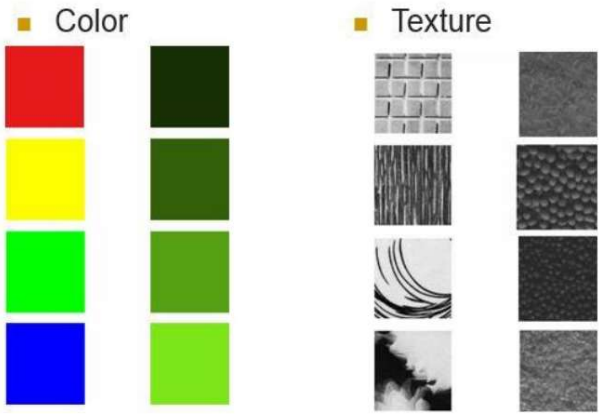
Data Mining for Visual Data

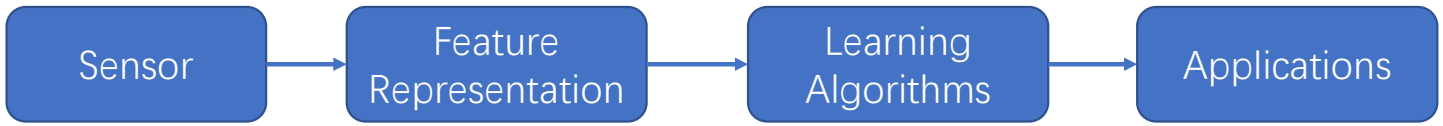
Outline

- Introduction
- Color Histogram
- LBP
- HOG

Introduction

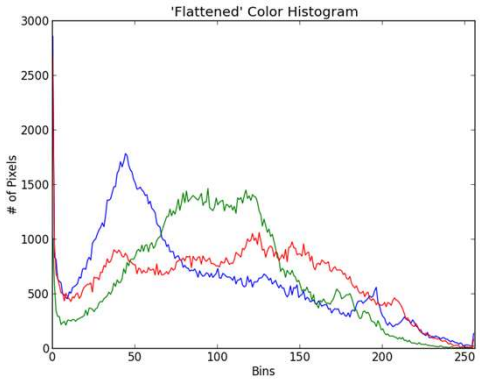


Supervised:
Classification (Recognition) or Regression/Prediction
Unsupervised:
Clustering or Grouping

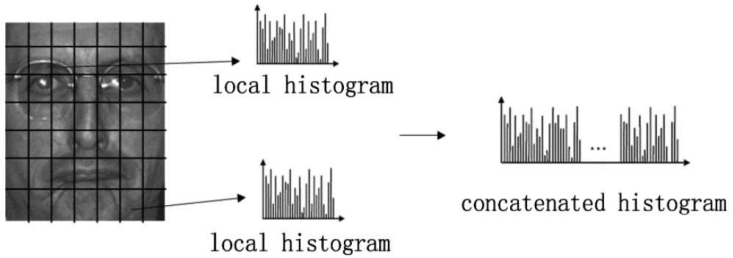


e.g., Face or Human Detectors

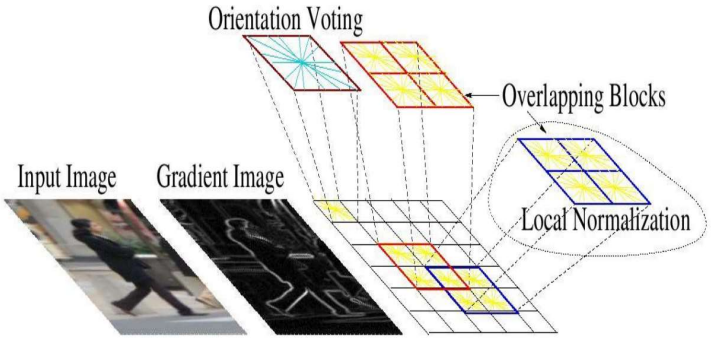
Common Visual Features



Color Histogram



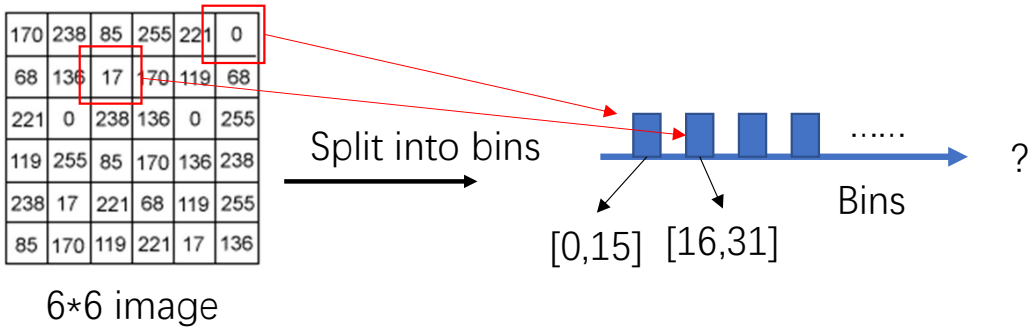
LBP



HOG

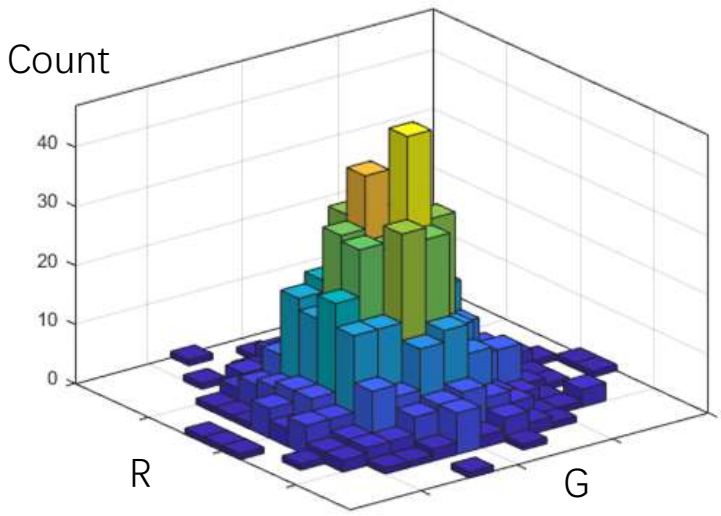
Color Histogram

- Gray-scale image: [0,255]
  - Uniformly divide the range into bins, e.g., 16 bins
  - Count how many pixels for each bin
  - Normalize



Color Histogram

- How about RGB image?
  - 1. Get histogram of each channel then concatenate.
    - What is the drawback?
  - 2. Split bins based on three channels.
    - What is the drawback?



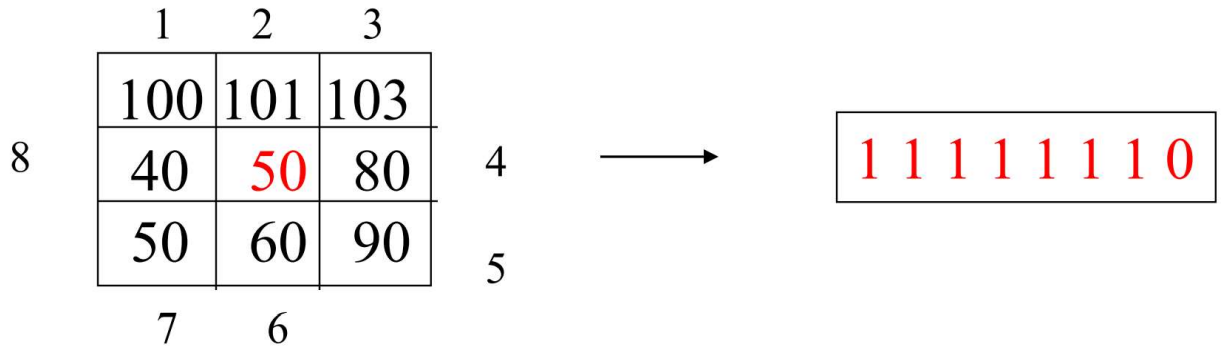
Take R and G channels for example

Color Histogram

- Spatial invariant
  - Good or not good?
  - Can we make it spatial variant?
- No texture information
  - Can be used to identify different people?
  - Can be used to identify different races (e.g, yellow, white, black)?

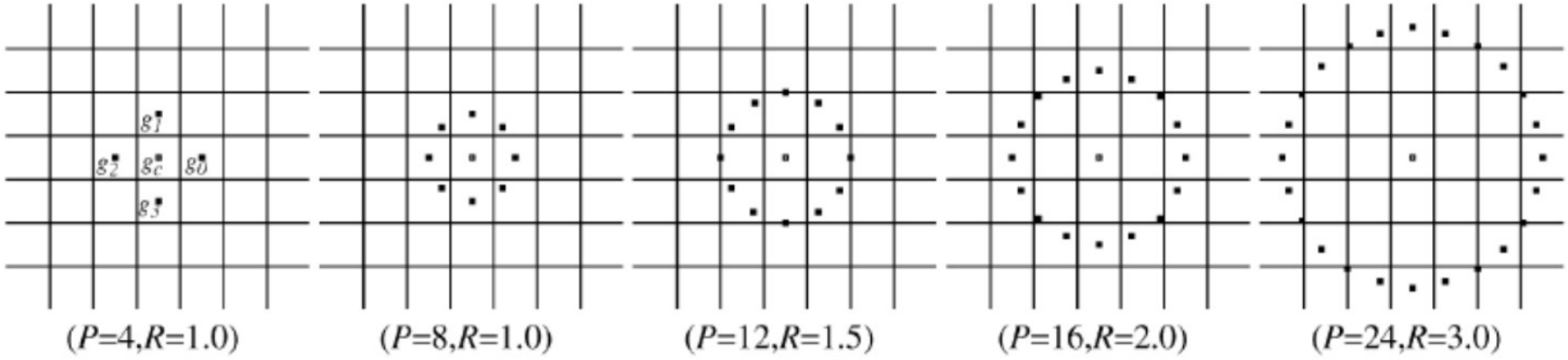
Local Binary Pattern (LBP)

- For each pixel p, create an 8-bit number [b1 b2 b3 b4 b5 b6 b7 b8], where bi = 0 if neighbor i has value less than p's value and 1 otherwise.
- Represent the texture in the image (or a region) by the histogram of these numbers.

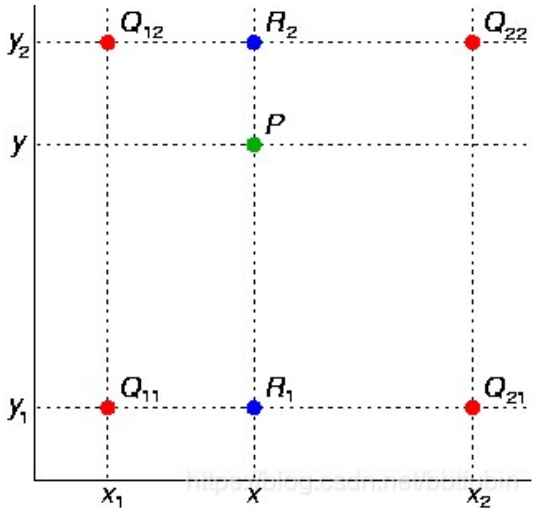


Variations of LBP

- Multi-scale LBP

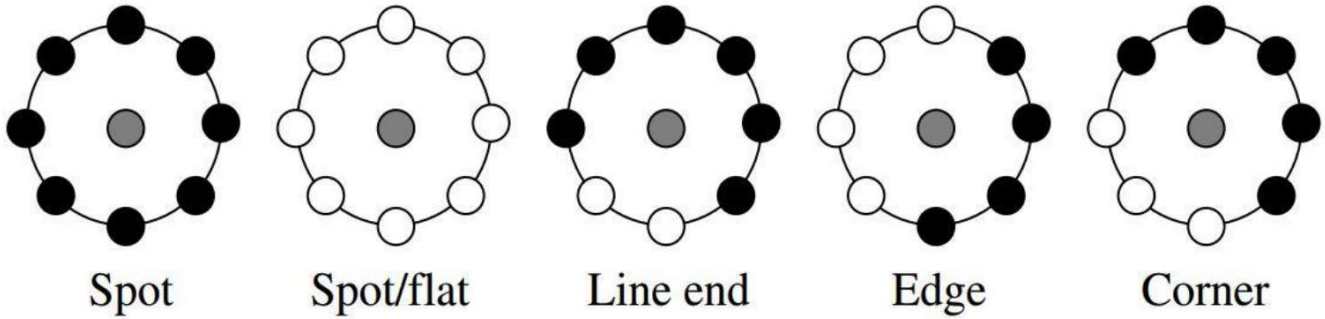


Bilinear Interpolation

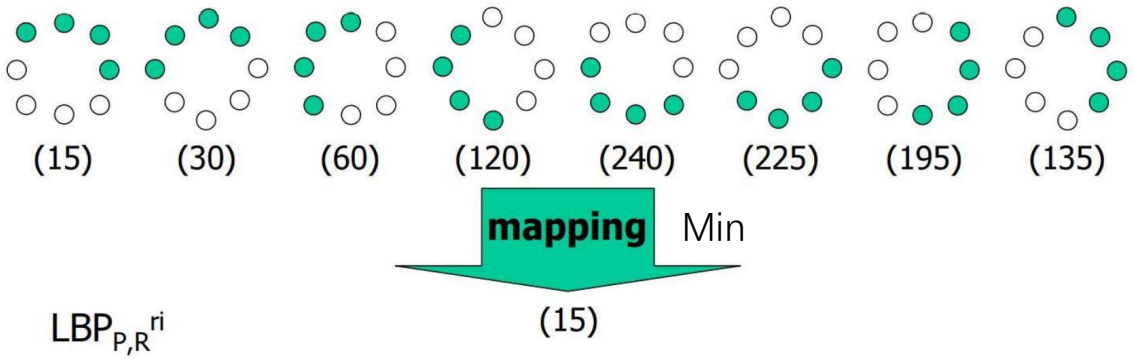


f(P) $$\frac{1}{(x_2-x_1)(y_2-y_1)}\left(f(Q_{11})(x_2-x)(y_2-y)+f(Q_{21})(x-x_1)(y_2-y)+f(Q_{12})(x_2-x)(y-y_1)+f(Q_{22})(x-x_1)(y-y_1)\right)$$

Variations of LBP

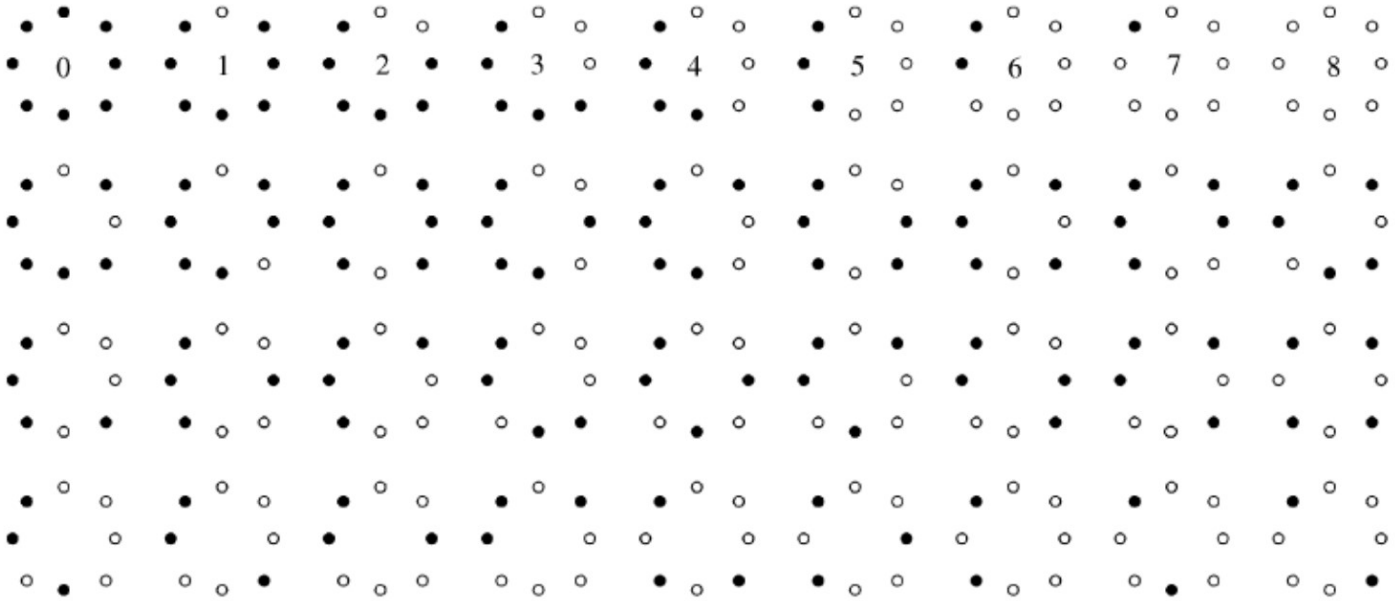


Rotation invariance

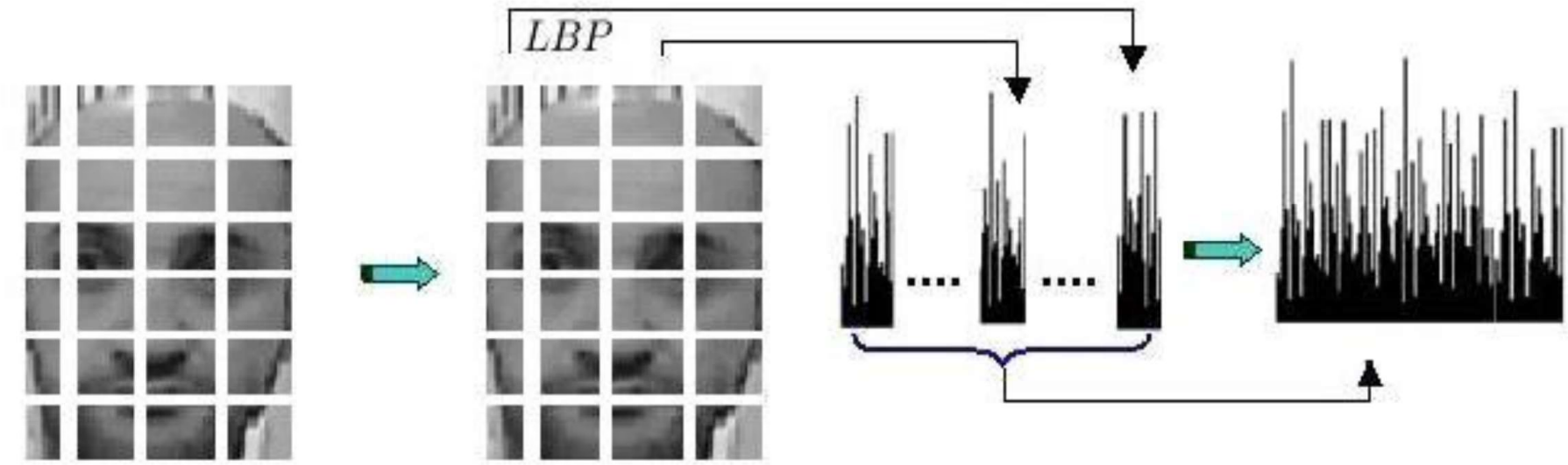


How many patterns?

LBP with Rotation Invariance



LBP Histogram



Gradient Operators

-1	0	1

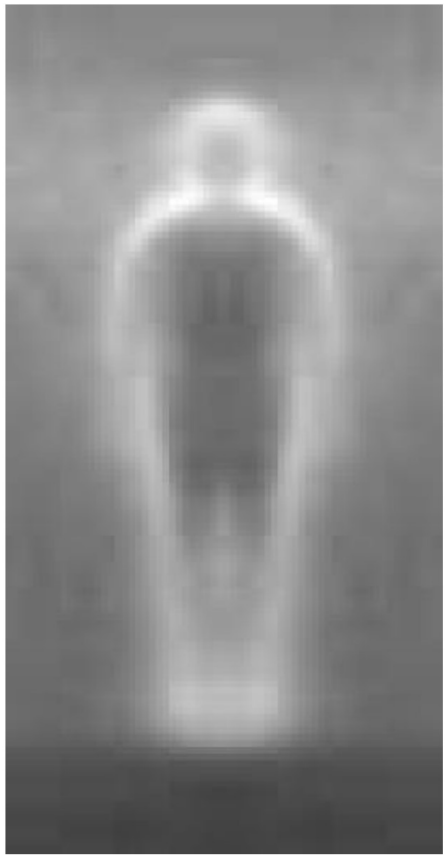
centered

-1	1

uncentered

1	-8	0	8	-1

cubic-corrected



0	1
-1	0

diagonal

-1	0	1
-2	0	2
-1	0	1

Sobel

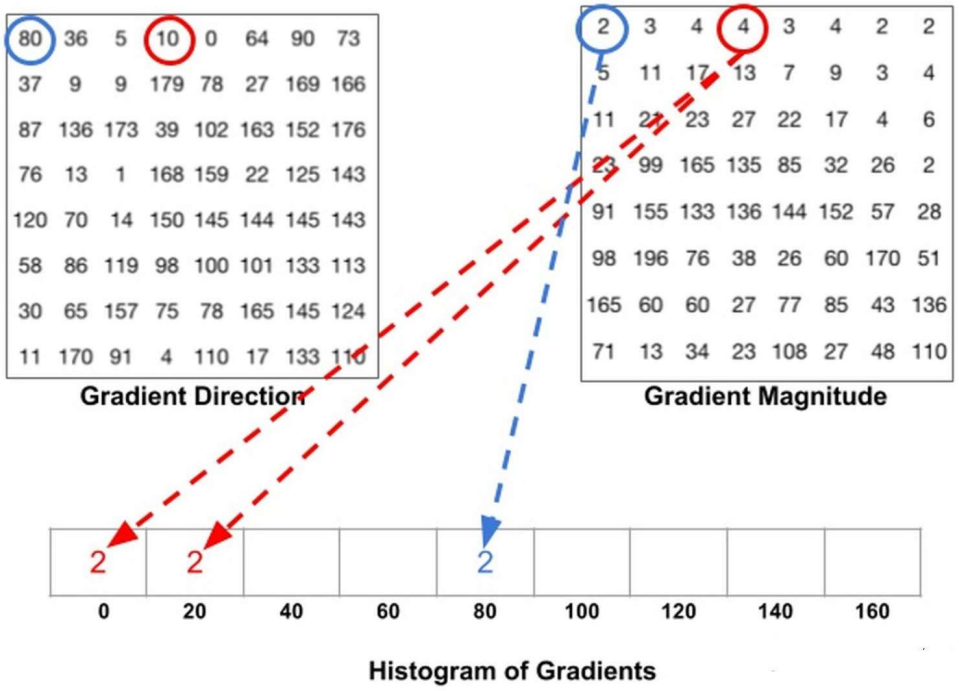
Histogram of Oriented Gradient (HOG)

- Get direction and magnitude from 8*8 cell

$$g = \sqrt{g_x^2 + g_y^2}$$

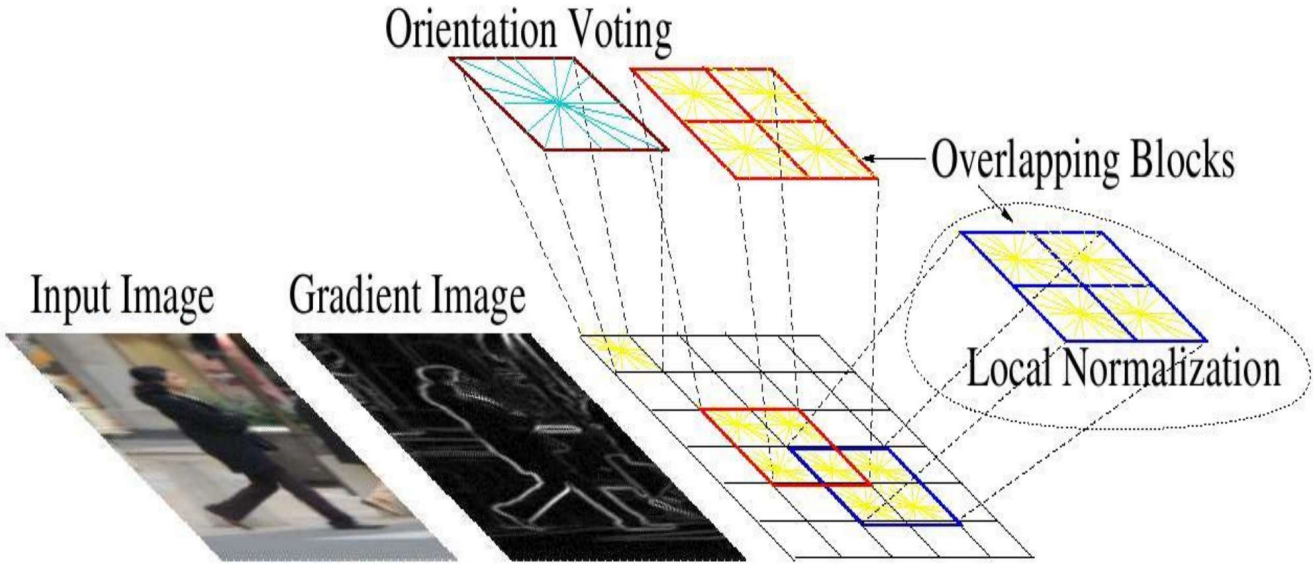
$$\theta = arctan\frac{g_x}{g_y}$$

- Split into 9 bins
  - [0,20,⋯,160]



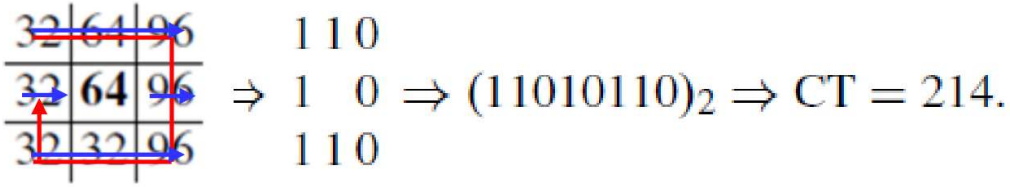
Histogram of Oriented Gradient (HOG)

- Normalization based on 16*16 blocks (4 cells)
  - Normalize on vector with size 4*9=36



Centrist

- Centrist descriptor [Wu and Rehg, '11]
- Census Transform (CT):
  - encodes the "signs of neighboring comparisons" information of a gradient image



- For an image, apply the CT to each pixel, and create the histogram of CT values.