

阅读与思考

(1) SIFT 中的哪些设计有助于实现特征的：尺度不变性、光照不变性、旋转不变性。

尺寸不变性：

通过尺度空间方法 (Detection of scale-space extrema), 先计算尺度可变的高斯函数 $G(x, y, \sigma)$ 和输入图像 $I(x, y)$ 的卷积, 其公式为:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

之后, 高斯差分算子 DoG (Difference of Gaussians) 被运用, 其函数为 $D(x, y, \sigma)$ 。

DoG 与尺度归一化的高斯拉普拉斯函数 $\sigma^2 \nabla^2 G$ 非常近似。此函数结果可以用两个相邻的尺度 $k\sigma$ 和 σ 来计算差分近似, 得到高斯差分图像, 其公式为:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma)$$

之后, 为了检测 $D(x, y, \sigma)$ 的局部最大值和最小值, 每个采样点都会和他当前图像的 8 个邻近点以及上下尺度上的各 9 个邻近点进行比较, 比所有邻近点都大或都小的点才会被选择。

光照不变性：

在关键点被计算后, 我们需要一组向量描述关键点。此描述子不但包括关键点, 也包括关键点周围对其有贡献的像素点, 用来作为目标匹配的依据。之前的操作已经指定了图像的位置、尺度和每个关键点的方向, 下一步就是为局部图像区域计算描述子, 使其不随各种变化而改变, 比如光照变化、视角变化等。首先, 图像的梯度量级和方向是在关键点周围的采样, 使用关键点的尺度来选择图像高斯模糊的程度。为了达到方向不变性, 描述子的坐标和梯度方向会随关键点方向进行旋转的。

描述子在关键点尺度空间内的窗口中计算 8 个方向的梯度信息，得出向量表征。通过对关键点周围图像区域分块，计算块内梯度直方图，生成具有独特性的向量，这个向量是该区域图像信息的一种抽象。特征向量形成后，为了去除光照变化的影响，向量会做归一化处理，对于图像灰度值整体飘移，图像各点的梯度是领域像素相减得到，所以也能去除。

旋转不变性：

通过局部图像属性，每个关键点被指定一个方向，关键点描述子可以与这个方向相关，从而实现图像的旋转不变性。关键点的尺度是用来选择尺度最接近的高斯平滑图像 L ，这样所有的计算都是在一个尺度不变条件下进行的。对于每个图像样本 $L(x, y)$ ，在这个尺度下，梯度量级 $m(x, y)$ 和方向 $\theta(x, y)$ 是用像素差预计算出来。方向直方图会记录关键点区域窗口周围样本点的梯度方向，且直方图的峰值会确保关键点的方向。选取的窗口大小和关键点尺度成正比。这样提取的信息，就具有尺度不变性。图像的关键点检测已完毕，每个关键点有三个信息：位置、尺度、方向，同时也就使关键点具备了平移、缩放和旋转不变性。

(2) SIFT 兴趣点的亚像素精准定位如何实现？高精度的特征点定位有什么实用价值？

获得各个关键点的候选值后，对关键点数据，如位置、尺度和主曲率等进行细节匹配，这样低对比度（对噪声过敏）或定位在边缘的点会被淘汰。此外，局部样本点会经过 3D 二次方程式来决定最大值的插值位置，这样的处理会提高关键点的稳定性和匹配。通过对 $D(x, y, \sigma)$ 进行泰勒级数展开，把样本点作为原点。之

后，求 $D(x, y, \sigma)$ 进行泰勒级数展开后关于 x 的偏导为零，即得极值位置。由于 DoG 函数在图像边缘有较强的边缘响应，需要排除边缘响应。

高精度的特征点定位的用处为物体识别、机器人定位与导航、图像拼接、三维建模、手势识别、视频跟踪、笔记鉴定、指纹与人脸识别、犯罪现场特征提取等。