D E P A R T M E N T O F C O M P U T E R S C I E N C E
U N I V E R S I T Y O F C O P E N H A G E N

Corners and Interest Points Matching

François Lauze

Today

- Corners – statistical analysis of gradient directions
- Harris, Plessey mf.
- Interest points matching
- Introduction to descriptors

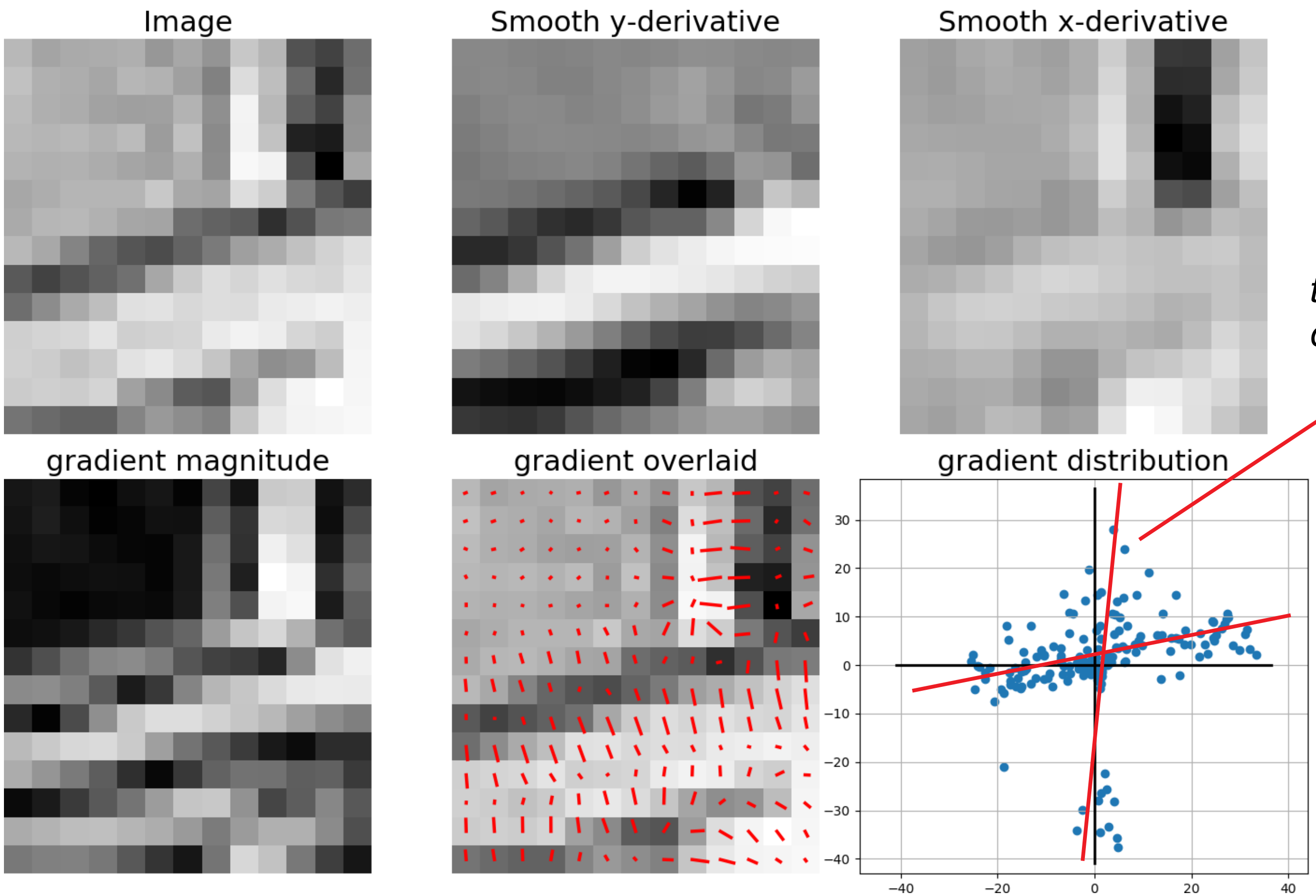
Corners and Edges

- There exist a number of approaches to corner detecting. Here we will focus on a single type based on analyzing the statistics of gradients.

A corner

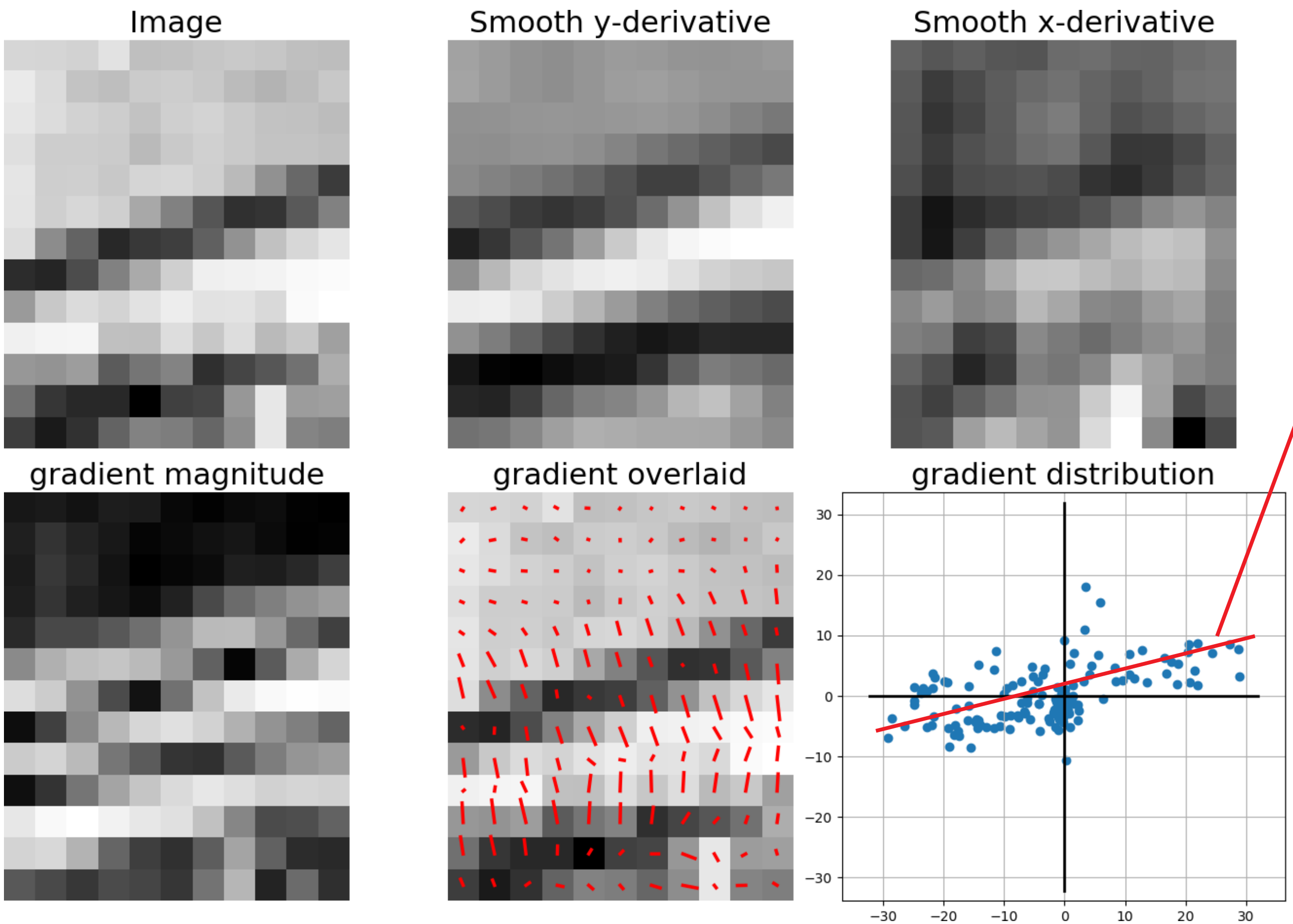
an edge

Gradient Distribution of Corners



two main directions

Gradient Distribution of Edges



one main direction

The structure tensor

Let (G_x^i, G_y^i) be the i'th gradient in a square area around a potential corner point. In practice each gradient is weighted by a Gaussian of the distance to the point.

Let $A_{[Nx2]}$ be the stack of all N gradients

$$A = \begin{bmatrix} G_x^1 & G_y^1 \\ G_x^2 & G_y^2 \\ \vdots & \vdots \\ G_x^n & G_y^n \end{bmatrix}$$

Let $T = \frac{1}{n} A^T A$ be the structure tensor, i.e. a 2x2 symmetric matrix of sums of products of gradient components.

T captures the second order statistics (i.e. variances and covariance's) of the uncentered gradient point cloud.

Structure tensor computations

$$A = \begin{bmatrix} G_x^1 & G_y^1 \\ G_x^2 & G_y^2 \\ \vdots & \vdots \\ G_x^n & G_y^n \end{bmatrix} \qquad A^T = \begin{bmatrix} G_x^1 & G_x^2 & \cdots & G_x^n \\ G_y^1 & G_y^2 & \cdots & G_y^n \end{bmatrix}$$

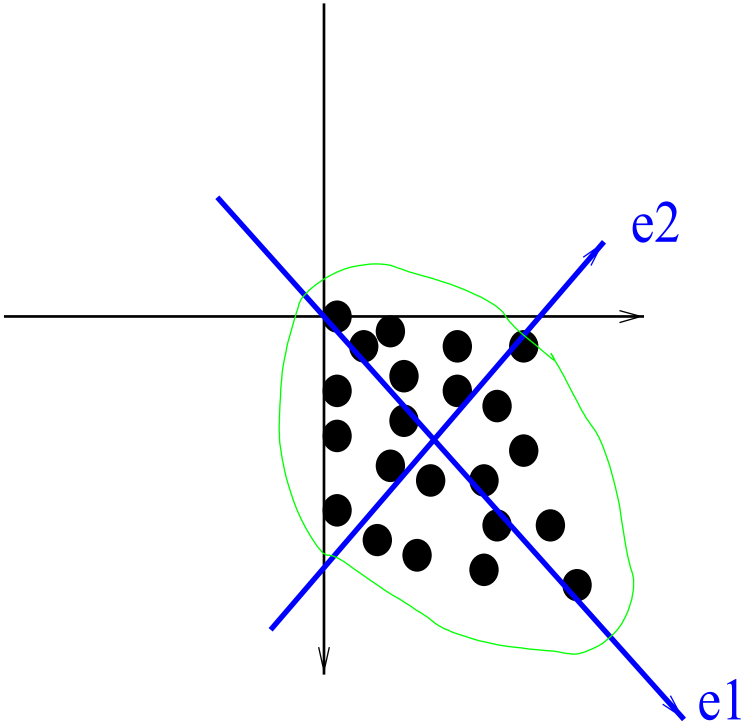
$$T = \frac{1}{n} A^T A = \frac{1}{n} \begin{bmatrix} \sum_{i=1}^n (G_x^i)^2 & \sum_{i=1}^n G_x^i G_y^i \\ \sum_{i=1}^n G_x^i G_y^i & \sum_{i=1}^n (G_y^i)^2 \end{bmatrix}$$

T is a symmetric matrix!

Structure tensor analysis

- The first eigenvector of *T* points in the direction where the gradient point cloud is broadest (has the largest variance given by the first eigenvalue).

- The second eigenvector is orthogonal to the first and points in the direction where the gradient point cloud is most condensed (has smallest variance = the 2. eigenvalue).



Corners by maximum of smallest eigenvalues

- Only if the smallest eigenvalue is large the local gradient points in several dominant directions.

With $T = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$,

$$\lambda_2 = \frac{a + c - \sqrt{(a-c)^2 + 4b^2}}{2}$$

To verify this solve $Tx = \lambda x$ or $\det(T - \lambda I_2) = 0$.

Please note that we have to take a square root

Eigenvalues

$$T = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad T - \lambda I_2 = \begin{bmatrix} a & b \\ b & c \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} a-\lambda & b \\ b & c-\lambda \end{bmatrix}$$

$$\begin{aligned} 0 = \det(T - \lambda I_2) &= (a-\lambda)(c-\lambda) - b^2 \\ &= \lambda^2 - \lambda\underbrace{(a+c)}_{\mathrm{Tr}(T)} + \underbrace{ac - b^2}_{\det(T)} \end{aligned}$$

The roots are the eigenvalues of
T

$$\lambda_1 = \frac{a + c + \sqrt{(a-c)^2 + 4b^2}}{2}$$

$$\lambda_2 = \frac{a + c - \sqrt{(a-c)^2 + 4b^2}}{2}$$

Eigenvectors

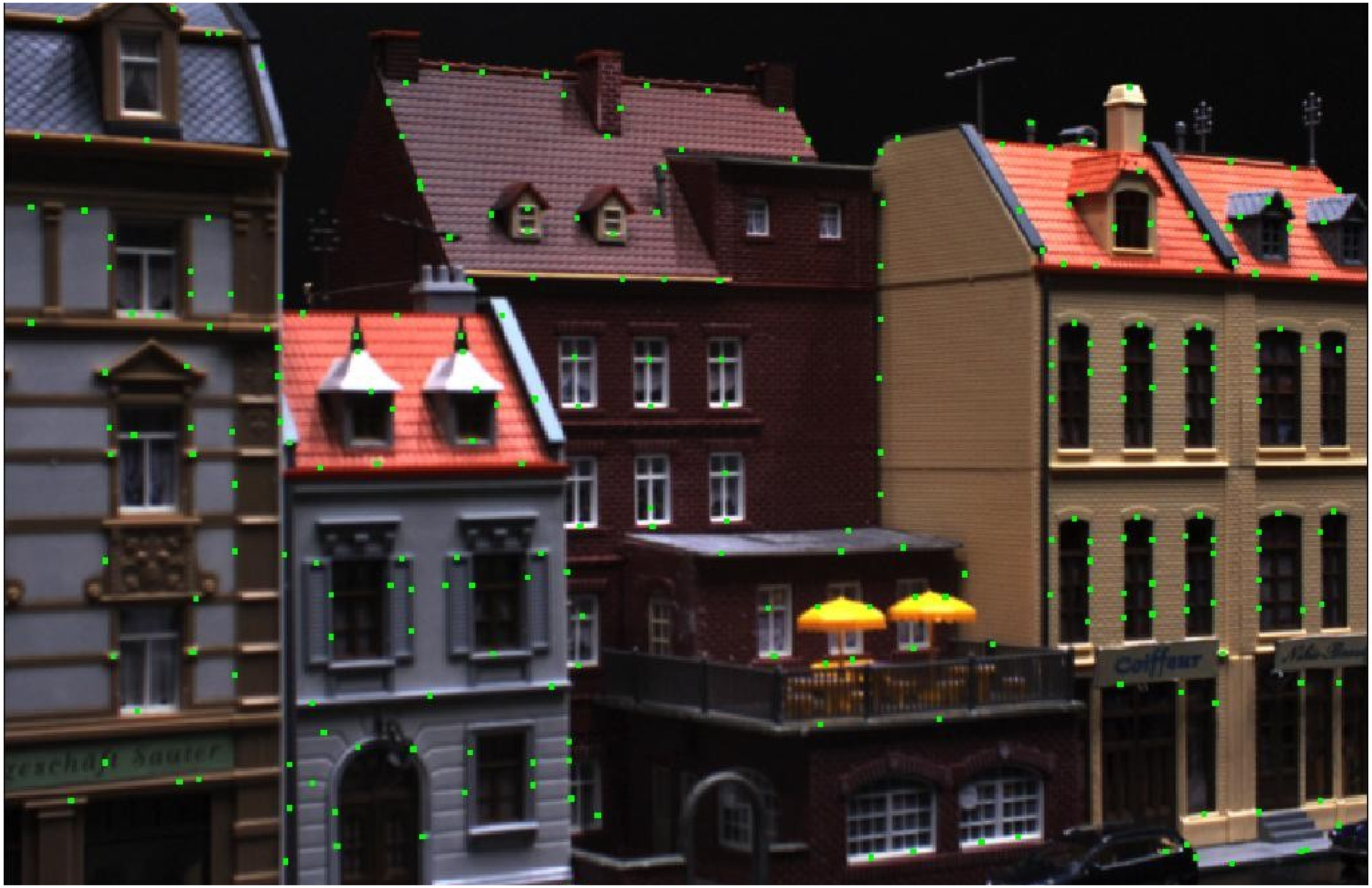
- They are the vectors *x* for which $Tx = \lambda x$, with λ one of the *eigenvalues* i.e. one of the λ_1 or λ_2 from previous slide.

- There are infinitely many solutions!

- If $\lambda_1 \neq \lambda_2$, there is one *vector line* of solution per eigenvalue, i.e. two eigenvectors for the same eigenvalue are multiple of each others. So we can choose an eigenvector with norm 1. Still there are two choices: x or $-x$

- This is what Python or Matlab returns when asked to compute eigenvectors (x or $-x$ just by chance).

Example: Shi-Tomasi corner detection



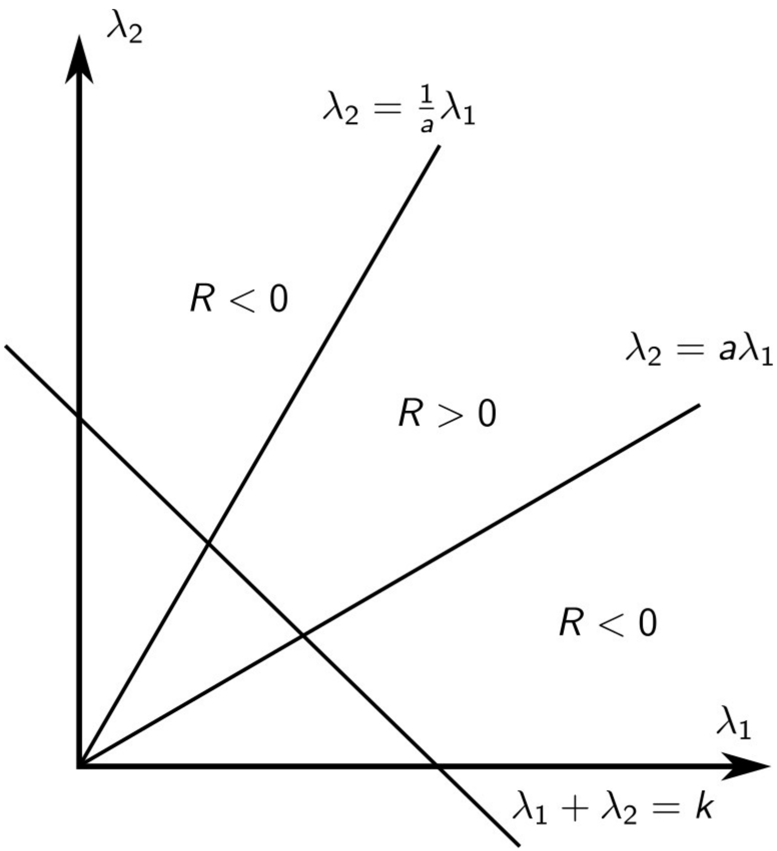
Harris corners

- The Harris detector builds on the same structure tensor T. Harris and Stephens wanted to speed up and avoid the square root extraction.

- We want:

$$a\lambda_1 < \quad \lambda_2 < \frac{1}{a}\lambda_1$$

- Where 0 < a < 1
- Corresponding to the
- central area of the figure



- A function that is positive in the central area is:

$$F(\lambda_1, \lambda_2) = (\lambda_1 - a\lambda_2)(\lambda_2 - a\lambda_1)$$

- Rewrite:

$$(1+a^2)\lambda_1\lambda_2 - a(\lambda_1^2 + \lambda_2^2) =$$

$$\lambda_1\lambda_2 - k(\lambda_1 + \lambda_2)^2$$

with $k = \dfrac{a}{(1+a)^2}$

and use that $Det(T) = \lambda_1\lambda_2 = AC - B^2$

and $tr(T) = \lambda_1 + \lambda_2 = A + C$

Harris detector

- We get $R = \det(T) - k\,\mathrm{Tr}(T)^2$
- We mark corner points as the ones with locally maximum value of R > 0.
- The value of k must be < 0.25. Often values between 0.15 and 0.20 are used.
- The detector may (as the other detectors) be extended to multiple scales.

Example Harris corner detection

Interest point detectors: Detecting Harris corners (fixed scale)

Plessey corner detector

- The Plessey approach avoid the constant *k* using*:*

$$\text{Plessey} = \frac{2\det(T)}{\mathrm{Tr}(T)+\varepsilon}$$

- Where the determinant and trace is defined as before.

- Corner points are marked when the measure locally is maximal.

- The measure *Det/Tr* is found to be the harmonic mean of the two eigenvalues, i.e.:

$$\frac{2}{\frac{1}{\lambda_1}+\frac{1}{\lambda_2}}$$

Example: Plessey corner detection

Many other corner/interest point detectors exist

- You should be aware that you have only heard a fraction of the story.
- Another detector mark points where is iso-intensity curve through the point bends locally maximal.
- You may check:
  1. The Kitchen operator
  2. The Förstner corner detector
  3. The Susan corner detector

Whatever you choose, be prepared to get something else compared to what you expected.

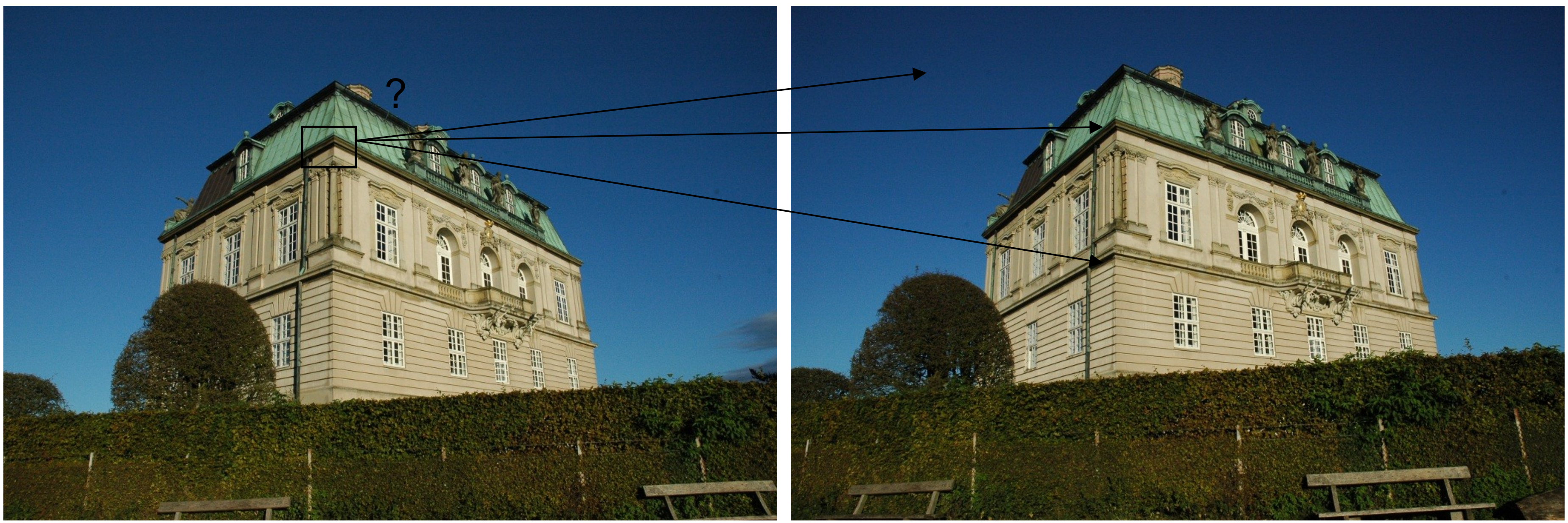
Now we are going to use the interest points

- Consider the problem of matching two or more images

- Relevant for
  - 3D reconstruction
  - Content-based image retrieval
  - Object detection and recognition
  - More tasks...

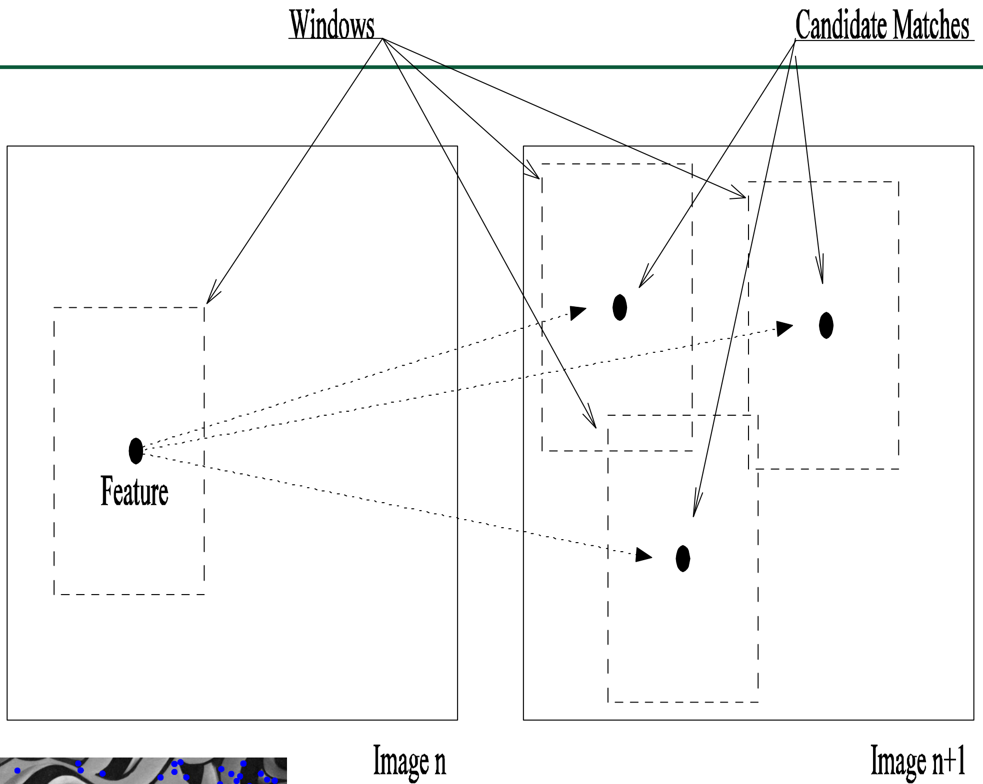
In general (if we do not know a lot about from where the images are taken) a point in one image may match any point in the other image.

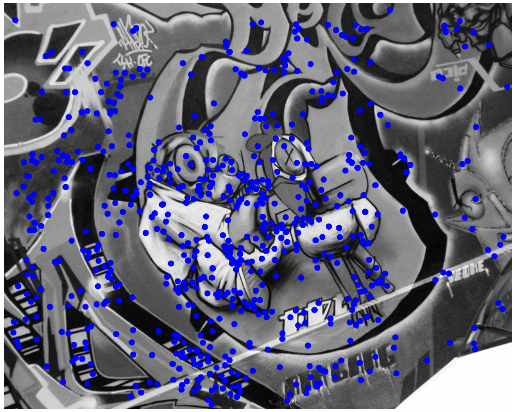
Fundamental Question:
Where did the underlying feature go?



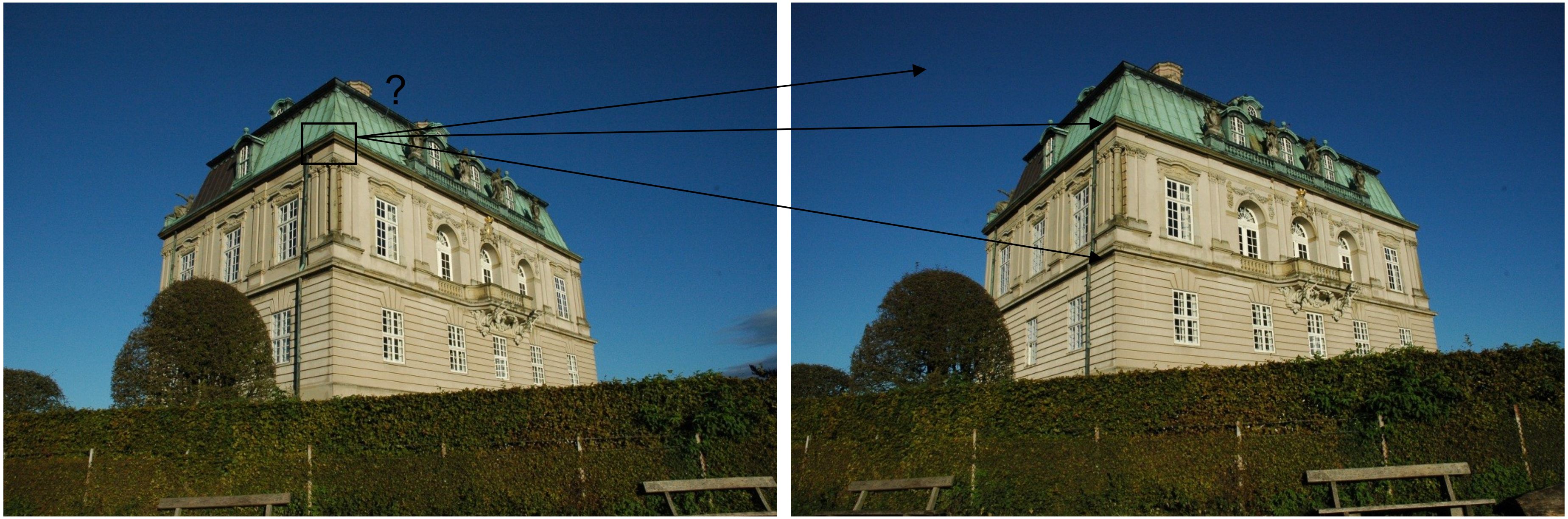
Matching: Ideally, we want to find the location in the other image which corresponds to the same physical location on the building

Matching Strategy Illustrated





Not all image patches can be matched



To be able to figure out which points belongs to which we need a bit of structure:
Sky patches are impossible to match. Occluded areas cannot be matched.
Repetitive structures like windows may easily be mismatched.

Dense versus sparse matching?

Dense : Match all pixels in the images
Sparse: Match a subset of positions in the images
Both approaches have merit, but dense is computationally expensive

Salient points = Interests points

- **Interest or Salient points**: Local structure that appear distinct from the image in the surrounding region of the salient point.
- We already know to detect such points!
- Terminology: Salient points aka interest points aka key points aka features (old confusing terminology)
- Problem: How dense do we need the points?

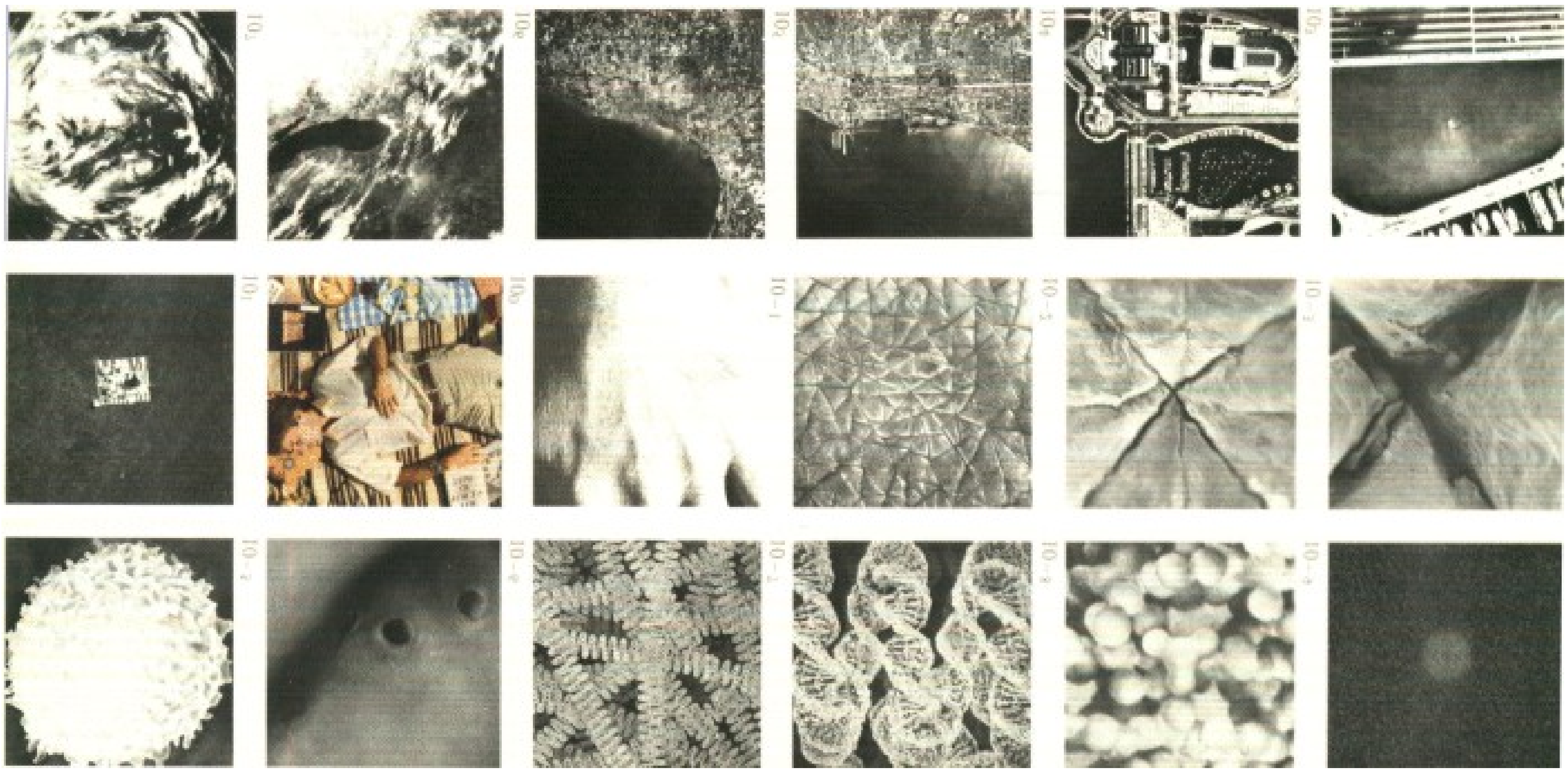
Interest point detectors: Detecting blobs

How do we know the size of things?

Multi-Scale analysis

An Introduction to Scale Space Theory

Measurements
What do we measure?



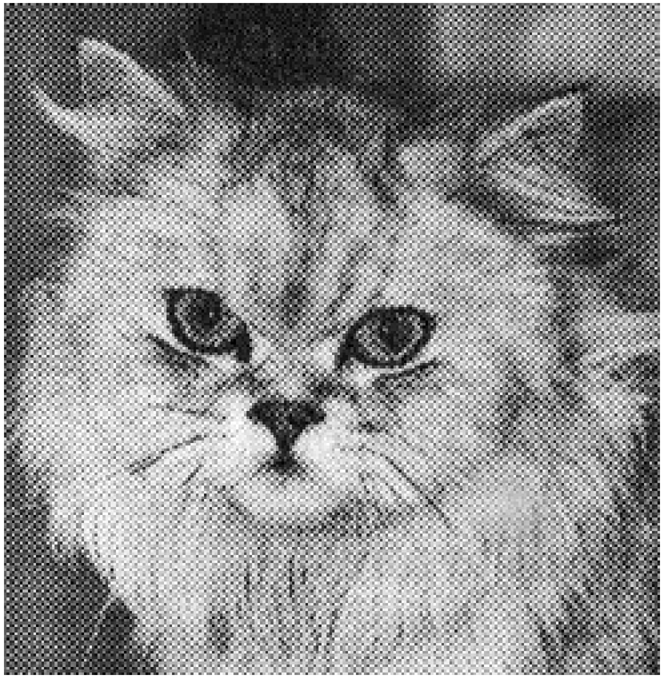
Slide by Arjan Kuijper, 2003

Observations

- Objects have a size / scale.
- Objects consists of objects of various sizes.
    - They contain several scales.
- Objects are measured by some device.
    - Cameras, the eye, …
- Devices are finite.
    - They have a minimum and a maximum detection range: the inner and outer scale. They determine the spatial resolution.
- The device must allow multi-scale structures.
    - It has to respect the various sizes of the object. The inner scale isn't always the best scale.

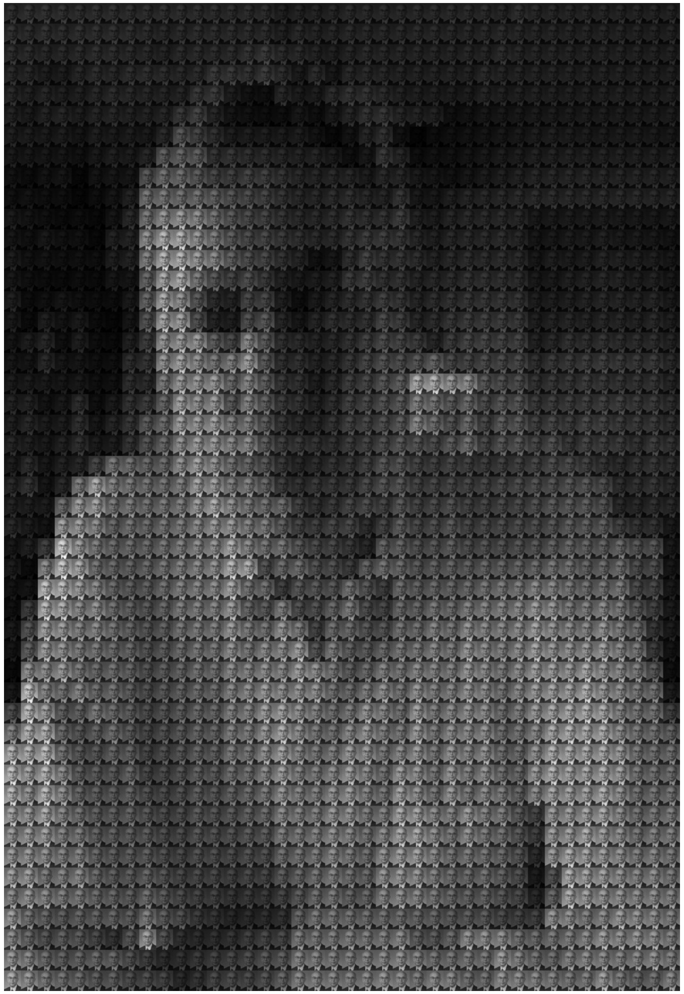
The visual system

- We see multi-scale:
  - The images only contain two values (black and white).
  - We regard them as grey level images, or see structure.

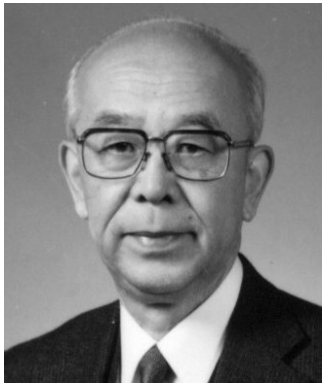


Slide by Arjan Kuijper, 2003

The visual system
(Founding fathers of scale space theory)



Jan J. Koenderink

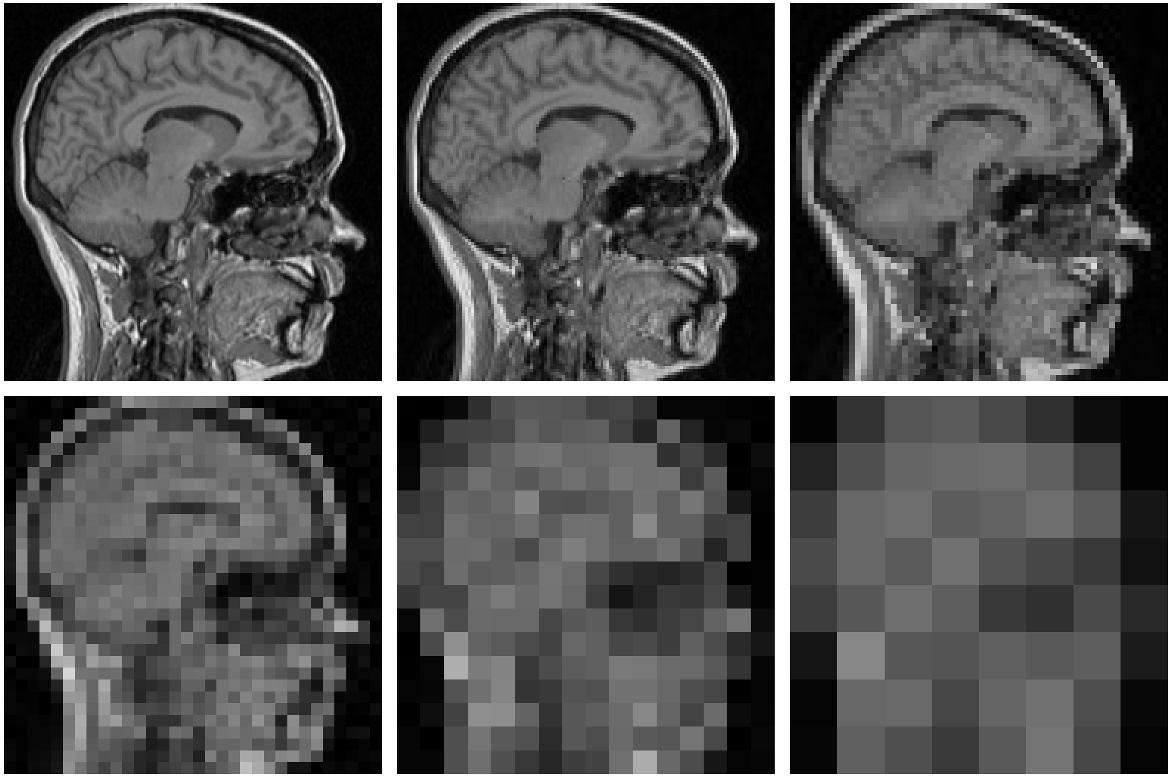


Taizo Iijima

Slide by Arjan Kuijper, 2003

To model (II)

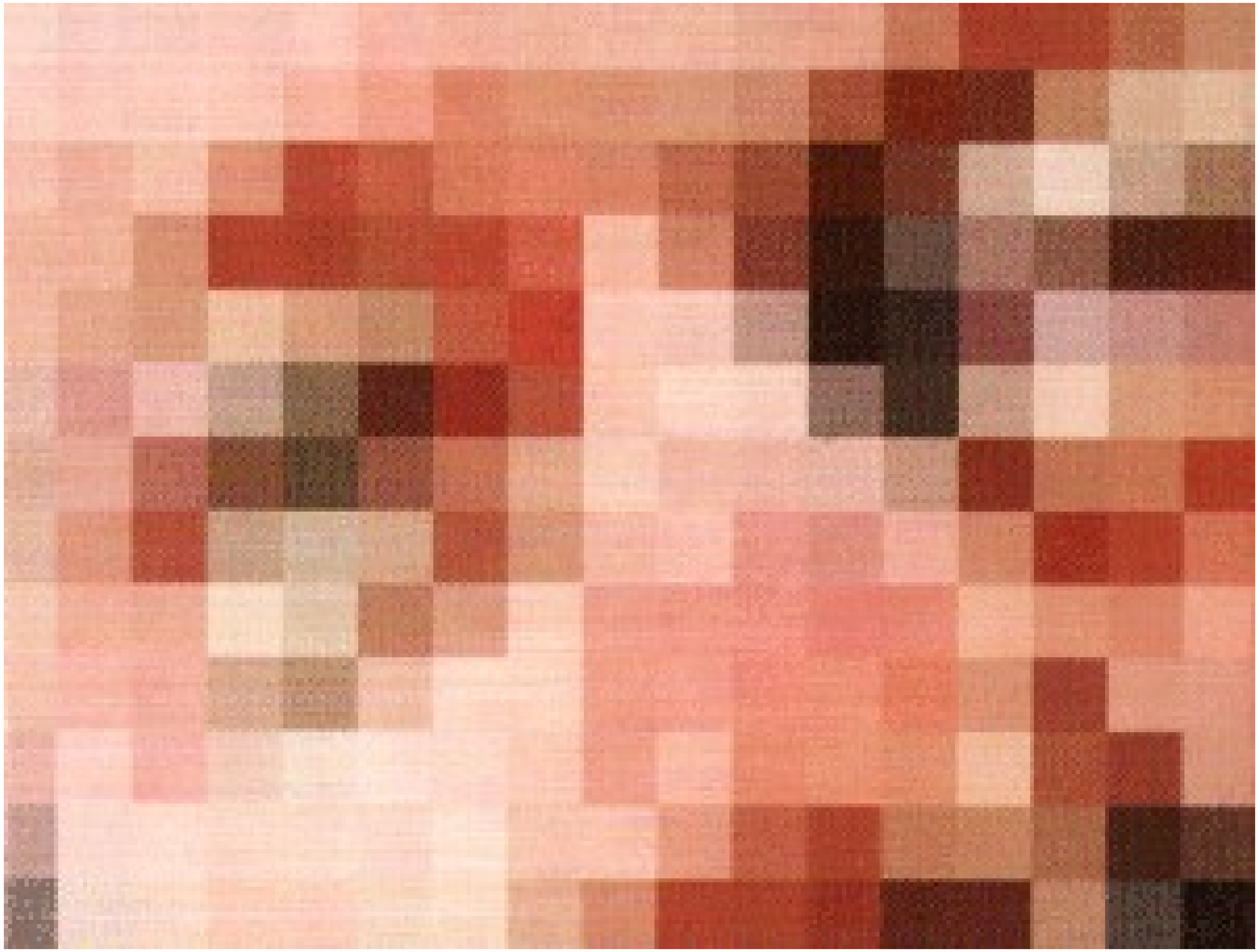
- Don’t trust the resolution.
  - What does a detector of a 3 pixels circular size detect?

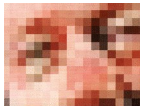


Slide by Arjan Kuijper, 2003

To model (III)

- Don’t trust the grid.





Slide by Arjan Kuijper, 2003

To model a device

- It has finite resolution.
  - Infinite resolution is impossible.
- Take uncommitted observations
  - There is no bias, no knowledge, no memory.
- We know nothing.
  - At least, at the first stage. Refine later on.
- Allow different scales.
  - There’s more than just pixels .
- View them simultaneously.
  - There is no preferred size.
- Noise is part of the measurement.
  -

Deep structure

The challenge is to understand the image
really on all the levels simultaneously,
and not as an unrelated set of derived images
at different levels of blurring.

Jan Koenderink (1984)

Our choice of model: Linear scale space

- The scale space of *I* is a 1-parameter family

$$L(x, y; \sigma) = (I * G)(x, y; \sigma)$$

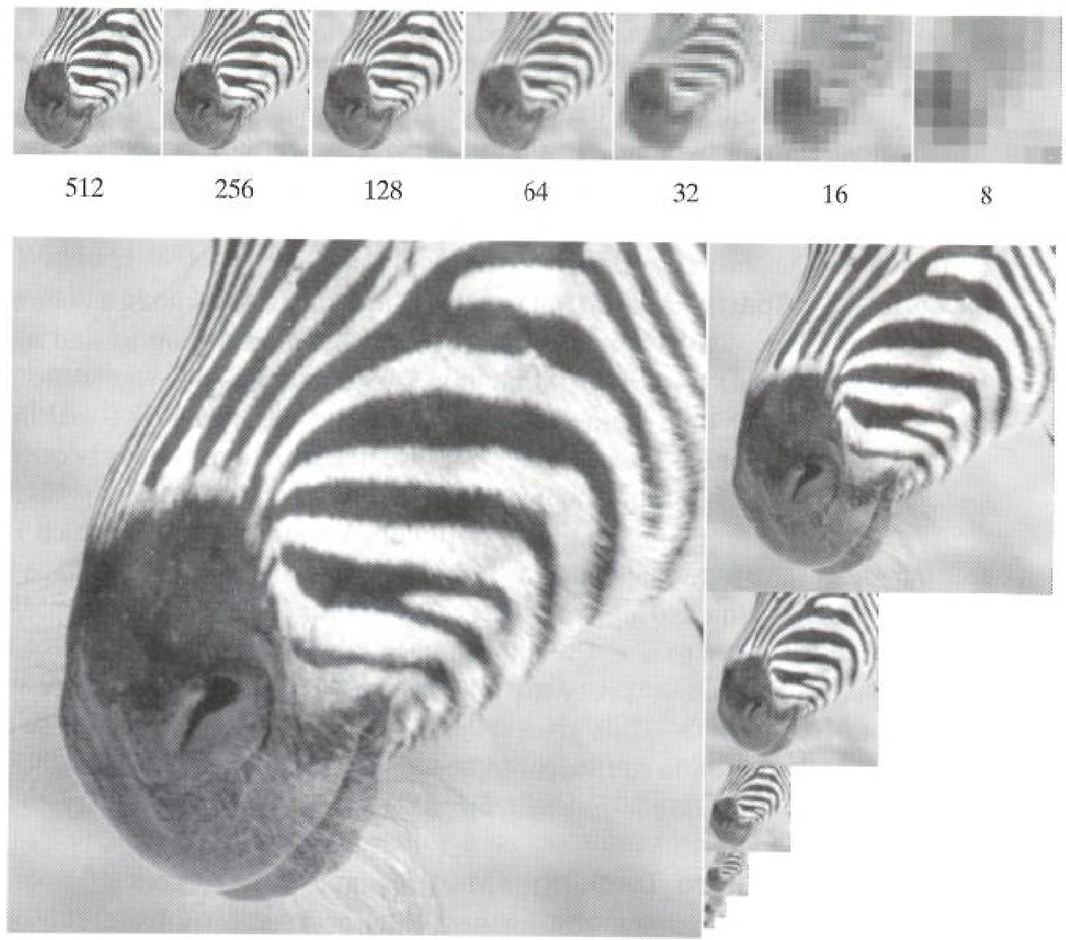
$$L(x, y; \sigma = 0) \simeq I(x, y)$$

$$G(x, y; \sigma) = \frac{{2\pi\sigma^2}^{-\frac{x^2+y^2}{2\sigma^2}}}{e}$$

- Scale is given by σ and is an important parameter in Computer Vision algorithms

- As the scale increases details in the image disappear and we focus on the large scale structures that are left.

Multi-scale versus multi-resolution: The Gaussian pyramid



Multi scale analysis

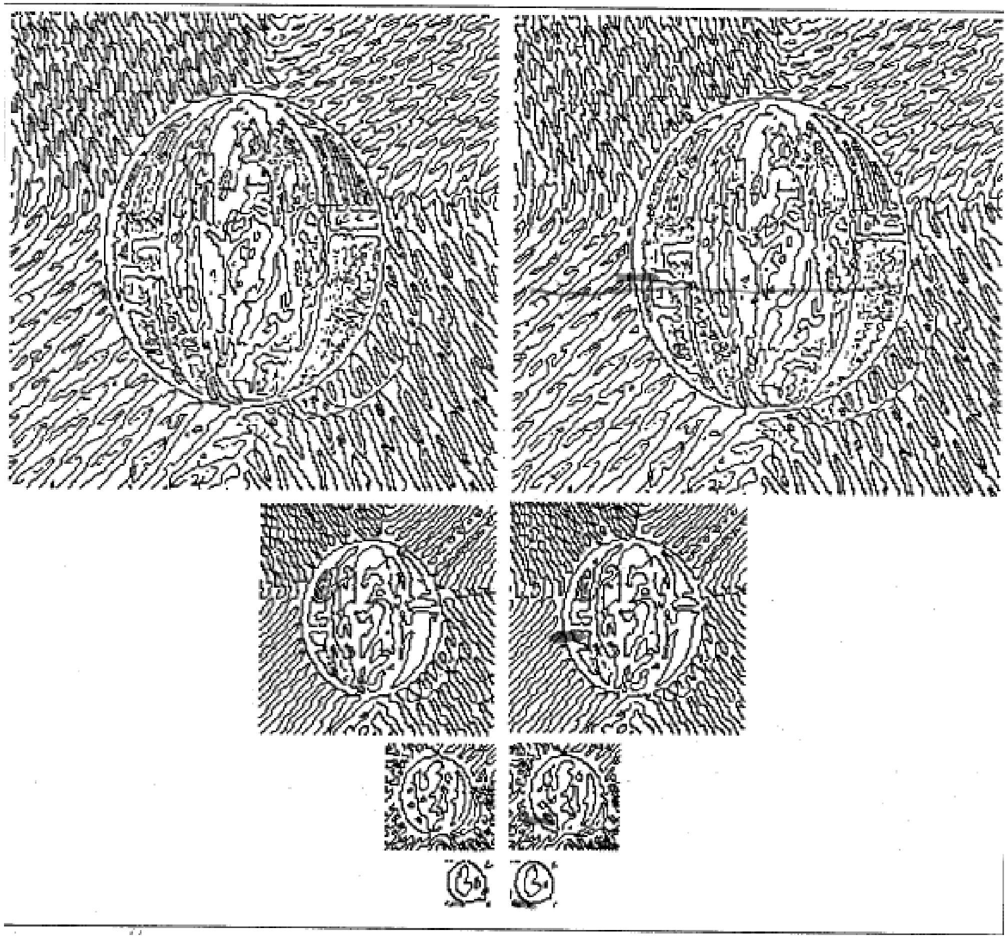
- There is an obvious need because we do not know the size of things.
- We may have to match images obtained from different distances.
- We may easily reduce resolution but we cannot increase
- Multi scale analysis may be used to speed up using a coarse to fine approach

Coarse to fine analysis

1. Instead of solving the original problem reduce resolution and solve at coarse scale.
2. Use coarse solution to constrain the solution at finer scales.
3. Iterate until problem solved at full resolution.

- **Application within**: Image matching, Flow computation/video coding, Adaptive thresholding or image normalization, Image registration etc. etc.

Example:
Coarse-to-fine edge-based stereo analysis



Multi scale feature detectors

- All the mentioned interest point detectors may be extended to operate on all scale levels.
- Multi scale detectors must compare the detection strength at neighboring scale levels and suppress the non-maximal.
- The detection strength evaluation must be scale invariant. This is a non-trivial matter (see paper by Tony Lindeberg in the supplementary material).

Introduction to descriptors

- We have to attribute each feature point with a descriptor in order to compare if the a candidate match is good or not.

- Descriptors have to be local, informative, insensitive to noise, luminance variations, perspective deformation including rotation, to pixel quantization etc.

- We like compact (low dimensional) descriptors but recognize that high dimensional ones may be more discriminative.

Descriptors

- Most modern descriptors build on
  1. Statistics / histogramming
  2. Spatial sampling of the gradients is some neighborhood to the interest point.

- Some descriptors include color, e.g. color histograms (but we skip this here)

- Few descriptors use raw image data as used in correlation

- Next lecture will dwell with descriptor construction and matching

Literature

Reading material:

- Forsyth and Ponce: Ch. 7.1 – 7.2, 7.7 (Filtering basics)
- Lowe IJCV 2004, Sec. 1 – 3 (DoG, Laplacian)
- Schmid, Mohr and Bauckhage IJCV 2000, Sec. 1 – 2 (Harris corner)
- Lindeberg 1996 (Scale-space theory)

Additional material:

- C. Harris and M. Stephens: A Combined Corner and Edge Detector. 4th Alvey Vision Conference, 147—151, 1988.