Introduction

chap 1

Introduction

- Pattern Recognition

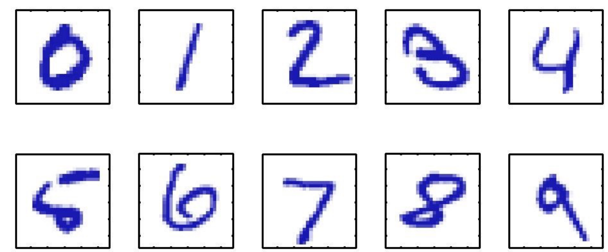
the act of taking in raw data and making an action based on the “category” of the pattern

the automatic discovery of regularities in data through the use of computer algorithms and with the use of these regularities to take actions such as classifying the data into different categories.

  - has been crucial for our survival, and over the past tens of millions of years we have evolved highly sophisticated neural and cognitive systems for such tasks.
  - Recognizing a face, understanding spoken words, reading handwritten characters, and deciding whether an apple is ripe by its smell…

Machine Perception

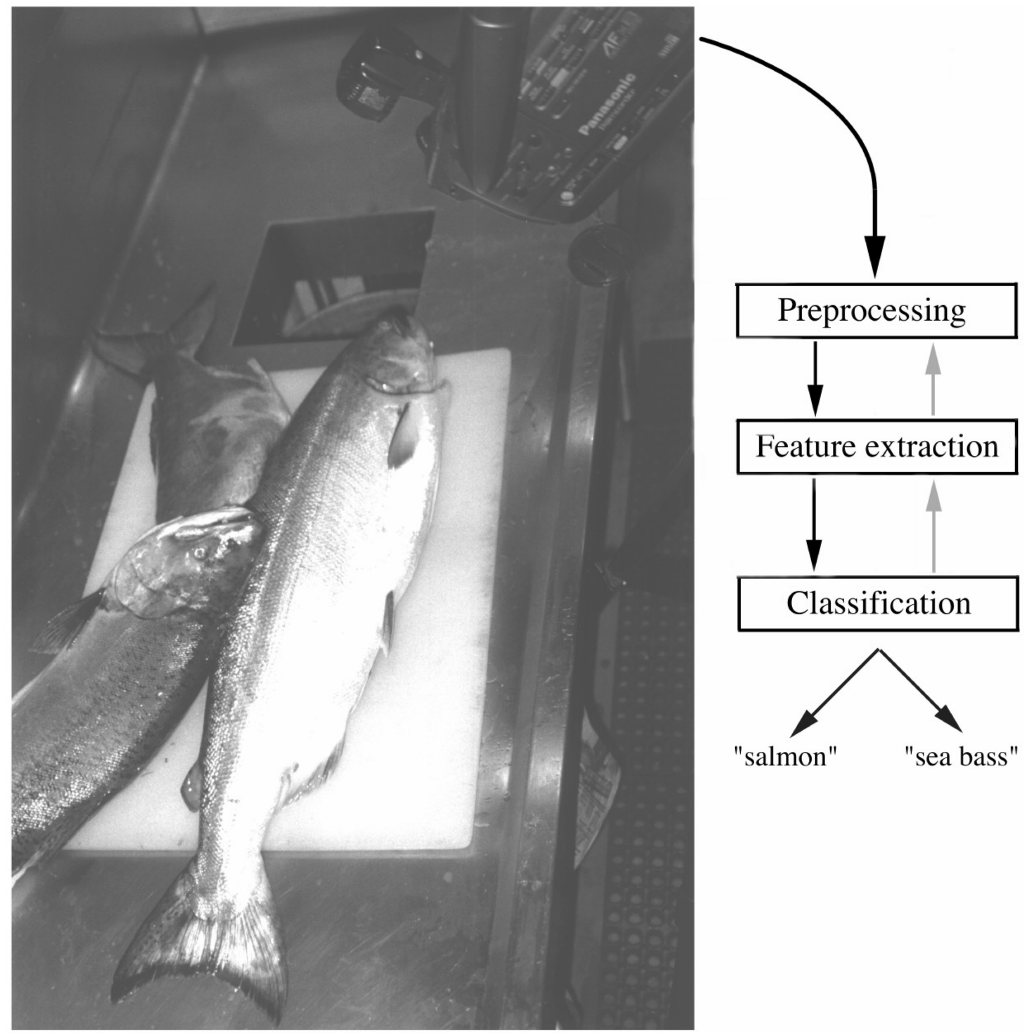
- to design and build machines that can recognize patterns.
  - reliable and accurate pattern recognition by machine would be useful.
    - speech recognition
    - fingerprint identification
    - optical character recognition
    - DNA sequence identification.



  - deeper understanding and appreciation for pattern recognition systems in the natural world – particularly in humans – is required
    - the design efforts may be influenced by *knowledge of how these are solved in nature*
      - both in algorithms and in the special-purpose hardware design

An Example

- **<u>A fish-packing plant</u>:**
  - separating sea bass from salmon using optical sensing automatically.
- *preprocessor*
  - to simplify subsequent operations without losing relevant information
  - segmentation
- *feature extractor*
  - to reduce the data by measuring certain *features* or *properties*
- *classifier*
  - evaluates the evidence presented and makes a final decision as to the species



An Example

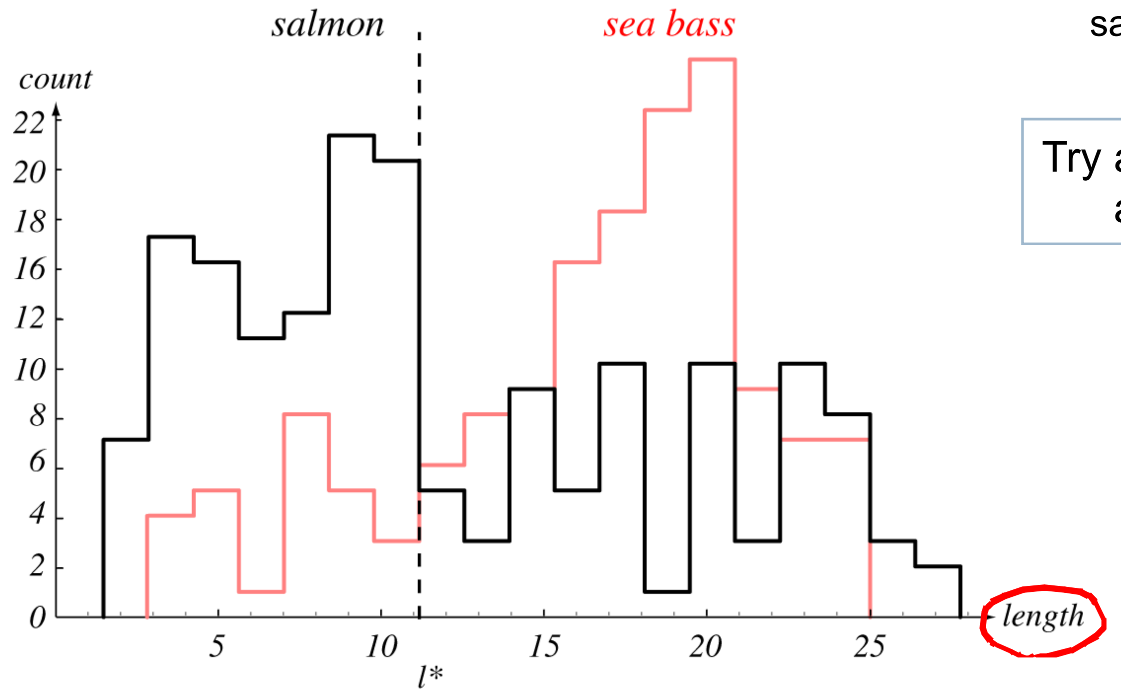
- Features
  - physical differences between the two types of fish
    - length, lightness, width, number and shape of fins, position of mouth….
  - noise or variations in the images
    - variations in lighting, position of the fish
- Model
  - different descriptions in mathematical form
  - Suppose somebody at the plant tell us that a sea bass is generally longer than a salmon:
    - tentative models for the fish: *sea bass have some typical length, and this is greater than that for salmon*

An Example

- Training Samples
  - to choose some critical value we could obtain some design or training samples of the different types of fish

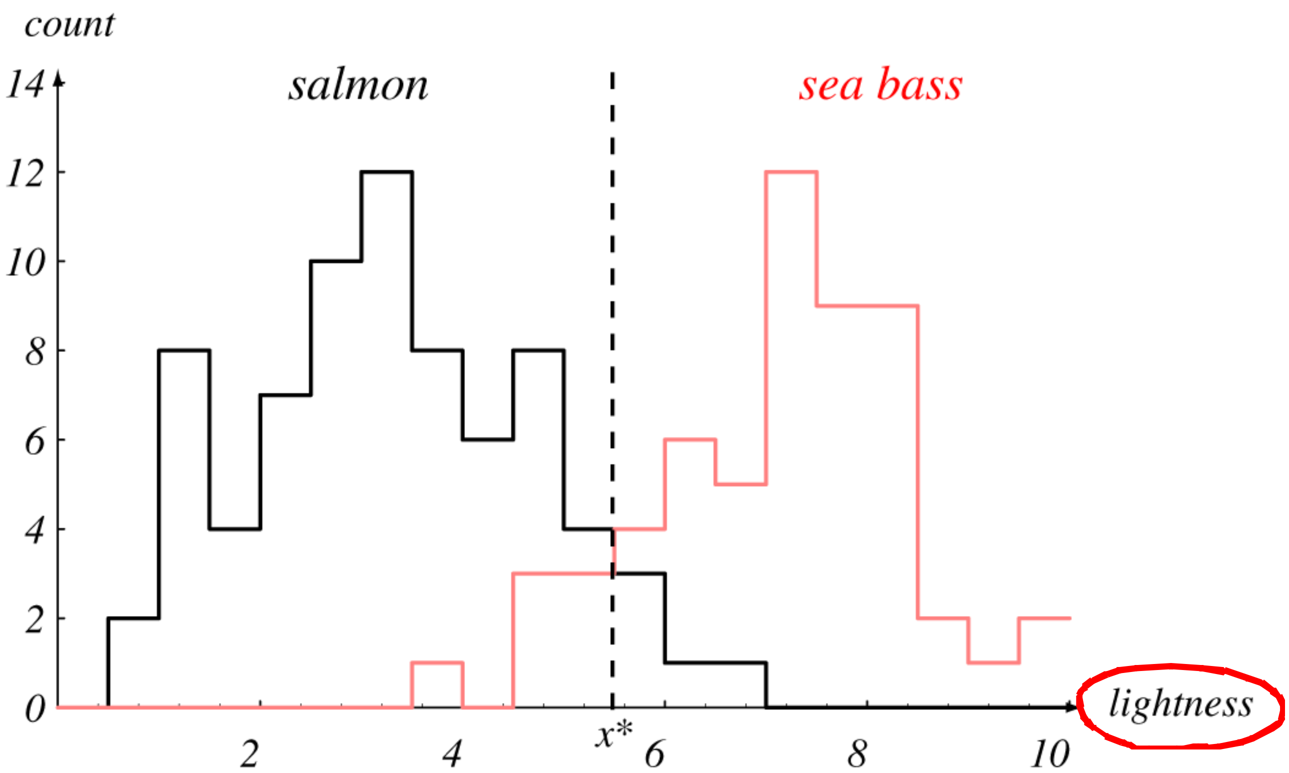
the single criterion is quite poor!

sea bass are somewhat longer than salmon *on average*



Try another feature:
average lightness of the fish

An Example



Much better?

An Example

- Cost
  - salmon in cans labeled "sea bass" vs. sea bass in cans labeled "salmon"
    - consequences of the decisions are not equally costly
    - proper setting of the decision threshold

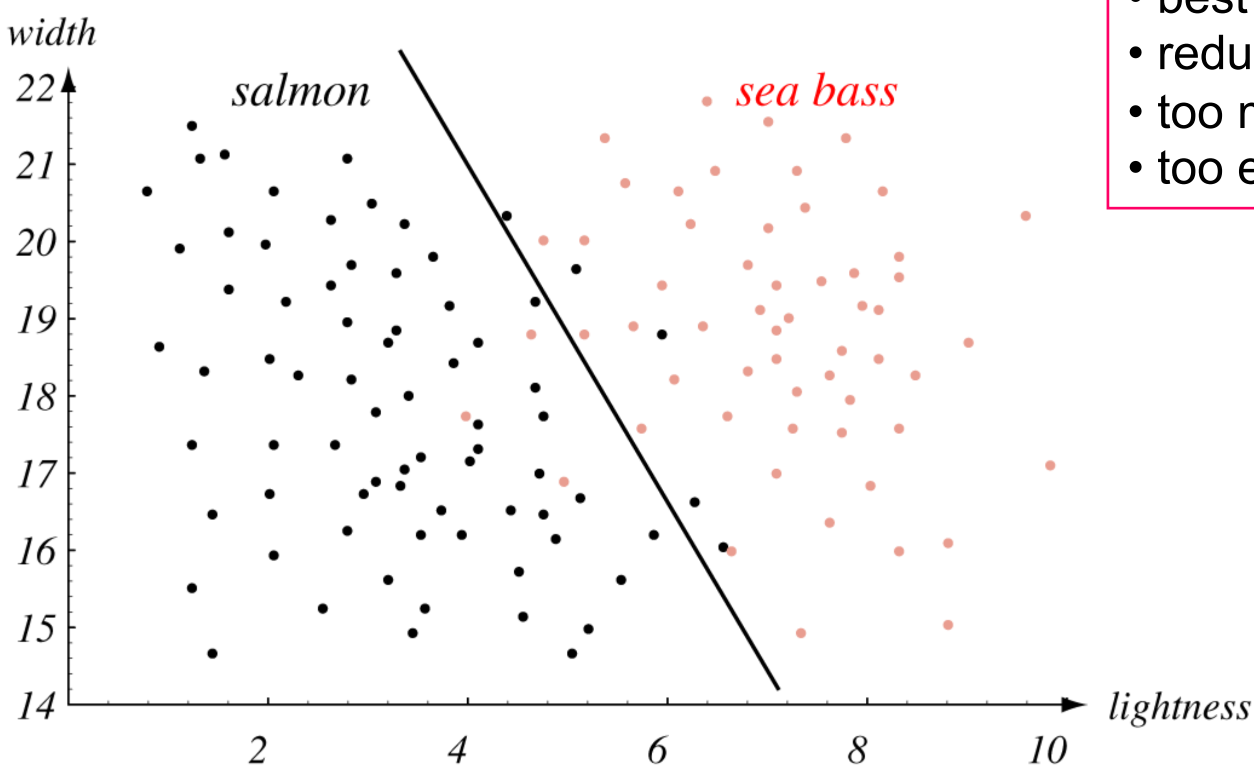
- Decision Theory
  - the central task: to make a decision rule so as to minimize such a cost.
  - to improve recognition, we must resort to the use of more than one feature at a time.

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

x_1: the lightness
x_2: the width

An Example

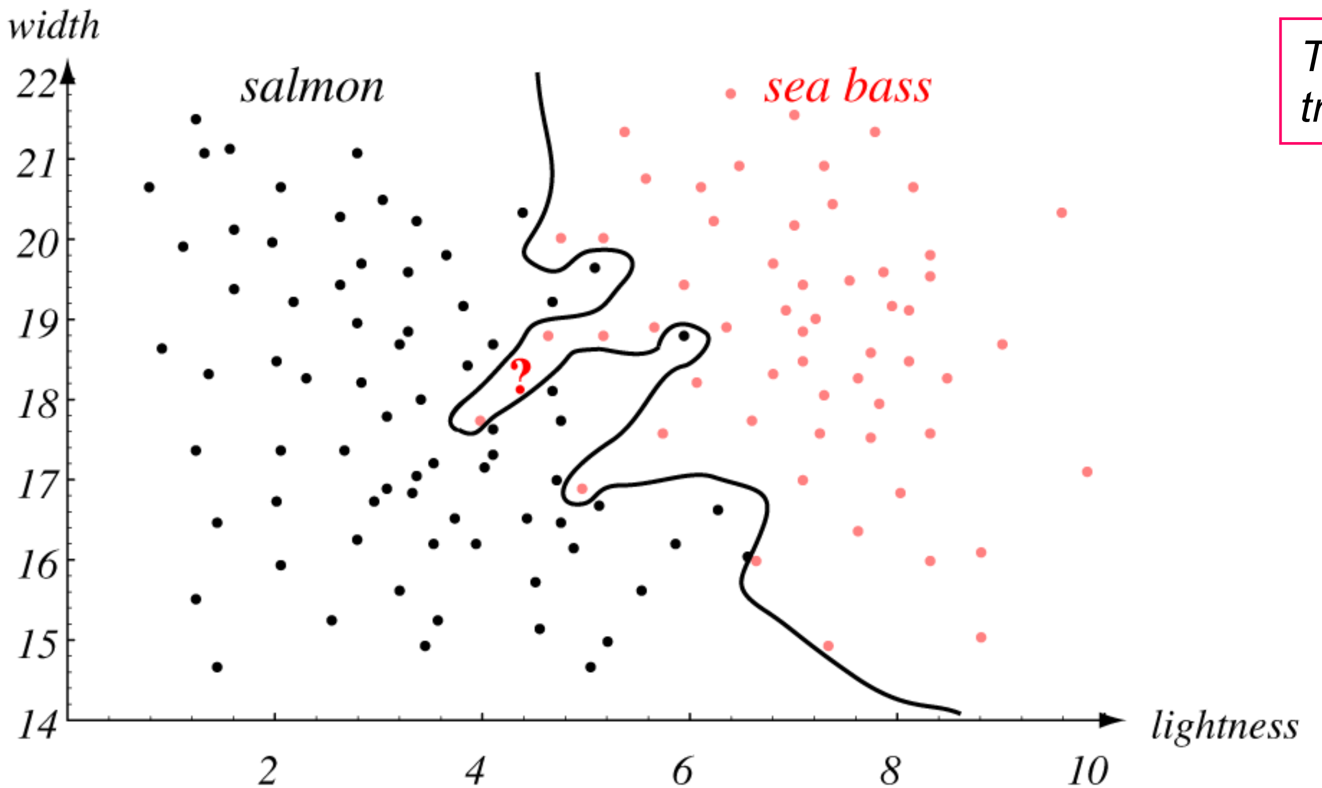
- Decision boundary
  - to partition the feature space



- best features?
- redundant features?
- too many features?
- too expensive to extract?

An Example

- **Generalization**
  - Will the designed classifier work when *novel* patterns are presented?

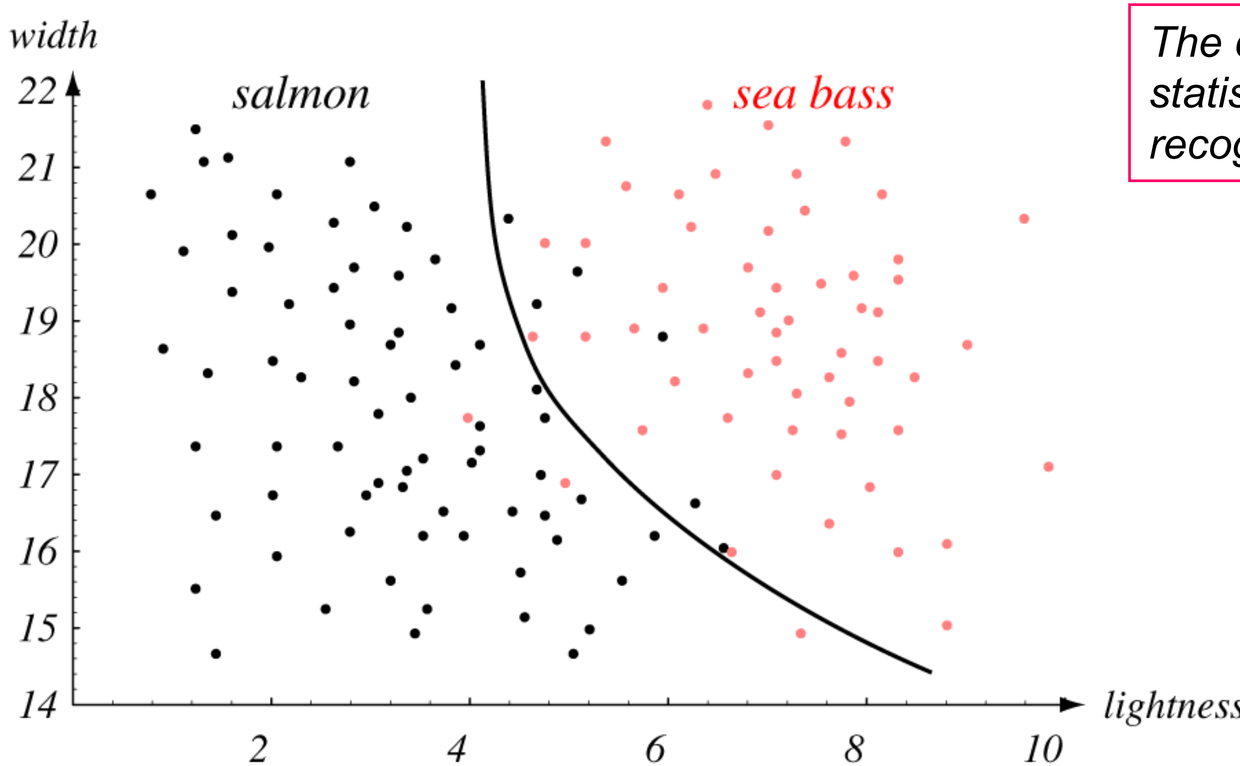


Tuned to the particular training samples

An Example

- How would our system automatically determine that the *simple curve (below)* is preferable to
  - the straight line *or* the complicated boundary ?

The central problems in statistical pattern recognition



An Example

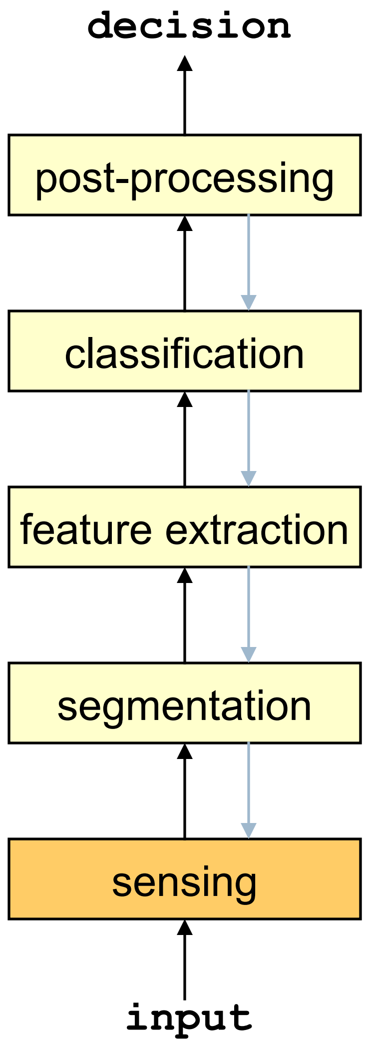
- Good representation by careful feature selection
  - We seek a representation in which the patterns that lead to the same action are close to one another, yet far from those that demand a different action.
  - Additional desirable characteristics:
    - small number of features
      - simpler decision regions
      - easier to train
    - robustness
      - Insensitive to noise or other errors

- Analysis by Synthesis
  - insufficient training data
  - to incorporate knowledge of the problem domain
    - a model of how each pattern is generated

An Example

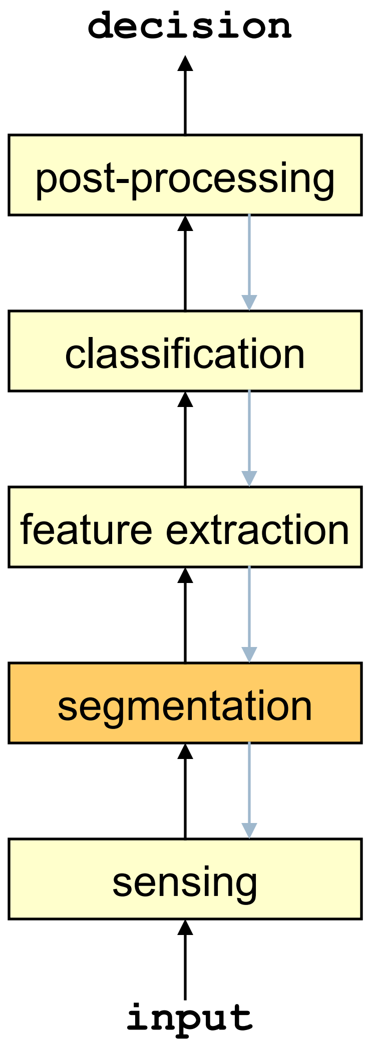
- Related fields
  - image processing
  - associative memory
  - regression
  - interpolation
  - density estimation

Pattern Recognition Systems



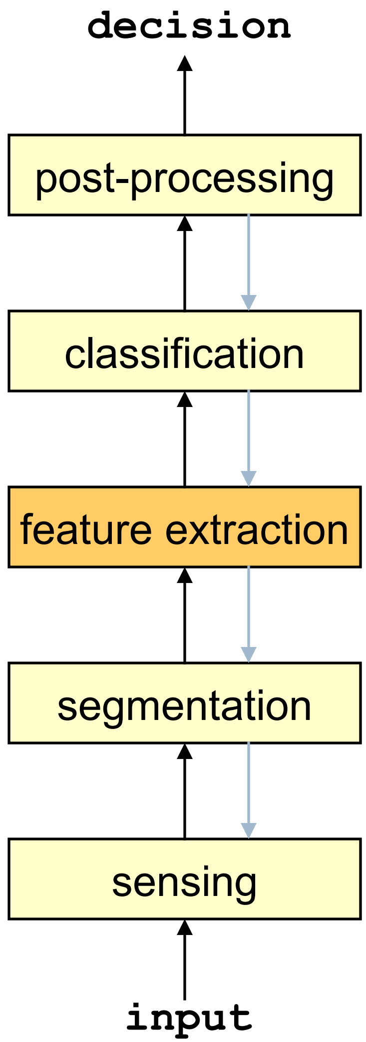
- Sensing
  - input to a pattern recognition system
    - camera
    - microphone array
  - characteristics and limitations of the transducer
    - bandwidth
    - resolution
    - sensitivity
    - distortion
    - signal-to-noise ratio
    - latency

Pattern Recognition Systems



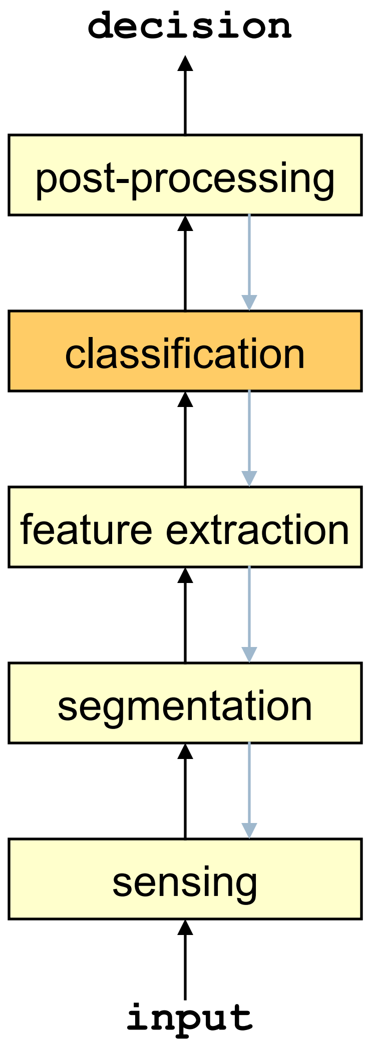
- Segmentation
  - the individual patterns have to be segmented
    - segmentation before categorized?
    - categorized before segmentation?
  - one of the deepest problems in PR
    - speech recognition
    - i vs. =
  - Mereology
    - the problem of *subsets* and *supersets*
    - the study of part/whole relationship

Pattern Recognition Systems



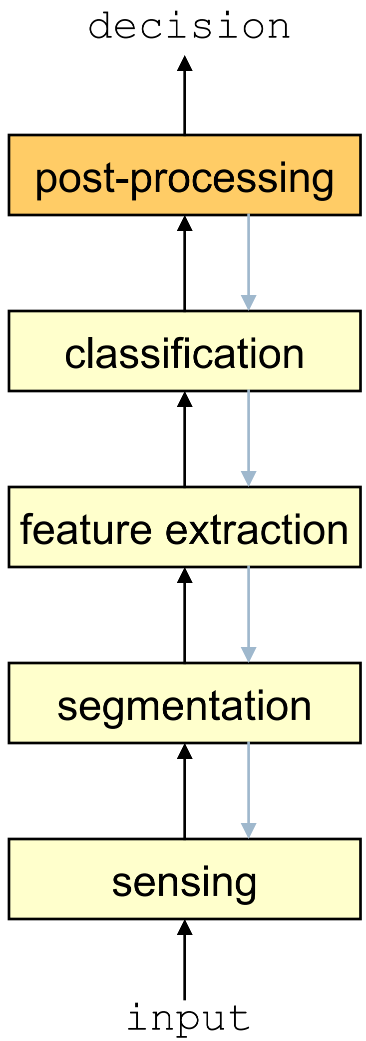
- Feature extraction
  - the boundary between feature extraction and classification is somewhat arbitrary.
  - goal: to characterize a object to be recognized by measurement whose values are very similar for objects in the same category, and very different for objects in different categories
  - invariant features
    - translation, rotation, scale
  - occlusion, projective distortion, rate (in speech), deformation
  - feature selection
    - select the most valuable features from a larger set of candidate.

Pattern Recognition Systems



- Classification
  - perfect classification performance is often impossible
    - to determine *the probability* for each of the possible categories
  - *variability* of feature values for objects in the same category may be due to
    - complexity and noise
  - noise (in general terms)
    - Any property of the sensed pattern which is not due to the true underlying model but instead to randomness in the world or the sensors.
  - to determine the values of all of the features for a particular input is not always possible.
    - In our example, it may not be possible to determine the width of the fish because of *occlusion* by another fish.

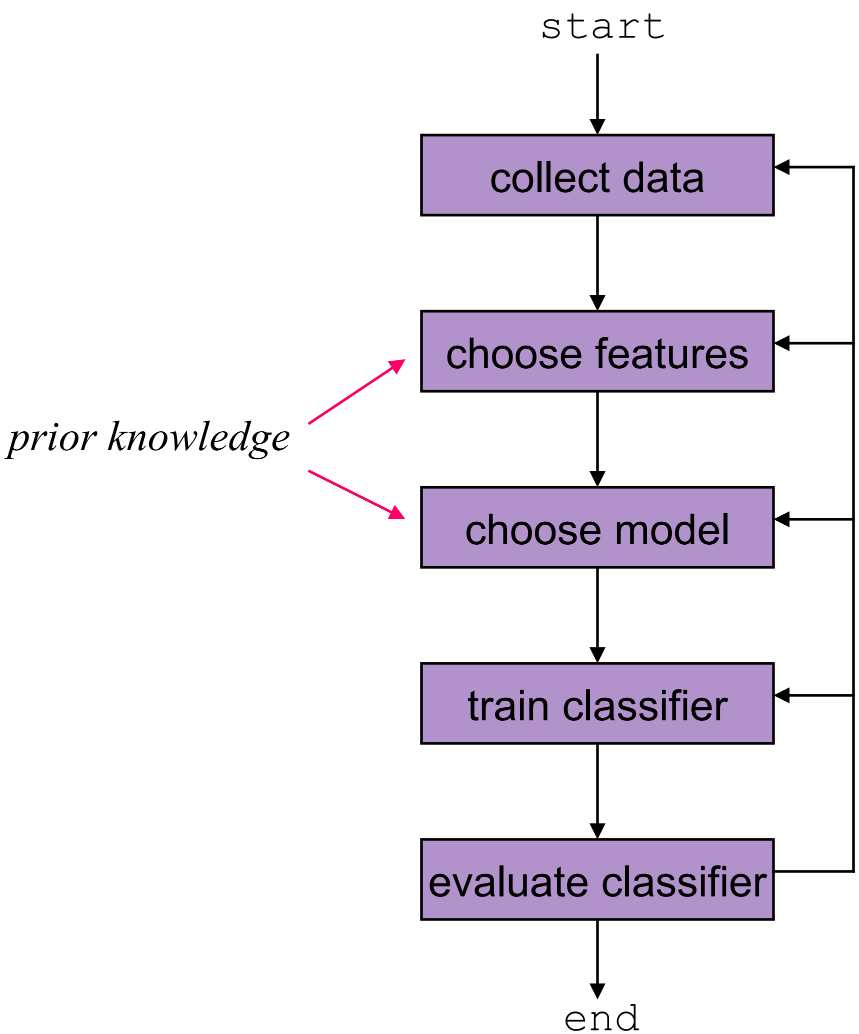
Pattern Recognition Systems



- **Post-processing**
  - The post-processor uses the output of the classifier to decide on the recommended action.
  - error rate
    - the percentage of new patterns that are assigned to the wrong category.
    - minimum-error-rate classification
    - to minimize risk is the better recommendation.
  - context
    - input-dependent information other than from the target pattern itself.

      THE/CHAT
    - context can be highly complex and abstract.
  - multiple classifiers
    - speech recognizer: acoustic recognition + lip reading
    - agreement/disagreement

The Design Cycle



computational complexity:
to design an excellent recognizer, but not within the engineering constraints?

overfitting: perfect classification of the training sample

Learning and Adaptation

- Supervised Learning
  - A teacher provides a category label or cost for each pattern in a training set, and seeks to reduce the sum of the costs for these patterns.
- Unsupervised Learning
  - no explicit teacher
  - The system forms clusters or natural groupings of the input patterns.
- Reinforcement Learning
  - *Learning with a critic*
  - No desired category signal is given; instead only teaching feedback is that the tentative category is *right or wrong*.

Conclusion

- Pattern Recognition: looks hard to learn?

- The good news is at least threefold:
  - There is an existence proof that many of the problems can indeed be solved
    - as demonstrated by humans and other biological systems.
  - Mathematical theories solving some of these problems have in fact been discovered
  - There remain many fascinating unsolved problems providing opportunities for progress.