SYDE 372 - Winter 2011
Introduction to Pattern Recognition

Basic Concepts

Alexander Wong

Department of Systems Design Engineering
University of Waterloo

Outline

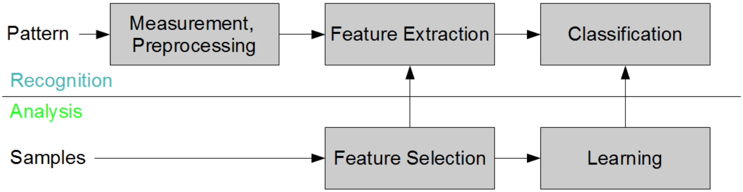
What is pattern recognition?

- process in which some input is measured, analysed, and classified as being more or less similar to a learned model.
- Example of pattern recognition task: optical character recognition:
  - A sensor measures a scene depicting a character.
  - Measured image is analysed and important features are extracted from the image (e.g., edge orientation, number of black pixels, etc.)
  - The extracted features are evaluated based on some prior knowledge/model to associate a label and meaning to the character.

General pattern recognition framework

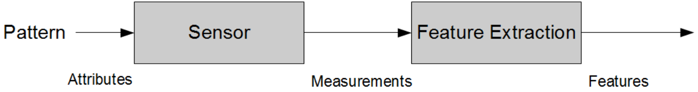
- The general pattern recognition framework consists of the following components:
  - **Measurement/Preprocessing**: input pattern is measured and preprocessed for improved recognition (e.g., data binarization)
  - **Feature Extraction**: Features are extracted to create concise representation of pattern
  - **Classification**: Assign a class label to the pattern based on some learned model.
  - **Feature selection**: Select a set of meaningful features to represent samples used for training the classifier
  - **Learning**: Learn a model based on features representing the samples.

General pattern recognition framework



Patterns

- **Patterns** have properties or attributes which distinguishes it from other patterns (e.g., apples vs. oranges)
- **Measurements** taken of a pattern should reflect either directly or indirectly the attributes associated with pattern (e.g., images)
- **Features** provide a concise representation of measurements to facilitate classification (e.g., shape, color, size, etc.)



Statistical Pattern Recognition

- **Measurements** represented by a vector, $\underline{x}$, consisting of m measurements obtained from sensor (e.g., color intensities at each pixel in image)
- **Features** represented by a vector, $\underline{y}$, consisting of n features obtained from measurements (e.g., overall color, shape, size, texture, etc.)

$$\underline{x} = [x_1 \quad x_2 \quad \ldots \quad x_m]^T \tag{1}$$

$$\underline{y} = [y_1 \quad y_2 \quad \ldots \quad y_n]^T \tag{2}$$

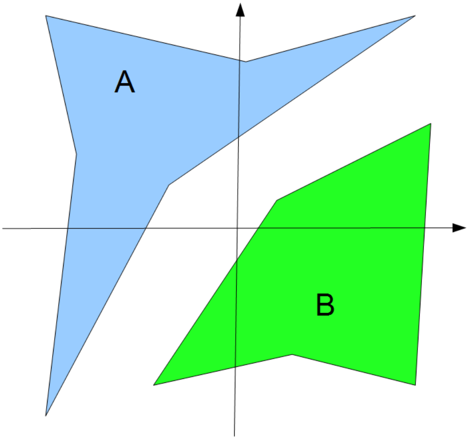
Classes

- Goal of pattern recognition is to assign input to a class.
- A **class** is a group or set of patterns which are similar in some sense (i.e., they share some common properties or attributes.
- Need to represent class somehow:
  - Prototype: idealized form that boils down the “essence” of a class (e.g., mean apple).
  - Set of samples known to belong to a class (e.g., a bunch of samples known to be apples)

Feature space

- Note: patterns doe not need to be identical to belong to the same class (e.g., not all apples have to be the exact same color)
- Patterns within the same class may different due to:
  - Noise or random variations in measurement process (e.g., apple imaged at different perspectives)
  - Inherent variability within a class (e.g., some apples are larger or rounder than others)
- A class typically spans a region within feature space to account for differences in measurement and feature values amongst its members.

Feature space



Feature extraction and classification

- **Feature extraction**: process of transforming from measurements to features
- Attempts to recover defining attributes from patterns to facilitate classification
- **Classification**: process of transforming features into class name or labels
- Strong relationship between feature extraction and classification. For example:
  - Good features allow for simple classifiers (e.g., best feature is just the class label!)
  - Complex classifiers compensate for features that are not linearly separable.

Classification problems

- Three conceptually different types of classification problems we wish to tackle:
  - **1. Probability model is known for each class**:
    - Typical in cases where the physical process is known and provides a probability model (e.g., Gaussian noise), or reasonable assumption can be made about the probabilistic behavior (e.g., car arrival time as a Poisson process)
    - Statistical decision theory may be used to find optimal classification in the sense of minimizing probability of error.

Classification problems

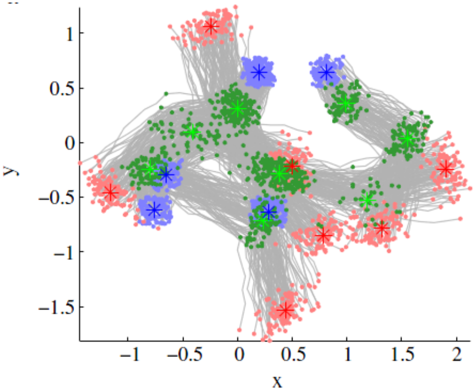
- Three conceptually different types of classification problems we wish to tackle:
  - **2. Samples with known class labels**:
    - Typical in cases where we wish to use samples that have been manually labeled for classification (e.g., face images tagged by human observer)
    - Two possible approaches: i) learn empirical probability model based on samples, and ii) derive classifiers directly from distribution of samples in feature space.

Classification problems

- Three conceptually different types of classification problems we wish to tackle:
  - **3. Samples with no other known label information**:
    - Need to not only determine the definition of each class, but also determine the number of classes!
    - Commonly referred to as a clustering problem
    - Approach is to look for naturally occurring order, groupings or clusters in the data.

Samples with no other known information

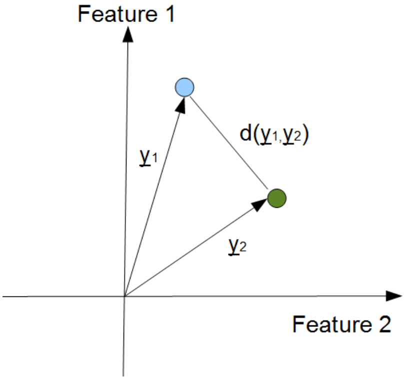
- Example: Crowd motion trajectory from video footage.



Similarity

- Regardless of initial class definitions, pattern recognition system must assign unknown pattern to known class.
- To make decision, classifier needs to assess similarity of unknown patter to known class
- Two patterns are similar of they share common properties
- In vector space representation, sharing common properties implies closeness in feature space
- Such closeness can be measured quantitatively using a distance metric $d(\underline{y}_1, \underline{y}_2)$ (the lower the d, the greater the similarity)

Distance metric



Challenges with designing similarity measures

- Features may have fundamentally different natures that are hard to compare as a whole (or even hard to quantify)
- For example, some features are continuous (e.g., height),
- Some features are discrete (number of children)
- Some features are unordered states (marital status, sex, race, religious denomination, etc.)
- What similarity measure is appropriate for classifying based on such features?