EE 6885 Statistical Pattern Recognition

Fall 2005
Prof. Shih-Fu Chang
http://www.ee.columbia.edu/~sfchang

Lecture 1 (9/7/05)

dvmm

Introduction

- Reading
  - Textbook DHS Chapter 1
  - References:
    - DIP Gonzalez and Woods, Textbook, Chapter 12, Object Recognition
    - Paper:
      Anil K. Jain, etc., "Statistical Pattern Recognition: A Review," IEEE Tran. on Pattern Analysis and Machine Intelligence, vol 22, No 1, Jan. 2000.

dvmm

About the course

- Objectives:
  - Learn how to formulate and solve pattern recognition problems
    - Feature extraction, object/event recognition, structure detection, video search and retrieval
  - Get insights and experience of recent machine learning techniques
    - Bayesian, SVM, Neural Network, HMM, feature selection, clustering
  - Have fun in experimenting with actual visual classification/indexing problems
  - Demos:
    - Google Yahoo Image search, Goggle News search, Columbia video search
- Intended Audience
  - Beginning graduate students or professionals
  - familiar with signal/image processing
  - comfortable with probability, statistics, linear algebra, and some machine learning concepts

Sample Applications: Video Indexing and Analysis

- Indexing, search, and retrieval for images and videos
  - See Columbia's WebSEEk and VideoQ http://www.ee.columbia/edu/{webseek, videoq}
  - Goggle image search
  - "find video clips of basketball going through the hoop"
  - "find images containing shape shown in the sketch"
- Automatic annotation of visual content
  - recognition of text, face, scene, vehicle, location, etc
- Automatic parsing of video programs into structures
  - break videos into shots, scenes, and stories
- Event detection
  - sports events, human activities, meetings, medical, and other spatio-temporal patterns
- Summary
  - topic clustering, highlight generation
  - See Columbia's sports highlight, news topic clustering demo

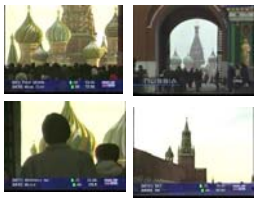
What do we search in video? (examples)

- TRECVID 2003 **Event**
  - "Find shots of an airplane taking off."

- TRECVID 2004 **Named-person Location**
  - "Find shots of Bill Clinton speaking with a US flag visible behind him."

- IBM Speech Group **Objects**
  - "Find shots containing monkeys or gorillas."

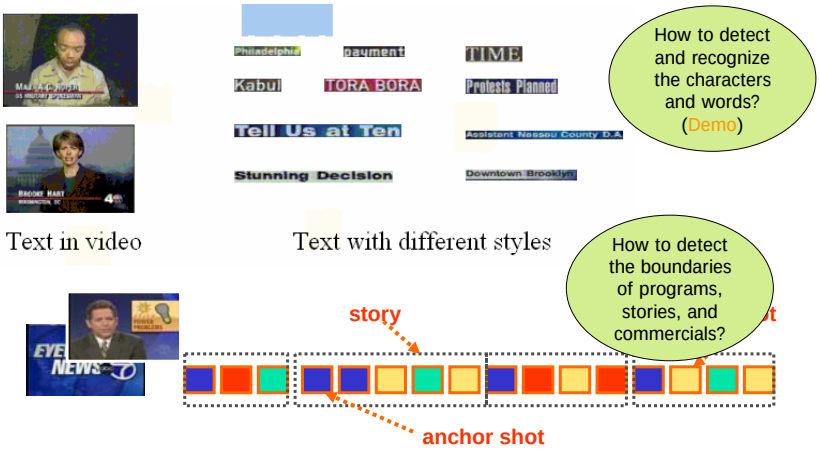
- BBC Logs **Named-location**
  - "Find shots of the Kremlin."



Related Activities NIST TRECVID

- Low-level feature detection (motion, shot etc)
- High-level feature detection
  - Image →classifier→ {'people', 'vehicle', 'explosion', etc}
- Story boundary detection
- Search : fully automatic, manual, interactive
- 2005 Data
  - 6 channels in English, Chinese, Arabic
  - >170 hours, 126,000 subshots
  - 39 concepts manually annotated over >80 hours (LSCOM-Lite) : very valuable resource for researchers!
- Participating groups
  - 62 groups (due: high-level feature 8/22, search 9/21/05)

Examples of object & structure recognition



A multi-modal pattern recognition problem

- Objective: a story boundary at time τ_k?
  - τ_k = { shot boundaries or significant pauses}

(Hsu and Chang)

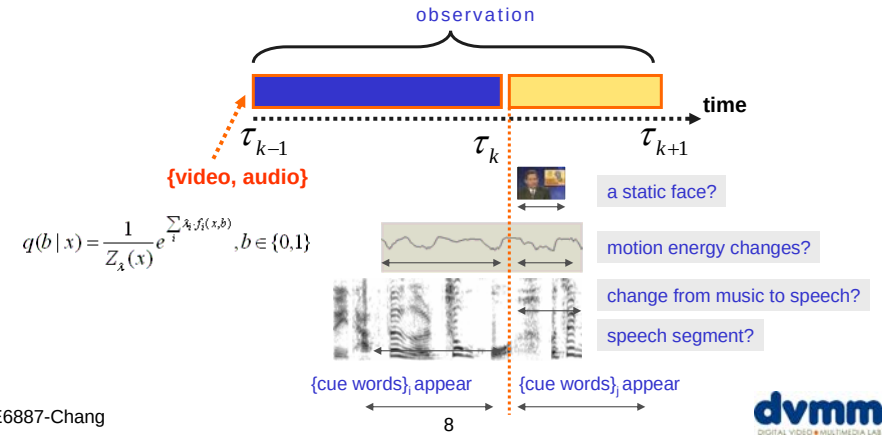
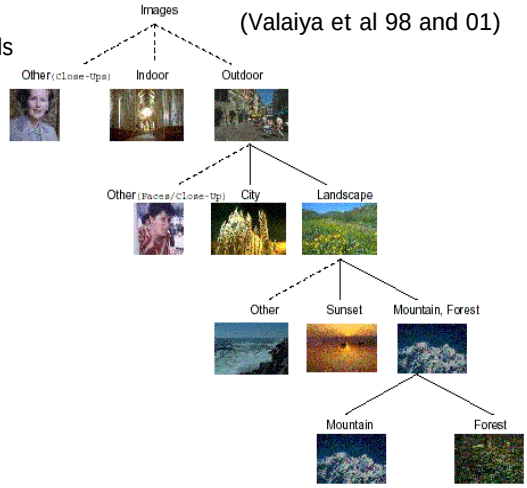
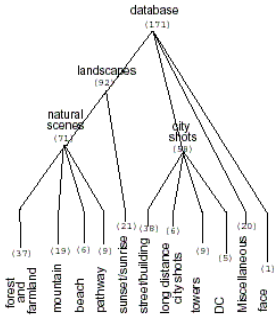


Image to Keyword Annotation

- Tag images with category labels
- Organize the photos into hierarchical categories

(Valaiya et al 98 and 01)



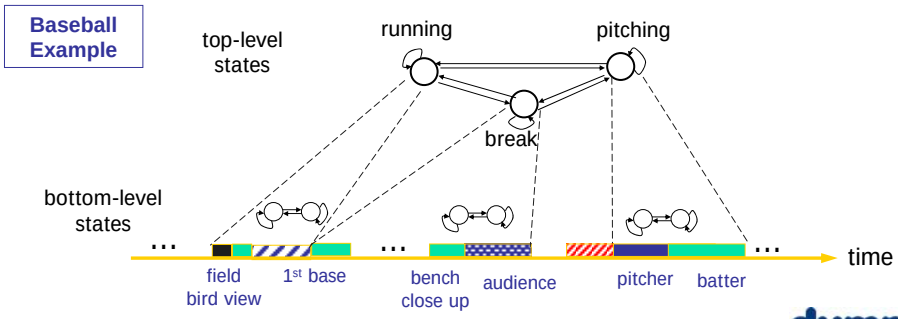


Unsupervised Video Structure Discovery: Hierarchical Hidden Markov Model

[Demo]

(Xie et al)

- Inference Multi-Level Markovian Temporal Dependence
  - High-level states represent distinct events
  - Presence of each event produces observations modeled by low-level HMMs



A Very High-Level Stat. Pattern Recog. Architecture

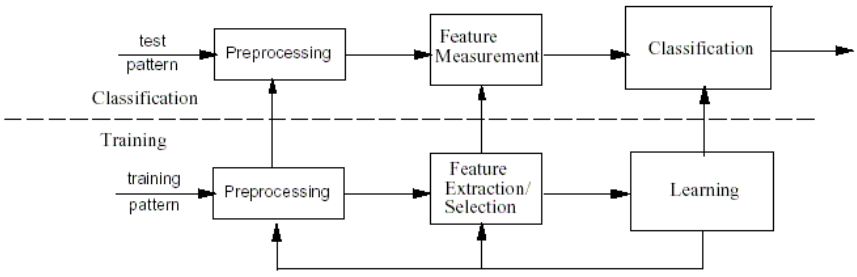


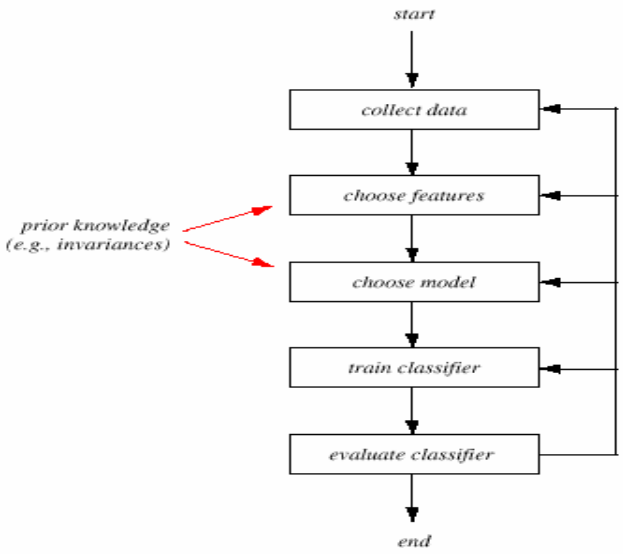
Figure 1: Model for statistical pattern recognition.

(From Jain, Duin, and Mao, SPR Review, '99)

Important issues

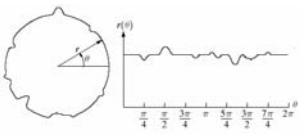
- Image/video pre-processing – quality, resolution etc
- Feature extraction
  - Color, texture, motion, shape, layout, regions, parts, etc
- Feature representation
  - Discrete vs. continuous, vectorization, dimension
  - Invariance to scale, rotation, translation …
- Feature selection
  - PCA, MDS, Kernel PCA, etc
- Classification models
  - Generative vs. discriminative
  - Multi-modal fusion, early fusion vs. late fusion
- Size of training/test data and manual supervision efforts
- Validation and evaluation processes
- Complexity

Design Cycle



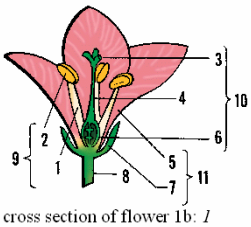
Some examples of feature representation

- Features determine the patterns and their separability
- E.g.,
  - Angular distance for closed shapes
  - Part features for iris flowers



a b

FIGURE 12.2 A noisy object and its corresponding signature.



cross section of flower 1b: *1* filament, *2* anther, *3* stigma, *4* style, *5* petal, *6* ovary, *7* sepal, *8* pedicel, *9* stamen, *10* pistil, *11* perianth

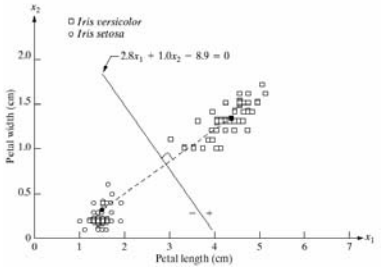
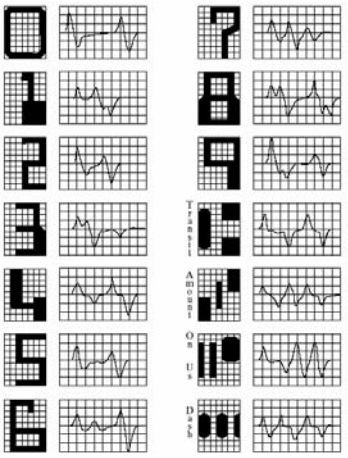


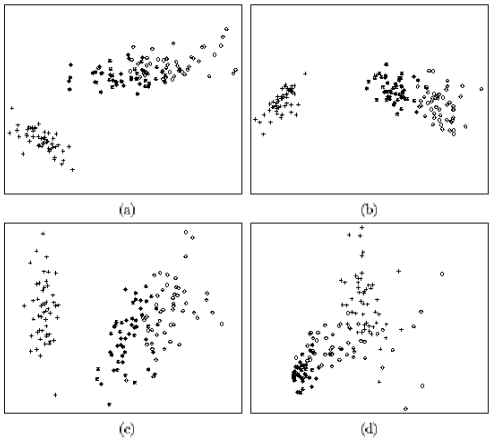
FIGURE 12.6 Decision boundary of minimum distance classifier for the classes of *Iris versicolor* and *Iris setosa*. The dark dot and square are the means.

Another example of feature



- Bankers Asso. Font used on personal checks
- Use magnetic ink and reader to simplify segmentation
- Feature: the horizontal scan of the rate of increase/decrease of the character area
- Peaks and zeros are arranged to be located at the vertical grid lines → can be sampled accurately
- Patterns can be easily distinguished

Feature Representation Extraction/Selection



(Jain et al 99)

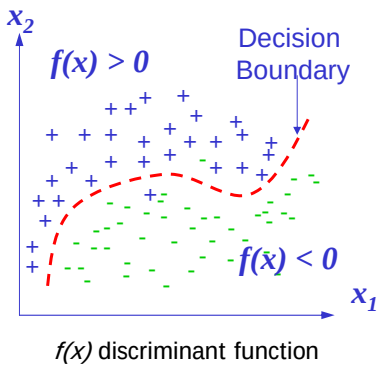
PCA

Fischer Analysis

MDS

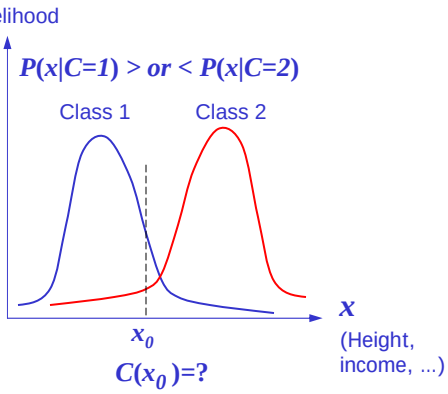
Kernel PCA

Classification Paradigms



f(x) discriminant function

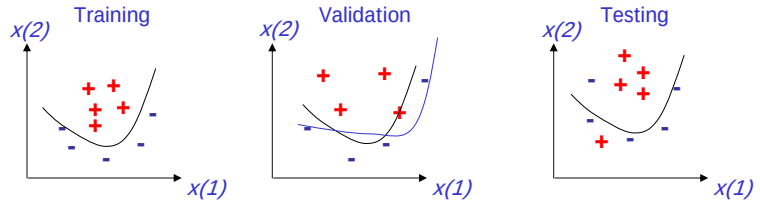
Discriminative



(Height, income, ...)

Generative

Training / Validation / Testing



optimal features, models, parameters

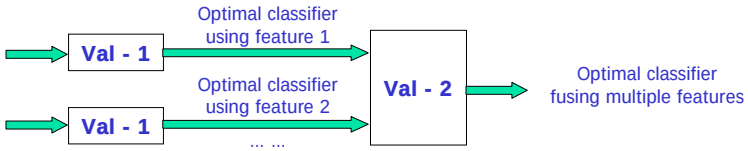
Select optimal hypothesis through validation

Evaluate performance over test data

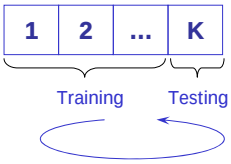
- Assume the same distribution in different set, otherwise the optimal solution from validation may not be optimal in test data

Training / Validation / Testing (cont.)

- Multiple validation sets can be used for different optimization steps.



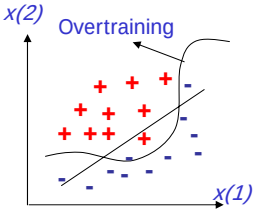
- Cross validation, leave-one-out

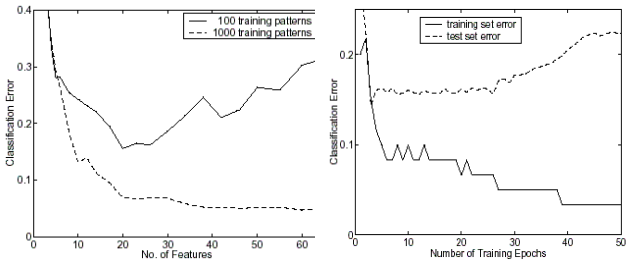


Rotate the choice of the test set and average the performance over runs

Curse of Dimensionality and Overfitting

A case of overtraining





Rule of thumb – (# of training patterns per class) / (# of features) > 10

Topics (tentative)

- Image/Video feature extraction
- Bayesian Decision Theory
- Parameter Estimation (ML and non-parametric)
- Linear Discriminant Functions
- Support Vector Machines
- Neural Networks
- Fusion, evaluation
- Feature Selection
- Clustering

Course Format

- Weekly Lectures
- Textbook
  - Pattern Classification, 2nd Edition, Richard O. Duda, Peter E. Hart, and David G. Stork, 2000, Wiley
- Reference Books
  - Related research papers, distributed online
  - The Elements of Statistical Learning, Trevor Hastie, Robert Tibshirani & Jerome Friedman, 2001, Springer Verlag
  - Machine Learning, Tom Mitchell, 1997, McGraw Hill
- Grading
  - 6 homeworks (30%), midterm (25%), final (25%), mini-project (20%)
  - Homeworks include both analytical problems and programming assignments
  - Matlab is recommended for programming

Resources and Matlab

- Links on the class web site
  - Tutorials on paper writing, Matlab, etc
- Software links on web site to
  PR Tools, Spider (SVM), Netlab, HMM, etc
- EE6887 SPR Class Dataset
  - Benchmark data set, a few thousands of images from broadcast news and stock photos
  - Extracted features and labels
  - Will distribute for class project use only
- Matlab is recommended for programming
  - Accessible in Mudd 251 Computer Lab
    (Need CU ACIS account)
  - Student edition available

- Next Lecture
  - Image feature extraction
  - Bayesian Decision and Bayesian Minimal Risk Classifiers
    Textbook DHS Chap. 2.1 – 2.6