Pattern Recognition Basics

Xiaojun Qi

Pattern Class vs. Pattern

- Pattern Class: It is a category determined by some given common attributes.
- Pattern: It is the description of any member of a category representing a pattern class. In other words, it is the description of an object.
- We may divide our acts of recognition into two major types:
  – The recognition of concrete items
  – The recognition of abstract items

Pattern Recognition

- Pattern recognition can be defined as the categorization of input data into identifiable classes via the extraction of significant features or invariant attributes of the data from a background of irrelevant detail.

Two Categories of Pattern Recognition

- The study of the pattern recognition capability of human beings and other living organisms. (Psychology, Physiology, and Biology)
- The development of theory and techniques for the design of devices capable of performing a given recognition task for a specific application. (Engineering, Computer, and Information Science)

Task of Classification	Input Data	Output Response
Character Recognition	Optical signals or strokes	Name of character
Speech Recognition	Acoustic waveforms	Name of word
Speaker Recognition	Voice	Name of speaker
Weather Prediction	Weather maps	Weather forecast
Medical Diagnosis	Symptoms	Disease
Stock Market Prediction	Financial news and charts	Predicted market ups and downs.

Fundamental Issues in Pattern Recognition System Design

- The representation of input data which can be measured from the objects to be recognized.
  – The pattern vectors contain all the measured information available about the patterns.
  – When the measurements yield information in the form of real numbers, it is often useful to think of a pattern vector as a point in an n-dimensional Euclidean space.
  – The set of patterns belonging to the same class corresponds to an ensemble of points scattered within some region of the measurement space.

Fundamental Issues in Pattern Recognition System Design

- The extraction of characteristic features or attributes from the received input data and the reduction of the dimensionality of pattern vectors.
  - Intraset features: characterizing attributes common to all patterns belonging to that class.
  - Interset features: reprenting the differences between pattern classes.

7

Fundamental Issues in Pattern Recognition System Design

- The determination of optimum decision procedures, which are needed in the identification and classification process.
  - If completed *a prior* knowledge about the patterns to be recognized is available, the decision functions may be determined with precision.
  - If only qualitative knowledge about the patterns is available, reasonable guesses of the forms of the decision functions can be made.
  - If there exists little *a priori* knowledge about the patterns to be recognized, a training or learning procedure is needed.

8

Summary

- The patterns to be recognized and classified by an automatic pattern recognition system must possess a set of measurable characteristics.
- Correct recognition will depend on
  - The amount of discriminating information contained in the measurements;
  - The effective utilization of this information.

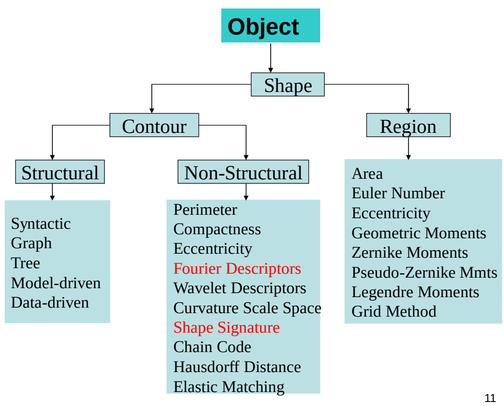
9

Sample Object Features: Shape

- **What features can we get from an Object?**

- Perimeter
- Area
- Eccentricity: The ratio of the major to the minor axis
- Curvature: The rate of change of slope. That is: Use the difference between the slopes of adjacent boundary segments as a descriptor of curvature at the point of intersection of the segments.
- Chain Code…

10



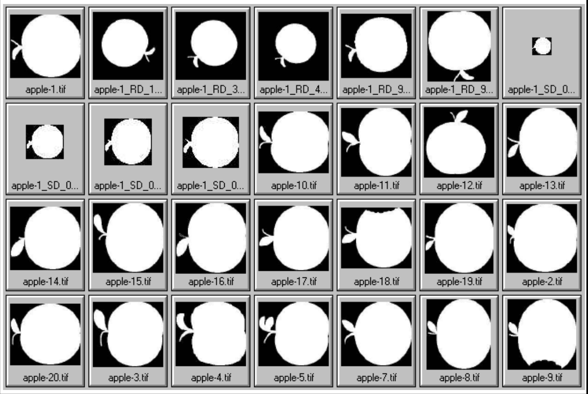
11

Object Representation and Recognition

- Representing a region involves two choices:
  - We can represent the region in terms of its external characteristics (its boundary)
    - Focus on Shape Characteristics.
  - We can represent the region in terms of its internal characteristics (the pixels comprising the region)
    - Focus on regional properties such as color and texture.
- Describe the region based on the chosen representation.
  - A region may be represented by its boundary, and the boundary can be described by features such as:
    - Length,
    - The orientation of the straight line joining its extreme points,
    - The number of concavities in the boundary.

12

The features selected as descriptors should be as insensitive as much as possible to variations in size, translation, and rotation.

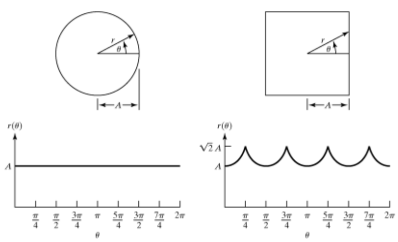


Boundary Features (1) – Shape Signatures

- A signature is a 1-D functional representation of a boundary and may be represented in various ways.
- Regardless of how a signature is generated, the basic idea is to reduce the boundary representation to a 1-D function, which presumably is easier to describe than the original 2-D boundary.

14

A Shape Signature Example



Signature: Distance versus angle

It is invariant to translation. Normalization with respect to rotation can be achieved by finding a way to select the same starting point to generate the signature, regardless of the shape's orientation. Scale all functions so that the values always span the same range of values [0, 1]

15

Other Shape Signatures

- Complex Coordinates
- Central Distance
- Central Complex Coordinates
- Chordlength
- Cumulative Angular Function
- Curvature Function
- Area Function

➔ Normally, we do not consider boundaries with self-intersections, boundaries with deep, narrow concavities, or boundaries with thin, long protrusions.

16

Shape Signatures – Complex Coordinates

- The boundary can be represented as the sequence of coordinates s(t) = [x(t), y(t)] for t = 0, 1, 2, …, N-1, where $x(t) = x_t$ and $y(t) = y_t$; (x_t, y_t)'s are encountered in traversing the boundary in the counterclockwise direction and N is the total number of points on the boundary.

$Z(t) = x(t) + iy(t)$

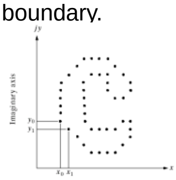


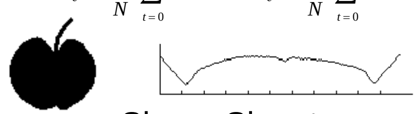
FIGURE 11.13 A digital boundary and its representation as a complex sequence. The points (x_0, y_0) and (x_1, y_1) shown are (arbitrarily) the first two points in the sequence.

7

Shape Signatures – Central Distance

$$r(t) = ([x(t) - x_c]^2 + [y(t) - y_c]^2)^{1/2}$$

$$x_c = \frac{1}{N}\sum_{t=0}^{N-1} x(t), \quad y_c = \frac{1}{N}\sum_{t=0}^{N-1} y(t)$$



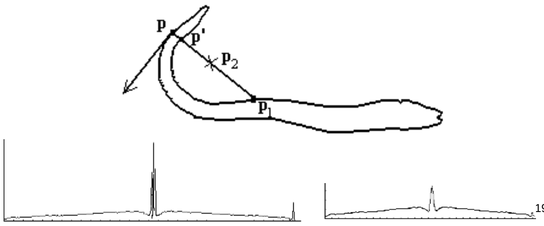
Shape Signatures --Central Complex Coordinates

$$z(t) = [x(t) - x_c] + i[y(t) - y_c]$$

18

Shape Signatures – Chordlength

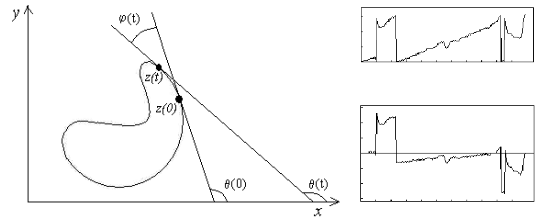
- R*(t) = length of chord in object perpendicular to tangent at p, as a function of p.
- The *chord length function r*(t)* is derived from shape boundary without using any reference point



19

Shape Signatures – Cumulative Angular Function

- It is also called turning angle function.
- $\varphi(t) = [\theta(t) - \theta(0)] \bmod (2\pi)$



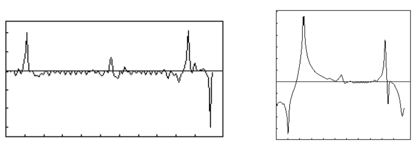
20

Shape Signatures – Curvature Function

- $K(t) = \theta(t) - \theta(t-1)$

$$\theta(t) = \arctan \frac{y(t) - y(t+w)}{x(t) - x(t+w)}$$

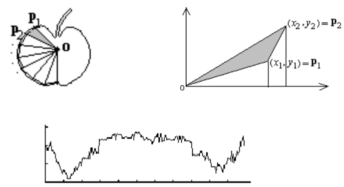
w is the jumping step in selecting next pixel



21

Shape Signatures – Area Function

$$A(t) = \frac{1}{2} | x_1(t) y_2(t) - x_2(t) y_1(t) |$$



22

Boundary Features (2) – Fourier Descriptors

- Fourier Transform of the Signature s(t):

$$u_n = \frac{1}{N} \sum_{t=0}^{N-1} s(t) e^{-j2\pi nt/N}$$

$$\text{for} \quad n = 0, 1, \ldots, N-1$$

The complex coefficients u_n are called the Fourier descriptors of the boundary, and are denoted as FDn.

23

Boundary Features (2) – Fourier Descriptors

- The inverse Fourier transform of these coefficients restores s(t). That is:

$$s(t) = \sum_{n=0}^{N-1} u_n e^{j2\pi nt/N} \quad \text{for } t = 0, 1, \ldots, N-1$$

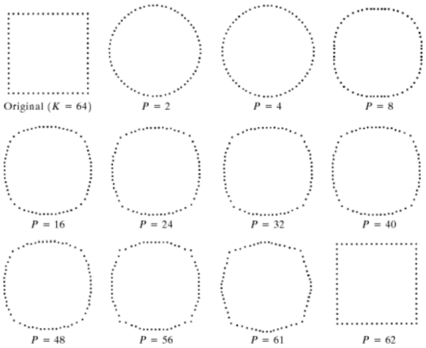
- Suppose that only the first P coefficients are used (that is, setting $u_n = 0$ for n > P-1). The result is the following approximation to s(k):

$$\hat{s}(t) = \sum_{n=0}^{P-1} u_n e^{j2\pi nt/N}$$

for t = 0, 1, 2, …, N-1.

24

FIGURE 11.14 Examples of reconstruction from Fourier descriptors. P is the number of Fourier coefficients used in the reconstruction of the boundary.



The goal is to use a few Fourier descriptors to capture the gross essence of a boundary. These coefficients carry shape information and can be used as the basis for differentiating between distinct boundary shapes.

Transformation	Boundary	Fourier Descriptor
Identity	$s(k)$	$a(u)$
Rotation	$s_r(k) = s(k)e^{j\theta}$	$a_r(u) = a(u)e^{j\theta}$
Translation	$s_t(k) = s(k) + \Delta_{xy}$	$a_t(u) = a(u) + \Delta_{xy}\delta(u)$
Scaling	$s_s(k) = \alpha s(k)$	$a_s(u) = \alpha a(u)$
Starting point	$s_p(k) = s(k - k_0)$	$a_p(u) = a(u)e^{-j2\pi k_0 u/K}$

$$s_t(k) = [x(k) + \Delta x] + j[y(k) + \Delta y]$$

$$s_p(k) = x(k - k_0) + jy(k - k_0)$$

1) Magnitude |FDn| is translation and rotation invariant
2) |FD0| carries scale-information
3) "Low-frequency" terms (*t* small): smooth behavior
4) "High-frequency" terms (*t* large): jaggy, bumpy behavior

Normalized Fourier Descriptor

$$f = \left[\frac{|FD_1|}{|FD_0|}, \frac{|FD_2|}{|FD_0|}, \ldots, \frac{|FD_m|}{|FD_0|}\right]$$

Why?

When two shapes are compared, m=N/2 coefficients are used for central distance, curvature and angular function.

m=N coefficients are used for complex coordinates.

$$d = \sqrt{\sum_{i=1}^{m} (f_i^q - f_i^t)^2}$$

where $f_q = (f_q^1, f_q^2, \ldots, f_q^m)$ and $f_t = (f_t^1, f_t^2, \ldots, f_t^m)$

are the feature vectors of the two shapes respectively.

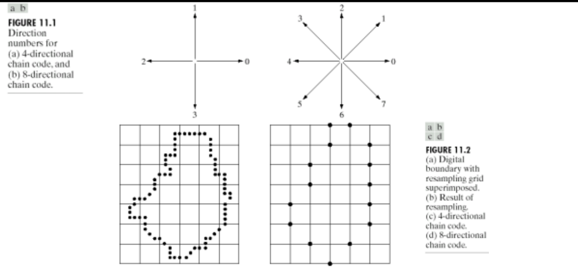
- Complex FFT example:

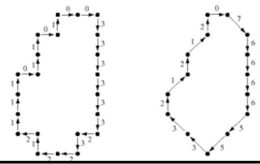
A = [2 3 4 4] ;
B = [1 3 4 7] ;
C = A + B * i ;
fft(A) = [13 -2 + i -1 -2 – i] ;
fft(B) = [15 -3 + 4i -5 -3 - 4i] ;
fft(C) = [13 +15i -6 - 2i -1 - 5i 2-4i] ;

- Criteria for shape representation
  - Rotation, scale and translation Invariant
  - Compact & easy to derive
  - Perceptual similarity
  - Robust to shape variations
  - Application Independent
- FD satisfies all these criteria
- Problem
  - Different shape signatures can be used to derive FD, which is the best?

Boundary Features (3) – Chain Codes

- A chain code could be generated by following a boundary in a clockwise direction and assigning a direction to the segments connecting every pair of pixels.
- Disadvantage:
  - The resulting chain of codes tends to be quite long
  - Any small disturbance along the boundary due to noise or imperfect segmentation cause changes in the code that may not be related to the shape of the boundary.





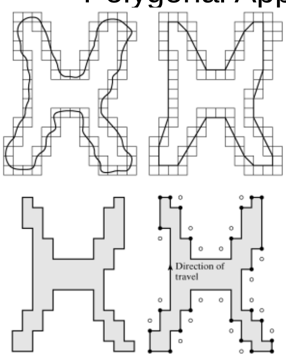
The accuracy of the resulting code representation depends on the spacing of the sampling grid.

31

- The chain code of a boundary depends on the starting point.
  - Treat the chain code as a circular sequence of direction numbers and redefine the starting point so that the resulting sequence of numbers forms an integer of minimum magnitude.
- Rotation Invariance:
  - Normalize for rotation by using the first difference of the chain code instead of the code itself. This difference is obtained by counting the number of direction changes (in a counterclockwise direction) that separate two adjacent elements of the code.
- Scaling Invariance:
  - Normalize for scaling is achieved by altering the size of the resampling grid.

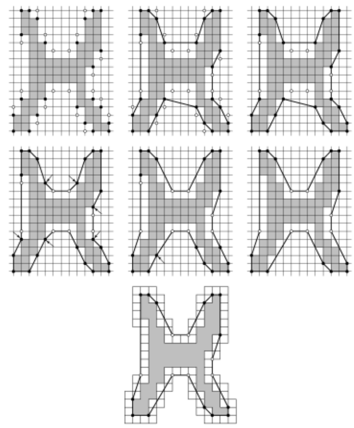
32

Boundary Features (4) – Polygonal Approximation



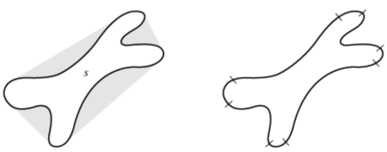
The goal of polygonal approximation is to capture the “essence” of the boundary shape with the fewest possible polygonal segments. → Minimum Perimeter Polygons Approach

33



34

Boundary Features (5) – Boundary Segments



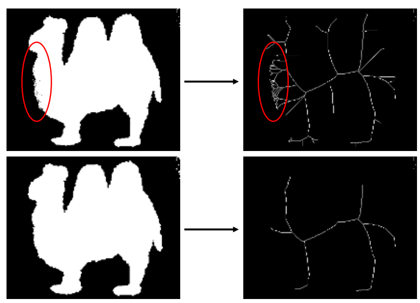
Check out Matlab built-in function:

regionprops

35

Boundary Features (6) – Skeletons

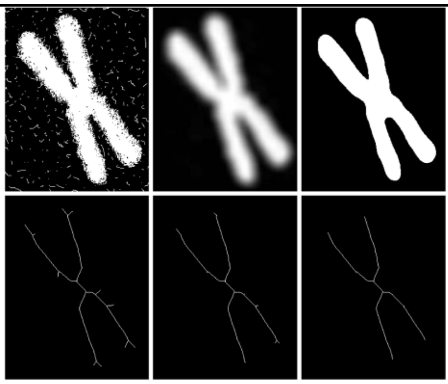
- “eat away” pixels from all sides until there is one left → Sensitive to noise



On original

After 1 *dilation* (adding 1 pixel at border)

36



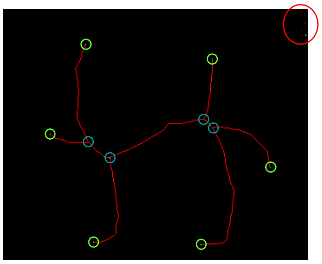
a b c
d e f

FIGURE 11.13 (a) Segmented human chromosome. (b) Image smoothed using a 25 × 25 Gaussian averaging mask with sig = 15. (c) Thresholded image. (d) Skeleton. (e) Skeleton after 8 applications of spur removal. (f) Result of 7 additional applications of spur removal.

37

Boundary Features (7) – Skeleton Features

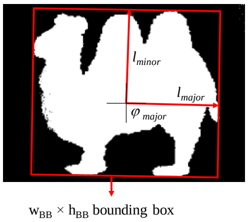
- Count #endpoints
- Count #branches

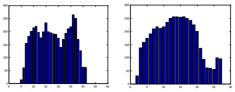


38

Boundary Features ()

- **Simple features**
  - Area : A
  - Circumference: r
  - Euler's number: #parts - #holes
  - Direction: φ_{major}
  - Eccentricity: $||l_{major}|| / ||l_{minor}||$
  - Elongatedness: w_{BB} / h_{BB}
  - Rectangularity: A / A_{BB}
  - Compactness: r^2 / A
  - Greyvalue/colour/texture statistics
  - Projections
- Not all invariant

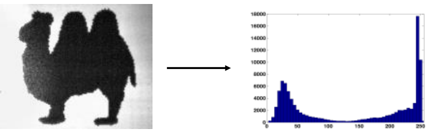




39

Regional Color Features

- **Histogram** (grey-scale images)



- Invariant to translation, rotation, small variations
- Normalized histogram is invariant to scale
- Not very sensitive to noise
- But: removes a lot of information!

Which features should we choose?

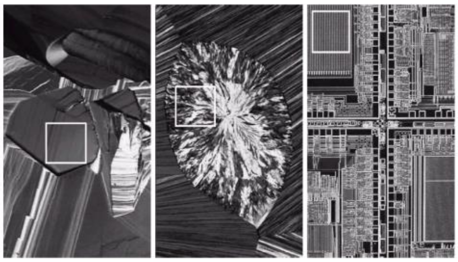
40

Regional Texture Features

- The three principal approaches used in image processing to describe the texture of a region are statistical, structural, and spectral.
  - Statistical approaches yield characterizations of textures as smooth, coarse, grainy, and so on.
  - Structural techniques deal with the arrangement of image primitives, such as the description of texture based on regularly spaced parallel lines.
  - Spectral techniques are based on properties of the Fourier spectrum and are used primarily to detect global periodicity in an image by identifying high-energy, narrow peaks in the spectrum.

41

Regional Texture Features – Statistical Approaches



a b c

FIGURE 11.22 The white squares mark, from left to right, smooth, coarse, and regular textures. These are optical microscope images of a superconductor, human cholesterol, and a microprocessor. (Courtesy of Dr. Michael W. Davidson, Florida State University.)

42

TABLE 11.2
Texture measures for the subimages shown in Fig. 11.22.

Texture	Mean	Standard deviation	R (normalized)	Third moment	Uniformity	Entropy
Smooth	82.64	11.79	0.002	−0.105	0.026	5.434
Coarse	143.56	74.63	0.079	−0.151	0.005	7.783
Regular	99.72	33.73	0.017	0.750	0.013	6.674

Mean : $m = \sum_{i=0}^{L-1} z_i p(z_i)$

Standard Deviation : $\sigma = \sqrt{\sum_{i=0}^{L-1} (z_i - m)^2 p(z_i)}$

R : $R = 1 - \frac{1}{1+\sigma^2(z)}$

Third Moment : $\mu_3(z) = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$

Uniformity : $U = \sum_{i=0}^{L-1} p^2(z_i)$

Entropy : $e = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$

Standard Deviation is a measure of gray-level contrast that can be used to establish descriptors of relative smoothness.

R: Normalized variance in the range of [0, 1]

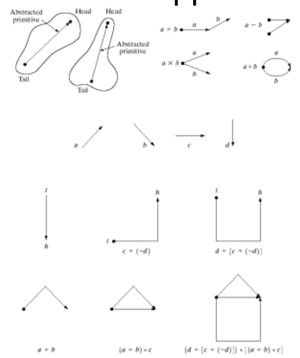
Third moment is a measure of the skewness of the histogram.

Uniformity: Histogram based measure.

Average Entropy is a measure of variability and is 0 for a constant image.

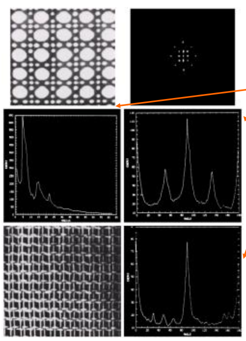
43

Regional Texture Features – Structural Approaches



44

Regional Texture Features – Spectral Approaches



$$S(r) = \sum_{\theta=0}^{\pi} S_\theta(r)$$

$$S(\theta) = \sum_{r=1}^{R_0} S_r(\theta)$$

where S is the spectrum function

r and θ is radius and angle

in the polar coordinates

45

Regional Moment-Based Features

$$m_{pq} = \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} x^p y^q I(x,y)\,dxdy$$

–0[th] order : size
–1[st] order : center-of-mass
–2[nd] order : orientation
–higher order : shape

- Central moments: translation invariant

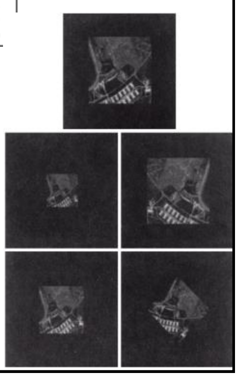
$$\mu_{pq} = \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} x^p y^q I\left(x - \frac{m_{10}}{m_{00}}, y - \frac{m_{01}}{m_{00}}\right) dxdy$$

- Hu moments (7): translation, rotation, scale invariant
- Zernike moments (47): translation, rotation, scale invariant

46

TABLE 11.3
Moment invariants for the images in Figs. 11.25(a)-(e).

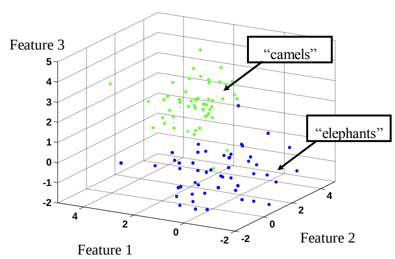
Invariant (Log)	Original	Half Size	Mirrored	Rotated 2°	Rotated 45°
ϕ_1	6.249	6.226	6.919	6.253	6.318
ϕ_2	17.180	16.954	19.955	17.270	16.803
ϕ_3	22.655	23.531	26.689	22.836	19.724
ϕ_4	22.919	24.236	26.901	23.130	20.437
ϕ_5	45.749	48.349	53.724	46.136	40.525
ϕ_6	31.830	32.916	37.134	32.068	29.315
ϕ_7	45.589	48.343	53.590	46.017	40.470



1) Seven moments invariants are calculated for each of these images, and the logarithm of the results are taken to reduce the dynamic range.
2) As the Table shows, the results are in reasonable agreement with the invariants computed for the original image.
3) The major cause of error can be attributed to the digital nature of the data, especially for the rotated images.

Feature Space: Multi-Dimension

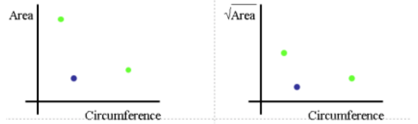
- End result: a *k*-dimensional space,
  - –in which each dimension is a **feature**
  - –containing *N* (labeled) **samples** (objects)



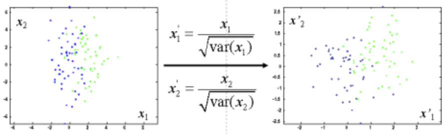
48

Feature Space: Normalization

Different features have different scale



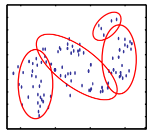
Solution: normalize variance in each direction



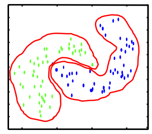
49

Feature Space: Pattern Recognition

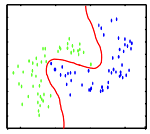
Clustering: find natural groups of samples in unlabelled data



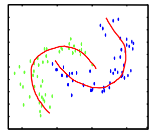
Density estimation: make a statistical model of the data



Classification: find functions separating the classes



Regression: fit lines or other functions to data (not in this course)



What is our basic problem?

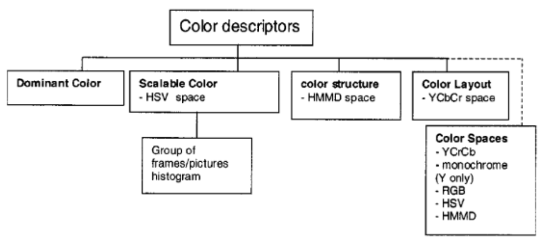
50

Summary

- Features are derived from measurements
- Application-dependent knowledge tells what features are important
- Invariance is important to make discrimination easier
- Recognition:
  - Noise removal
  - Shading removal
  - Segmentation and labeling
  - Features: Simple, Skeletons, Moments, Polygons, Chain Code, Fourier descriptors, ….

51

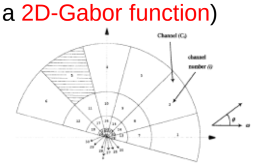
Other Sample Features: MPEG-7 Color Descriptors



52

Other Sample Features: MPEG7 Homogenous Texture Descriptor

- Partition the frequency domain into 30 channels (modeled by a 2D-Gabor function)



- Compute the energy and energy deviation for each channel
- Compute mean and standard variation of frequency coefficients

53

Other Sample Features: MPEG7 Non-Homogenous Texture Descriptor

- Represent the spatial distribution of five types of edges
  - vertical, horizontal, 45°, 135°, and non-directional
- Dividing the image into 16 (4x4) blocks
- Generating a 5-bin histogram for each block
- It is scale invariant

54