

高级特征：

HOG：

- HOG特征提取算法的实现大致过程：HOG特征提取方法就是将一个image（你要检测的目标或者扫描窗口）：
 1. 灰度化（将图像看做一个x,y,z（灰度）的三维图像）；
 2. 采用Gamma校正法对输入图像进行颜色空间的标准化（归一化）；目的是调节图像的对比度，降低图像局部的阴影和光照变化所造成的影响，同时可以抑制噪音的干扰；
 3. 计算图像每个像素的梯度（包括大小和方向）；主要是为了捕获轮廓信息，同时进一步弱化光照的干扰。
 4. 将图像划分成小cells（例如6*6像素/cell）；
 5. 统计每个cell的梯度直方图（不同梯度的个数），即可形成每个cell的descriptor；
 6. 将每几个cell组成一个block（例如3*3个cell/block），一个block内所有cell的特征descriptor串联起来便得到该block的HOG特征descriptor。
 7. 将图像image内的所有block的HOG特征descriptor串联起来就可以得到该image（你要检测的目标）的HOG特征descriptor了。这个就是最终的可供分类使用的特征向量了。
- 代码实现：

python代码实现：

1、使用scikit-image库：

```
1 from skimage import feature as ft
2 img = cv2.imread('img.jpg', cv2.IMREAD_GRAYSCALE)
3 features = ft.hog(img,orientations=6,pixels_per_cell=[20,20],cells_per_block=[2,2],visualize=True)
4 plt.imshow(features[1],cmap=plt.cm.gray)
5 plt.show()
```

其中：

- 1、image:输入图像
 - 2、orientation: 指定块的个数。把所有的方向都转换为0°~180°内，然后按照orientation划分块，如果你选定的orientation= 6, 则bin一共有6个，每30°有一个。
 - 3、pixels_per_cell : 每个细胞单元的像素数，是一个tuple类型数据,例如(20,20)
 - 4、cell_per_block : 每个块内有多少个细胞单元, tuple类型, 例如(2,2), 意思是将block均匀划分为2x2的块。
- 源代码实现：https://github.com/coco-create/feature_extraction/blob/master/HOG_Extraction.py
 - 参考资料：<https://zhuanlan.zhihu.com/p/166146662>
https://blog.csdn.net/weixin_44791964/article/details/103549605

SIFT：尺度不变特征变换

• SIFT特征提取的实质

在不同的尺度空间上查找关键点(特征点)，并计算出关键点的方向。SIFT所查找到的关键点是一些十分突出、不会因光照、仿射变换和噪音等因素而变化的点，如角点、边缘点、暗区的亮点及亮区的暗点等。

• SIFT特征提取的方法

1. 构建DOG尺度空间:

模拟图像数据的多尺度特征, 大尺度抓住概貌特征, 小尺度注重细节特征。通过构建高斯金字塔(每一层用不同的参数 σ 做高斯模糊(加权)), 保证图像在任何尺度都能有对应的特征点, 即保证**尺度不变性**。

2. 关键点搜索和定位:

确定是否为关键点, 需要将该点与同尺度空间不同 σ 值的图像中的相邻点比较, 如果该点为max或min, 则为一个特征点。找到所有特征点后, 要去除低对比度和不稳定的边缘效应的点, 留下具有代表性的关键点(比如, 正方形旋转后变为菱形, 如果用边缘做识别, 4条边就完全不一样, 就会错误; 如果用角点识别, 则稳定一些)。去除这些点的好处是**增强匹配的抗噪能力和稳定性**。最后, 对离散的点做曲线拟合, 得到精确的关键点的位置和尺度信息。

3. 方向赋值:

为了实现**旋转不变性**, 需要根据检测到的关键点的局部图像结构为特征点赋值。具体做法是用梯度方向直方图。在计算直方图时, 每个加入直方图的采样点都使用圆形高斯函数进行加权处理, 也就是进行高斯平滑。这主要是因为SIFT算法只考虑了尺度和旋转不变性, 没有考虑**仿射不变性**。通过高斯平滑, 可以使关键点附近的梯度幅值有较大权重, 从而部分弥补没考虑仿射不变性产生的特征点不稳定。注意, 一个关键点可能具有多个关键方向, 这**有利于增强图像匹配的鲁棒性**。

4. 关键点描述子的生成:

关键点描述子不但包括关键点, 还包括关键点周围对其有贡献的像素点。这样可使关键点有**更多的不变特性, 提高目标匹配效率**。在描述子采样区域时, 需要考虑旋转后进行双线性插值, 防止因旋转图像出现白点。同时, 为了保证**旋转不变性**, 要以特征点为中心, 在附近领域内旋转 θ 角, 然后计算采样区域的梯度直方图, 形成n维SIFT特征矢量(如128-SIFT)。最后, 为了**去除光照变化的影响**, 需要对特征矢量进行归一化处理。

- 参考资料: <https://www.jianshu.com/p/03449c6b42c4>

<https://www.jianshu.com/p/e25562a87cca>

总结和理解: 采用一些方法(梯度直方图等)对于每帧图像获取一个特征向量(描述子), 根据图像间的描述子计算某种距离(欧式/余弦/汉明)即可比较两者的差异 或者是 采用SVM等分类方法进行分类

问题: 一开始考虑使用自变量(如f1分数、帧率、分辨率)来刻画video和video之间的overlap的函数关系具体形式, 但这些变量与视频内容本身没有太大关系, 用这些变量来刻画的效果并不好(与算法中使用的描述子没有直接的关系), 但是算法中使用的描述子是通过计算梯度等方法得到, 与视频内容直接相关, 但无法直接用函数关系去刻画