

行人计数主要包括行人检测和行人追踪两部分，而行人检测是功能实现的第一步，因此这周我重新看了一遍《Monocular Pedestrian Detection Survey and Experiments》这篇 survey 以加深对行人检测的理解。这篇 survey 主要介绍了行人检测的一般过程以及主流的算法，最后比较了这些算法的性能、特点等。

行人检测的主要流程如下：

(1) 目标区域的选择 (ROI Selection)：其中一种做法是使用各种大小移动窗口在图片上的不同方向进行扫描。但这种方法的计算复杂度较高，不适用的实时处理的应用。另一种方法是使用一系列的特征表示图像，这些特征是从图像数据中抽取出来的。

(2) 分类 (Classification)：对于分类结果只有简单的两类，即行人与非行人，而分类结果取决于这两类的后验概率 (posterior probability)。分类的模型有两种，分别是 **Generative Model** 和 **Discriminative Model**。

这篇 survey 介绍的行人检测的方法有 **Haar Wavelet-Based Cascade**、**Neural Network Using Local Receptive Fields (NN/LRF)**、**Histograms of Oriented Gradients with Linear SVM (HOG/LinSVM)** 以及 **Combined Shape-Texture-Based Pedestrian Detection**。

最后经过实验，HOG 比其他方法取得的效果要好，基于这个原因我着重看了下面这篇关于 HOG 的论文，并打算采用这一方法来实现行人检测功能。

《Histograms of Oriented Gradients for Human Detection》

HOG 思想是：在一副图像中，局部目标的表象和形状 (appearance and shape) 能够被梯度或边缘的方向密度分布很好地描述。

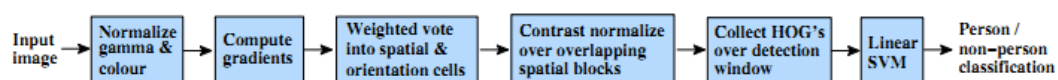
具体的实现方法是：首先将图像分成小的连通区域 (细胞单元)。然后采集细胞单元中各像素点的梯度的或边缘的方向直方图。最后把这些直方图组合起来就可以构成特征描述器。为了提高性能，还可以把这些局部直方图在图像的更大的范围内 (叫区间或 block) 进行对比度归一化 (contrast-normalized)，所采用的方法是：先计算各直方图在这个区间 (block) 中的密度，然后根据这个密度对区间中的各个细胞单元做归一化。通过这个归一化后，能对光照变化和阴影获得更好的效果。

与其他的特征描述方法相比，HOG 有很多优点：

(1) 由于 HOG 方法是在图像的局部细胞单元上操作，所以它对图像几何的 (geometric) 和光学的 (photometric) 形变都能保持很好的不变性，这两种形变只会出现在更大的空间领域上。

(2) 作者通过实验发现，在粗的空域抽样 (coarse spatial sampling)、精细的方向抽样 (fine orientation sampling) 以及较强的局部光学归一化 (strong local photometric normalization) 等条件下，只要行人大体上能够保持直立的姿势，就容许行人有一些细微的肢体动作，这些细微的动作可以被忽略而不影响检测效果。

HOG 方法的流程如下：



(1) 色彩和伽马归一化 (Colour/Gamma Normalization)

作者分别在灰度空间、RGB 色彩空间和 LAB 色彩空间上对图像进行色彩和伽马归一化，但实验结果显示，这个归一化的预处理工作对最后的结果没有影响，原因可能是：在后续步骤中也有归一化的过程，那些过程可以取代这个预处理的归一化。所以，在实际应用中，这一步可以省略。

(2) 梯度计算 (Gradient Computation)

最常用的方法是：简单地使用一个一维的离散微分模板在一个方向上或者同时在水平和垂直两个方向上对图像进行处理。作者也尝试其他更复杂的模板，但效果较差。因此作者的结论是：模板越简单，效果反而越好。

(3) Spatial/Orientation Binning: 为图像的每个细胞单元构建梯度方向直方图

细胞单元中的每一个像素点都为某个基于方向的直方图通道投票。投票是采取加权投票的方式，即每一票都是带权值的，这个权值是根据该像素点的梯度幅度计算出来。可以采用幅值本身或者它的函数来表示这个权值，实际测试表明：使用幅值来表示权值能获得最佳的效果，除此之外，也可以选择幅值的函数来表示，比如幅值的平方根 (square root)、幅值的平方 (square of the gradient magnitude)、幅值的截断形式 (clipped version of the magnitude) 等。

(4) Normalization and Descriptor Blocks: 把细胞单元组合成大的区间

由于局部光照的变化以及前景-背景对比度的变化，使得梯度强度的变化范围非常大。这就需要对梯度强度做归一化，作者采取的办法是：把各个细胞单元组合成大的、空间上连通的区间 (blocks)。这样一来，HOG 描述器就变成了由各区间所有细胞单元的直方图成分所组成的一个向量。这些区间是互有重叠的，这就意味着：每一个细胞单元的输出都多次作用于最终的描述器。

(5) Block Normalization: 区间归一化

作者采用了四种不同的方法对区间进行归一化，并对结果进行了比较。引入 v 表示一个还没有被归一化的向量，它包含了给定区间 (block) 的所有直方图信息。 $\|v\|_k$ 表示 v 的 k 阶范数，这里的 k 取 1、2。用 ε 表示一个很小的常数。这时，归一化因子可以表示如下：

$$\text{L2-norm: } v \rightarrow v / \sqrt{\|v\| + \varepsilon^2}$$

$$\text{L1-norm: } v \rightarrow v / (\|v\| + \varepsilon)$$

$$\text{L1-sqrt: } v \rightarrow \sqrt{v / (\|v\| + \varepsilon)}$$

还有第四种归一化方式：L2-Hys，它可以通过先进行 L2-norm，对结果进行截短 (clipping)，然后再重新归一化得到。作者发现：采用 L2-Hys、L2-norm 和 L1-sqrt 方式所取得的效果是一样的，L1-norm 稍微表现出一点点不可靠性。但是对于没有被归一化的数据来说，这四种方法都表现出来显著的改进。

(6) Classifier: 分类

作者使用 SVM 来分类，把提取的 HOG 特征输入到 SVM 分类器中，寻找一个最优超平面作为决策函数。