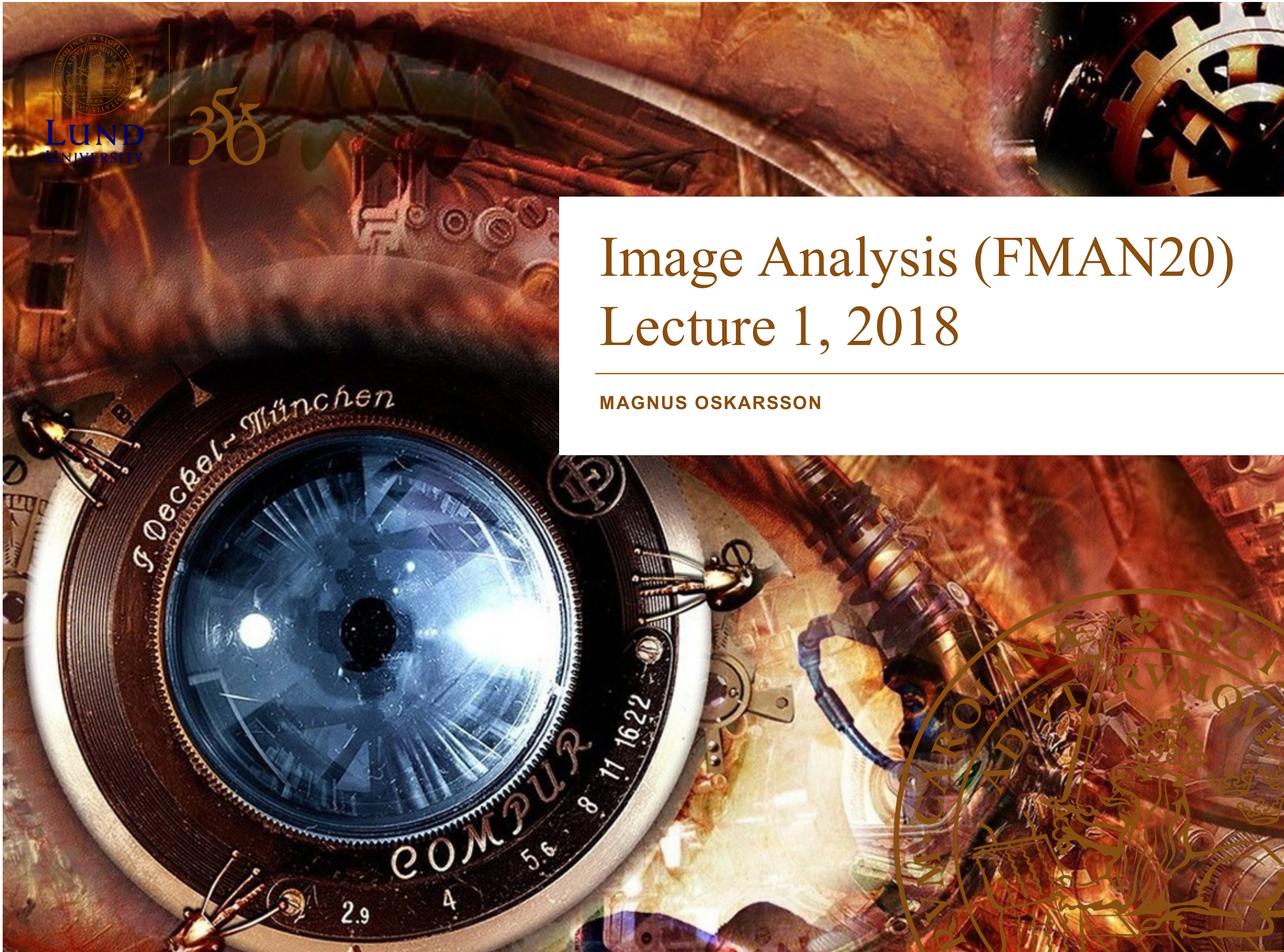
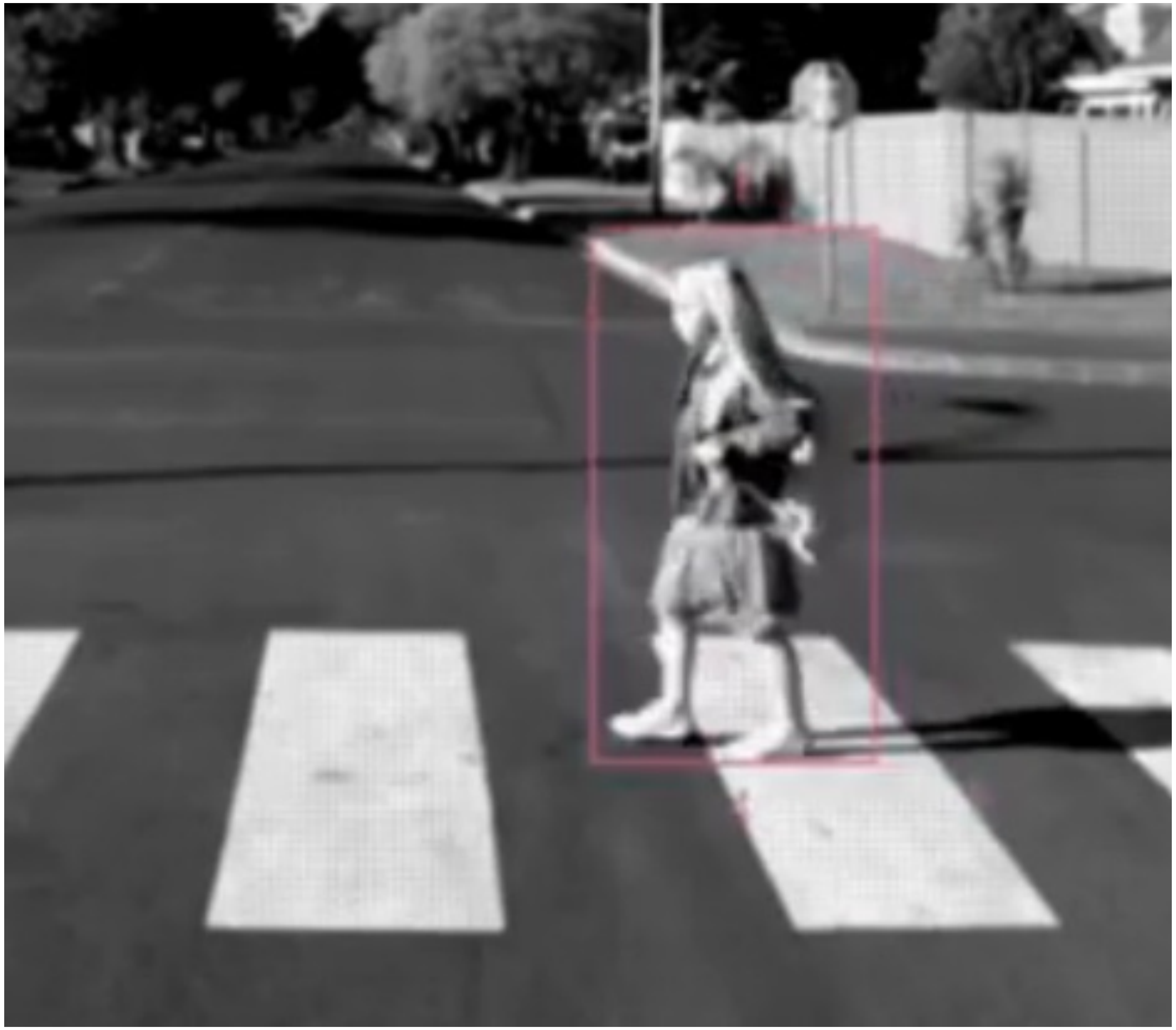


Image Analysis (FMAN20) Lecture 1, 2018

MAGNUS OSKARSSON

Image Analysis - Motivation

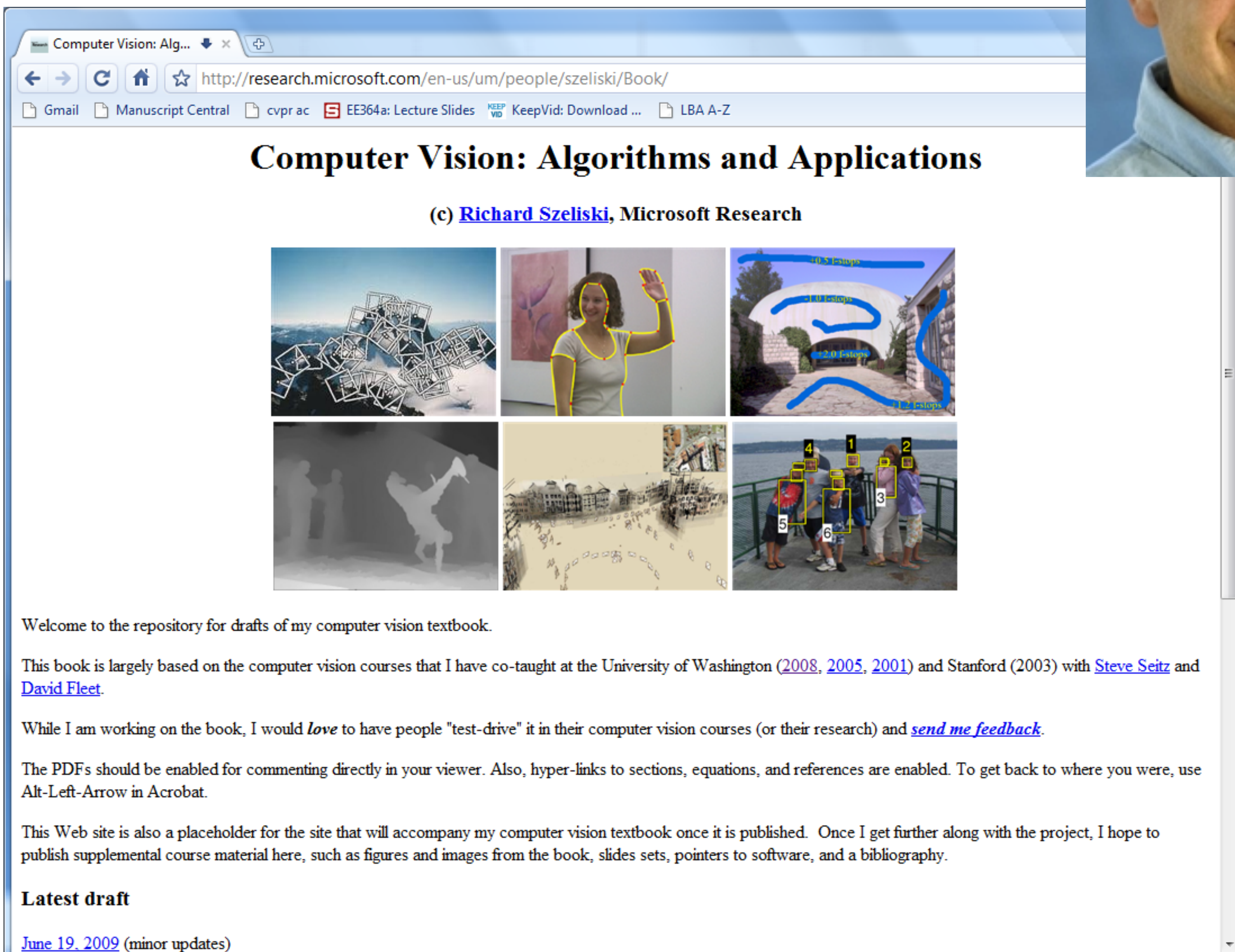
Image Analysis - Motivation



Overview

1. Image Analysis – examples
2. Image Models (continuous vs discrete)
3. Sampling and interpolation
4. Discrete geometry

PrimaryText



Computer Vision: Algorithms and Applications

(c) Richard Szeliski, Microsoft Research

Welcome to the repository for drafts of my computer vision textbook.

This book is largely based on the computer vision courses that I have co-taught at the University of Washington (2008, 2005, 2001) and Stanford (2003) with Steve Seitz and David Fleet.

While I am working on the book, I would ***love*** to have people "test-drive" it in their computer vision courses (or their research) and ***send me feedback***.

The PDFs should be enabled for commenting directly in your viewer. Also, hyper-links to sections, equations, and references are enabled. To get back to where you were, use Alt-Left-Arrow in Acrobat.

This Web site is also a placeholder for the site that will accompany my computer vision textbook once it is published. Once I get further along with the project, I hope to publish supplemental course material here, such as figures and images from the book, slides sets, pointers to software, and a bibliography.

Latest draft

June 19, 2009 (minor updates)

Mathematical Imaging Group

- Mathematical Imaging Group
  - 3 prof, 4 lecturers, 15-20 phd students
  - Mathematics and mathematical statistics
- Centre for Mathematical Sciences
  - Mathematics (appr. 80 employees)
  - Mathematical statistics(appr. 30 employees)
  - Numerical Analysis (appr. 10 employees)

Research

- Geometry (3D shape, camera calibration, camera motion , structure and motion, robotics)
- Medical Image Analysis (Shape variation, segmentation, tomography, decision support)
- Cognitive Vision (recognition, detection, scene interpretation, attention, segmentation, handwriting recognition)

The goal of computer vision

- To bridge the gap between pixels and “meaning”

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

The goal of computer vision

- To bridge the gap between pixels and “meaning”

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

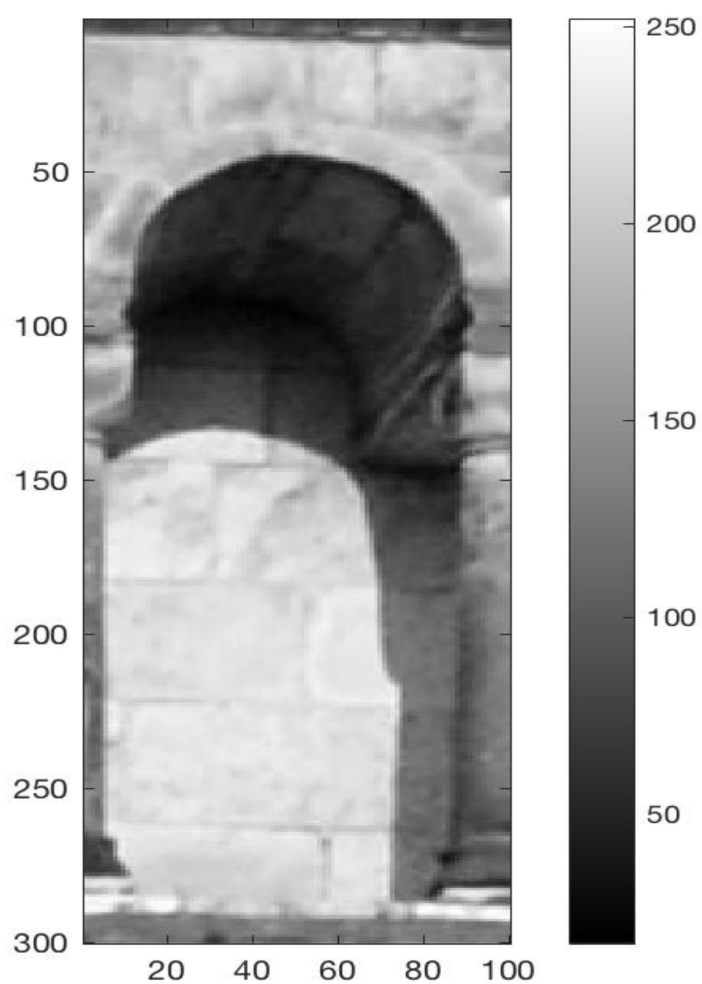
The goal of computer vision

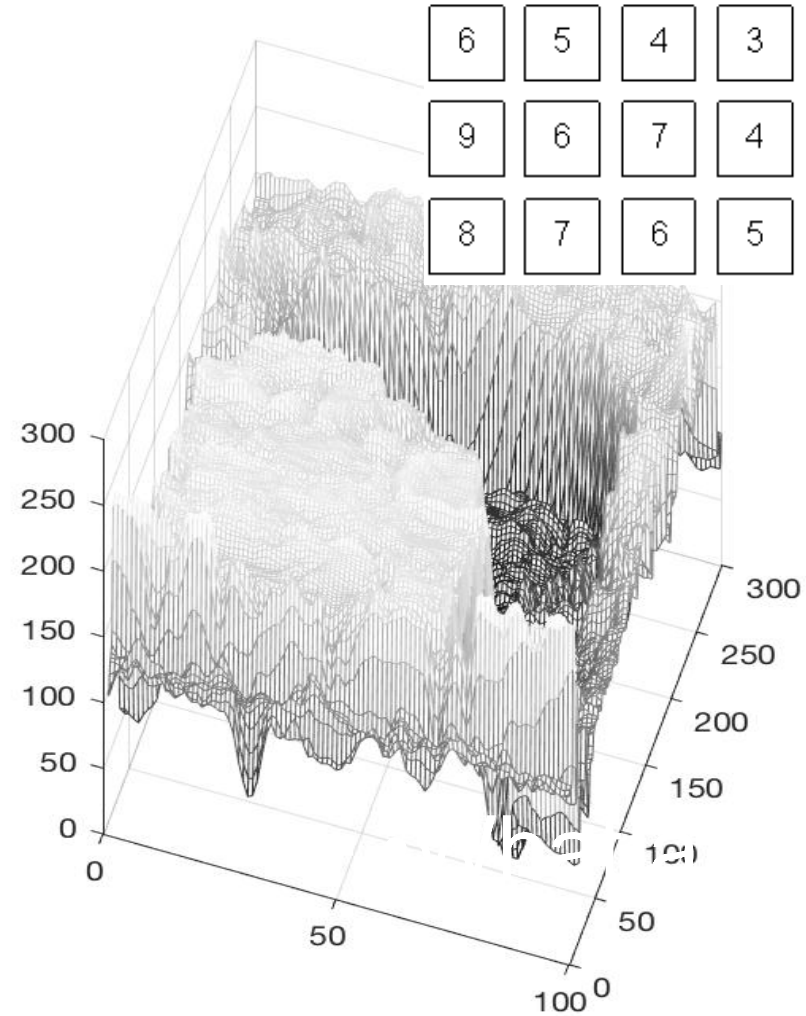
- Images are functions. Each pixel measures brightness

$$f : \mathbf{Z}^2 \rightarrow \mathbf{Z}$$

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 we see





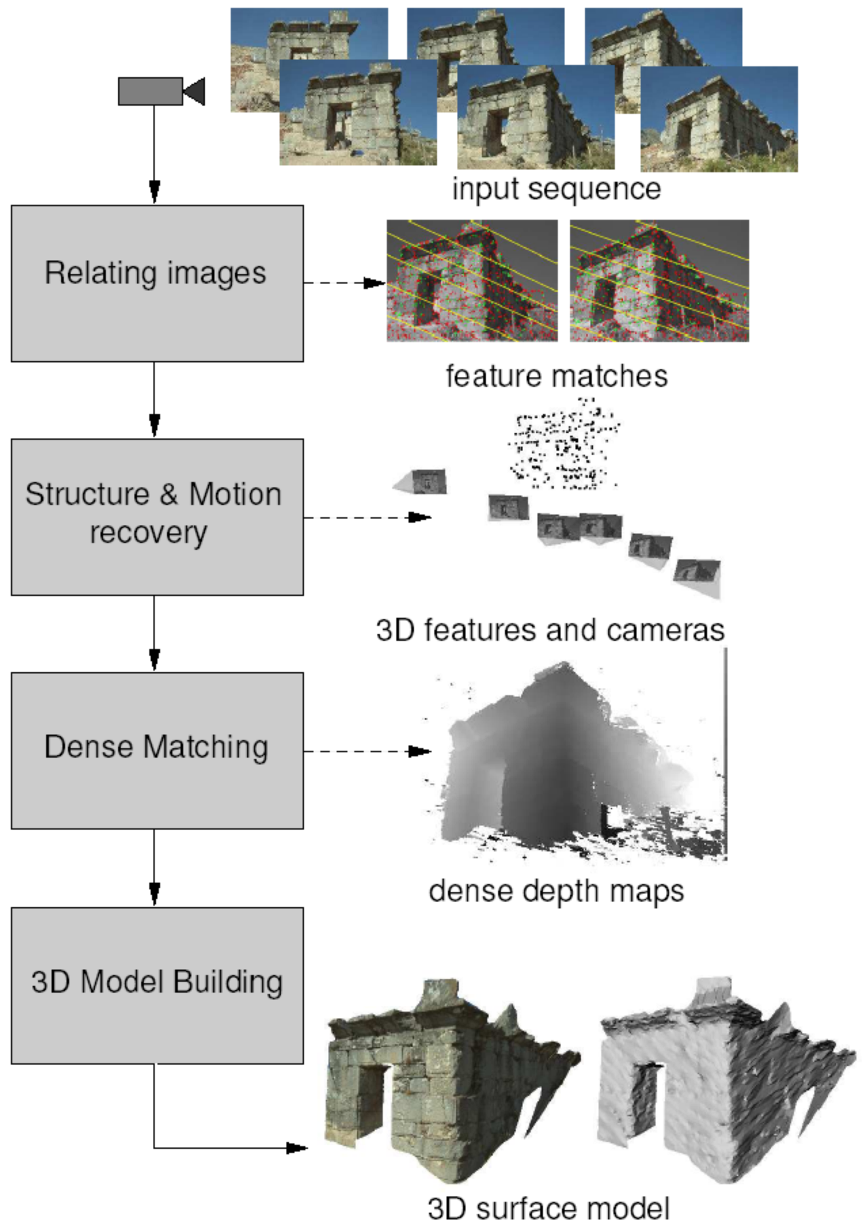
Why vision?

- As image sources multiply, so do applications
  - Relieve humans of boring, easy tasks
  - Enhance human abilities: human-computer interaction, visualization
  - Perception for robotics / autonomous agents
  - Organize and give access to visual content

What kind of information can we extract from an image?

- Metric 3D information
- Semantic information

Vision as measurement device



•Vision as a source of semantic information

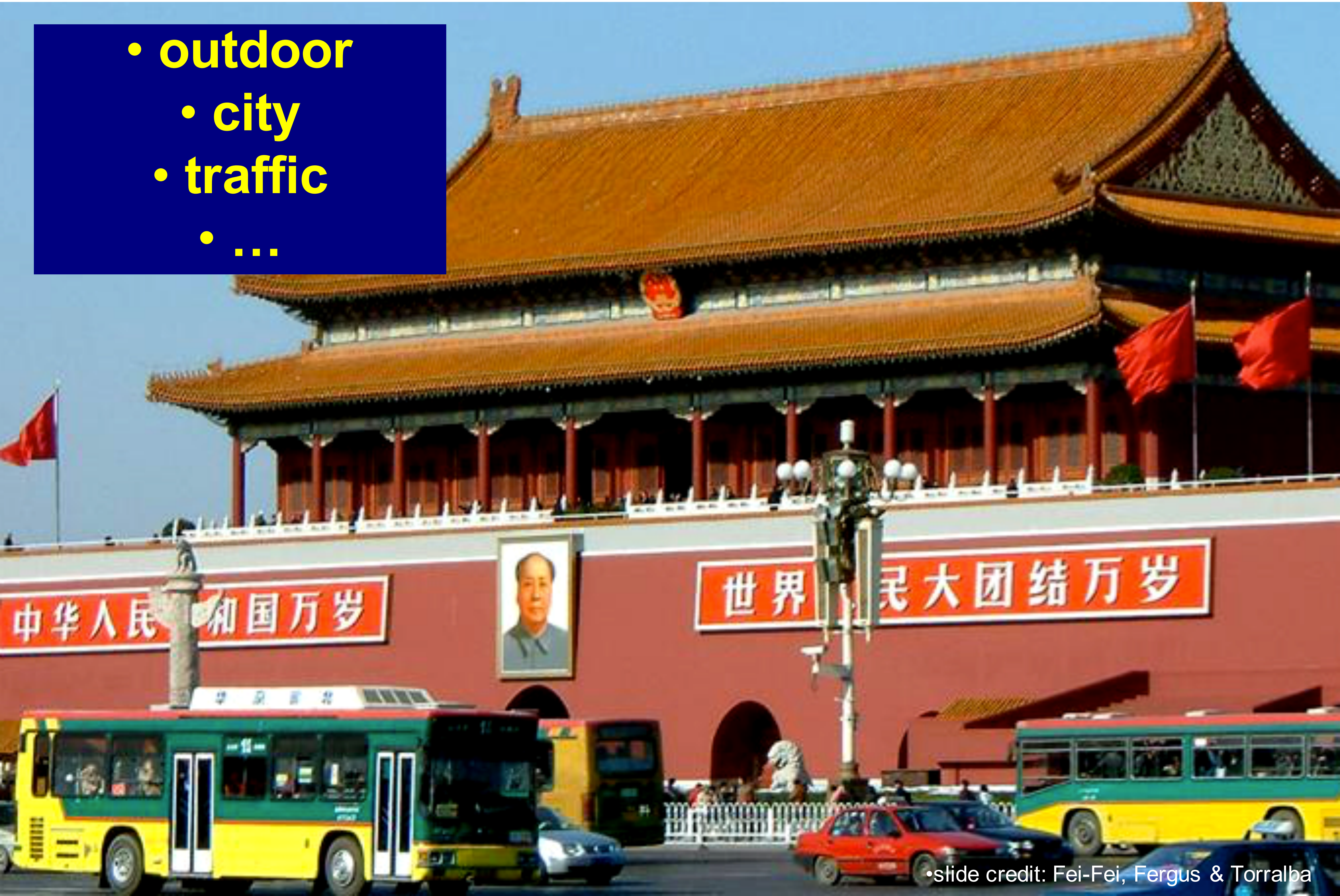
•slide credit: Fei-Fei, Fergus & Torralba

•Object categorization

•slide credit: Fei-Fei, Fergus & Torralba

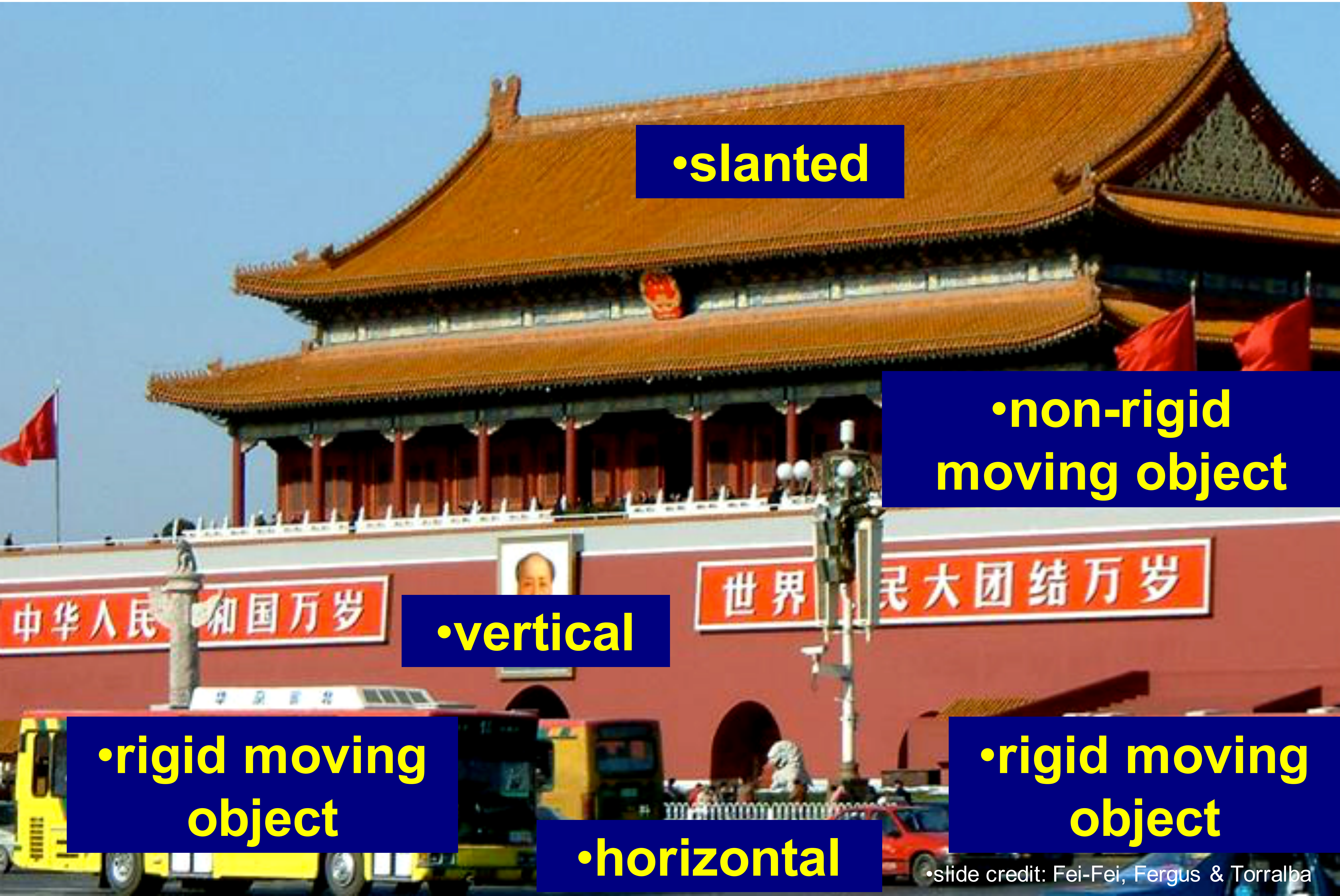
•Scene and context categorization

- outdoor
- city
- traffic
- …



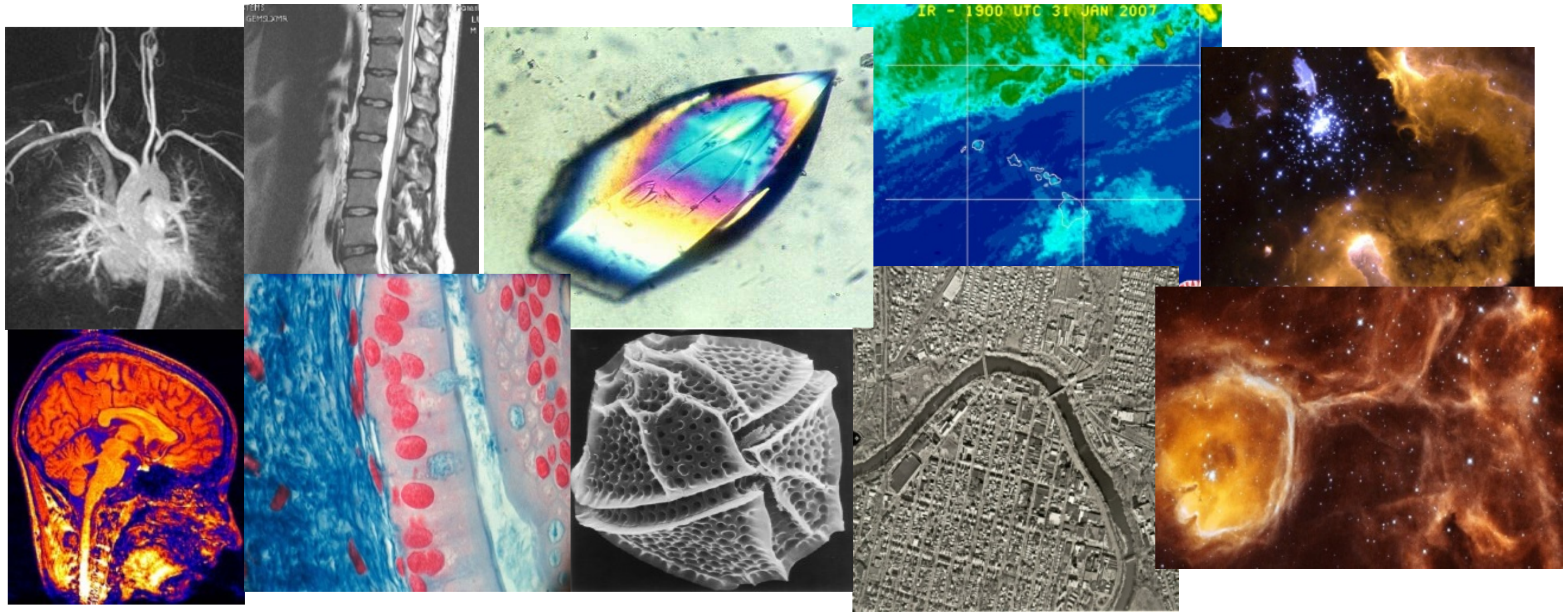
•slide credit: Fei-Fei, Fergus & Torralba

•Qualitative spatial information



- Vision is useful: Images and video are everywhere!

Google Image Search Picasa flickr webshots picsearch YouTube Broadcast Yourself



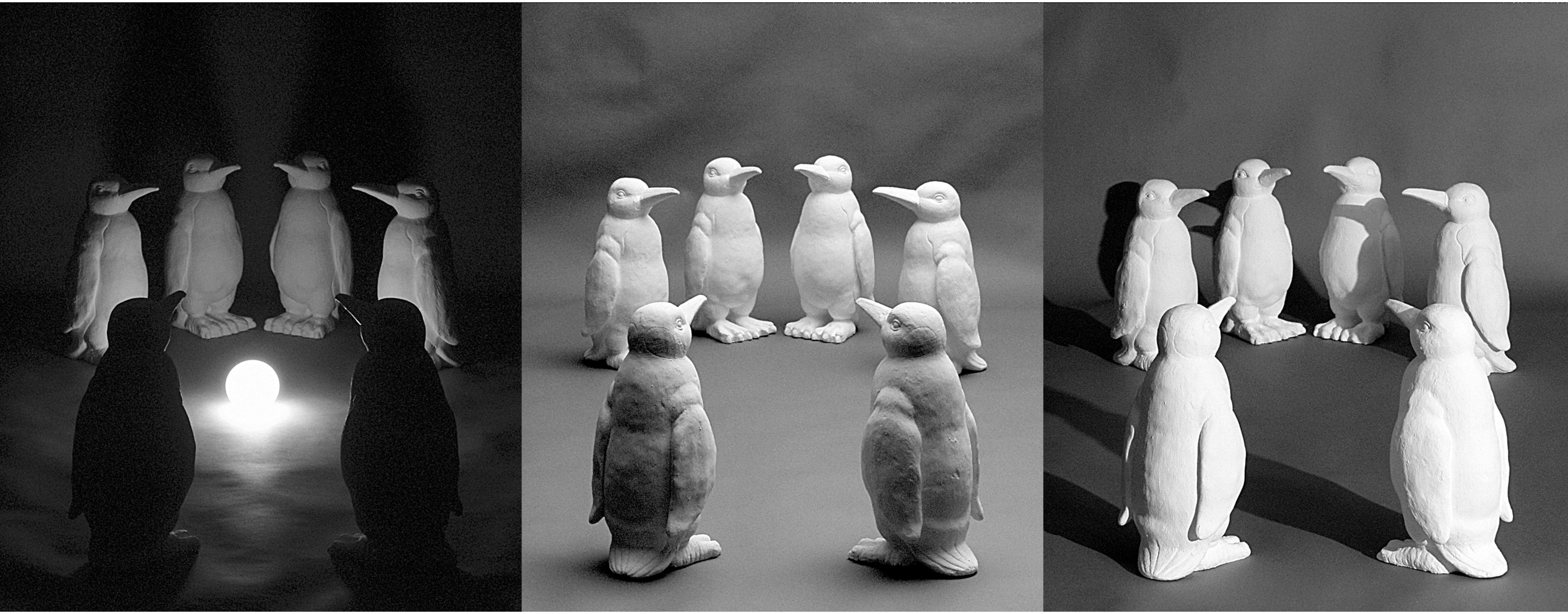
Why is computer vision difficult?

•Challenges: viewpoint variation

•Michelangelo 1475-1564

•slide credit: Fei-Fei, Fergus & Torralba

•Challenges: illumination

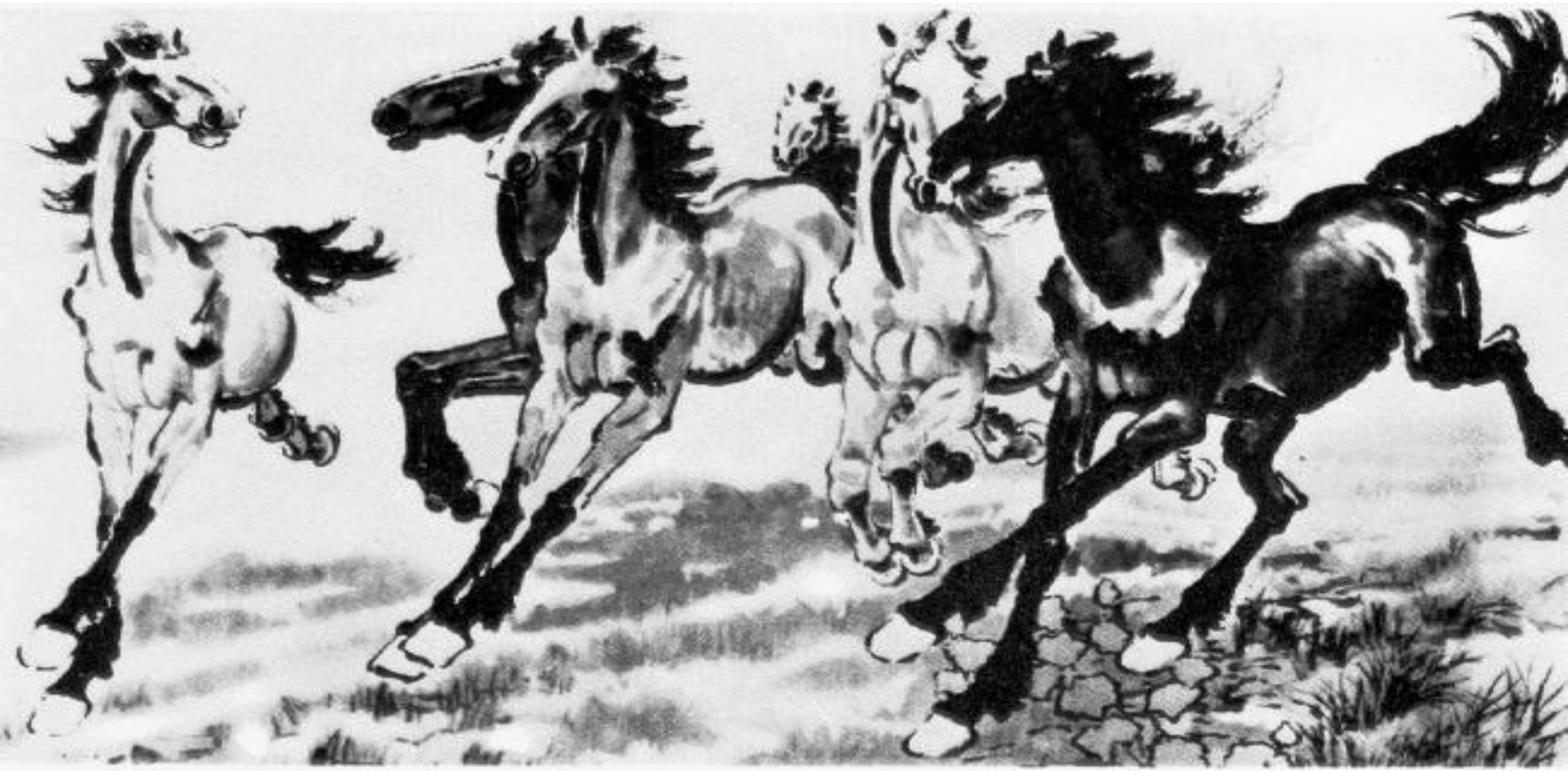


•image credit: J. Koenderink

•Challenges: scale

•slide credit: Fei-Fei, Fergus & Torralba

•Challenges: deformation



•Xu, Beihong 1943

•slide credit: Fei-Fei, Fergus & Torralba

•Challenges: occlusion

•Magritte, 1957

•slide credit: Fei-Fei, Fergus & Torralba

•Challenges: background clutter

•Challenges: Motion

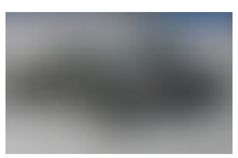
Challenges: object intra-class variation

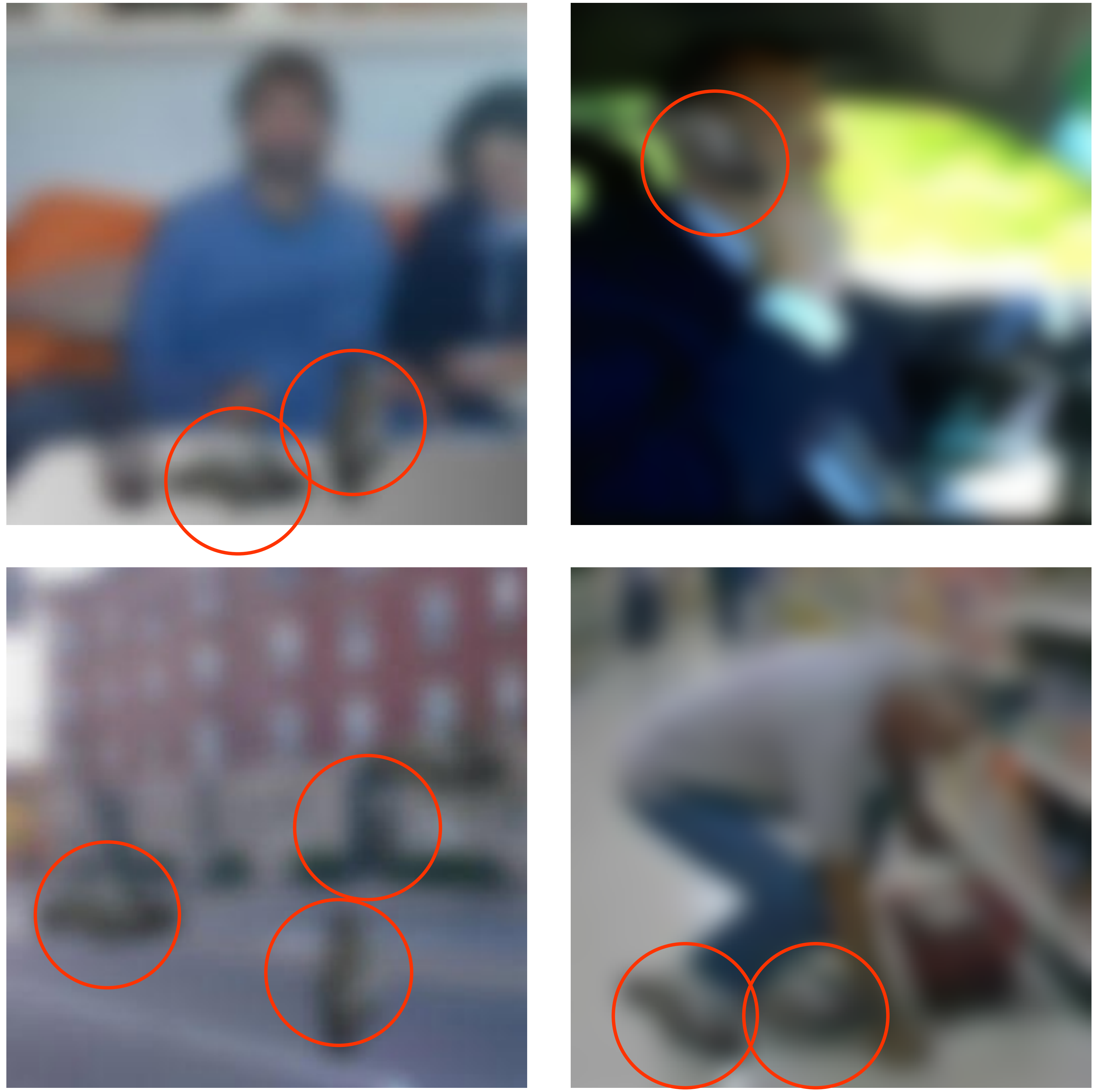





•slide credit: Fei-Fei, Fergus & Torralba

Challenges: local ambiguity



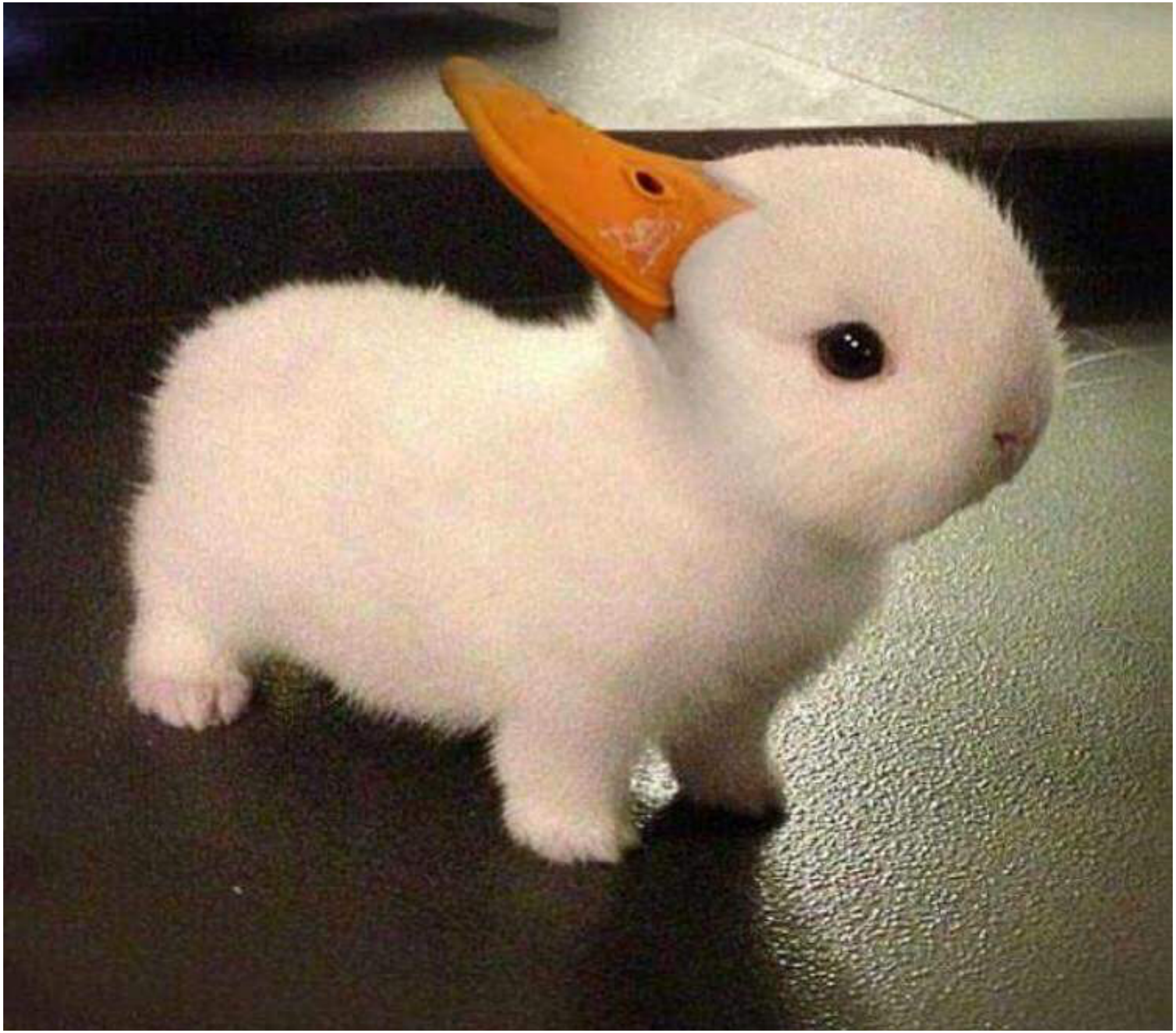


•slide credit: Fei-Fei, Fergus & Torralba

Challenges: context

Challenges: context

Challenges: context



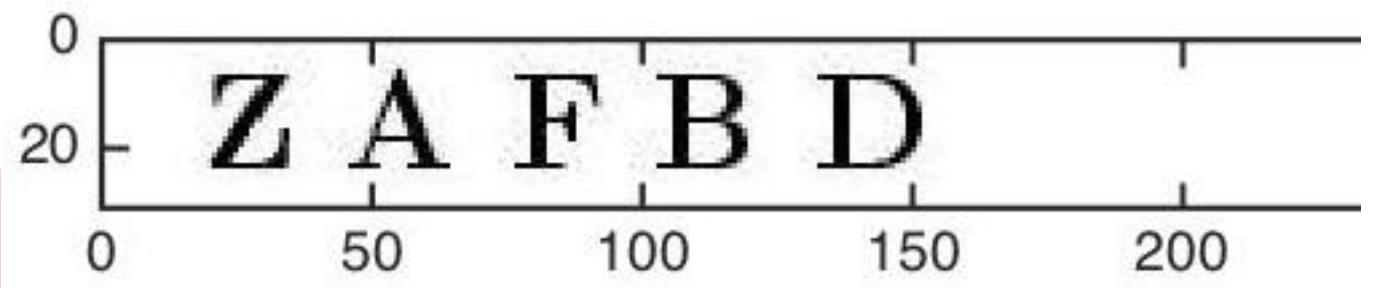
In this course

- Tools:
  - Basics of image analysis
  - Linear Algebra, Linear System Theory
  - Mathematical Statistics
  - Machine Learning
  - Segmentation
- System development
  - Based on the tools
  - Ground truth, evaluation, benchmarking

After the course

- You should be able to develop and test your own image analysis system
- You should have tools for understanding and working with big data
- You should have improved your skills in programming and modelling.

OCR project

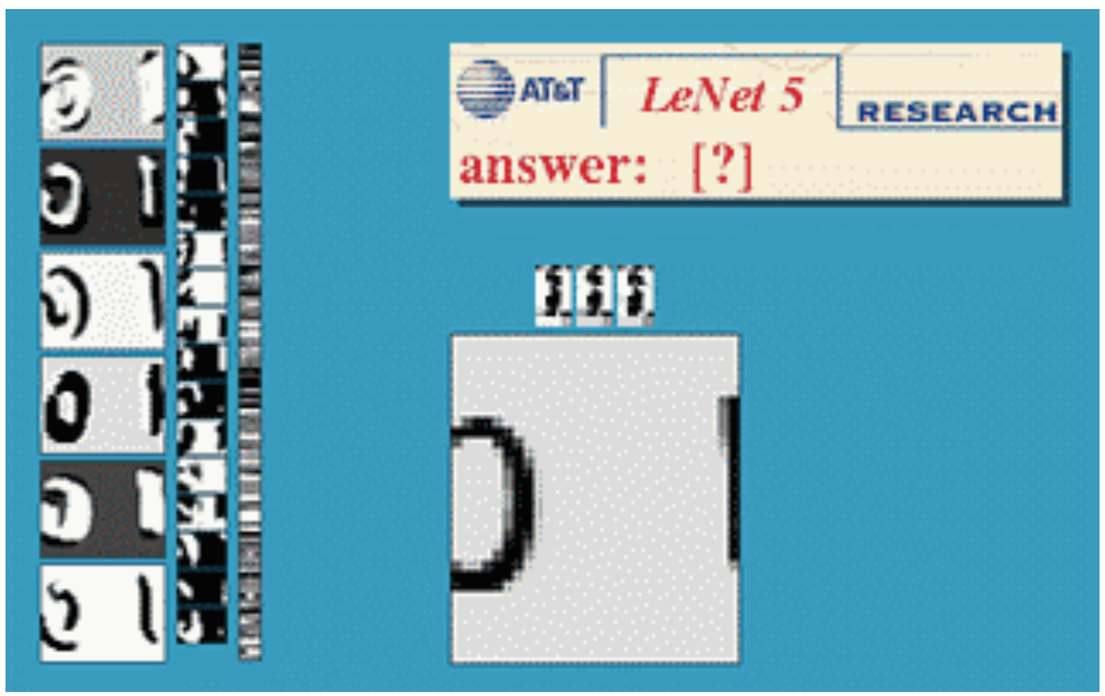


- Input: Image
- Output: Text ’ZAFBD’
- 1. Segmentation
- 2. Features
- 3. Classification

- Evaluation, benchmark

Optical character recognition (OCR)

- Technology to convert scanned docs to text
  - If you have a scanner, it probably came with OCR software



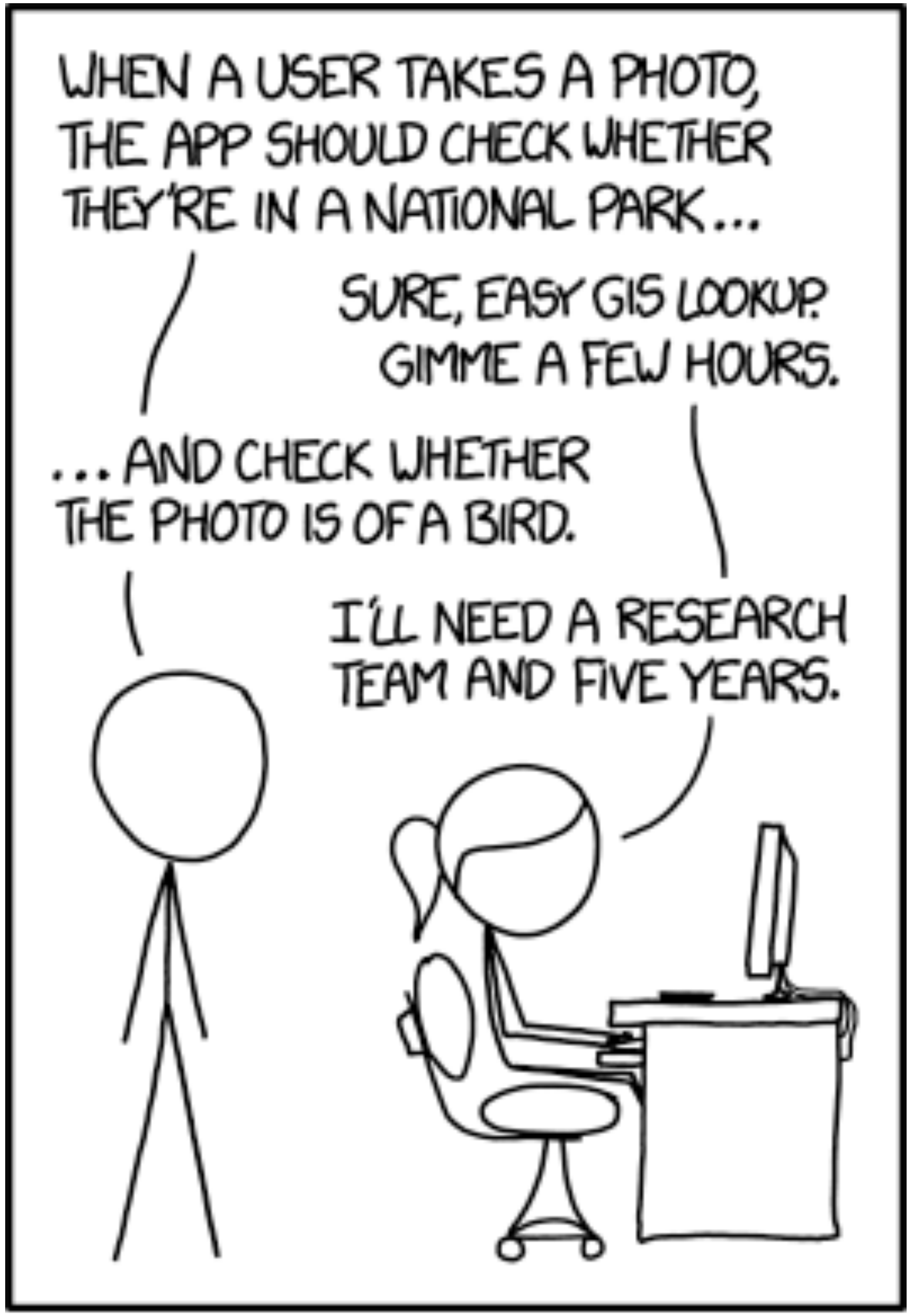
•Digit recognition, AT&T labs

•http://www.research.att.com/~yann/

•License plate readers

•http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Increasing difficulty



IN CS, IT CAN BE HARD TO EXPLAIN THE DIFFERENCE BETWEEN THE EASY AND THE VIRTUALLY IMPOSSIBLE.

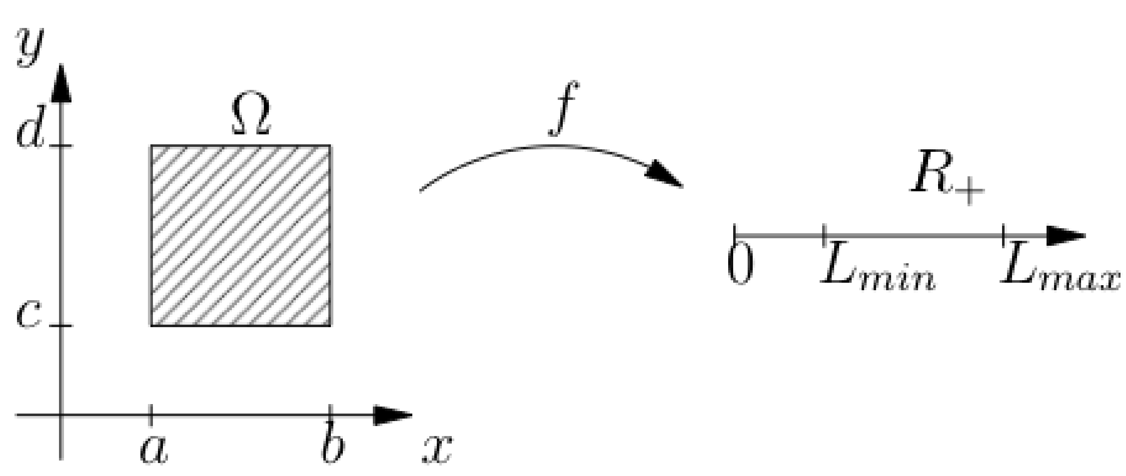
Continuous Model

An image can be seen as a function

$$f : \Omega \mapsto \mathbb{R}_+ \quad ,$$

where $\Omega = \{\,(x, y) \mid a \le x \le b,\, c \le y \le d\} \subseteq \mathbb{R}^2$ and $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \ge 0\}$. $f(x, y)$ = intensity at point (x, y) = gray-level
(f does not have to be continuous)

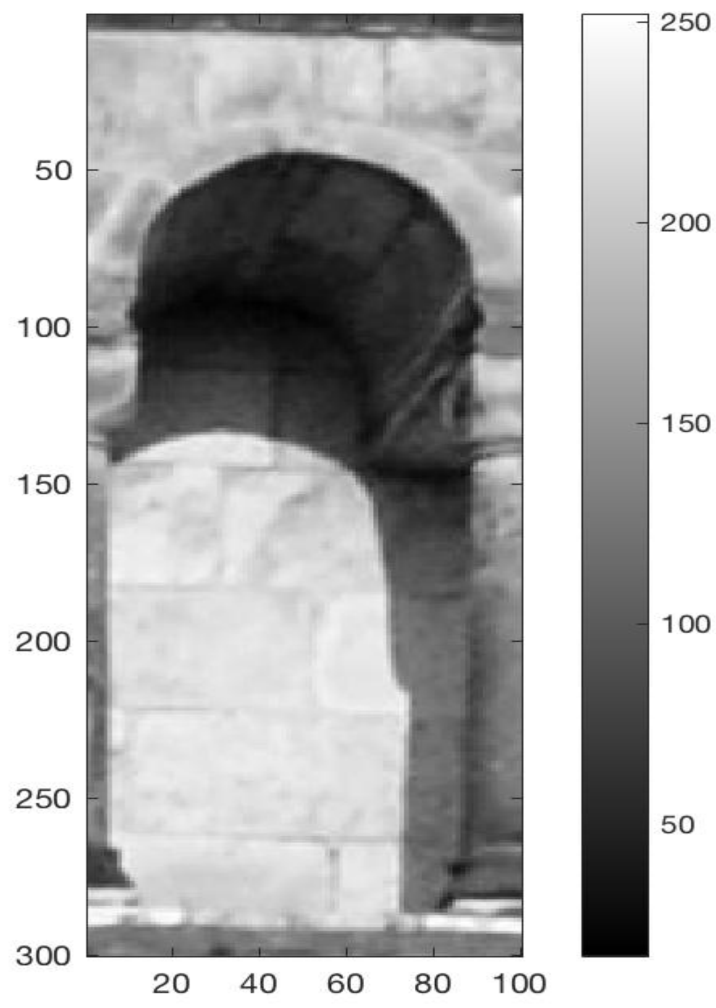
$0 \le L_{min} \le f \le L_{max} \le \infty$
$[L_{min}, L_{max}]$ = gray-scale

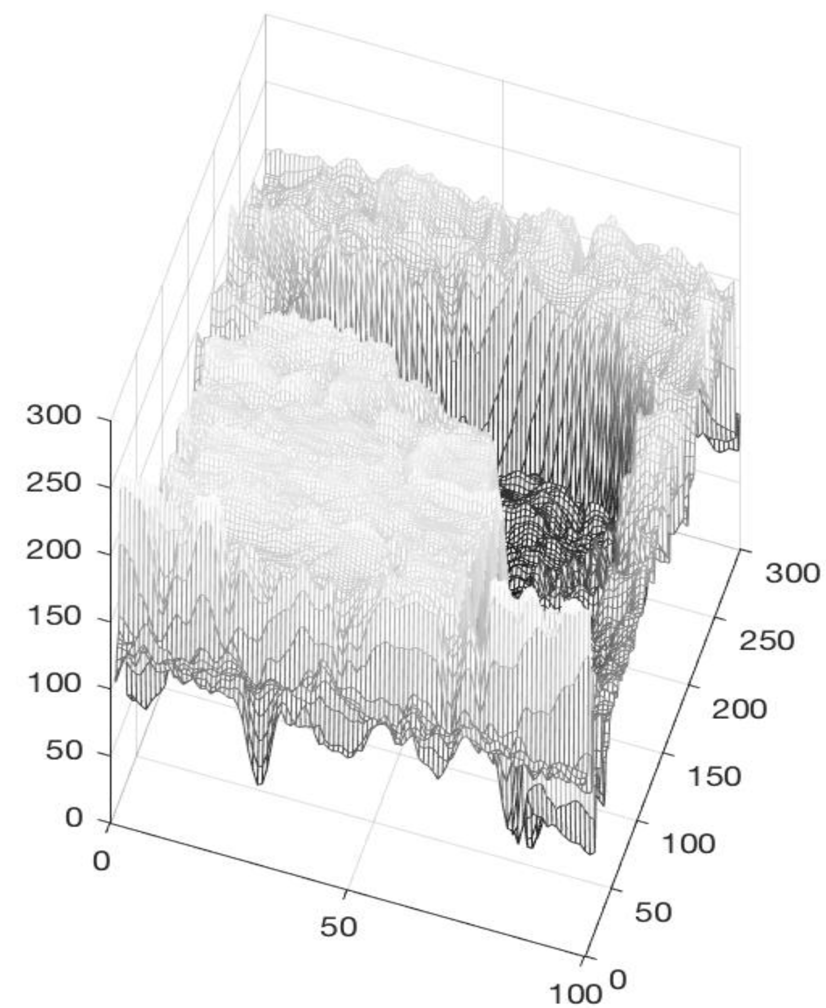


Continuous Model

An image can be seen as a function

$$f : \Omega \mapsto \mathbb{R}_+$$



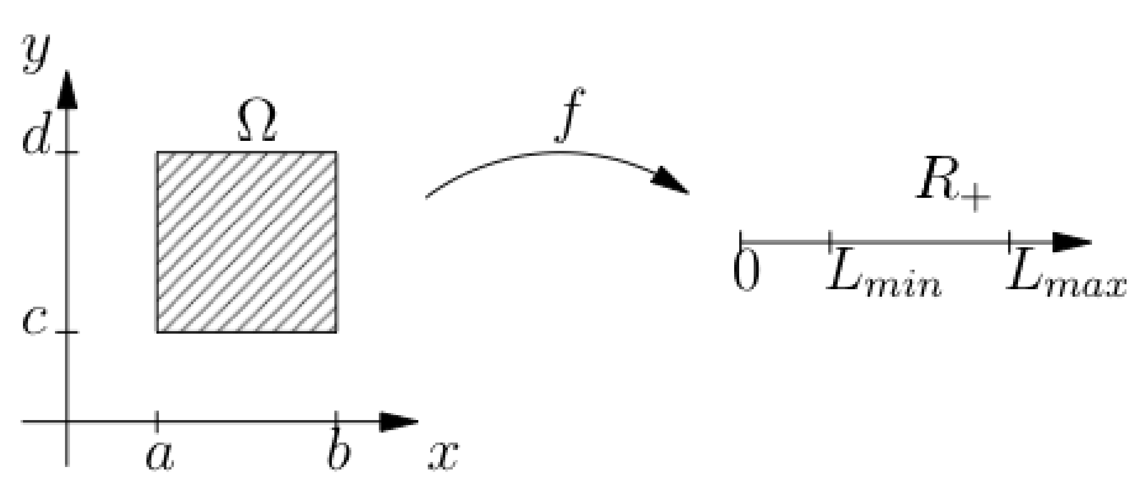


Discrete Image Model

- Discretize x,y -> sampling M rows, N columns
- Discretize f -> quantization
  - (often in 2^m levels)
  - Color depth
  - ”8 bit grayscale”, 2^8 = 256 levels, 0-255

$$f : \Omega \mapsto \mathbf{Z} \qquad \Omega \subset \mathbf{Z}^2$$

- Decreasing M and N
  - Chess patterns
- Decreasing m
  - False contours



Sampling, decreasing M and N

1

Discretize X and Y -> Sampling M rows, N colomns

Sampling, decreasing M and N

2

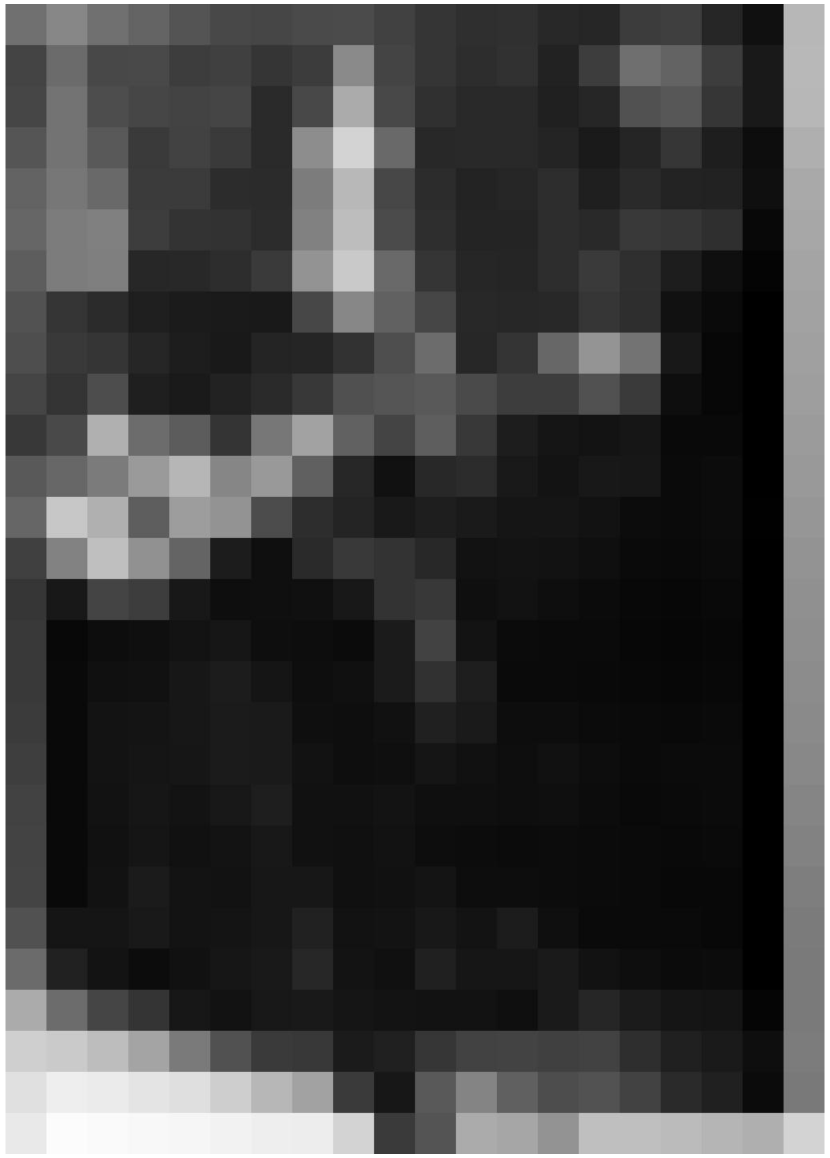
Sampling, decreasing M and N

3

Sampling, decreasing M and N

4

Sampling, decreasing M and N



$\boxed{5}$

Chess patterns

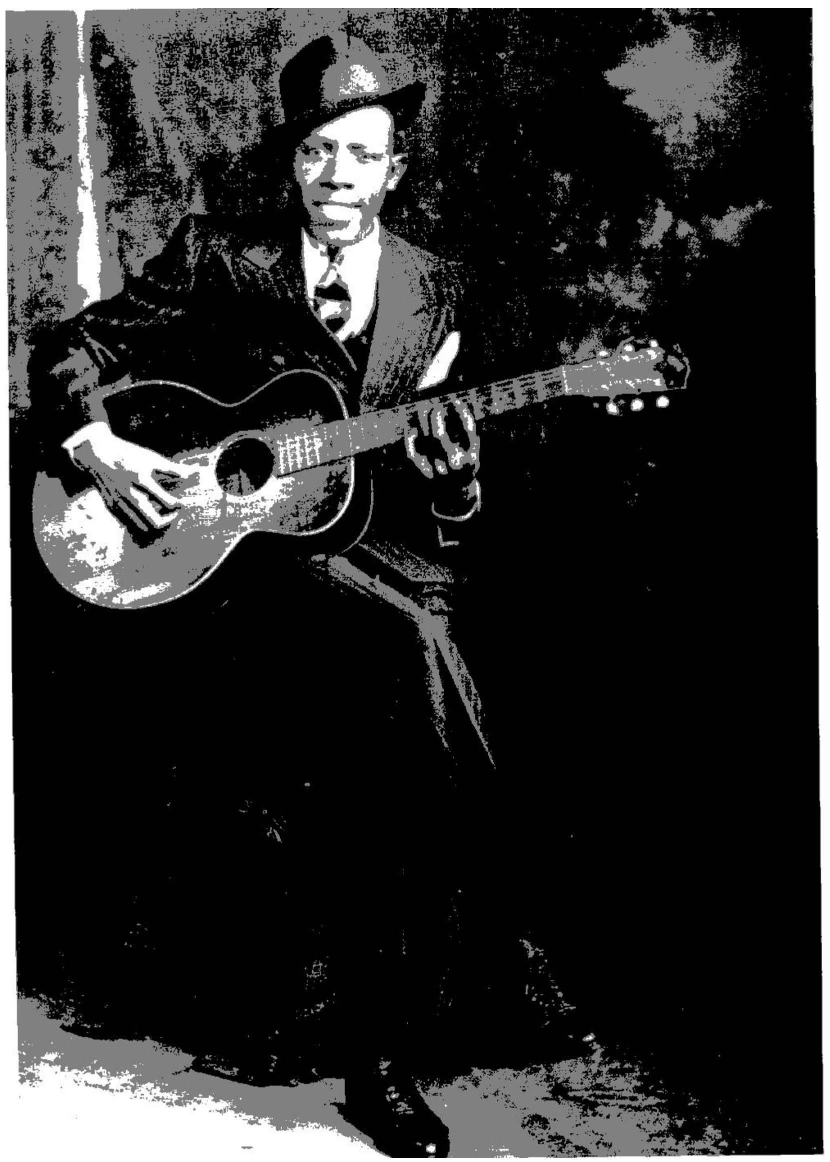
Sampling, decreasing M and N

Quantization, decreasing m

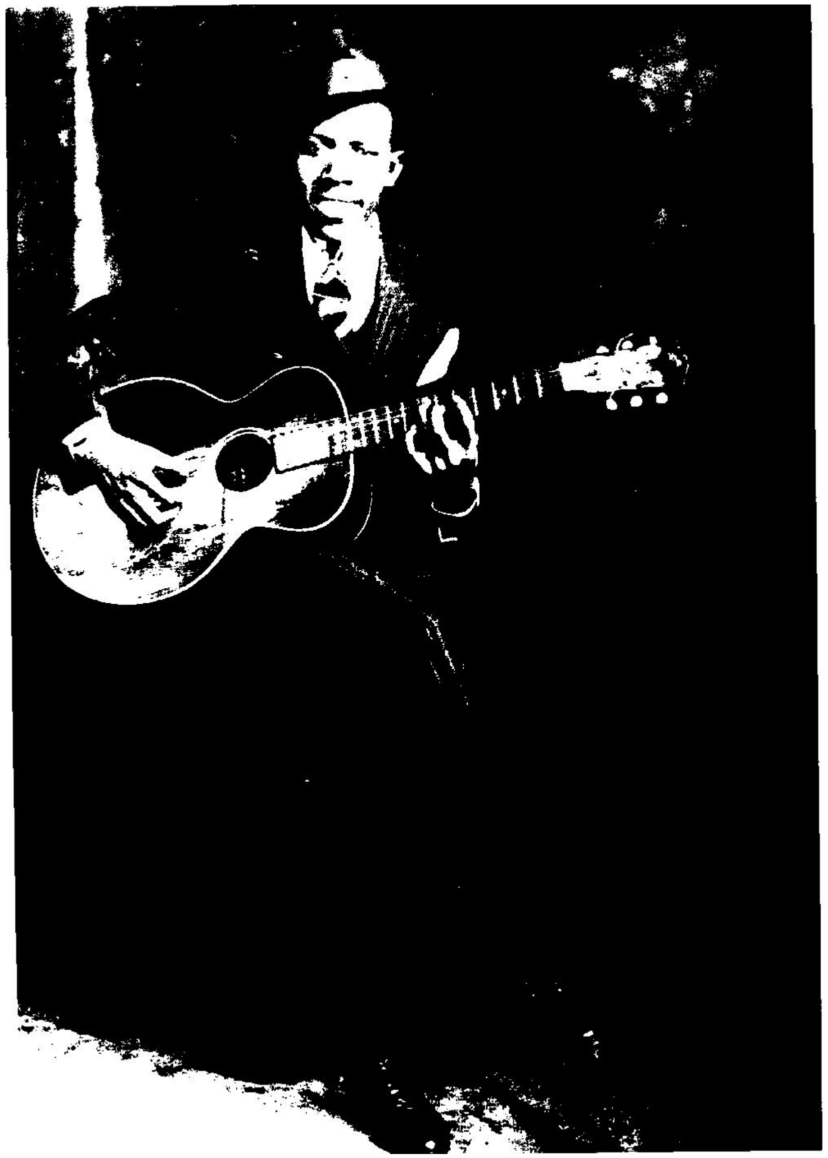
Discretize f -> Quantization in 2^m levels

Quantization, decreasing m

Quantization, decreasing m



Quantization, decreasing m



False contours

Interpolation

- Discrete image f $f : \mathbb{Z}^2 \to \mathbb{R}$
- Continuous image F $F : \mathbb{R}^2 \to \mathbb{R}$
- Going from F to f (sampling)

$$f(i,j) = D(F)(i,j) = F(i,j)$$

- Going from f to F (interpolation)

$$F_h(x,y) = I_h(f)(x,y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h(x-i, y-j) f(i,j)$$

Interpolation

- Discrete image f $f : \mathbb{Z}^2 \rightarrow \mathbb{R}$
- Continuous image F $F : \mathbb{R}^2 \rightarrow \mathbb{R}$
- Going from F to f (sampling)

$$f(i,j) = D(F)(i,j) = F(i,j)$$

- Going from f to F (interpolation)

$$F_h(x,y) = I_h(f)(x,y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h(x-i, y-j) f(i,j)$$

Interpretation:

Place a hump h at each pixel
Scale the hump by f(i,j)
Add together

Interpolation

- Discrete image f $f : \mathbb{Z}^2 \to \mathbb{R}$
- Continuous image F $F : \mathbb{R}^2 \to \mathbb{R}$
- Going from F to f (sampling)

$$f(i,j) = D(F)(i,j) = F(i,j)$$

- Going from f to F (interpolation)

$$F_h(x,y) = I_h(f)(x,y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h(x-i, y-j) f(i,j)$$

Different choices of h (different humps)
-> different types of interpolation

Interpolation – what is h?

- How can you find h from method?????

$$F_h(x,y) = I_h(f)(x,y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h(x-i, y-j) f(i,j)$$

$$f : \mathbb{Z}^2 \to \mathbb{R}$$

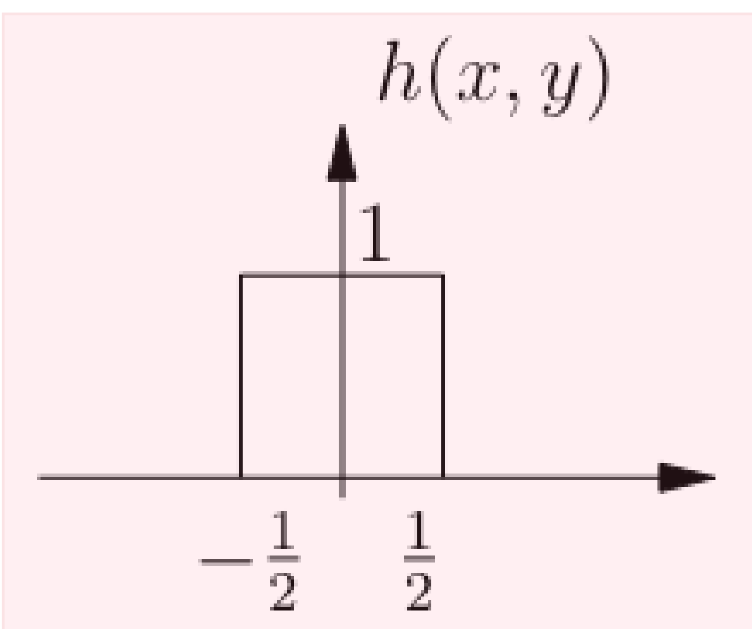
$$F : \mathbb{R}^2 \to \mathbb{R}$$

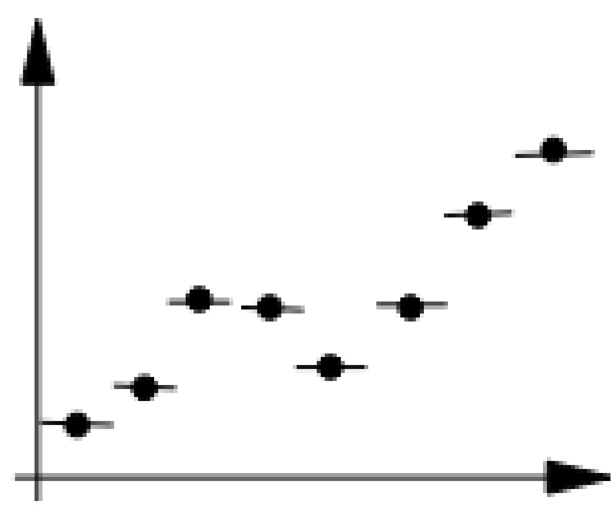
Interpolation

- Going from f to F (interpolation)

$$F_h(x,y) = I_h(f)(x,y) = \sum^{\infty} \sum^{\infty} h(x-i, y-j) f(i,j)$$

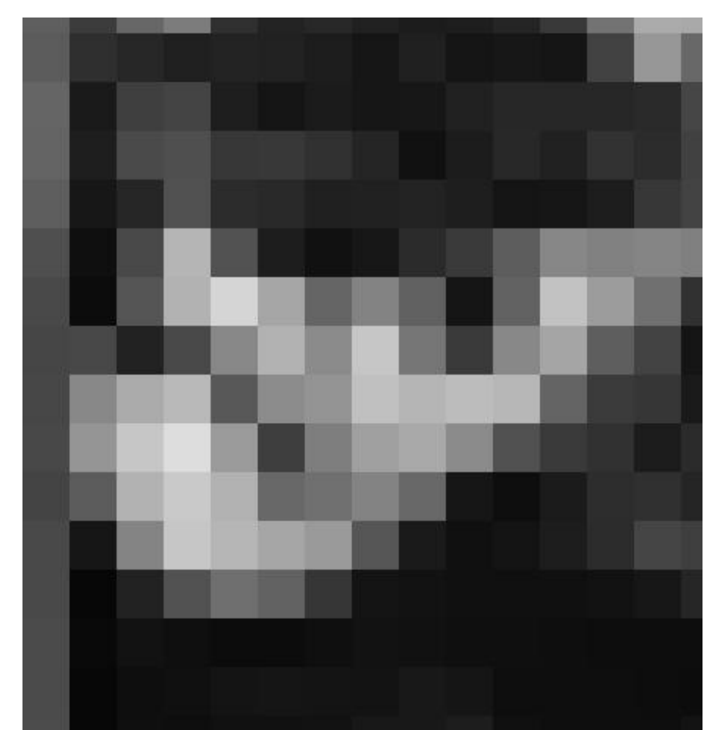
- Example 1 – Pixel Replication





In 1D

In 2D



$$f : \mathbb{Z}^2 \rightarrow \mathbb{R}$$

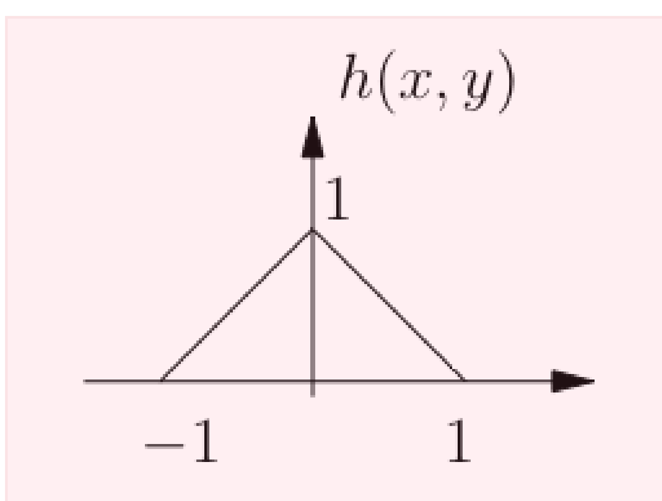
$$F : \mathbb{R}^2 \rightarrow \mathbb{R}$$

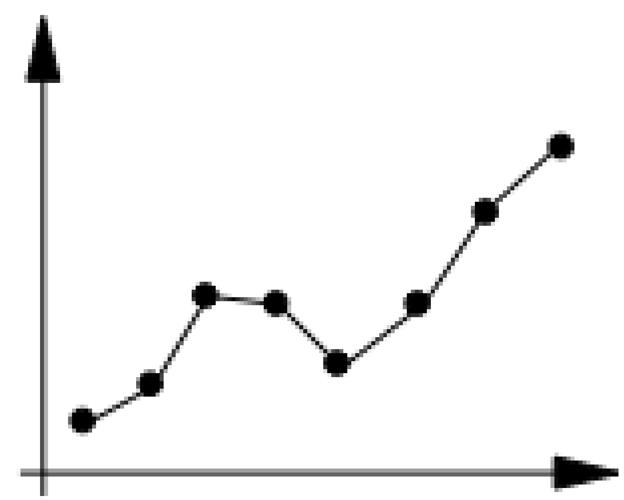
Interpolation

- Going from f to F (interpolation)

$$F_h(x,y) = I_h(f)(x,y) = \sum^{\infty} \sum^{\infty} h(x-i, y-j) f(i,j)$$

- Example 2 – Linear interpolation
- (In two dimensions the corresponding function is bilinear)





$$f : \mathbb{Z}^2 \to \mathbb{R}$$

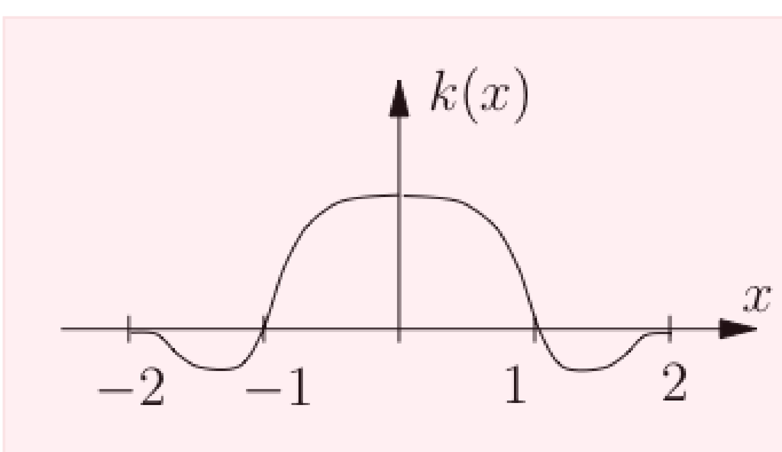
$$F : \mathbb{R}^2 \to \mathbb{R}$$

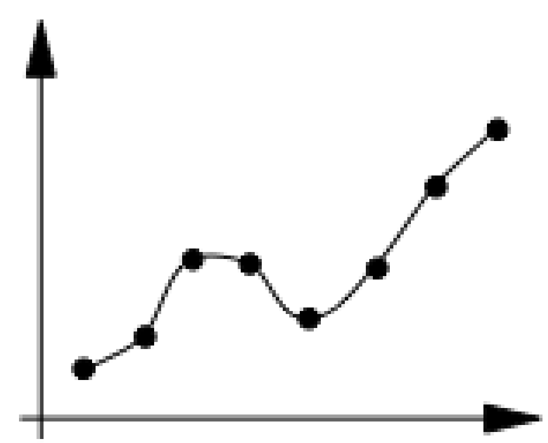
Interpolation

- Going from f to F (interpolation)

$$F_h(x,y) = I_h(f)(x,y) = \sum^{\infty} \sum^{\infty} h(x-i, y-j) f(i,j)$$

- Example 3 – Cubic interpolation
- (In two dimensions the corresponding function is bicubic)





$$f : \mathbb{Z}^2 \to \mathbb{R}$$

$$F : \mathbb{R}^2 \to \mathbb{R}$$

Interpolation

- Going from f to F (interpolation)

$$F_h(x,y) = I_h(f)(x,y) = \sum^{\infty} \sum^{\infty} h(x-i, y-j) f(i,j)$$

- Example 4 – Ideal Interpolation

$$\operatorname{sinc}(x) = \begin{cases} \frac{\sin \pi x}{\pi x}, & x \neq 0 \\ 1, & x = 0 \end{cases}$$

$$F(x,y) = I(f)(x,y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \operatorname{sinc}(x-i) \operatorname{sinc}(y-j) f(i,j).$$

Interpolation

- Discrete image f $f : \mathbb{Z}^2 \rightarrow \mathbb{R}$
- Continuous image F $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$
- If the function F is square integrable, i.e.

$$\int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} |F(x,y)|^2 dx dy$$

- Is bounded.
- If also the fourier transform is zero outside $[-\pi, \pi] \times [-\pi, \pi]$.
- Then

$$I(D(F)) = F.$$

Digital Geometry

Let $\mathbb{Z}$ be the set of integers $0, \pm 1, \pm 2, \ldots$.

Grid: $\mathbb{Z}^2$,
$$\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array}$$

Grid point: (x, y)

Definition

4-neigbourhood to (x, y):

$$N_4(x, y) = \begin{pmatrix} \cdot & \times & \cdot \\ \times & (x, y) & \times \\ \cdot & \times & \cdot \end{pmatrix} \quad .$$

Digital Geometry

Definition

p and q are **4-neighbours** if $p \in N_4(q)$. ■

Definition

A 4-**path** from p to q is a sequence

$$p = r_0, r_1, r_2, \ldots, r_n = q \ ,$$

such that r_i and r_{i+1} are 4-neighbours. ■

Definition

Let $S \subseteq \mathbb{Z}^2$. S is 4-**connected** if for every $p, q \in S$ there is a 4-path in S from p to q. ■

There are efficient algorithms for dividing sets $M \subseteq \mathbb{Z}^2$ in connected components. (For example, see MATLAB's bwlabel).

Digital Geometry

Similar definitions with other neighbourhood structures

Definition

***D*-neighbourhood** to (x, y):

$$N_D(x,y) = \begin{pmatrix} \times & \cdot & \times \\ \cdot & (x,y) & \cdot \\ \times & \cdot & \times \end{pmatrix} .$$

■

Definition

8-neighbourhood to (x, y):

$$N_8(x,y) = N_4(x,y) \cup N_D(x,y) = \begin{pmatrix} \times & \times & \times \\ \times & (x,y) & \times \\ \times & \times & \times \end{pmatrix} .$$

Digital Geometry (bwlabel)

```
>> bild = [1 1 0 0 0;1 0 0 0 1;0 0 0 0 1;1 1 0 1 1]

bild =

     1     1     0     0     0
     1     0     0     0     1
     0     0     0     0     1
     1     1     0     1     1

>> segmentering = bwlabel(bild)

segmentering =

     1     1     0     0     0
     1     0     0     0     3
     0     0     0     0     3
     2     2     0     3     3
```

LUND UNIVERSITY

Gray level transformations Pixelwise operations

A simple method for image enhancement

Definition

Let $f(x, y)$ be the intensity function of an image. A **gray-level transformation**, T, is a function (of one variable)

$$g(x, y) = T(f(x, y))$$
$$s = T(r) \quad ,$$

that changes from gray-level f to gray-level g. T usually fulfils

- $T(r)$ increasing in $L_{min} \leq r \leq L_{max}$,
- $0 \leq T(r) \leq L$.

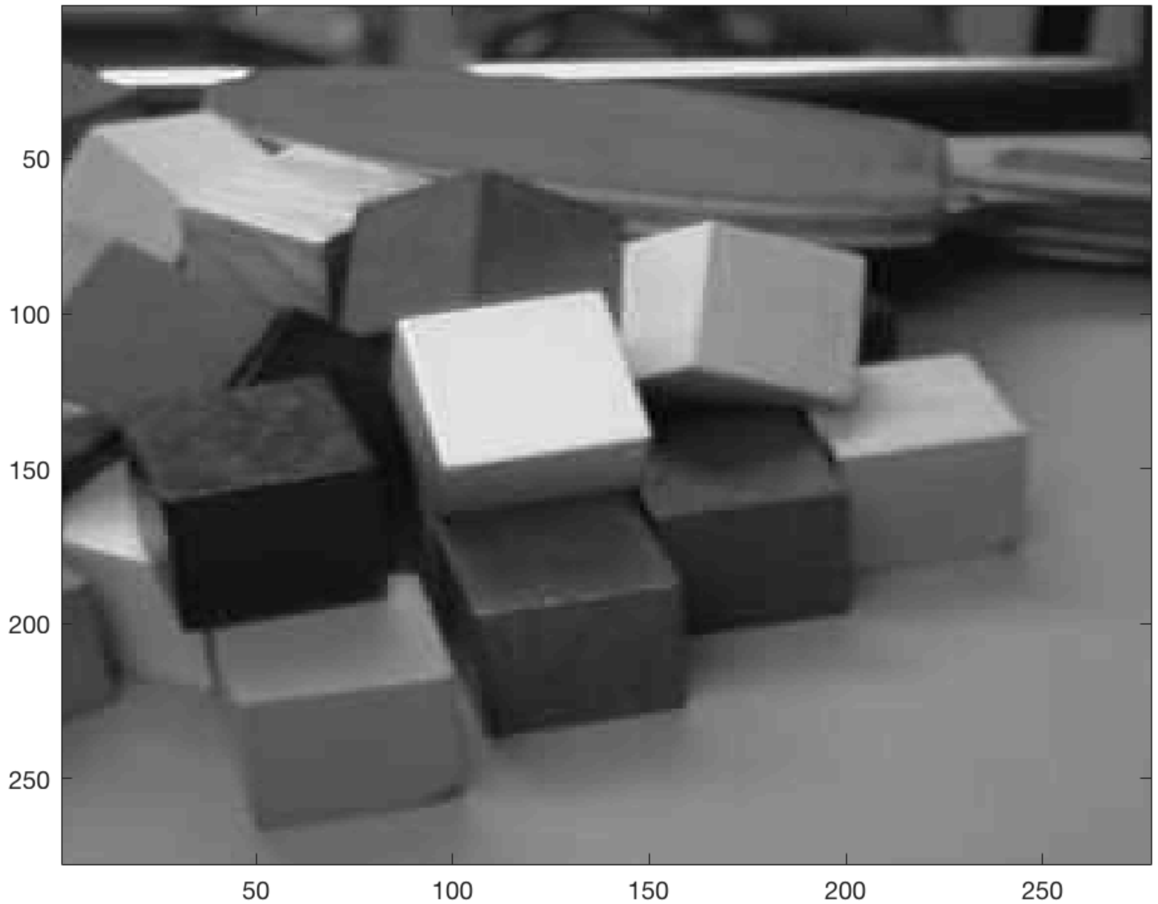
In many examples we assume that $L_{min} = 0$ och $L_{max} = L = 1$. The requirements on T being increasing can be relaxed, e.g. with inversion.

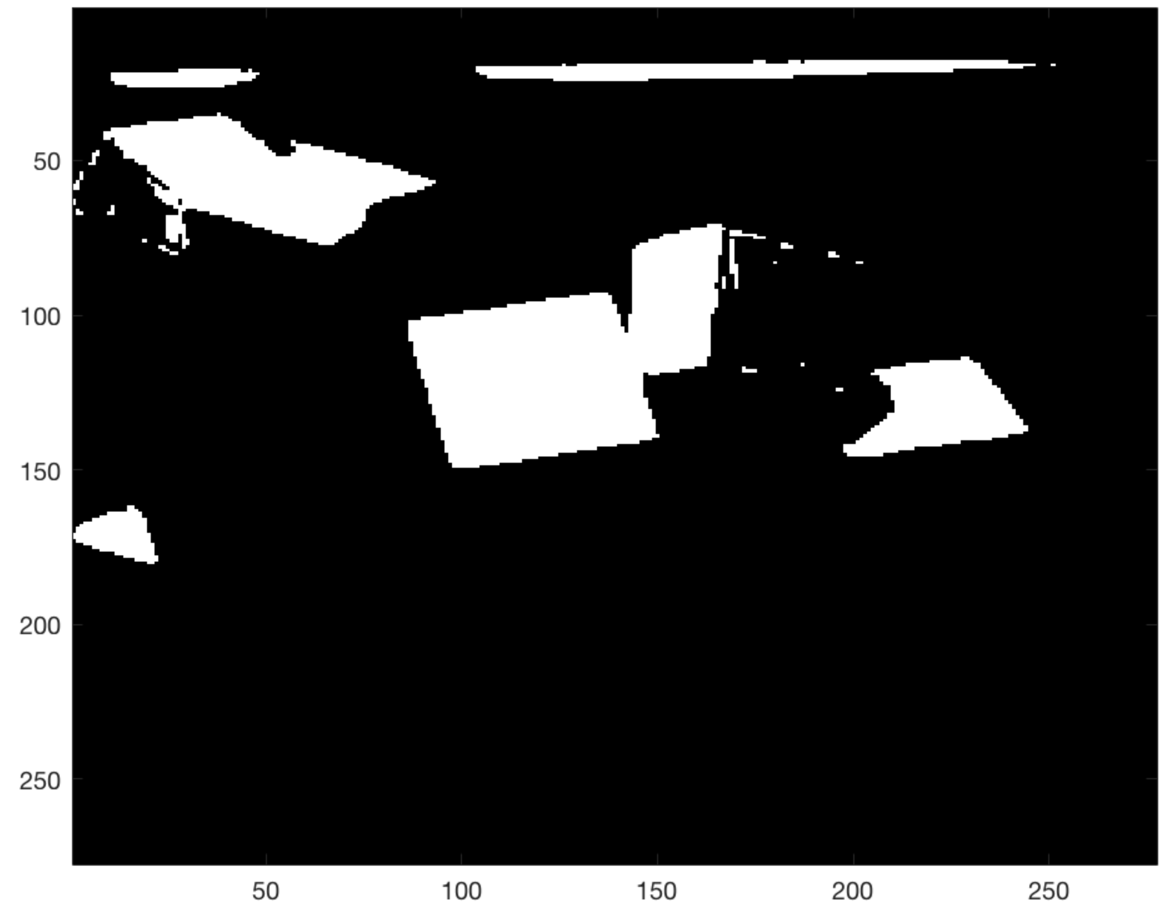
Gray level transformations
Pixelwise operations

Let

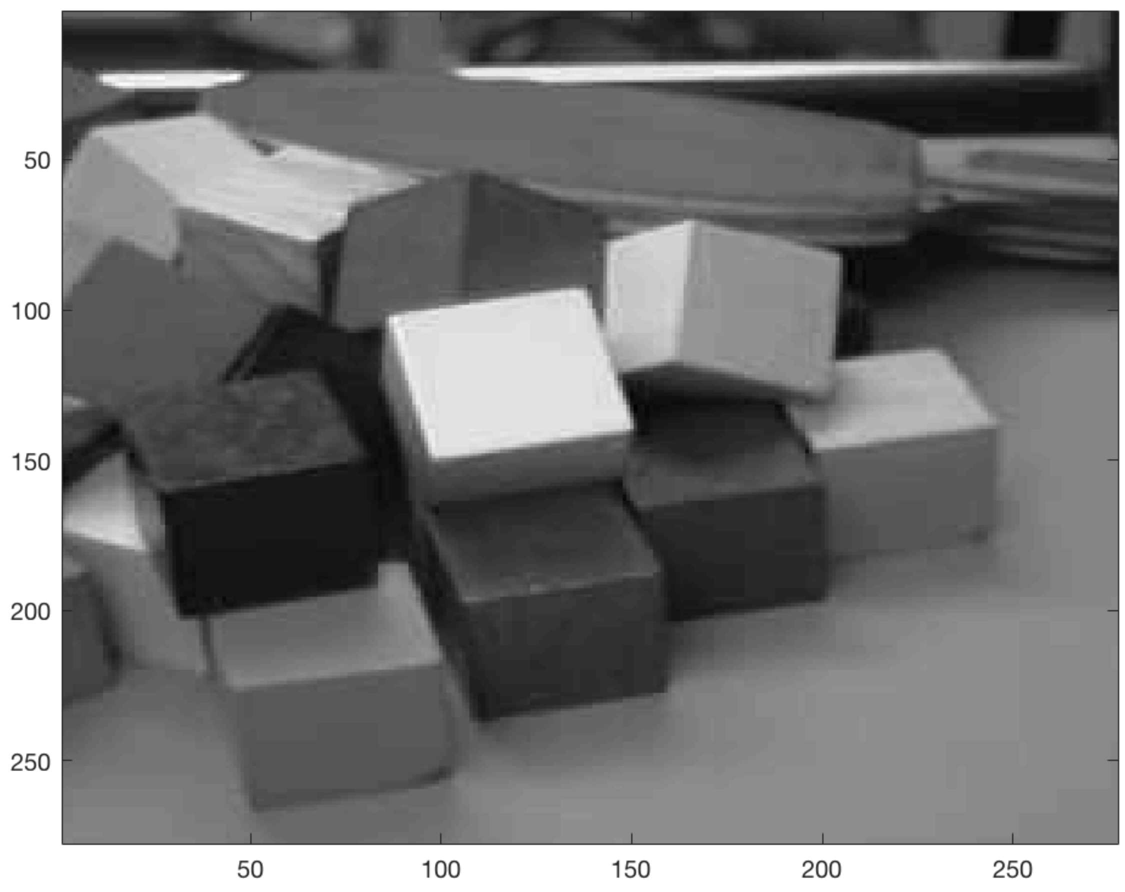
$$T(r) = \begin{cases} 0 & r \leq m \\ 1 & r > m, \end{cases}$$

for some $0 < m < 1$.

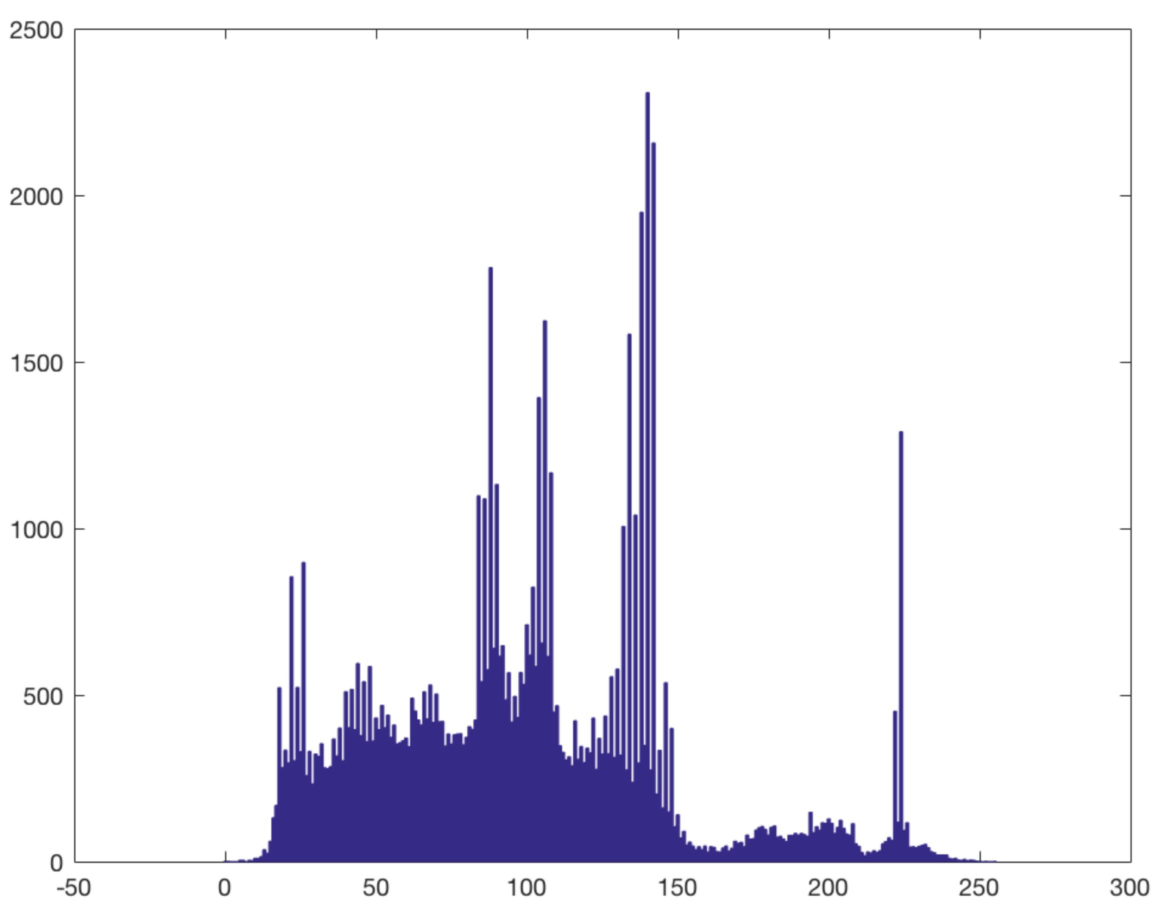




Histograms



50
100
150
200
250
50
100
150
200
250



2500
2000
1500
1000
500
0
-50
0
50
100
150
200
250
300

Histograms

- Let $s = T(r)$ be a gray level transformation
- Let p_r be the histogram before the transformation
- Let p_s be the histogram after the transformation
- Assume that T is a monotonically increasing function.
- The pixels that were darker than level r before are darker than s after.

It follows that

$$\int_0^s p_s(t)dt = \int_0^r p_r(t)dt.$$

Histograms

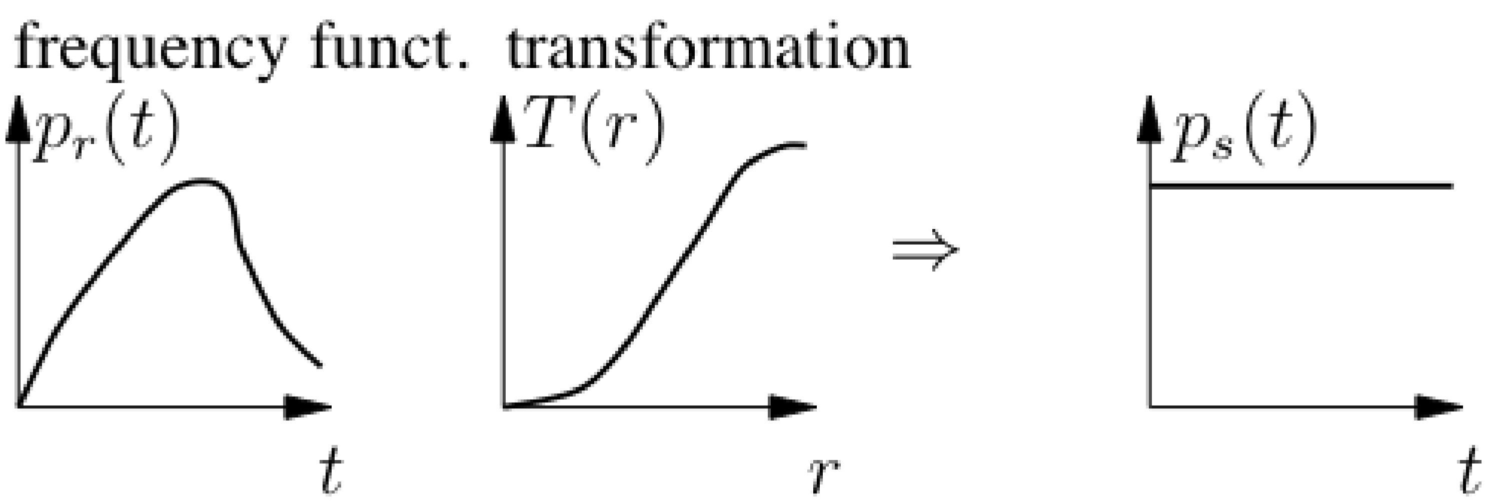
$$\int_0^s p_s(t)dt = \int_0^r p_r(t)dt.$$

Take T so that $p_s(s) = 1$ (constant).

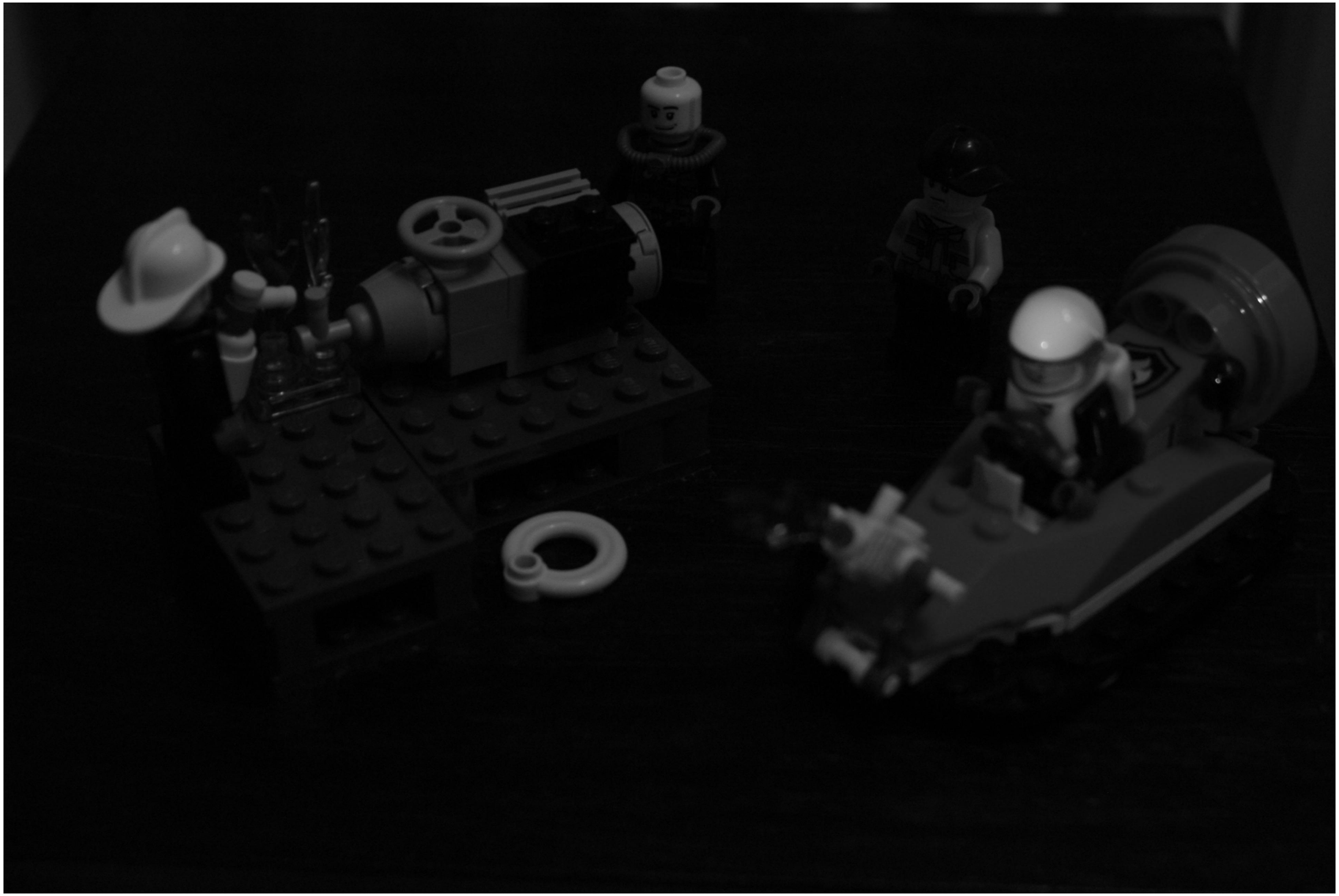
$$\int_0^r p_r(t)dt = \int_0^s 1\,dt = s \Rightarrow s = T(r) = \int_0^r p_r(t)dt$$

or

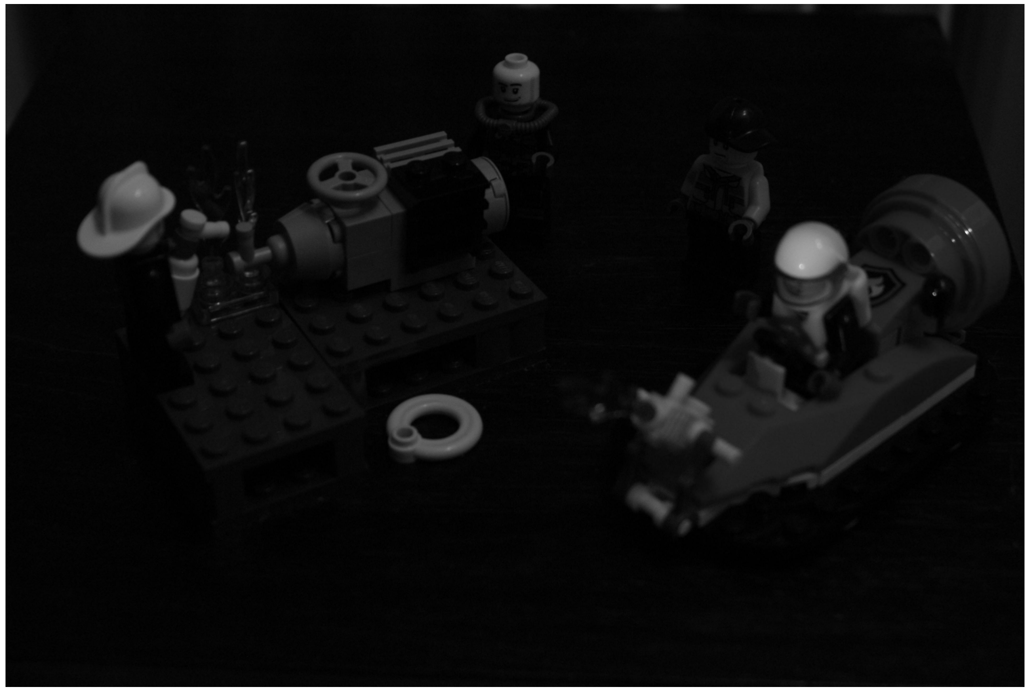
$$\frac{ds}{dr} = p_r(r)$$

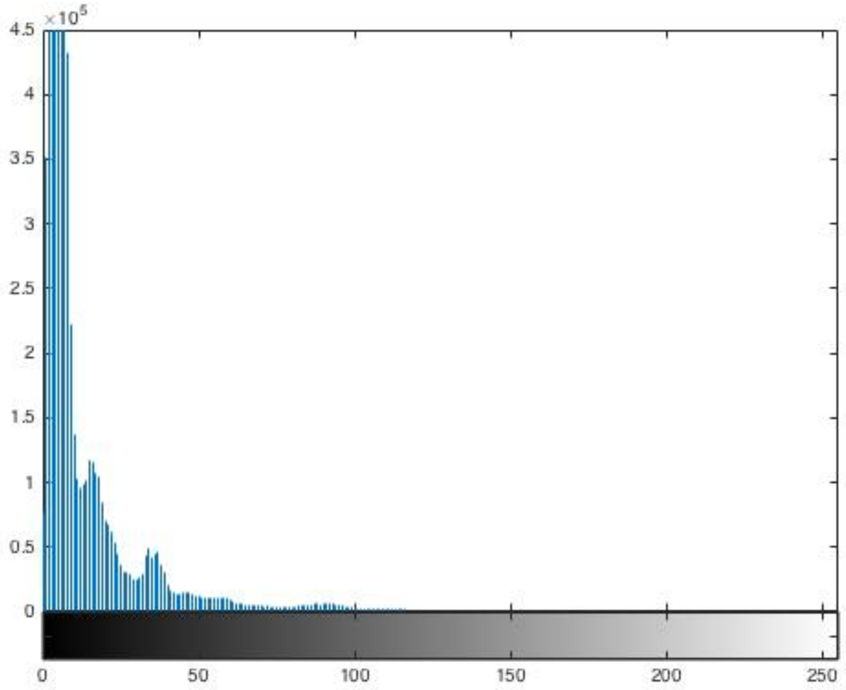


Histogram equalization

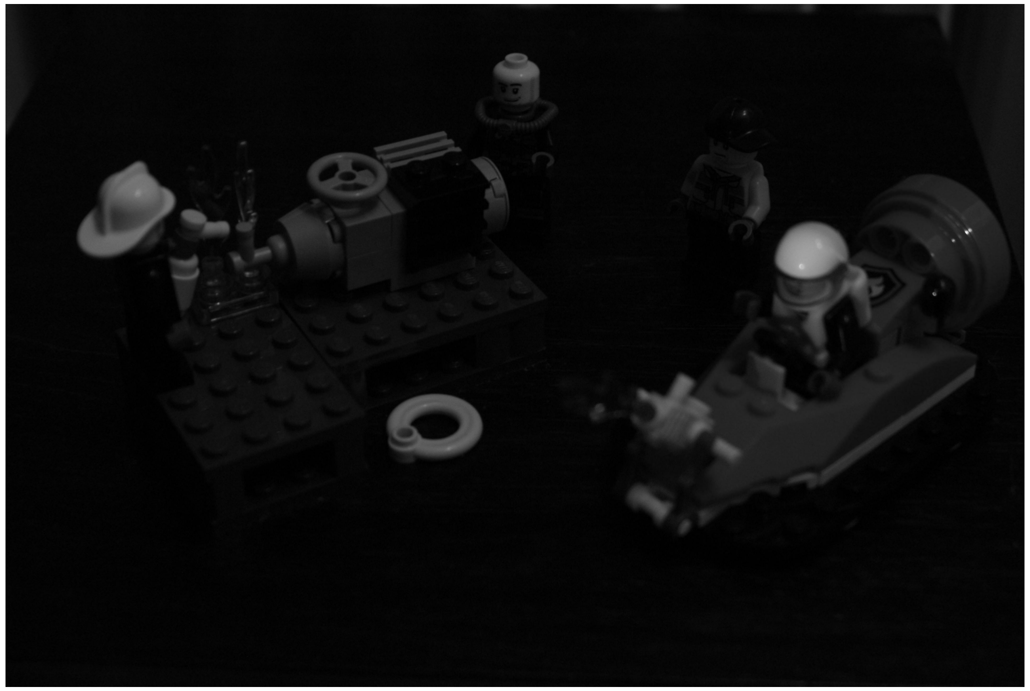


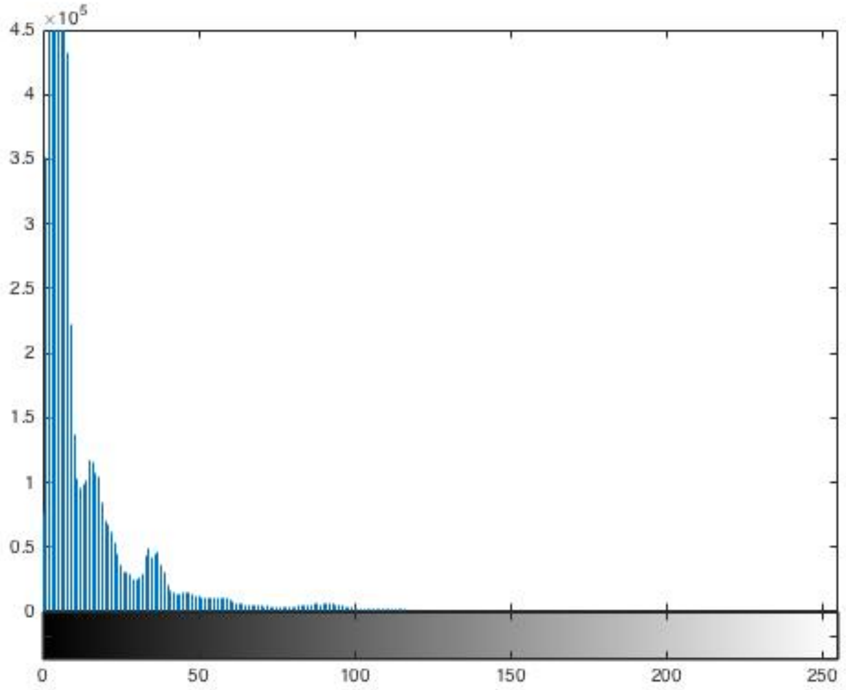
Histogram equalization

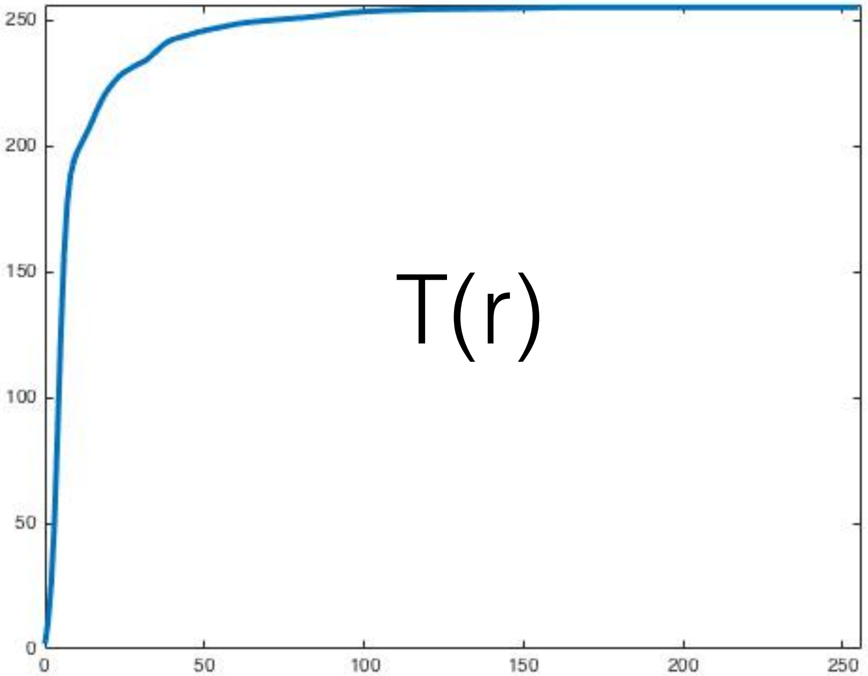




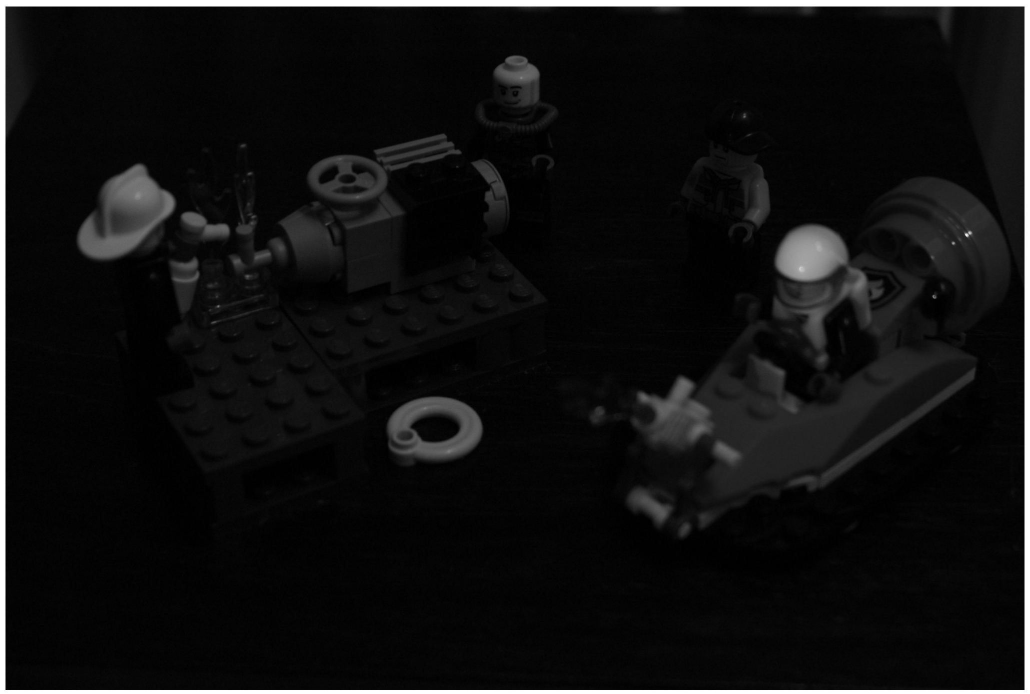
Histogram equalization

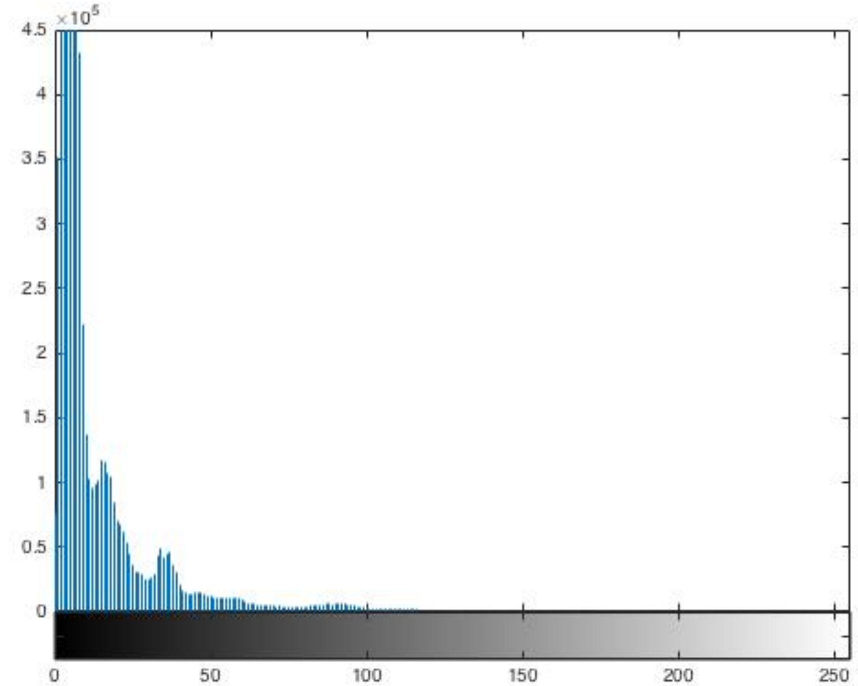


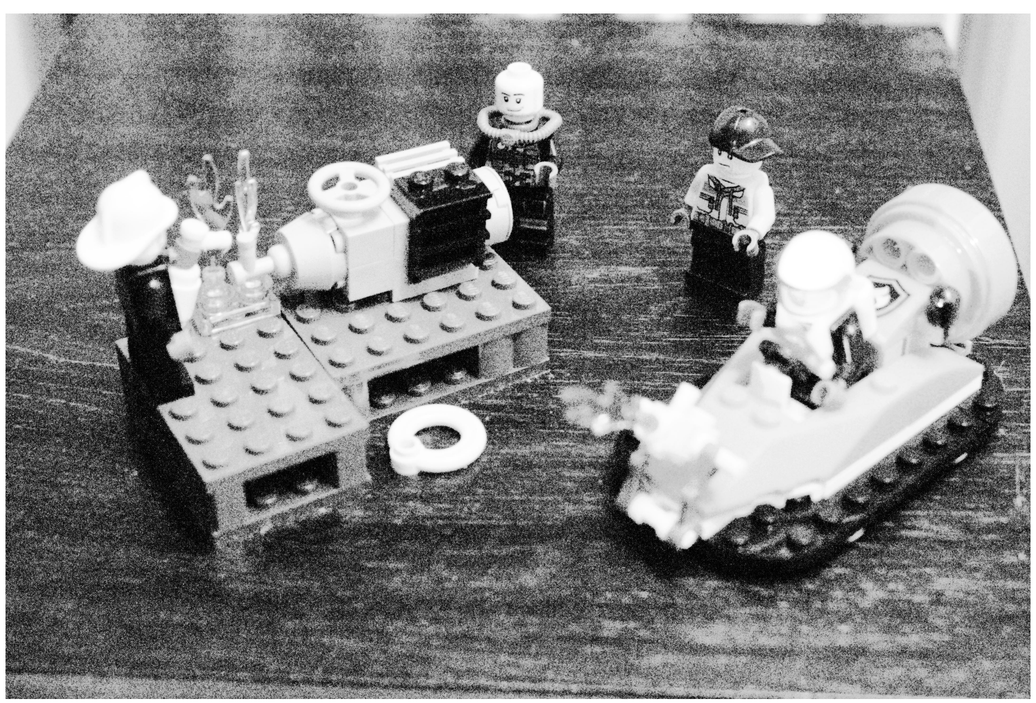


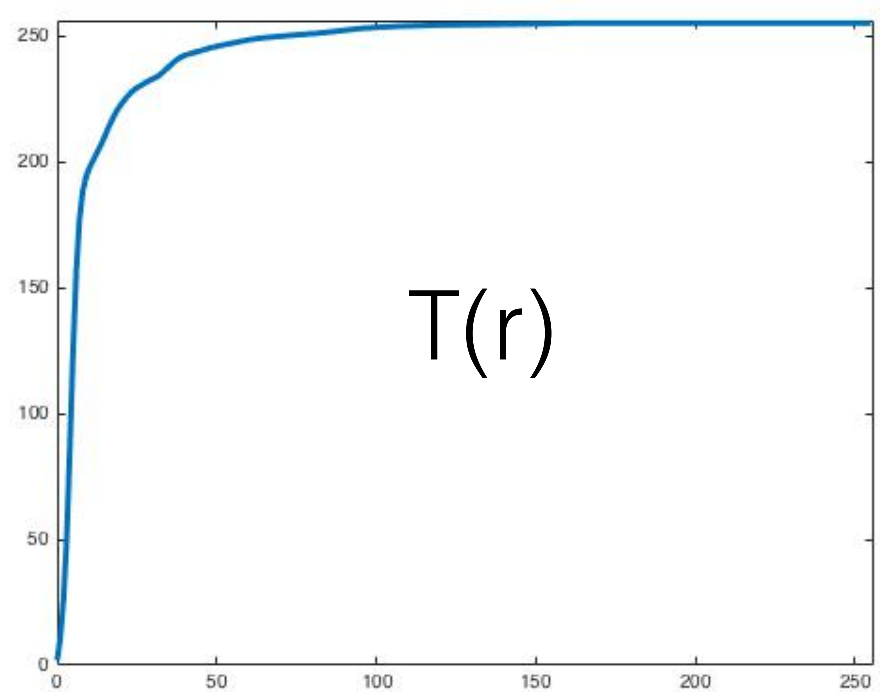


Histogram equalization

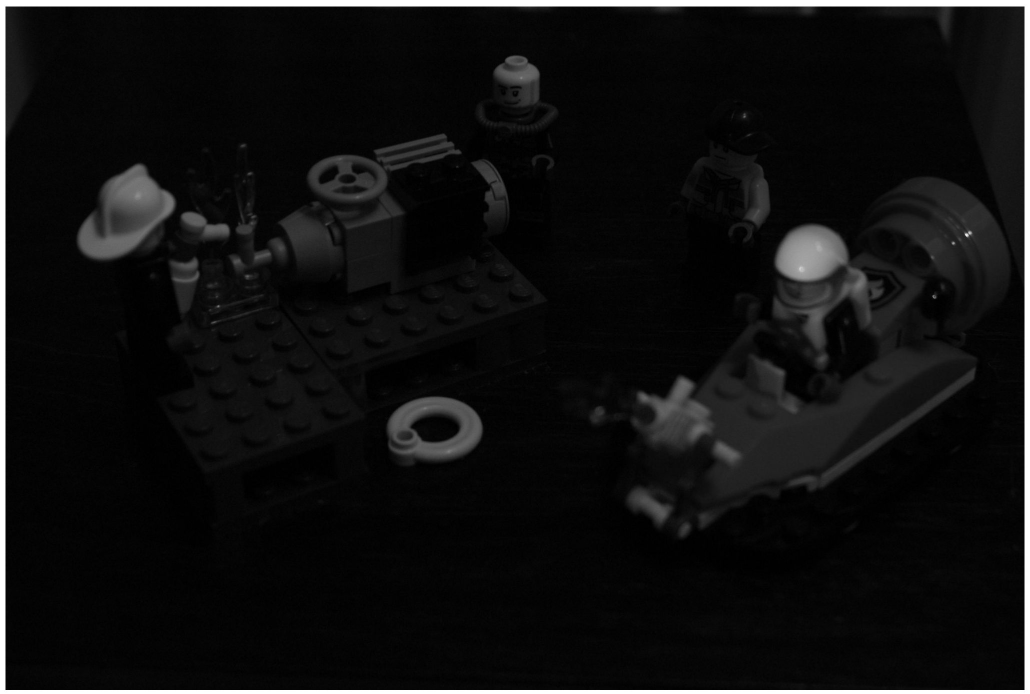


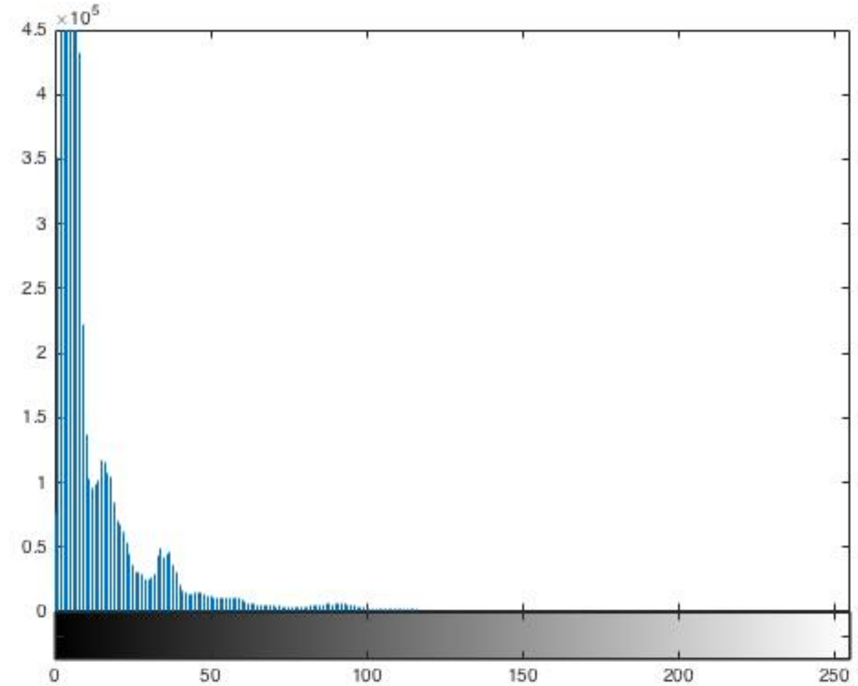


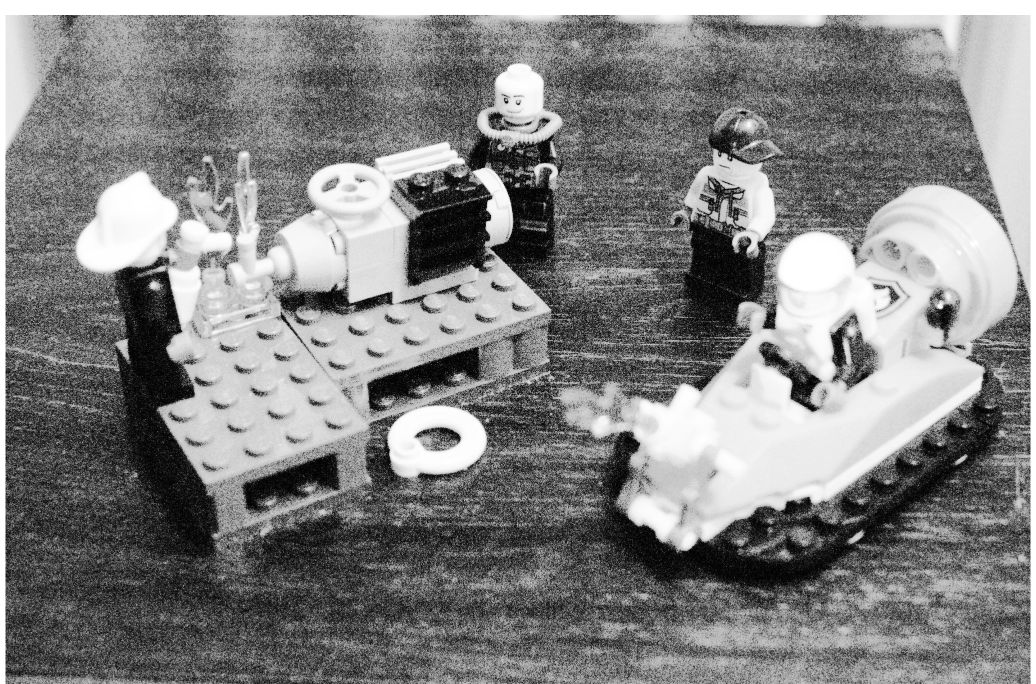


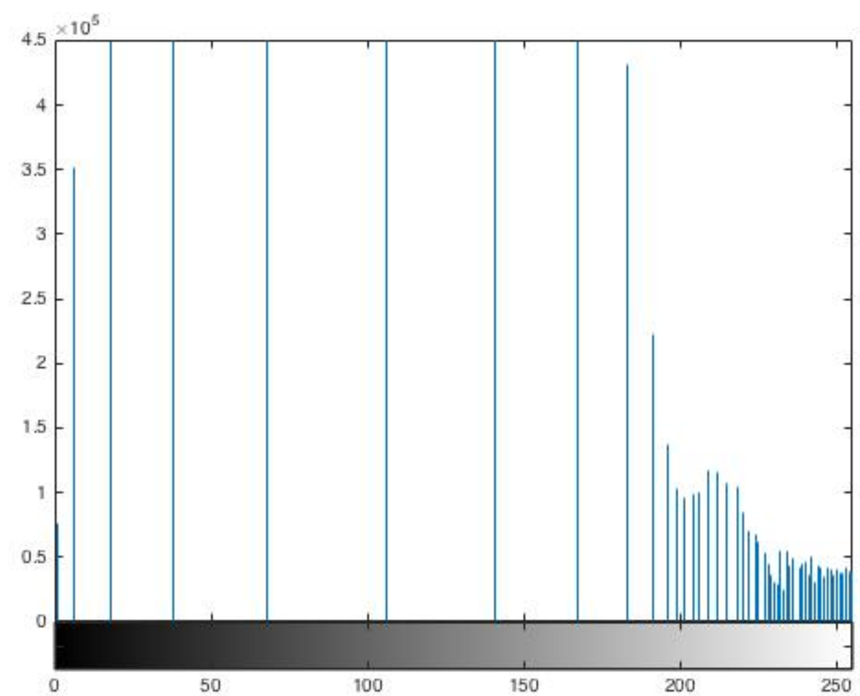


Histogram equalization









Review

- What is image analysis
- Image models
- Sampling and Interpolation
- Discrete Geometry and ’bwlabel’
- Gray-level transformations, histograms and histogram equalization

- Read lecture notes
- Experiment with matlab demo scripts
- Start working on assignment 1

Master’s Thesis Suggestion of the day

- Make a system that takes inventory of a bookshelf
- I want a drone that takes inventory every night and an app that can be used to search for the right book. The drone should fly and point at the right book, when I ask for it. Voice interface.
- Help the professor. Where is my book?

LUND UNIVERSITY

350

Collaborations

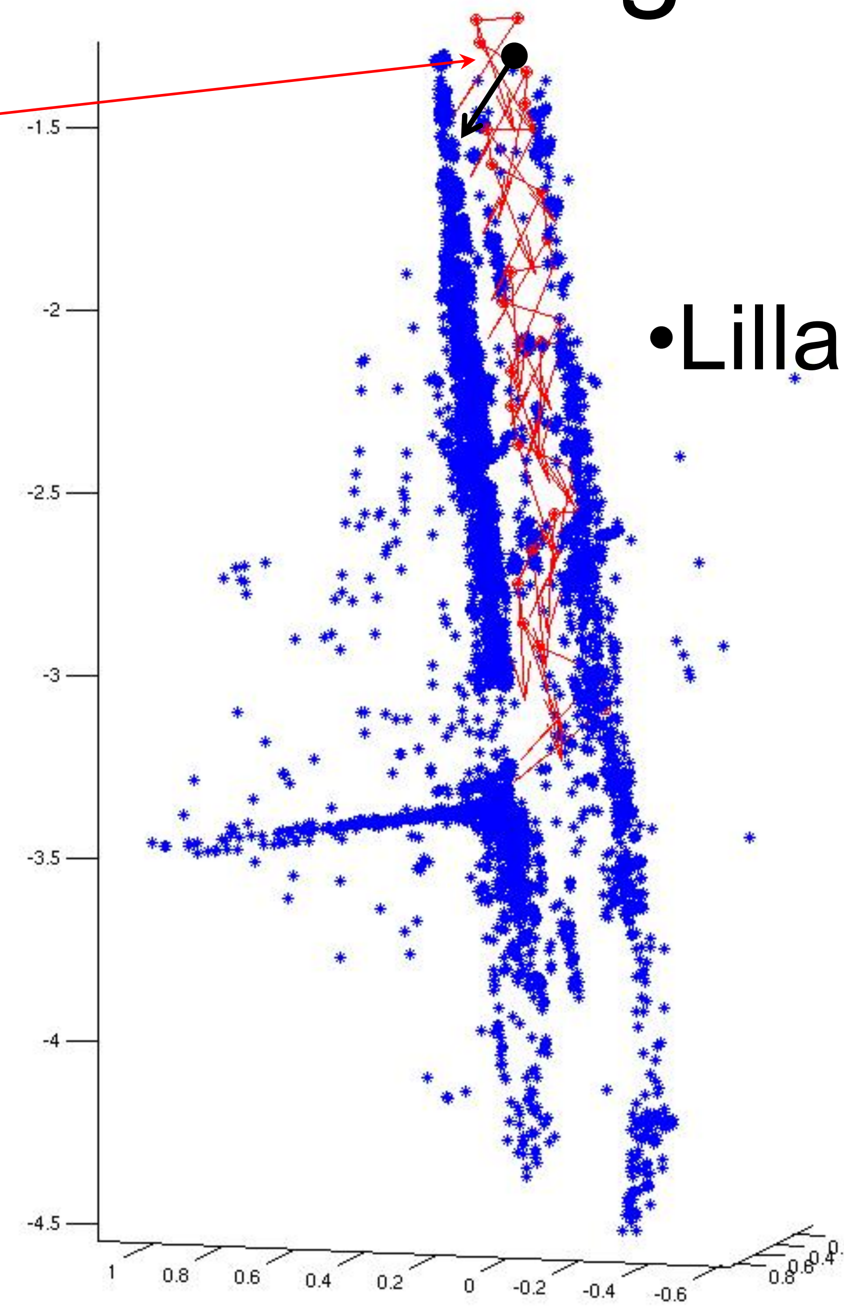
- Automatic control
- Robotics
- Traffic safety anlysis
- MR
- Orthopaedics
- Radiology
- Cancer research
- Computer Science
- EIT
- Architecture
- Food (Livsmedelsteknik)
- SLU
- …

- Sony
- Ericsson
- Axis
- Precise Biometrics
- Cellavision
- Anoto
- Exini
- Apple
- Google
- Danaher motion
- Cognimatics
- Decuma
- Polar Rose
- Spiideo
- Nocturnal Vision
- …

Where was this image taken?

•Where was this image taken?

•Stortorget

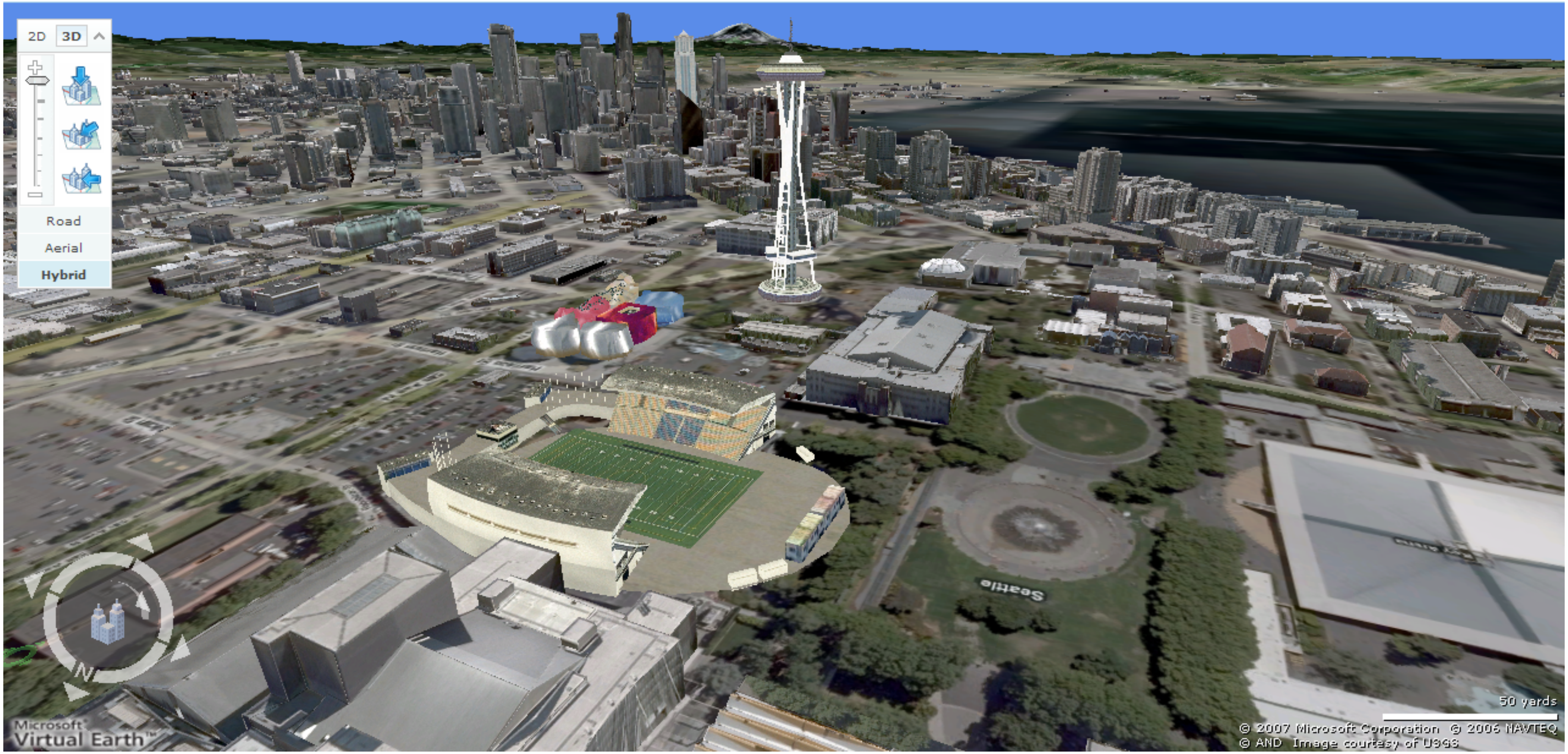


•Lilla Fiskaregatan

Object recognition (in mobile phones)

- This is becoming real:
  - Lincoln Microsoft Research
  - [Point & Find](), [Nokia]()
  - SnapTell.com (now amazon)

Earth viewers (3D modeling)



- Image from Microsoft's Virtual Earth
- (see also: Google Earth)

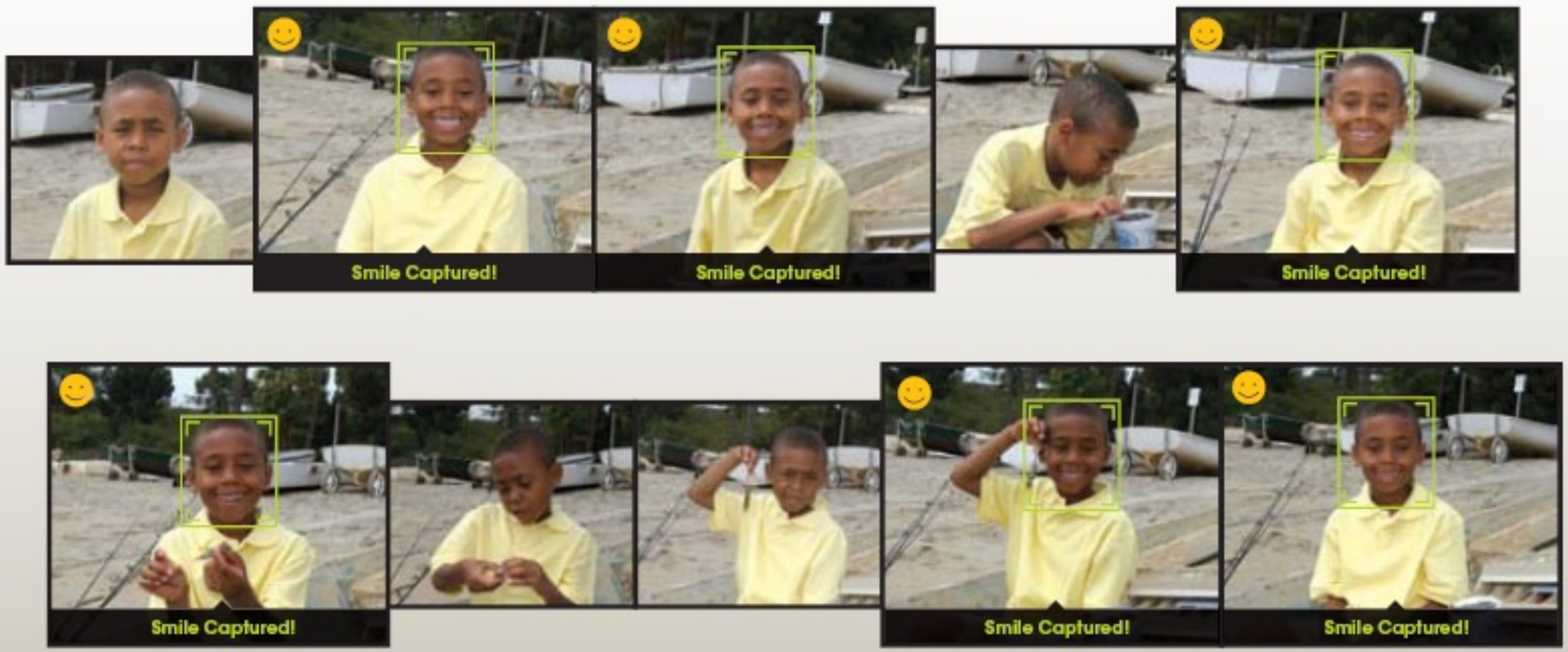
Face detection

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

Smile detection?

The Smile Shutter flow

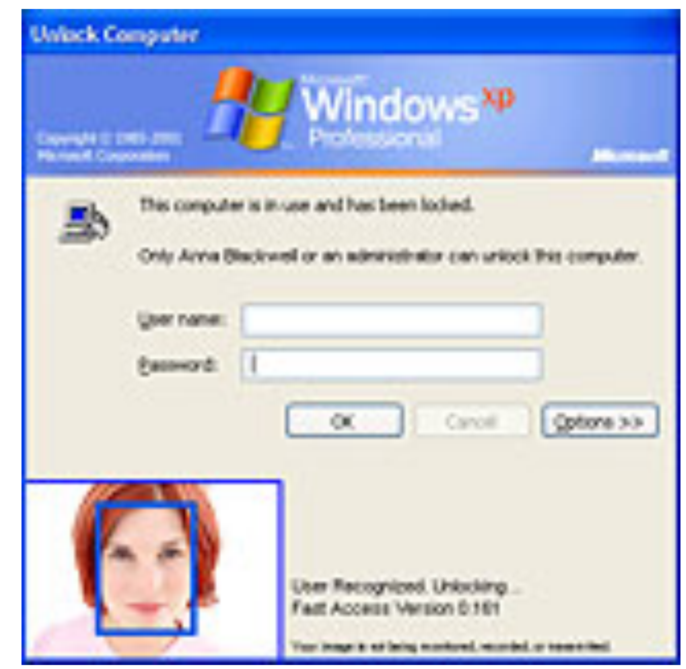
Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



• [Sony Cyber-shot® T70 Digital Still Camera]

Login without a password…

•Fingerprint scanners on many new laptops, other devices



•Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Special effects: shape capture

• *The Matrix* movies, ESC Entertainment, XYZRGB, NRC

Special effects: motion capture

•*Pirates of the Carribean*, Industrial Light and Magic

•Click here for interactive demo

Sports

•*Sportvision* first down line

•Nice explanation on www.howstuffworks.com

Smart cars

- [Mobileye]()
  - Vision systems currently in high-end BMW, GM, Volvo models
  - By 2010: 70% of car manufacturers.
  - [Video demo]()

•Slide content courtesy of Amnon Shashua

Smart cars

- Mobileye
  - Vision systems currently in high-end BMW,
  - By 2010: 70% of car manufacturers.
  - Video demo

•Slide content courtesy of Amnon Shashua

Vision-based interaction (and games)

•Nintendo Wii has camera-based IR tracking built in. See [Lee's work at CMU](#) on clever tricks on using it to create a [multi-touch display](#)!

•[Digimask](#): put your face on a 3D avatar.

•[*"Game turns moviegoers into Human Joysticks"*,](#) CNET

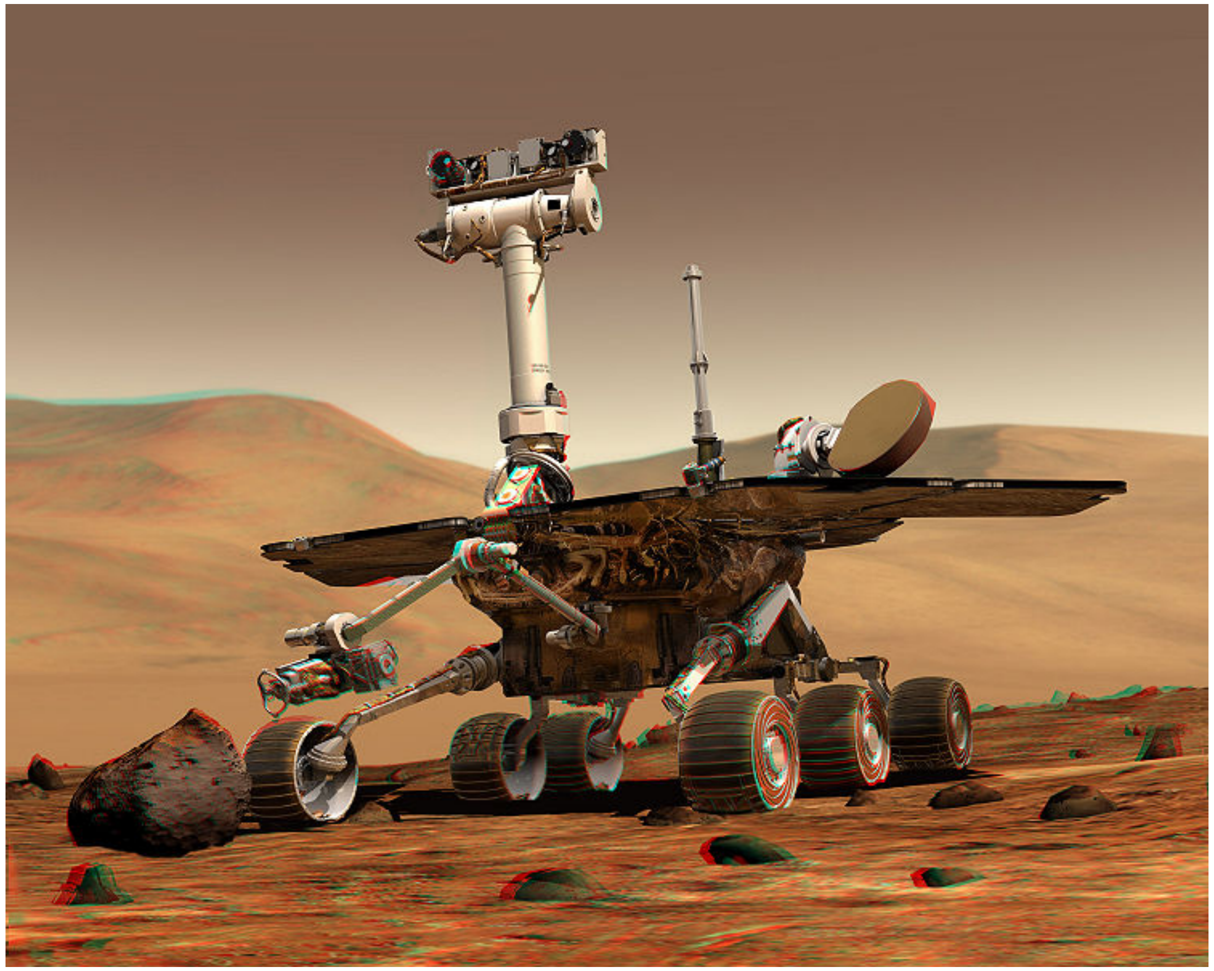
•Camera tracking a crowd, based on [this work](#).

Vision in space

•NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

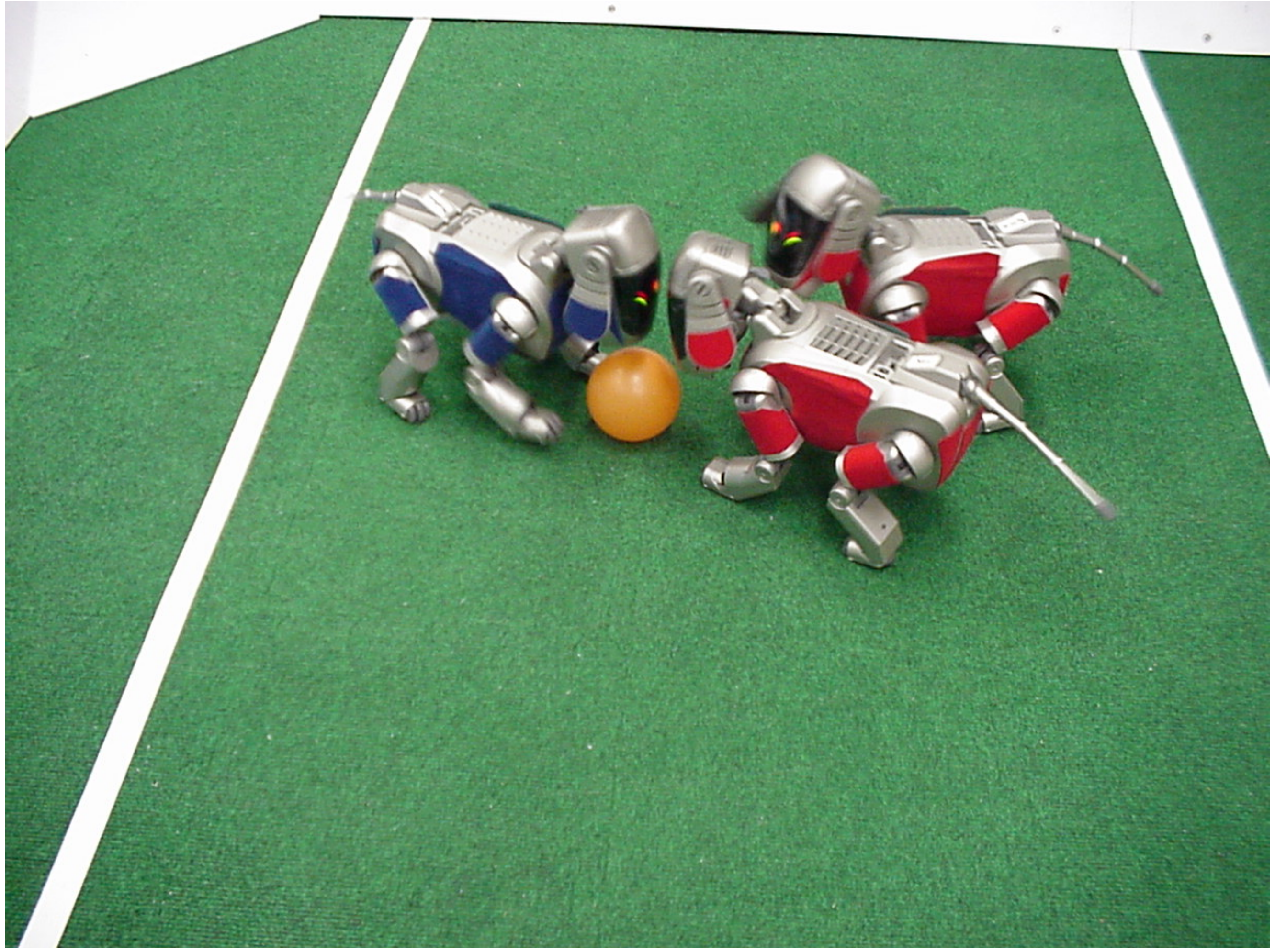
- Vision systems (JPL) used for several tasks
  - Panorama stitching
  - 3D terrain modeling
  - Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Robotics



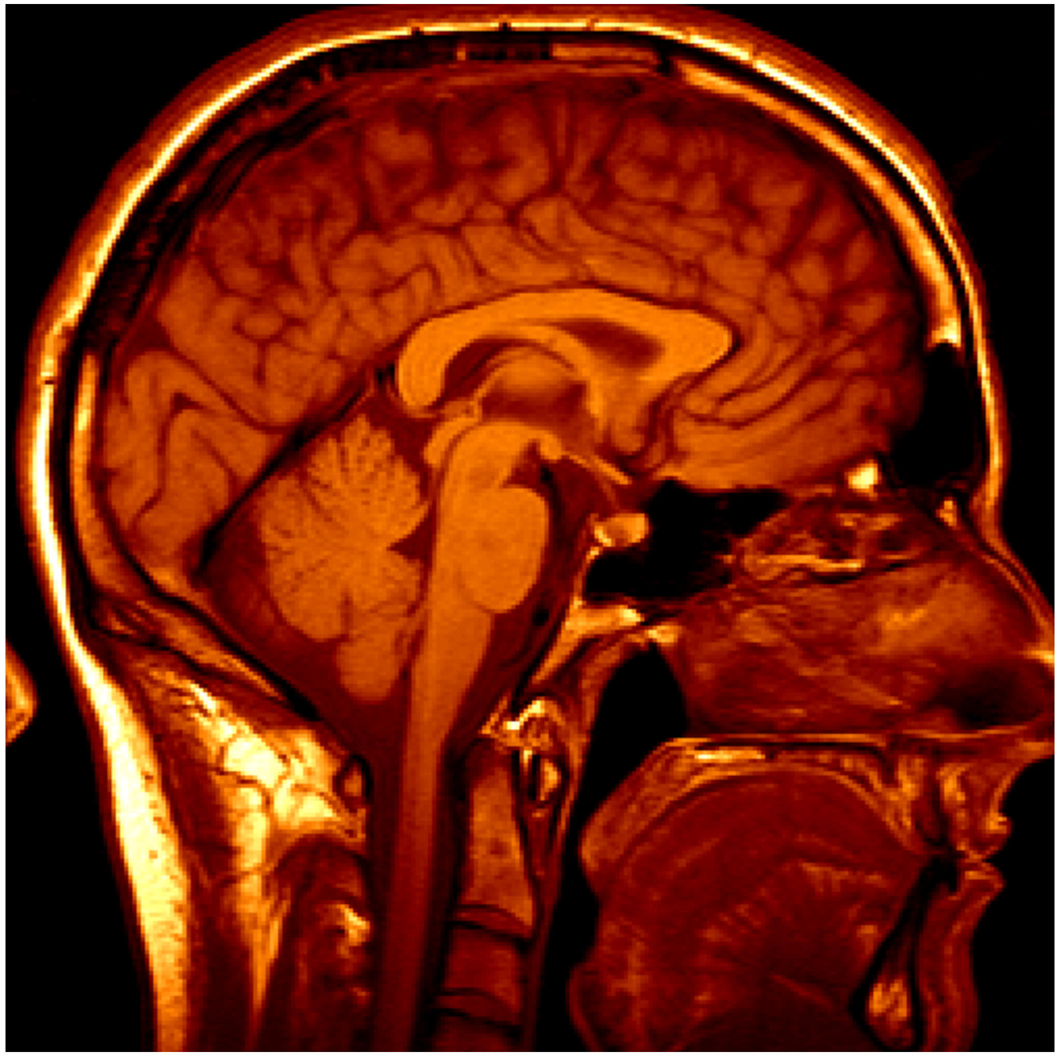
•NASA's Mars Spirit Rover

•http://en.wikipedia.org/wiki/Spirit_rover



•http://www.robocup.org/

Medical imaging



•3D imaging

•MRI, CT

•Image guided surgery

•Grimson et al., MIT