

一、SIFT、BOW、SVM 的基本理论

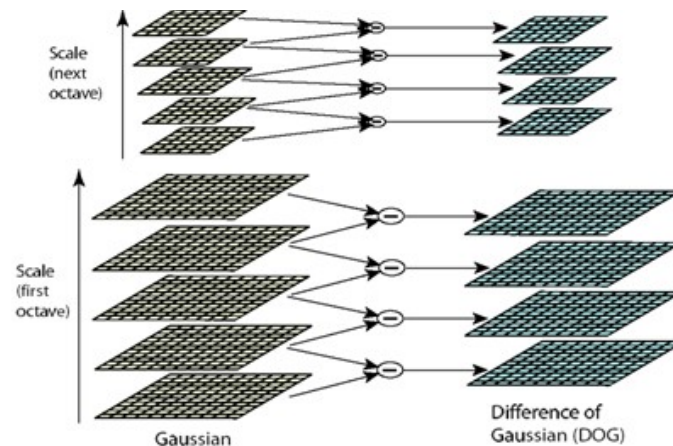
我没有看算法的具体公式只是大体了解并且整理了一下这三个部分的流程和基本概念。

1.SIFT 算法 Scale-invariant feature transform 尺度不变特征转换

检测与描述影像中的局部特征，他在空间尺度中寻找极值点，并提取出其位置、尺度、旋转不变量
整个算法分为以下几个部分：

*构建尺度空间

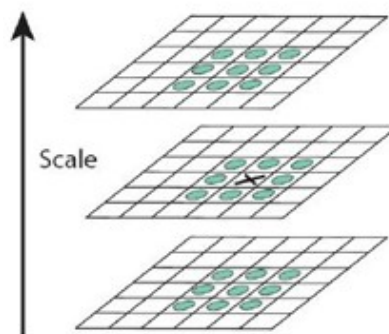
初始化操作，尺度空间理论的目的是模拟图像数据的多尺度特征。可以利用高斯卷积实现尺度变换，定义一幅二维图像的尺度空间。为了有效的在尺度空间检测到稳定的关键点，提出了高斯差分尺度空间（DOG scale-space）。利用不同尺度的高斯差分核与图像卷积生成。并且进行图像金字塔建立，得到其在不同尺度的图像。如图一所示。



图一

*LOG 近似 DOG 找到关键点<检测尺度空间极值点>

寻找尺度空间极值点，每个采样点要和它所有的相邻点比较，看其是否比它的图像域的相邻点大或者小。
如图 2 所示。当为最值时认为该点为图像在该尺度下的一个特征点。



图二

*去除不好的特征点

本质就是去除 DOG 局部曲率非常不对称的像素

*为每个关键点制定方向参数

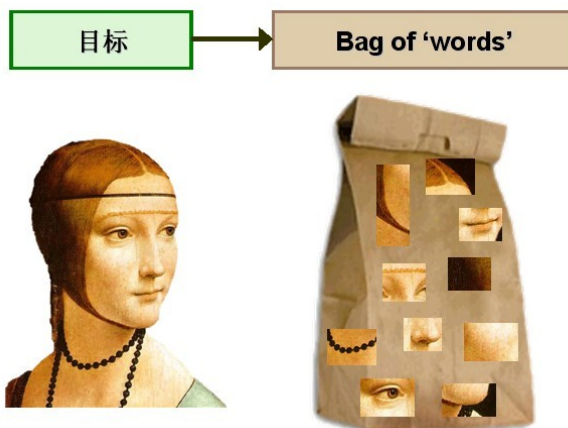
利用关键点邻域像素的梯度方向分布特征为每个关键点制定方向参数，使算子具备旋转不变性。至此，图像的关键点就检测完毕，每个关键点有三个信息：位置、所处尺度、方向。由此可以确定一个 sift 特征区域。

*关键点描述子的生成

将坐标轴旋转为关键点的方向，以确保旋转不变性。生成描述子，最终形成 128 维的 SIFT。最后进行归一化

2.BOW 模型 Bag-of-words model

最初词袋模型是信息检索领域常用的文档表示方法。后来，将词袋模型应用于图像表示。为了表示一幅图像，可以将图像看作文档，即若干个“视觉词汇”的集合，同样的，视觉词汇相互之间没有顺序。如图三所示。



图三

之后，从图像中提取出相互独立的视觉词汇，通常需要三个步骤：

(1) 特征检测 (2) 特征表示 (3) 单词本的生成

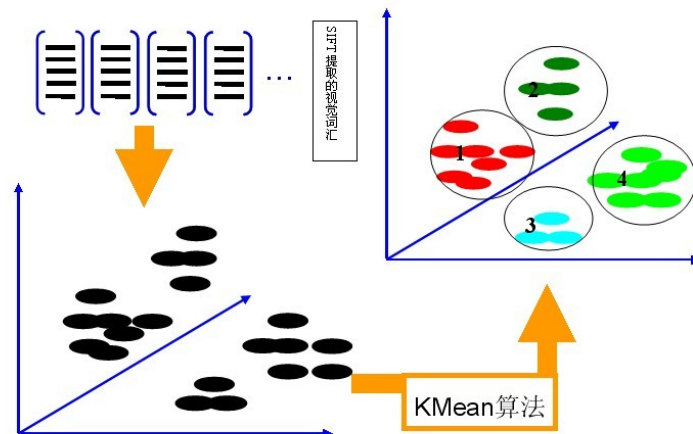
利用 Bag-of-words 模型将一幅图像表示成为数值向量：

前一步用到 SIFT 算法从图像中提取不变的特征点，作为视觉词汇，将所有的视觉词汇集合在一起。也就是先对每一幅图像划分 patch, 每一个 patch 用一个特征向量来表示，就是用 SIFT 来表示，所以每一个 patch 都是 128 维。如图四所示



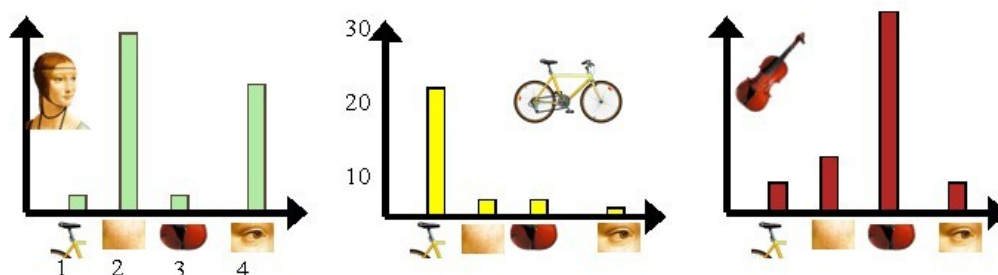
图四

接下来利用 K-Means 算法构造单词表。K-Means 算法是一种基于样本间相似性度量的间接聚类方法，此算法以 K 为参数，把 N 个对象分为 K 个簇，以使簇内具有较高的相似度，而簇间相似度较低。SIFT 提取的视觉词汇向量之间根据距离的远近，可以利用 K-Means 算法将词义相近的词汇合并，作为单词表中的基础词汇，假定我们将 K 设为 4，那么单词表的构造过程如下图五所示。



图五

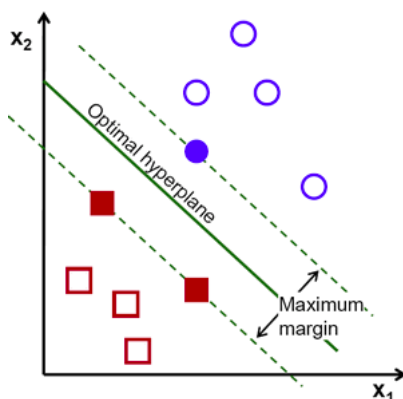
最后，统计单词表中每个单词在图像中出现的次数，从而将图像表示成为一个K维数值向量。也就是先初始化一个100个bin的初始值为0的直方图h。每一幅图像有很多patch，计算这些patch和每一个质心的距离，看看每一个patch离哪一个质心最近，那么直方图h中相对应的bin就加1，然后计算完这幅图像所有的patches之后，就得到了一个bin=100的直方图，然后进行归一化，用这个100维向量来表示这幅图像。对所有图像计算完成之后，就可以进行分类聚类训练预测之类的了。如图六所示。



图六

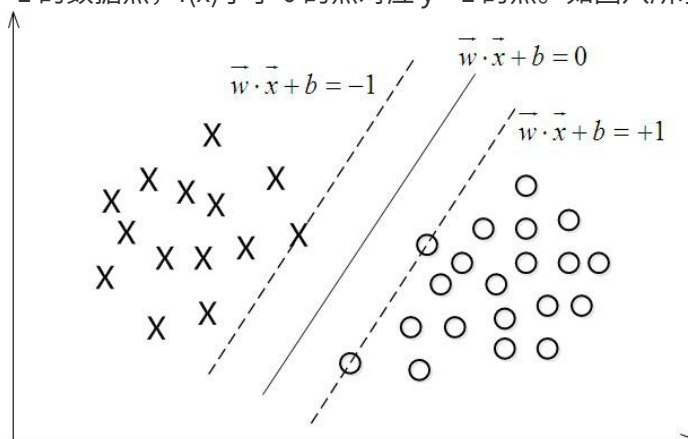
3.SVM 分类器 Support Vector Machine

svm 是一个由分类超平面定义的判别分类器，算法会输出一个最优超平面对新样本（测试样本）进行分类。所谓的超平面就是找到一条分割线，它要离所有的样本点都尽可能的远。如图七所示



图七

这个超平面可以用分类函数 $f(x) = w^T x + b$ 表示，当 $f(x)$ 等于 0 的时候， x 便是位于超平面上的点，而 $f(x)$ 大于 0 的点对应 $y=1$ 的数据点， $f(x)$ 小于 0 的点对应 $y=-1$ 的点。如图八所示



图八

确定这个超平面要用函数间隔 Function margin 与几何间隔 Geometrical margin，在超平面 $w^*x+b=0$ 确定的情况下， $|w^*x+b|$ 能够表示点 x 到距离超平面的远近，而通过观察 w^*x+b 的符号与类标记 y 的符号是否一致可判断分类是否正确，所以，可以用 $(y^*(w^*x+b))$ 的正负性来判定或表示分类的正确性。由此引

出函数间隔的概念，通过对法向量 w 加些约束条件，从而引出真正定义点到超平面的距离--几何间隔 (geometrical margin) 的概念。