1. 主要思想:

HOG(方向梯度直方图),在一幅图像中,局部目标的表象和形状能够被梯度方向密度分布很好的描述。

1. 主要流程
2. 灰度化
3. 采用Gmama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，同时可以抑制噪声的干扰
4. Sobel核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。
5. 将图像划分成小cells(例如8x8)。统计每个cell的梯度直方图(一般为180°划分为9个bins, 梯度值按落在方向区间内与两端点的距离成反比累加梯度值)。形成每个cell的描述子。
6. 将几个cell组成一个block(例如2\*2个cell/block), 一个block内所有的cell的特征串联起来就得到该block的特征描述子。
7. 将block按照一定的步距滑动得到所有block的描述子, 串联起来就是该扫描窗口的HOG特征。

在目标检测中，通常以固定形状不同尺度的窗口在整幅图像上滑动得到多个HOG特征描述子，将每个HOG描述子送入训练好的svm进行评分, 最后进行NMS抑制得到最合适的检测框。

优点： 算法比较简单直观，归一化抑制了光照颜色等影响、

缺点： 算法运行非常慢，不具备旋转不变特征，基本不能运用于视频上，遮挡对检测效果有很大的影响。

**图像预处理**

预处理包括**灰度化**和**Gamma变换**。

灰度处理是可选操作，因为灰度图像和彩色图像都可以用于计算梯度图。对于彩色图像，先对三通道颜色值分别计算梯度，然后取梯度值最大的那个作为该像素的梯度。

然后进行伽马矫正，调节图像对比度，**减少光照对图像的影响**（包括光照不均和局部阴影），使**过曝或者欠曝**的图像恢复正常，更接近人眼看到的图像。

**计算图像梯度**

为了得到梯度直方图，那么首先需要计算图像水平方向和垂直方向梯度。一般使用特定的卷积核对图像滤波实现，可选用的卷积模板有：soble算子、Prewitt算子、Roberts模板等等

**计算梯度直方图**

经过上一步计算，每一个像素点都会有两个值：**梯度幅值/梯度方向。**

在这一步中，图像被分成若干个8×8的cell，例如我们将图像resize至64x128的大小，那么这幅图像就被划分为8x16个8x8的cell单元，并为每个8×8的cell计算梯度直方图。当然，cell的划分也可以是其他值：16x16，8x16等，根据具体的场景确定。

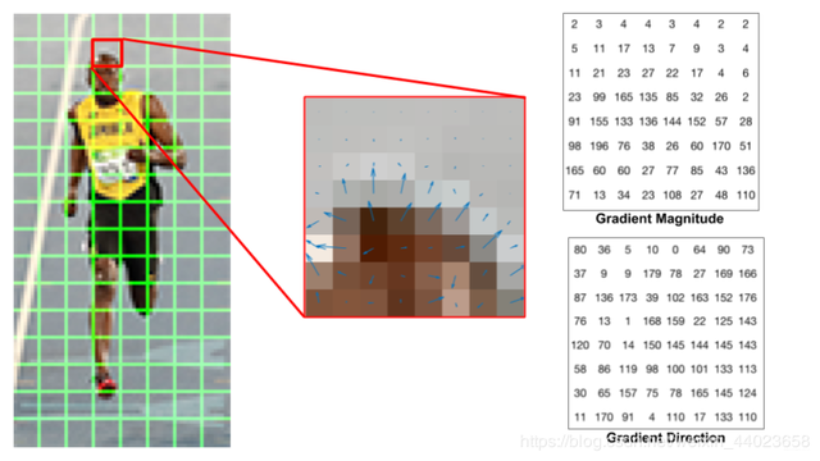
在计算梯度直方图，让我们先了解一下为什么我们将图像分成若干个cell？

这是因为如果对一整张梯度图逐像素计算，其中的**有效特征是非常稀疏**的，不但**运算量大**，而且会受到一些**噪声干扰。**于是我们就使用**局部特征描述符**来表示一个更紧凑的特征，计算这种局部cell上的梯度直方图更具鲁棒性。

以8x8的cell为例，一个8x8的cell包含了8x8x2 = 128个值，因为每个像素包括梯度的大小和方向。

在HOG中，每个8x8的cell的梯度直方图本质是一个由9个数值组成的向量， 对应于0、20、40、60…160的梯度方向(角度)。那么原本cell中8x8x2 = 128个值就由长度为9的向量来表示，用这种梯度直方图的表示方法，**大大降低了计算量**，同时又对光照等环境变化更加地鲁棒。

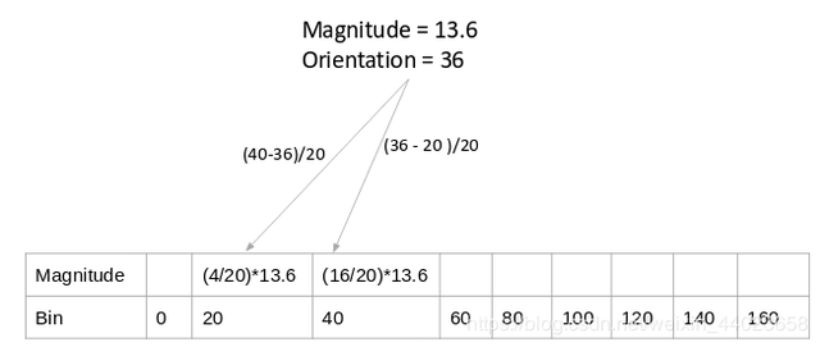
首先，看下图：



左图是一副64x128的图像，被划分为8x16个8x8的cell；中间的图像表示一个cell中的梯度矢量，箭头朝向代表梯度方向，箭头长度代表梯度大小。

右图是 8×8 的cell中表示梯度的原始数值，注意角度的范围介于**0到180度**之间，而不是0到360度， 这被称为“无符号”梯度，因为两个完全相反的方向被认为是相同的。

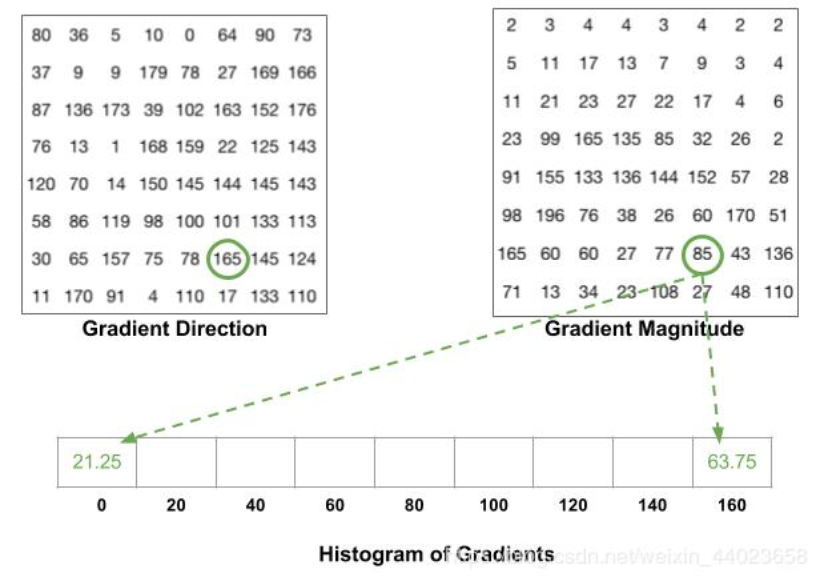
接下来，我们来计算cell中像素的梯度直方图，将0-180度分成9等份，称为9个bins，分别是0，20，40...160。然后对每个bin中梯度的贡献进行统计：



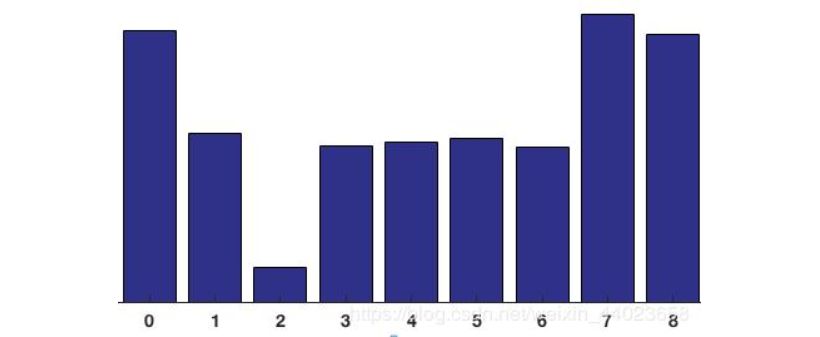
统计方法是一种**加权投票统计**，如上图所示，某像素的梯度幅值为13.6，方向为36，36度两侧的角度bin分别为20度和40度，那么就按一定加权比例分别在20度和40度对应的bin加上梯度值，加权公式为：

40度对应的bin：(（40-36）/20) \* 13.6，分母的20表示20等份，而不是20度；20度对应的bin：(（36-20）/20) \* 13.6，分母的20表示20等份，而不是20度；

还有一个细节需要注意，如果某个像素的梯度角度大于160度，也就是在160度到180度之间，那么把这个像素对应的梯度值按比例分给0度和160度对应的bin。如左下图绿色圆圈中的角度为165度，幅值为85，则按照同样的加权方式将85分别加到0度和160度对应的bin中。



对整个cell进行投票统计，正是在HOG特征描述子中创建直方图的方式，最终得到由9个数值组成的向量—梯度方向图：



Block 归一化

HOG特征将8×8的一个局部区域作为一个cell，再以2×2个cell作为一组，称为一个block，也就是说一个block表示16x16的区域。

我们可能会想，为什么又需要分block呢？

这是因为，虽然我们已经为图像的8×8单元创建了HOG特征，但是图像的梯度对整体光照很敏感。这意味着对于特定的图像，图像的某些部分与其他部分相比会非常明亮。

我们不能从图像中完全消除这个。但是我们可以通过使用16×16个块来对**梯度进行归一化来减少这种光照变化。**

由于每个cell有9个值，一个block（2×2个cell）则有36个值，HOG是通过滑动窗口的方式来得到block的.

前面已经说明，归一化的目的是为了**降低光照**的影响，因为梯度对整体光照非常敏感，比如通过将所有像素值除以2来使图像变暗，那么梯度幅值将减小一半，因此直方图中的值也将减小一半，我们就需要将直方图“归一化”。

归一化的方法有很多：L1-norm、L2-norm、max/min等等，一般选择**L2-norm**。

采用同样的方法，一个cell有一个梯度方向直方图，包含9个数值，一个block有4个cell，那么一个block就有4个梯度方向直方图，将这4个直方图拼接成长度为36的向量，然后对这个向量进行归一化。

而每一个block将按照上图滑动的方式进行重复计算，直到整个图像的block都计算完成。

获得HOG描述子

每一个16 \* 16大小的block将会得到一个长度为36的特征向量，并进行归一化。 那会得到多少个特征向量呢？

例如，对于上图被划分8 \* 16个cell ，每个block有2x2个cell的话，那么cell的个数为：(16-1)x(8-1)=105。即有7个水平block和15个竖直block。

每个block有36个值，整合所有block的特征值，最终获得由36 \* 105=3780个特征值组成的特征描述符，而这个特征描述符是一个一维的向量，长度为3780。

获得HOG特征向量，就可以用来可视化和分类了。对于多维的HOG特征，SVM就可以排上用场了。