Introduction to Computer Vision

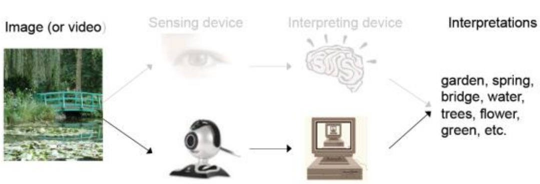
One picture is worth more than
ten thousand words

Xiaojun Qi

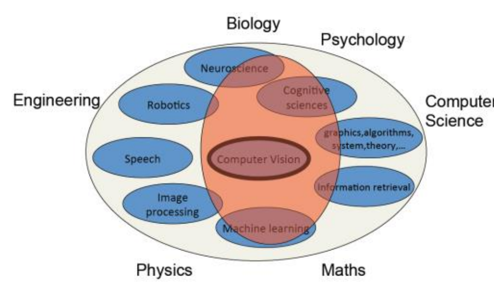
Computer Vision

- Focus on view analysis using techniques from Image Processing (IP), Pattern Recognition (PR), and Artificial Intelligence (AI). It is the area of AI concerned with modeling and replicating human vision using computer software and hardware.

What is Computer Vision?



What Is It Related To?



The Goal of Computer Vision

- To bridge the gap between pixels and “meaning”

What we see

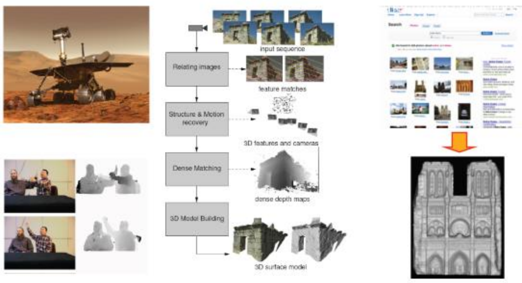
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

What Kind of Information Can We Extract from An Image

- Metric 3D Information
- Semantic Information

Vision as Measurement Device



Pollefeys et al.

Goesele et al.

7

Vision as a Source of Semantic Information

8

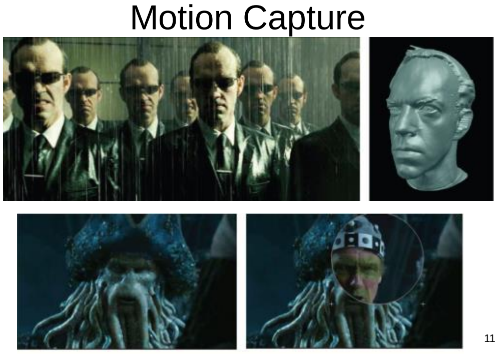
Why Study Computer Vision?

- Vision is useful: Images and video are everywhere

Surveillance and security

Medical and scientific images

Special Effects: Shape and Motion Capture



11

3D Urban Modeling

Bing maps: https://www.microsoft.com/maps/

Google Streetview: https://www.google.com/streetview/

12

Face Detection

- Many digital cameras now detect faces
  - Canon, Sony, Fuji

13

Smile Detection

The Smile Shutter Flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Sony Cyber-shot® T70 digital still camera can automatically trip the shutter at just the right instant to catch the perfect expression.

14

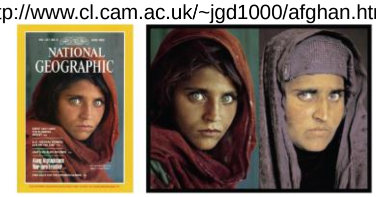
Face Recognition: Apple iPhone Software

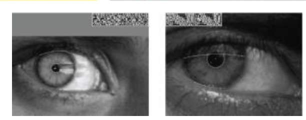
http://www.apple.com/mac/iphoto/

15

Biometrics: Iris

- How the Afghan Girl was Identified by Her Iris Patterns (http://www.cl.cam.ac.uk/~jgd1000/afghan.html)





16

Biometrics: Fingerprint and Face

Fingerprint scanner on many new laptops, other devices

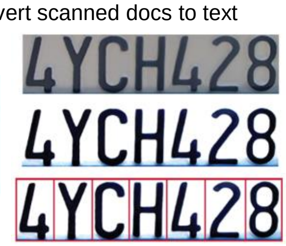
Face recognition systems begin to appear more widely
http://www.sensiblevision.com/

17

Optical Character Recognition (OCR)

- Technology to convert scanned docs to text

Digit Recognition, AT&T Labs



License Plate Readers
https://en.wikipedia.org/wiki/Automatic_number_plate_recognition/

18

Toys and Robots

19

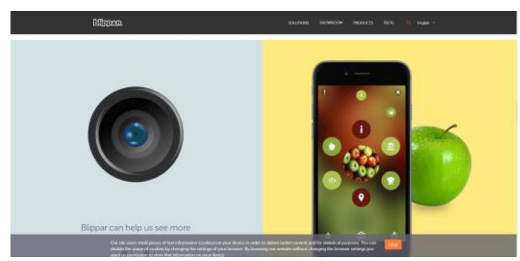
Mobile Visual Search: Google Goggles

20

Mobile Visual Search: iPhone Apps

21

Blippar: Augmented Reality & Visual Discovery Solutions



22

Automotive Safety

- Mobileye: http://www.mobileye.com/en-us/
  - Vision systems in high-end BMW, GM, Volvo models

23

Automotive Safety

- Forward Collision Warning
- Lane Departure Warning
- Pedestrian & Cyclist Collision Warning
- Headway Monitoring & Warning
- Speed Limit Indication

24

Vision in Supermarkets

- LaneHawk LH5000 (https://www.datalogic.com/eng/retail/fixed-retail-scanners/lanehawk-lh5000-pd-830.html) is a loss-prevention solution that turns bottom-of-basket (BOB) into profits in real time. LaneHawk detects and recognizes items, sends their UPC codes to the POS, and includes those items as part of the transaction. LaneHawk makes sure that stores get paid for their BOB items.
- A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk, you are assured to get paid for it …

25

Vision-Based Interaction (Games)

Microsoft's Kinect

Sony EyeToy

Assistive technologies

26

Vision for Robotics, Space Exploration

- NASA's Mars Exploration Rover Spirit captured this westward view from atop, a low plateau where Spirit spent the closing months of 2007.
- Vision systems (JPL) used for several tasks:
  - Panorama stitching
  - 3D terrain modeling
  - Obstacle detection, position tracking
  - For more, read "Computer Vision on Mars" by Matthies et al.

27

Computer Vision Applications

- The applications applied to almost every area of human activities
  - Application categorized by sources: Electromagnetic Energy Spectrum, Electron Microscopy, and Synthetic.
  - Application categorized by fields: Biological Research, Defense/Intelligence, Document Processing, Factory Automation, Law Enforcement, Medical Diagnostic Imaging, Photography, Astronomy, Image Database Retrieval and etc.

28

Types of Digital Images Based on the Sources

- Electromagnetic Energy Spectrum
  - Gamma Rays
  - X-Rays
  - Ultraviolet
  - Visible
  - Infrared
  - Microwaves
  - Radio Waves
- Electron Microscopy
  - Acoustic
  - Ultrasonic
  - Electronic
- Synthetic
  - Fractals and 3D modeling

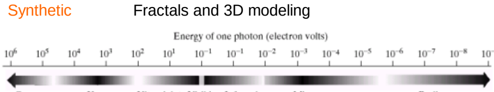


FIGURE 1.5 The electromagnetic spectrum arranged according to energy per photon.

29

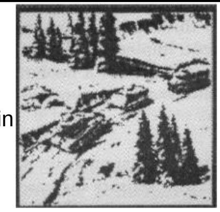
Sample Computer Vision Applications

1. Biological Research

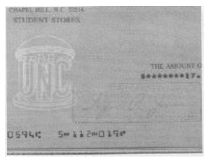
- Automatic analysis of a biological example (specimen analysis)
- Bone, tissue, and cell analysis (counting and classification)
- Analysis, classification, and matching of DNA material

30

2. Defense/Intelligence
- Automatic interpretation of earth satellite imagery
- Recognize and track targets in real time
- Security and surveillance



3. Document Processing
- Scanning, archiving, and transmission of documents
- Automatic detection and recognition of printed characters



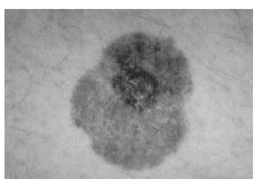
31

4. Factory Automation
- Visual inspection and assembly
- Industrial Inspection

5. Law Enforcement
- Fingerprint feature extraction, classification, and identification
- DNA Matching

32

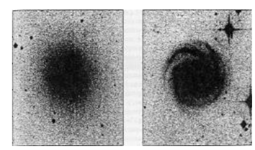
6. Medical Diagnostic Imaging
- Digital Angiography
- Skin Cancer Detection
- Computed Tomography
- Brain Tumor
- Mammography (Breast Cancer)



7. Photography
- Add/Subtract objects to and from a scene
- Special effects (Morphing, Warping)

33

8. Astronomy
- Separating stars from galaxies
- Galaxy classification



9. Image Database Retrieval
- Shape Retrieval
- Color Retrieval
- Texture Retrieval
- Content-based Image Retrieval

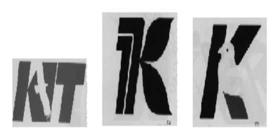


Image query by example: Query Image (left), and two most similar images produced by an image database system

34

Course Overview

- Learn basic concepts and methodologies for digital image processing and pattern recognition (machine learning)
- Learn Matlab programming language (6 programming assignments)
- Work on an exciting final project (CS6680 students), which may lead to your plan A thesis or plan B report
- Understand the solution to an exciting final project (CS5680 students)

35

Digital Image Processing (DIP) Techniques

2D Image →

Low-level Process--produce better images Reduce Noise Enhancement Transforms (Fourier, Wavelet, and etc.)	2D Image
Mid-level Process: Morphology (Extract Image Components and Attributes) Segmentation (Edge Detection; Recognition of Individual Objects)	Attributes
High-level Process: Make Sense of an ensemble of recognized objects (Perform the cognitive functions associated with vision)	Y/N

36

DIP: Low-Level Processing

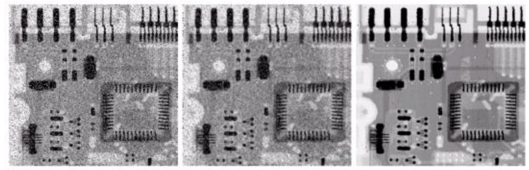
1. Standard procedures are applied to improve image quality

2. Procedures are required to have no intelligent capabilities

37

DIP Low-Level Processing Examples

Ex 1: Remove Noise – Mean and Median Filter Techniques



a b c

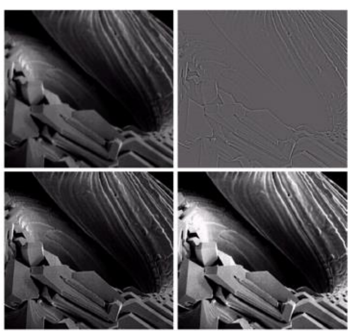
FIGURE 3.37 (a) X-ray image of circuit board corrupted by salt-and-pepper noise. (b) Noise reduction with a 3 × 3 averaging mask. (c) Noise reduction with a 3 × 3 median filter. (Original image courtesy of Mr. Joseph E. Pascente, Lixi, Inc.)

38

Ex 2: Enhancement – Generalized Laplacian Filter Technique

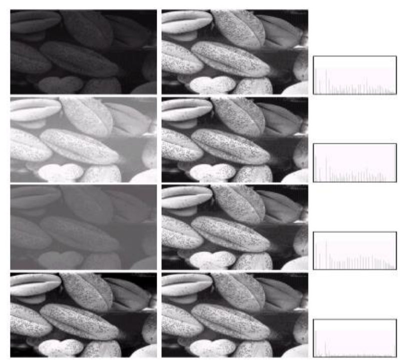
a b
c d

FIGURE 3.43 (a) Same as Fig. 3.41(c), but darker. (a) Laplacian of (a) computed with the mask in Fig. 3.42(b) using $A = 0$. (c) Laplacian enhanced image using the mask in Fig. 3.42(b) with $A = 1$. (d) Same as (c), but using $A = 1.7$.



39

Ex 3: Enhancement – Histogram Equalization Technique



40

Ex 4: Enhancement – Power-Law Transformation Technique

a b
c d

FIGURE 3.9 (a) Aerial image. (b)–(d) Results of applying the transformation in Eq. (3.2-3) with $c = 1$ and $\gamma = 3.0, 4.0$, and 5.0, respectively. (Original image for this example courtesy of NASA.)

41

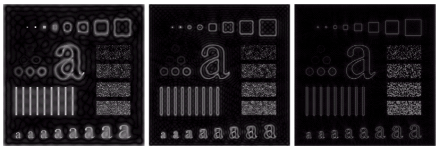
Ex 5: Smooth – Frequency-Domain Lowpass Filter Technique

FIGURE 4.18 (a) Original image. (b)–(f) Results of filtering with Gaussian lowpass filters with cutoff frequencies set at radii values of 5, 15, 30, 80, and 230, as shown in Fig. 4.11(b). Compare with Figs. 4.12 and 4.15.

a b
c d
e f

42

Ex 6: Sharpen – Frequency-Domain Highpass Filter Technique



a b c

FIGURE 4.24 Results of ideal highpass filtering the image in Fig. 4.11(a) with D_0 = 15, 30, and 80, respectively. Problems with ringing are quite evident in (a) and (b).

43

DIP Mid-Level Processing

1. Extract and characterize components in an image
2. Some intelligent capabilities are required

44

DIP Mid-Level Processing Examples

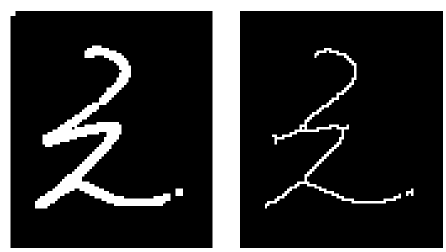
Ex 1: Boundary Extraction – Binary Morphological Erosion

a b

FIGURE 9.14 (a) A simple binary image, with 1's represented in white. (b) Result of using Eq. (9.5-1) with the structuring element in Fig. 9.13(b).

45

Ex 2: Structure Finding – Binary Morphological Thinning



46

Ex 3: Appearance Uniform – Gray-Level Morphological Erosion

Ex 4: Special Effects – Gray-Level Morphological Opening

47

Ex 5: Segmentation – Edge Detection

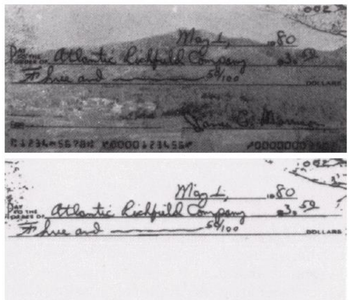
a b
c d

FIGURE 10.11 Same sequence as in Fig. 10.10, but with the original image smoothed with a 5 × 5 averaging filter.

8

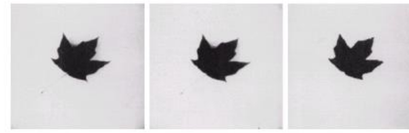
Ex 6: Segmentation – Local Thresholding Technique

a
b
FIGURE 10.37
(a) Original image. (b) Image segmented by local thresholding. (Courtesy of IBM Corporation.)

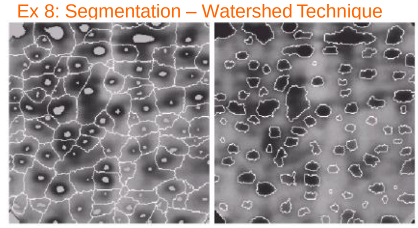


Ex 7: Segmentation – Split and Merge Technique

a b c
FIGURE 10.43
(a) Original image. (b) Result of split and merge procedure. (c) Result of thresholding (a).



Ex 8: Segmentation – Watershed Technique



a b
FIGURE 10.48
(a) Image showing internal markers (light gray regions) and external markers (watershed lines). (b) Result of segmentation. Note the improvement over Fig. 10.47(b). (Courtesy of Dr. S. Beucher, CMM/Ecole des Mines de Paris.)

DIP High-Level Processing

1. Recognition and interpretation
2. Procedures require high intelligent capabilities

Part of Pattern Recognition or Machine Learning

DIP High-Level Processing Examples

Ex 1: Object Recognition – Matching Technique

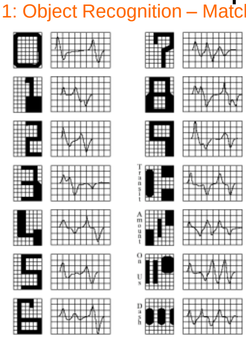
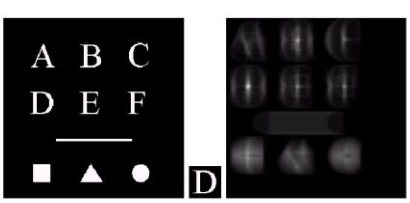


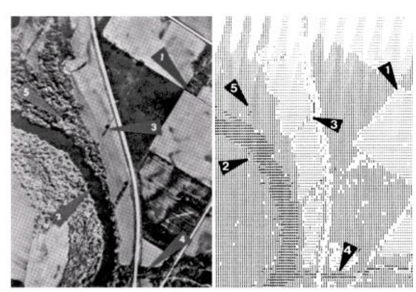
FIGURE 12.7
American Bankers Association E-13B font character set and corresponding waveforms.

Ex 2: Object Recognition – Correlation Technique



a b c
FIGURE 12.9
(a) Image. (b) Subimage. (c) Correlation coefficient of (a) and (b). Note that the highest (brighter) point in (c) occurs when subimage (b) is coincident with the letter "D" in (a).

Ex 3: Object Recognition – Bayes Classifier Technique



a b
FIGURE 12.13 (a) Multispectral image. (b) Printout of machine classification results using a Bayes classifier. (Courtesy of the Laboratory for Applications of Remote Sensing, Purdue University.)

Pattern Recognition (PR) a.k.a Machine Learning (ML)

- Classify what inside of the image
- Applications:
  - Speech Recognition/Speaker Identification
  - Fingerprint/Face Identification
  - Signature Verification
  - Character Recognition
  - Biomedical: DNA Sequence Identification
  - Remote Sensing
  - Meteorology
  - Industrial Inspection
  - Robot Vision

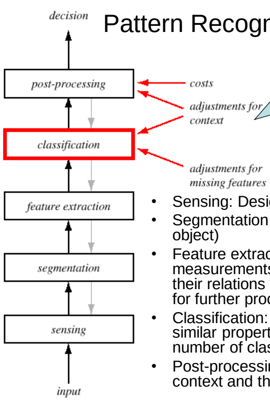
55

PR or ML

- It deals with classification, description, and analysis of measurements taken from physical or mental processes.
- Pattern recognition
  - Take in raw data
  - Determine the category of the pattern
  - Take an action based on the category of the pattern

56

Pattern Recognition Main Components



How many classes are needed in an English character recognition system?
How many classes are needed to distinguish Chinese from English?

- Sensing: Design of transducers
- Segmentation and grouping (into a composite object)
- Feature extraction: A set of characteristic measurements (numerical or non-numerical), and their relations are extracted to represent patterns for further process.
- Classification: The process or events with same similar properties are grouped into a class. The number of classes is task-dependent.
- Post-processing: Considering the effects of context and the cost of errors

57

PR Example

- Fish-packing plant -- species determination
  - Separate sea bass from salmon using optical sensing
  - Image features
    - Length
    - Lightness
    - Width
    - Number and shape of fins, etc.
  - Establish models for objects to be classified
    - Descriptions in mathematics form
    - Considering noise or variations in population itself and sensing

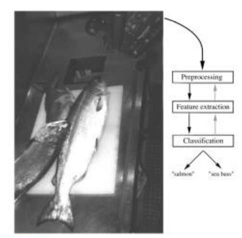


FIGURE 1.1. The objects to be classified are first sensed by a transducer (camera), whose signals are preprocessed. Next the features are extracted and finally the classification is emitted; here either "salmon" or "sea bass." Although the information flow is often chosen to be from the source to the classifier, some systems employ information flow in which earlier levels of processing can be altered based on the tentative or preliminary response in later levels (gray arrows). Yet others combine two or more stages into a unified step, such as simultaneous segmentation and feature extraction. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

58

First Feature Extraction

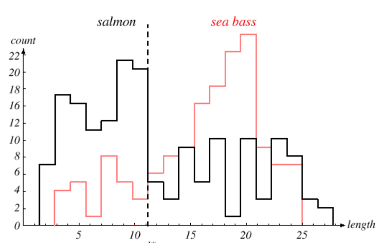


FIGURE 1.2. Histograms for the length feature for the two categories. No single threshold value of the length will serve to unambiguously discriminate between the two categories; using length alone, we will have some errors. The value marked l^* will lead to the smallest number of errors, on average. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

59

Second Feature Extraction

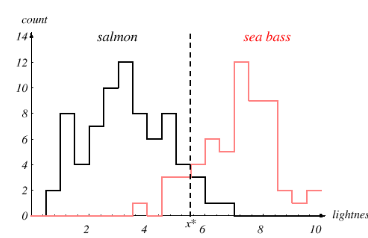


FIGURE 1.3. Histograms for the lightness feature for the two categories. No single threshold value x^* (decision boundary) will serve to unambiguously discriminate between the two categories; using lightness alone, we will have some errors. The value x^* marked will lead to the smallest number of errors, on average. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

60

Classification -- Classifier Design

- Feature space
  - Feature vector

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

- Scattering plot for training samples
- Classifier : design of decision boundary on scattering plot
  - Partition the feature space into several regions.

61

Linear Classifier

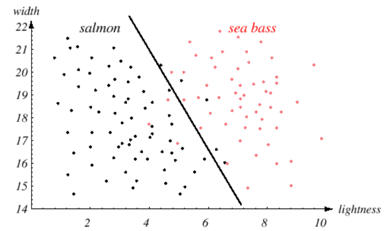


FIGURE 1.4. The two features of lightness and width for sea bass and salmon. The dark line could serve as a decision boundary of our classifier. Overall classification error on the data shown is lower than if we use only one feature as in Fig. 1.3, but there will still be some errors. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

62

The Best Classifier

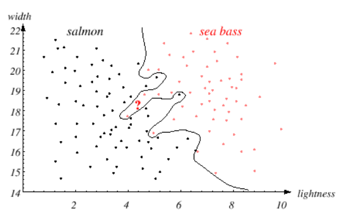


FIGURE 1.5. Overly complex models for the fish will lead to decision boundaries that are complicated. While such a decision may lead to perfect classification of our training samples, it would lead to poor performance on future patterns. The novel test point marked **?** is evidently most likely a salmon, whereas the complex decision boundary shown leads it to be classified as a sea bass. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

63

The Optimal Classifier

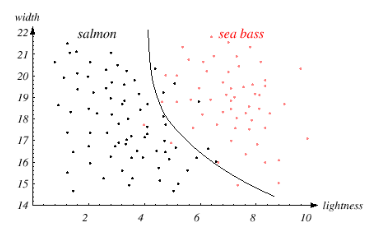


FIGURE 1.6. The decision boundary shown might represent the optimal tradeoff between performance on the training set and simplicity of classifier, thereby giving the highest accuracy on new patterns. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

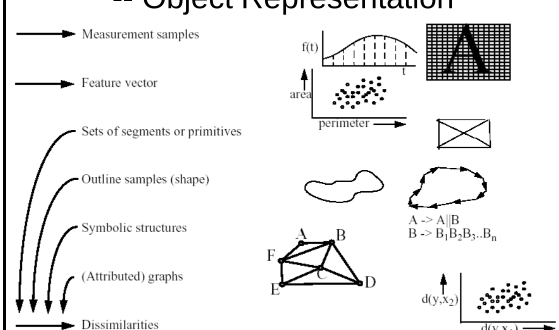
64

PR Feature Extraction

- Seek distinguishing features that are invariant to irrelevant transformations.
  - Distinguishing features
    - Feature values are similar in the same category and very different in different categories.
  - Irrelevant transformations
    - Rotation, Scale, and Translation, (RST invariance, major concern)
    - Occlusion
    - Projective distortion
    - Non-rigid deformations
- Feature selection (those are most effective)

65

PR Feature Extraction Techniques -- Object Representation



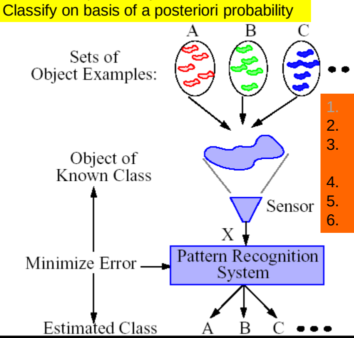
PR Classification

- Assign an object to a category by using the feature vector.
- Difficulty of classification depends on the variability of the feature values in the same category relative to the difference between feature values in different categories.
- The variability of feature values in the same category may come from noise

67

PR Classification Techniques -- Statistics (Distribution) Based PR

Define/Find probability model
Classify on basis of a posteriori probability

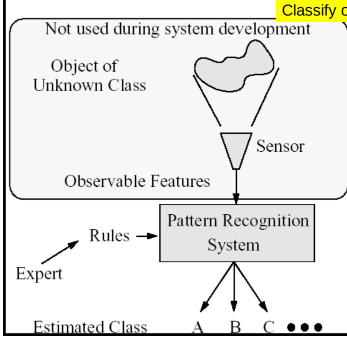


1. Bayes Classification
2. Perception Algorithm
3. Least-Mean-Square Error Algorithm
4. Neural Network (NN)
5. Deep Convolutional NN
6. Support Vector Machines

68

PR Classification Techniques -- Expert (Model) Based PR

Define/find discriminant function G(x)
Classify on basis of value (sign) of G(x)

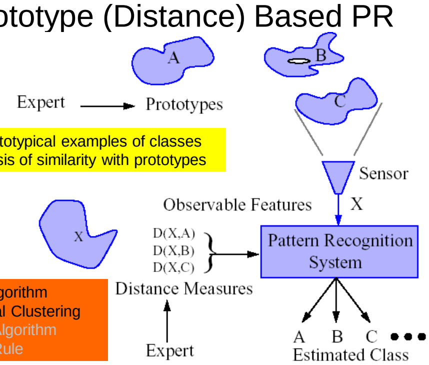


1. Linear Discriminant Analysis (LDA)
2. Principal Component Analysis (PCA)
3. Decision Functions

69

PR Classification Techniques --Prototype (Distance) Based PR

Define/find prototypical examples of classes
Classify on basis of similarity with prototypes



1. K-Mean Algorithm
2. Hierarchical Clustering
3. ISO Data Algorithm
4. Syntactic Rule

PR Post-Processing

- Consider the cost of action
  - Minimize classification error rate
  - Minimize risk (total expected cost)
- Exploit context (input-dependent information) to improve system performance
  - E.g., use the context information for OCR or speech recognition
- Multiple classifier (different from multiple features)
  - Each classifier operates on different aspects of the input (e.g., speech recognition = acoustic recognition + lip reading)
  - Decision fusion

71

Areas Related to PR

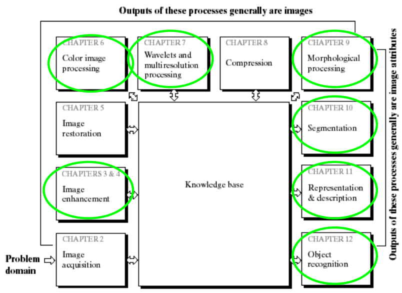
- Image Processing
- Speech Processing
- Artificial Intelligence
- Associate Memory
- Neural and Fuzzy
- Probability and Statistics (Statistical)
  - Regression (find functional description of data for new input prediction)
  - Interpolation (infer the function for intermediate ranges of input)
  - Density estimation (for ML and MAP classification)
- Formal language (Syntactic)
- Neural network architecture design and training (Neural)

72

DIP/PR Main Components

- Software (Algorithms)
- Hardware (Camera, Lights, Frame grabber, Processor)

FIGURE 1.23 Fundamental steps in digital image processing.



73

Prerequisite

- Need good background in mathematics: Linear algebra, statistics, matrix theory, calculus, and etc.
- Need good background in algorithm design and analysis.
- Need good background in programming.

74

CS5680/CS6680

- CS5680/CS6680
  - Enthusiastic undergraduates and graduates
  - Want to get to know this exciting technology

75