

图像特征提取

- [图像内容识别流程](#)
- [图像特征提取](#)
 - [像素特征](#)
 - [纹理特征](#)
 - [共生矩阵描述子](#)
 - [局部二值模式](#)
 - [区域特征](#)
 - [梯度方向直方图](#)
 - [关键点特征](#)
 - [Harris角点检测](#)
 - [SIFT](#)

图像内容识别流程

1. 图像内容识别：（1）图像分类；（2）图像目标检测；（3）图像语义分割。
2. 主要阶段：感知 -> 预处理 -> 特征提取 -> 分类。

图像特征提取

1. 特征：待标记或区分的**明显属性或描述**。
 - 特征提取：
 - 特征检测：在图片、区域或边界中**发现特征**；
 - 特征描述：**将定量属性分配给检测到的特征**，使特征能够有效区分图像或图像集合。
 - 理想属性：
 - **紧密性**：无冗余信息、特征向量维度小；
 - **鲁棒性**：不受物体的简单形变、颜色变化及位置变化影响；
 - **区分性**：类内具有聚合性、类间具有区分性。
 - 表达方式：特征向量形式矩阵 $(1 \times n \text{ 或 } n \times 1)$ 。
 - 应用对象：**像素特征、纹理特征、区域特征、关键点特征**。
 - 特征范围：（1）局部特征；（2）全局特征。

像素特征

2. **像素特征**：（1）RGB色彩空间；（2）HSI色彩空间。

- RGB模型： $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ ， x_1 是某点红色图像的灰度值，其余分量是同点处绿色和蓝色图像的灰度值。
 - 应用：彩色显示器、彩色摄像头。
- HSI（色调、饱和度、亮度）：
 - 应用：灰度级处理技术。

纹理特征

3. 纹理特征：（1）共生矩阵描述子；（2）局部二值模式（LBPs）。

共生矩阵描述子

4. 共生矩阵描述子：

○ 共生矩阵：

设 Q 是定义两个像素相对位置的一个算子（如“右边的一个像素”），令 G 是一个矩阵，其元素 g_{ij} ，是图像 f 中灰度为 z_i, z_j 的像素对在 Q 规定的位置上出现的次数，其中 $i \leq i, j \leq L$ 。以这种方式形成的矩阵称为灰度（或亮度）公式矩阵。

○ 共生矩阵描述子：用来描述共生矩阵的统计特性。

设共生矩阵 $G = \{g_{ij}\} \in \mathbb{Z}^{K \times K}$ ，定义：

$$p_{ij} = \frac{g_{ij}}{\sum_{x=1}^K \sum_{y=1}^K g_{xy}}$$

则共生矩阵 G 的均匀性（uniformity）或能量为：

$$\sum_{i=1}^K \sum_{j=1}^K p_{ij}^2$$

熵值（entropy）为：

$$-\sum_{i=1}^K \sum_{j=1}^K p_{ij} \log_2 p_{ij}$$

局部二值模式

5. 局部二值模式（Local Binary Pattern, LBP）：

○ 主要思想：根据中心像素的灰度值对邻居像素的亮度进行**局部阈值化**，来形成一个二值模式。

像素纹理：对于一个中心像素，其纹理通过一个联合分布来描述：

$$T = t(g_c, g_0, g_1, \dots, g_{t-1})$$

其中 g_c 是中心像素的灰度值， $\{g_0, g_1, \dots, g_{t-1}\}$ 是相邻像素的灰度值。

亮度缩放差值不变性：

$$T = t(s(g_0 - g_c), s(g_1 - g_c), \dots, s(g_{t-1} - g_c))$$

$$\text{其中 } s(x) = \begin{cases} 1, & \text{when } x \geq 0 \\ 0, & \text{when } x < 0 \end{cases}$$

当错佐夫的元素按照排序后形成一个**包含零一值的循环链**时，则特定的方向就能够被一致地加权，从而形成一个**标量链状码**描述符，即：

$$LBP_{P,R} = \sum_{i=0}^{P-1} s(g_i - g_c) 2^i$$

其中 R 为局部圆形范围的半径， P 为该范围内选取的相邻灰度值的个数。

旋转不变性：对圆链码进行归一化，使得到的LBP值最小。

- 即不断对圆链逐位循环右移，各求其圆链码，直至计算出最小值。
- 在实际应用中，一般都不将LBP图谱作为特征向量用于分类识别，而是采用**LBP特征谱的统计直方图作为特征向量**用于分类识别。

区域特征

梯度方向直方图

6. 区域特征：梯度方向直方图（HOG）。

步骤：

1. 确定窗口、胞体、块大小/形状和重叠大小；
2. 全局光度归一化（ γ 校正）；
3. 计算方向直方图：

- 计算水平梯度 g_x 和垂直梯度 g_y ，则梯度强度 g 和方向 θ 为：

$$g = \sqrt{g_x^2 + g_y^2}, \quad \theta = \arctan \frac{g_x}{g_y}$$

4. 构建方向直方图：

- 对于胞体内的每个像素，其梯度强度用来增加它的梯度方向所对应的桶的大小。梯度强度对每个桶的贡献值可以用邻域桶的中心值进行线性或双线性的内插。
- 若是**无符号梯度**，则桶均匀地分布在 $(0^\circ, 180^\circ)$ 之间，间隔 20° 分为9个桶。

5. 对比度归一化；

6. 形成最终的HOG描述子。

关键点特征

7. 关键点特征：（1）Harris角点检测；（2）尺度不变特征变换（SIFT）。

Harris角点检测

8. Harris角点检测：

- 原理：在图像上移动一个窗口用于计算灰度变化：
 1. 在各个方向上零或小灰度变化的区域：检测器窗位于一个（几乎）恒定的区域中；
 2. 在某个方向上变化但在其正交方向上不变化的区域：检测器窗横跨两个区域的边界；
 3. 所有方向发生重大变化的区域：检测器窗包含一个角（或孤立点）。
- HS角检测器：即试图区分上述三个条件的数学公式。
- 计算方法：

用 $f(s, t)$ 表示一个“窗口”中心的灰度值，其移动了 (x, y) 后的窗口中心灰度值为 $f(s + x, t + y)$ 。因此，可通过两灰度值的差的平方的加权和来表示窗口移动的灰度变化：

$$C(x, y) = \sum_{s, t} w(s, t) [f(s + x, t + y) - f(s, t)]^2$$

$w(s, t)$ 为加权函数（可取 $w(s, t) = e^{-(s^2 + t^2)/2\sigma^2}$ ），表示窗口内各元素的权重。

利用泰勒级数展开的线性项，上式可以写为：

$$\begin{aligned}
C(x, y) &= \sum_{s, t} w(s, t) [x f_x(s, t) + y f_y(s, t)]^2 \\
&= \sum_{s, t} w(s, t) \cdot (x \ y) \begin{pmatrix} f_x \\ f_y \end{pmatrix} \cdot (f_x \ f_y) \begin{pmatrix} x \\ y \end{pmatrix} \\
&= (x \ y) \left[\sum_{s, t} w(s, t) \begin{pmatrix} f_x^2 & f_x f_y \\ f_x f_y & f_y^2 \end{pmatrix} \right] \begin{pmatrix} x \\ y \end{pmatrix}
\end{aligned}$$

令 $M = \sum_{s, t} w(s, t) A$, $A = \begin{pmatrix} f_x^2 & f_x f_y \\ f_x f_y & f_y^2 \end{pmatrix}$, 其中 M 称为**哈里斯矩阵**。

通过 M 的特征值, 即可完成前述三种情况的区分:

- **两个小特征值**: 灰度恒定;
 - **一个大特征值、一个小特征值**: 存在垂直边界或水平边界;
 - **两个大特征值**: 表示存在角或孤立点。
- **角响应测度**:

$$R = \lambda_x \lambda_y - k(\lambda_x + \lambda_y)^2 = \det(M) - k \cdot \text{trace}(M)。$$

- 当两个特征值都较小时, **测度 R 值较小**, 表示当前窗口处的图像是平坦的;
- 当一个特征值大、一个特征值小时, **测度 R 具有较大的负值**, 表示窗口跨界;
- 当两个特征值都较大时, **测度 R 具有较大的正值**, 表示存在角点。

其中, k 是一个经验常数, 表示HS检测器的**敏感因子**: k 越小, 检测器越可能找到角。

SIFT

9. 尺度不变特征变换 (SIFT) :

1. 构件尺寸空间:

- 定义: 一种多尺度表示, 以一致的方式处理不同尺度的图像结构;
- 思想: 无约束场景中的目标会以不同的方式出现, 具体方法取决于获取图像的尺度。由于事先无法确定尺度, 因此**一种合理的方法是同时处理所有相关的尺度, 实现尺度不变性**。

尺度空间将图像表示为平滑后图像的一个**参数簇**, 以**模逆图像尺寸减小时出现的细节损失**。

控制平滑的参数称为尺度参数。在SIFT中, 高斯核用于实现平滑, **因此尺度参数是其标准差**。

灰度图像 $f(x, y)$ 的**尺度空间** $L(x, t, \sigma)$ 是 f 与一个**可变尺度高斯核** $G(x, y, \sigma)$ 的卷积:

$$L(x, y, \sigma) = G(x, y, \sigma) * f(x, y), \quad G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-\frac{m}{2})^2 + (y-\frac{m}{2})^2}{2\sigma^2}}$$

2. 得到初始关键点:

查找与尺度空间中倍频程1的前两层相关的关键点位置, 可查找以下函数的极值:

$$D(x, y, \sigma) = [G(x, y, k\sigma) - G(x, y, \sigma)] * f(x, y) = L(x, y, k\sigma) - L(x, y, \sigma)$$

3. 提高关键点位置的精度;

4. 删除不合适的关键点;

5. 计算关键点方向:

- **根据局部图像性质为每个关键点分配一个一致的方向**, 以便能够表示相对于其方向的关键点, **实现图像旋转的不变性**。

6. 计算关键点描述子。

- 为了实现方向不变性，描述子的坐标和梯度方向相对于关键点方向进行了旋转。