



DOI: 10.12086/oe.2021.210310

融合 HSV 与方向梯度特征的多尺度图像检索

江曼¹, 张皓翔¹, 程德强^{1,2*},
郭林¹, 寇旗旗³, 赵雷¹

¹中国矿业大学信息与控制工程学院, 江苏 徐州 221116;

²中国矿业大学地下空间智能控制教育部工程研究中心, 江苏 徐州 221116;

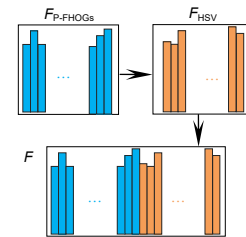
³中国矿业大学计算机科学与技术学院, 江苏 徐州 221116

摘要: 针对现有彩色图像检索算法存在旋转变化鲁棒性差、特征维度高和检索时间长的问题, 通过融合主曲率的改进方向梯度特征与 HSV 颜色特征, 提出了一种创新的多尺度图像检索方法。该方法从多个尺度将图像表面的几何曲率信息融合到 FHOG 描述符中, 得到基于主曲率的改进方向梯度算法(P-FHOG), 在此基础上进一步融合图像的颜色信息, 得到基于颜色特征与改进方向梯度特征的多尺度图像检索方法(CP-FHOG)。在 Corel-1000 与 Coil-100 数据集上与先进的图像检索方法进行对比实验, 分别取得了 85.89% 和 93.38% 的平均准确率, 该算法相比其他算法准确率更高、旋转变化鲁棒性更强、检索时间更短, 提高了检索效率。

关键词: 图像检索; 颜色信息; 方向梯度; 多尺度; 特征融合

中图分类号: TP391.4

文献标志码: A



江曼, 张皓翔, 程德强, 等. 融合 HSV 与方向梯度特征的多尺度图像检索[J]. 光电工程, 2021, 48(11): 210310

Jiang M, Zhang H X, Cheng D Q, *et al.* Multi-scale image retrieval based on HSV and directional gradient features[J].

Opto-Electron Eng, 2021, 48(11): 210310

Multi-scale image retrieval based on HSV and directional gradient features

Jiang Man¹, Zhang Haoxiang¹, Cheng Deqiang^{1,2*}, Guo Lin¹, Kou Qiqi³, Zhao Lei¹

¹School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China;

²Engineering Research Center of Intelligent Control for Underground Space, Ministry of Education, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China;

³School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China

Abstract: Aiming at the problems of poor robustness of rotation change, high feature dimension, and long retrieval time of existing color image retrieval algorithms, this paper proposed an innovative image retrieval method by fusing color features and improved directional gradient features. It proposed an improved directional gradient algorithm based on the principal curvatures (P-FHOG) by combining the geometric curvature information of the image surface

收稿日期: 2021-09-24; 收到修改稿日期: 2021-11-05

基金项目: 国家自然科学基金资助项目(51774281)

作者简介: 江曼(1996-), 女, 硕士研究生, 主要从事图像处理与模式识别方面的研究。E-mail: jiangman@cumt.edu.cn

通信作者: 程德强(1979-), 男, 博士, 教授, 博士生导师, 主要从事机器视觉与模式识别、图像处理与视频编码、图像智能检测与信息处理方面的研究。E-mail: cdqcumt@126.com

版权所有©2021 中国科学院光电技术研究所

into the FHOG descriptor from multiple scales. At the same time, the color information of the image was further fused to obtain the multi-scale image retrieval method based on the color features and the improved directional gradient features (CP-FHOG). The experiment was compared with the advanced image retrieval methods on the Corel-1000 and Coil-100 data sets, and the average accuracy rates of 85.89% and 93.38% were achieved, respectively. The results show that the proposed algorithm is more accurate and robust (in rotation change) than other algorithms.

Keywords: image retrieval; color information; directional gradient; multiple scales; features fusion

1 引言

图像检索的目的是从大规模数字图像数据库中查看、搜索和恢复图像,随着计算机视觉与数字媒体的迅速发展,图像检索已广泛应用于搜索引擎、数字图书馆、医学图像管理等领域^[1]。图像检索方法可以分为两类:基于文本的图像检索(text-based image retrieval, TBIR)方法和基于内容的图像检索(content-based image retrieval, CBIR)方法。TBIR方法通过人工标注将关键字、字幕等描述作为属性添加到图像中,图像所附的属性很大程度上取决于标注者的主观感知,无法涵盖查询的全部范围,该方法不仅耗费人力,且无法准确检索图像。CBIR方法从图像中提取颜色、纹理和形状等视觉特征^[2],以此作为标准结合机器学习进行图像检索,无需人工标注,提高了检索准确率,节省检索时间,成为该领域的研究热点。

特征提取是通过图像整体和局部的特征信息,获取特征描述符用于检索^[3],是CBIR方法的关键步骤,决定着检索的效率与准确率。近年来,尺度不变特征(scale-invariant feature transform, SIFT)、加速鲁棒特征(speeded up robust features, SURF)、方向梯度直方图(histogram of oriented gradient, HOG)、局部二值模式(local binary pattern, LBP)、区域特征提取算法(maximally stable external regions, MSER)、高斯差分算子(difference of Gaussian, DoG)等特征描述符表现优秀,在CBIR中得到了广泛的应用^[4]。Agarwal等^[5]将方向梯度直方图(HOG)应用于CBIR方法中,有效地提高了检索的准确率。Rui等^[6]提取梯度场方向梯度直方图(gradient field histogram of oriented gradient, GF-HOG)实现了有效的草图检索。Joolee等^[7]将梯度方向的三值直方图(ternary histogram of oriented gradient, THOG)用于视频检索,达到了良好的效果。基于描述符的算法对检索精度有很大程度的提高,但由于图像中包含多种视觉特征,仅使用一种特征难以达到精确的检索效果。为了解决这一问题,学者提出了将多种特征融合的图像检索系统。其中,颜色特征作

为最常用的视觉特征,在图像检索中发挥重要的作用。Cheng等^[8]提出一种基于结合颜色特征与HOG特征结合的图像检索算法,取得了不错的效果。Pavithra等^[9]利用颜色矩信息、LBP特征和Canny边缘检测方法,提取图像的纹理和边缘特征,实现了对相似图像的快速响应和有效检索。Bella等^[10]提出了一种基于灰度共生矩阵和HSV(Hue, Saturation, Value)颜色矩的图像检索系统,融合图像的信息特征,增强了系统的检索性能。Dhiman等^[11]利用RGB通道对图像进行离散小波变换,再利用灰度共生矩阵与改进的LBP特征进行纹理信息融合,达到了高效的图像检索效果。Danapur等^[12]将RGB颜色空间转换为HSV颜色空间,利用旋转局部二值模式(robust local binary pattern, RLBP)描述符提取纹理特征。Khwildi等^[13]提出了一种基于HSV颜色直方图和颜色矩的高动态范围(high dynamic range imaging, HDR)图像检索方法,利用曼哈顿距离度量图像之间的相似度,具有很强的稳定性。随着CBIR方法的快速发展,将其与深度学习算法相结合逐渐成为图像检索的主流。搭建网络模型自动学习图像特征,利用网络的多层次性将图像特征关联,并完成图像分类。AlexNet使用非线性激活函数,提取丰富的特征,提高检索准确率。VGGNet加深网络结构提升性能,通过足量数据刻画模型在空间上的分布。ResNet与Inception系列模型共同展示了基于深度学习的图像检索新模式。

将改进的颜色特征与其他特征相结合的系统虽然有较好的鲁棒性,但由于待检索图像存在目标相似、背景复杂等问题,导致图像的多特征信息提取时间长、特征维度高。并且对于不同角度旋转变化的数据集,图像的纹理特征难以提取,导致现有的图像检索系统对此类数据集检索性能不佳。针对目前图像检索算法存在的旋转变鲁棒性差、特征维度高和检索时间长的的问题,本文提出一种基于颜色特征与改进方向梯度特征的多尺度图像检索方法。首先,考虑到传统HOG方法在检索过程中对图像旋转变化的不适应性,本文引入融合主曲率的多尺度改进方向梯度直方图算法

(principal curvatures-fast histogram of oriented gradients, P-FHOGs), 在多尺度高斯空间内分析图像表面的局部几何曲率信息, 以提高检索对于目标物体不同角度的旋转鲁棒性。同时, 利用 HSV 颜色空间提取图像整体的颜色信息, 并将颜色特征与改进方向梯度特征融合, 得到基于颜色与改进方向梯度特征的图像检索方法(color and principal curvatures-fast histogram of oriented gradients, CP-FHOG), 进一步提高了特征提取的有效性。实验证明, 本文提出的方法可以达到快速、有效的检索需求, 在图像标准数据集中检索准确率显著提高。

2 CP-FHOG 算法与基本原理

2.1 基本原理

2.1.1 多尺度主曲率

在三维欧氏空间中, 给定曲面上一个点 p_0 , 它的法向量 n 的切平面在一个可微曲面上是唯一的。在点 p_0 上有无穷多个包含 n 法向量平面, 在所有的法向量平面与曲面的交线中, 有两条交线的曲率 k 在所有曲率中最大与最小, 分别用 k_1 和 k_2 表示, 这两个曲率称为主曲率(principal curvatures, PCs)。如图 1 所示为 p_0 点的空间主曲率示意图, 表征了图像旋转不变的特征信息, 即无论图像以任何角度旋转, 主曲率都将保持不变。

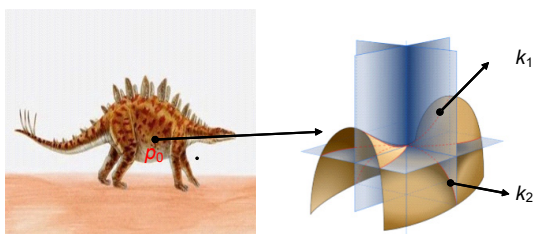


图 1 图像某一点的空间主曲率
Fig. 1 The spatial principal curvatures of an image at a point

要得到彩色图像的主曲率特征, 首先要计算每个像素点的海森矩阵, 然后根据矩阵的行列式和迹得出特征值, 进而计算图像表面几何结构的主曲率信息。对于灰度图像 $I(x, y)$ 上任意一点的海森矩阵 H 定义如下:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} = \begin{bmatrix} (p \otimes I) \otimes d_2 & (p \otimes I') \otimes d_1 \\ (p \otimes I') \otimes d_1 & (d_2 \otimes I) \otimes p \end{bmatrix}, \quad (1)$$

其中: $I' = (d_1 \otimes I) \otimes p$, 式中 \otimes 代表卷积, p 代表插值器, d_1 和 d_2 代表滤波器。以 D_{xx} 为例, $(p \otimes I) \otimes d_2$ 表

示插值器 p 沿列先进行 1D 卷积, 然后滤波器 d_2 沿行再进行 1D 卷积, 由此可得海森矩阵的值^[14]。如图 2 所示为 p_0 点的海森矩阵。

海森矩阵特征值的和等于矩阵的迹, 矩阵特征值的乘积等于矩阵的行列式:

$$\text{tr } H = k_1 + k_2 = D_{xx} + D_{yy}, \quad (2)$$

$$\det H = k_1 k_2 = D_{xx} D_{yy} - (D_{xy})^2. \quad (3)$$

利用 H 的行列式和迹可以得到两个特征值:

$$k_1 = \frac{(D_{xx} + D_{yy}) + \sqrt{(D_{xx} - D_{yy})^2 + 4D_{xy}^2}}{2}, \quad (4)$$

$$k_2 = \frac{(D_{xx} + D_{yy}) - \sqrt{(D_{xx} - D_{yy})^2 + 4D_{xy}^2}}{2}. \quad (5)$$

海森矩阵的特征值 k_1 和 k_2 即为主曲率。经过不同高斯空间后得到不同尺度的主曲率信息, 主曲率描述了纹理特征的局部极大值与极小值, 反映了图像多个尺度的宏观和微观纹理信息。

2.1.2 FHOG 描述符

HOG 描述符自提出以来, 在目标检测与图像检索领域得到了广泛的应用。HOG 描述符对图像的几何形变和光学形变具有很好的鲁棒性, 将图像分割成许多小单元格, 并计算每个小单元格中的梯度和像素, 通过合并小单元格提取像素信息来构建图像的方向梯度直方图。HOG 描述符为了保证提取纹理特征的准确性, 常常要减小滑动图像块的步长, 这样会导致最终提取的特征维数高、时间长和过拟合。

为了解决上述问题, Felzenszwalb 等^[15]提出了快速的方向梯度直方图(fast histogram of oriented gradients, FHOG), 最初应用于目标检测中, 通过把高维特征融入到相关滤波框架中来有效地提取特征, 事实证明, FHOG 描述符可以更少的计算时间获得优于传统 HOG 描述符的性能。如图 3 所示为 FHOG 描述符提取特征的流程图。FHOG 描述符的滑动图像块中, 九个小单元格被分成四组, 在每组中都能提取两个 9

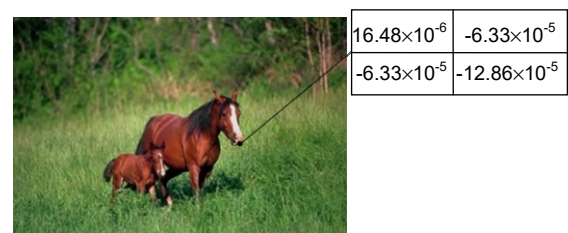


图 2 图像某一点处的海森矩阵
Fig. 2 The Hessian matrix at some point in the image

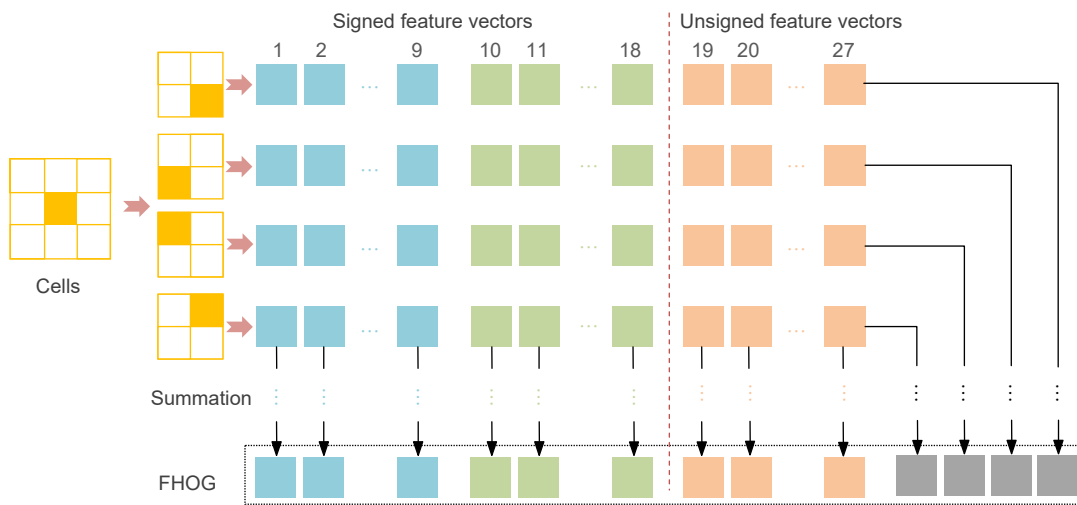


图 3 FHOG 描述符提取特征流程图

Fig. 3 Flow chart of the FHOG descriptor extraction feature

维有符号 HOG 特征向量和一个 9 维无符号 HOG 特征向量, 组合成一个 27 维特征向量, 四组特征向量组合成一个 4×27 特征矩阵。

将该矩阵的列相加得到 1×27 的列向量, 矩阵的行相加得到 4×1 的行向量。通过连接该行向量与列向量得到一个降维后的 31 维特征向量, 此向量即描述了具有 9 个小单元格的滑动图像块的特征。

2.1.3 HSV 颜色空间

颜色特征是图像检索中最常用的特征之一, 对图像平移、失真等相关变化具有很强的鲁棒性。颜色特征通常在颜色空间中用颜色直方图来描述, 常用的颜色空间有 HSV、RGB、YCbCr、Lab、Luv 等, 颜色直方图是图像颜色特征检索的关键, 有较强的稳定性。

HSV 颜色空间由三个分量组成: 色调 H、饱和度 S 和亮度 V。由于 H 和 S 两个分量十分接近人类视觉感知的颜色, HSV 颜色空间被广泛应用于图像检索中。V 分量与其他两个分量分开, 使得颜色表示更加直观, 因此 HSV 颜色空间更具有统一性、整体性、紧凑性和自然性。彩色图像从 RGB 颜色空间转换到 HSV 颜色空间的公式如下:

$$I_H = \arccos \left\{ \frac{1}{2} \times \frac{(R-G)+(R-B)}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right\}, \quad (6)$$

$$I_S = 1 - \frac{3}{R+G+B} [\min(R,G,B)], \quad (7)$$

$$I_V = \frac{1}{3}(R+G+B) \quad (8)$$

在 HSV 颜色空间中, I_H 描绘色调, I_S 描绘饱和度,

I_V 描绘亮度^[12]。

2.2 CP-FHOG 算法

2.2.1 CP-FHOG 算法流程

为了解决现有彩色图像检索算法旋转变化鲁棒性差、特征维度高和检索时间长等问题, 本文提出了基于多尺度主曲率方向梯度与颜色特征融合的图像检索算法。首先, 为了更好地表征图像的纹理特征, 提出基于主曲率的方向梯度特征描述符, 并构造多层高斯尺度空间, 提高特征描述符的旋转与尺度鲁棒性; 其次, 将图像的颜色特征与改进的方向梯度特征描述符融合, 提高检索描述符的稳定性。

如图 4 所示为 CP-FHOG 算法流程图, 首先, 输入的彩色图像通过高斯空间转换为灰度高斯图像, 计算灰度高斯图像的主曲率信息融入 FHOG 描述符, 同时对主曲率信息进行多采样构造基于主曲率的混合采样方向梯度特征(P-FHOG1、P-FHOG2、P-FHOG3), 再将三个尺度的特征融合得到图像的多尺度方向梯度特征直方图 P-FHOGs。其次, 将 RGB 彩色图像通过 HSV 颜色空间进行量化, 提取图像的颜色信息以构造颜色特征直方图。最后, 将图像的方向梯度特征与颜色特征相融合, 得到图像的特征融合直方图。

2.2.2 多尺度主曲率的方向梯度特征

如图 4 所示, 输入图像分别通过三个尺度的高斯空间得到灰度图像, 进而提取灰度图像的主曲率信息, FHOG 描述符通过不同尺度的采样窗口提取灰度图像的方向梯度信息, P-FHOG 算法将主曲率信息与方向梯度信息融合, 得到图像的方向梯度特征直方图。

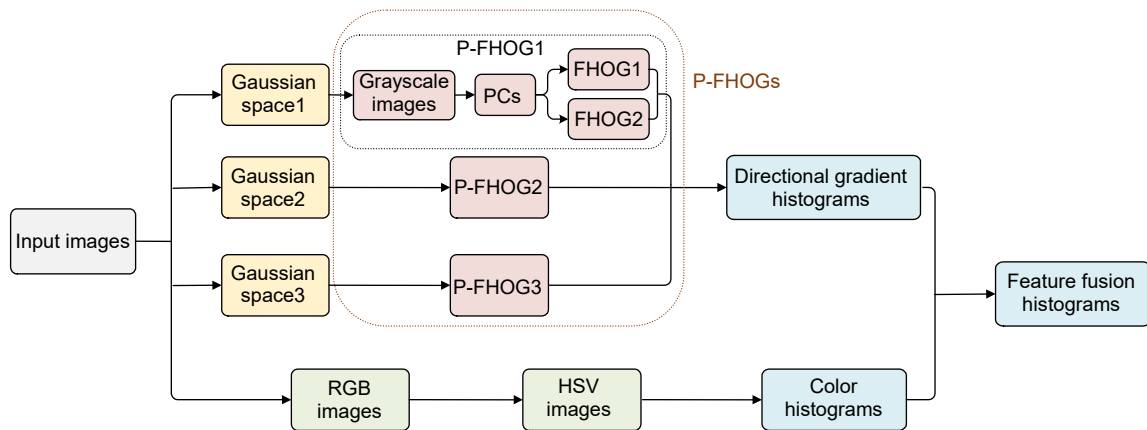


图 4 CP-FHOG 算法流程图
Fig. 4 Flow chart of the CP-FHOG algorithm

关于海森矩阵的计算方法,已在本文 2.1.1 节中详细叙述,海森矩阵的特征值 k_1 和 k_2 即为图像的主曲率,代表每个像素点的最大值与最小值曲率,描述了图像几何旋转不变性质。得到图像的主曲率信息之后,为了提高区分度与对比度,便于计算融合主曲率的方向梯度信息,利用比值公式与幂指数转换公式^[16]对 k_1 和 k_2 进行量化:

$$\eta = \frac{k_1}{k_2}, \tag{9}$$

$$\zeta = \frac{2}{1 + \exp(-2\eta)} - 1, \tag{10}$$

式中: η 和 ζ 分别为比值量化与幂指数量化后的主曲率信息,接着再用 FHOG 描述符对主曲率信息 ζ 进行统计。

本文中 FHOG 描述符采用两种采样模式,称为 FHOG1 和 FHOG2。首先将每个图像块分割成九个小单元格,再将这些小单元格分为四组像素块,每组像素块包含四个小单元格,每个小单元格有 8×8 或 16×16 个像素点,即在提取特征的过程中, FHOG1 和 FHOG2 分别使用 $8 \text{ pixels} \times 8 \text{ pixels}$ 和 $16 \text{ pixels} \times 16 \text{ pixels}$ 的采样窗口。同时,将 2.1.2 中提到的两个 9 维有符号 HOG 特征向量和一个 9 维无符号 HOG 特征向量,改进成两个 30 维有符号特征和一个 30 维无符号特征向量。因此无论使用哪种采样模式,每个像素块总是提供 3×30 维特征向量。最后将四组像素块的四个 3×30 维特征向量通过列累加并降维形成一个 3×30 维的特征向量,其包含所有像素块的特征信息,因此可以完整描述图像的方向梯度特征。将两种采样模式下的特征向量 α_{FHOG1} 与 α_{FHOG2} 线性组合,得出 P-FHOG 算法的特征向量 $\alpha_{\text{P-FHOG1}}$:

$$\alpha_{\text{P-FHOG1}} = \frac{1}{2}(\alpha_{\text{FHOG1}} + \alpha_{\text{FHOG2}}) \tag{11}$$

将 P-FHOG 算法在三个尺度空间中进行联接,从而构建 P-FHOGs 算法模型。P-FHOG 算法在每个尺度空间内有 90 维特征向量,则三个不同的尺度空间构建的 P-FHOGs 算法共统计 270 维特征向量:

$$F_{\text{P-FHOGs}} = [\alpha_{\text{P-FHOG1}} \quad \alpha_{\text{P-FHOG2}} \quad \alpha_{\text{P-FHOG3}}], \tag{12}$$

式中: α 为 P-FHOG 算法在提取特征过程中产生的向量。因此,本文所构建的改进方向梯度特征直方图有 270 bins。在下面的章节中,使用 $F_{\text{P-FHOGs}}$ 来描述算法的方向梯度特征。

2.2.3 颜色特征

为了提取图像的颜色特征,首先在 HSV 颜色空间中进行量化。色调、饱和度和亮度分别在 $[0, 360]$ 、 $[0, 1]$ 和 $[0, 1]$ 的范围内。HSV 颜色空间的量化式^[17]如下:

$$Q_H = \begin{cases} 0, & I_H \in [0, 20] \\ 0, & I_H \in (315, 360] \\ 1, & I_H \in (20, 40] \\ 2, & I_H \in (40, 75] \\ 3, & I_H \in (75, 155] \\ 4, & I_H \in (155, 190] \\ 5, & I_H \in (190, 270] \\ 6, & I_H \in (270, 295] \\ 7, & I_H \in (295, 315] \end{cases}, \tag{13}$$

$$Q_S = \begin{cases} 0, & I_S \in [0, 0.2] \\ 1, & I_S \in (0.2, 0.7] \\ 2, & I_S \in (0.7, 1] \end{cases}, \tag{14}$$

$$Q_V = \begin{cases} 0, & I_V \in [0, 0.2] \\ 1, & I_V \in (0.2, 0.7] \\ 2, & I_V \in (0.7, 1] \end{cases}, \tag{15}$$

式中: Q_H, Q_S, Q_V 分别为每个通道的量化等级, 将 I_{HSV} 量化为 Q_{HSV} 后, 使用下列等式进行图像的颜色特征构造:

$$F_{HSV} = L_H \times Q_H + L_S \times Q_S + L_V \times Q_V, \quad (16)$$

式中: L_H, L_S, L_V 为量化后每个通道的权重值, 分别设置为 9, 3, 1。则等式即为

$$F_{HSV} = 9 \times Q_H + 3 \times Q_S + 1 \times Q_V, \quad (17)$$

式中: $F_{HSV} \in [1, 2, \dots, 71]$ 表示图像的颜色特征。本文算法所构建的图像颜色特征直方图有 72 bins, 如图 5。

2.2.4 特征融合

融合主曲率的方向梯度特征直观地反映了图像局部的纹理信息, 对旋转变换、尺度变化有很强的鲁棒性, 其提取的灰度信息不能完整反映图像的全局特征。图像的颜色直方图属于全局特征, 对图像整体的平移、几何变化有更强的稳定性。如图 6 所示, 本文将基于主曲率的方向梯度特征与颜色特征, 经过同比例级联加权融合, 得到最终的图像特征 F , 提高了检索精度。

3 实验与分析

本文利用 Matlab2019 在两个公开的图像检索数据集 Corel-1000^[18] 和 COIL-100^[19] 上进行对比实验, 来验证 CP-FHOG 算法的有效性。采用检索准确率 AP, 平均准确度 mAP, 平均召回率 Recall, 特征提取时间

SFET, 检索时间 RT, 特征维度 Dimension 作为算法性能的评价指标, 数据集采用的评价方法为欧氏距离。

3.1 实验数据

实验数据集 Corel-1000, 包含 1000 张图像, 由 10 个不同的类别组成, 每类包含 100 幅大小为 256 pixels×348 pixels 或 348 pixels×256 pixels 的图像, 广泛应用于图像检索和分类。如图 7 所示, 数据集包括背景和复杂对象, 如 African, sea, architecture, bus, dinosaur, elephant, flower, horse, mountain, food, 类别间有很强的相似性, 通常用来测试算法的检索精度和效率。

实验数据集 Coil-100, 包含 7200 张图像, 由 100 个不同的类别组成, 每个图像类别都包含 72 幅大小为 128 pixels × 128 pixels, 0~360° 不同视角下旋转的相同对象。由于这种类别中旋转视角的差异, 该数据库常用于图像检索的准确性评价^[20]。如图 8 所示为数据集中 statue, car, frog, white cup, rolaid, mud pot, truck, cat, tomato, jug, pink cup, stick 类别的成像情况。

3.2 参数设置

高斯空间的尺度标准差 δ 和 FHOG 描述符的采样模式 m 影响滤波输出的灰度图像大小与提取特征的完整性, 为了避免图像尺度失真或特征提取不完全导

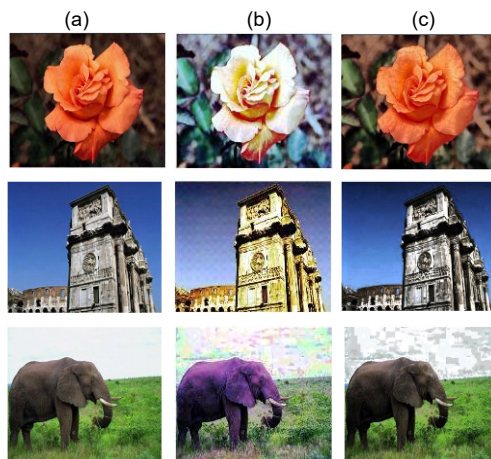


图 5 颜色特征提取。(a) 输入图像; (b) RGB 转换图像; (c) HSV 转换图像

Fig. 5 Extraction of the color features. (a) Input images; (b) RGB converted images; (c) HSV converted images

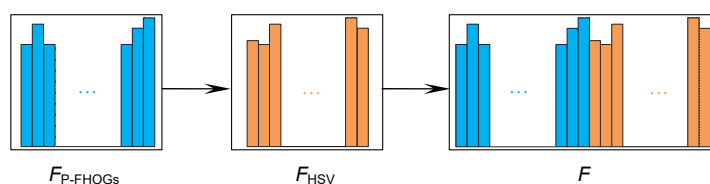


图 6 特征融合级联直方图

Fig. 6 Feature fusion cascade histogram



图 7 Corel-1000 数据集的样本图像
Fig. 7 Sample images of the Corel-1000 dataset

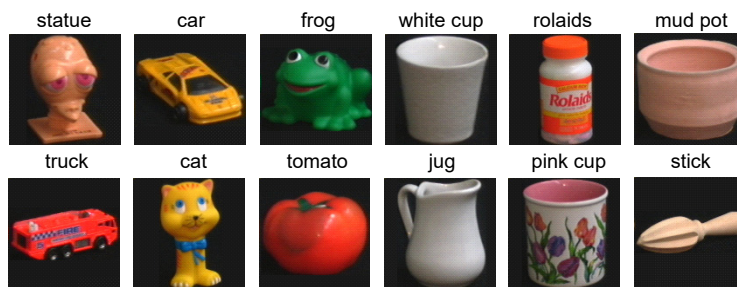


图 8 Coil-100 数据集的样本图像
Fig. 8 Sample images of the Coil-100 dataset

致检索准确率降低, δ 和 m 应当合理选择。图像的局部信息特征与方向直方图数量 b 有关, b 的选取也会对检索的精度与效率产生影响。本文在实验数据集 Corel-1000 上采用交叉验证法^[21]选取以上 3 个参数。

如图 9 所示为 δ 和 m 对图像检索平均准确率的影响。为验证参数, 令方向直方图数量 $b=9$, 适当的改变采样模式 m 与尺度空间标准差 δ , 即算法在不同的尺度空间与不同的采样模式下进行实验。可以看出, 在同一尺度空间中, 采样模式中的像素值过大或过小都会使检索准确率下降。在尺度空间 $\delta=(0.2,0.5,1)$ 中, 采样模式 m 设置为(8,16)算法表现最好, 检索平均准确率为 84.75%。

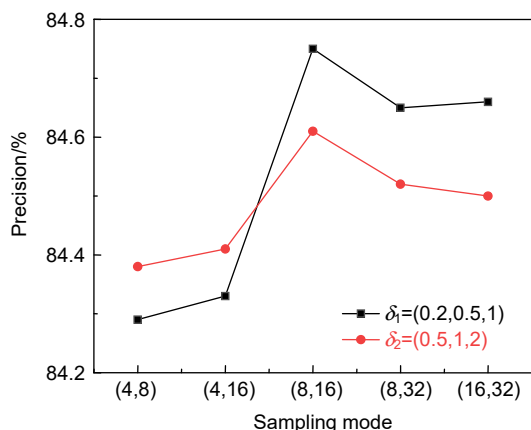


图 9 不同的 δ 和 m 对准确率的影响
Fig. 9 Influence for different δ and m on accuracy

另外, 检索平均准确率还与方向直方图数量 b 有关, 指定最优的采样模式 $m=(8,16)$ 与尺度空间 $\delta_1=(0.2,0.5,1)$ 的条件下, 方向直方图数量 b 对图像检索平均准确率的影响如图 10 所示。

适当增加 b 有利于像素块内的特征提取, 平均准确率有较大的上升, 而 b 过大时直方图对像素块的分辨率产生影响, 从而准确率下降, 当方向直方图数量 $b=30$ bins 时, 得到算法最佳的平均准确率, 可得 CP-FHOG 算法的最佳参数设置如表 1 所示。

3.3 结果与分析

为了评估 CP-FHOG 算法对图像检索的性能, 本文分别在两个数据集上进行对比实验, 在数据集中通

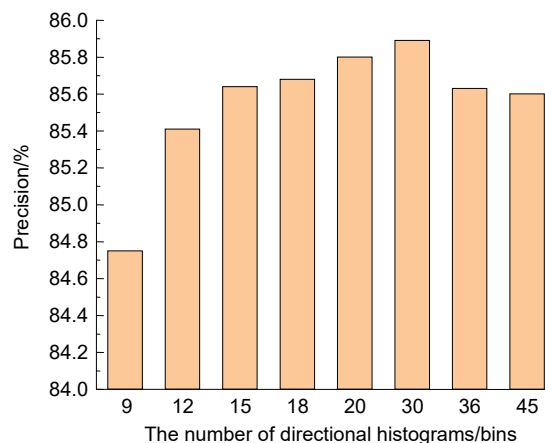


图 10 不同的 b 对准确率的影响
Fig. 10 Influence for different b on accuracy

表 1 实验参数设置

Table 1 Experimental parameter setting

参数	δ	m	b
CP-FHOG	(0.2,0.5,1)	(8,16)	30

过欧氏距离, 取与查询图像最接近的 top20 张图像分析实验结果, 对比实验的结果如表 2~5 所示。

3.3.1 特征融合的有效性

对于 Corel-1000 数据集, 将 CP-FHOG 算法与其他 CBIR 先进算法进行对比试验, 并取每种算法检索结果的前 20 张进行性能评价。如图 11 所示为 CP-FHOG 算法在 African, flower 类别中的检索结果, 表 2 和表 3 可见每个类别图像的检索准确率与召回率。CP-FHOG 算法提取颜色和纹理特征用于彩色图像检索, 在超过半数的图像类别中都具有较好的精度, 尤

其在 African, bus, dinosaur 类别中获得了最高的准确率。对比 Pavithra 等^[9]提出的融合纹理特征与边缘特征的算法, CP-FHOG 算法加入了颜色特征, 有更有效特征表达。其他算法在一些类别中也得到了很好的检索结果, 如 Kundu 等^[22]的算法在大象这一类别中有更高的准确率。

如图 12 所示为单一特征 HSV 颜色空间、改进方向梯度直方图特征 P-FHOGs 算法与融合特征 CP-FHOG 算法的实验结果对比图。由图可以看出, 当使用单一特征进行检索时, 效果远不如特征融合后的算法结果。同时对比表 2 的最后三列, 当仅使用 HSV 空间提取颜色特征进行检索时, 各类别的准确率均低于 CP-FHOG 算法检索的准确率。当去除颜色特征仅使用 P-FHOG 算法进行检索时, 除 sea 类别之外其余类别的准确率仍低于 CP-FHOG 算法。而 sea 类别由

表 2 数据集 Corel-1000 上的各类别检索准确率/%

Table 2 Retrieval accuracy of each category on the Corel-1000 dataset/%

Category	Pavithra ^[9]	Kundu ^[22]	Dubey ^[23]	Sonug ^[25]	Xiao ^[26]	HSV	P-FHOGs	CP-FHOG
African	81.0	44.0	75.0	67.6	67.0	93.4	62.5	98.6
Sea	66.0	32.0	55.0	59.8	60.0	55.5	77.4	69.7
Architecture	78.8	52.0	67.0	58.0	56.0	58.7	64.8	66.7
Bus	96.3	62.0	95.0	94.0	96.0	91.5	99.0	99.6
Dinosaur	100.0	40.0	97.0	99.8	98.0	99.7	100.0	100.0
Elephant	70.8	80.0	63.0	58.0	53.0	54.6	58.2	70.4
Flower	95.8	57.0	93.0	88.6	93.0	87.5	89.3	95.8
Horse	98.8	75.0	89.0	93.8	82.0	97.6	81.7	98.7
Mountain	67.8	57.0	45.0	47.8	46.0	57.6	54.2	73.0
Food	77.3	56.0	70.0	49.2	58.0	79.3	65.1	86.5

表 3 数据集 Corel-1000 上的各类别检索召回率/%

Table 3 Retrieval recall rate of each category on Corel-1000 dataset/%

Category	Pavithra ^[9]	Kundu ^[22]	Dubey ^[23]	Sonug ^[25]	Xiao ^[26]	HSV	P-FHOGs	CP-FHOG
African	16.2	8.8	15.0	13.5	13.4	18.6	12.5	19.7
Sea	13.2	6.4	11.0	12.0	12.0	11.1	15.4	13.9
Architecture	15.8	10.4	13.4	11.6	11.2	11.7	12.9	13.3
Bus	19.3	12.4	19.0	18.8	19.2	18.3	19.0	19.9
Dinosaur	20.0	8.0	19.4	20.0	98.0	19.9	20.0	20.0
Elephant	14.2	16.0	12.6	11.6	10.6	10.9	11.6	14.1
Flower	19.2	11.4	18.6	17.7	18.6	17.5	17.8	19.2
Horse	19.8	15.0	17.8	18.8	16.4	19.5	16.3	19.7
Mountain	13.6	11.4	9.0	9.6	9.2	11.5	10.8	14.6
Food	15.5	11.2	14.0	9.8	11.6	15.8	13.0	17.3

表 4 数据集 Corel-1000 上的各参数对比

Table 4 Comparison of parameters on the Corel-1000 dataset

Algorithm	mAP/%	Recall/%	SFET/s	RT/s	Dimension
Pavithra ^[9]	83.26	16.65	0.671	1.108	768
Kundu ^[22]	55.50	11.10	0.400	-	99
Sun ^[24]	83.50	16.70	9.150	1.027	900
Dubey ^[23]	74.90	14.98	102.400	16.490	1024
Sonug ^[25]	71.66	14.33	-	-	4096
Xiao ^[26]	70.10	14.02	-	-	63
HSV	77.54	14.18	0.020	0.023	72
P-FHOGs	75.22	14.02	0.053	0.021	270
CP-FHOG	85.89	17.18	0.067	0.048	342

表 5 数据集 Corel-1000 上与深度学习算法对比各类别检索准确率/%

Table 5 Retrieval accuracy of each category compared with the deep learning algorithm on the Corel-1000 dataset/%

Category	AlexNet ^[24]	GoogleNet	VGG-19	ResNet-50	CP-FHOG
African	33.0	65.0	68.0	78.0	98.6
Sea	22.0	75.0	79.0	77.0	69.7
Architecture	40.0	90.0	90.0	99.0	66.7
Bus	23.3	87.0	88.0	90.0	99.6
Dinosaur	71.0	88.0	90.0	88.0	100.0
Elephant	27.5	80.0	85.0	87.0	70.4
Flower	50.0	91.0	93.0	95.0	95.8
Horse	59.2	83.0	88.0	93.0	98.7
Mountain	26.7	80.0	90.0	98.0	73.0
Food	65.0	80.0	81.0	85.0	86.5

于场景中颜色特征提取相对较弱, 因此融合颜色之后的效果不如单独使用方向梯度特征。由此可以看出, 将颜色特征与改进方向梯度方向融合, 显著提高了图像检索的准确率, 达到了精准检索的目的。

在数据集 Corel-1000 上对各检索算法的平均准确率、召回率、单张特征提取时间、单张检索时间、特征维度进行比较。如表 4 所示, CP-FHOG 算法的平均准确率达到最优, 相比 HSV 颜色特征与 P-FHOGs 算法分别提升了 8.35%、10.67%。仅使用 HSV 颜色空间的特征提取时间最短, 仅使用 P-FHOGs 时的检索时间最短。Dubey 等^[23]算法引入了基于局部二进制模式的多通道加法器和解码器用于图像检索, 导致特征维度过大, 检索用时最长。与该算法相比, CP-FHOG 算法利用多尺度高斯滤波输出灰度图像, 迭代次数更少, 提取特征维度降低至 342 维, 检索时间明显减少。总的来看, 特征融合后的 CP-FHOG 算法不仅检索精

度高, 在检索时间、特征维度方面也都有显著的改善。

如表 5 所示, 将融合后的算法与 AlexNet、GoogleNet、VGGNet、ResNet 这些经典的深度学习算法进行对比实验, CP-FHOG 算法检索明显高于 AlexNet^[24]与 GoogleNet 模型的准确率, 与 VGG-19 网络对比, 本文算法在大多数类别也取得了较高的准确率, 与 ResNet 网络相对比, 本文算法与其在不同类别各有所长, 且总体平均准确率略高, 可以看出本文算法用于图像检索的有效性。

3.3.2 多尺度主曲率的旋转鲁棒性

如图 13 所示, 数据集 Coil-100 包含不同旋转角度的检索目标。任意选取数据集中 12 类的检索结果, 将 CP-FHOG 算法与主流的检索算法 SIFT、SURF、LBP 进行比较。

如表 6 所示为数据集 Coil-100 上各类别的检索准确率与平均准确率, 结果表明 CP-FHOG 算法在大多

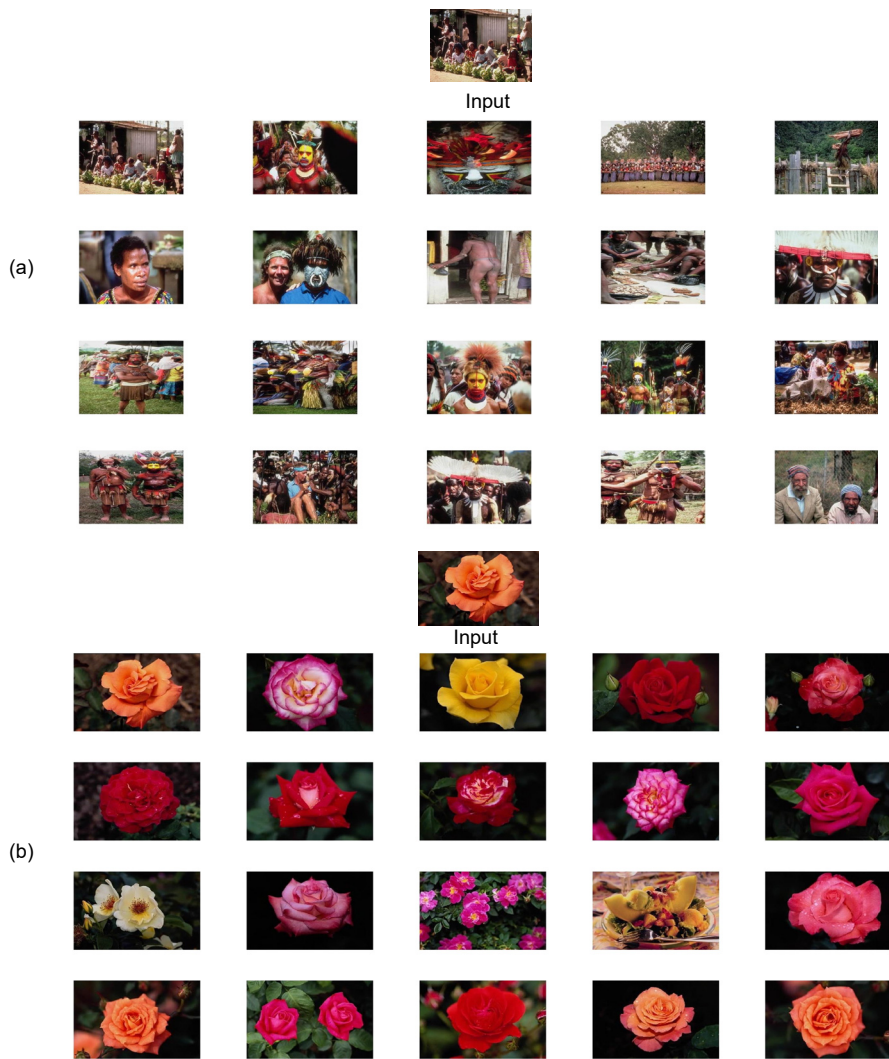


图 11 Corel-1000 数据集的检索结果。(a) Africans; (b) Flowers
 Fig. 11 The retrieval results of the Corel-1000 dataset. (a) Africans; (b) Flowers

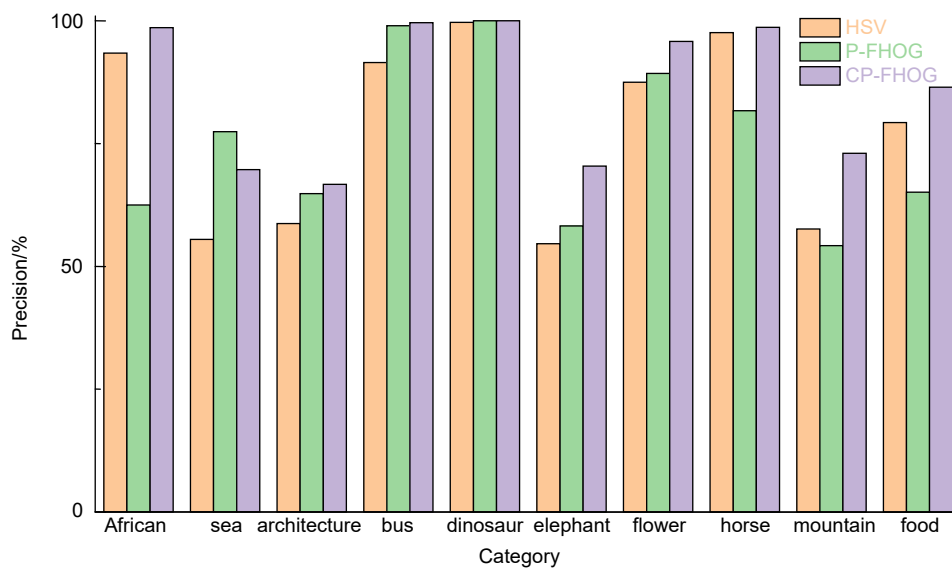


图 12 消融实验结果对比图
 Fig. 12 Comparison of the ablation experiment results

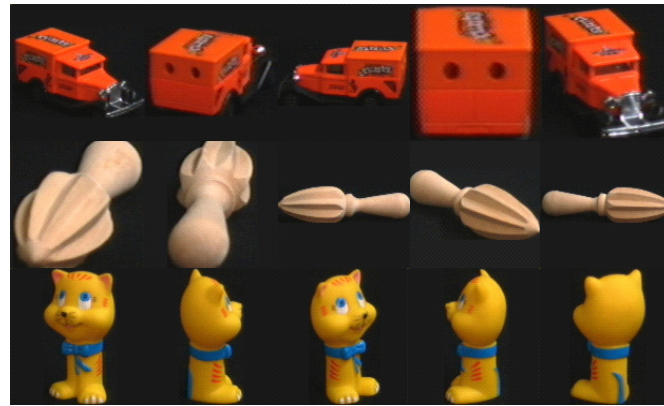


图 13 不同旋转角度的检索目标
Fig. 13 Retrieval targets with different rotation angles

表 6 数据集 Coil-100 上的各类别检索准确率/%
Table 6 Retrieval accuracy of each category in the COIL-100 dataset/%

Category	CP-FHOG	HSV	P-FHOGs	Ahmed ^[27]	SIFT	SURF	MSER	LBP	RGLBP
Tomato	98.7	93.5	89.3	93.0	15.0	75.0	15.0	35.0	20.0
Cat	100.0	100.0	86.3	90.0	32.0	45.0	55.0	40.0	25.0
Statue	100.0	100.0	63.2	100.0	35.0	30.0	45.0	25.0	55.0
Stick	60.9	52.8	25.8	93.0	30.0	35.0	90.0	50.0	10.0
Rolaids	100.0	100.0	95.3	65.0	20.0	60.0	40.0	65.0	85.0
Mud pot	100.0	100.0	99.8	100.0	20.0	45.0	90.0	70.0	50.0
Frog	99.0	91.2	60.8	95.0	20.0	65.0	45.0	55.0	45.0
Jug	98.8	98.2	57.3	100.0	20.0	45.0	70.0	65.0	60.0
Car	93.3	98.7	16.9	98.0	22.0	65.0	22.0	60.0	55.0
Pink cup	100.0	100.0	70.1	88.0	40.0	50.0	35.0	60.0	50.0
White cup	100.0	100.0	96.8	94.0	45.0	40.0	60.0	25.0	50.0
Truck	69.9	52.1	30.8	90.0	15.0	35.0	35.0	30.0	60.0

数类别中具有最高的检索精度，并且在 cat, statue, rolaids, mud pot, pink cup, white cup 类别中准确率达到 100%。对比经典的 SIFT, SURF, MSER 等图像检索算法，CP-FHOG 算法结合主曲率，提取的特征保留了显著的纹理信息，有效地抵抗旋转变换，性能有很大的提升，获得了更准确的检索结果。对比 Ahmed 等^[27]提出的算法，同样结合了颜色和纹理特征，但是对于该数据集中存在视角差异，CP-FHOG 算法结合主曲率，提取的特征保留了显著的纹理信息，更有效地抵抗旋转变换，获得了更准确的检索结果。对比 HSV, P-FHOGs, CP-FHOG 算法的结果，HSV 颜色空间在此数据集上的检索平均准确率为 90.5%，融合 P-FHOGs 后得到的 CP-FHOG 算法准确率为

93.4%，提升了 2.9%。不难发现，相比于使用单一特征进行检索，特征融合后的检索准确率有明显的提升，性能也更加稳定，进一步证明了本文算法的旋转鲁棒性以及检索有效性。

4 结 论

针对目前彩色图像检索存在的特征提取时间长、特征维度高，对于不同视角的样本数据集检索性能差的问题，本文提出了一种基于改进方向梯度与颜色特征融合的多尺度图像检索方法，通过 HSV 颜色空间获取图像的颜色信息，有效提取图像的整体特征。针对提取单一颜色特征存在的局部特征信息丢失的问题，在多个尺度空间中将主曲率信息融入 FHOG 描述符

中, 在保留颜色特征的同时提取图像局部纹理特征, 有效地区分了相似类别间颜色差异, 同时避免了图像纹理信息的混淆, 降低了特征的维度。利用欧氏距离在 Corel-1000 和 Coil-100 数据集上检索, 将提出的算法与主流的特征提取和图像检索算法进行比较与性能评估。实验结果表明, 本文提出的 CP-FHOG 算法与 HSV 和 P-FHOGs 算法相比, 提取的细节特征更加丰富, 旋转鲁棒性更强, 在包含复杂背景与不同旋转角度目标的数据集中检索准确率得到了显著的提升, 检索时间与特征维度也有很大的改善。下一步研究内容是提取图像的边缘特征并结合机器学习方法, 实现更高效的彩色图像检索。

参考文献

- [1] Yan C G, Gong B, Wei Y X, *et al.* Deep multi-view enhancement hashing for image retrieval[J]. *IEEE Trans Pattern Mach Intell*, 2021, **43**(4): 1445–1451.
- [2] Kou Q Q, Cheng D Q, Yu W J, *et al.* Texture target classification with CLBP and local geometric features[J]. *Opto-Electron Eng*, 2019, **46**(11): 180604.
寇旗旗, 程德强, 于文洁, 等. 融合 CLBP 和局部几何特征的纹理目标分类[J]. *光电工程*, 2019, **46**(11): 180604.
- [3] Liu F, Wu Z W, Yang A Z, *et al.* Multi-scale feature fusion based adaptive object detection for UAV[J]. *Acta Opt Sin*, 2020, **40**(10): 1015002.
刘芳, 吴志威, 杨安喆, 等. 基于多尺度特征融合的自适应无人机目标检测[J]. *光学学报*, 2020, **40**(10): 1015002.
- [4] Celik C, Bilge H S. Content based image retrieval with sparse representations and local feature descriptors: a comparative study[J]. *Pattern Recognit*, 2017, **68**: 1–13.
- [5] Agarwal M, Maheshwari R P. HOG feature and vocabulary tree for content-based image retrieval[J]. *Int J Signal Imaging Syst Eng*, 2011, **3**(4): 246–254.
- [6] Hu R, Barnard M, Collomosse J. Gradient field descriptor for sketch based retrieval and localization[C]//*Proceedings of 2010 IEEE International Conference on Image Processing*, Hong Kong, China, 2010: 1025–1028.
- [7] Joolee J B, Lee Y K. Video retrieval based on image queries using THOG for augmented reality environments[C]//*Proceedings of 2018 IEEE International Conference on Big Data and Smart Computing*, Shanghai, China, 2018: 557–560.
- [8] Cheng D Q, Zhang H X, Jiang M, *et al.* Color image retrieval method fusing principal curvature and color information[J]. *J Comput Aided Des Comput Graph*, 2021, **33**(2): 223–231.
程德强, 张皓翔, 江曼, 等. 融合主曲率与颜色信息的彩色图像检索算法[J]. *计算机辅助设计与图形学学报*, 2021, **33**(2): 223–231.
- [9] Pavithra L K, Sharmila T S. An efficient framework for image retrieval using color, texture and edge features[J]. *Comput Elect Eng*, 2018, **70**: 580–593.
- [10] Bella M I T, Vasuki A. An efficient image retrieval framework using fused information feature[J]. *Comput Elect Eng*, 2019, **75**: 46–60.
- [11] Garg M, Dhiman G. A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants[J]. *Neural Comput Appl*, 2020, **33**(4): 1311–1328.
- [12] Danapur N, Dizaj S A A, Rostami V. An efficient image retrieval based on an integration of HSV, RLBP, and CENTRIST features using ensemble classifier learning[J]. *Multimed Tools Appl*, 2020, **79**(33): 24463–24486.
- [13] Khwildi R, Ouled Zaid A. HDR image retrieval by using color-based descriptor and tone mapping operator[J]. *Vis Comput*, 2020, **36**(8): 1111–1126.
- [14] Farid H, Simoncelli E P. Differentiation of discrete multidimensional signals[J]. *IEEE Trans Image Process*, 2004, **13**(4): 496–508.
- [15] Felzenszwalb P F, Girshick R B, McAllester D, *et al.* Object detection with discriminatively trained part-based models[J]. *IEEE Trans Pattern Anal Mach Intell*, 2010, **32**(9): 1627–1645.
- [16] Kou Q Q, Cheng D Q, Zhuang H D, *et al.* Cross-complementary local binary pattern for robust texture classification[J]. *IEEE Signal Process Lett*, 2018, **26**(1): 129–133.
- [17] Zhang H X, Jiang M, Kou Q Q. Color image retrieval algorithm fusing color and principal curvatures information[J]. *IEEE Access*, 2020, **8**: 184945–184954.
- [18] Wang J Z, Li J, Wiederhold G. SIMPLcity: semantics-sensitive integrated matching for picture libraries[J]. *IEEE Trans Pattern Anal Mach Intell*, 2001, **23**(9): 947–963.
- [19] Nene S A, Nayar S K, Murase H. Columbia object image library (COIL-100)[R]. New York: Columbia University, 1996.
- [20] Kavitha H, Sudhamani M V. Object Based Image Retrieval from Database Using Combined Features[C]//*Proceedings of the 2014 Fifth International Conference on Signal and Image Processing*. IEEE, Bangalore, INDIA, 2014: 161–165.
- [21] Lv C, Cheng D Q, Kou Q Q, *et al.* Target tracking algorithm based on YOLOv3 and ASMS[J]. *Opto-Electron Eng*, 2021, **48**(2): 200175.
吕晨, 程德强, 寇旗旗, 等. 基于 YOLOv3 和 ASMS 的目标跟踪算法[J]. *光电工程*, 2021, **48**(2): 200175.
- [22] Kundu M K, Chowdhury M, Bulo S R. A graph-based relevance feedback mechanism in content-based image retrieval[J]. *Knowl-Based Syst*, 2015, **73**: 254–264.
- [23] Dubey S R, Singh S K, Singh R K. Multichannel decoded local binary patterns for content-based image retrieval[J]. *IEEE Trans Image Process*, 2016, **25**(9): 4018–4032.
- [24] Sun Q P. Research on image retrieval based on deep learning[J]. *Jingdezhen Compr Coll J*, 2018, **33**(3): 15–18.
孙奇平. 基于深度学习的图像检索研究[J]. *景德镇学院学报*, 2018, **33**(3): 15–18.
- [25] Somnugpong S, Khiewwan K. Content-based image retrieval using a combination of color correlograms and edge direction histogram[C]//*Proceedings of the 2016 13th International Joint Conference on Computer Science and Software Engineering*, Khon Kaen, Thailand, 2016: 1–5.
- [26] Xiao Y, Wu J X, Yuan J S. mCENTRIST: a multi-channel feature generation mechanism for scene categorization[J]. *IEEE Trans Image Process*, 2014, **23**(2): 823–836.
- [27] Ahmed K T, Ummesafi S, Iqbal A. Content based image retrieval using image features information fusion[J]. *Inf Fusion*, 2019, **51**: 76–99.

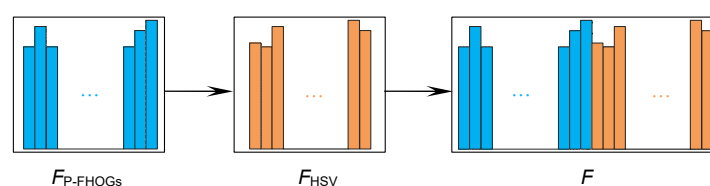
Multi-scale image retrieval based on HSV and directional gradient features

Jiang Man¹, Zhang Haoxiang¹, Cheng Deqiang^{1,2*}, Guo Lin¹, Kou Qiqi³, Zhao Lei¹

¹School of Information and Control Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China;

²Engineering Research Center of Intelligent Control for Underground Space, Ministry of Education, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China;

³School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China



Feature fusion cascade histogram

Overview: With the rapid development of computer vision and digital media, image retrieval has been successfully applied to search engines, digital libraries, medical image management, and other fields. For current color image retrieval, the extraction of a single image feature is often too limited, and it is difficult to achieve the purpose of efficient and fast retrieval. Color feature and directional gradient feature are two important features of an image, which are widely used in the field of image retrieval. Color information represents the overall features of the image, and the directional gradient feature represents the partial features information of the image by extracting the texture information of the image. Aiming at the problems of poor rotation change robustness, high feature dimension, and long retrieval time in current retrieval methods, a color image retrieval method that combines color feature with improved directional gradient feature is proposed. First, the input color image is converted into a grayscale image through Gaussian space, and the surface geometric curvature information and texture information of the grayscale image are extracted and integrated into the FHOG descriptor, and the main curvature information is multi-sampled to construct a mixed sampling directional gradient feature (P-FHOG1, P-FHOG2, P-FHOG3) based on the main curvature, and the improved directional gradient feature (P-FHOGs) based on the main curvature is obtained by merging the features of three scales. At the same time, the image is converted from RGB color space to HSV color space and the color information of the image is extracted after quantization to construct the color feature histogram, and the color feature of the image is obtained. On this basis, the two features are merged to obtain an image retrieval method based on color feature and improved directional gradient feature (CP-FHOG). The experiment was compared with the advanced image retrieval methods on the Corel-1000 and Coil-100 data sets, and the average accuracy rates of 85.89% and 93.38% were achieved, respectively. On the Corel-1000 data set, the features extraction time and retrieval time of the algorithm in this paper are 0.067 s and 0.048 s, respectively, which are improved by 0.075 s and 1.06 s, respectively, compared with the second-performing algorithm. At the same time, ablation experiments were performed in the two data sets to verify the effectiveness of the fusion algorithm. The experimental results show that, compared with HSV and P-FHOGs algorithms, CP-FHOG extracts richer detailed features, has stronger rotation robustness, and significantly improves retrieval accuracy in datasets containing complex backgrounds and targets with different rotation angles. Besides, retrieval time and feature dimension have also been greatly improved. The color image retrieval method proposed in this paper introduces main curvature information and color information based on FHOG descriptors, combines the advantages of color feature and directional gradient feature, and extracts rich overall and detailed features. The experimental result proves that the retrieval accuracy of the method in this paper is higher and the method has rotation robustness.

Jiang M, Zhang H X, Cheng D Q, *et al.* Multi-scale image retrieval based on HSV and directional gradient features[J]. *Opto-Electron Eng*, 2021, 48(11): 210310; DOI: 10.12086/oe.2021.210310

Foundation item: National Natural Science Foundation of China (51774281)

* E-mail: cdqcumt@126.com