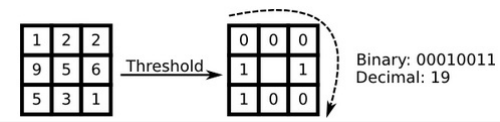


LBP算法

最初的LBP是定义在像素3x3邻域内的，以邻域中心像素为阈值，将相邻的8个像素的灰度值与其进行比较，若周围像素值大于中心像素值，则该像素点的位置被标记为1，否则为0。这样，3x3邻域内的8个点经比较可产生8位二进制



用比较正式的公式来定义的话：

$$LBP(x_c, y_c) = \sum_{p=0}^{P-1} 2^p s(i_p - i_c)$$

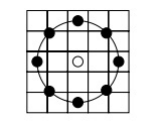
其中 (x_c, y_c) 代表3x3邻域的中心元素，它的像素值为 i_c ， i_p 代表邻域内其他像素的值。 $s(x)$ 是符号函数，定义如下：

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{else} \end{cases} \quad (1)$$

LBP的改进版本

(1) 圆形LBP算子

基本的 LBP算子的最大缺陷在于它只覆盖了一个固定半径范围内的小区域，这显然不能满足不同尺寸和频率纹理的需要。为了适应不同尺度的纹理特征，并达到灰度和旋转不变性的要求，Ojala等对 LBP 算子进行了改进，将 3×3邻域扩展为任意邻域，并用圆形邻域代替了正方形邻域，改进后的 LBP 算子允许在半径为 R 的圆形邻域内有任意多个像素点。从而得到了诸如



上图内有八个黑色的采样点，每个采样点的值可以通过下式计算：

$$x_p = x_c + R \cos(\frac{2\pi p}{P})$$
$$y_p = y_c - R \sin(\frac{2\pi p}{P})$$

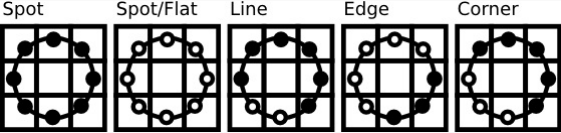
其中 (x_c, y_c) 为邻域中心点， (x_p, y_p) ， $p \in P$ 为某个采样点。通过上式可以计算任意个采样点的坐标，但是计算得到的坐标未必完全是整数，所以可以通过双线性插值来得到该采样点的像素值：

$$f(x, y) \approx \begin{bmatrix} 1-x & x \end{bmatrix} \begin{bmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1) \end{bmatrix} \begin{bmatrix} 1-y \\ y \end{bmatrix}.$$

(2) LBP等价模式

一个LBP算子可以产生不同的二进制模式，对于半径为R的圆形区域内含有P个采样点的LBP算子将会产生 2^P 种模式。很显然，随着邻域集内采样点数的增加，二进制模式的种类是急剧增加的。例如：5×5邻域内20个采样点，有 2^{20} 种二进制模式。如此多的二值模式无论对于纹理的提取还是对于纹理的识别、分类及信息的存取都是不利的。同时，过多的模式种类对于纹理的表达是不利的。例如，将LBP算子用于纹理分类或人脸识别时，常采用LBP模式的统计直方图来表达图像的信息，而较多的模式种类将使得数据量过大，且直方图过于稀疏。因此，需要对原始的LBP模式进行降维，使得数据量减少的情况下能最好的代表图像的信息。

为了解决二进制模式过多的问题，提高统计性，Ojala提出了采用一种“等价模式”（Uniform Pattern）来对LBP算子的模式种类进行降维。Ojala等认为，在实际图像中，绝大多数LBP模式最多只包含两次从1到0或从0到1的跳变，00000000（0次跳变），00000111（只含一次从0到1的跳变），10001111（先由1跳到0，再由0跳到1，共两次跳变）都是等价模式类。除等价模式类以外的模式都归为另一类，称为混合模式类，例如10010111（共四次跳变）。比如下图给出了几种等价模式的示意图。



通过这样的改进，二进制模式的种类大大减少，而不会丢失任何信息。模式数量由原来的 2^P 种减少为 $P(P-1)+2$ 种，其中P表示邻域集内的采样点数。对于3×3邻域内8个采样点来说，二进制模式由原始的256种减少为54种。通过上述方法，每个像素都会根据邻域信息得到一个LBP值，如果以图像的形式显示出来可以得到下图，明显LBP对光照有较强的鲁棒性。



来源： <<http://blog.csdn.net/pi9nc/article/details/26678691>>

LBP的改进版本：

原始的LBP提出后，研究人员不断对其提出了各种改进和优化。

(1) 圆形LBP算子：

基本的 LBP算子的最大缺陷在于它只覆盖了一个固定半径范围内的小区域，这显然不能满足不同尺寸和频率纹理的需要。为了适应不同尺度的纹理特征，并达到灰度和旋转不变性的要求，Ojala等对 LBP 算子进行了改进，将 3×3邻域扩展到任意邻域，并用圆形邻域代替了正方形邻域，改进后的 LBP 算子允许在半径为 R 的圆形邻域内有任意多个像素点。从而得到了诸如半径为R的圆形区域内含有P个采样点的LBP算子：

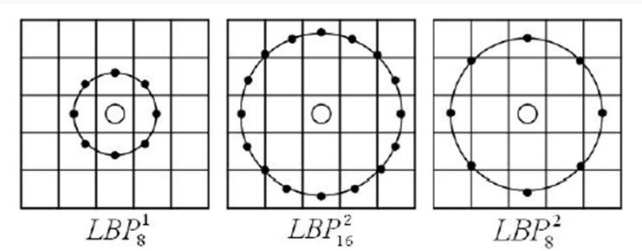


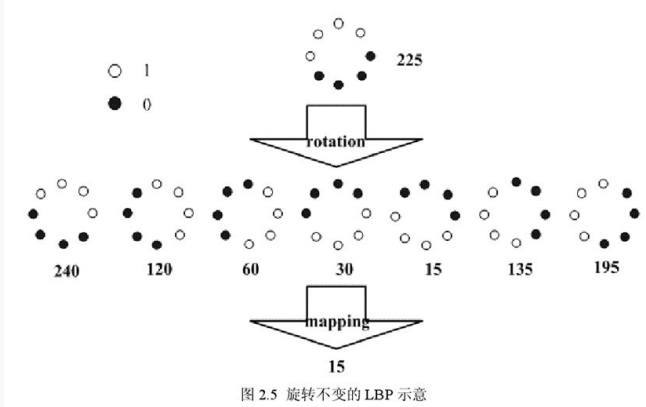
图2.2 几种LBP算子

(2) LBP旋转不变模式

从 LBP 的定义可以看出，LBP 算子是灰度不变的，但却不是旋转不变的。图像的旋转就会得到不同的 LBP 值。

Maenpaa等人又将 LBP算子进行了扩展，提出了具有旋转不变性的 LBP 算子，即不断旋转圆形邻域得到一系列初始定义的 LBP值，取其最小值作为该邻域的 LBP 值。

图 2.5 给出了求取旋转不变的 LBP 的过程示意图，图中算子下方的数字表示该算子对应的 LBP值，图中所示的 8 种 LBP 模式，经过旋转不变的处理，最终得到的具有旋转不变性的 LBP 值为 15。也就是说，图中的 8 种 LBP 模式对应的旋转不变的 LBP模式都是 00001111。



(3) LBP等价模式

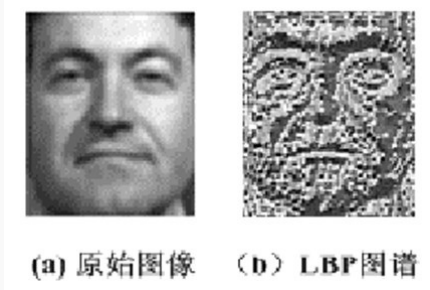
一个LBP算子可以产生不同的二进制模式，对于半径为R的圆形区域内含有P个采样点的LBP算子将会产生 P^2 种模式。很显然，随着邻域集内采样点数的增加，二进制模式的种类是急剧增加的。例如：5×5邻域内20个采样点，有 $2^{20}=1,048,576$ 种二进制模式。如此多的二值模式无论对于纹理的提取还是对于纹理的识别、分类及信息的存取都是不利的。同时，过多的模式种类对于纹理的表达是不利的。例如，将LBP算子用于纹理分类或人脸识别时，常采用LBP模式的统计直方图来表达图像的信息，而较多的模式种类将使得数据量过大，且直方图过于稀疏。因此，需要对原始的LBP模式进行降维，使得数据量减少的情况下能最好的代表图像的信息。

为了解决二进制模式过多的问题，提高统计性，Ojala提出了采用一种“等价模式”（Uniform Pattern）来对LBP算子的模式种类进行降维。Ojala等认为，在实际图像中，绝大多数LBP模式最多只包含两次从1到0或从0到1的跳变。因此，Ojala将“等价模式”定义为：当某个LBP所对应的循环二进制数从0到1或从1到0最多有两次跳变时，该LBP所对应的二进制就称为一个等价模式类。如00000000（0次跳变），00000111（只含一次从0到1的跳变），10001111（先由1跳到0，再由0跳到1，共两次跳变）都是等价模式类。除等价模式类以外的模式都归为另一类，称为混合模式类，例如10010111（共四次跳变）（这是我的个人理解，不知道对不对）。

通过这样的改进，二进制模式的种类大大减少，而不会丢失任何信息。模式数量由原来的 2^P 种减少为 $P (P-1)+2$ 种，其中P表示邻域集内的采样点数。对于3×3邻域内8个采样点来说，二进制模式由原始的256种减少为58种，这使得特征向量的维数更少，并且可以减少高频噪声带来的影响。

2、LBP特征用于检测的原理

显而易见的是，上述提取的LBP算子在每个像素点都可以得到一个LBP“编码”，那么，对一幅图像（记录的是每个像素点的灰度值）提取其原始的LBP算子之后，得到的原始LBP特征依然是“一幅图片”（记录的是每个像素点的LBP值）。



LBP的应用中，如纹理分类、人脸分析等，一般都不将LBP图谱作为特征向量用于分类识别，而是采用LBP特征谱的统计直方图作为特征向量用于分类识别。

因为，从上面的分析我们可以看出，这个“特征”跟位置信息是紧密相关的。直接对两幅图片提取这种“特征”，并进行判别分析的话，会因为“位置没有对准”而产生很大的误差。后来，研究人员发现，可以将一幅图片划分为若干的子区域，对每个子区域内的每个像素点都提取LBP特征，然后，在每个子区域内建立LBP特征的统计直方图。如此一来，每个子区域，就可以用一个统计直方图来进行描述；整个图片就由若干个统计直方图组成；

例如：一幅100*100像素大小的图片，划分为10*10=100个子区域（可以通过多种方式来划分区域），每个子区域的大小为10*10像素；在每个子区域内的每个像素点，提取其LBP特征，然后，建立统计直方图；这样，这幅图片就有10*10个子区域，也就有了10*10个统计直方图，利用这10*10个统计直方图，就可以描述这幅图片了。之后，我们利用各种相似性度量函数，就可以判断两幅图像之间的相似性了；

3、对LBP特征向量进行提取的步骤

- （1）首先将检测窗口划分为16×16的小区域（cell）；
- （2）对于每个cell中的一个像素，将相邻的8个像素的灰度值与其进行比较，若周围像素值大于中心像素值，则该像素点的位置被标记为1，否则为0。这样，3*3邻域内的8个点经比较可产生8位二进制数，即得到该窗口中心像素点的LBP值；
- （3）然后计算每个cell的直方图，即每个数字（假定是十进制数LBP值）出现的频率；然后对该直方图进行归一化处理。
- （4）最后将得到的每个cell的统计直方图进行连接成为一个特征向量，也就是整幅图的LBP纹理特征向量；

然后便可利用SVM或者其他机器学习算法进行分类了。

各类特征运行速度

特征	用时（毫秒）	特征数目
Sobel算子	5.99420	n/a
拉普拉斯算子	3.13440	n/a
Canny算子	3.41160	n/a
霍夫线变换	13.70790	10
霍夫圆变换	78.07720	0
Harris角点	9.41750	n/a
ShiTomasi角点	16.98390	18
亚像素级角点	3.63360	18
SURF角点	266.27000	151
Star关键点	14.82800	56
FAST角点	31.29670	159
SIFT角点	287.52310	54
		2
MSER区域	40.62970	

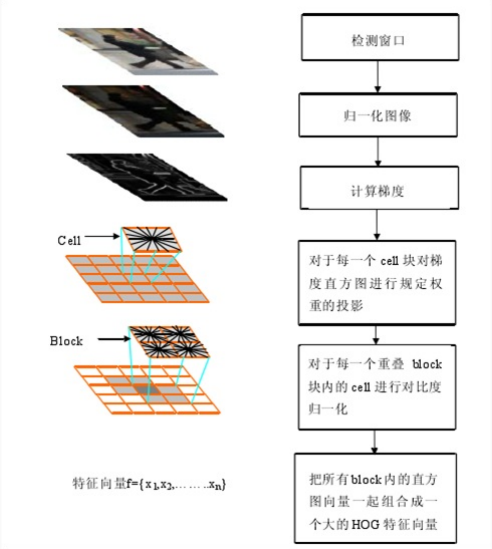
（图片尺寸：583x301，处理器：AMD ATHLON IIx2 240，内存：DDR3 4G，显卡：GeForce 9500GT，操作系统：Windows 7）

来源：<<http://www.cnblogs.com/xrwang/archive/2010/03/03/ImageFeatureDetection.html>>

特征提取HOG

1.介绍
HOG（Histogram of Oriented Gradient）是2005年CVPR会议上，法国国家计算机科学与自动控制研究所的Dalal等人提出的一种解决人体目标检测的图像描述子，该方法使用梯度方向直方图（Histogram of Oriented Gradients,简称HOG）特征来表达人体，提取人体的外形信息和运动信息，形成丰富的特征集。

2.生成过程



1) 图像归一化
归一化图像的主要目的是提高检测器对光照的鲁棒性，因为实际的人体目标可能出现各种不同的场合，检测器，必须对光照不太敏感才会有好的效果。

2)利用一阶微分计算图像梯度
图像平滑
对于灰度图像，一般为了去除噪点，所以会先利用离散高斯平滑模板进行平滑：高斯函数在不同平滑的尺度下进行对灰度图像进行平滑操作，Dalal等实验表明 在下，人体检测效果最佳（即不做高斯平滑），使得错误率缩小了约一倍。不做平滑操作，可能原因：图像时基于边缘的，平滑会降低边缘信息的对比度，从而减少 图像中的信号信息。

梯度法求图像梯度
一阶微分处理一般对灰度阶梯有较强的响应
一阶微分：

$$\frac{\partial f}{\partial x} = f(x+1) - f(x)$$

对于函数f(x,y)，在其坐标（x,y）上的梯度是通过如下二维列向量定义的：
这个向量的模值由下式给出：

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\delta f}{\delta x} \\ \frac{\delta f}{\delta y} \end{bmatrix}$$

$$|\nabla f| = || \nabla f ||_2 = \left[G_x^2 + G_y^2 \right]^{1/2} = \left[\left(\frac{\delta f}{\delta x} \right)^2 + \left(\frac{\delta f}{\delta y} \right)^2 \right]^{1/2}$$

因为模值的计算开销比较大，一般可以按如下公式近似求解：

$$|\nabla f| \approx |G_x| + |G_y|$$

Dalal等人利用许多一阶微分模板进行求梯度近似值，但在实验中表明模板[-1,0,1]效果最好。
采用模板[-1,0,1]为例计算图像梯度以及方向，通过梯度模板计算水平和垂直方向的梯度分别如下：

$$\begin{aligned} G_h(x,y) &= f(x+1,y) - f(x-1,y) \quad \forall x,y \\ G_v(x,y) &= f(x,y+1) - f(x,y-1) \quad \forall x,y \end{aligned}$$

其中，分别表示该像素点的水平，垂直梯度值。计算该像素点的梯度值（梯度强度）以及梯度方向：

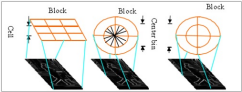
$$M(x,y)=\sqrt{G_h(x,y)^2+G_v(x,y)^2}+|G_h(x,y)|+|G_v(x,y)|$$
$$\theta(x,y)=\arctan(G_h(x,y)/G_v(x,y))$$

对于梯度方向的范围限定，一般采用无符号的范围，故梯度方向可表示为：

$$\theta(x,y)=\begin{cases} \theta(x,y)+\pi, & \theta(x,y)<0 \\ \theta(x,y), & others \end{cases}$$

3)基于梯度幅值的方向权重投影
HOG结构

通常使用的HOG结构大致有三种：矩形HOG（简称为R-HOG），圆形HOG和中心环绕HOG。它们的单位都是Block（即块）。Dalal的试验证明矩形HOG和圆形HOG的检测效果基本一致，而环绕形HOG效果相对差一些。



矩形HOG块的划分：

一般一个块（Block）都由若干单元（Cell）组成，一个单元都有若干个像素点组成。

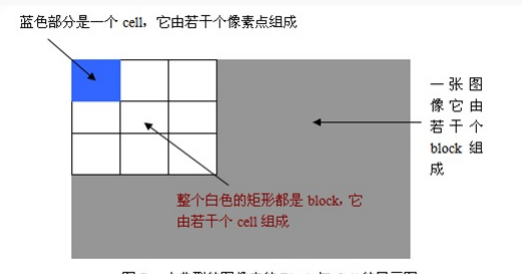
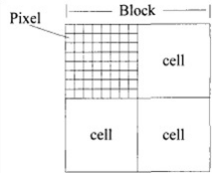
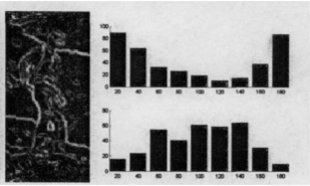


图 7 一个典型的图像中的 Block 与 Cell 的展示图

在每个Cell中有独立做梯度方向统计，从而以梯度方向为横轴的的直方图，前面我们已经提到过，梯度方向可取0度到180度或0度~360度，但 dalal实验表明，对于人体目标检测0度~180度这种忽略度数正负级的方向范围能够取得更好的结果。然后又将这个梯度分布平均分成 个方向角度（orientation bins），每个方向角度范围都会对应一个直方柱。



根据Dalal等人实验，在人体目标检测中，在无符号方向角度范围并将其平均分成9份（bins）能取得最好的效果，当bin的数目继续增大效果改变不明显，故一般在人体目标检测中使用bin数目为9范围0~180度的度量方式。

Block中各个参数的最终选取：

对于人体对象检测，块的大小为3×3个单元格，单元格的大小为6×6个像素时，检测效果是最好的，错误率约为10%左右。块的大小为2×2个单元格，单元格大小为8×8个像素时，也相差无几。6—8个像素宽的单元格，2—3个单元格宽的块，其错误率都在最低的一个平面上。块的尺寸太大时标准化的作用被削弱了从而导致错误率上升，而如果块的尺寸太小时，有用的信息反而会被过滤掉。

在实际应用中，在Block和Cell划分之后，对于得到各个像区域中，有时候还会为了进行一次高斯平滑，但是对于人体目标检测等问题，该步骤往往可以忽略，实际应用效果不大，估计在主要还是去除区域中噪点，因为梯度对于噪点相当敏感。

对梯度方向的投影权重方式的选取：

对于梯度方向的加权投影，一般都采用一个权重投影函数，它可以是像素点的梯度幅值，梯度幅值的平方根或梯度幅值的平方，甚至可以使梯度幅值的省略形式，它们都能够一定程度上反应出像素上一定的边缘信息。根据Dalal等人论文的测试结果，采用梯度幅值量级本身得到的检测效果最佳，使用量级的平方根会轻微降低检测结果，而使用二值的边缘权值表示会严重降低效果（约为5%个单位10~4FPPW（False Positives Per Window））。

4)HOG特征向量归一化

对block块内的HOG特征向量进行归一化。对block块内特征向量的归一化主要是为了使特征向量空间对光照，阴影和边缘变化具有鲁棒性。还有归一化是针对每一个block进行的，一般采用的归一化函数有以下四种：

- a) $L2-norm, v \leftarrow v / \sqrt{\|v\|_2^2 + \epsilon^2}$ (ϵ 是个很小的常值，避免分母为0)

b) $L2-Hys$, (方法同上，只是限制v的最大值到0.2，然后再次重新标准化)

c) $L1-norm, v \leftarrow v / (\|v\|_1 + \epsilon)$

d) $L1-sqrt, v \leftarrow \sqrt{v / (\|v\|_1 + \epsilon)}$

在人体检测系统中进行HOG计算时一般使用L2-norm，Dalal的文章也验证了对于人体检测系统使用L2-norm的时候效果最好。

5)得出HOG最终的特征向量

最终我可以通过以上步骤得到一个由 $\beta \times \varsigma \times \eta$ 个数据组成的高维向量，其中 β 表示每个 cell 中方向单元（bin）的数目， ς, η 分别表示 block 的个数以及一个 block 中 cell 的数目。至此 HOG 对于图像的描述向量生成。

3.HOG的应用：

主要用在object detection 领域，特别是行人检测，智能交通系统，当然也有文章提到把HOG用在手势识别，人脸识别等方面。

4.HOG与SIFT区别

HOG和SIFT都属于描述子，以及由于在具体操作上有很多相似的步骤，以致使很多人误认为HOG是SIFT的一种，其实两者在使用目的和具体处理细节上是有很大的区别的。HOG与SIFT的主要区别如下：

- ① SIFT是基于关键点特征向量的描述。

② HOG是将图像均匀的分成相邻的小块，然后在所有的小块内统计梯度直方图。

③ SIFT需要对图像尺度空间下对像素求极值点，而HOG中不需要。

④ SIFT一般有两大步骤，第一个步骤是对图像提取特征点，而HOG不会对图像提取特征点。

5.HOG的优点：

HOG表示的是边缘（梯度）的结构特征，因此可以描述局部的形状信息；

位置和方向空间的量化一定程度上可以抑制平移和旋转带来的影响；

采取在局部区域归一化直方图，可以部分抵消光照变化带来的影响。

由于一定程度忽略了光照颜色对图像造成的影响，使得图像所需要的表征数据的维度降低了。

而且由于它这种分块分元的处理方法，也使得图像局部像素点之间的关系可以很好得到的表征。

6.HOG的缺点：

描述子生成过程冗长，导致速度慢，实时性差；

很难处理遮挡问题。

由于梯度的性质，该描述子对噪点相当敏感

从 RGB 到 HSI 的彩色转换

给定一幅 RGB 彩色格式的图像，每个 RGB 像素的 H 分量可用下式得到：

$$H = \begin{cases} \theta, & B \leq G \\ 360 - \theta, & B > G \end{cases} \tag{6.2-2}$$

由 RGB 到 HSI 或从 HSI 到 RGB 的计算是逐个像素执行的。为表述清晰，我们省略了转换公式对 (x, y) 依赖性。

其中^①，

$$\theta = \arccos \left\{ \frac{\frac{1}{2} [(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \right\}$$

饱和度分量由下式给出：

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)]$$

(6.2-3)

最后，强度分量由下式给出：

$$I = \frac{1}{3} (R + G + B)$$

(6.2-4)

如图6.13中指出的那样，假定 RGB 值已归一化到区间[0, 1]内，且角度 θ 根据 HSI 空间的红轴来度量。色调可以用式(6.2-2)得到的所有值除以 360° 归一化为[0, 1]范围内。如果给定的 RGB 值在区间[0, 1]内，则其他两个 HSI 分量已经在区间[0, 1]内了。

式(6.2-2)到式(6.2-4)中的结果可由图6.12和图6.13所示的几何关系推得。这一推导很烦琐，而且对现在的讨论没有太大意义。对这些公式的证明及对下面的 HSI 至 RGB 转换结果感兴趣的读者，可以参考本书的参考文献或网站。

📖 关于 RGB 和 HSI 之间及 HSI 和 RGB 之间的转换公式的详细推导过程，请读者参阅本网站上的 Tutorials 部分。