Image Analysis - Lecture 1

Magnus Oskarsson

Lecture 1

- Administrative things
- What is image analysis?
- Examples of image analysis
- Image models

Information

Lectures: 14 $\times$ 2h,
Exercises: 7 $\times$ 2h,
Lab sessions: 4 $\times$ 2h, weeks 2, 3, 4 and 6 (compulsory)
Handins: 5 (compulsory)
Project: Next study period (optional)
Credits: 6 without project, 9 with project
Pass on course (grade 3): Laboratory sessions ok + handins ok
Pass on course (grades 4 and 5): Laboratory sessions ok + handins ok + Written exam (hemtenta) + Oral exam

The Course

- contains information about constructing image systems
- contains general mathematical tools

F1 - Introduction, image models
F2 - Linear algebra, algebra of images, Fourier transform
F3 - Linear filters, convolution
F4 - Scale space theory, edge detection
F5 - Texture, segmentation, clustering
F6 - Segmentation: graph-based methods
F7 - Fitting, Hough transform and robust estimators
F8 - Active contours, snakes
F9 - Recognition and classification
F10 - Statistical image analysis
F11 - Multispectral images
F12 - Model selection, image search and applications
F13 - Computer Vision
F14 - Repetition.

Image analysis

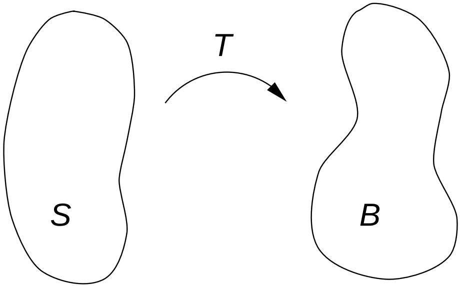


Image processing: Enhance the image (image -> image)
Image analysis: Interpret the image (image -> interpretation)
Computer vision: Mimic human vision, geometry, interpretation
Computer graphics: Generate images from models

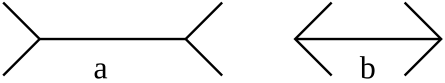
Computer vision

Computer vision - attempt to mimic human visual function
Examples:

- Recognition
- Navigation
- Reconstruction
- Scene understanding

Perceptual problems

Example 1:

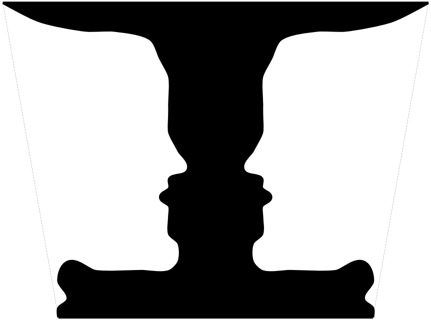


What is true ?

1. In the figure $a = b$.
2. In the figure $a > b$.

Perceptual problems (ctd.)

Exemple 2:



1. This is an image of a vase
2. This is an image of two faces.

Mathematical Imaging Group, Centre for mathematical sciences

- **Research projects**: EU, VR, SSF, Industry
- **Masters thesis projects**
- **SSBA**
- **Industry research**: NDC, Decuma, Ludesi, Gasoptics, Exini, Cellavision, Precise Biometrics, Anoto, Wespot, Cognimatics, Polar Rose, Nocturnal Vision

Related courses

- Multispectral imaging 6hp
- Computer vision 6hp
- Statistical Image Analysis 6hp
- Digital pictures – compression 9hp
- Courses in Copenhagen and Malmö

Research areas

- Geometry and computer vision
- Medical image analysis
- Cognitive vision

Continuous model

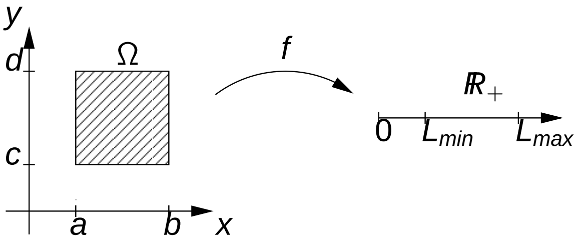
An image can be seen as a function

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

where $\Omega = \{\,(x, y) \mid a \leq x \leq b,\, c \leq y \leq d\} \subseteq \mathbb{R}^2$ and $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$. $f(x, y)$ = intensity at point (x, y) = gray-level
(f does not have to be continuous)
$0 \leq L_{min} \leq f \leq L_{max} \leq \infty$
$[L_{min}, L_{max}]$ = gray-scale

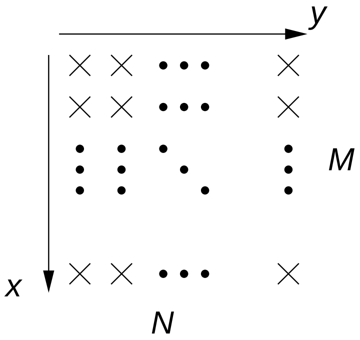
Continuous model (ctd.)

Change to gray-scale $[0, L]$ where 0='black' and L='white'.



Discrete model

Discretise x, y, called **sampling**.
Discretise f, called **quantization**.
Sampling:
Point grid in xy-plane.



Sampling

$$f(x,y) \mapsto \begin{pmatrix} f_{0,0} & \cdots & f_{0,N-1} \\ \vdots & f_{j,k} & \vdots \\ f_{M-1,0} & \cdots & f_{M-1,N-1} \end{pmatrix}$$

Quantization

Use G gray-levels
Usually $G = 2^m$ for some m.
NMm bits are required for storing an image
Ex: $512 \cdot 512 \cdot 8 \sim 262$kB
(256 gray-levels)
M, N decreased $\Rightarrow$ Chess-pattern
m decreased $\Rightarrow$ False contours

Sampling

Given an image with **continuous** representation it is straightforward to convert it into a **discrete** one by sampling.

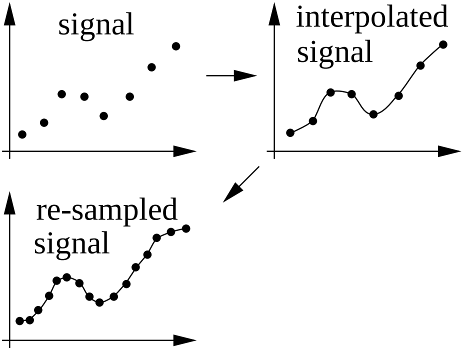
Common model for image formation is **smoothing** followed by **sampling**

Interpolation

Given an image with **discrete** representation one can obtain a **continuous** version by interpolation.

Problem: (Interpolation)
Given $f(i,j), \quad i,j \in \mathbb{Z}^2$.
"compute" $f(x,y), \quad x,y \in \mathbb{R}^2$

Re-sampling



Re-sampling (ctd.)

Problem: (Re-sampling)
Given $f(i,j), \quad i,j \in \mathbb{Z}^2$.
"Compute" $f(x,y), \quad x,y \in \mathbb{R}^2$

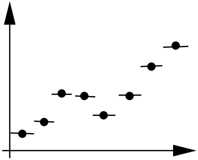
Discrete image -> Interpolation -> continuous image -> sampling -> New discrete image in different resolution

Used frequently on computers when displaying an image in a different size, thus needing a different resolution.

Nearest neighbour (pixel replication)

$$f(x,y) = f(i,j),$$

where (i,j) is the grid point closest to (x,y).



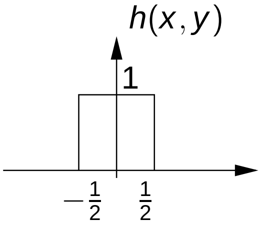
Called **pixel replication**.

Nearest neighbour

Pixel replication can be seen as interpolation with

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

where

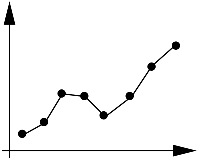


Notice the similaraty to **convolution**.

Linear interpolation

In one dimension

$$f(x) = (x - i)f(i+1) + (i+1-x)f(i), \quad i < x < i+1$$



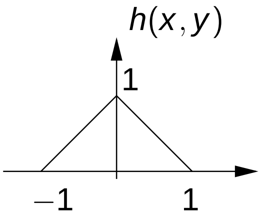
Called **linear inperpolation.**

Linear interpolation (ctd.)

Linear interpolation can bee seen as convolution

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

with a different interpolation function h:



Two dimensions

$$
\begin{aligned}
f(x,y) =&(i+1-x)(j+1-y)f(i,j)+ \\
&+(x-i)(j+1-y)f(i+1,j)+ \\
&+(i+1-x)(y-j)f(i,j+1)+ \\
&+(x-i)(y-j)f(i+1,j+1), \\
&i<x<i+1, j<y<j+1
\end{aligned}
$$

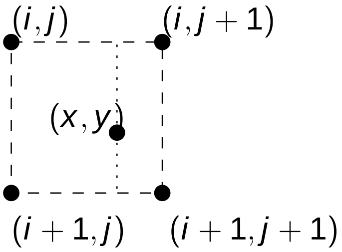
Called **bilinear interpolation.** Between grid points the intensity is

$$
f(x,y) = ax + by + cxy + d,
$$

where a, b, c, d is determined by the gray-levels in the corner points.

Bilinear interpolation

For two-dimensional signals (images) we can apply linear interpolation, first in x-direction and then y-direction.



Cubic interpolation (Cubic spline)

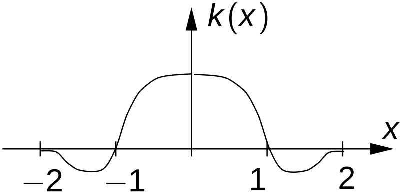
Define a function k such that

$$k(x) = \begin{cases} a_3x^3 + a_2x^2 + a_1x + a_0 & x \in [0,1] \\ b_3x^3 + b_2x^2 + b_1x + b_0 & x \in [1,2] \\ 0 & x \in [2,\infty) \end{cases}$$

and

- k symmetric around the origin
- $k(0) = 1$, $k(1) = k(2) = 0$
- k and k' continuous at $x = 1$
- $k'(0) = k'(2) = 0$

Cubic spline function



Determination of a_i and b_i

These conditions give

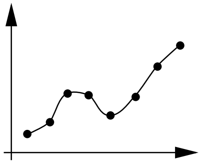
$$k(x) = \begin{cases} (a+2)x^3 - (a+3)x^2 + 1 & x \in [0,1] \\ ax^3 - 5ax^2 + 8ax - 4a & x \in [1,2] \end{cases}$$

where a is a free parameter.
Common choice is $a = -1$.
Interpolation is done using the convolution

$$f(x) = \sum_i f(i)k(x-i)$$

Cubic interpolation for images

For images one first interpolates in x-direction and then in y-direction.



Sinc interpolation

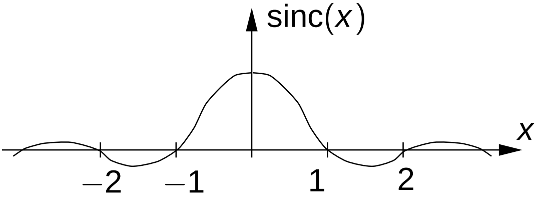
Assume that $f(x)$ is a band-limited signal.
Sampling theorem:

$$f(x) = \sum_k \operatorname{sinc}(2\pi(x-k))f(k)$$

Sketch of proof: Fouriertransform $F(\omega)$ is band limited. Thus it can be written as a fourier series, where the coefficients are $f(k)$. Inverse fouriertransform completes the proof.
Drawback: sinc has unlimited support $\Rightarrow$ large filter $\Rightarrow$ time consuming.
Solution: Cut sinc after the first or the first few oscilations $\Rightarrow$ almost like cubic interpolation.

Sinc interpolation for images

For images one interpolates first in x-direction and then in y-direction.

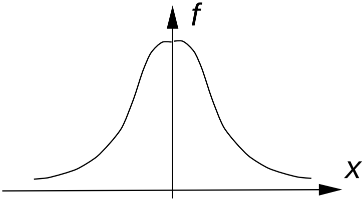


Gauss interpolation

Interpolate with

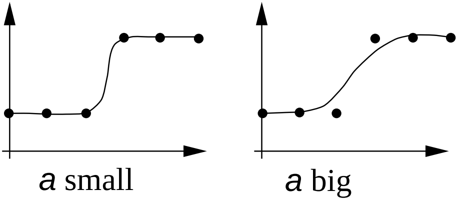
$$f(x) = \sum_k e^{-(x-k)^2/a^2} f(k)$$

where *a* determines 'scale/resolution/blurriness'.



Scale selection

Gives a scale-space pyramid with the same image at different scales by changing *a*. More about this later.

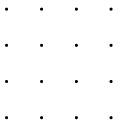


This is called **Gaussian pyramid** or **scale space pyramid** or **scale space representation**.

Digital Geometry

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

Grid: $\mathbb{Z}^2$,



Grid point: (x, y)

Definition

4-**neigbourhood** to (x, y):

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

■

Neighbours, connectedness, paths

Definition

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

Definition

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

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

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).

D- and 8-neighbourhoods

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):

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

■

Gray-level transformation

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)

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

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

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

■

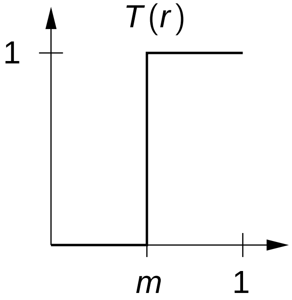
In many examples we assume that $L_{min} = 0$ och $L_{max} = L = 1$. The requirements on T being increasing can be relaxed, e.g. with inversion.

Thresholding

Let

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

for some $0 < m < 1$.



Thresholding (ctd.)

i.e.

$$f(x,y) \leq m \Rightarrow g(x,y) = 0 \quad \text{(black)},$$
$$f(x,y) > m \Rightarrow g(x,y) = 1 \quad \text{(white)}.$$

The result is an image with only two gray-levels, 0 and 1. This is called a **binary image**.

The operation is called **thresholding**. ■

Continuous images

- Let $s = T(r)$ be a gray-scale transformation ($r = T^{-1}(s)$)
- Let $p_r(r)$ be the frequency function for the original image.
- Let $p_s(s)$ be the frequency function for the resulting image.

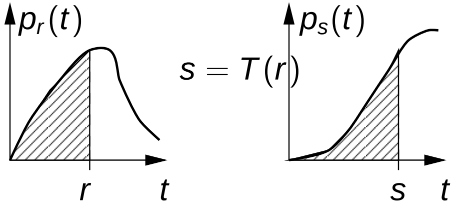
It follows that

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

Continuous images (ctd.)

Differentiate with respect to s

$$p_s(s) = p_r(r)\frac{dr}{ds} \quad (s = T(r)) \; .$$



Histogram equalization

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

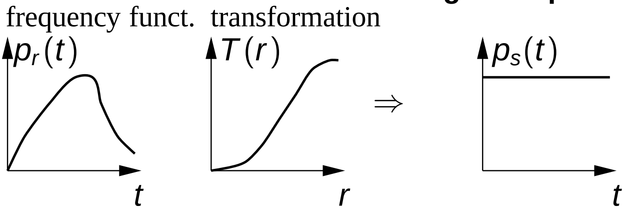
$$\int_0^r p_r(t)dt = \int_0^s 1dt = s \Rightarrow s = T(r) = \int_0^r p_r(t)dt$$

or

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

Histogram equalization (ctd.)

This transformation is called **histogram equalization**.



Histogram equalization for digital images

$$p_r(r_k) = \frac{n_k}{n} \quad ,$$

where

- n=number of pixels
- n_k=number of pixels with intensity r_k

i.e. a histogram.
Histogram equalization is obtained by

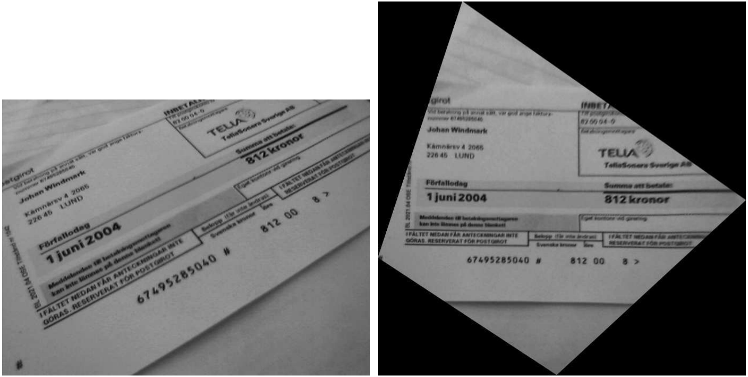
$$s_k = T(r_k) = \sum_{j=0}^{k} \frac{n_j}{n} \quad .$$

Note that s_k does not have to be an allowed gray-scale $\Rightarrow$ perfect equalization cannot be obtained.

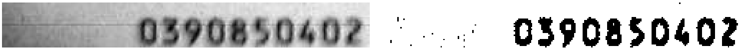
Example OCR (Optical Character Recognition)

- Image of text
- Image enhancement, filtering.
- Segmentation
  - Thresholding
  - Connected components with digital metrics.
- Classification

Images show how a system for OCR (Optical Character Recognition) can be used in a mobile telephone.
The binary image is interpreted into ascii characters.



Original image and rectified image.



Cut-out of OCR number after thresholding.

Masters thesis suggestion of the day: The automatic book database

Create a system for taking inventory of your books by taking images of them and analysing the images.
Images - segmentation - OCR - Database - Search - what is missing? - etc.

Repetition - Lecture 1

- What is image analysis?
- Image models (continuous - discrete - sampling - quantization, sampling and interpolation)
- Digital geometry (4-, D-, 8- neighbours, paths, connected components)
- Gray-level transformations (thresholding, histogram equalization)

Recommended reading

- Forsyth & Ponce: **1. Cameras**.
- Szeliski: **1. Introduction** and **3.1 Point operators**.

i.e.	id est	that is	det vill säga
e.g.	exempli gratia	for example	till exempel
cf.	confer	compare with (see)	jämför, se