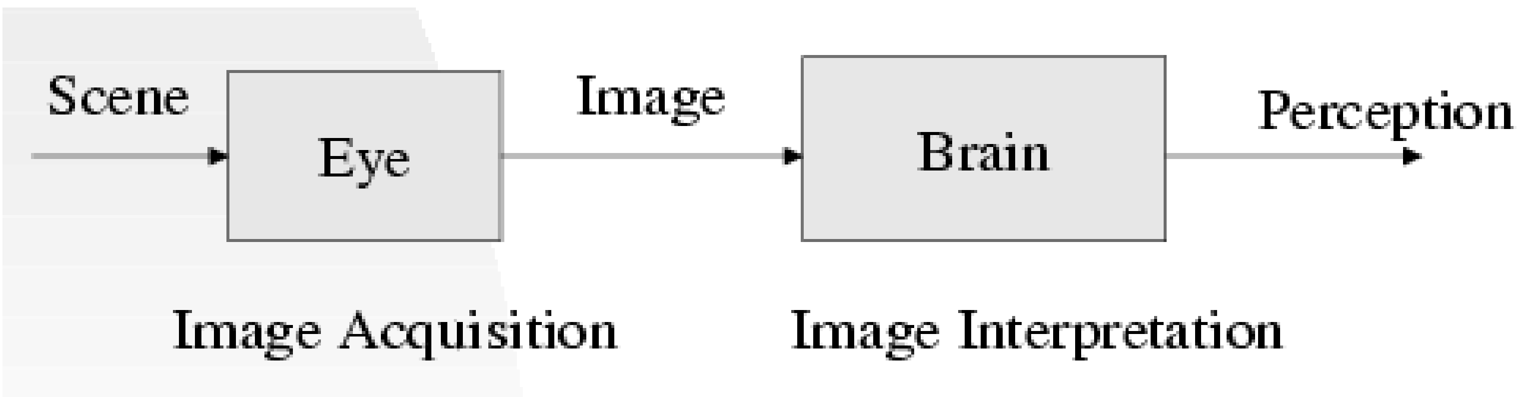
Vision

- Vision (biological) is the process of discovering what is present in the world and where it is by looking.



Biological and Computer Vision

Vision allows humans to perceive and understand as well as recognize the 3D real world surrounding them.

Biological vision → **Eyes** + **Brain**

Eyes → **Sensing**

Brain → **Processing, perception and understanding/recognition**

Computer vision is a duplication of human vision. it is a great trick in extracting description of the 3D world from 2D pictures or sequences of pictures.

The "real world" → **Image capturing** → **Image preprocessing** → **Feature extraction** → **Pattern classification, description as well as interpretation and understanding/recognition**

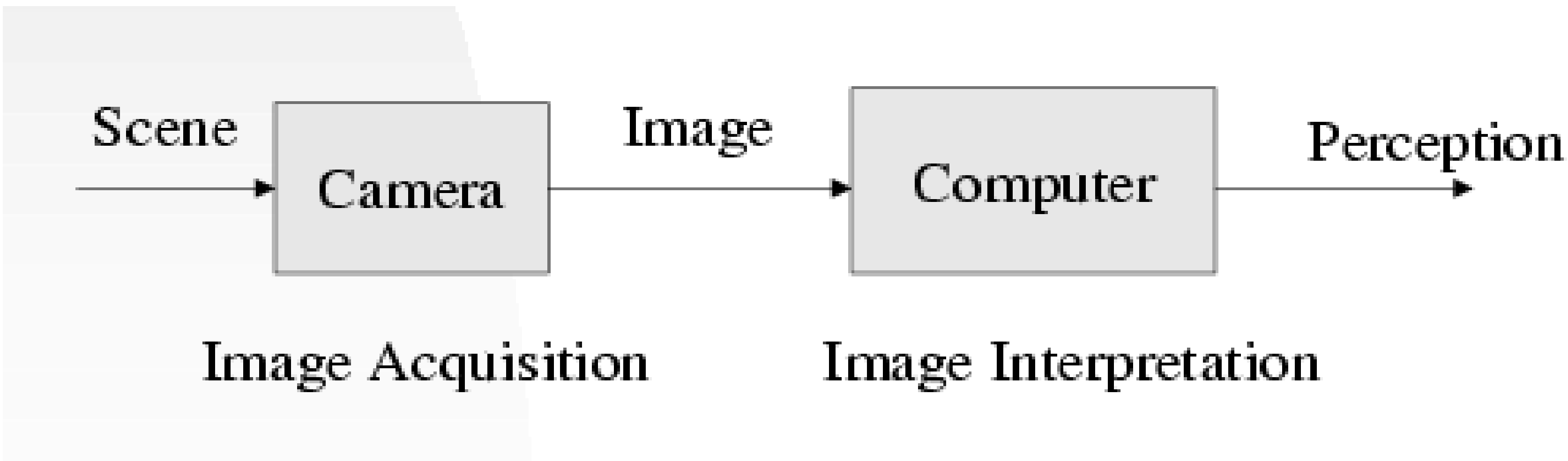
Image capturing: Sensors, camera

Image preprocessing: Enhancing, noise filtering, normalization

Feature extraction: Edge, area, shape, size, texture etc

Computer Vision (1)

- Computer Vision is the study of analysis of pictures and videos in order to achieve results like those as by human.



Computer Vision (2)

- ✓ **Biological vision uses eyes and brain to perceive and recognize 3D world**
- ✓ **Computer vision uses cameras, image processing and AI techniques**

Computer vision is multidisciplinary subject that uses the tools and techniques of mathematics, image processing, pattern recognition, artificial intelligence, psychology, computer science, electronics and other disciplines to obtain similar understanding or recognition results like human vision.

The task of computer vision consists of two steps:

- **Low-level image processing**
  - **uses very little knowledge**
  - **mostly deals with image preprocessing and compression**

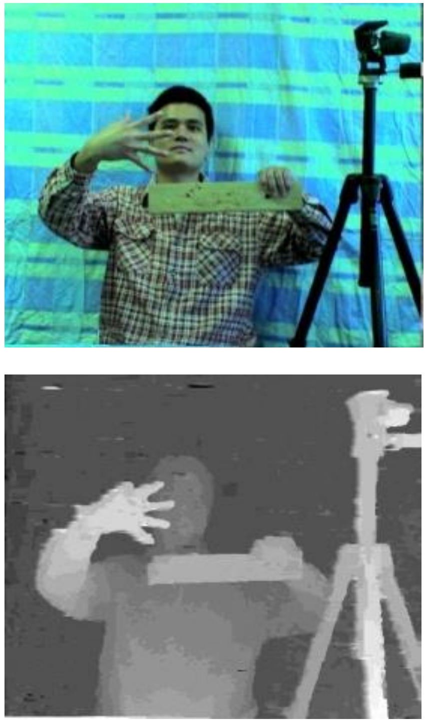
- **High-level image understanding**
  - **based on knowledge, goals, plans and AI tools**
  - **mostly deals with feature extraction and classification, analysis, interpretation and recognition**

Computer Vision (3)

- Automatic understanding of images and video
  1. Computing properties of the 3D world from visual data *(measurement)*
  2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. *(perception and interpretation)*
  3. Algorithms to mine, search, and interact with visual data (*search and organization*)

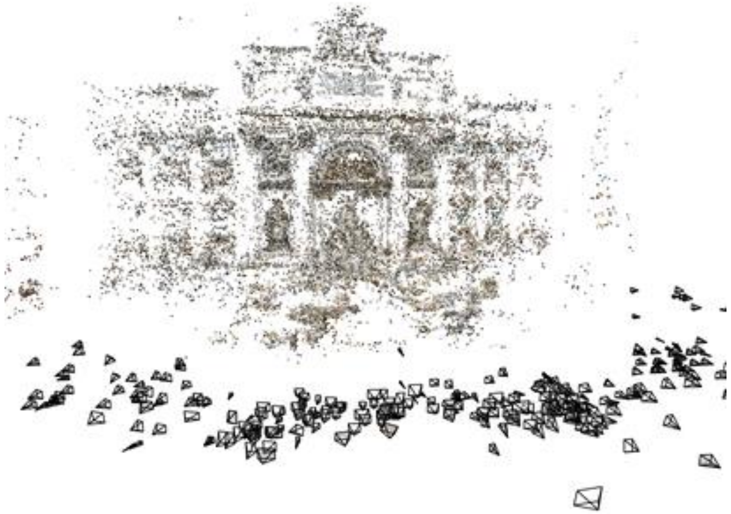
1. Vision for measurement

Real-time stereo



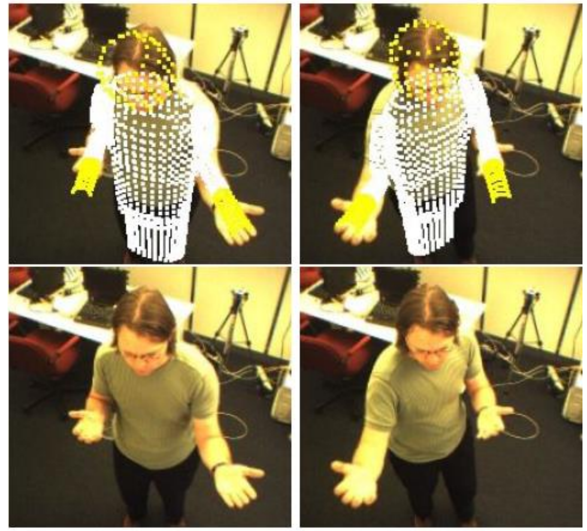
Wang et al.

Structure from motion



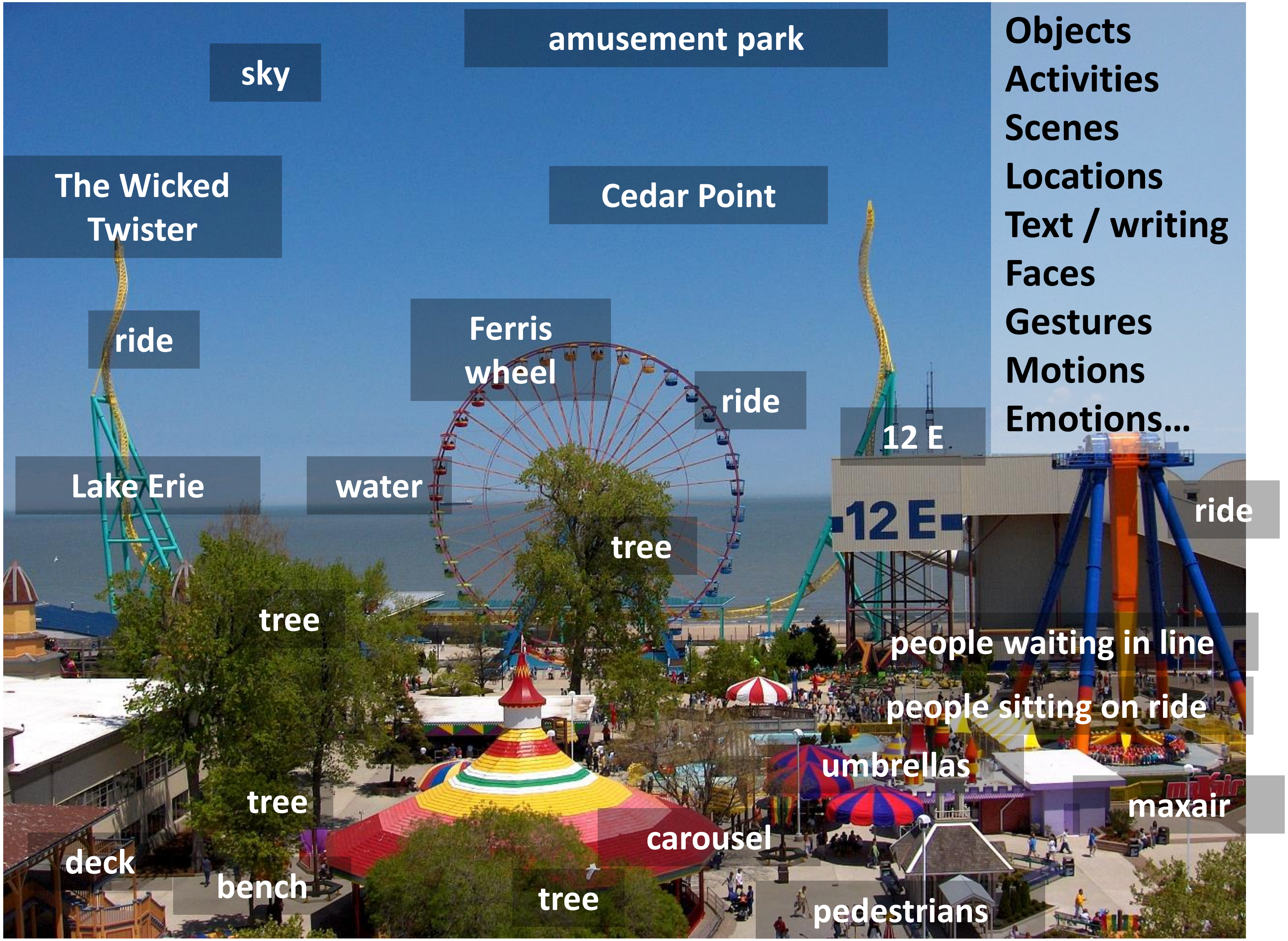
Snavely et al.

Tracking

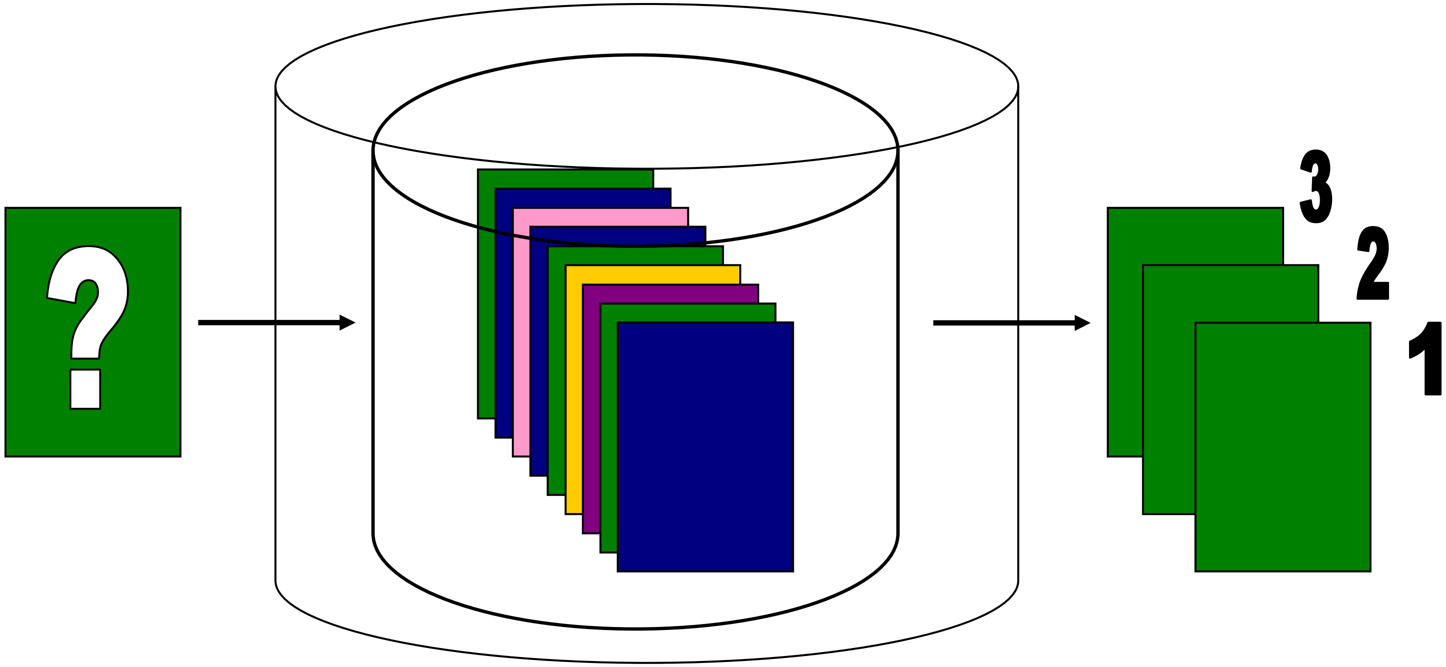


Demirdjian et al.

2. Vision for perception, interpretation



3. Visual search, organization



Query **Image or video archives** **Relevant content**

Example

Finding People in images

Problem 1: Given an image I

Question: Does I contain an image of a person?

“Yes” Instances

Phil Noble / AP


Mike Hewitt / Allsport


Patrick Gardin / AP


Andy Barron / Reno Gazette-Journal


Sydney Morning Herald

“No” Instances


Eric Miller / Reuters


Mark Garkfinkel / The Boston Herald

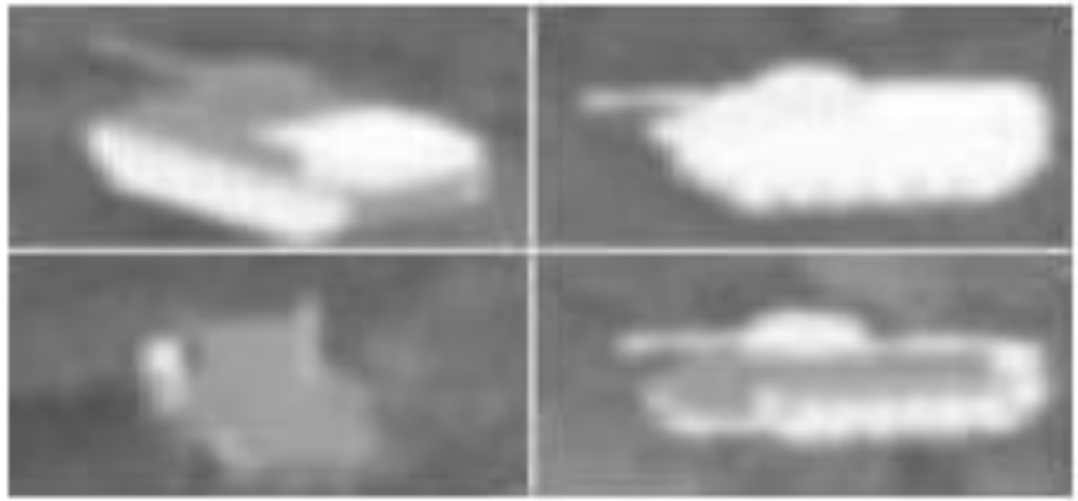
Jeff J. Mitchell / Reuters

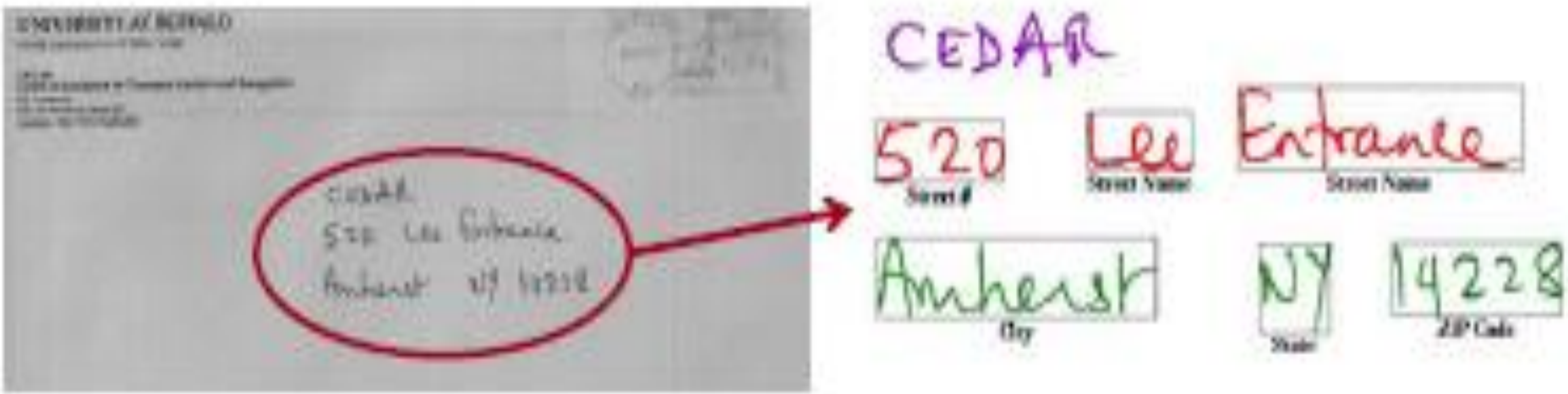

Monroe County Sheriffs Department / Newsmakers


Uno Andersson / AP


NASA via AFP

Applications of Computer Vision (1)





Applications of Computer Vision (2)

- Robot vision: Recognition and interpretation of objects in a scene, auto-navigation
- Industrial automation and machine vision: Parts identification, detection of faults
- Biometrics: Finger-print matching and analysis, face recognition
- Surveillance systems: Intruder tracking and detection, automatic target (missiles, aircrafts, satellites) detection and recognition
- Radar imaging: Automated target (missiles, aircrafts, satellites) detection and identification, Aircraft navigation and landing
- Remote sensing: Weather prediction, environment monitoring, urban growth estimation
- Character recognition: Automated mail sorting, processing of bank checks
- Medical diagnosis: Tumor detection, measurement of size of shape of internal organs etc.

Computer Vision: System Development

- **Data collection**
  - Probably the most time-intensive component of a PR project
  - How many examples are enough?
- **Feature choice**
  - Critical to the success of the PR problem
    - "Garbage in, garbage out"
  - Requires basic prior knowledge
- **Model choice**
  - Statistical, neural and structural approaches
  - Parameter settings
- **Training**
  - Given a feature set and a "blank" model, adapt the model to explain the data
  - Supervised, unsupervised and reinforcement learning
- **Evaluation**
  - How well does the trained model do?
  - Overfitting vs. generalization

Practical Situation Handling

- Consider the following scenario
  - A fish processing plant wants to automate the process of sorting incoming fish according to species (salmon or sea bass)
  - The automation system consists of
    - a conveyor belt for incoming products
    - two conveyor belts for sorted products
    - a pick-and-place robotic arm
    - a vision system with an overhead CCD camera
    - a computer to analyze images and control the robot arm

Cont. (1)

- Problem Analysis
  - Set up a camera and take some sample images to extract features
    - Length
    - Lightness
    - Width
    - Number and shape of fins
    - Position of the mouth, etc…

    - This is the set of all suggested features to explore for use in our classifier!

Cont. (2)

- Preprocessing
  - Use a segmentation operation to isolate fishes from one another and from the background
- Information from a single fish is sent to a feature extractor whose purpose is to reduce the data by measuring certain features
- The features are passed to a classifier

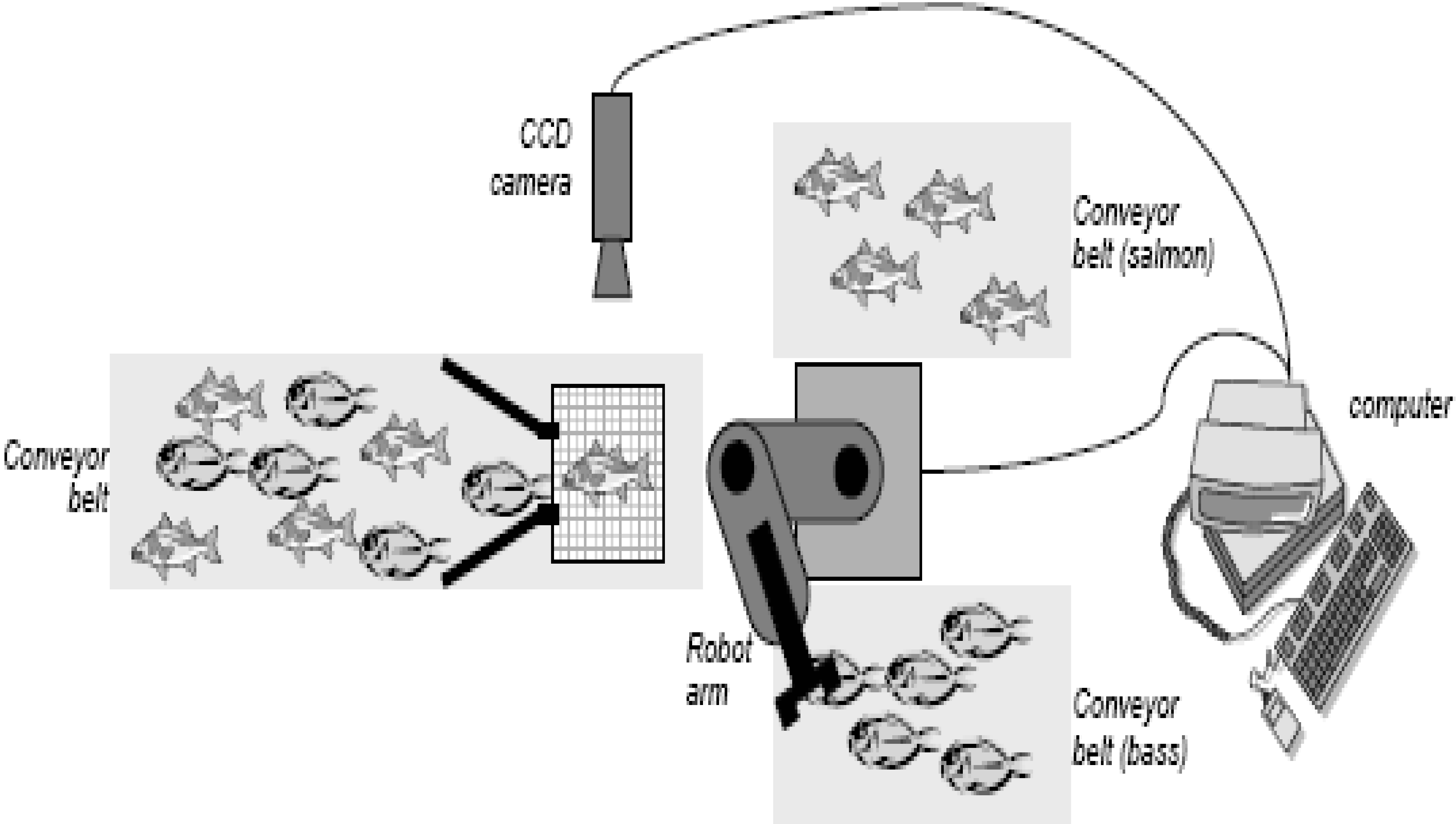
Cont. (3)


Preprocessing
Feature extraction
Classification
"salmon"
"sea bass"

Cont. (4)

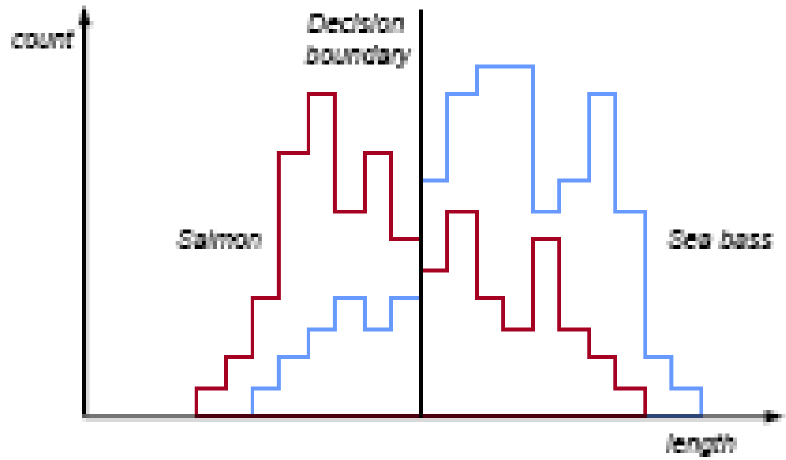
- Classification
  - Select the length of the fish as a possible feature for discrimination

Cont. (5)

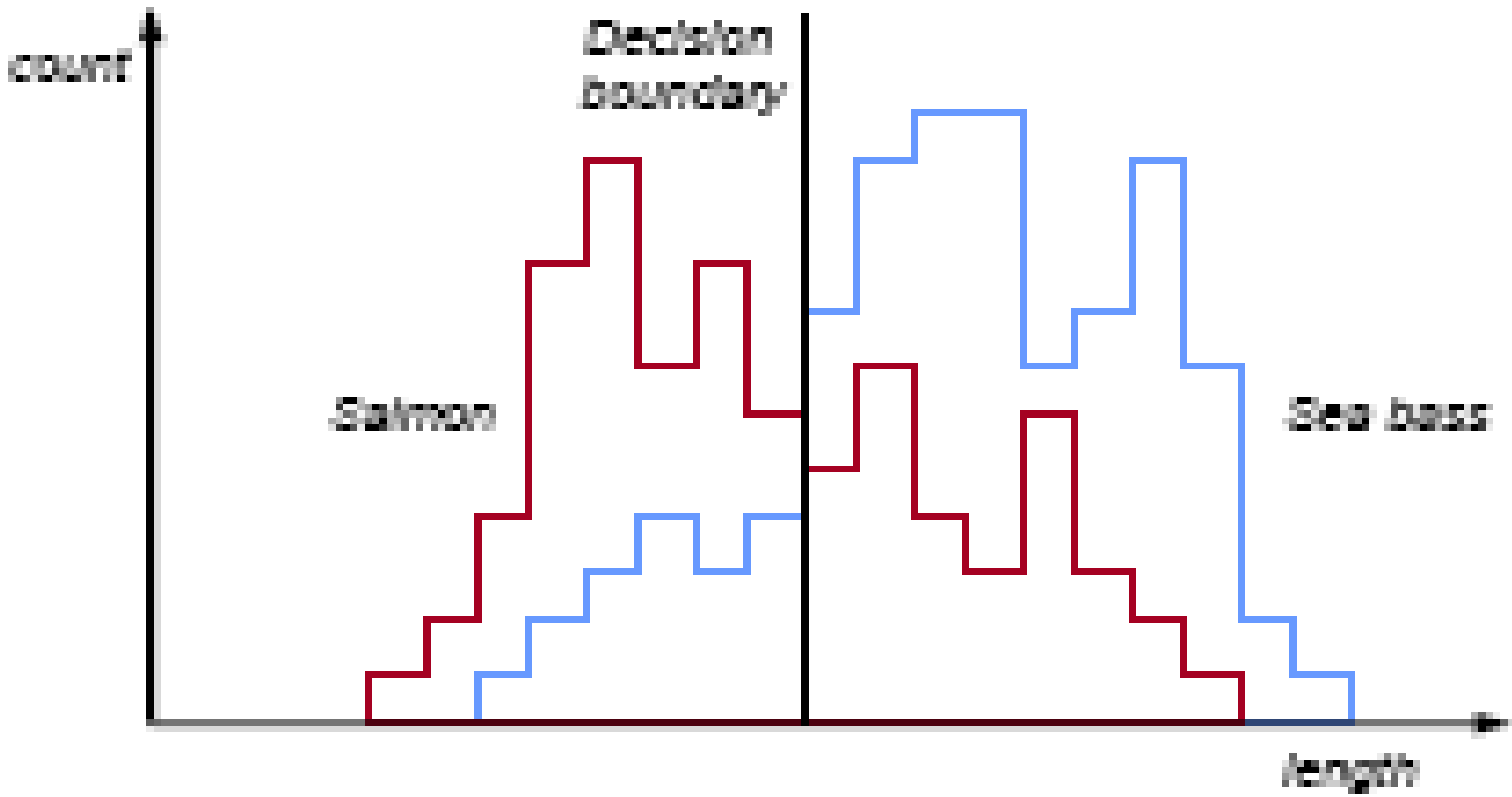

CCD
camera
Conveyor
belt (salmon)
computer
Conveyor
belt
Robot
arm
Conveyor
belt (bass)

Cont. (6)

- **Sensor**
  - The vision system captures an image as a new fish enters the sorting area
- **Preprocessing**
  - Image processing algorithms
    - adjustments for average intensity levels
    - segmentation to separate fish from background
- **Feature Extraction**
  - Suppose we know that, on the average, sea bass is larger than salmon
    - From the segmented image we estimate the length of the fish
- **Classification**
  - Collect a set of examples from both species
  - Compute the distribution of lengths for both classes
  - Determine a decision boundary (threshold) that minimizes the classification error
  - We estimate the classifier's probability of error and obtain a discouraging result of 40%
  - What do we do now?



Cont. (7)



Classification rate of success= 40%

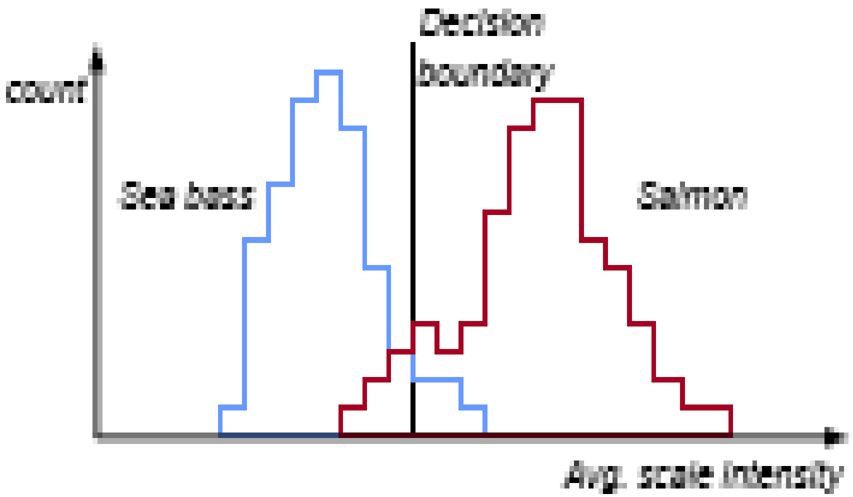
using only length feature!

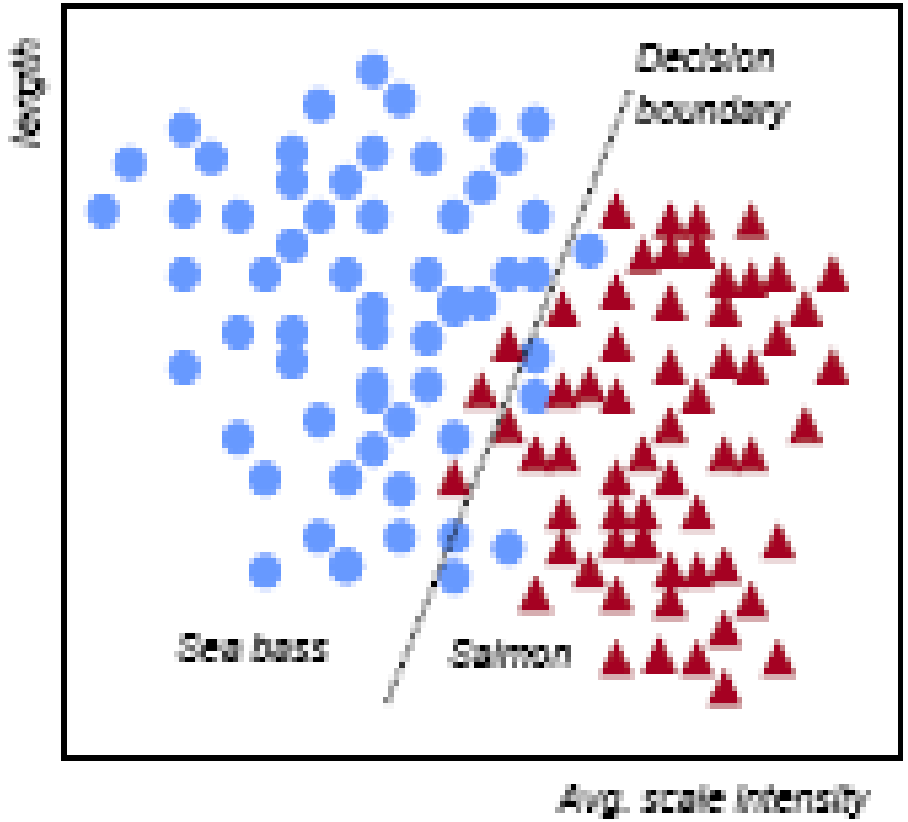
Very discouraging results!!!

What do we do now???

Cont. (8)

- Improving the performance of our PR system
  - Determined to achieve a recognition rate of 95%, we try a number of features
    - Width, Area, Position of the eyes w.r.t. mouth...
    - only to find out that these features contain no discriminatory information
  - Finally we find a "good" feature: average intensity of the scales
  - We combine "*length*" and "*average intensity of the scales*" to improve class separability
  - We compute a linear discriminant function to separate the two classes, and obtain a classification rate of 95.7%

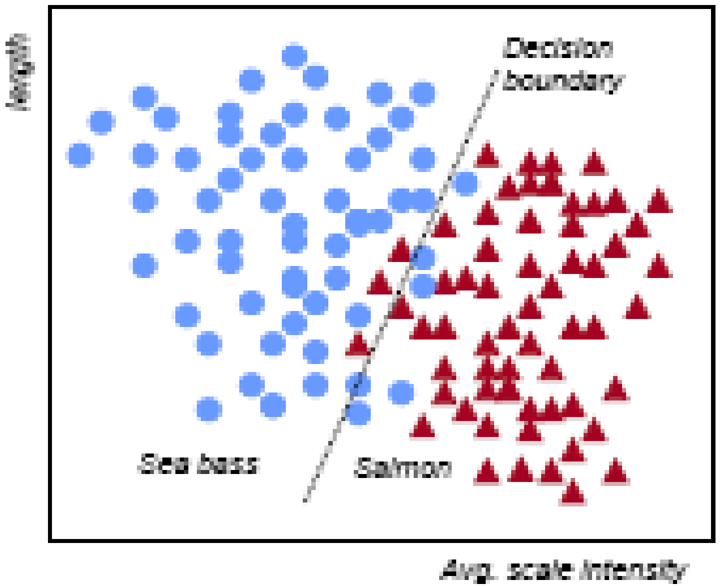


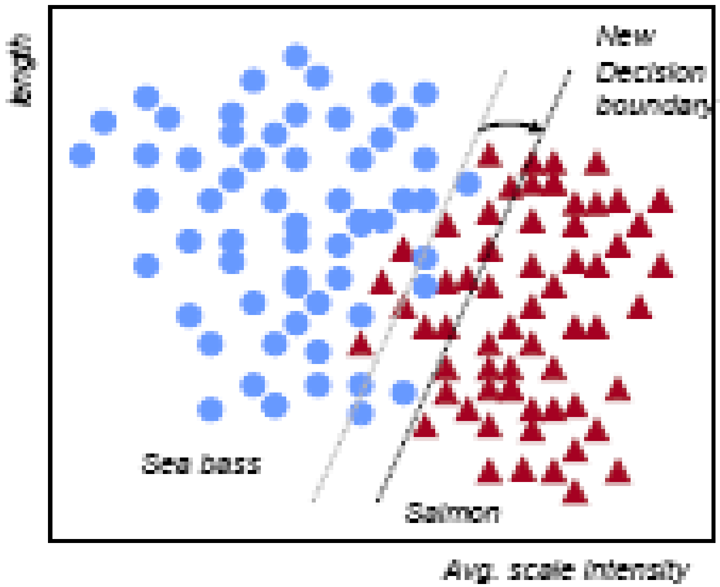


Classification rate of success= 95.7%
using two features (length and Av. Intensity)

Cont. (10)

- **Cost Versus Classification rate**
  - Our linear classifier was designed to minimize the overall misclassification rate
  - Is this the best objective function for our fish processing plant?
    - The cost of misclassifying salmon as sea bass is that the end customer will occasionally find a tasty piece of salmon when he purchases sea bass
    - The cost of misclassifying sea bass as salmon is an end customer upset when he finds a piece of sea bass purchased at the price of salmon
  - Intuitively, we could adjust the decision boundary to minimize this cost function





Cont. (11)

- **The issue of generalization** **Machine Learning**
  - The recognition rate of our linear classifier (95.7%) met the design specs, but we still think we can improve the performance of the system
    - We then design an artificial neural network with five hidden layers, a combination of logistic and hyperbolic tangent activation functions, train it with the Levenberg-Marquardt algorithm and obtain an impressive classification rate of 99.9975% with the following decision boundary