## HOG特征

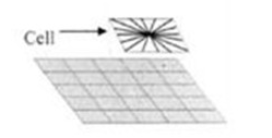
### Introduction

方向梯度直方图（Histogram of Oriented Gradient, HOG）特征最初是一种在计算机视觉和图像处理中用来进行物体检测的特征描述子，在05年的CVPR文章[1]中结合SVM完成行人检测任务时提出。后来被广泛由于各种视觉任务中，包括分类，认证等。

### Notations

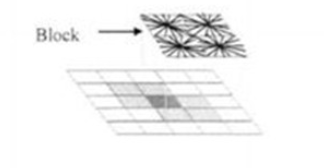
在正式介绍算法流程之前，首先需要了解一些概念：

* 像素(pixel)：一幅图像由一个个像素组成，通常说图像的大小就是指的这个像素矩阵的维度。如果一幅图像由RGB表示，每个像素点都可以三个数值表示，分别代表该像素点在RGB通道上的数值大小。图像的大小在HOG中经常就称为窗口的大小(m\*n)。
* 胞元（cell）：HOG算法里面将一个c\*c的像素矩阵称为一个胞元。经常使用6\*6，8\*8等大小的胞元，这个大小可以自己定义。如图：



下面图像每个方块代表一个像素。

* 块（block）：块可以看成胞元的矩阵，每个块由b\*b个胞元组成。



下面每个块代表一个胞元。

* 区间（Bin）：对于一个实数范围，在统计样本的直方图时得先定义每个区间的范围。由于这里是对每个像素点上的梯度进行加权统计，算法采用的策略是对梯度方向来划分区间，然后使用该点的像素大小来进行加权统计直方图。区间bin就是划分区间数，整个梯度方向的范围可以考虑[0,2\*pi],也可以只考虑[0,pi]。对于行人检测的时候采用无符号的往往可以取得较好效果，但也有可能有物体检测在有符号的时候效果更好。Bin的数目也会对效果产生影响，一般增加到9的时候效果就变化不大了。

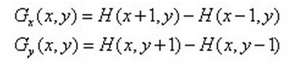
上面三者的关系可以看成是胞元由像素组成，块由胞元组成。下面再简单介绍算法里面用到的一些操作。

* Gamma变换和归一化：

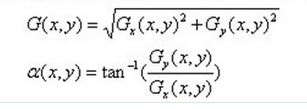
Gamma压缩公式：

http://img.my.csdn.net/uploads/201208/31/1346396938_1310.jpg比如可以取Gamma=1/2；首先需要将整个图像进行规范化（归一化）。这种压缩处理能够有效地降低图像局部的阴影和光照变化。因为颜色信息作用不大，通常先转化为灰度图

* 求像素的梯度：分别求x和y方向上的梯度值以及梯度方向：
  + x,y方向上的梯度值：



* + 梯度的大小以及梯度的方向可以分别表示为：

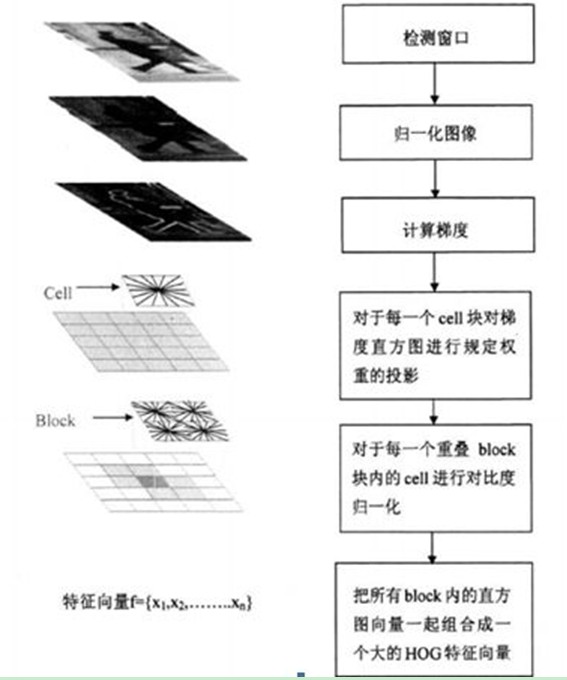


实现的时候，x方向的梯度值相当于用[-1,0,1]对原图像做一个卷积，y方向相当于用[-1,0,1]’对原图像做掩膜。也有地方计算梯度用[-1,1]掩膜算子,这个应该和应用场景有关。

* Overlapping：它的中文翻译是重叠，我的理解是它指的是在由像素构成胞元，由胞元构成块的过程中，胞元和块的重叠。类似于两个滑动窗口的公共部分。由像素构成胞元的时候没有重叠，也就是胞元的数目最后只和窗口大小以及胞元大小决定；而胞元构成块的时候就可能出现重叠，所以块的数目不仅和胞元总数目，每个块的胞元数有关，还和这个overlapping 的比例有关。

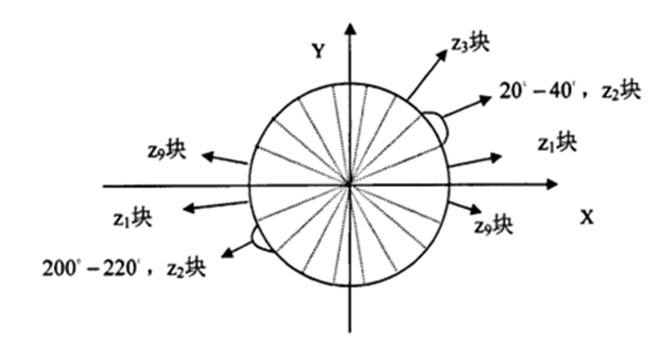
## 算法流程

算法的具体流程可用下面的图来表示：



每一步具体的实现可以参考源代码，里面的细节部分有些不一样，但并不影响结果的正确性。下面对每一步进行简单介绍。

1. 检测窗口：这个相当于输入一个图像进行尺度的resize处理；
2. 归一化图像：对图像每个像素进行伽马变换；
3. 计算梯度的方法在上面介绍了，可以直接应用；
4. 构造descriptor: 这一步是整个图像构造非线性特征的关键。可以包括图中的最后三步。
   1. 构造cell的梯度直方图特征：按照上面的介绍将图像划分为一个个cell，图像最终分的cell数只和图像大小（m\*n）和cell(c\*c) 大小有关：m\*n/(c\*c);对于每一个cell，根据bin统计它们的梯度方向的直方图分布，并根据梯度的大小进行加权统计。如果bin的数量是9，如下图所示划分区间，（这里是[0,pi]的梯度统计,如果是[0,2\*pi]的统计，每个区间为连续的40度角），

显然每个*胞元*的特征维数就是梯度方向划分的区间数（bin），这里就是9，每个维度的值就是梯度方向在这个区间上的像素对应的梯度值的带权求和。这样就为每个胞元构造了特征。

* 1. 构造block的特征：这一步主要是用来去掉光照、阴影等影响的，对于光照影响不剧烈的图像，例如很小区域内的字母，数字图像，可以不做这一步。而且论文中也提及了，这一步的对于最终分类准确率的影响也不大。通过胞元的特征来构造block的特征，每个block是通过多个胞元来构成的，block的特征通过直接串联相应的胞元特征构成。所以block的特征维度为（在上面的胞元上，假设每个block有2\*2个胞元构成）9\*2\*2= 36，每一列的数值即为每一个胞元的特征。最后可以将block特征化成向量形式。
  2. 构造图像的特征：图像的特征是通过block特征的串联构成的。Block的数目不仅跟胞元总数以及每个block中的胞元数有关，还和overlapping的系数有关，假设每个block overlap是参数是3个像素，原始图像为81\*81像素大小的，那么图像可以分为27\*27的胞元矩阵，这个胞元矩阵就可以划分为(（27-2）/1)\* (（27-2）/1)的block矩阵，最终得到625个block的特征，最后图像的特征就是625\*36=22500维。这里取的胞元数较小，所以得到的特征数较高，可以调整这些参数来改变特征维度。

在[4]中提到了patch就是本文中胞元(cell)的概念，但他的分割的大小是指将图像分成多少个胞元矩阵，下图分别是2\*2以及4\*4的分割。

## Reference

[1] Navneet Dalal and Bill Triggs，《Histograms of Oriented Gradients for Human Detection》，2005

[2] <http://blog.csdn.net/zouxy09/article/details/7929348>;

[3] <http://blog.csdn.net/abcjennifer/article/details/7365651>

[4] http://www.cnblogs.com/hrlnw/archive/2013/08/06/2826651.html

[5]<http://blog.csdn.net/liulina603/article/details/8291093>;

[6] www.cnblogs.com/tiandsp/archive/2013/05/24/3097503.html