CSCM77
HoG: detecting human

Prof. Xianghua Xie

x.xie@swansea.ac.uk

http://csvision.swan.ac.uk

Human Detection

- Combination of the following concepts
  - Histograms
  - Filtering, edge detection in particular
  - Unsupervised learning: clustering
  - Supervised learning: classification
- We will discuss how we can use these knowledge to perform human detection in complex scenes
  - HoG: Histogram of Oriented Gradients
  - Involves decision making, i.e. detection or classification
    - K-NN: K Nearest Neighbour
    - SVM: Support Vector Machine

Human Detection

- Why is it a difficult problem?
  - Large variability in pose
  - Variable appearance
  - Complex background
  - Unconstrained illumination
  - Occlusions
  - Different scales

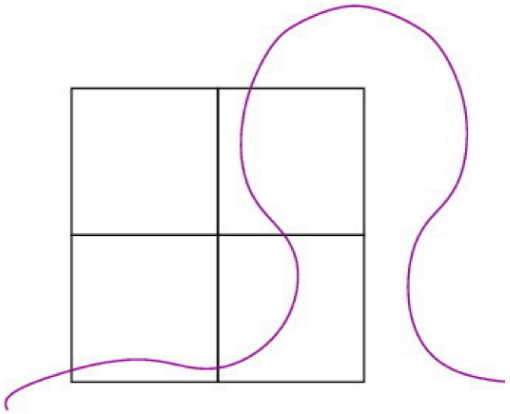
- Constraint in this work:
  - Upright fully visible

Credits: Navneet Dalal, Peter Barnum,

Human Detection

- Histogram of Oriented Gradients
- Assumption:
  - Local shape information is often well described by the distribution of intensity gradients or edge directions
  - Even without the precise information about the location of edges themselves

"Miss a, b, c, d, e"



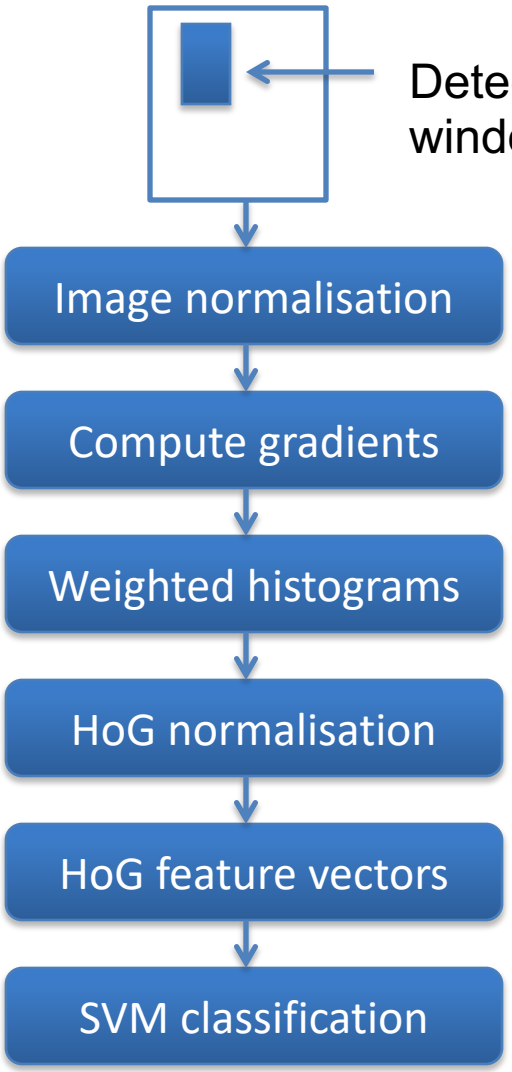
- Head, should, leg silhouettes are likely useful;
- Vertical edges for upright poses;

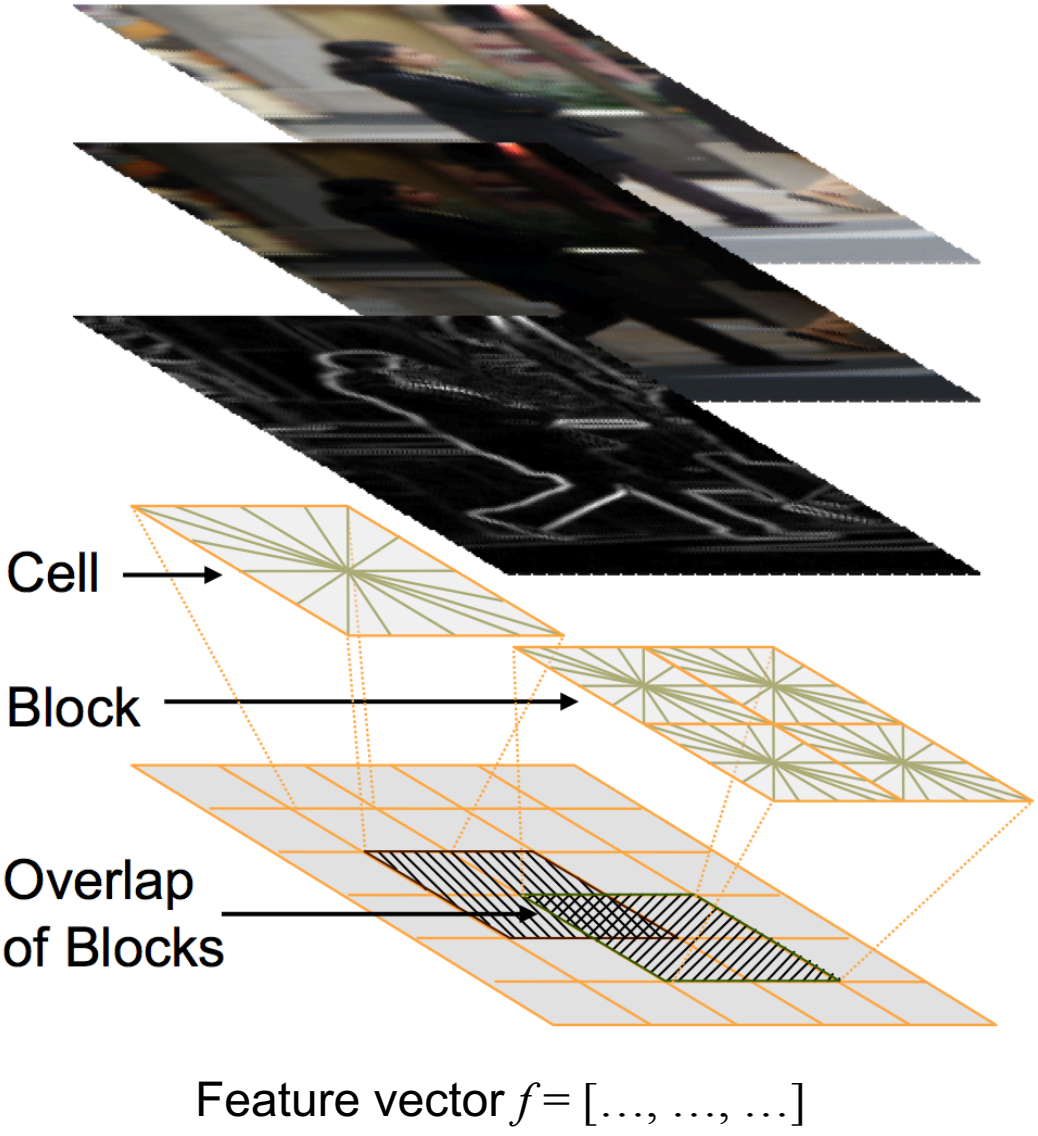
But we are not explicitly searching for these…

Dalal, N.; Triggs, B., "Histograms of oriented gradients for human detection," Computer Vision and Pattern Recognition 2005

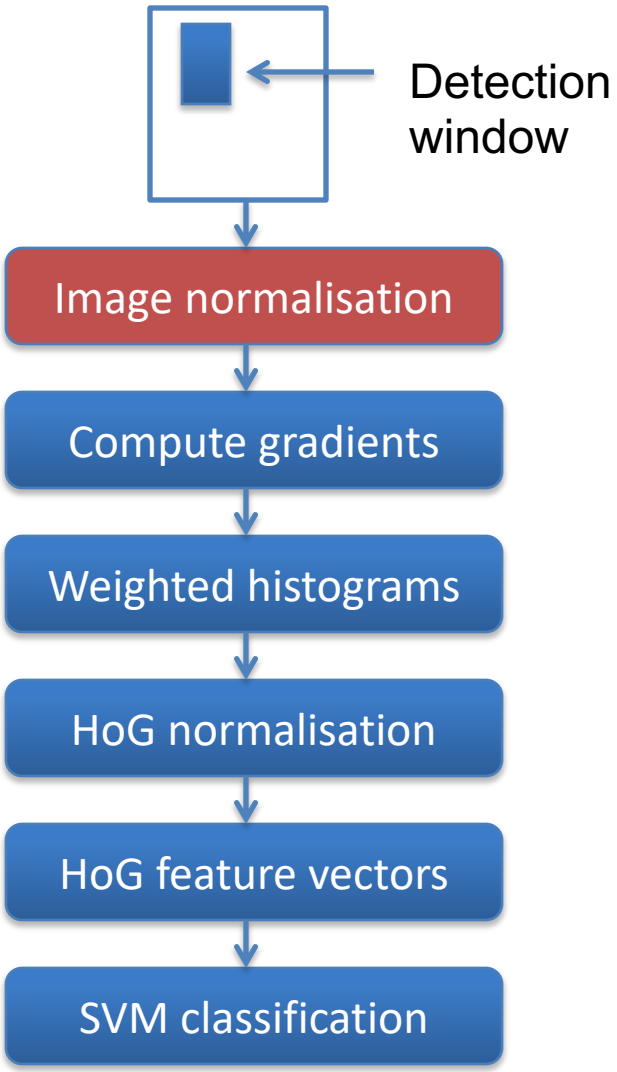
Histogram of Oriented Gradients

- Overview



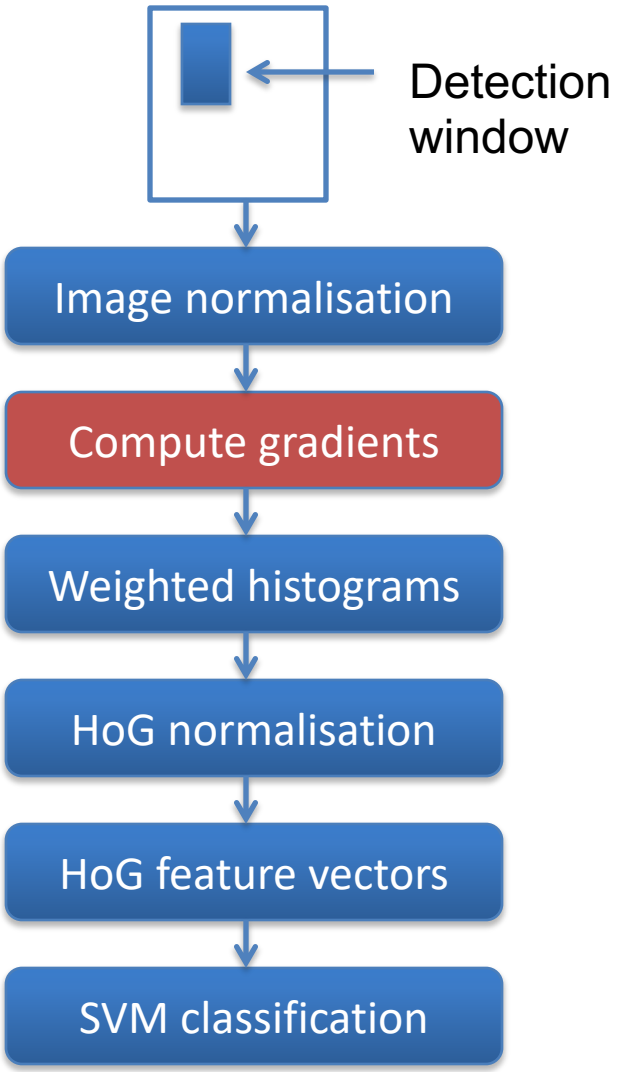


Histogram of Oriented Gradients



- Image normalisation
  - Different pixel representations
    - Gray level
    - Colour: selection of colour space
      - RGB colour space
      - L*a*b colour space – more perceptually uniform, i.e. Euclidean distance metric is more applicable
  - Gamma normalisation and compression
    - Square root, Logarithmic
    - modest effect on performance (possibly due to subsequent feature normalisation)

Histogram of Oriented Gradients

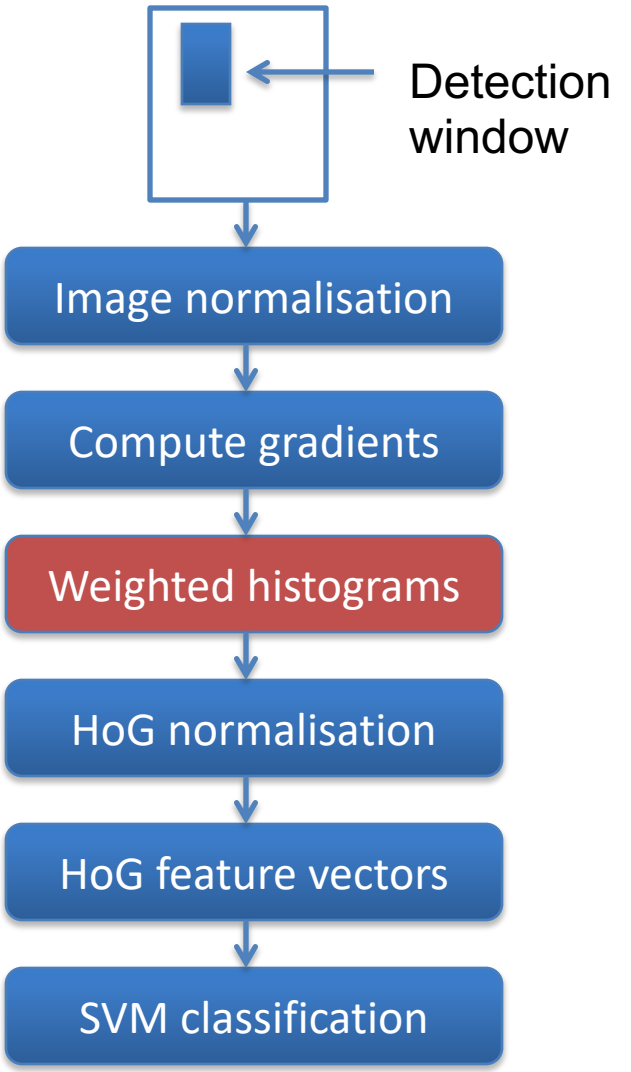


- Gradient computation
  - Filtering with discrete derivative masks
    - Prewitt
    - Sobel
    - Uncentred kernel [1, -1]
    - Diagonal

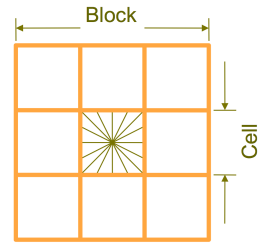
$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

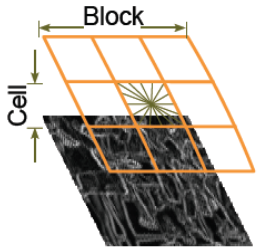
    - Often Prewitt kernel is used, and has been shown in many works it outperforms others in HoG

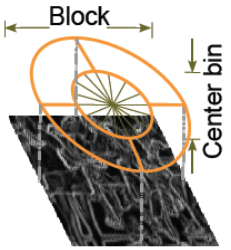
Histogram of Oriented Gradients



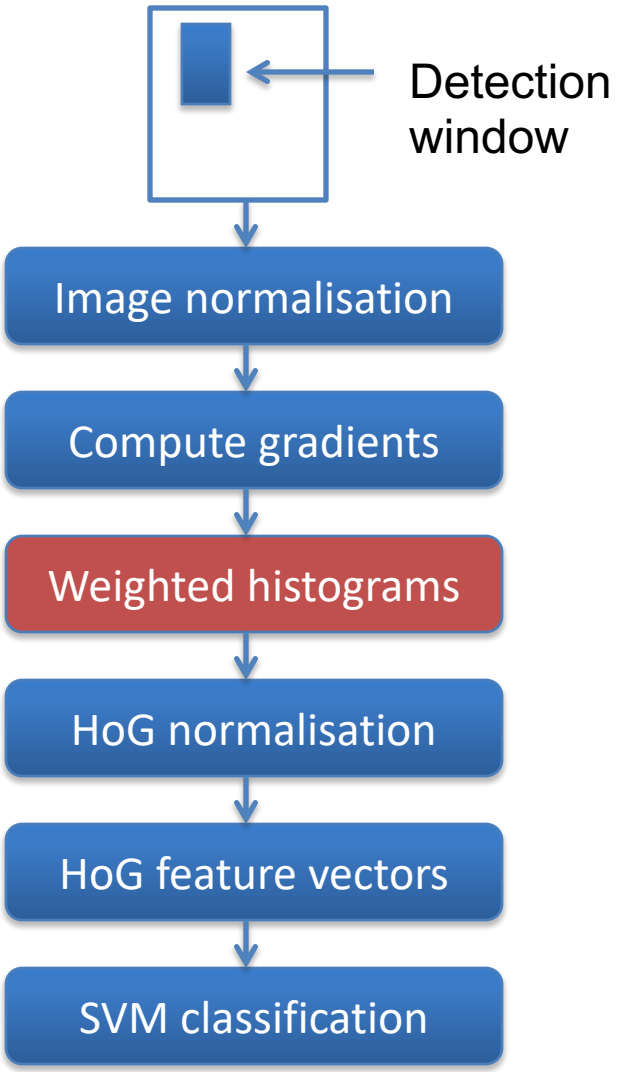
- Weighted histogram of gradient orientations
  - Gradient orientation bins
    - Signed orientation [-180, 180]
    - Unsigned [0, 180]
  - Create histograms of these quantised orientations
    - Rectangular blocks
    - Circular blocks







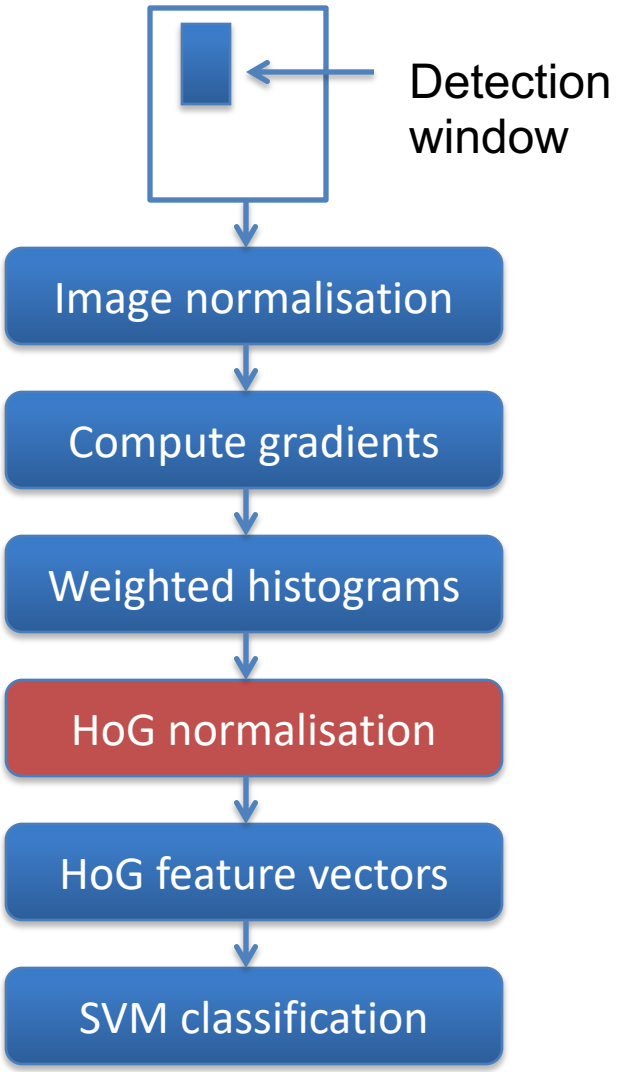
Histogram of Oriented Gradients



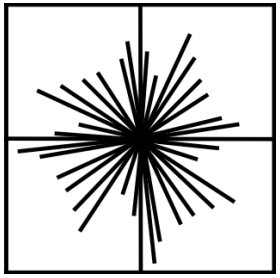
- Weighted histogram of gradient orientations **cont.**
  - Weight the orientation with
    - Gradient magnitude (or its transformation, such as square of magnitude
    - Weighted orientation histogram at each block

Only dominant orientation in each cell is shown

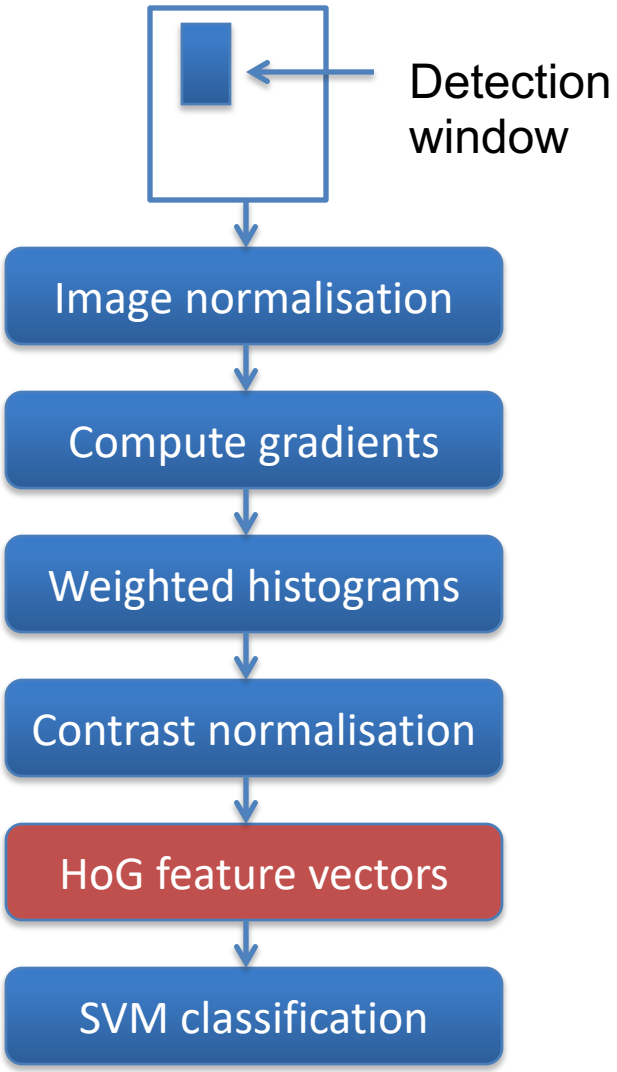
Histogram of Oriented Gradients



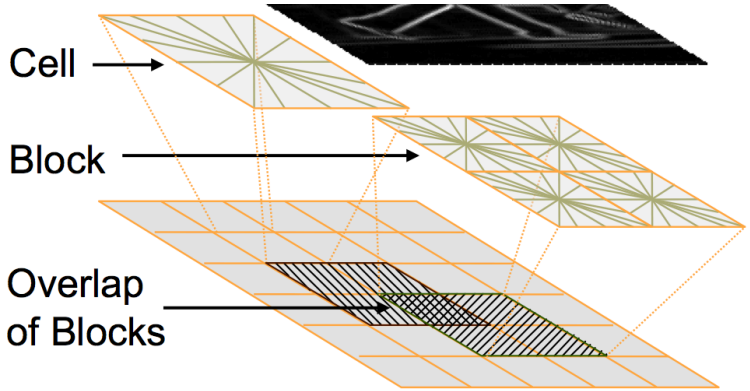
- HoG normalisation
  - Normalising contrast within overlapping blocks of cells
  - Normalisation results in better invariance to changes in illumination and shadowing
  - Common normalisation of histogram, e.g.
    - Based on L1 norm
    - Based on L2 norm



Histogram of Oriented Gradients



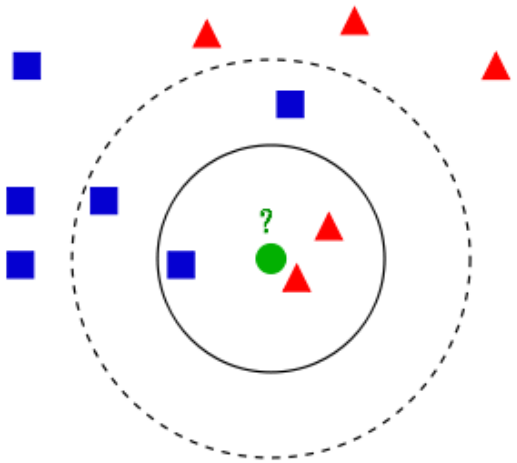
- HoG feature vector
  - Collect normalised histograms in every detection window
  - The detection window convolve across the image in order to detect human
  - Each detection window thus produce a vector (feature vector)
  - These feature vectors are then fed into supervised classifier to make a decision: human or non-human



Histogram of Oriented Gradients

- SVM: Support Vector Machine for Classification

- However, first to mention KNN: K-Nearest Neighbour
  - Classifying testing samples based on closest training samples in the feature space
  - Testing sample is classified by a majority vote of its K nearest neighbours
  - Typically using Euclidean distance metric (L2 norm) to measure closest neighbours



May be sensitive to the number K

- Solid line: classified as red triangle;
- Dashed line: classified as blue rectangle.

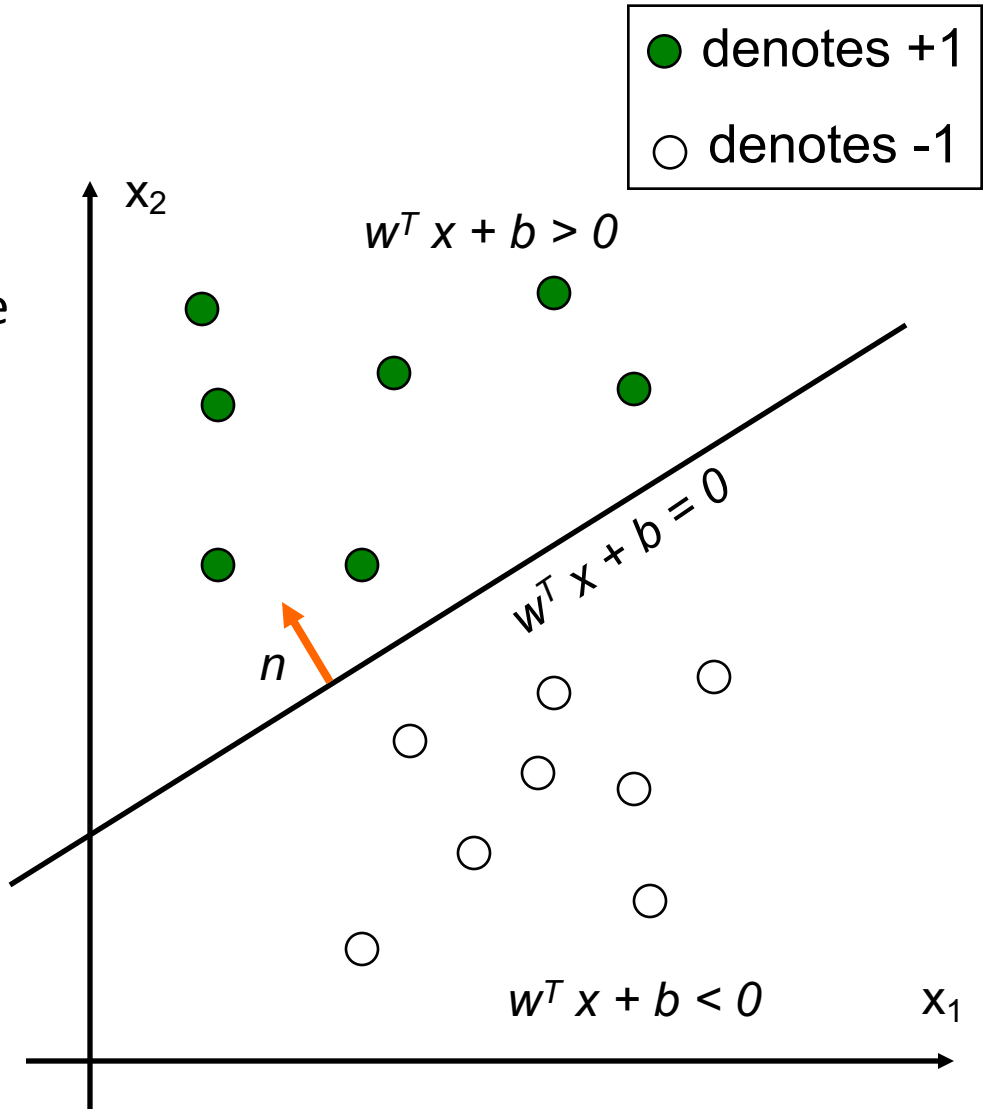
Histogram of Oriented Gradients

A hyper-plane in the feature space (w is the column vector of coeffici ents)

$$\mathbf{w}^{\mathrm{T}}\mathbf{x} + b = 0$$

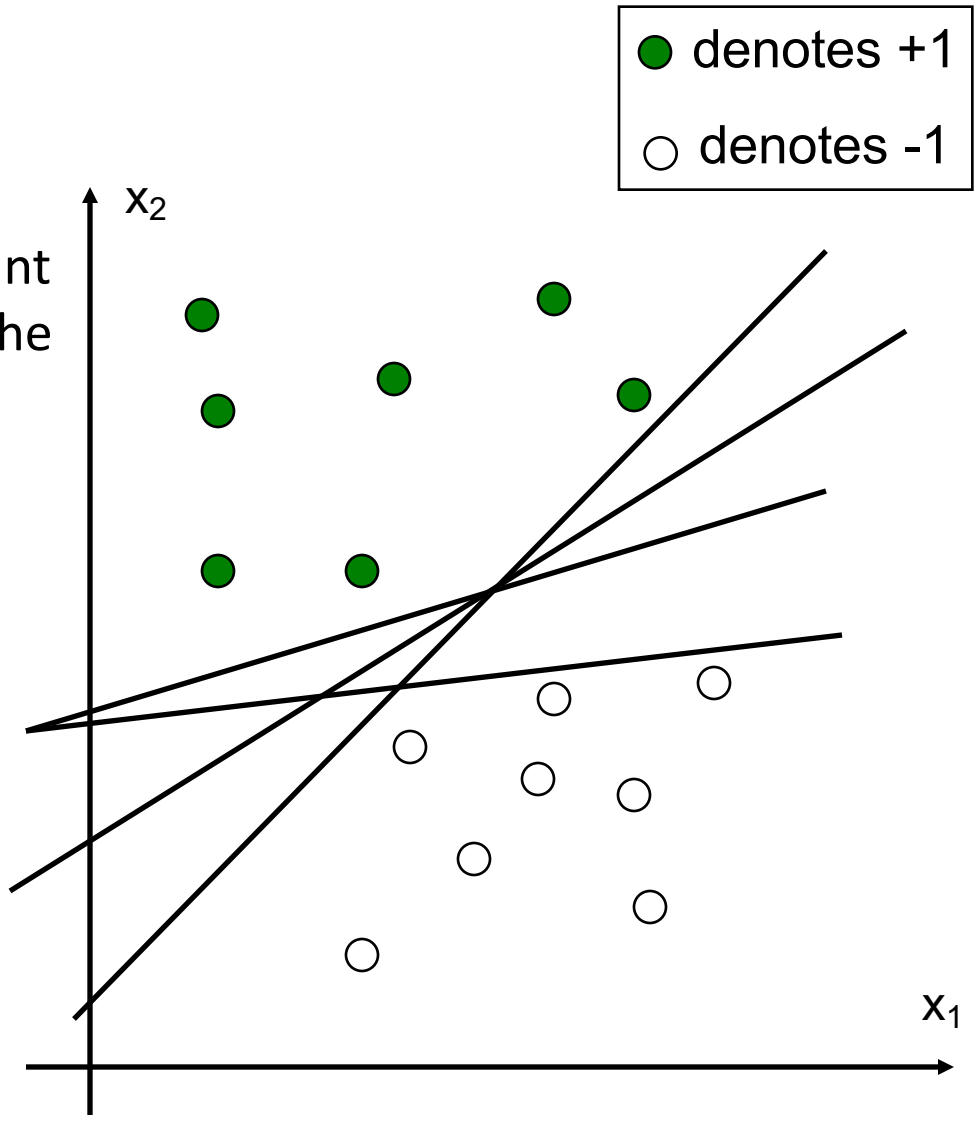
(Unit-length) normal vector of the hyper-plane:

$$\mathbf{n} = \frac{\mathbf{w}}{\|\mathbf{w}\|}$$



SVM: Large Margin Linear Classifier

- How would you classify these points using a linear discriminant function in order to minimize the error rate?
  - Infinite number of answers
  - Which one is the best?



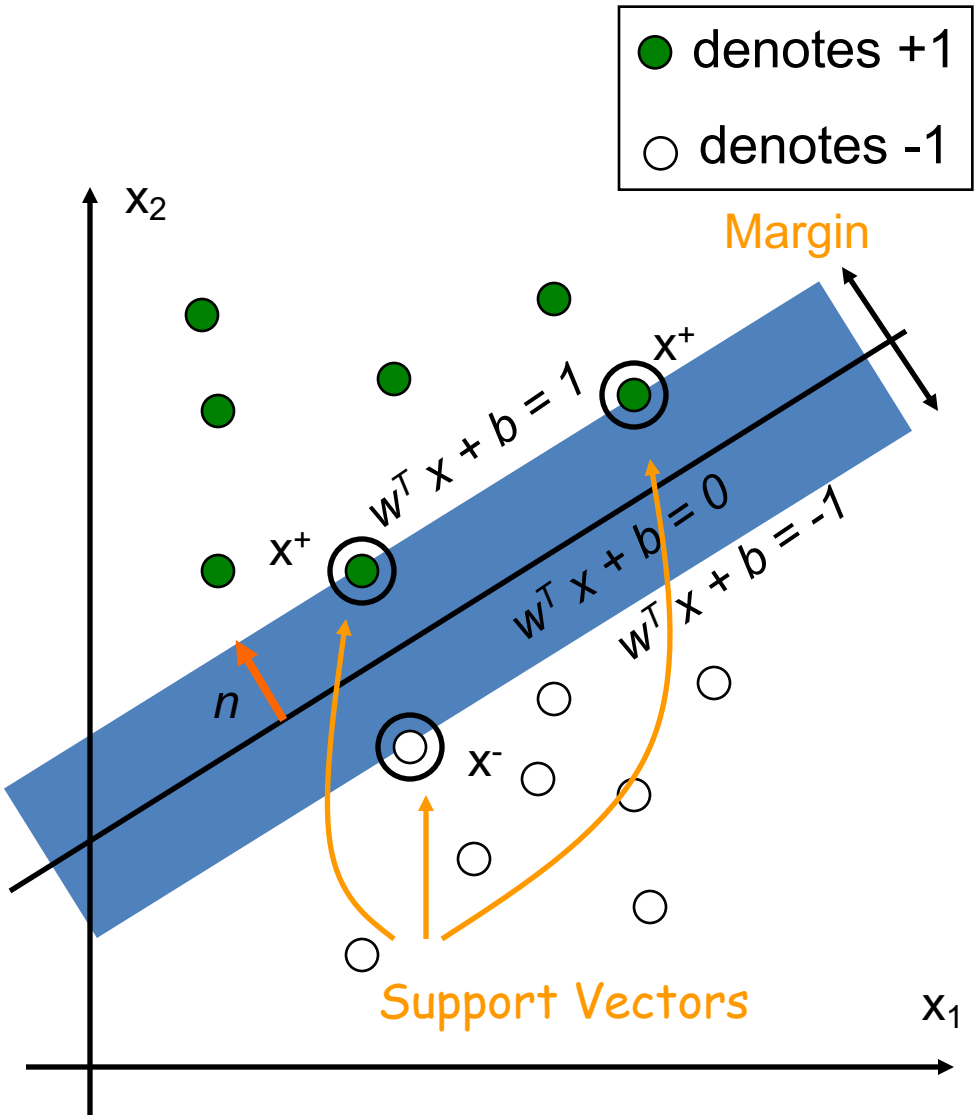
SVM: Large Margin Linear Classifier

- We know that

$$\mathbf{w}^T \mathbf{x}^+ + b = 1$$

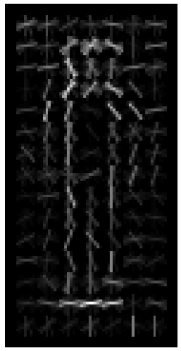
$$\mathbf{w}^T \mathbf{x}^- + b = -1$$

- Maximise the margin width to find the optimal hyper-plane
- It can be proved that this is eq uivalent to minimise $\|\mathbf{w}\|^2$
- The final solution rests on the support vectors (hence the na med SVM)

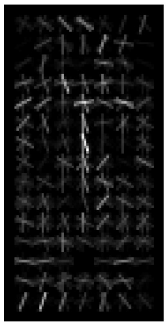


Histogram of Oriented Gradients

- Train the linear SVM with positive and negative samples



Positive samples



Negative samples

Histogram of Oriented Gradients

- SVM effectively determines which HoG features are useful in detecting human and which are not (importance sampling/feature selection)

In each triplet: original image, the corresponding HoG feature vector (with only the dominant orientation of each cell is shown), and the dominant orientation selected by SVM.

Histogram of Oriented Gradients

- Example results

Other applications of HoG