Inputs/outputs/assumptions

• What is the goal?
  – Say yes/no as to whether an object present in image
  And/or:
  – Determine pose of an object, e.g. for robot to grasp
  – Categorize all objects
  – Forced choice from pool of categories
  – Bounding box on object
  – Full segmentation
  – Build a model of an object category

Detection via classification: Main idea

Basic component: a binary classifier

Car/non-car Classifier

If object may be in a cluttered scene, slide a window around looking for it.
Detection via classification: Main idea

Fleshing out this pipeline a bit more, we need to:
1. Obtain training data
2. Define features
3. Define classifier

Feature extraction:

- Sample at multiple scales and positions (and orientations)
- “Does this contain object category X or not?”

Feature extraction: global appearance

Simple holistic descriptions of image content
- grayscale / color histogram
- vector of pixel intensities

Eigenfaces: global appearance description

An early appearance-based approach to face recognition

Generation low-dimensional representation of appearance with a linear subspace.

Recognition via nearest neighbors in face space

Gradient-based representations

- Consider edges, contours, and (oriented) intensity gradients
HOG

(one of the most widely used features)

Gradient-based representations: Histograms of oriented gradients (HoG)

Map each grid cell in the input window to a histogram counting the gradients per orientation.

Code available: http://pascal.inrialpes.fr/software/olt/

K. Grauman, B. Leibe

• Histogram of gradient orientations
  -Orientation
  -Position

  – Weighted by magnitude