Recognition I

Computer Vision

CSE 190-B

Lecture 18

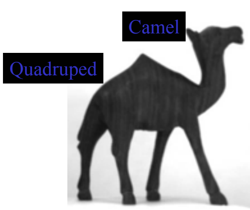
Announcements

- Fourth (last) homework due on Thursday
- In problem 7.6, second to last line should have:

$$\hat{p}_i = \left[(x_i - \bar{x})/\bar{d}, (y_i - \bar{y})/\bar{d}, 1\right]^T$$

- **Makeup class, Thursday, June 5, 11:00-12:30, APM 4218**

- Final Exam, Thursday, June 12, 3:00-6:00, WLH 2205



Problem:

Recognizing instances

Recognizing categories

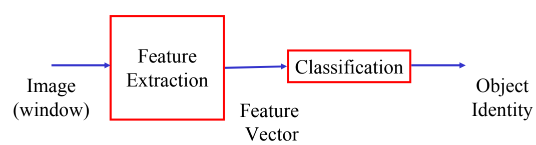
Many issues

- Instances vs. categories
- Object representation: 2-D, 3-D, primitives, volumes, points, lines, color, dynamics, function,
- Image features: Intensities, Color, Texture, Shape (2D or 3-D), Motion
- Classification method
- Search strategies
- Automatically learned or taught

Recognition

- Appearance-based vision
  - Often pattern recognition or learning-based
  - Pattern recognition
  - Image as a feature vector - nearest neighbor
  - PCA (Eigenspaces)
  - Appearance manifolds
- Feature-based recognition
  - "3-D Model-based"
  - Interpretation trees
  - Invariants
- Function

Sketch of a Recognition Architecture

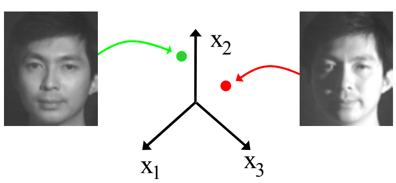


Bayesian Classification

Example: Face Detection

- Scan window over image.
- Classify window as either:
  - Face
  - Non-face

Image as a Feature Vector



- Consider an n-pixel image to be a point in an n-dimensional space, $\mathbf{x} \in \mathbf{R}^n$.
- Each pixel value is a coordinate of **x**.

The Problem of Recognition

Given an image I and a database of k objects and a representation R_j for object j in the database, recognition can be expressed as:

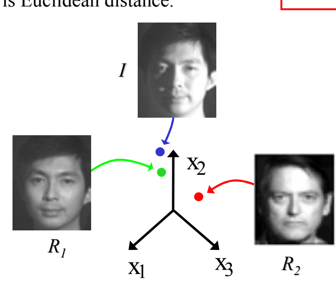
$$\boxed{i = \arg\min_{j \in [1,\cdots,k]} c(R_j, I)}$$

where $c(R_j, I)$ is a function which gives the compatibility or consistency of representation j with the image.

Simplest Recognition Scheme

- R_j is an image.
- $c(R_j, I)$ is Euclidean distance.

⇨ **Nearest Neighbor Classifier**



Face Recognition

- Whose face is this? (perhaps in a mugshot)
- Issue:
  - What differences are important and what not?
  - Reduce the dimension of the images, while maintaining the "important" differences.
- One strategy:
  - Principal components analysis
- Many face recognition strategies at http://www.cs.rug.nl/users/peterkr/FACE/face.html

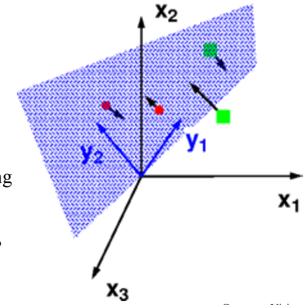
Eigenfaces: linear projection

- An n-pixel image $\boldsymbol{x} \in \mathbf{R}^n$ can be projected to a low-dimensional feature space $\boldsymbol{y} \in \mathbf{R}^m$ by

$$\boldsymbol{y} = W\boldsymbol{x}$$

where W is an n by m matrix.

- Recognition is performed using nearest neighbor in $\mathbf{R}^m$.
- How do we choose a good W?



Eigenfaces: Principal Component Analysis (PCA)

Assume we have a set of n feature vectors $\boldsymbol{x}_i$ $(i = 1, \ldots, n)$ in $\mathbb{R}^d$. Write

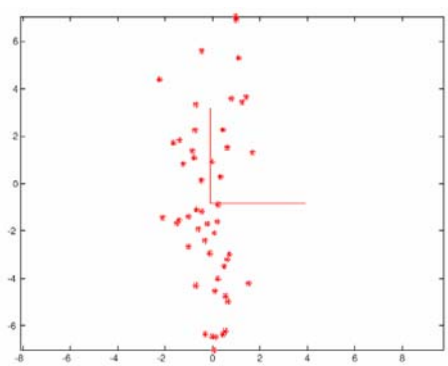
$$\boldsymbol{\mu} = \frac{1}{n}\sum_i \boldsymbol{x}_i$$

$$\Sigma = \frac{1}{n-1}\sum_i (\boldsymbol{x}_i - \boldsymbol{\mu})(\boldsymbol{x}_i - \boldsymbol{\mu})^T$$

The unit eigenvectors of Σ — which we write as $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_d$, where the order is given by the size of the eigenvalue and $\boldsymbol{v}_1$ has the largest eigenvalue — give a set of features with the following properties:

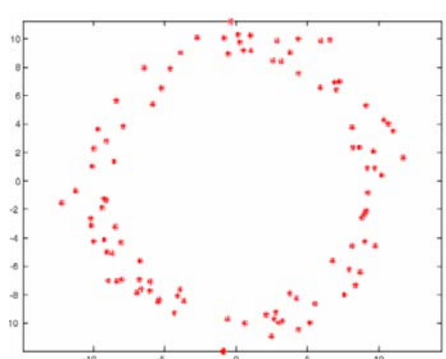
- They are independent.
- Projection onto the basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ gives the k-dimensional set of linear features that preserves the most variance.

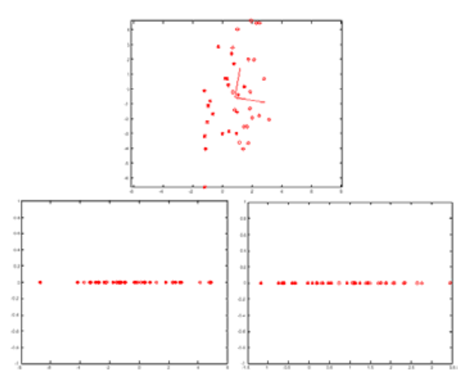
Algorithm 22.5: *Principal components analysis identifies a collection of linear features that are independent, and capture as much variance as possible from a dataset.*



Difficulties with PCA

- Projection may suppress important detail
  - smallest variance directions may not be unimportant
- Method does not take discriminative task into account
  - typically, we wish to compute features that allow good discrimination
  - not the same as largest variance





Illumination Variability

"The variations between the images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity."
-- Moses, Adini, Ullman, ECCV '94

Fisherfaces: Class specific linear projection

P. Belhumeur, J. Hespanha, D. Kriegman, *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*, PAMI, July 1997, pp. 711--720.

- An n-pixel image $\boldsymbol{x} \in \mathbf{R}^n$ can be projected to a low-dimensional feature space $\boldsymbol{y} \in \mathbf{R}^m$ by

$$\boldsymbol{y} = W\boldsymbol{x}$$

where W is an n by m matrix.

- Recognition is performed using nearest neighbor in $\mathbf{R}^m$.

- How do we choose a good W?