Foundations of Information Processing

LUT University

Introduction to Information Processing

Course: Information Processing

Algorithmic problem solving:

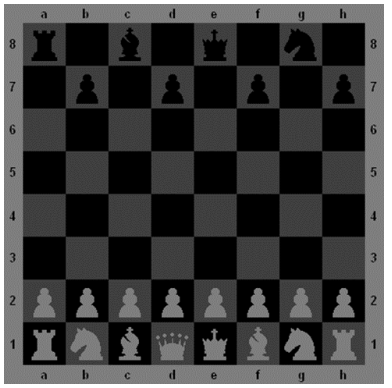
- Introduction to Information Processing.
- General and algorithmic problem solving.
- Algorithms: development and parts.
- Modular, iterative, and recursive algorithms, and design principles.
- Complexity of algorithms.
- Search problems and playing games.

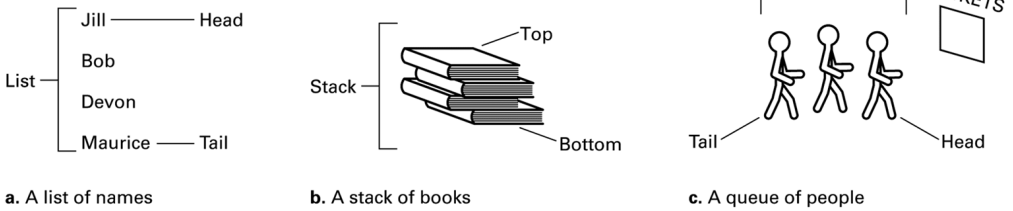
Data and its conversion:

- Data encoding.
- Information and data compression.
- Data structures.
- Data encryption.
- Propositional logic and reasoning.
- Compilation.

Information Processing: fundamentals to learn

```
MODULE factorial(n)
   k := 1
   WHILE n > 1 DO
      k := k * n
      n := n –1
   ENDWHILE
RETURN k
ENDMODULE
```

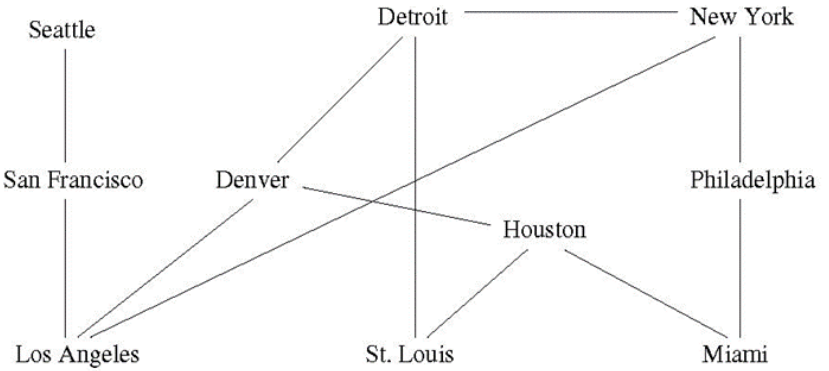




a. A list of names

b. A stack of books

c. A queue of people



$$i(A) = \log_b \frac{1}{P(A)} = -\log_b P(A)$$

```
  0      1      0      1
+ 0    + 0    + 1    + 1
---    ---    ---    ---
  0      1      1     10
```

p	q	p→q	q→p	(p→q) ∧ (q→p)
E	E	T	T	T
E	T	T	E	E
T	E	E	T	E
T	T	T	T	T

The university system: a short introduction

The introduction to the university system, especially including

- education and research.

LUT as an example about the universities.

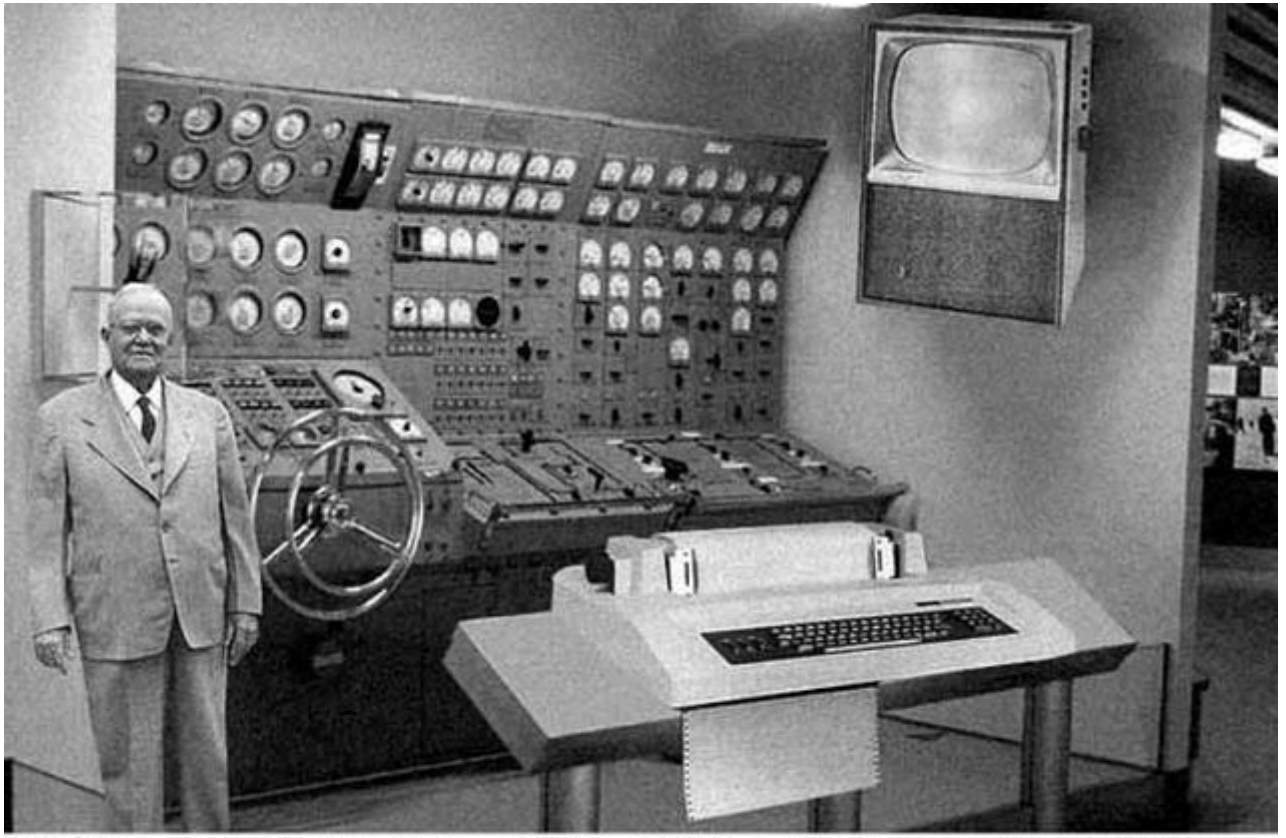
- Get familiar with the LUT page.
- What kind of education is given?
- What are the research areas?

Department of Computational Engineering as an example about the departments: get familiar with the COPE page, especially with the focus areas.

Computer Vision and Pattern Recognition Laboratory (CVPRL) as an example about the information processing research groups.

- Get familiar with the CVPRL page.
- Education: Computer Vision and Pattern Recognition at the master and doctoral level.
- Research: Computer Vision, Machine Learning, Pattern Recognition, Data Analytics, and Applications of Digital Image Processing and Analysis.

Who could have succeeded to predict the development of information processing?



Scientists from the RAND Corporation have created this model to illustrate how a "home computer" could look like in the year 2004. However the needed technology will not be economically feasible for the average home. Also the scientists readily admit that the computer will require not yet invented technology to actually work, but 50 years from now scientific progress is expected to solve these problems. With teletype interface and the Fortran language, the computer will be easy to use.

Figure: The vision about the future computer in 2004, predicted in 1954. (Source: Home computer year 2004. Popular Mechanics, 1954.)

Nowadays information processing is everywhere, as integrated and usually in a very small scale.

Large complicated digital solutions are also needed: For example:

https://www.youtube.com/watch?v=WhZVOR9_6Rg

https://www.youtube.com/watch?v=1g4l89SajcA

What is data? How to process data?

Senses, memory, cognition.

Data, information, knowledge:

For example, 39 (data), temperature (information), high fever (knowledge).

Hypotheses, experiments, observations, analyses.

Objects, interpretations, conclusions.

What is the meaning of data?

Does a piece of data contain value?

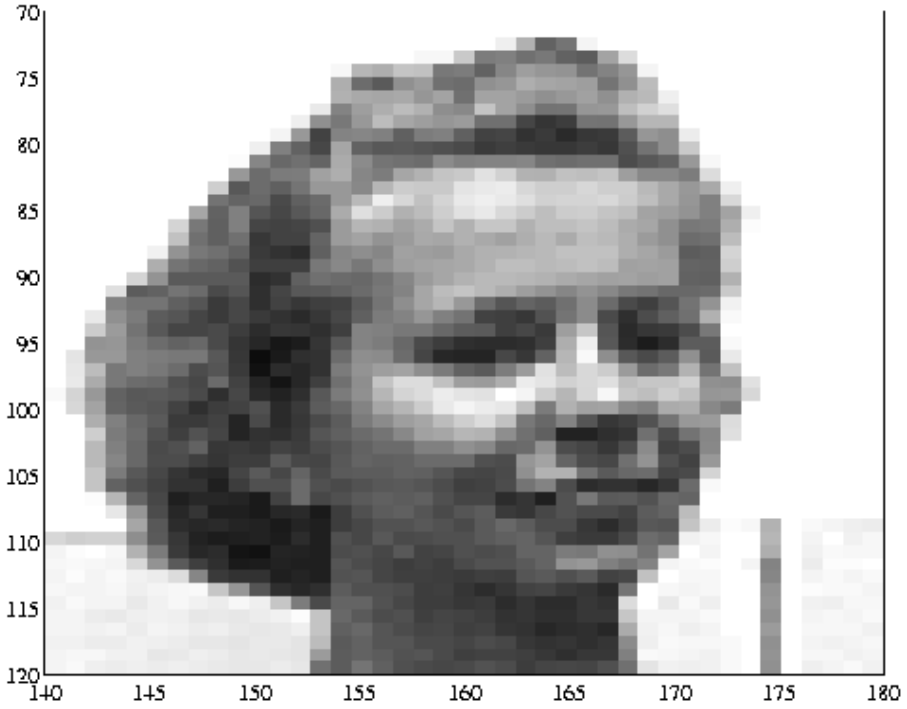
Does the human being act and conclude only based on data?

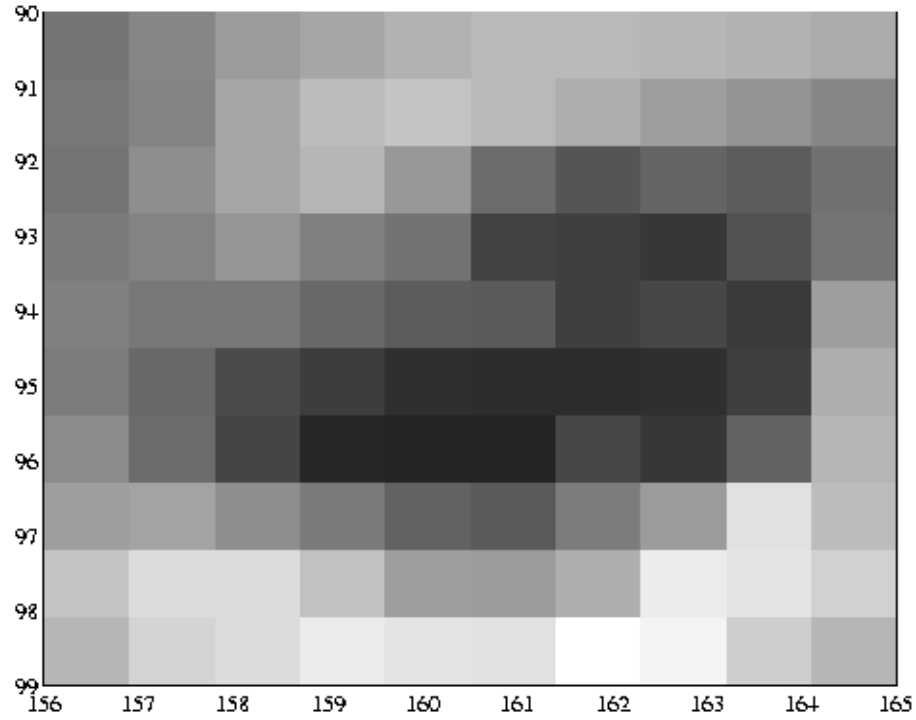
What does information processing mean actually,

and what is included into it?

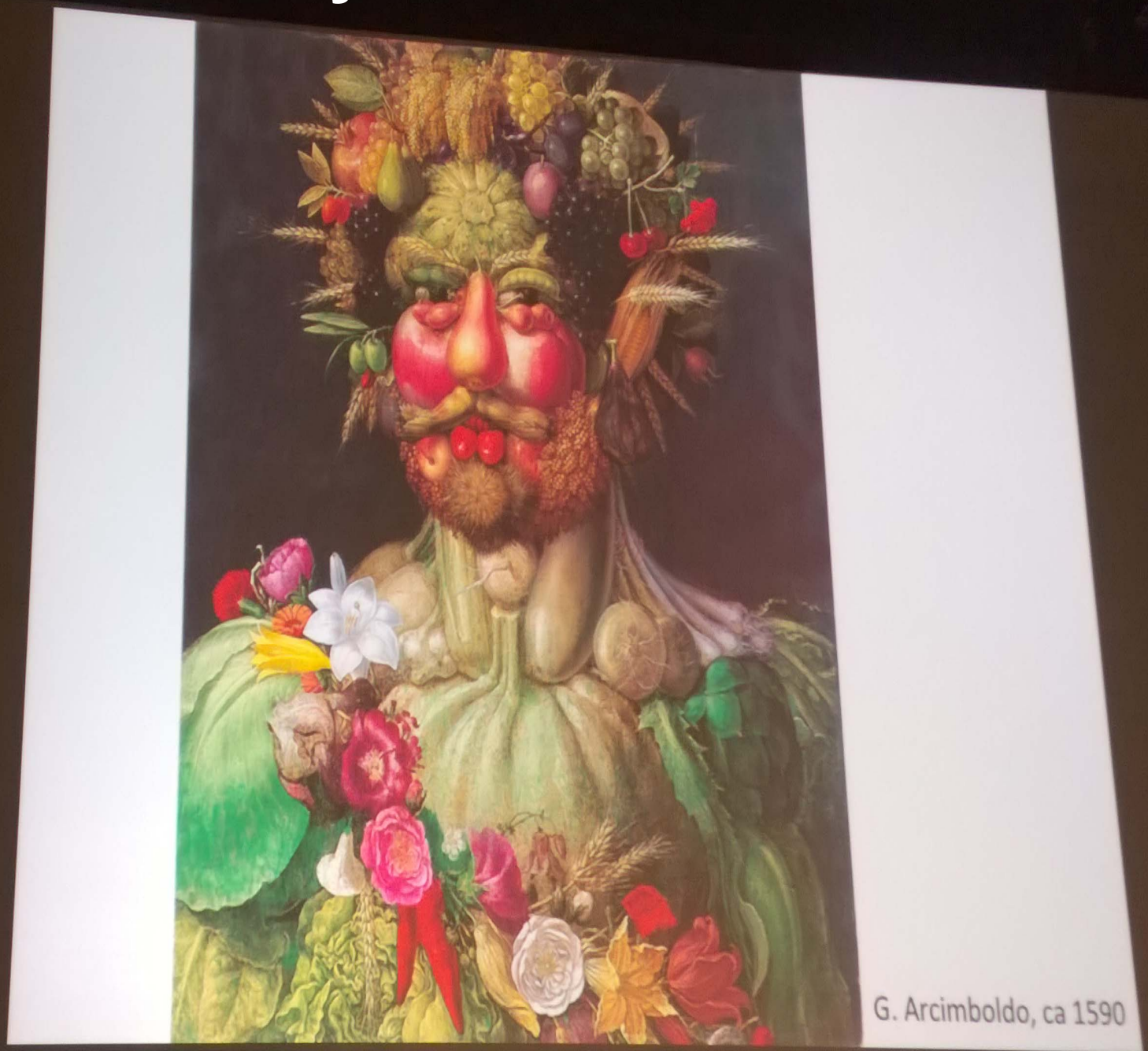
Information processing is challenging

For example, Computer Vision (also called Machine Vision) “sees” differently than the human, but it should understand the image as the human. What is useful data?



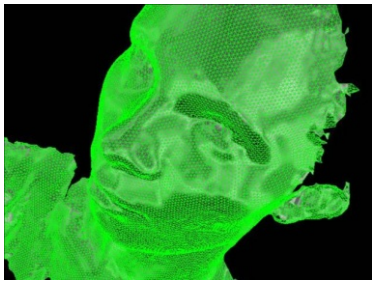


What do you see? What information?



Source: Prof. Fei-Fei Li at ICPR 2014

The world is three-dimensional: 3D Face Recognition

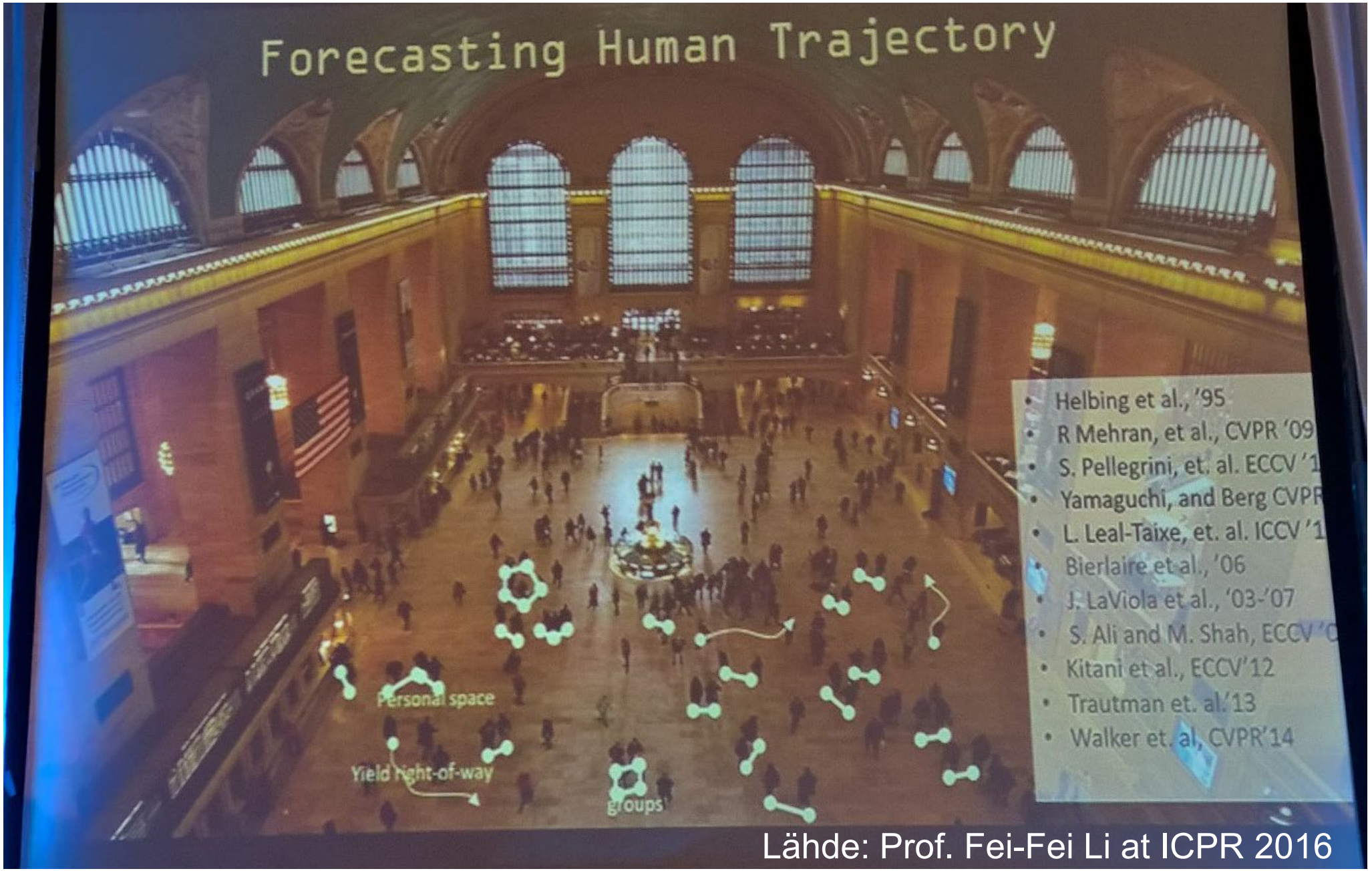


Future biometric passport and video analysis in surveillance and security

Link to the video

Link to the video

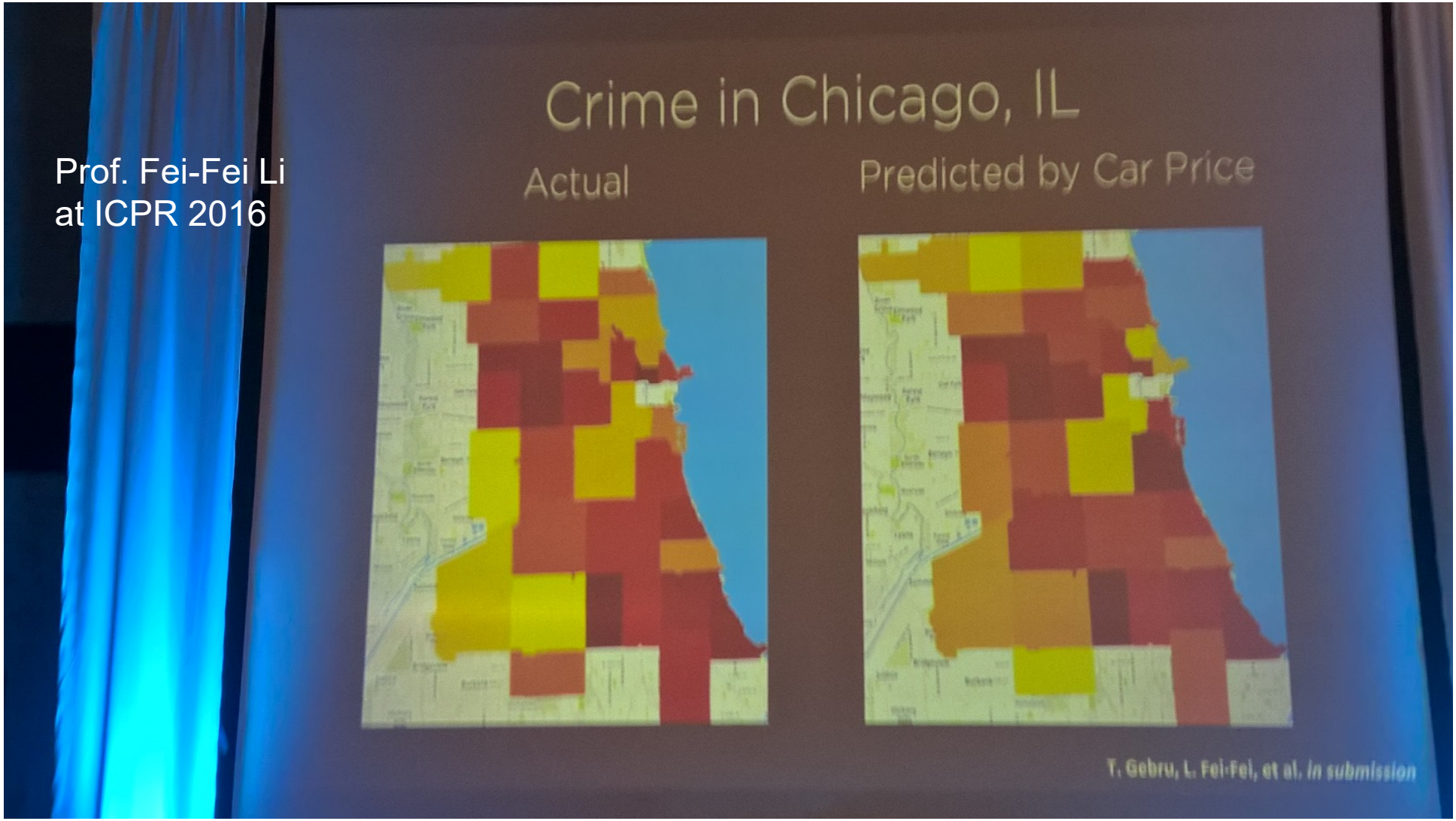
Extra challenge: many different moving objects



Lähde: Prof. Fei-Fei Li at ICPR 2016

[Link to the video]

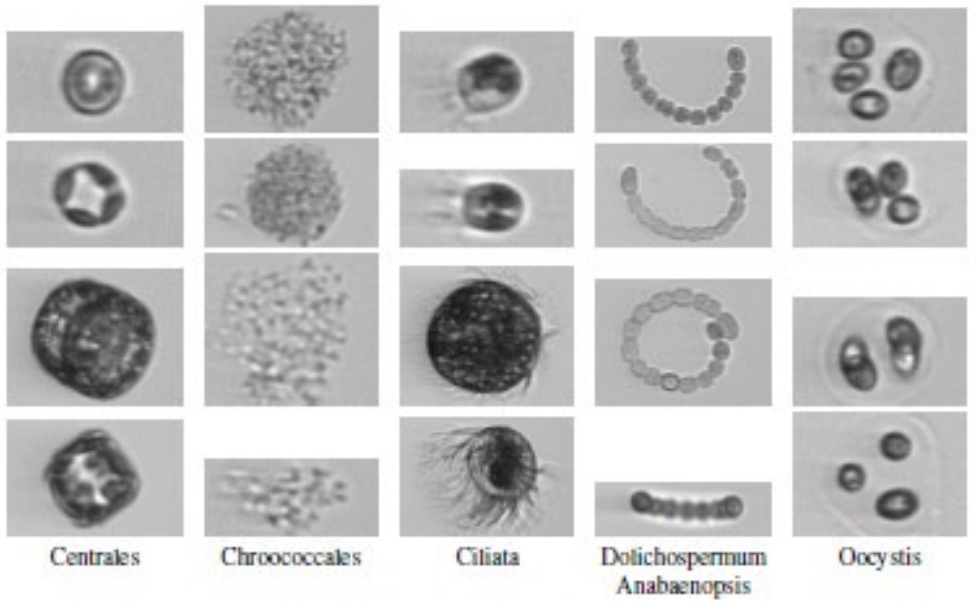
Another challenge in data mining:
What are correlations? What are causalities?

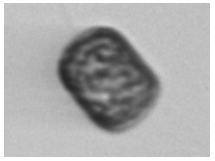


More challenges in information processing:

Who does know the true content of the data (ground truth) and can/is fast enough to mark it (annotation)?

FastVision: automatic recognition of plankton types by learning from annotations by biologists, training convolutional neural networks (machine learning).

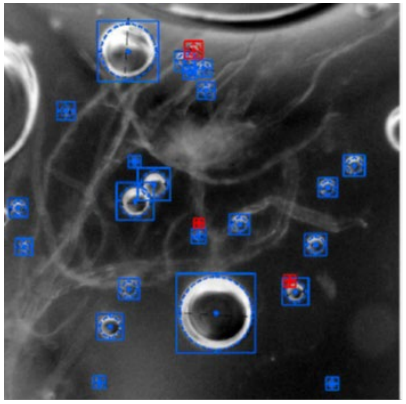


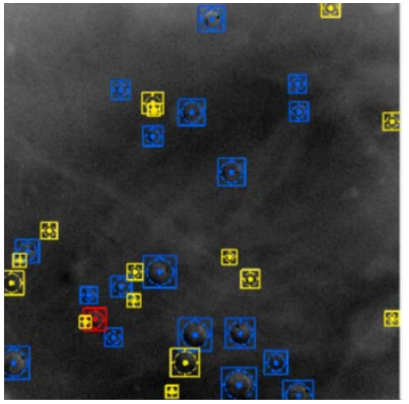


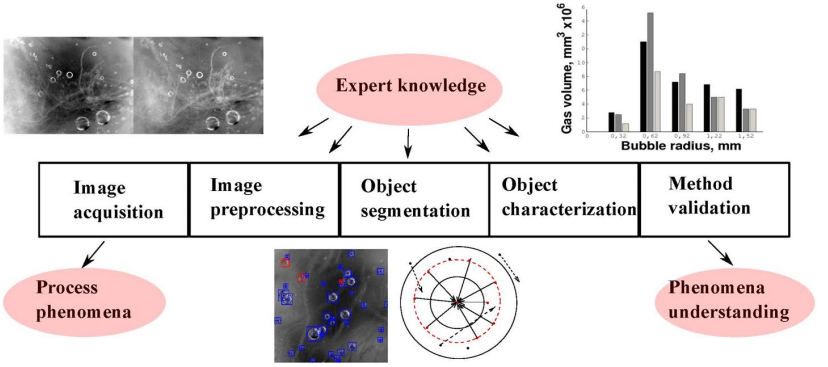
=> what plankton type?

Further in future:
machine teaching

PulpVision: automatic detection of bubbles and their distributions based on annotations by engineers of pulping.

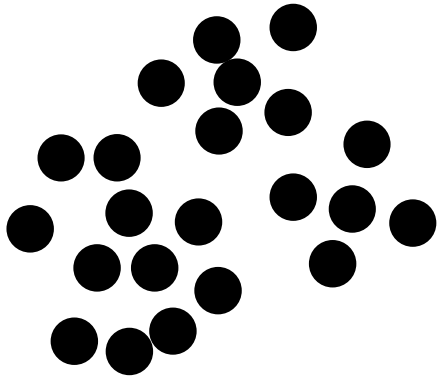






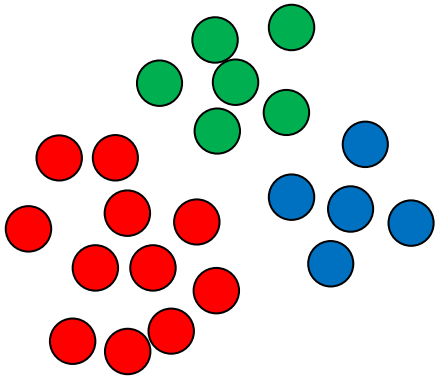
Clustering and Classification

Data



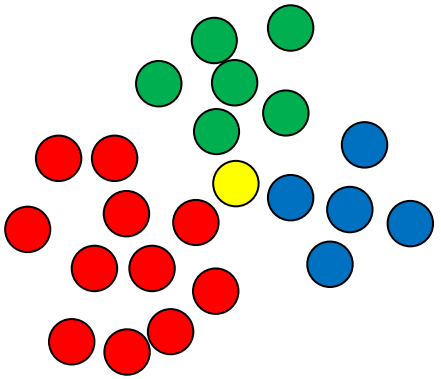
Expert knowledge about the classes
=>
3 classes

Clustering the data into the classes

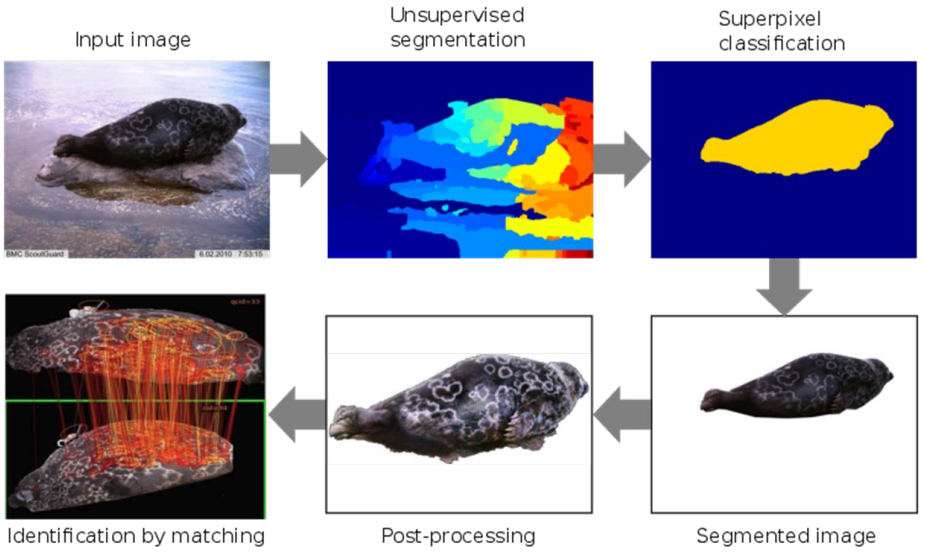


A new sample

Classification:
What is the class for the new sample?
Does it belong to , , or ?



SealVision& CoExist: could you help me by recognizing me?



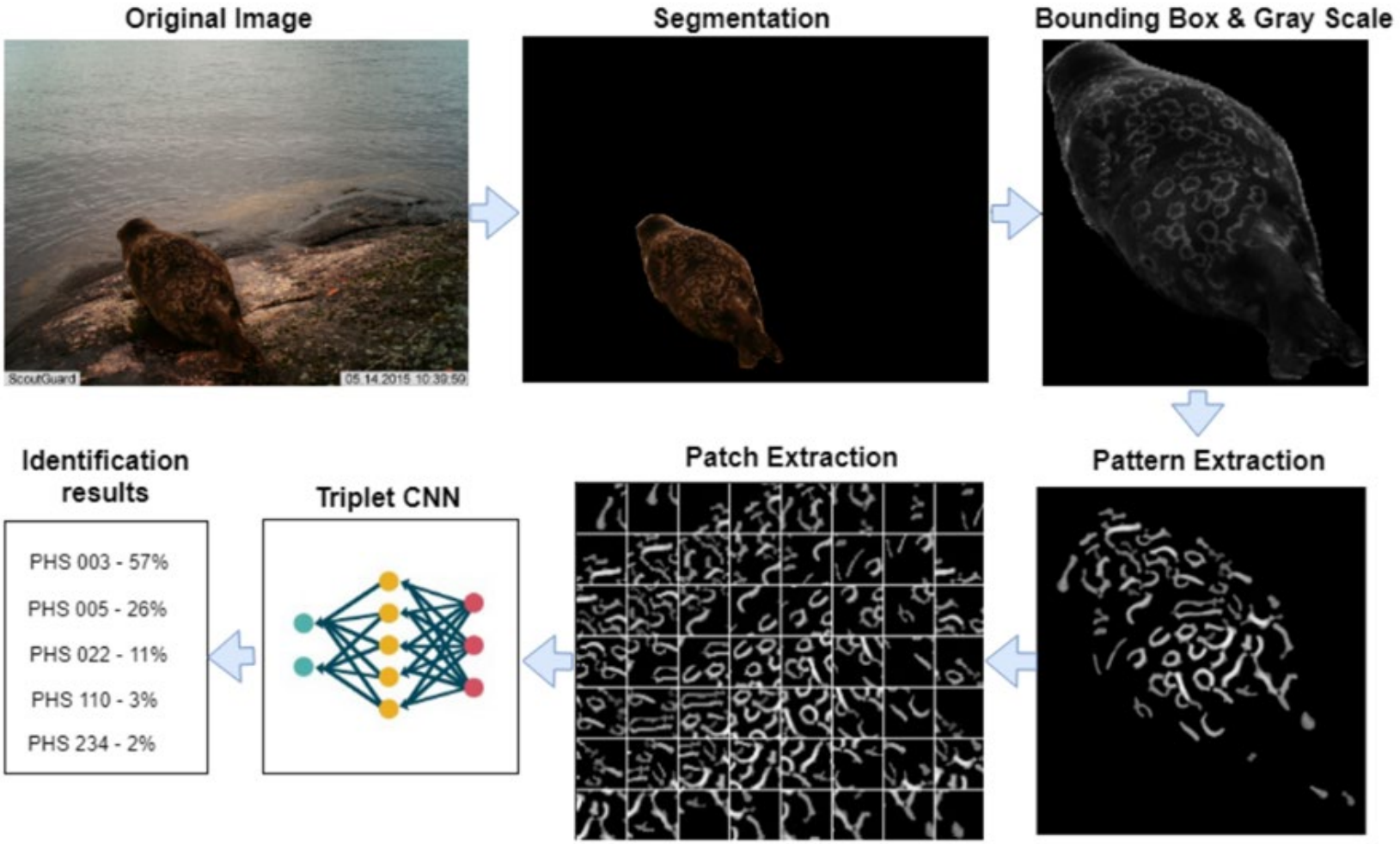
Detection and identification of **individual Saimaa ringed seals** based on the fur pattern using machine vision for conservation of nature.

Only around 400 Saimaa ringed seals (“saimaannorppa” in Finnish) left in Lake Saimaa.

“Biometric passport” for seals (wild life photo ID).

Collaboration: Department of Biology, University of Eastern Finland & BFNC, Russia.

Classification: training a Convolutional Neural Network (CNN) to learn fur patterns of each seal, by "showing" several images (samples) => automatic identification of ringed seal individuals (machine learning/deep learning)

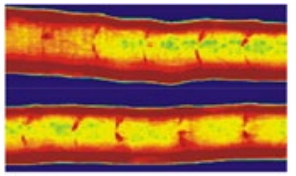


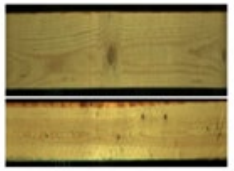
CVPRL: research projects at LUT

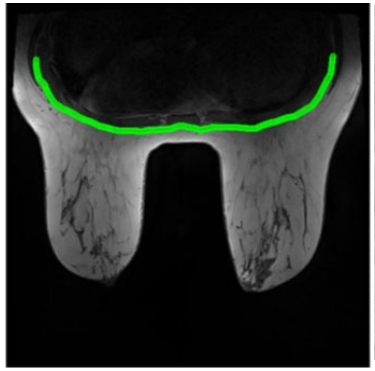
Computer Vision and Pattern Recognition Laboratory:

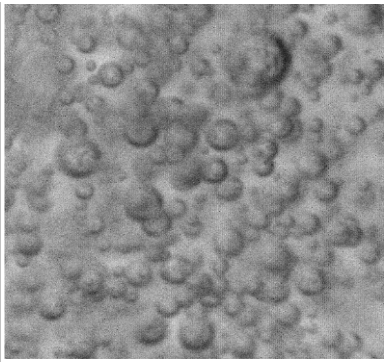
Applications of Computer Vision,
Digital Image Processing and Analysis,
Data Analytics.

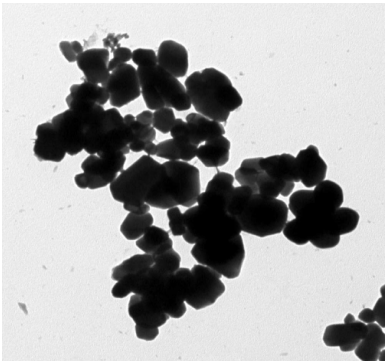

Laser scan (log surface)

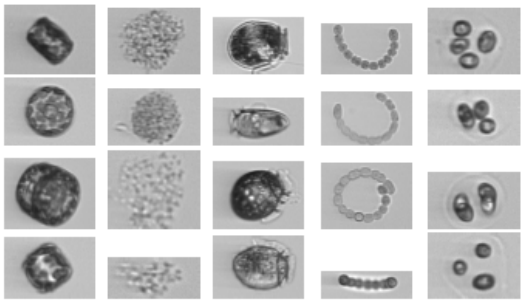

X-ray images (log)


RGB image (sawn timber)

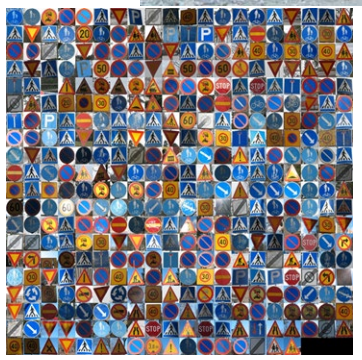


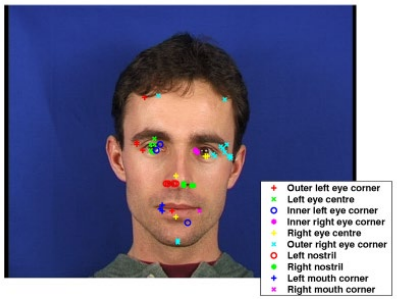


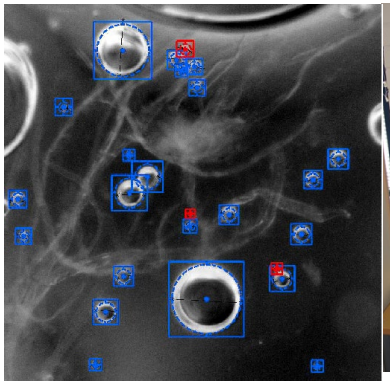


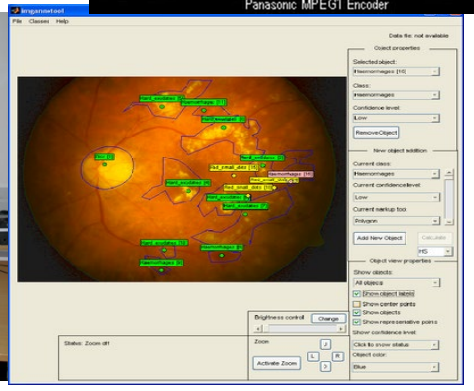


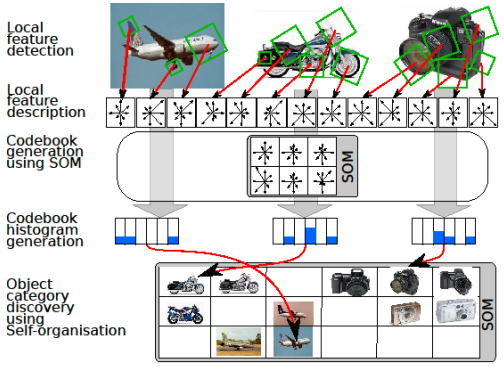

Panasonic MPEG1 Encoder











Information Processing

- The term “Information Processing” covers a wide range of different actions related to computers and their usage.
- History of definitions:
  - *Computer Science* (in the 1950s).
  - *Information Processing* (in the 1950s).
  - *Artificial Science* (vs. Natural Science, 1969, Herbert A. Simon: Nobel Prize, Turing Award, UCB, Carnegie Mellon U, Illinois IT).
  - *Algorithmics* (1968, Donald E. Knuth, 1968: Turing Award, Stanford U, Caltech).
- This course is about *Information Processing*, considering the content.
  - From an engineering point of view:

    Computer Science and Engineering.
- There is a very wide need of applications: Computer Vision, Pattern Recognition, Machine Learning, Data Science, Artificial Intelligence, IoT, Industrial Internet… Find more, for example, from the top conferences:
  - CVPR: http://cvpr2021.thecvf.com/
  - ICCV: http://iccv2021.thecvf.com/
  - ECCV: https://eccv2020.eu/

Definitions: Information Processing & Computer Science

- **Information Processing***:*
  ”The processing of information (by a machine or by an organism) so as to yield new or more useful information.”
  (Oxford English Dictionary ,OED)

- **Computer Science***:*
  ”The branch of knowledge concerned with the construction, programming, operation, and use of computers.” (OED)

Practical definition:

- Information processing = theories & hardware & software.

ACM/IEEE Computing Curricula:

Information Processing mostly in Computer Science.

ACM/IEEE Computing Curricula, Computer Science

2001:

- Algorithms and Complexity (AL) .
- Architecture and Organization (AR).
- Computational Science (CN).
- Discrete Structures (DS).
- Graphics and Visual Computing (GV).
- Human-Computer Interaction (HC).
- Information Management (IM).
- Intelligent Systems (IS).
- Net-Centric Computing (NC).
- Operating Systems (OS).
- Programming Fundamentals (PF).
- Programming Languages (PL).
- Social and Professional Issues (SP).
- Software Engineering (SE).

2013:

- Algorithms and Complexity (AL).
- Architecture and Organization (AR).
- Computational Science (CN).
- Discrete Structures (DS).
- Graphics and Visualization (GV).
- Human-Computer Interaction (HCI).
- Information Assurance and Security (IAS)
- Information Management (IM).
- Intelligent Systems (IS).
- Networking and Communications (NC).
- Operating Systems (OS).
- Platform-based Development (PBD).
- Parallel and Distributed Computing (PD).
- Programming Languages (PL).
- Software Development Fundamentals (SDF).
- Software Engineering (SE).
- Systems Fundamentals (SF).
- Social Issues and Professional Practice (SP).

ACM Computing Curricula 2020

Paradigms for global computing education:

1. Computer Engineering
2. Computer Science
3. Cybersecurity
4. Information Systems
5. Information Technology
6. Software Engineering
7. with data science

http://www.acm.org/education/curricula-recommendations

Summary

- The concept of *Information Processing* is very many-sided:
  - One of the main objectives of the course is to learn and understand *basic concepts* related to it.
  - After this course, the course “Fundamentals of Computer Science” will continue further in the spring term.
- The main goal is to apply computers and software for some selected purpose of information processing.
- The need for new applications and novel innovations has been increasing rapidly in the last decade.
  - The current ones are developing and the new ones coming.
  - Autonomous vehicles (cars, airplanes, drones, etc.), personally customized healthcare, smart houses, service robots, “genuinely” intelligent and integrated devices which can “sense” and “think”, etc.
  - Convolutional neural networks: machine learning (deep learning), machine teaching.