Outline

- Introduction to Object Recognition
- Decision Theoretic Methods for Object Recognition
  - Object Matching

Introduction to Object Recognition

- Goal is to recognize the existince of certain objects of interest in the image
- We look at the descriptors of certain objects in the image, and define "pattern" as an arrangement of these desriptors
- Some patterns will have common features, and these pattern form a "pattern class"
- Recognition is the task of automatically assigning each pattern to one of these classes
- Exampe: Pattern $\rightarrow$ objects with a certain area, pattern class $\rightarrow$ objects with area 3, objects with area 5, etc, classification $\rightarrow$ deciding that an object belongs to one of these groups

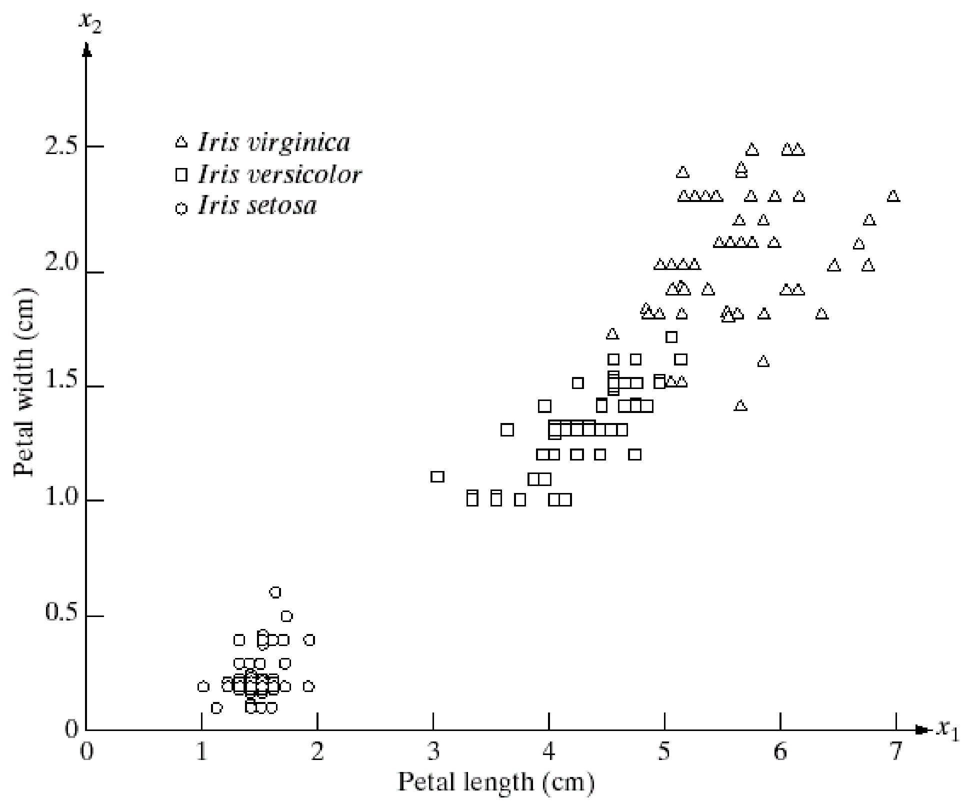
Introduction to Object Recognition (Cont.)

- We can represent patterns using one of three ways: vectors, strings, and trees
- Let us have a look at the vector representation:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

- Here $\boldsymbol{x}$ represents a pattern with n features denoted by x_i, $i = 1, \ldots, n$
- Example: pattern $\rightarrow$ flower, features $\rightarrow$ widths and lengths of the flower's petal, that is $\boldsymbol{x}$ has two components

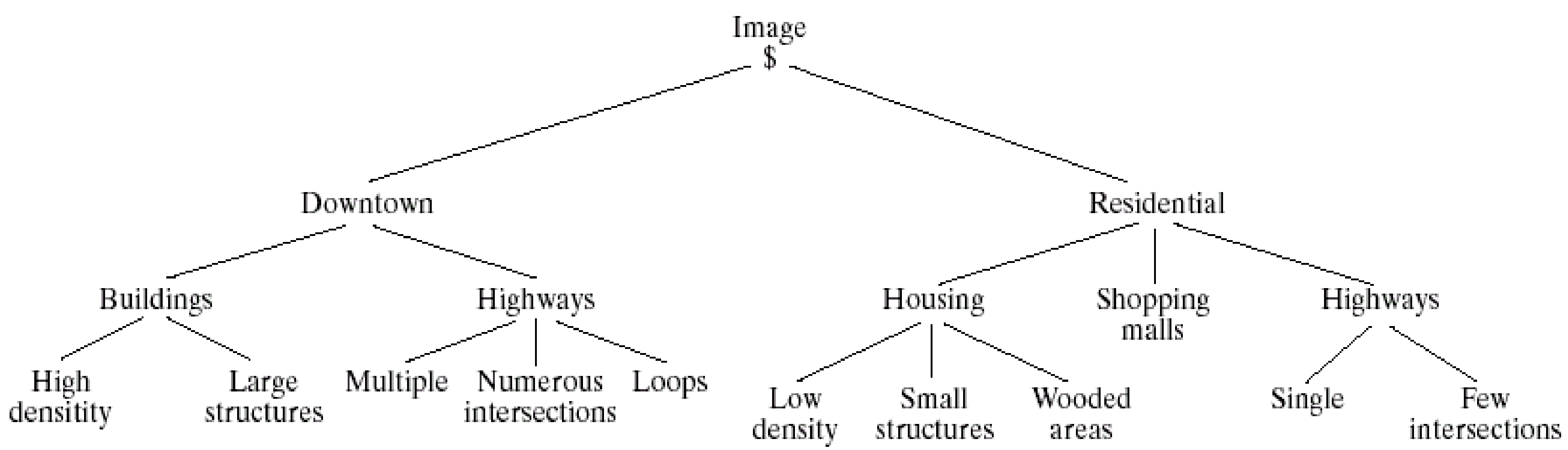
Introduction to Object Recognition (Cont.)



An example of pattern distribution on a space of features

Introduction to Object Recognition (Cont.)

- The second way to represent patterns is based on strings
- We can for example use repeating structures and use a series of these structures as a means of representing patterns with a certain boundary
- Finally, we can represent patterns using trees



Decision Theoretic Methods for Object Recognition

- Decision theoretic methods are mathematical tools to perform object/pattern classification
- Given a vector $\boldsymbol{x}$ representing a pattern we create decision functions d such that

$$d_i(\boldsymbol{x}) > d_j(\boldsymbol{x})$$

  when i is different than j, and $\boldsymbol{x}$ belongs to pattern i

- Then we switch our decision from i to j when the decision function values are equal
- That is $d_i(\boldsymbol{x}) = d_j(\boldsymbol{x})$ form the boundary between the classes i and j

Decision Theoretic Methods for Object Recognition: Matching

- We choose a single pattern that represents a particular class
- For a candidate to be classified, we compare to all representative patterns and choose the class that optimizes a cost function
- We can use the following cost functions for this purpose
  - Minimize the norm of the difference between the representative pattern vector and candidate pattern vector
  - Maximize the correlation between the representative pattern vector and candidate pattern vector

Decision Theoretic Methods for Object Recognition: Matching (Cont.)

- Let us choose the mean of a class as its representative function, then the decision function is defined as

$$d'_j(\boldsymbol{x}) = ||\boldsymbol{x} - \boldsymbol{m}_j||$$

where

$$\boldsymbol{m}_j = \frac{1}{N_j} \sum_{x \epsilon W_j} x_j$$

- Minimizing this cost function is equivalent to maximizing

$$d_j(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{m}_j - \frac{1}{2} \boldsymbol{m}_j^{\mathrm{T}} \boldsymbol{m}_j$$

- We simply calculate $d_j(\boldsymbol{x})$ for all j and select the class that corresponds to the maximum

Decision Theoretic Methods for Object Recognition: Matching (Cont.)

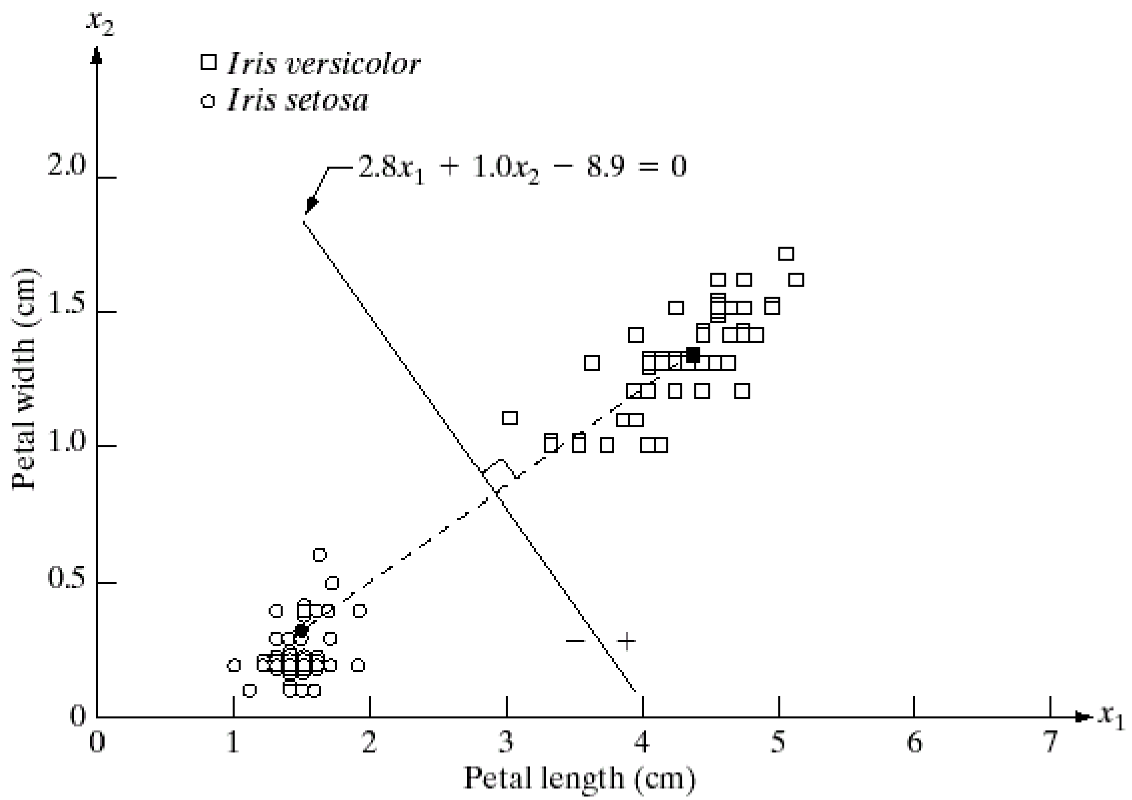
- The surface separating two classes i and j are given by

$$d_i(\boldsymbol{x}) - d_j(\boldsymbol{x}) = 0$$

resulting in

$$\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{m}_i - \boldsymbol{m}_j) - \frac{1}{2}(\boldsymbol{m}_i - \boldsymbol{m}_j)^{\mathrm{T}}(\boldsymbol{m}_i + \boldsymbol{m}_j) = 0$$

Decision Theoretic Methods for Object Recognition: Matching (Cont.)



An example of classification with norm of the difference

Decision Theoretic Methods for Object Recognition: Matching (Cont.)

- The second cost function for pattern classification we consider is the correlation
- The correlation between two pattern vectors is given by

$$d_j(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{m}_j$$

- We can select the class that maximizes this correlation
- However, alternatives that include normalizations can be used such as the correlation coefficient given by

$$d_j(\boldsymbol{x}) = \frac{[\boldsymbol{x} - \bar{\boldsymbol{x}}]^{\mathrm{T}}[\boldsymbol{m}_j - \bar{\boldsymbol{m}}_j]}{[\boldsymbol{x} - \bar{\boldsymbol{x}}]^{\mathrm{T}}[\boldsymbol{x} - \bar{\boldsymbol{x}}] \times [\boldsymbol{m}_j - \bar{\boldsymbol{m}}_j]^{\mathrm{T}}[\boldsymbol{m}_j - \bar{\boldsymbol{m}}_j]}$$