, 2004, 60(1): 63—86.
- [2] MOHSEN Z, DORAISAMY S H, HALIN A A. Texture Classification and Discrimination for Region-based Image Retrieval[J]. Journal of Visual Communication and Image Representation, 2015, 26: 305—316.
- [3] 周燕, 曾凡智. 基于二维压缩感知和分层特征的图像检索算法[J]. 电子学报, 2016, 44(2): 453—460. ZHOU Yan, ZENG Fan-zhi. An Image Retrieval Algorithm Based on Two-Dimensional Compressive Sensing and Hierarchical Feature[J]. Acta Electronica Sinica, 2016, 44(2): 453—460.
- [4] 于来行, 冯林, 张晶, 等. 自适应融合目标和背景的图像特征提取方法[J]. 计算机辅助设计与图形学学报, 2016, 28(8): 1250—1259. YU Lai-hang, FENG Lin, ZHANG Jing, et al. An Image Feature Extraction Method Based on Adaptive Fusion of Object and Background[J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(8): 1250—1259.
- [5] ATREYEE S, SUGATA B, LIU Cheng-jun. New Color GPHOG Descriptors for Object and Scene Image Classification[J]. Machine Vision and Applications, 2014, 25(2): 361—375.
- [6] BANERJI S, SINHA A, LIU Cheng-jun. New Image Descriptors Based on Color, Texture, Shape, and Wavelets for Object and Scene Image Classification[J]. Neurocomputing, 2013, 117: 173—185.
- [7] KARIM R, AZIZ M M A, SHATABDA S, et al. CoMOGrad and PHOG: from Computer Vision to Fast and Accurate Protein Tertiary Structure Retrieval[J]. Scientific Reports, 2015, 5: 13275.
- [8] BIASOTTI S, CERRI A, GIORGI D M. Spagnuolo, PHOG: Photometric and Geometric Functions for Textured Shape Retrieval[J]. Computer Graphics Forum, 2013, 32(5): 13—22.
- [9] SAVCHENKO A V, MILOV V R, BELOVA N S. Sequential Hierarchical Image Recognition Based on the Pyramid Histograms of Oriented Gradients with Small Samples[J]. Proceeding of 4th International Conference on Analysis of Images, Social Networks, and Texts (AIST), 2015, 25(4): 14—23.
- [10] THUBSAENG W, KAWEWONG A, PATANUKHOM K. Vehicle Logo Detection Using Convolutional Neural Network and Pyramid of Histogram of Oriented Gradients[C]// International Joint Conference on Computer Science and Software Engineering IEEE, 2014: 34—39.
- [11] LI Hao-jie, SUN Fu-ming, LIU Li-juan, et al. A Novel Traffic Sign Detection Method Via Color Segmentation and Robust Shape Matching[J]. Neurocomputing, 2015, 169: 77—88.
- [12] FIERRO-RADILLA A, PEREZ-DANIEL K, NAKANO-MIYATAKEA M, et al. An Effective Visual Descriptor Based on Color and Shape Features for Image Retrieval[C]// Mexican International Conference on Artificial Intelligence, Springer International Publishing, 2014: 336—348.
- [13] ASHA M M, RANJANI J J. Secure Image Retrieval Using Pyramid Histogram of Oriented Gradient Descriptor[C]// Proceeding of International Conference on Advanced Computing and Communication Systems (ICACCS), 2013: 19—21.
- [14] 袁杰, 魏宝刚, 王李冬. 一种综合 PHOG 形状和小波金字塔能量分布特征的图像检索方法[J]. 电子学报, 2011, 39(9): 2114—2119. YUAN Jie, WEI Bao-gang, WANG Li-dong. An Image Retrieval Method Synthesizing PHOG Shape Description and Wavelet Pyramid Energy Distribution[J]. Acta Electronica Sinica, 2011, 39(9): 2114—2119.
- [15] 潘泓, 夏良正. 具有平移、尺度和旋转不变性的小波变换[J]. 信号处理, 2004, 20(4): 147—151. PAN Hong, XIA Liang-zheng. A Wavelet Transform with Shift, Scale and Rotation Invariance[J]. Signal Processing, 2004, 20(4): 147—151.
- [16] BOSCH A, ZISSERMAN A, MUNOZ X. Representing Shape with a Spatial Pyramid Kernel[C]// Proceeding of the 6th ACM International Conference on Image and Video Retrieval Amsterdam, Netherlands, 2007: 401—408.
- [17] SWEDENS W. The Lifting Scheme: A Construction of Second Generation Wavelets[J]. SIAM Journal of Mathematical Analysis, 1998, 29(2): 511—546.
- [18] KOUROSH J K, HAMID S Z. Rotation-invariant Multiresolution Texture Analysis Using Radon and Wavelet Transform[J]. IEEE Trans on Image Process, 2005, 14(6): 783—794.
- [19] MANDAL M K, ABOULNASR T, PANCHANATHAN S. Fast Wavelet Histogram Techniques for Image Indexing[J]. Computer Vision and Image Understanding, 1999, 75(1/2): 99—110.
- [20] RADILLA A F, DANIEL K P, MIYATAKEA M N, et al. An Effective Visual Descriptor Based on Color and Shape Features for Image Retrieval[J]. Lecture Notes in Artificial Intelligence, 2014, 8856: 336—348.