梯度長條圖特徵(HOG) 是一種對圖像局部重疊區域的密集型描述符, 它通過計算局部區域的梯度方向長條圖來構成特徵。

HOG特徵是一種局部區域描述符,它通過計算局部區域上的梯度方向長條圖來構成人體特徵,能夠很好地描述人體的邊緣。它對光照變化和小量的偏移不敏感。

 Dalal提出的Hog特徵提取的過程：把樣本圖像分割為若干個圖元的單元（cell），把梯度方向平均劃分為9個區間（bin），在每個單元裡面對所有圖元的梯度方向在各個方向區間進行長條圖統計，得到一個9維的特徵向量，每相鄰的4個單元構成一個塊（block），把一個塊內的特徵向量聯起來得到36維的特徵向量，用塊對樣本圖像進行掃描，掃描步長為一個單元。最後將所有塊的特徵串聯起來，就得到了人體的特徵。例如，對於64\*128的圖像而言，每2\*2的單元（16\*16的圖元）構成一個塊，每個塊內有4\*9=36個特徵，以8個圖元為步長，那麼，水準方向將有7個掃描視窗，垂直方向將有15個掃描視窗。也就是說，64\*128的圖片，總共有36\*7\*15=3780個特徵。

在行人檢測過程中，除了上面提到的HOG特徵提取過程，還包括彩圖轉灰度，亮度校正等步驟。總結一下，在行人檢測中，HOG特徵計算的步驟：

（1）將輸入的彩圖轉換為灰度圖；

（2）採用Gamma校正法對輸入圖像進行顏色空間的標準化（歸一化）；   目的是調節圖像的對比度，降低圖像局部的陰影和光照變化所造成的影響，同時可以抑制噪音的干擾；

（3）計算梯度；主要是為了捕獲輪廓資訊，同時進一步弱化光照的干擾。

（4）將梯度投影到單元的梯度方向；目的是為局部圖像區域提供一個編碼，

（5）將所有儲存格在塊上進行歸一化；歸一化能夠更進一步對光照、陰影和邊緣進行壓縮，通常，每個儲存格由多個不同的塊共用，但它的歸一化是基於不同塊的，所以計算結果也不一樣。因此，一個儲存格的特徵會以不同的結果多次出現在最後的向量中。我們將歸一化之後的塊描述符就稱之為HOG描述符。

（6）收集得到檢測空間所有塊的HOG特徵；該步驟就是將檢測視窗中所有重疊的塊進行HOG特徵的收集，並將它們結合成最終的特徵向量供分類使用。

HOG即histogram of oriented gradient, 是用於目標檢測的特徵描述子，該技術將圖像局部出現的方向梯度次數進行計數，該方法和邊緣方向長條圖、scale-invariant feature transform類似，不同的是hog的計算基於一致空間的密度矩陣來提高準確率。[Navneet Dalal](http://en.wikipedia.org/w/index.php?title=Navneet_Dalal&action=edit&redlink=1" \o "Navneet Dalal (page does not exist)) and [Bill Triggs](http://en.wikipedia.org/w/index.php?title=Bill_Triggs&action=edit&redlink=1)首先在05年的CVPR中提出HOG，用於靜態圖像or視頻的行人檢測。

HOG特徵原理：

HOG的核心思想是所檢測的局部物體外形能夠被光強梯度或邊緣方向的分佈所描述。通過將整幅圖像分割成小的連接區域（稱為cells），每個cell生成一個方向梯度長條圖或者cell中pixel的邊緣方向，這些長條圖的組合可表示出（所檢測目標的目標）描述子。為改善準確率，局部長條圖可以通過計算圖像中一個較大區域(稱為block)的光強作為measure被對比標準化，然後用這個值(measure)歸一化這個block中的所有cells.這個歸一化過程完成了更好的照射/陰影不變性。

與其他描述子相比，HOG得到的描述子保持了幾何和光學轉化不變性（除非物體方向改變）。因此HOG描述子尤其適合人的檢測。

通俗的講：

HOG特徵提取方法就是將一個image：

1.            灰度化（將圖像看做一個x,y,z（灰度）的三維圖像）

2.            劃分成小cells（2\*2）

3.            計算每個cell中每個pixel的gradient（即orientation）

4.            統計每個cell的梯度長條圖（不同梯度的個數），即可形成每個cell的descriptor

再小談下Hog、SIFT與PCA-SIFT的應用與區別：

Hog沒有旋轉和尺度不變性，因此計算量小；而SIFT中每個feature需要用128維的向量來描述，因此計算量相對很大。

那麼行人檢測中怎麼應用HOG呢？

對於解決Scale-invariant 的問題：將圖片進行不同尺度的縮放，就相當於對範本進行不同尺度scale的縮放

對於解決Rotation-invariant 的問題：建立不同方向的模版（一般取15\*7的）進行匹配

總的來說，就是在不同尺度上的圖像進行不同方向的範本（15\*7）匹配，每個點形成一個8方向的梯度描述。

SIFT由於其龐大計算量不用與行人檢測，而PCA-SIFT的方法過濾掉很多維度的資訊，只保留20個主分量，因此只適用于行為變化不大的物體檢測。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| method | Time | Scale | Rotation | Blur | Illumination | Affine |
| Sift | common | best | best | common | common | good |
| PCA-sift | good | good | good | best | good | best |
| Surf | best | common | common | good | best | good |

1、HOG特徵：

       方向梯度直方圖（Histogram of Oriented Gradient, HOG）特徵是一種在計算機視覺和圖像處理中用來進行物體檢測的特徵描述子。它通過計算和統計圖像局部區域的梯度方向直方圖來構成特徵。Hog特徵結合SVM分類器已經被廣泛應用於圖像識別中，尤其在行人檢測中獲得了極大的成功。需要提醒的是，HOG+SVM進行行人檢測的方法是法國研究人員Dalal在2005的CVPR上提出的，而如今雖然有很多行人檢測算法不斷提出，但基本都是以HOG+SVM的思路為主。

（1）主要思想：

       在一副圖像中，局部目標的表象和形狀（appearance and shape）能夠被梯度或邊緣的方向密度分佈很好地描述。（本質：梯度的統計信息，而梯度主要存在於邊緣的地方）。

（2）具體的實現方法是：

       首先將圖像分成小的連通區域，我們把它叫細胞單元。然後採集細胞單元中各像素點的梯度的或邊緣的方向直方圖。最後把這些直方圖組合起來就可以構成特徵描述器。

（3）提高性能：

       把這些局部直方圖在圖像的更大的範圍內（我們把它叫區間或block）進行對比度歸一化（contrast-normalized），改採用的方法是：先計算各直方圖在這個區間（block）中的密度，然後根據這個密度對區間中的各個細胞單元做歸一化。通過這個歸一化後，能對光照變化和陰影獲得更好的效果。

（4）優點：

       與其他的特徵描述方法相比，HOG有很多優點。首先，由於HOG是在圖像的局部方格單元上操作，所以它對圖像幾何的和光學的形變都能保持很好的不變性，這兩種形變只會出現在更大的空間領域上。其次，在粗的空域抽樣、精細的方向抽樣以及較強的局部光學歸一化等條件下，只要行人大體上能夠保持直立的姿勢，可以容許行人有一些細微的肢體動作，這些細微的動作可以被忽略而不影響檢測效果。因此HOG特徵是特別適合於做圖像中的人體檢測的。

2、HOG特徵提取算法的實現過程：

大概過程：

HOG特徵提取方法就是將一個image（你要檢測的目標或者掃瞄窗口）：

1）灰度化（將圖像看做一個x,y,z（灰度）的三維圖像）；

2）採用Gamma校正法對輸入圖像進行顏色空間的標準化（歸一化）；目的是調節圖像的對比度，降低圖像局部的陰影和光照變化所造成的影響，同時可以抑制噪音的干擾；

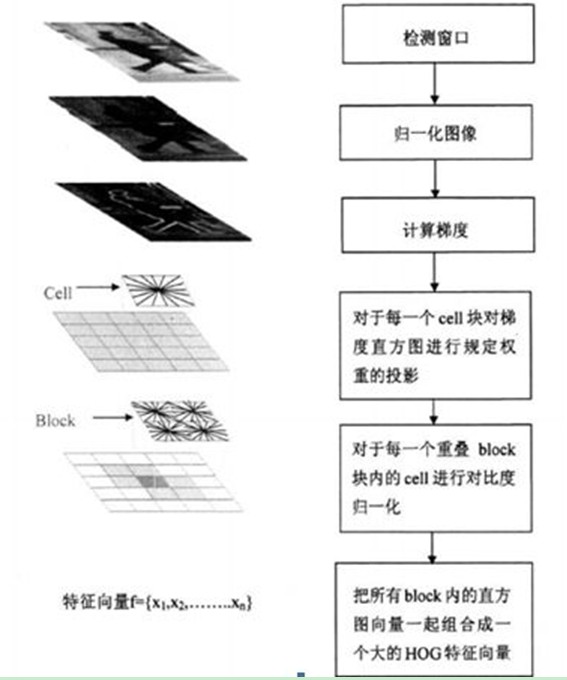
3）計算圖像每個像素的梯度（包括大小和方向）；主要是為了捕獲輪廓信息，同時進一步弱化光照的干擾。

4）將圖像劃分成小cells（例如6\*6像素/cell）；

5）統計每個cell的梯度直方圖（不同梯度的個數），即可形成每個cell的descriptor；

6）將每幾個cell組成一個block（例如3\*3個cell/block），一個block內所有cell的特徵descriptor串聯起來便得到該block的HOG特徵descriptor。

7）將圖像image內的所有block的HOG特徵descriptor串聯起來就可以得到該image（你要檢測的目標）的HOG特徵descriptor了。這個就是最終的可供分類使用的特徵向量了。



具體每一步的詳細過程如下：

（1）標準化gamma空間和顏色空間

     為了減少光照因素的影響，首先需要將整個圖像進行規範化（歸一化）。在圖像的紋理強度中，局部的表層曝光貢獻的比重較大，所以，這種壓縮處理能夠有效地降低圖像局部的陰影和光照變化。因為顏色信息作用不大，通常先轉化為灰度圖；

     Gamma壓縮公式：

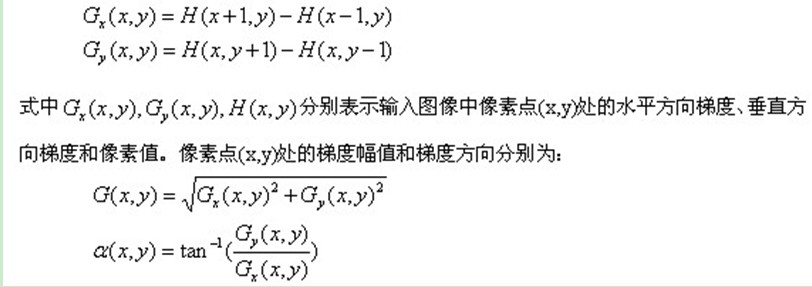
http://img.my.csdn.net/uploads/201208/31/1346396938_1310.jpg

     比如可以取Gamma=1/2；

（2）計算圖像梯度

        計算圖像橫坐標和縱坐標方向的梯度，並據此計算每個像素位置的梯度方向值；求導操作不僅能夠捕獲輪廓，人影和一些紋理信息，還能進一步弱化光照的影響。

圖像中像素點(x,y)的梯度為：



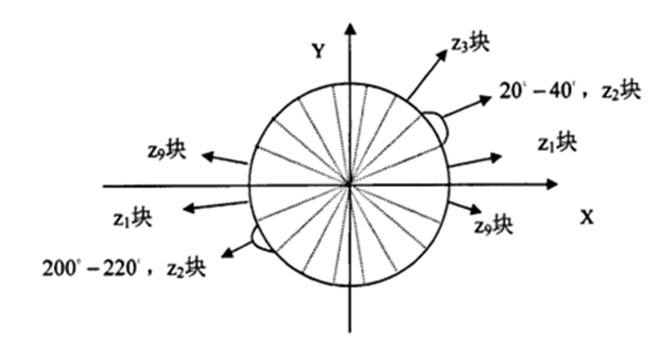
       最常用的方法是：首先用[-1,0,1]梯度算子對原圖像做卷積運算，得到x方向（水平方向，以向右為正方向）的梯度份量gradscalx，然後用[1,0,-1]T梯度算子對原圖像做卷積運算，得到y方向（豎直方向，以向上為正方向）的梯度份量gradscaly。然後再用以上公式計算該像素點的梯度大小和方向。

（3）為每個細胞單元構建梯度方向直方圖

        第三步的目的是為局部圖像區域提供一個編碼，同時能夠保持對圖像中人體對象的姿勢和外觀的弱敏感性。

我們將圖像分成若干個「單元格cell」，例如每個cell為6\*6個像素。假設我們採用9個bin的直方圖來統計這6\*6個像素的梯度信息。也就是將cell的梯度方向360度分成9個方向塊，如圖所示：例如：如果這個像素的梯度方向是20-40度，直方圖第2個bin的計數就加一，這樣，對cell內每個像素用梯度方向在直方圖中進行加權投影（映射到固定的角度範圍），就可以得到這個cell的梯度方向直方圖了，就是該cell對應的9維特徵向量（因為有9個bin）。

        像素梯度方向用到了，那麼梯度大小呢？梯度大小就是作為投影的權值的。例如說：這個像素的梯度方向是20-40度，然後它的梯度大小是2（假設啊），那麼直方圖第2個bin的計數就不是加一了，而是加二（假設啊）。

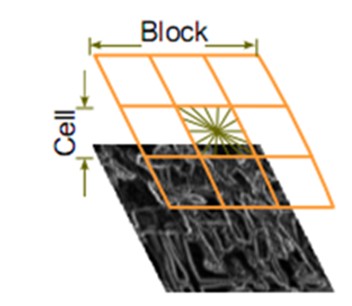


         細胞單元可以是矩形的（rectangular），也可以是星形的（radial）。

（4）把細胞單元組合成大的塊（block），塊內歸一化梯度直方圖

       由於局部光照的變化以及前景-背景對比度的變化，使得梯度強度的變化範圍非常大。這就需要對梯度強度做歸一化。歸一化能夠進一步地對光照、陰影和邊緣進行壓縮。

        作者採取的辦法是：把各個細胞單元組合成大的、空間上連通的區間（blocks）。這樣，一個block內所有cell的特徵向量串聯起來便得到該block的HOG特徵。這些區間是互有重疊的，這就意味著：每一個單元格的特徵會以不同的結果多次出現在最後的特徵向量中。我們將歸一化之後的塊描述符（向量）就稱之為HOG描述符。



        區間有兩個主要的幾何形狀——矩形區間（R-HOG）和環形區間（C-HOG）。R-HOG區間大體上是一些方形的格子，它可以有三個參數來表徵：每個區間中細胞單元的數目、每個細胞單元中像素點的數目、每個細胞的直方圖通道數目。

       例如：行人檢測的最佳參數設置是：3×3細胞/區間、6×6像素/細胞、9個直方圖通道。則一塊的特徵數為：3\*3\*9；

（5）收集HOG特徵

      最後一步就是將檢測窗口中所有重疊的塊進行HOG特徵的收集，並將它們結合成最終的特徵向量供分類使用。

（6）那麼一個圖像的HOG特徵維數是多少呢？

        順便做個總結：Dalal提出的Hog特徵提取的過程：把樣本圖像分割為若干個像素的單元（cell），把梯度方向平均劃分為9個區間（bin），在每個單元裡面對所有像素的梯度方向在各個方向區間進行直方圖統計，得到一個9維的特徵向量，每相鄰的4個單元構成一個塊（block），把一個塊內的特徵向量聯起來得到36維的特徵向量，用塊對樣本圖像進行掃瞄，掃瞄步長為一個單元。最後將所有塊的特徵串聯起來，就得到了人體的特徵。例如，對於64\*128的圖像而言，每16\*16的像素組成一個cell，每2\*2個cell組成一個塊，因為每個cell有9個特徵，所以每個塊內有4\*9=36個特徵，以8個像素為步長，那麼，水平方向將有7個掃瞄窗口，垂直方向將有15個掃瞄窗口。也就是說，64\*128的圖片，總共有36\*7\*15=3780個特徵。