Pattern Recognition

Introduction to Pattern Recognition

Wang, Yuan-Kai

王元凱

ykwang@fju.edu.tw

http://www.ykwang.tw

Department of Electrical Engineering, Fu Jen Catholic University

輔仁大學電機工程系

Cite this document as:

Wang, Yuan-Kai, “Introduction to Pattern Recognition,"
Lecture Notes of Wang, Yuan-Kai, Fu Jen University, Taiwan, 2016.

- This is the first lecture note of the course PATTERN RECOGNITION in English in 104-2 semester, EE, FJU.
- In this lecture note, I will introduce basic concept of an image recognition system.
- Web site of this course: http://pattern-recognition.weebly.com.

Goal of This Unit

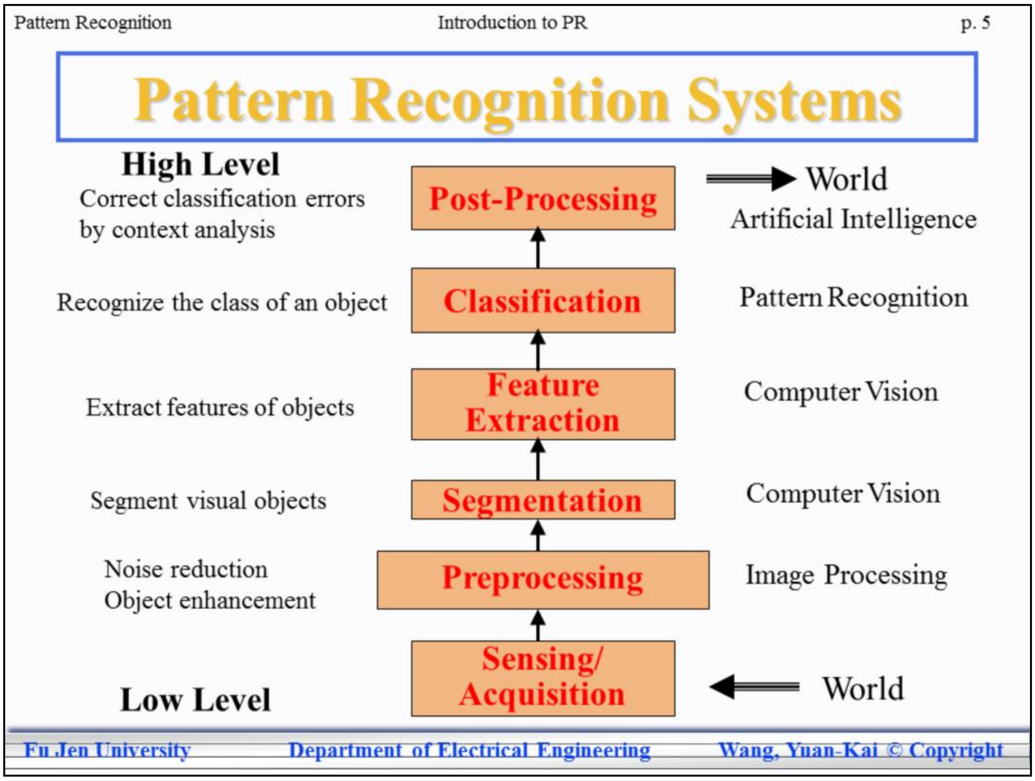
- **Explain PR with a systematic view**
  - **From sensor input to recognition output**
- **Discriminate among three courses**
  - **Pattern recognition, computer vision, digital image processing**

References

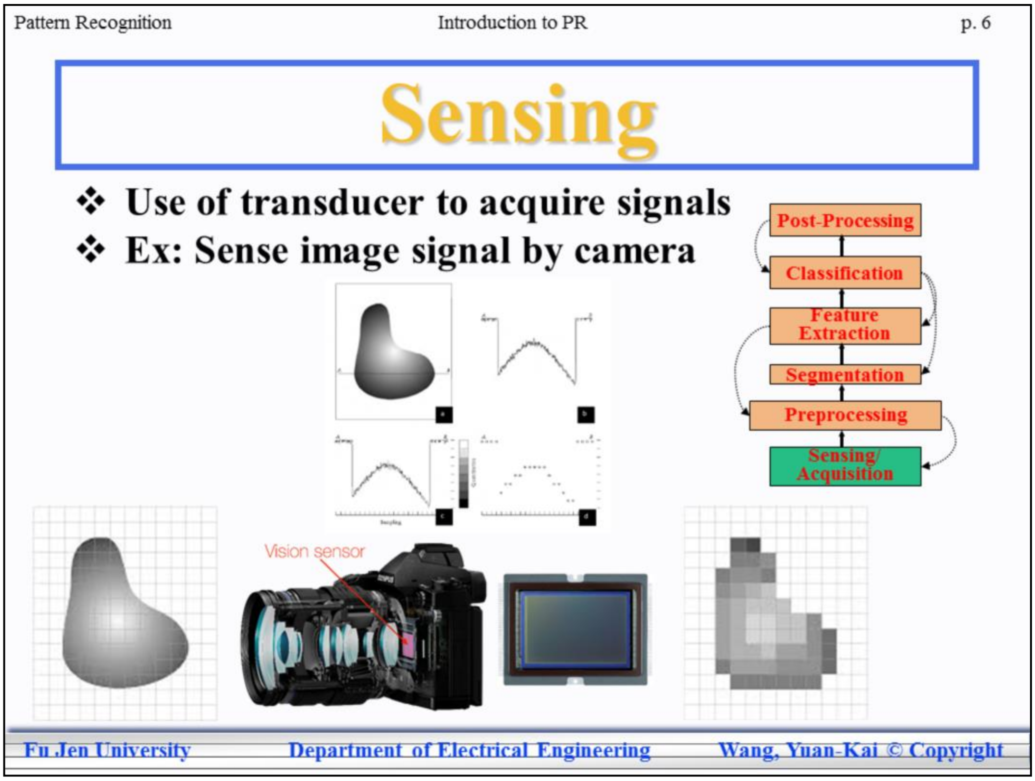
- **Pattern Recognition Engineering,**
  M. Nadler & E. P. Smith
  - **Chapter 1 - Introduction**
- **Pattern Recognition, 4th,**
  S. Theodoridis & K. Koutroumbas
  - **Chapter 1 - Introduction**
- **Pattern Classification, 2nd,**
  R. O. Duda, P. E. Hart & D. G. Stork
  - **Chapter 1 – Introduction**

Contents

- Pattern Recognition Systems
- Performance Evaluation



- A pattern recognition (PR) system is not identical to pattern recognition (PR) algorithms
  - A PR system is more complete than a PR algorithm.
  - A PR system needs to implement related techniques such as sensor, filtering and so on.
  - But PR algorithms are just a key component of a PR system.
- PR systems and algorithms can be used to a lot of signals, such as speech, image, video, music, text, and so on.
  - But this course will use only image signal as examples.



- Sensing is the first stage for a pattern recognition system.
  - It is usually hardware, but not software.
- Terminologies about transducer
  - Transduce: transform one energy form into another energy form
  - Transducer: an instrument to transduce energy
    - Ex.: microphone: a transducer to transform air vibration energy to electron energy
    - Ex.: camera: a transducer to transform photon energy into electron energy
  - Digital transducer: a transducer with the digitization of electron energy
    - Ex.: digital camera vs. analog camera: digital camera uses IC chip to digitize electron energy into digital signals(images), but traditional (analog) camera uses film to get images.
- Terminologies about image sensing
  - Image signals, image sensing, image sensor, camera
  - Could you differentiate the difference among these terminologies?
- Sampling and quantization: AD (Analog to Digital) Conversion
  - What is sampling: digitization in temporal domain
  - What is quantization: digitization of intensity/quantity

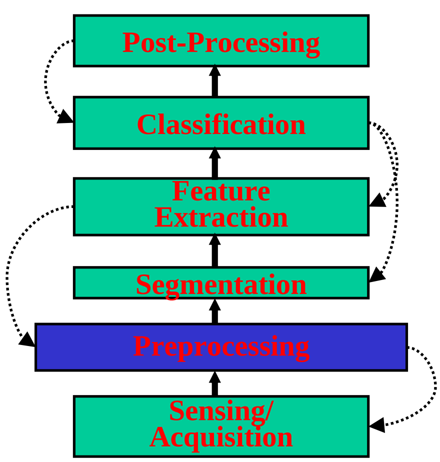
Problems of Sensing

- **Noises are mixed with signal**

- **Objects are not clear**

Unclear object Clear object

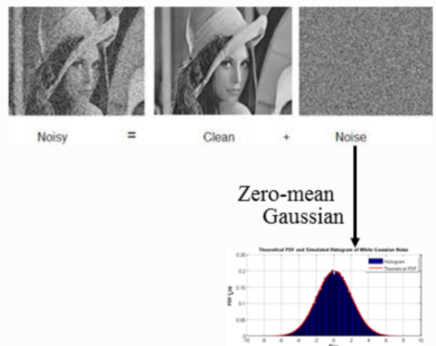
- Why are noises generated?
  - Dust, light, sampling, quantization, …
- Why are objects unclear?
  - Insufficient light(night), under exposure(wrong exposure setting of camera), de-focus (lens focus is not right), …
- Both noise removal and object enhancement are done in next stage: preprocessing



- Intrinsic noise
  - The measured physical signal is already noisy, because the sensor will have its own intrinsic noise level, from thermal and other noise sources.
  - Even the raw un-quantized signal is accompanied by noise.
  - Example noise intrinsic to the system
    - The hiss on a cassette recording
    - The rumble from a turntable
- Sampling and quantization noise
  - https://en.wikipedia.org/wiki/Quantization_(signal_processing)
  - It is the noise generated by the analog to digital signal process (AD conversion). It includes two steps: sampling and quantization.
  - Any physical measurement from sensors (image sensors such as Photoresistive target or vidicon-like, or solid-state arrays) are usually analog quantities that must be quantized in order to become machine variables.
  - Sampling and quantization generates error(noise)
    - Sampling and quantization produce approximate discrete data,
    - There are errors between original continuous signal and discrete data.
  - Sampling and quantization noise is produced by the approximation error
    - The finer the quantization, the smaller is the noise
    - The cost of equipment increases as some power of the fineness of quantization - resolution
- Interference noise
  - *Interference effects* cause slight variations of shape to occur between repeat scans of the same object
  - Ex. Electromagnetic interference https://en.wikipedia.org/wiki/Electromagnetic_interference

Noises with Different Characteristics

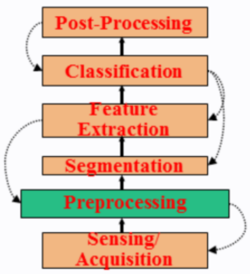
- **Mathematical**
  - **Additive noise**
  - **Mulplicative noise**
- **Statistical**
  - **Gaussian noise**
  - **Poisson noise**
  - **Impulse noise**



- Materials in Wikipedia to study noises
  - Gaussian noise: https://en.wikipedia.org/wiki/Gaussian_noise
  - Poisson noise: https://en.wikipedia.org/wiki/Shot_noise

Preprocessing

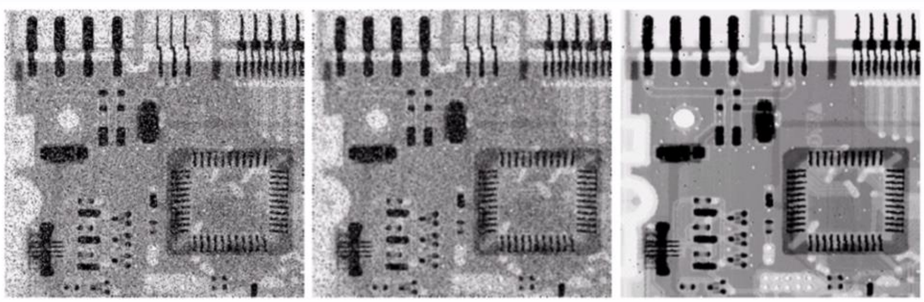
- Noise reduction
- Object enhancement



- Preprocessing is the second stage for a pattern recognition system.
- The following slides will give some examples to explain these two goals.

Noise Reduction

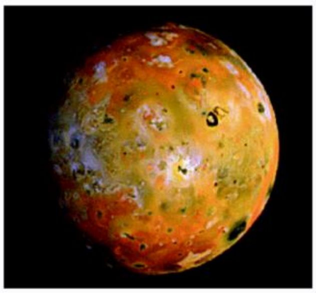
Noisy image	Denoised image by Gaussian filter	Denoised image By median filter



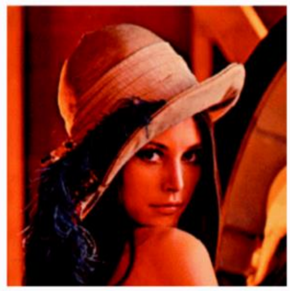
- Noise reduction is a process to reduce/remove noises in the image
  - It is also called: noise removal, denoise, …
- The noise in the left image of this slide is : salt-and-pepper noise (a kind of impulse noise)
  - Salt-and-pepper noise cannot be well-processed by Gaussian filter
  - Salt-and-pepper noise can be well-processed by median filter.

Object Enhancement

- What is an object?
  - Anything you want in the image



Black regions on the planet



Face

License plate

- What are the applications of these examples:
  - Black regions on the planet: astronomy research to study the history of the planet damaged by comet hits
  - Face object: face recognition
  - License plate: license plate recognition

Object Enhancement: License Plate

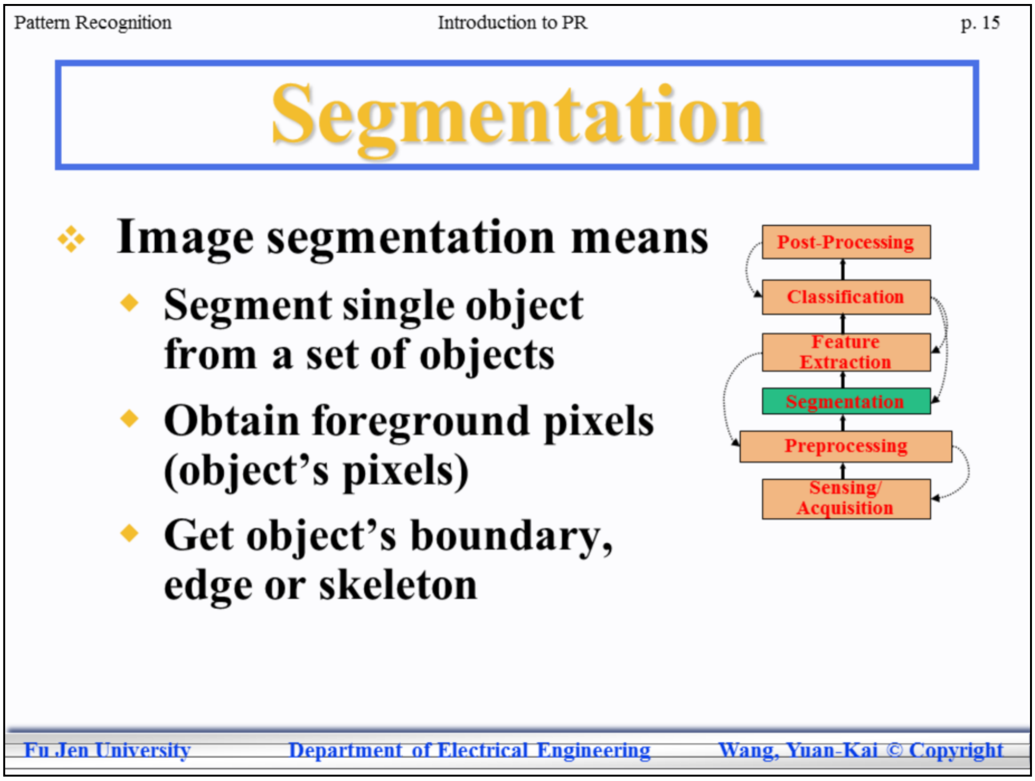
- License plate enhancement and recognition is a well-known and important application in ”intelligent video surveillance”.
- Two situations that make the object, license plate, looks bad
  - Blur due to high-speed car driving or low-speed of camera shutter
  - Strong light due to car head light

Methods for Denoise & Object Enhancement

- Histogram Processing
- Linear Filtering
  - Spatial domain (convolution)
  - Transform domain
    - Frequency domain by Fourier transform
    - Multiscale domain by Wavelet transform
- Nonlinear Filtering
  - Morphology, rank-based filter
  - Non-local mean
  - Sparse representation
  - Markov random field
  - *Neural network*

Digital Image Processing course

- You should learn preprocessing techniques in the “digital image proessing” course.
  - The course will teach histogram processing and linear filtering
  - You can then learn nonlinear filter by yourself based on the foundation of the course.
- This course is pattern recognition, so we will not teach preprocessing and image processing.

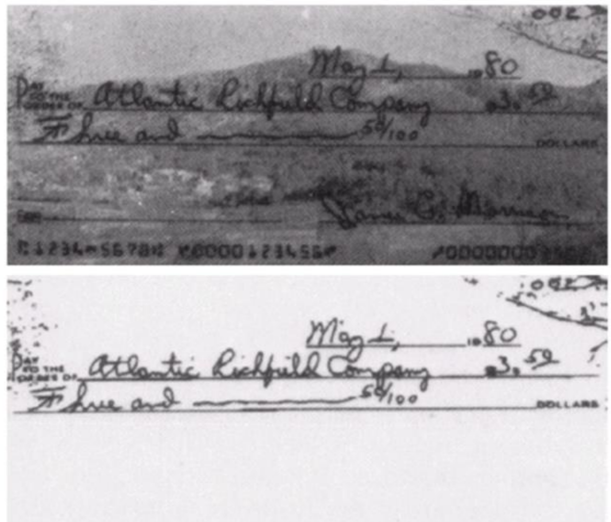


- Segmentation is the third stage of a pattern recognition system.
- Segmentation isolates the objects in the image into a new small image
  - In order to carry out segmentation, it is necessary to detect certain features that may not enter into the list of features utilized for recognition.
  - They are obtained from the direct (or preprocessed) measurements that are related to certain properties of pre-attentive vision.

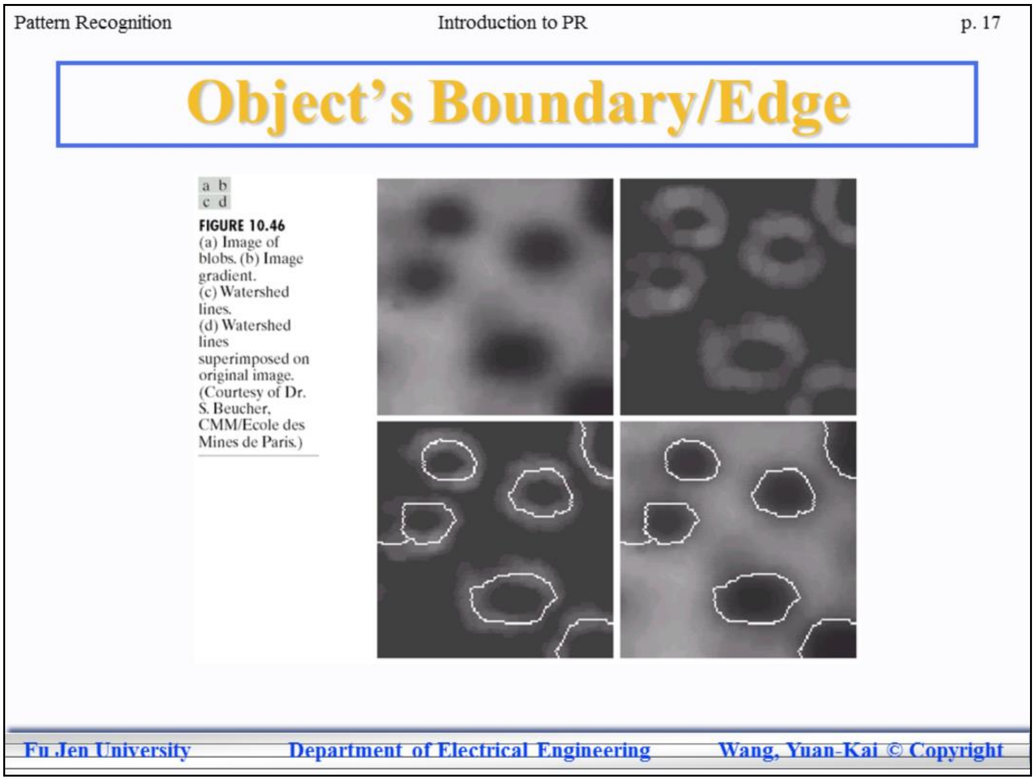
Object's Pixels (1/2)

a
b

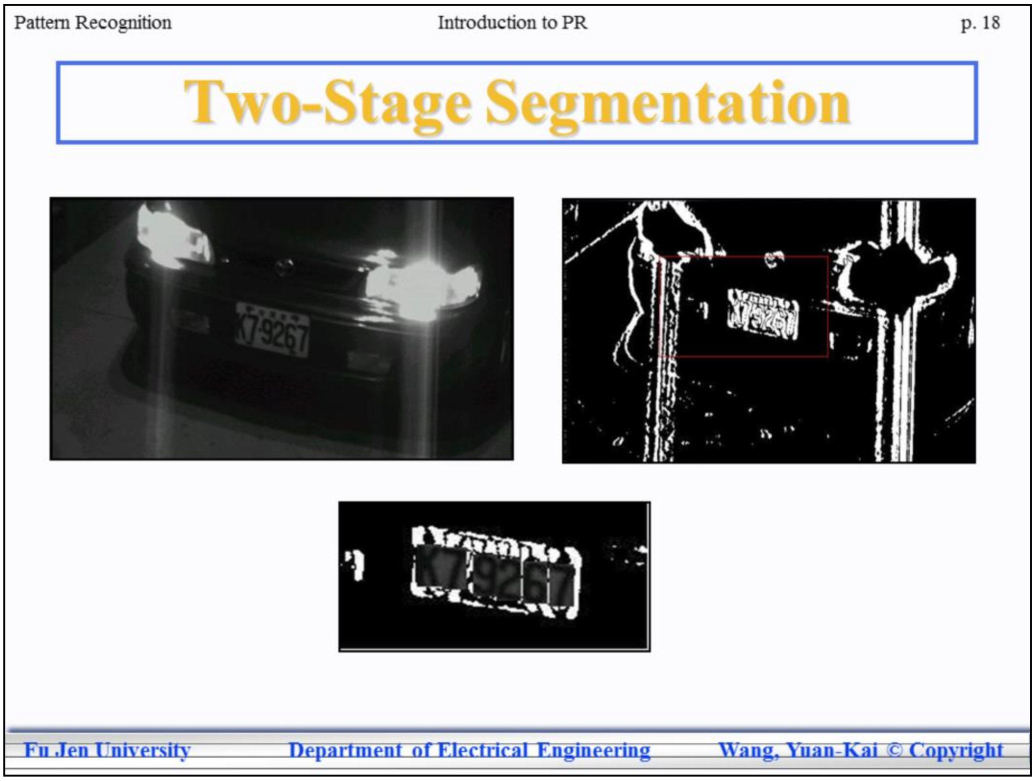
FIGURE 10.37 (a) Original image. (b) Image segmented by local thresholding. (Courtesy of IBM Corporation.)



- This figure comes from the textbook of digital image processing written by Gonzalez.
- In this example,
  - The top image is a check of a commercial bank.
  - The bottom image is a segmentation image with only signed name and numbers, without the watermark of the check (i.e., the background of the image).
  - This example uses a very simple segmentation technique: thresholding.



- This figure comes from the textbook of digital image processing written by Gonzalez.
- In this example
  - (a) image is a blur image. The black circles in this image are to be segmented.
  - (b) and (c) images are temporary processed results.
  - (d) image is segmentation result. The white contours of all black regions illustrate the segmentation results.

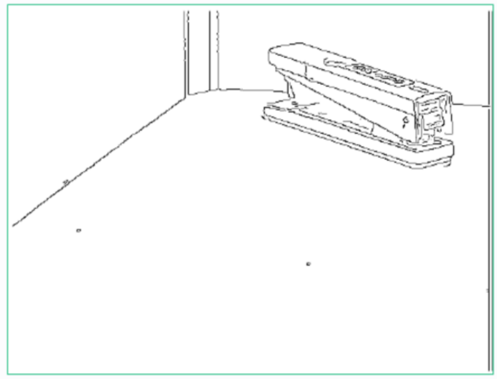


- To segment objects in complex images, we usually need to perform two segmentation steps: first step to find rough place of the object, and second step to find the exact locations of objects
- This example is for license plate recognition, and we need to firstly segment license plate, and then secondly segment the alphanumerics.
  - Left top image is the original image captures at night.
  - Right top image is a temporary processed image. Red rectangle in this image represents the first segmentation result that indicates the license plate.
  - Bottom image show the final results of second segmentation. Six alphanumeric in the license plate is isloated. These six characters can be later recognized.

Edge Detection (1/2)

- **What is edge detection**

Original image

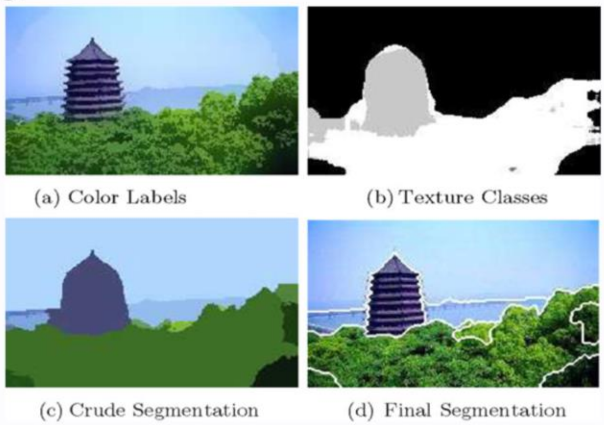


Edge image

- Edge detection is a very useful for image/object segmentation.
- Edge image can be obtained by linear or nonlinear filtering.

Color Segmentation

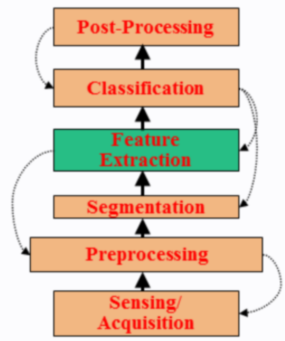
- **Usually use 3 planes (R,G,B) for segmentation**



- Color image segmentation is more difficult than grayscale image segmentation.
- We have to find texture, color and edge information in all three channels, and then use these information to segment the color image.

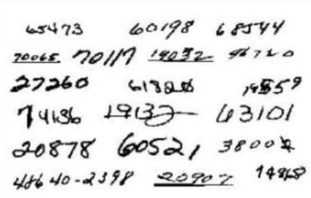
Feature Extraction

- Use high-level features to **describe and represent segmented objects**

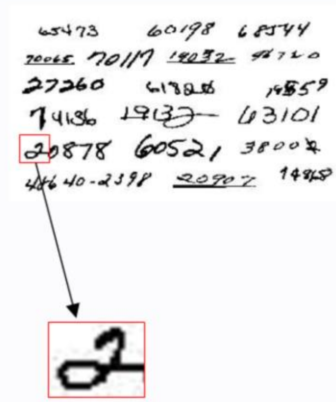


- Feature extraction is the fourth stage of a pattern recognition system.

Example 1



Segmentation →



Features

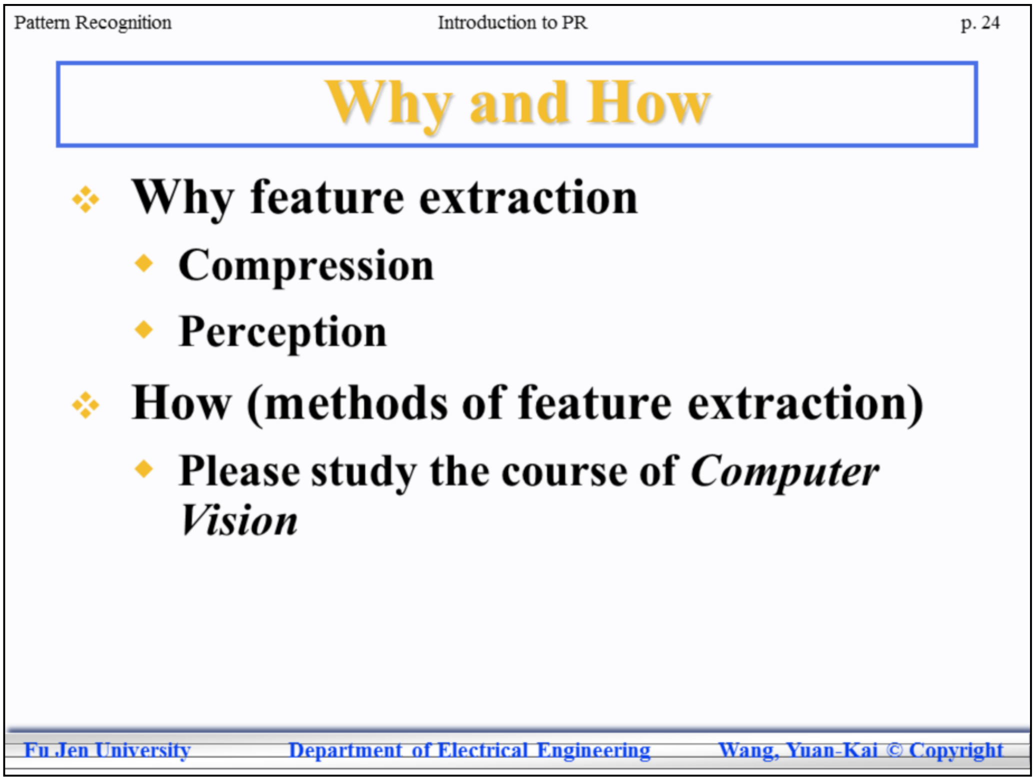
1. **Boundary**
2. **Line Segments**
3. **Shape: Fourier descriptors**
4. **Statistical moments**
5. **Topological descriptors**
6. **...**

Example 2

Segmentation

Features

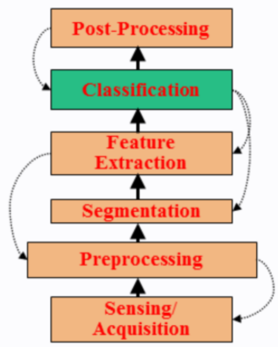
1. **Eyes: length, shape**
2. **Eyebrows: length, shape**
3. **Nose: length, shape**
4. **Mouse: length, shape**
5. **...**



- Two goals of feature extraction
  - **Compression:** Reduction of dimensionality in pattern space
    - There are too many measurements after acquisition/preprocessing/segmentation.
    - Many or most of them may not even help to distinguish the class of the object from other classes.
    - Feature extraction is the attempt to extract meaningful features from measurements.
  - **Perception:** Rendering the features more suitable for the decision process
    - When we look at a printed page, scene, or an electroencephalogram(EEG腦波圖), we don't see an array of optical values.
    - When we hear a speech, a siren, or an engine turning, we don't hear a time series of acoustic pressures.
    - Our primary sensory systems do these.
    - But we *perceive* letters, trees, α-waves, spoken words, loud high-pitched sounds, and so on.
- Feature extraction of image objects is a complex technique.
  - This course has not time to teach this technique.
  - Please go for the course of Computer Vision.

Classification

- ❖ Use the feature vector obtained by the feature extractor as input
- ❖ Then assign the object to a recognized class



- Classification is the fifth stage of a pattern recognition system.
- This is exactly the stage that will be explained in this course.

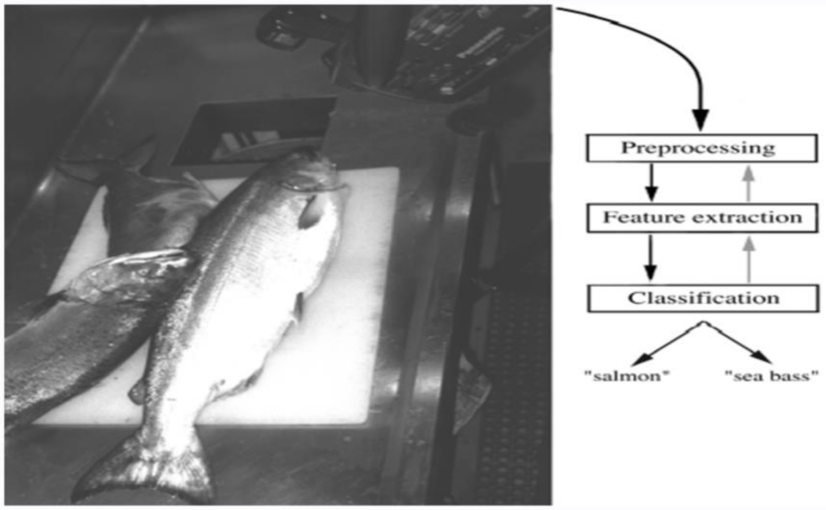
An Example

- Sort fish on a conveyor according to species using optical sensing

Species → Sea bass / Salmon

- How to do this by humans?
  - Classify fish species by its size, length, width, lightness, …
- How to do this by computers?
  - Image preprocessing, feature extraction, and pattern recognition: a pattern recognition system.

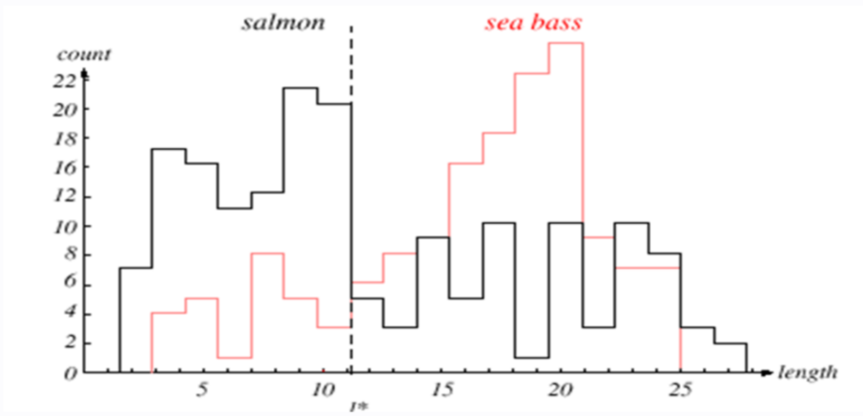
Automatic Fish Classification by Computers



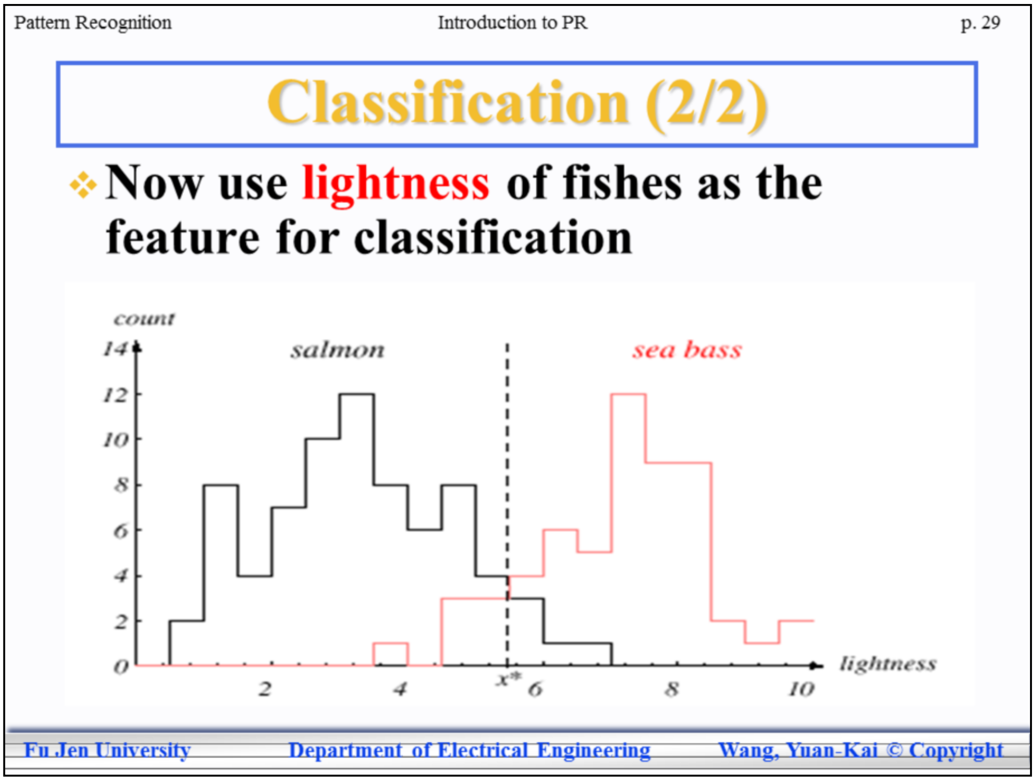
- Set up a camera on a platform that can places a fish
- Take a picture of the fish
- Write a program to
  - Segment the fish: Isolate fishes from backgrounds. It includes denoising, enhance the image by filtering, and segmentation.
    - This is *not* the topic of this class. But it will be taught by the class ”Digital Image Processing”
  - Extract features of fish: Length, lightness, width, number and shape of fins, position of the mouth, etc…
    - This is *not* the topic of this class. But it is taught by the class ”Computer Vision”
  - **Classify the species of the fish: Use pattern recognition algorithms to do this.**
    - **This is the topic of this course**
- Let us suppose that we can write a program to successfully
  - Segment the fish object in an image.
  - Extract the features of the fish, such as length, width, lightness, and so on.
- So next slide we go further to the ”classification”step.

Classification (1/2)

- **Select the length of the fish as a possible *feature* for discrimination**



- Horizontal axis is the length of fishes.
- Vertical axis is the number of fishes with respect to fish length.
- Black line shows the histogram of salmon. Red line shows the histogram of sea bass.
  - For example, the number of salmon with length 5 is 16. The number of sea bass with length 20 is 22.
- Black dotted line represents ”threshold” to classify salson and sea bass.
  - If we set the threshold to be 11.1, as shown in this figure.
  - Then all fishes with length less than 11.1 is classified to be salmon. All fishes with length larger than 11.1 is classified to be sea bass.
  - So, some sea bass fishes with length less than 11.1 is mis-classified, and so are sea bass fishes with length larger than 11.1
- Mis-classified fishes are called classification error. With these errors we can then calculate error rate or accuracy of this classification with respect to this threshold: 11.1.
  - Each threshold has an error rate.
  - Some thresholds have higher error rates, and they are not good thresholds.
  - The best threshold is the threshold with minimum error rate. But
- Two questions
  - What is the best threshold in the example?
  - If the error rate of the best threshold is not good enough, can we get better classification by other ways?

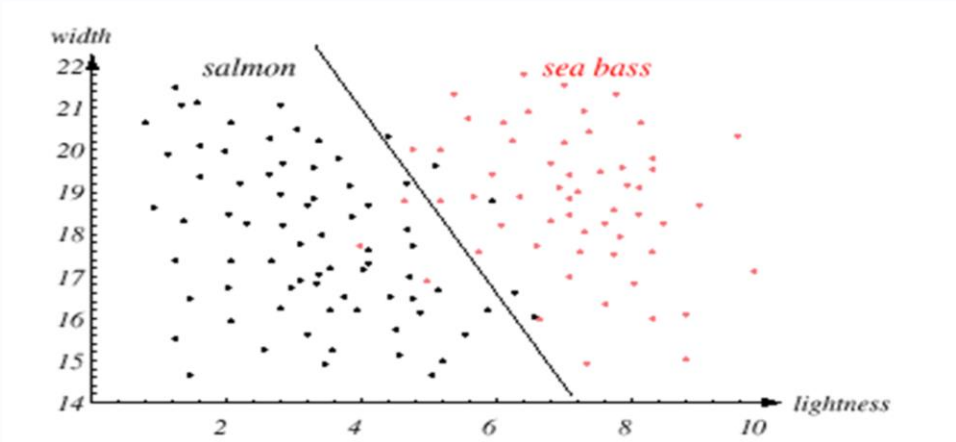


- The length feature in previous slide is a bad feature.
- Let us change the feature to be lightness of fishes.
  - The graph seems to be better than the graph of length, because the two histograms of two fish species are not "seriously overlapped".
- What is the best threshold in this example?
  - Suggestion: move the threshold (decision boundary) toward smaller values of lightness in order to minimize the cost (reduce the number of sea bass that are classified salmon!)
- Question
  - Can we get more better accuracy than using the lightness feature?

Add One More Feature

- Now let us use two features
  - Combine the *lightness* feature with the *width* feature of fishes

Classification in Two Dimensions

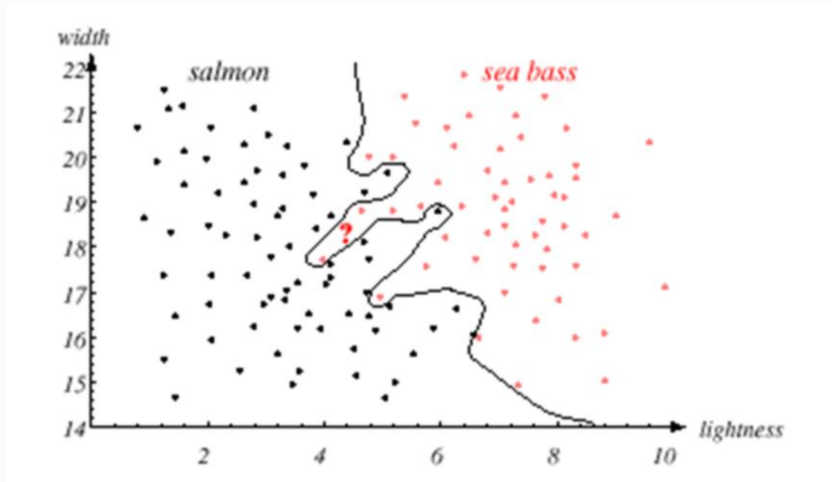


Decision boundary, discriminant, classifier

- Now both axes represent features for classification.
- Important concepts shown in this illustrated figure
  - Decision boundary (also called threshold in 1D example)
  - Noise (some salmons locate in the region of sea bass, and vice versa)
  - Currently we use **straight line** to be decision boundary
- We can move the line and change its direction(slope)
  - Lines at different location and with different slope have different error rate.
  - We have to find he best line (decision boundary) that has the minimum error rate.
- Classifier and discriminant
  - The decision boundary (straight line) is also called classifier and/or discriminant.
- Discussions
  - Could we add more features than two?
  - We might add other features that are not correlated with the ones we already have. A precaution should be taken not to reduce the performance by adding such “noisy features”
- Questions
  - How to find the discriminant/classifier/decision boundary?
    - **Machine learning and pattern recognition**
  - Could we use non-straight line or curve to be decision boundary?

Pattern Recognition Introduction to PR p. 32

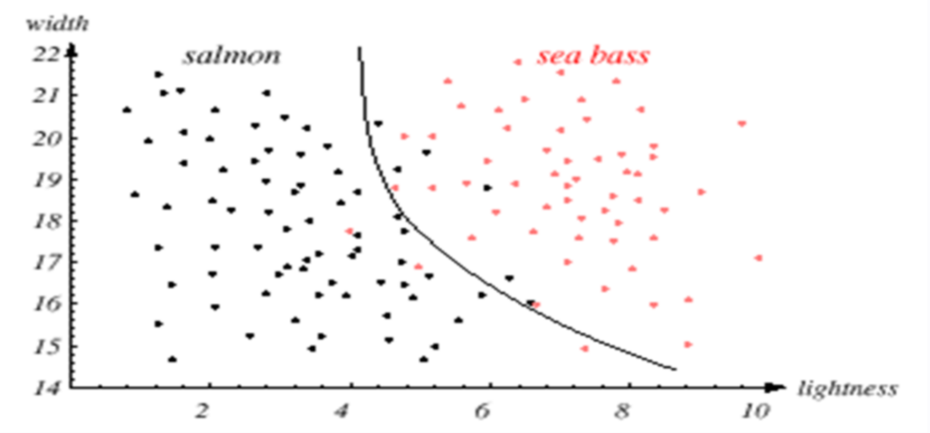
Classification in 2D: Perfect(Accurate) Classifier ?



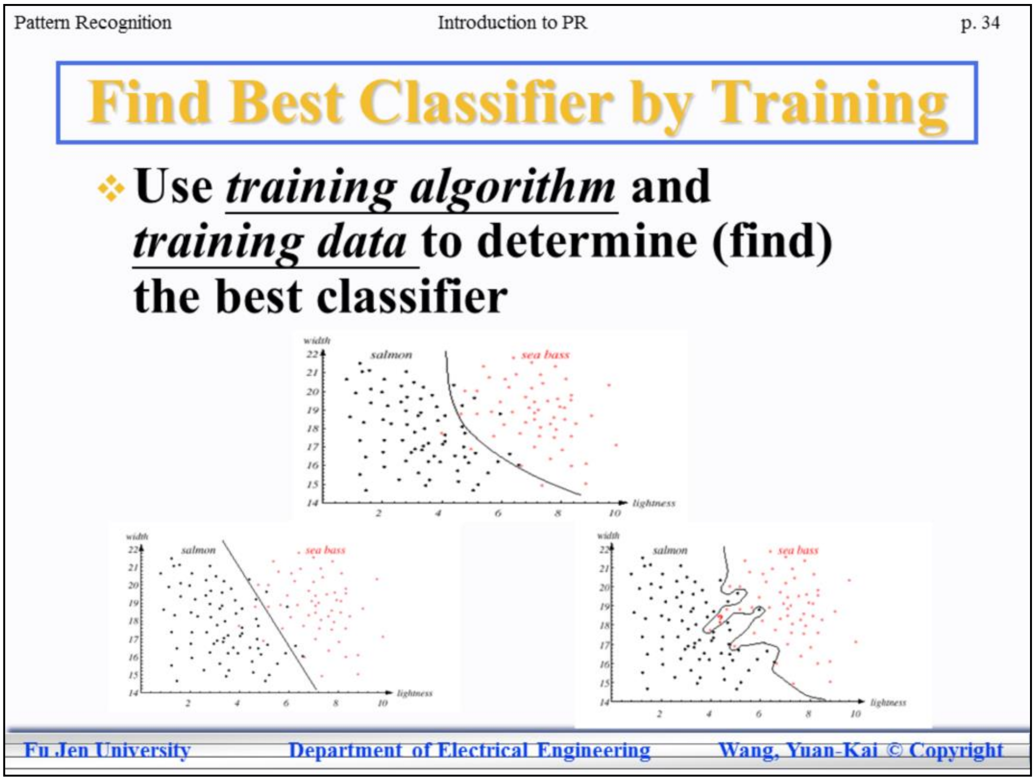
Fu Jen University Department of Electrical Engineering Wang, Yuan-Kai © Copyright

- Ideally, the best decision boundary should be the one which provides an optimal performance such as in this figure.
  - In this figure the decision boundary is a high-order curve, but not a straight line.
  - This curve looks like a very good classifier, because it has zero error.
- However, it is not a good classifier because of the issue of **generalization**
  - Its accuracy for current fishes is 100%.
  - Will it be still good for new unknown fishes? No, absolutely not.
  - Practically, this classifier is too specialized for current fishes, thus it has no generalization for future unknown fishes.

Classification in 2D: A Compromised Classifier



- In reality, this may be the best classifier.
  - This classifier(decision boundary) is a second-order curve.
- For current fishes
  - Its accuracy is better than straight lines, but is lower than the high-order curve in previous slide.
- For new unknown fishes
  - Its accuracy should be better than those accuracies of straight lines and high-order curve.



- How can a computer program "automatically" find the best classifier?
- We need two things
  - Training algorithm
  - Training data
- What is training algorithm
  - Training algorithm uses training data to find the best classifier. But how?
  - Remember that each classifier has an error rate. And the best classifier has the minimum error rate.
  - We have infinite number of classifiers: infinite straight lines and infinite curves. Each classifier has an error rate.
  - We can find the best classifier only if we calculate all of the error rates of classifiers and find the minimum of these error rate values.
  - But it is a mission impossible.
  - Therefore a lot of complex algorithm are developed to conquer this difficulty.
- Next slide shows some popular training algorithms.

Methods of Classification

- Linear separation
- Bayes classifier
- Support Vector Machines (SVM)
- Clustering
  - Kmeans, GMM(Gaussian Mixture Model)
- Hidden Markov models
- Neural networks
- Deep neural networks

- In this course, we will focus more on neural networks and deep neural networks.

Performance Evaluation

- **In real PR systems, it is not possible to achieve 100% correct recognition**
- **An easy metric to evaluate recognition results: error rate**
  - **The fraction of mis-classifications**
  - **It is the same with : accuracy**

- In most real PR system, it is not possible to achieve 100% correct recognition.
  - For example: OCR, Fingerprint Recognition
  - All we want is to achieve "the lowest possible" error rate.

Error Types

- **False Negative (Type-I Error)**
  - **An object is classified as not the class but it is .**
- **False Positive (Type-II Error)**
  - **An object is classified as the class that it is not.**

Conclusion

- This course teach only the classification step in a PR system
  - Sensing → Sensor course
  - Preprocessing → Digital Image Processing course
  - Segmentation, Feature extraction → Computer Vision course
  - Classification → This course
    - Focus on deep neural network
  - Post processing → Artificial Intelligence course