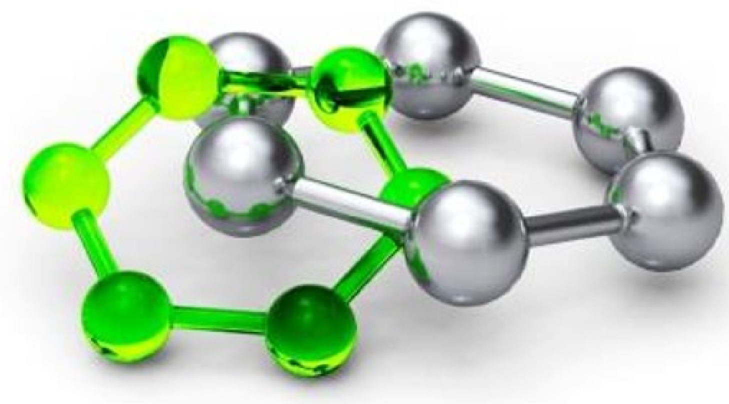
CEITEC

Central European Institute of Technology
BRNO | CZECH REPUBLIC

Image analysis II

C9940 3-Dimensional Transmission Electron Microscopy
S1007 Doing structural biology with the electron microscope

April 4, 2016



EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND
INVESTING IN YOUR FUTURE

2007-13
OP Research and
Development for Innovation

Outline

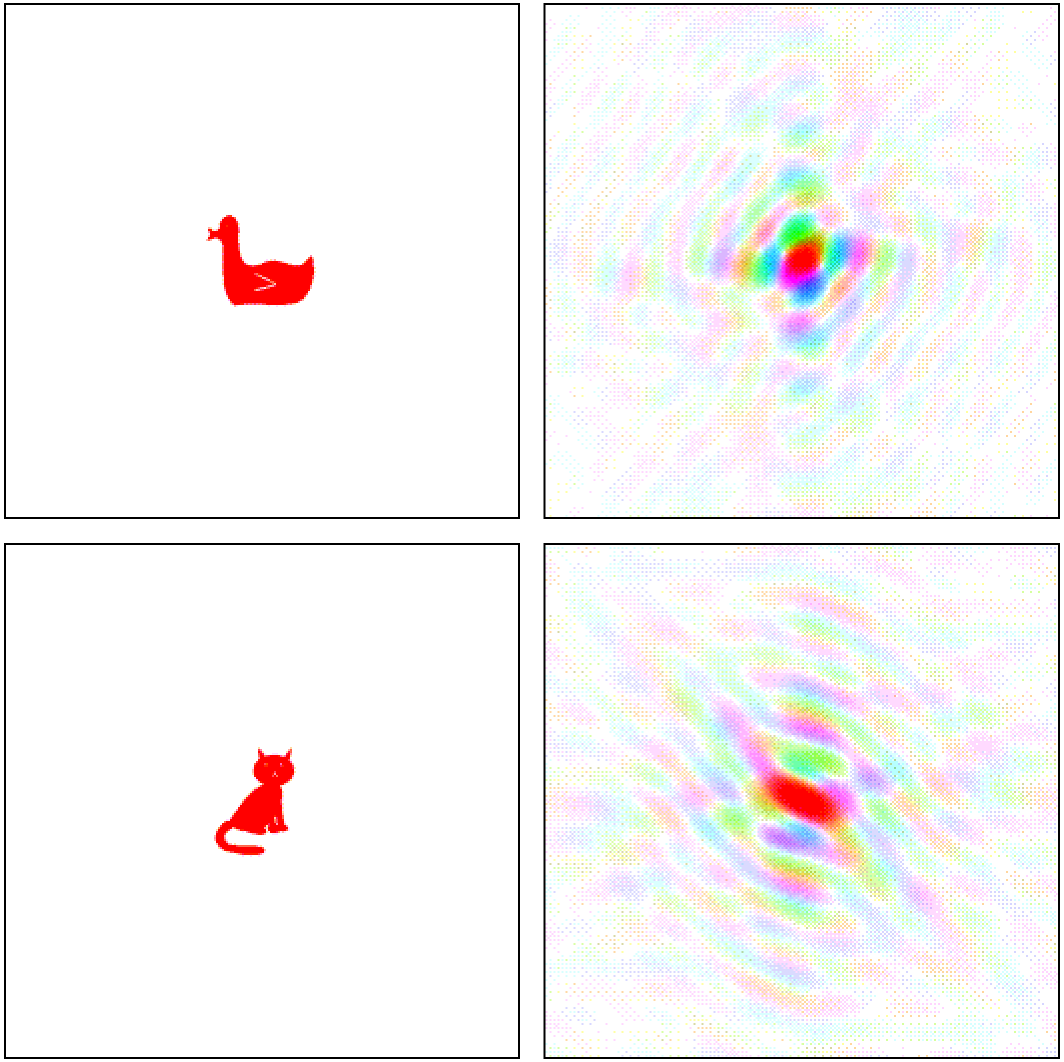
Image analysis II

- Fourier transforms revisited
  - Ducks and other animals
- Analogy to the Ewald sphere
- Aliasing
- Alignment
- Interpolation
- Multivariate data analysis

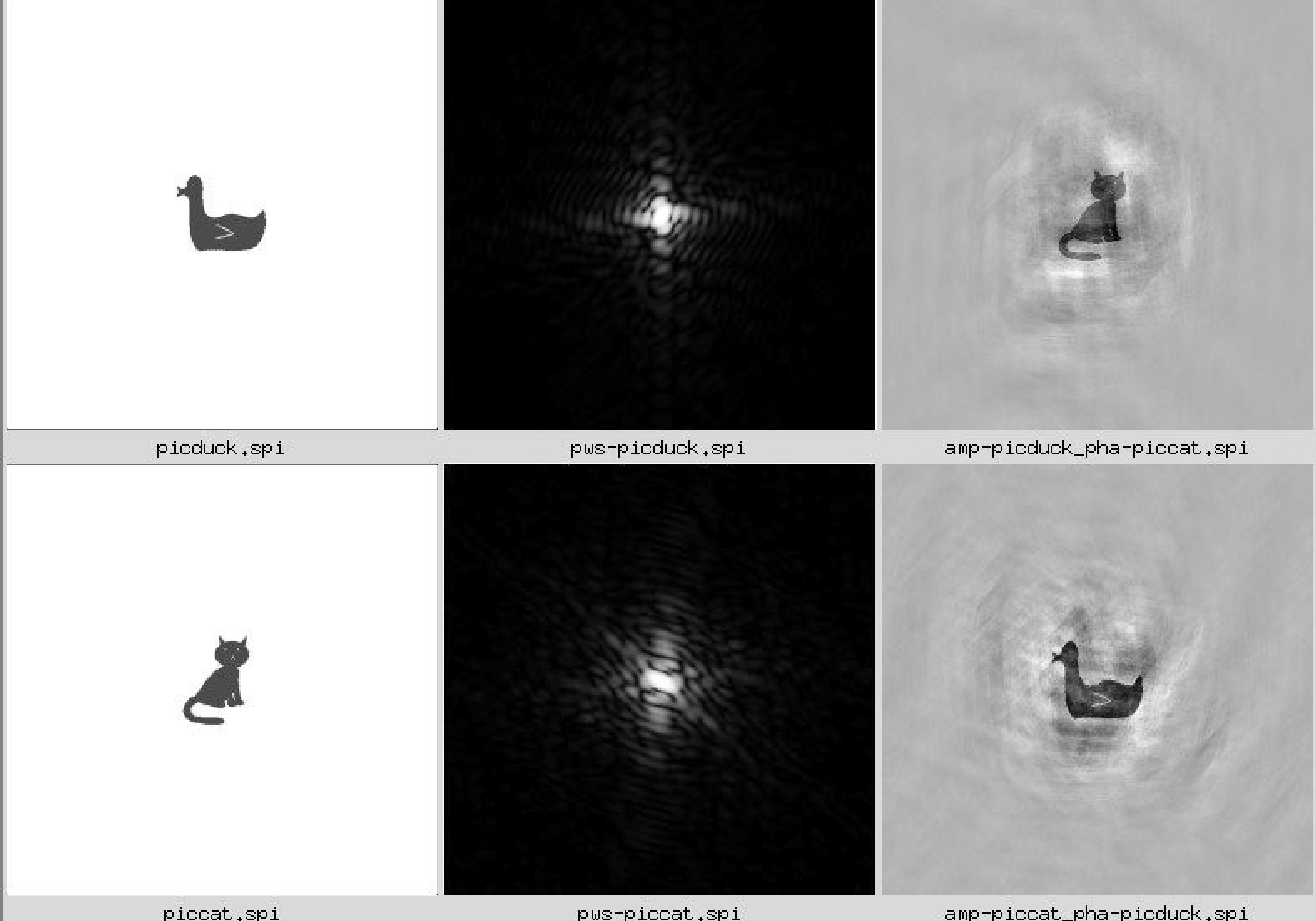
Outline

Image analysis II

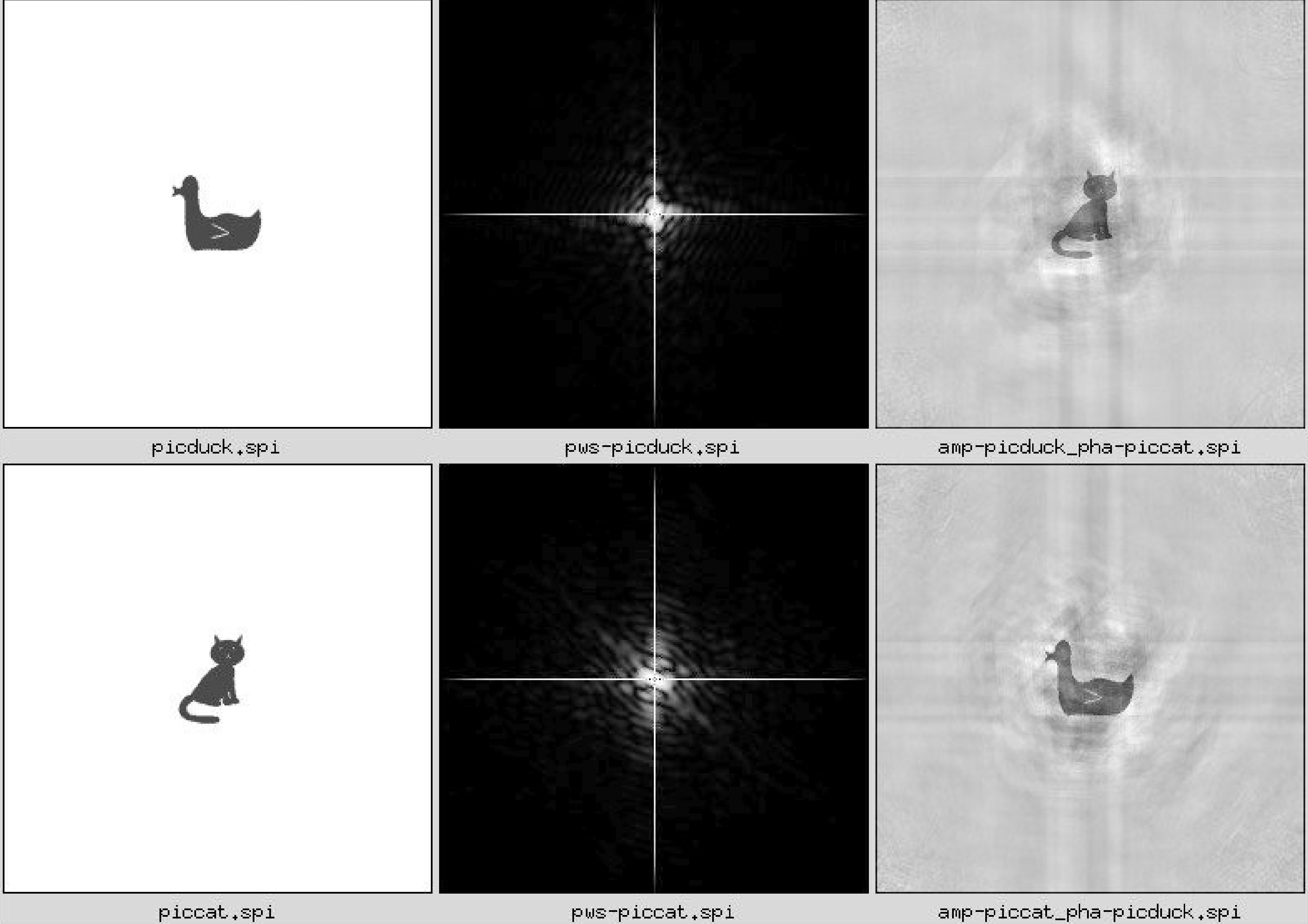
- Fourier transforms revisited
  - Ducks and other animals
- Analogy to the Ewald sphere
- Aliasing
- Alignment
- Interpolation
- Multivariate data analysis



http://www.ysbl.york.ac.uk/~cowtan/fourier/magic.html

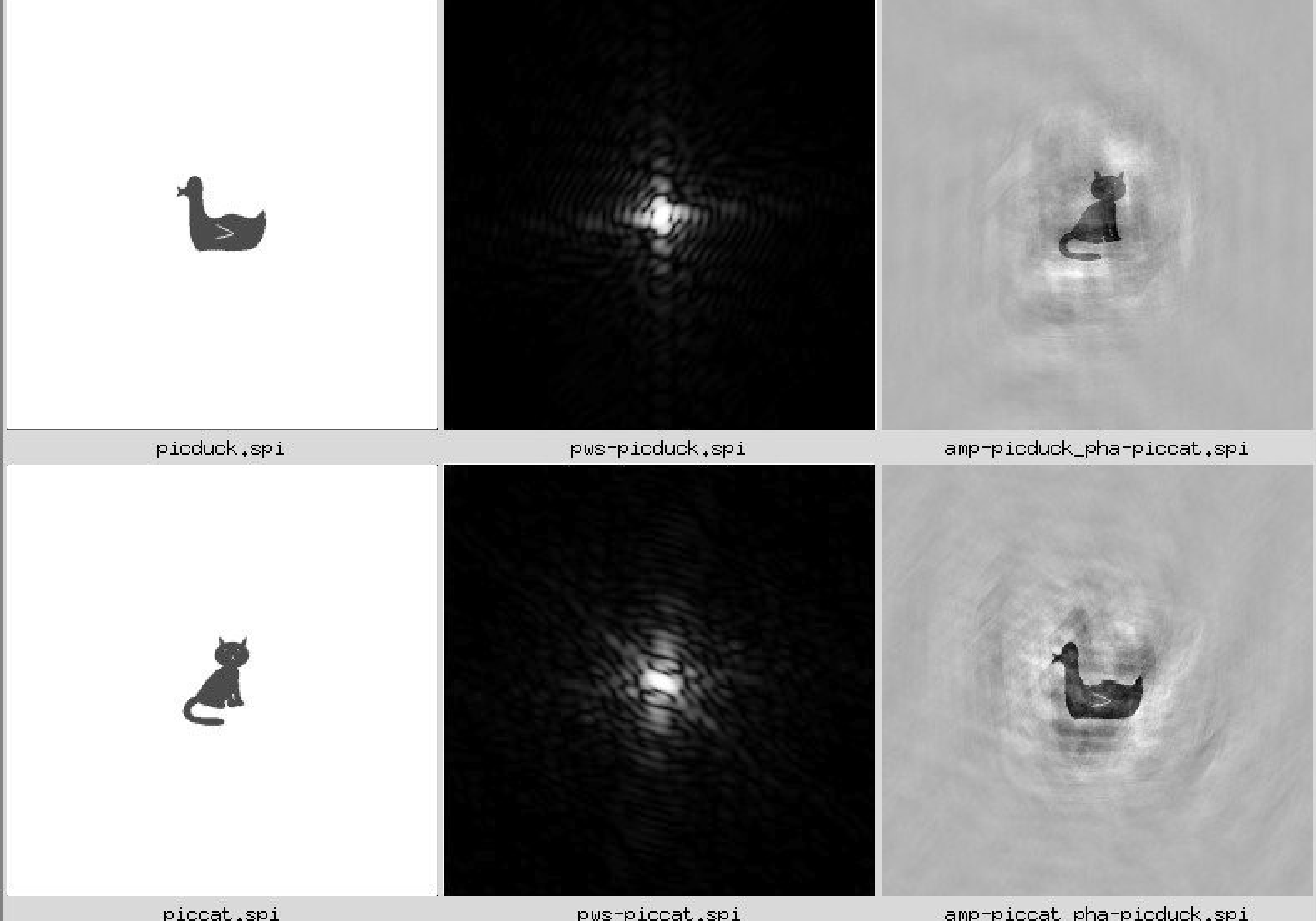

picduck.spi
pws-picduck.spi
amp-picduck_pha-piccat.spi
piccat.spi
pws-piccat.spi
amp-piccat_pha-picduck.spi

A quiz:
A problem I had with my first attempt



QUESTION: Where did that cross come from?

CEITEC



Apodization: Smoothly bring the signal to zero near an edge

CEITEC

Outline

Image analysis II

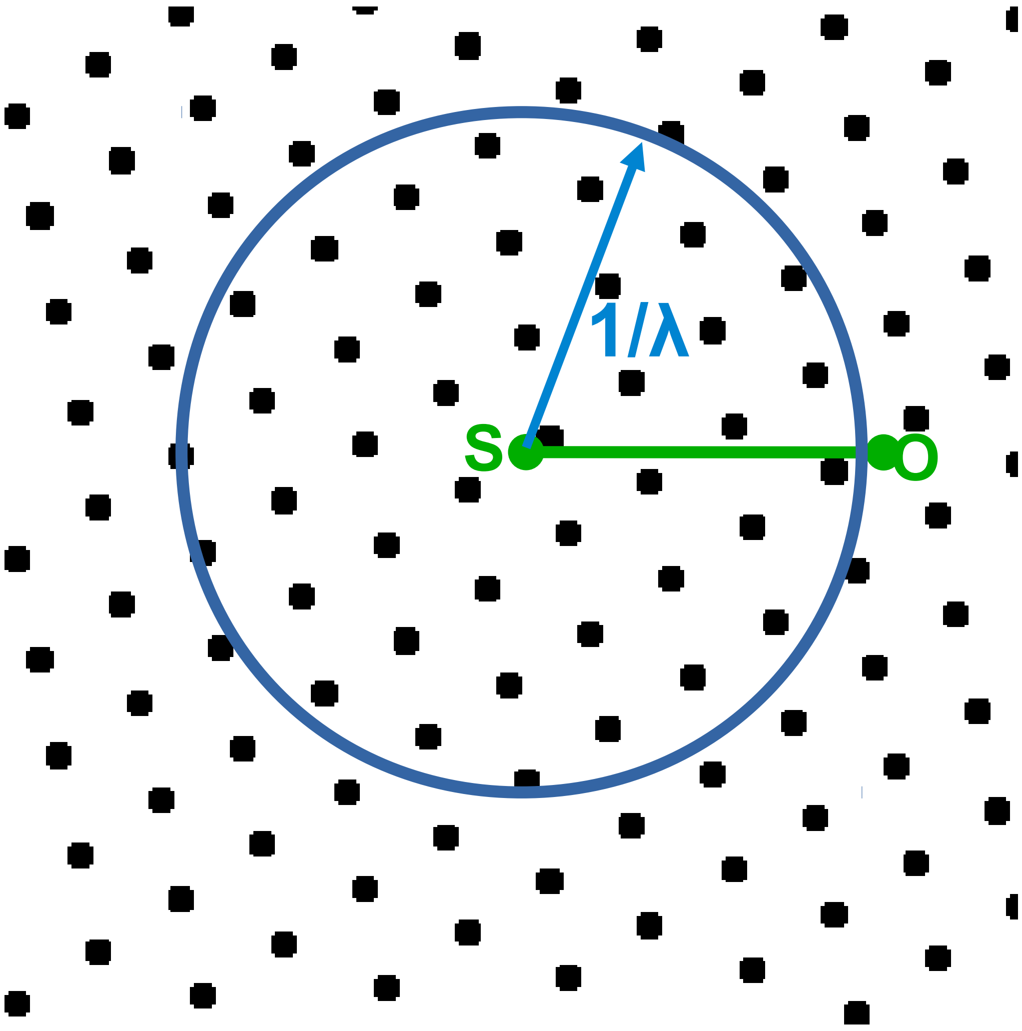
- Fourier transforms revisited
  - Ducks and other animals
- Analogy to the Ewald sphere
- Aliasing
- Alignment
- Interpolation
- Multivariate data analysis

Why are electrons useful?

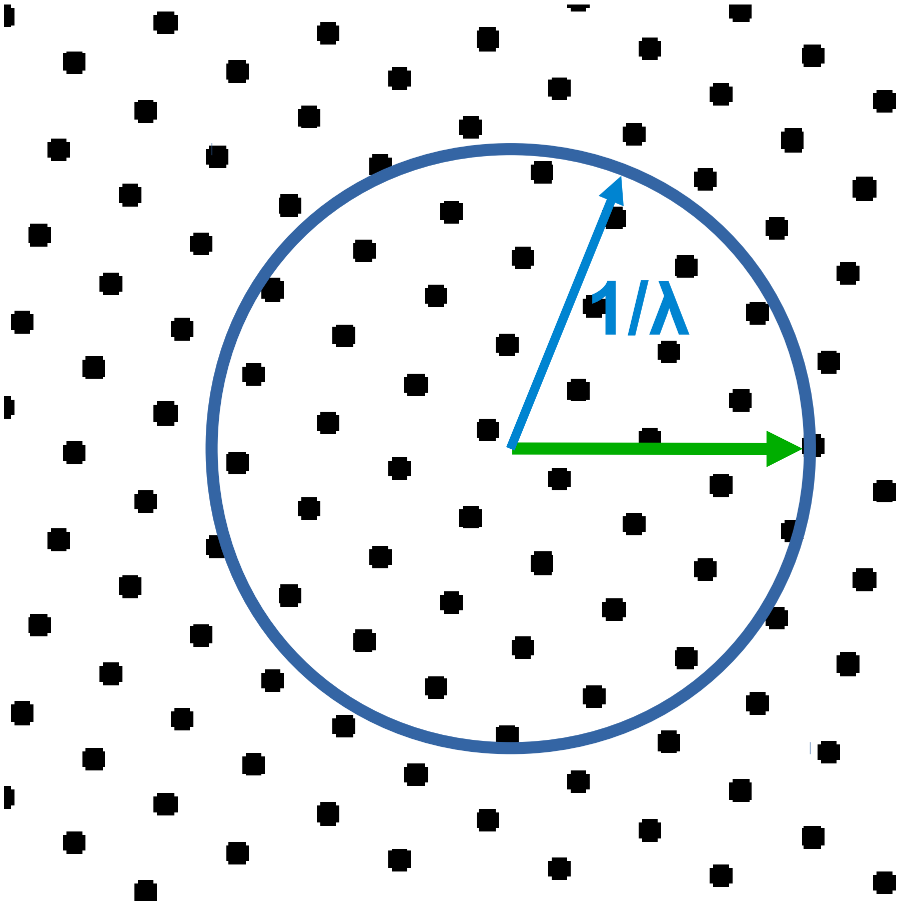
Wavelengths of various radiation types

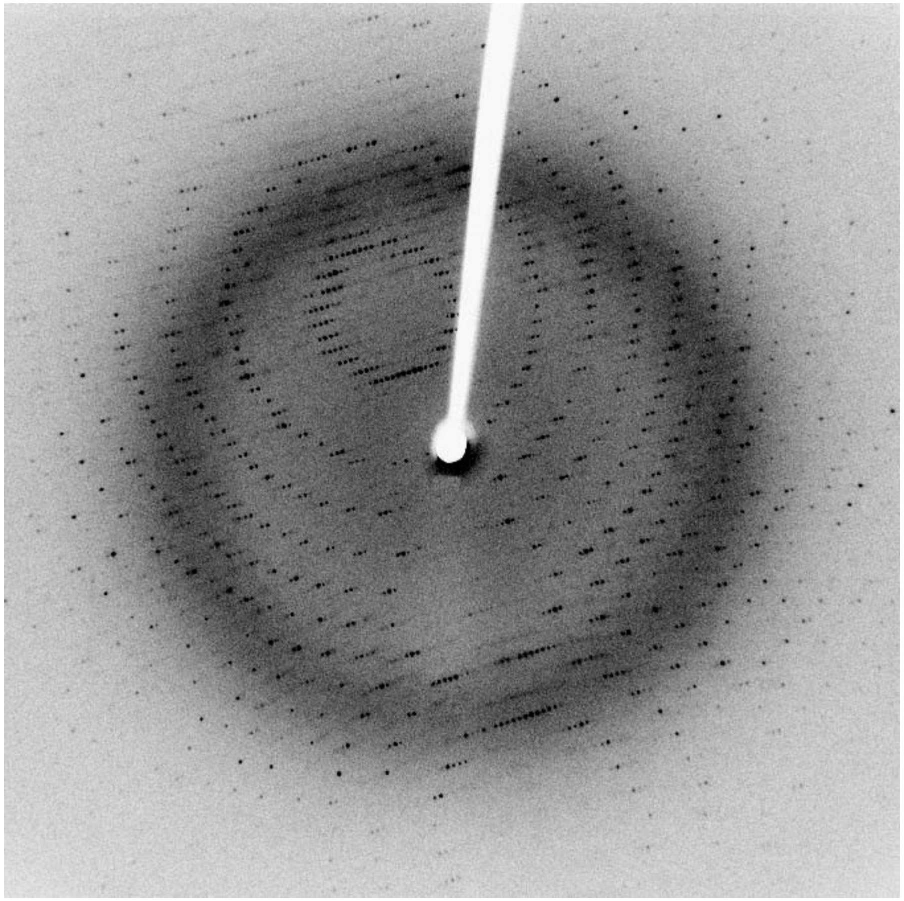
- Visible light: >380nm
- X-rays (copper K): 0.154nm (1.54Å)
- Electrons (300kV): 0.002nm (0.02Å)

How wavelength limits resolution

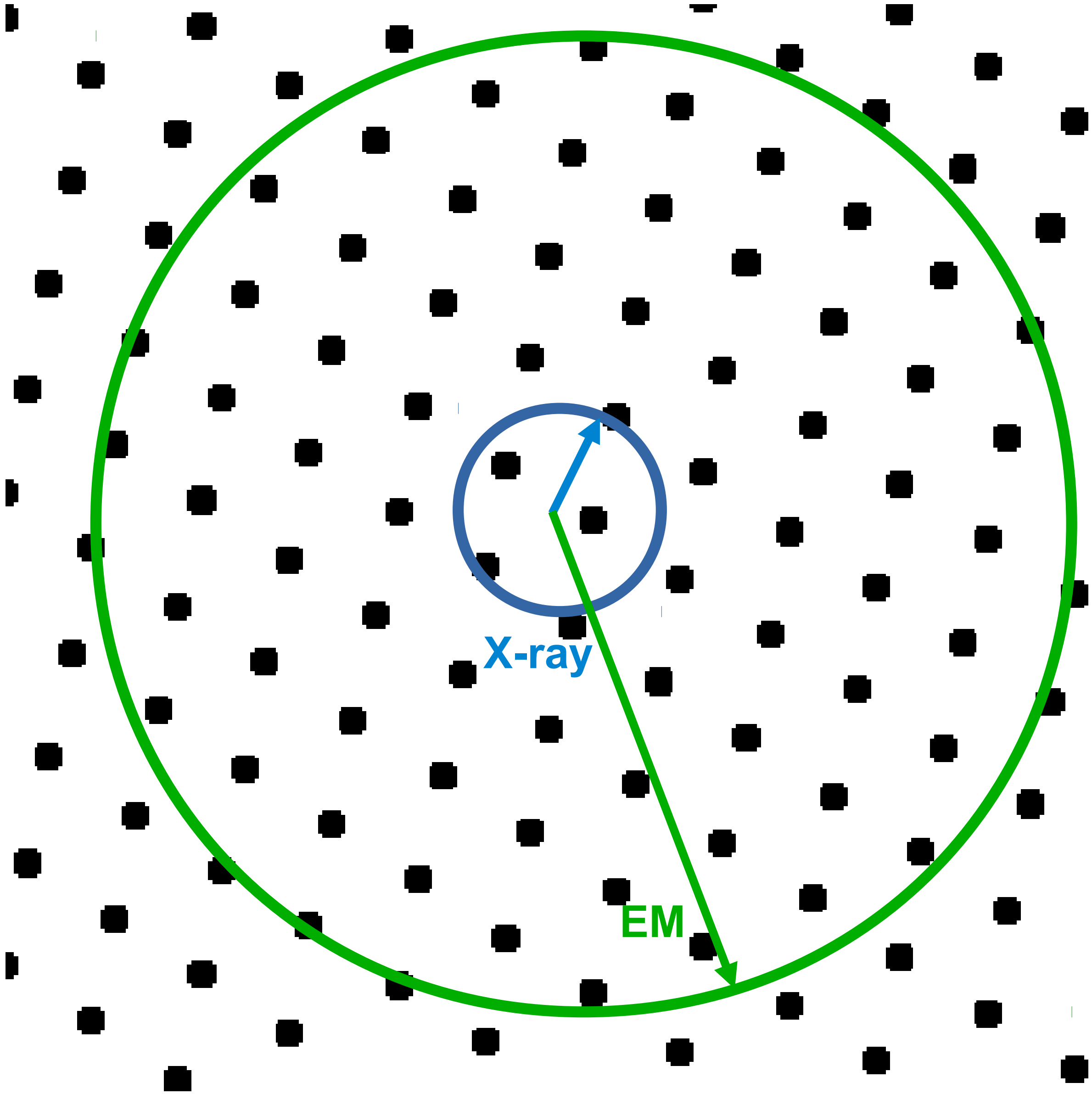


S: specimen origin
O: diffraction origin

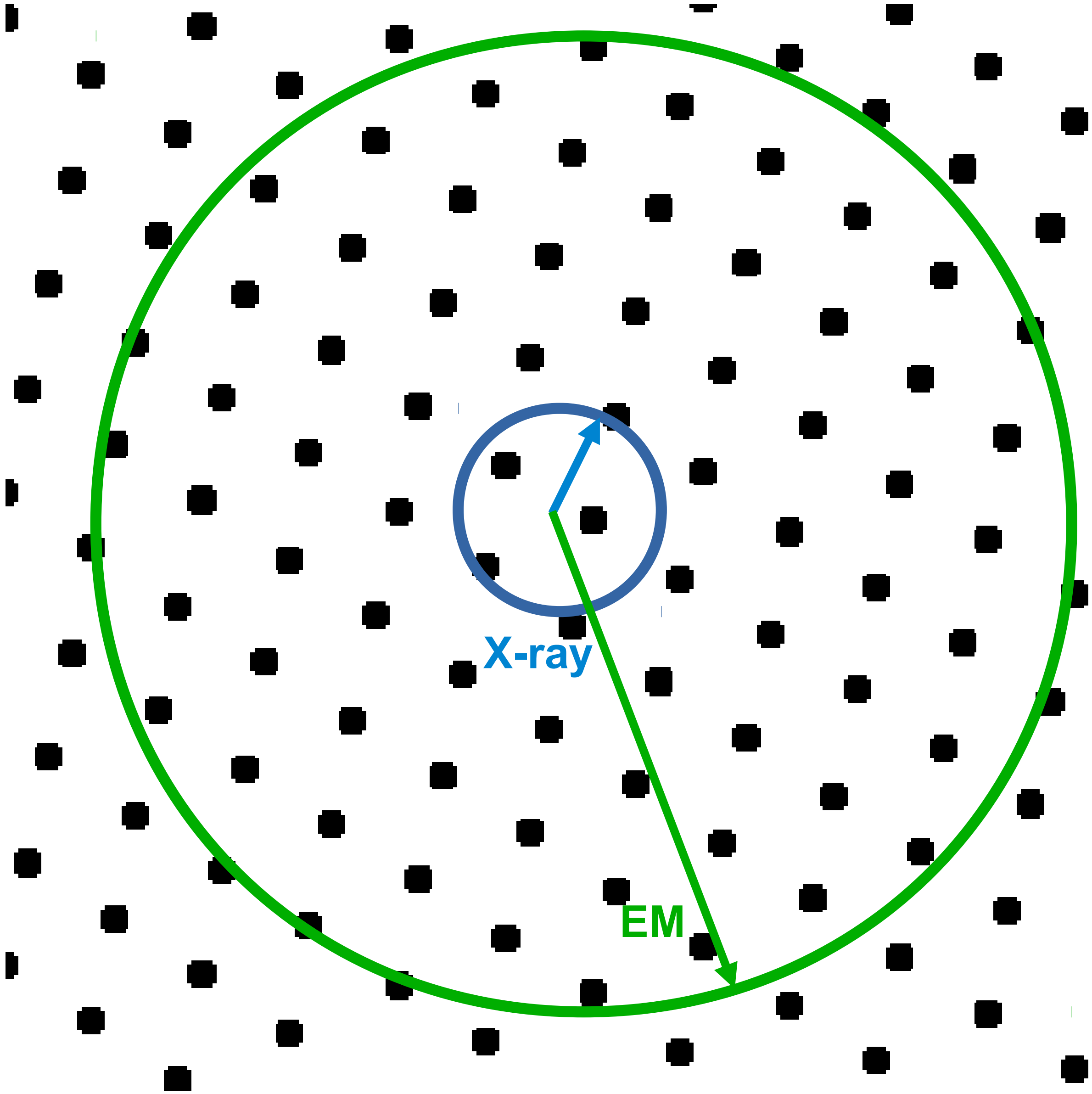




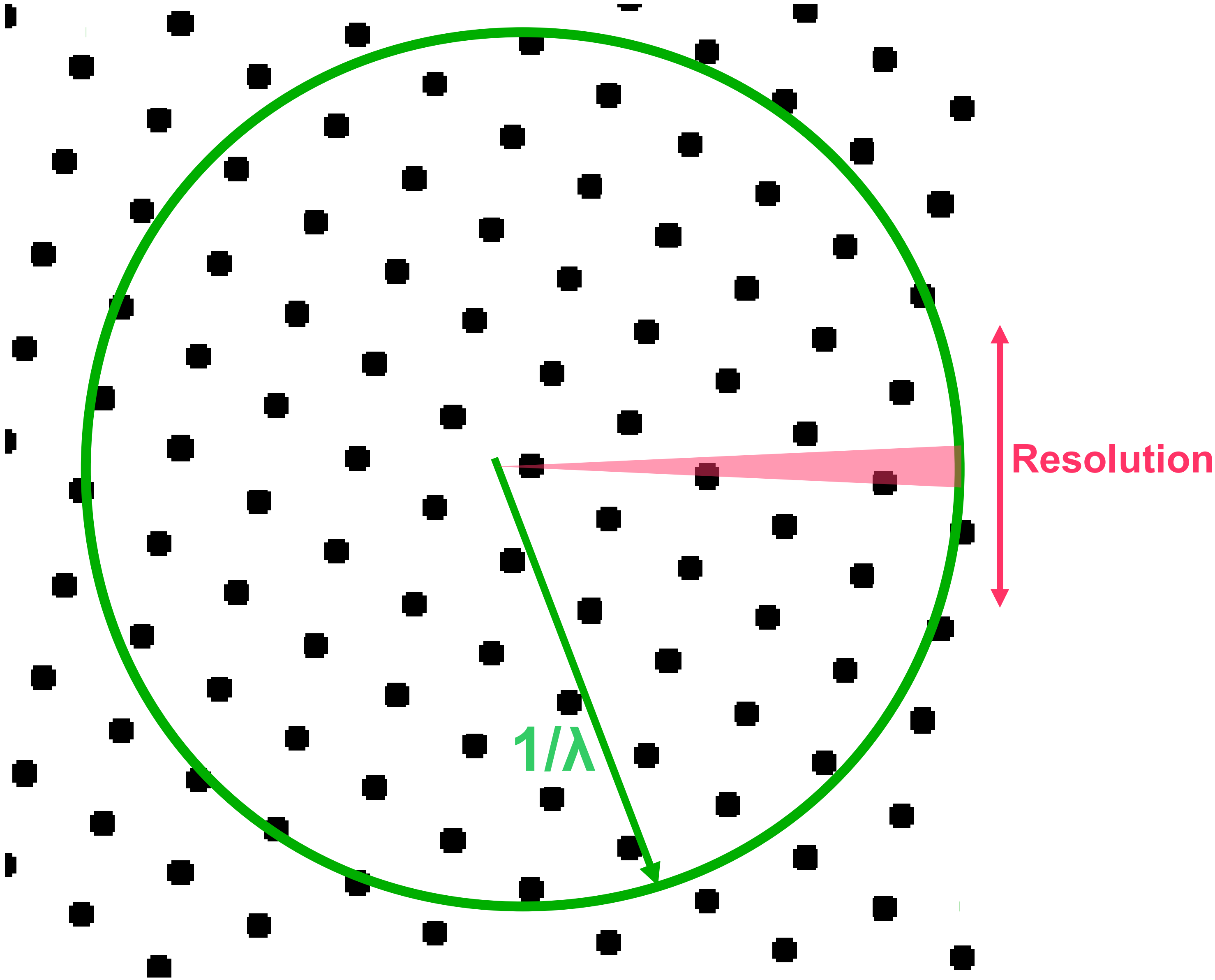
http://en.wikipedia.org



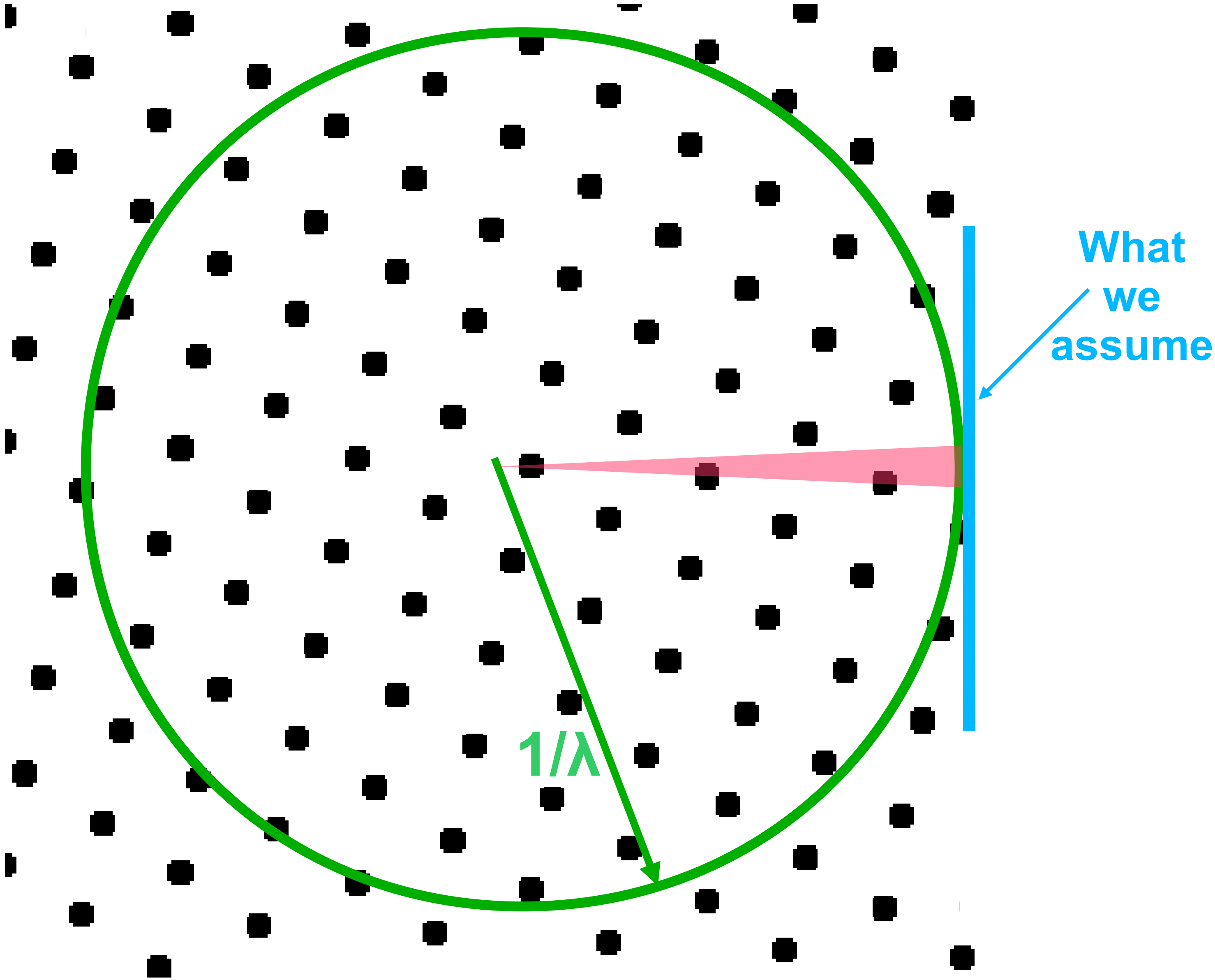
NOTE: Not to scale.
EM wavelength is ~80 smaller,
and therefore 1/λ would be 80X bigger.



NOTE 2: For practical purposes,
the radius of the Ewald sphere is so large that we ignore its curvature.

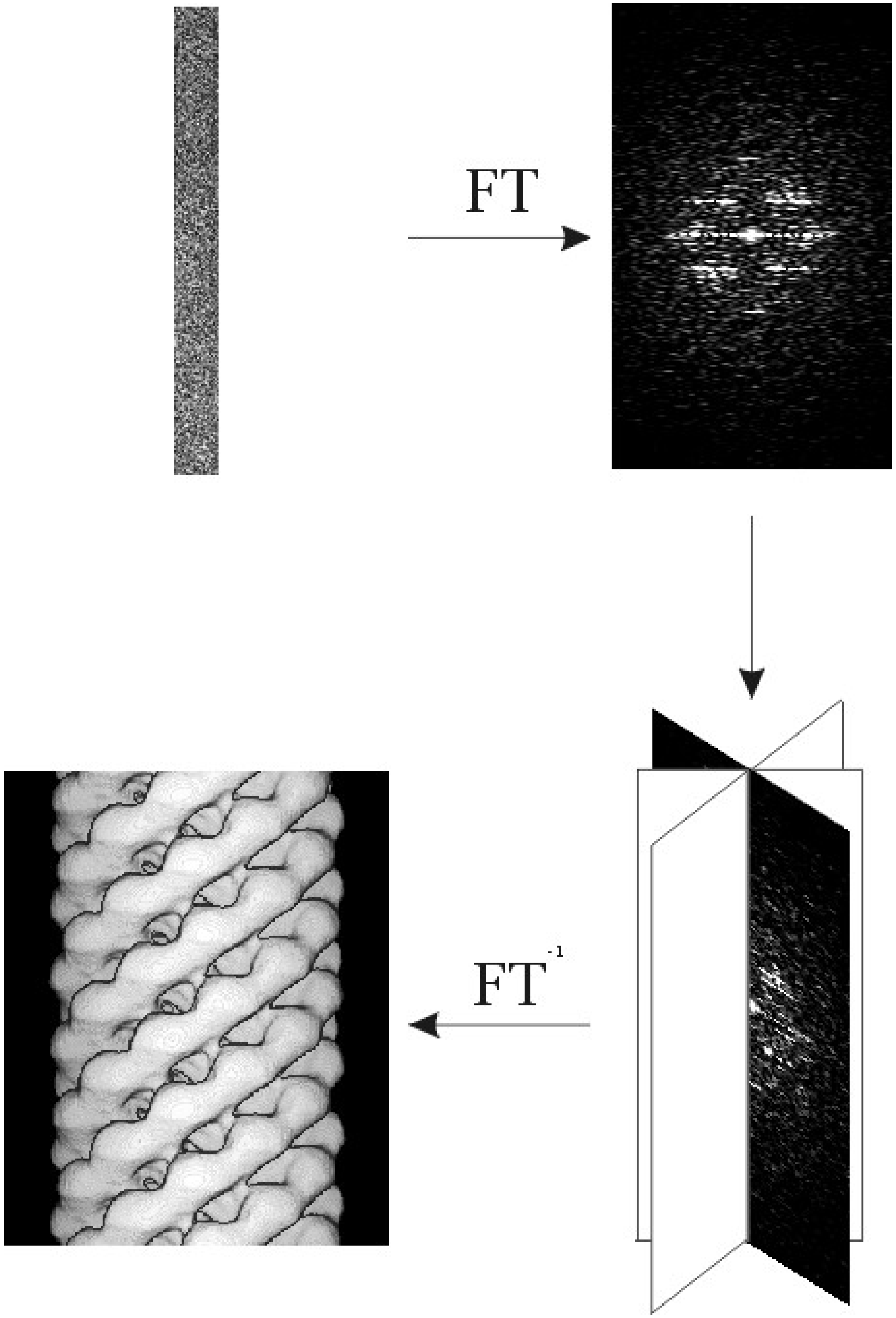


NOTE 3: Electron lenses are terrible, and biological samples are fragile, so in practice we'll see on a tiny fraction of the data we could theoretically get.



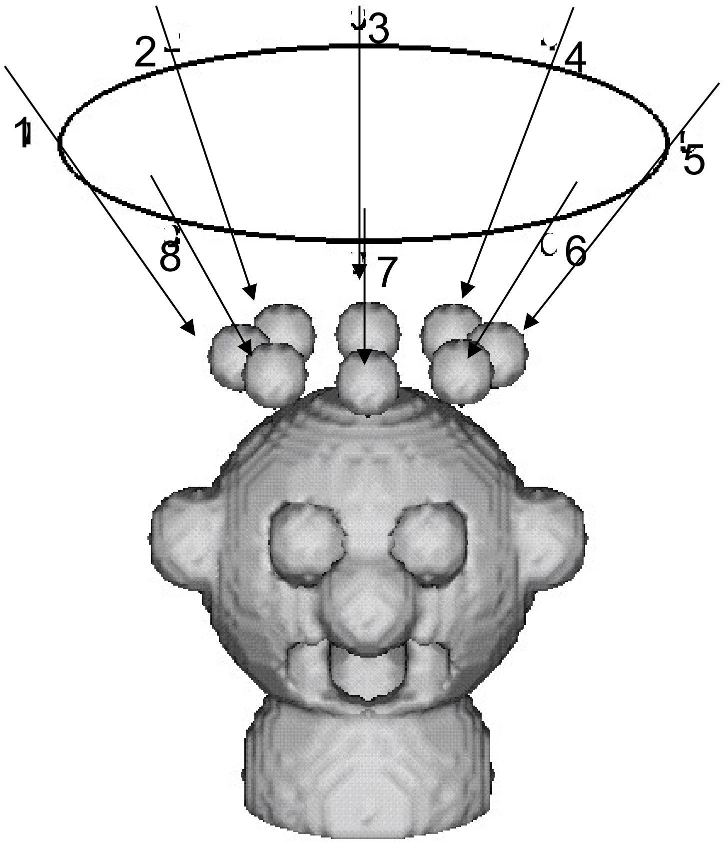
NOTE 4: For more information, see
DeRosier (2000) “Correction of high-resolution data for curvature of the Ewald sphere.”

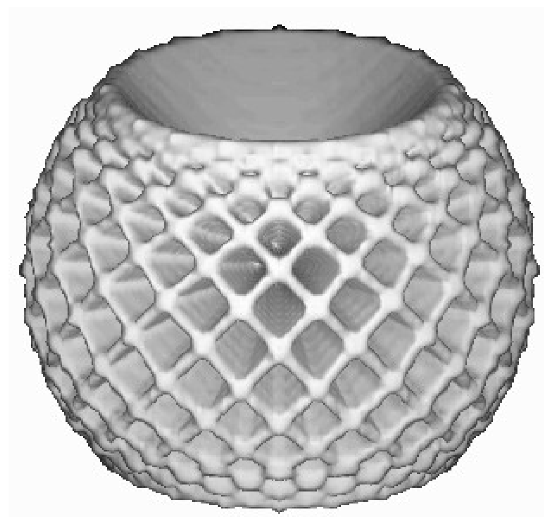
Preview: 3D reconstruction

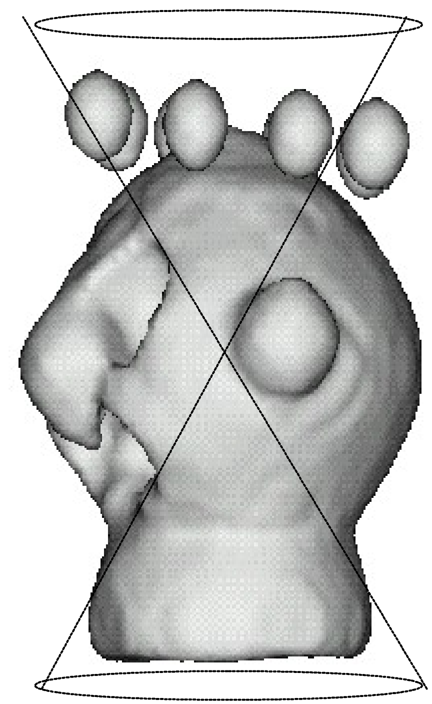


CEITEC

Preview: reconstruction





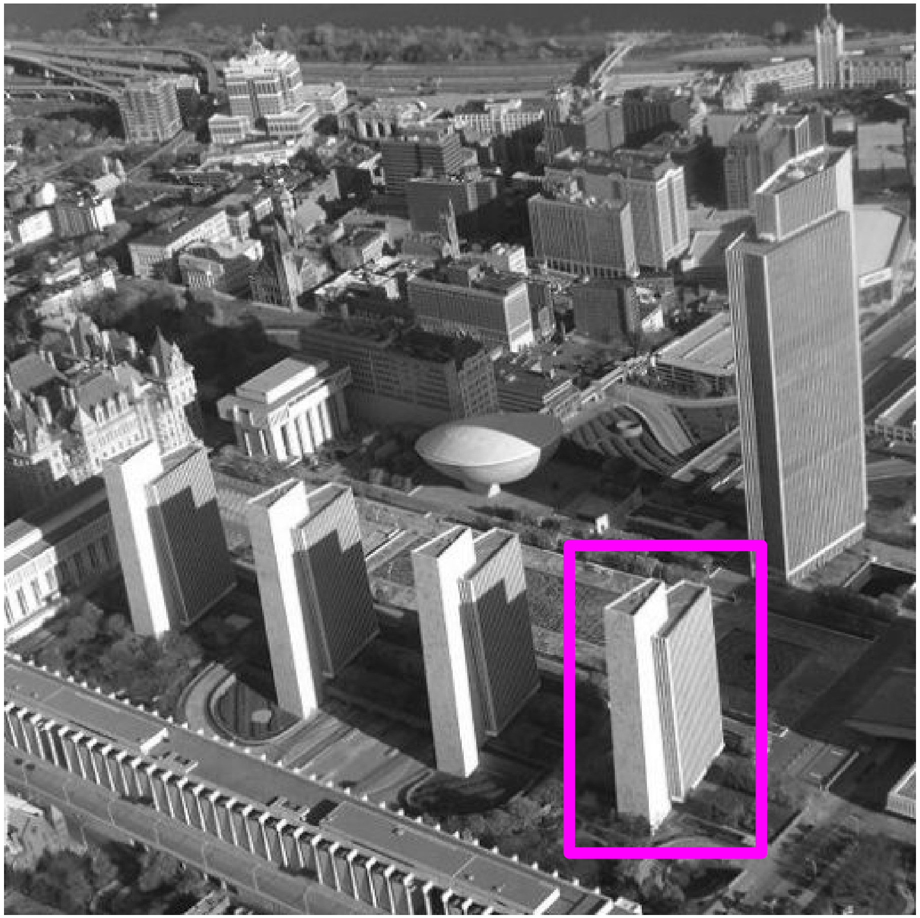


Outline

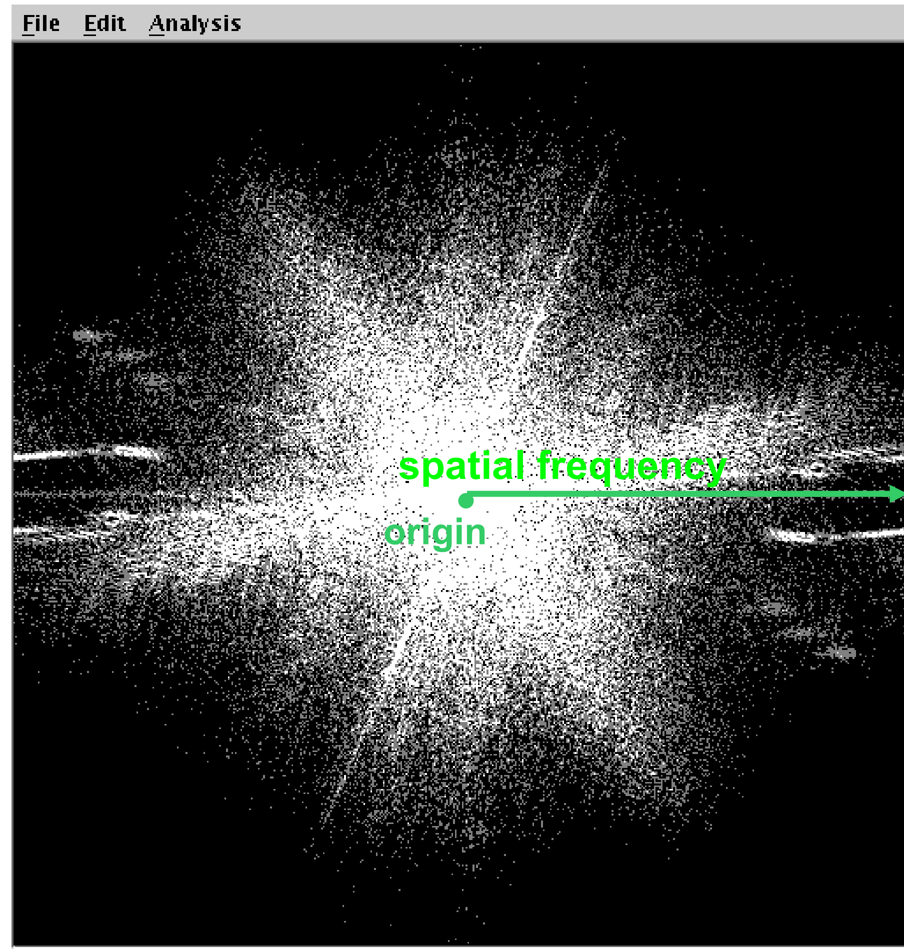
Image analysis II

- Fourier transforms revisited
  - Ducks and other animals
- Analogy to the Ewald sphere
- Aliasing
- Alignment
- Interpolation
- Multivariate data analysis

From last week...



From Wikipedia



This is an example of aliasing.

An example using SPIDER

CEITEC

https://youtu.be/6LzaPARy3uA?t=51

Outline

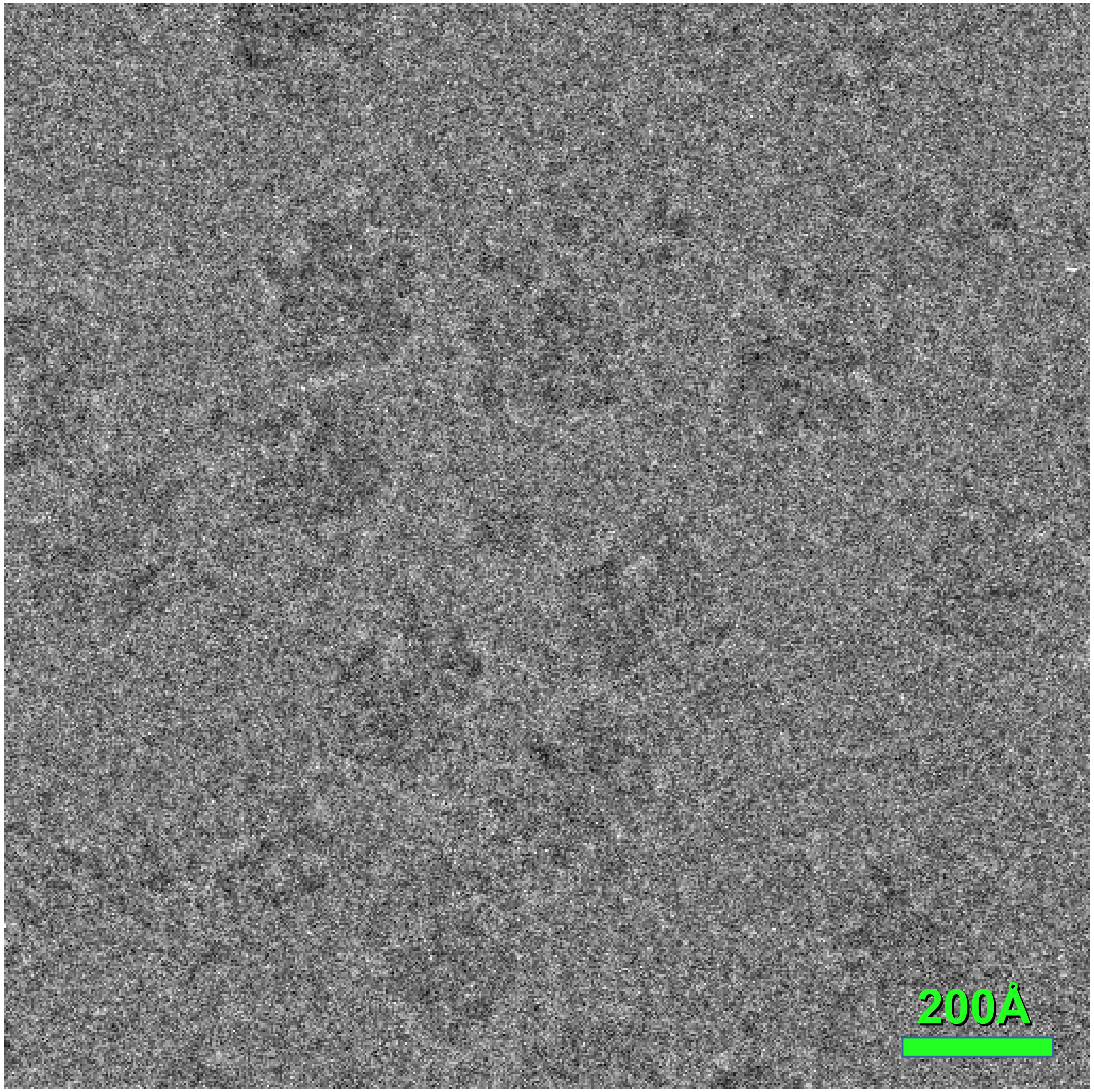
Image analysis II

- Fourier transforms revisited
  - Ducks and other animals
- Analogy to the Ewald sphere
- Aliasing
- Alignment
- Interpolation
- Multivariate data analysis

QUESTION:

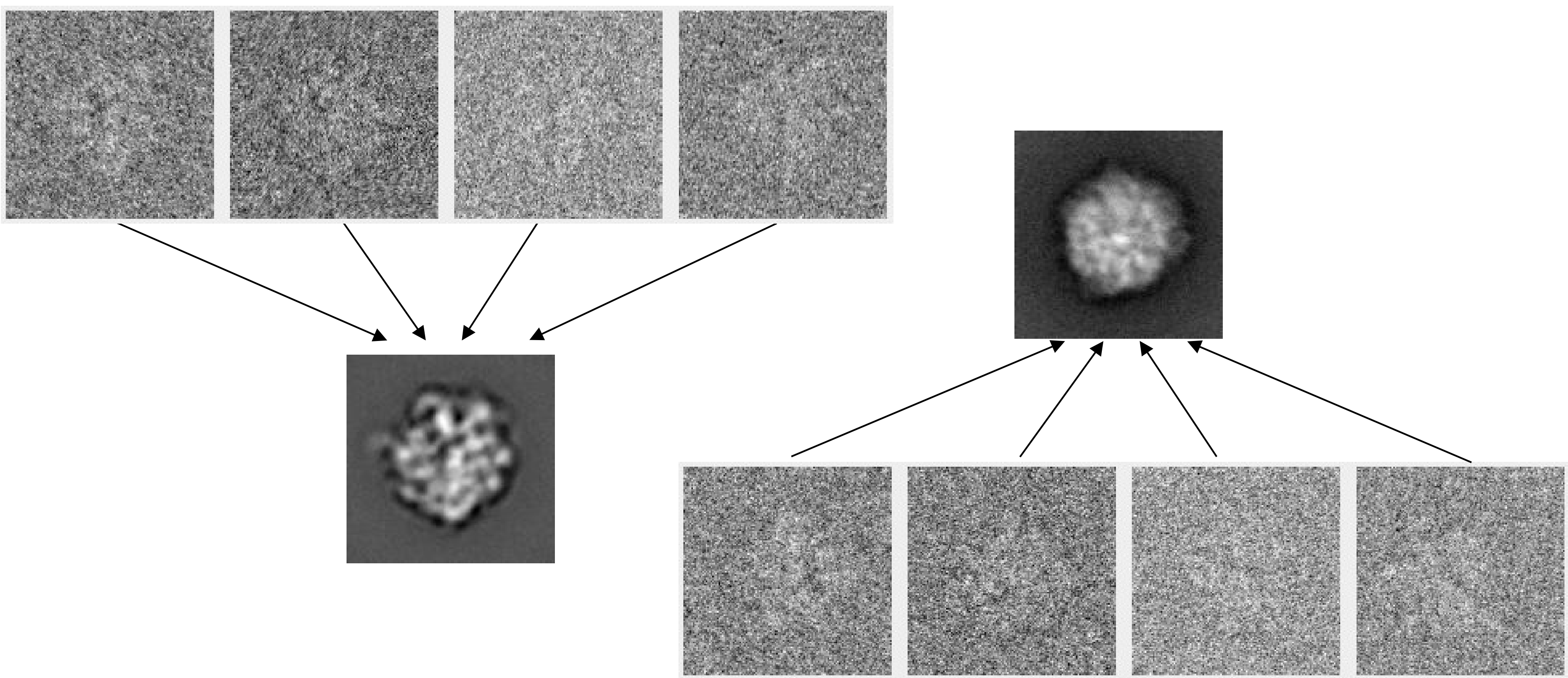
Why do we need to average the signal from many images?

ANSWER: Our signal-to-noise is poor



What happens if we don't align our images?

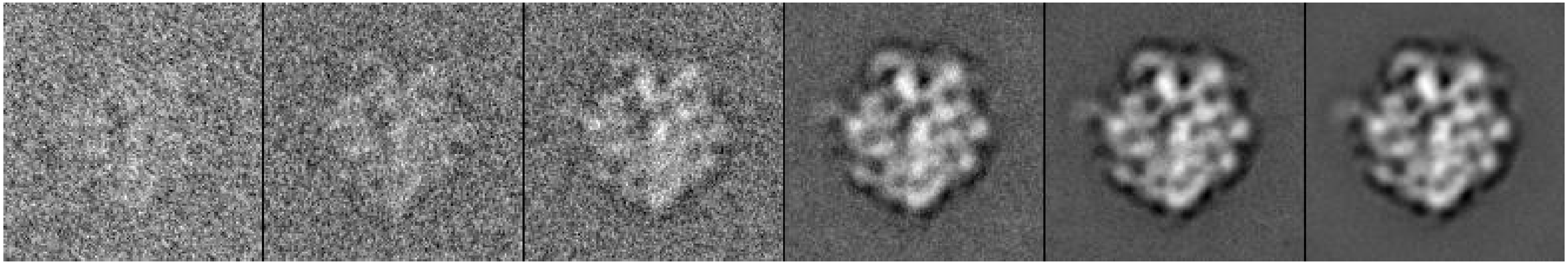
aligned images 1-4 of 4096 total



unaligned images 1-4 of 4096 total

This is a simple 2D case, but the effects are analogous in 3D.

What happens as we include more particles?



n=1 *n*=4 *n*=16 *n*=256 *n*=1024 *n*=4096

Signal-to-noise ratio increases with $\sqrt{n}$

(P)review of 3D reconstruction: The parameters required

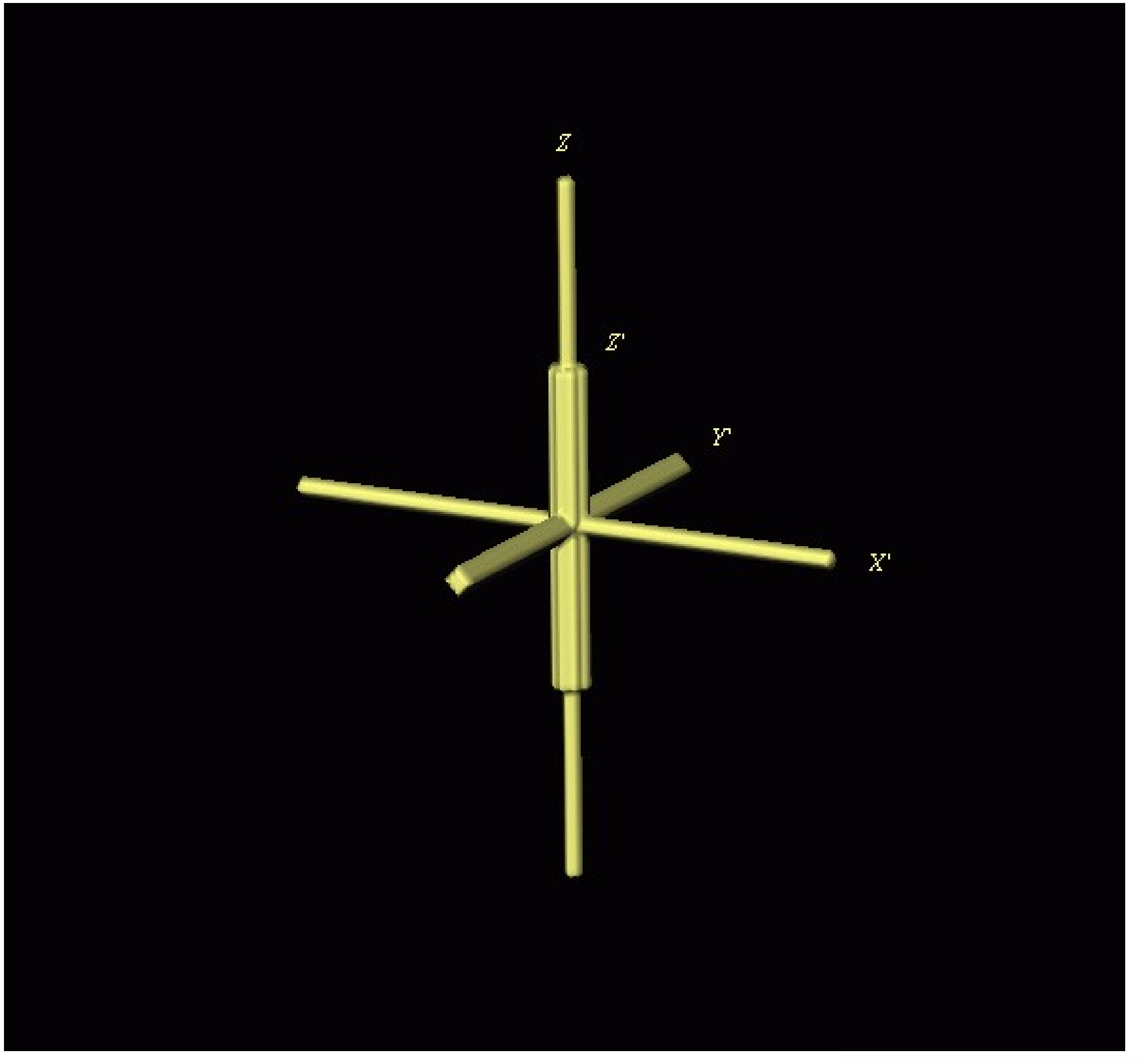
Two translational:

- *Δx*
- *Δy*

Three orientational (Euler angles):

- phi (about *z* axis)
- theta (about *y*)
- psi (about new *z*)

These are determined in 2D. We'll concentrate on these 1st.



http://www.wadsworth.org

CEITEC

How do find the relative translations between two images?

Cross-correlation

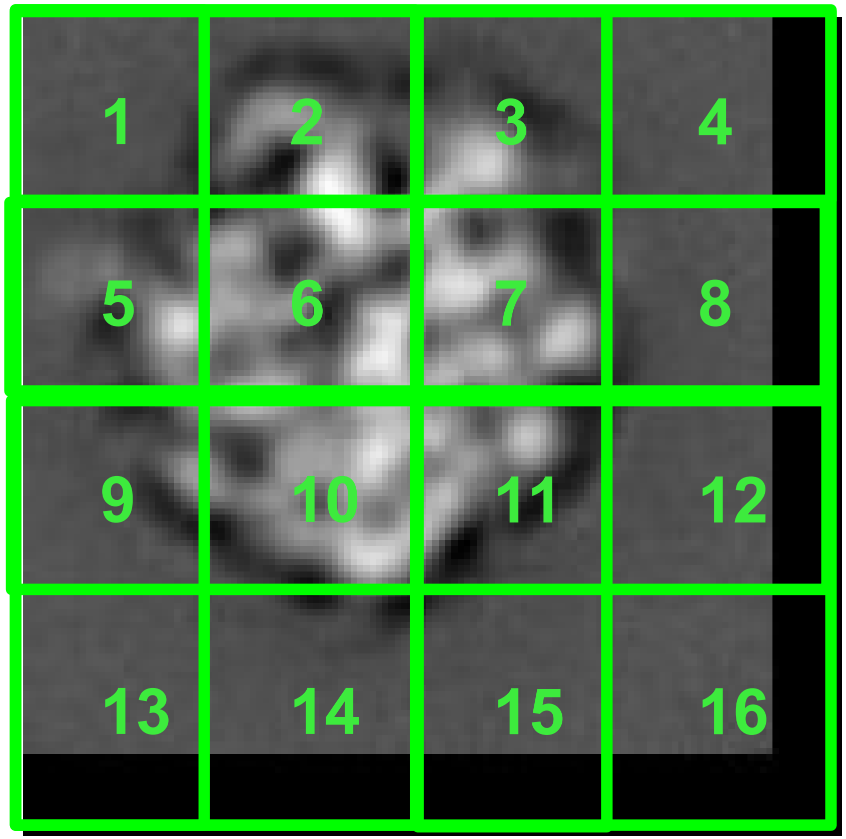


Image *f*

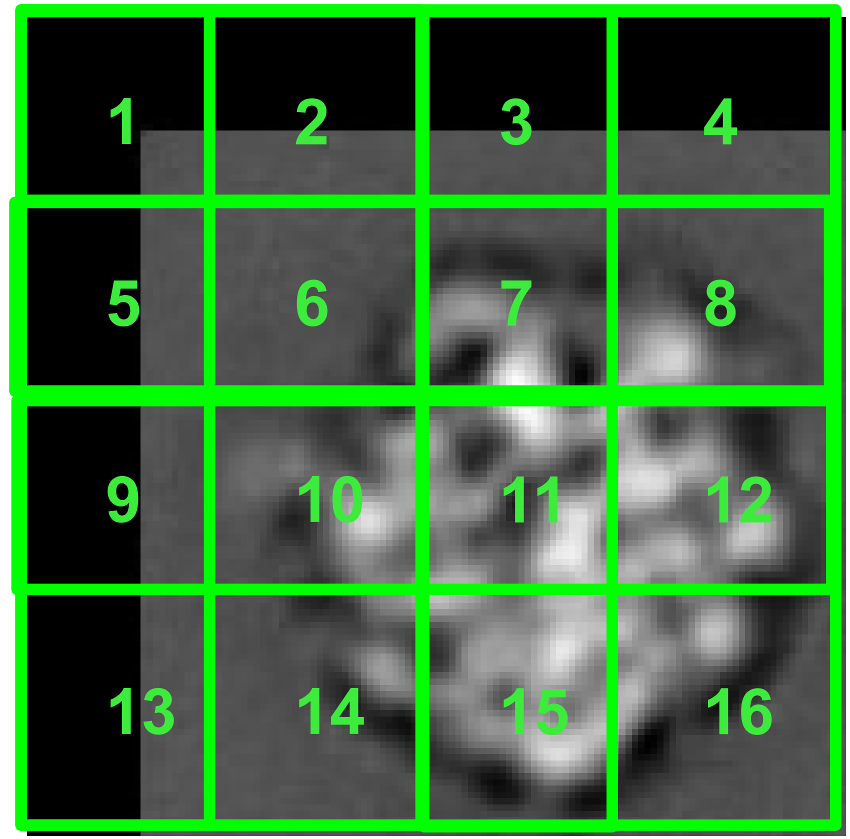


Image *g*

Cross-correlation coefficient:

$$\frac{\sum_{N=1}^{16} f(\vec{x})\, g(\vec{x})}{\sigma_f \sigma_g}$$

constant "normalization"

Cross-correlation

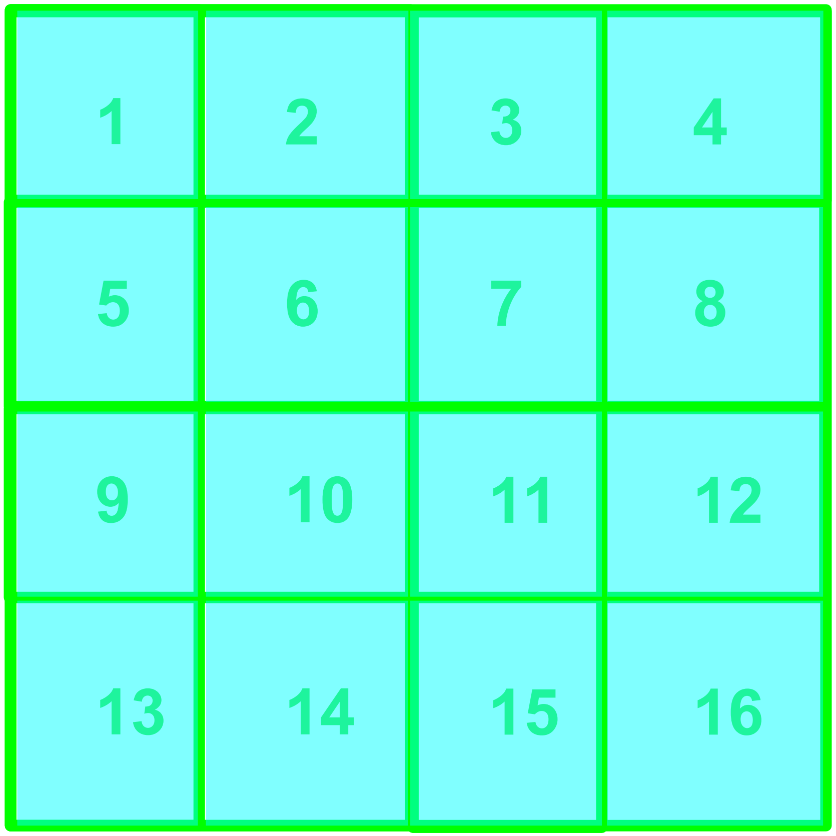


Image *f*

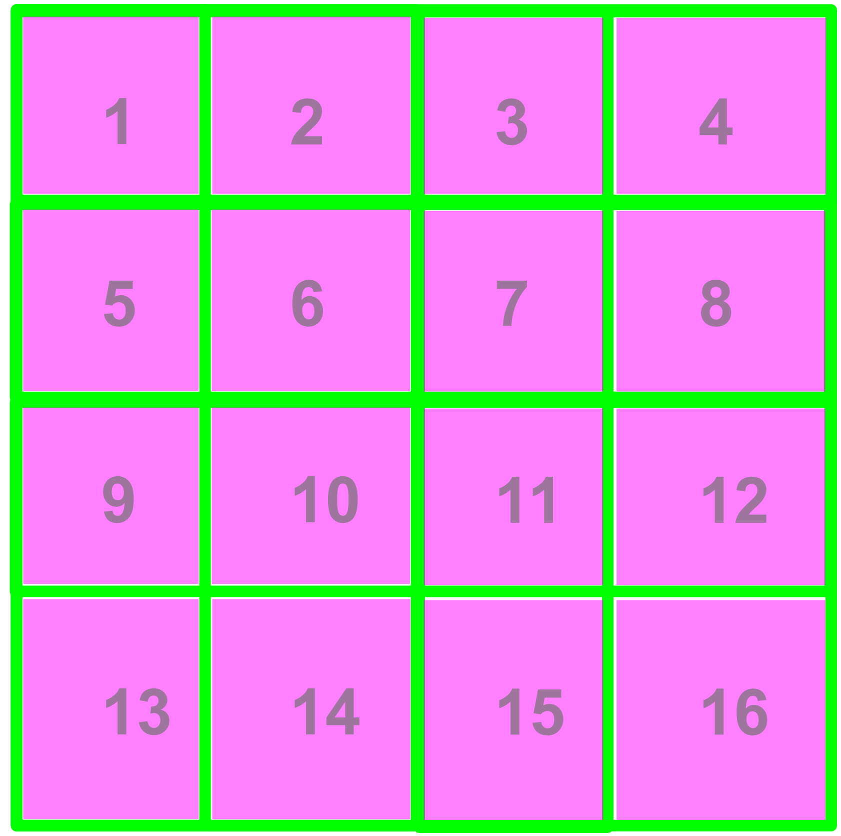


Image *g*

$$\textit{Unnormalized CCC} = f_1g_1 + f_2g_2 + f_3g_3 + f_4g_4 + f_5g_5 + f_6g_6 + f_7g_7 + f_8g_8 + f_9g_9 + f_{10}g_{10} + f_{11}g_{11} + f_{12}g_{12} + f_{13}g_{13} + f_{14}g_{14} + f_{15}g_{15} + f_{16}g_{16}$$

Cross-correlation

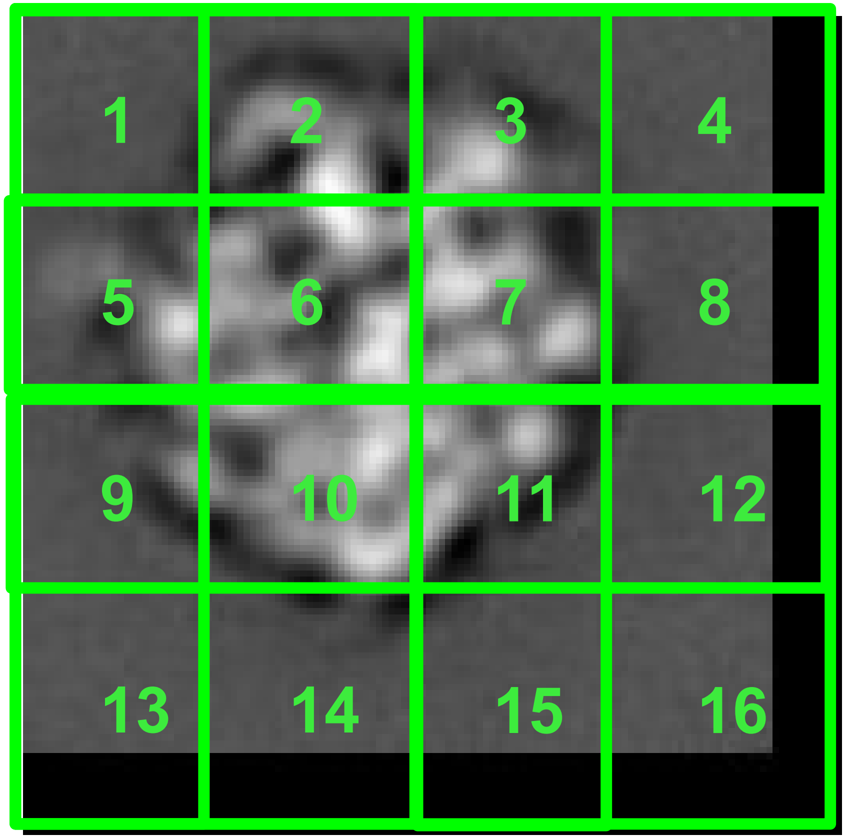


Image f

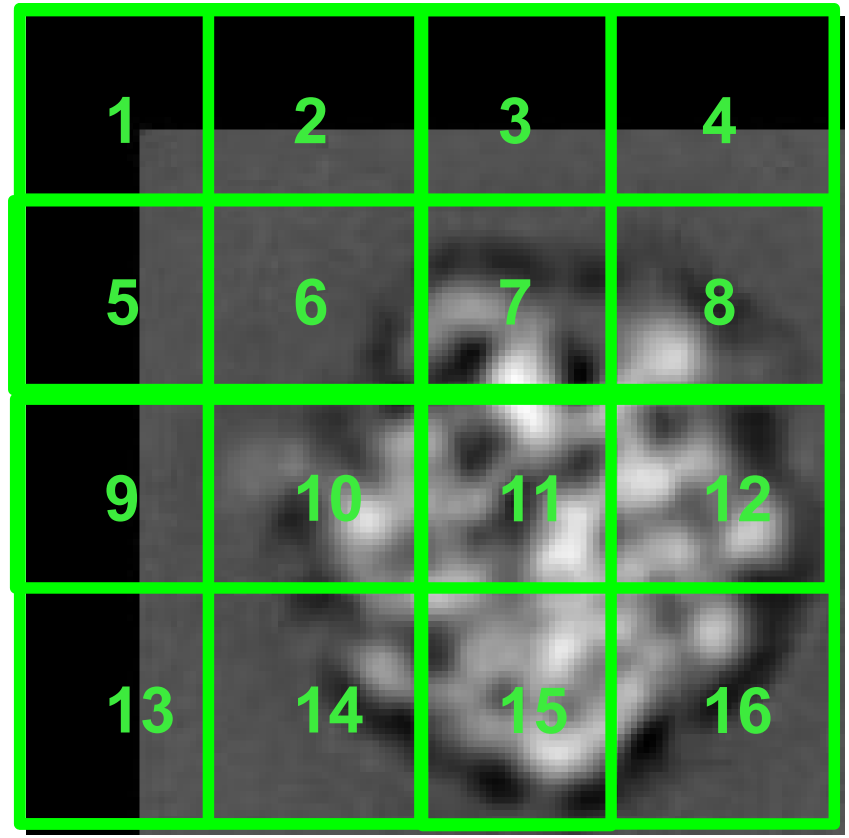


Image g

$$\textit{Unnormalized CCC} = f_1 g_1 + f_2 g_2 + f_3 g_3 + f_4 g_4 + f_5 g_5 + f_6 g_6 + f_7 g_7 + f_8 g_8 + f_9 g_9 + f_{10} g_{10} + f_{11} g_{11} + f_{12} g_{12} + f_{13} g_{13} + f_{14} g_{14} + f_{15} g_{15} + f_{16} g_{16}$$

CEITEC

Cross-correlation coefficient

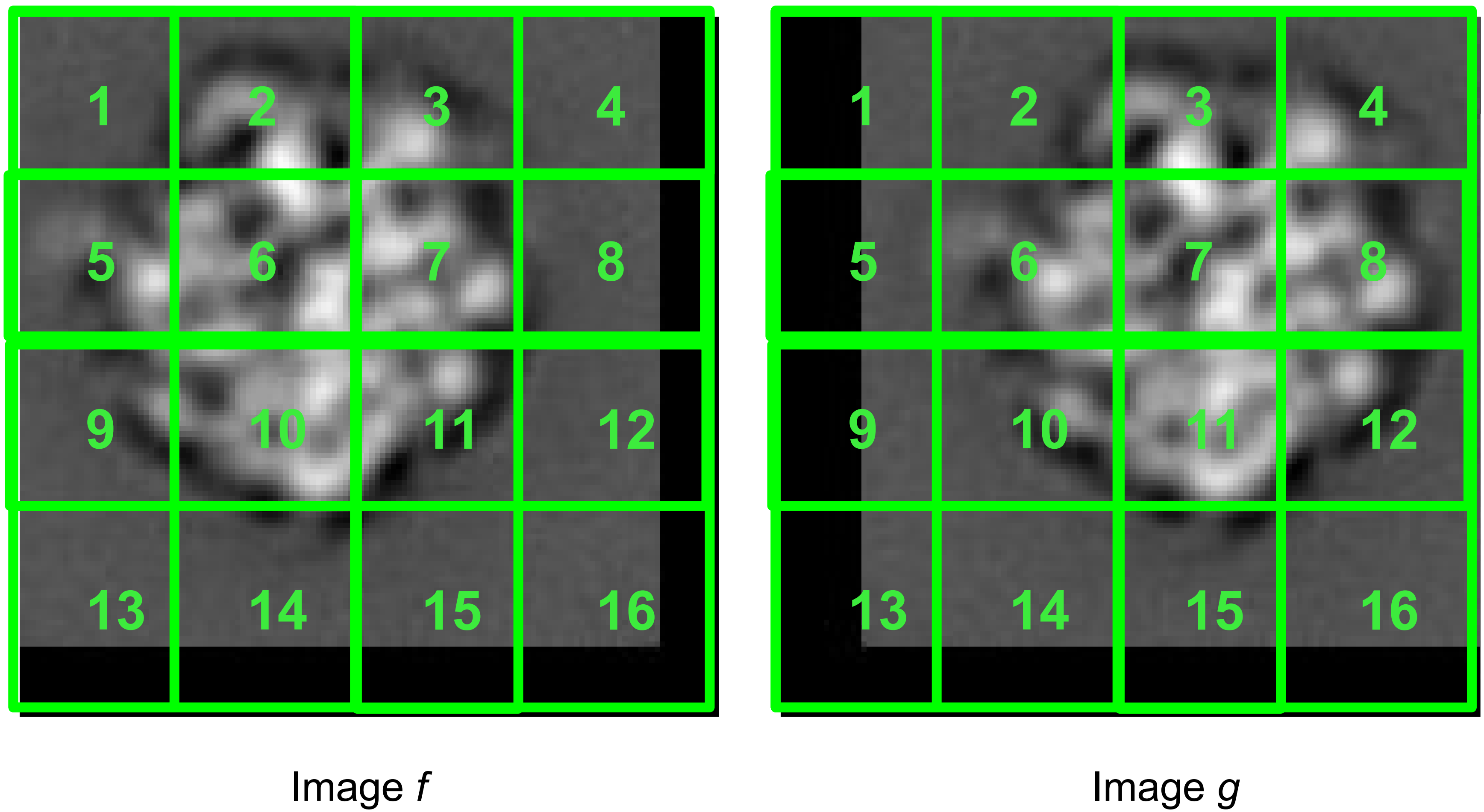
Cross-correlation coefficient: $$\frac{\sum_{N=1}^{16} f(\vec{x})\, g(\vec{x})}{\sigma_f \sigma_g}$$

If the alignment is perfect, the correlation value will be 1.

What if the correlation isn't perfect?

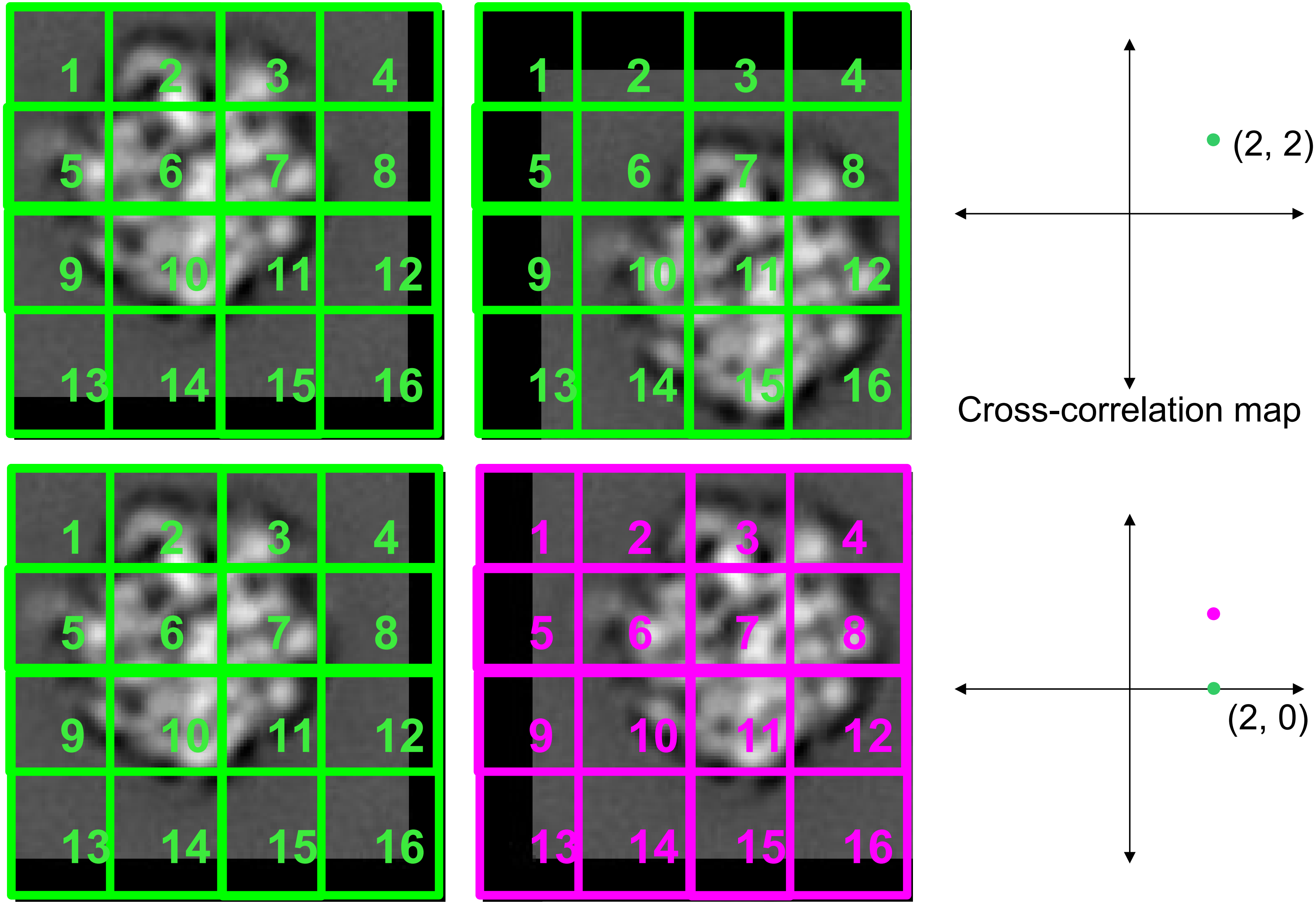
CEITEC

Cross-correlation



What if the correlation isn't perfect?

ANSWER: You try other shifts (perhaps all).



We would need to repeat this for all combinations of shifts.

Cross-correlation function (CCF)

Brute-force translational search is CPU-intensive

BUT

Fourier transforms can help us.

Some notation:

Real space	$f(x)$	$g(x)$
Fourier space	$F(X)$	$G(X)$

Cross-correlation function (CCF)

Brute-force translational search is CPU-intensive

BUT

Fourier transforms can help us.

Complex conjugate:
If a Fourier coefficient $F(X)$ has the form: $a + bi$
The complex conjugate $F^*(X)$ has the form: $a - bi$

$$F^*(X)\,G(X) = \text{F.T.(CCF)}$$

This gives us a map of all possible shifts.

Cross-correlation function (CCF)

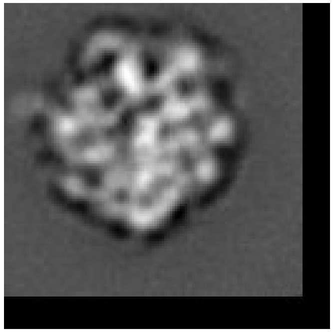


Image *f(x)*

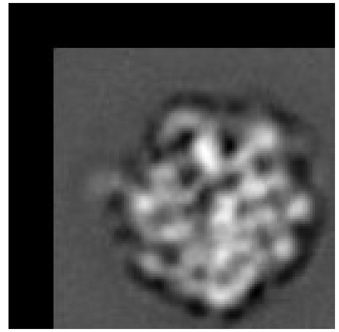
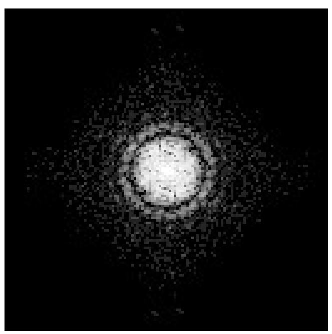
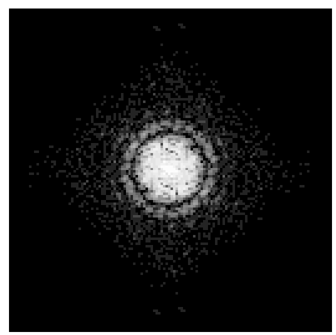


Image *g(x)*



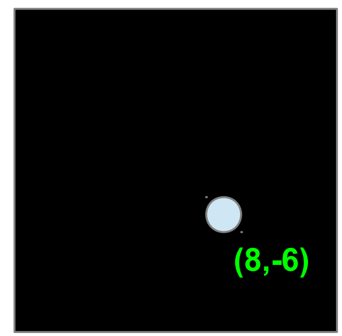
F.T. *F*(X)*
(complex conjugate)

x



F.T. *G(X)*

=



F.T. (CCF)

The position of the peak gives us the shifts that give the best match, *e.g.*, (8,-6).

Well, that was an easy case.
We only needed to do translational alignment.
What about orientation alignment?

Orientation alignment

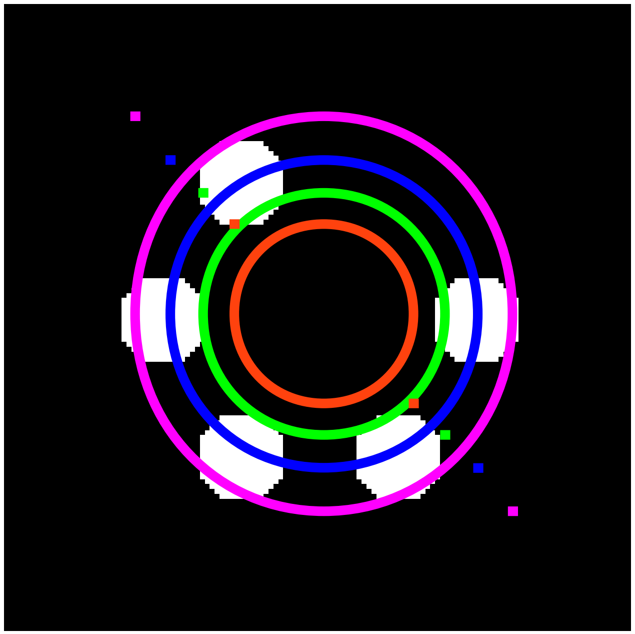


Image 1

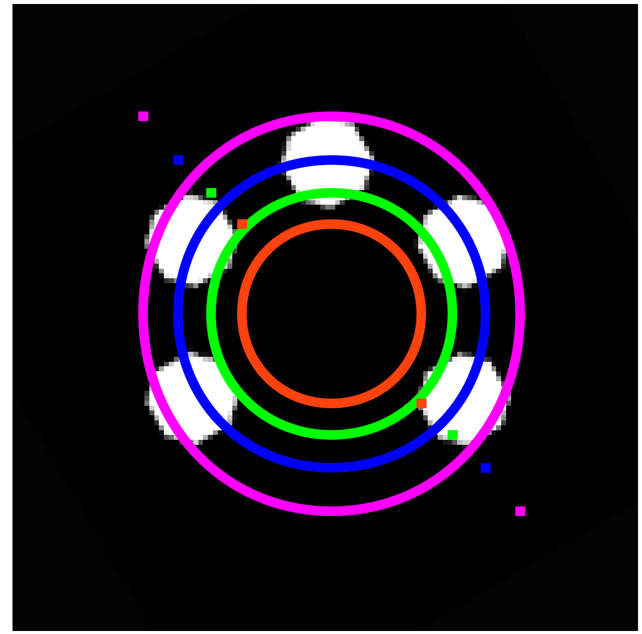
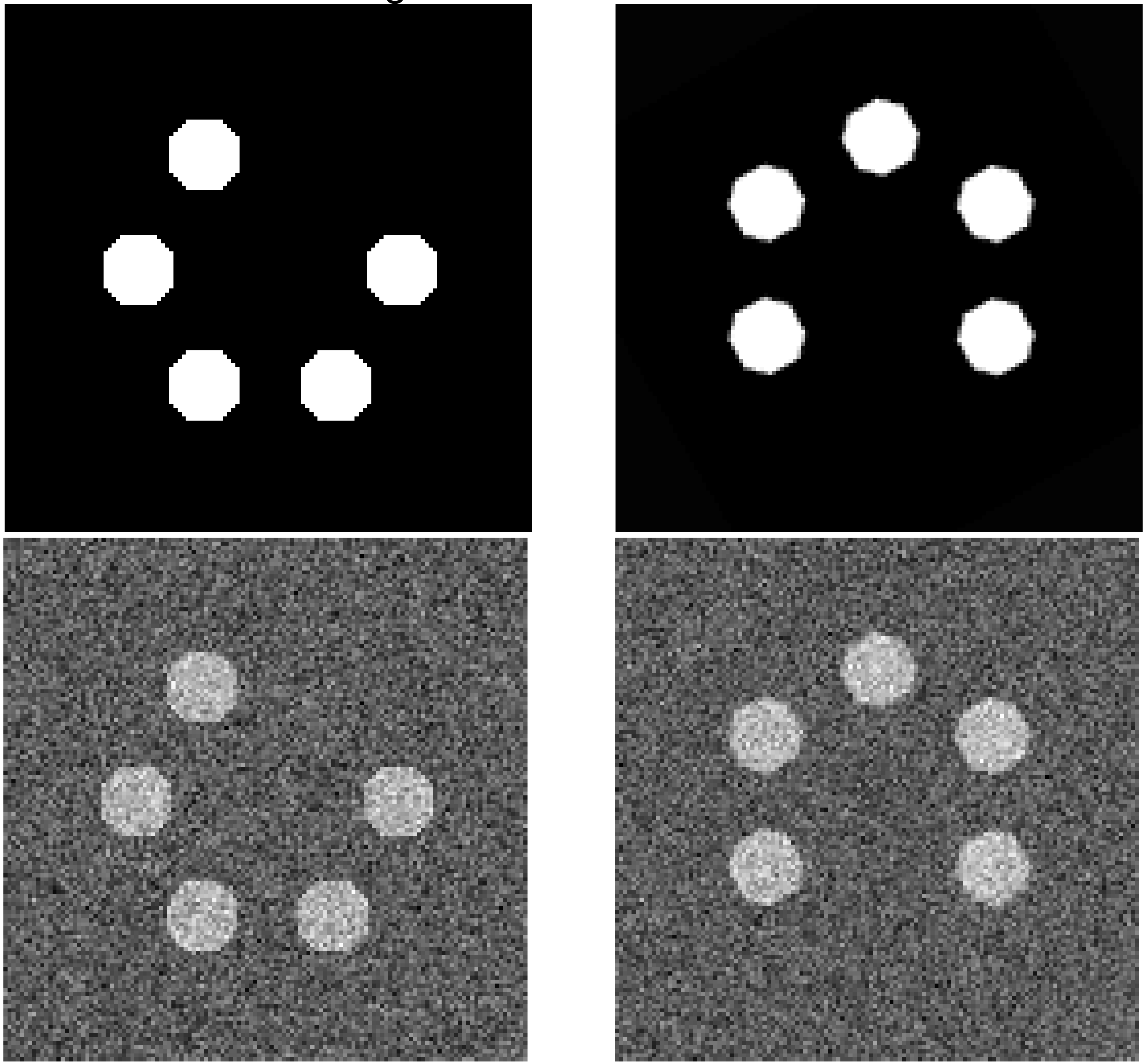


Image 2

We take a series of rings from each image, unravel them, and compute a series of 1D cross-correlation functions.

Shifts along these unraveled CCFs is equivalent to a rotation in Cartesian space.

Reference image



Noise added

Orientation alignment

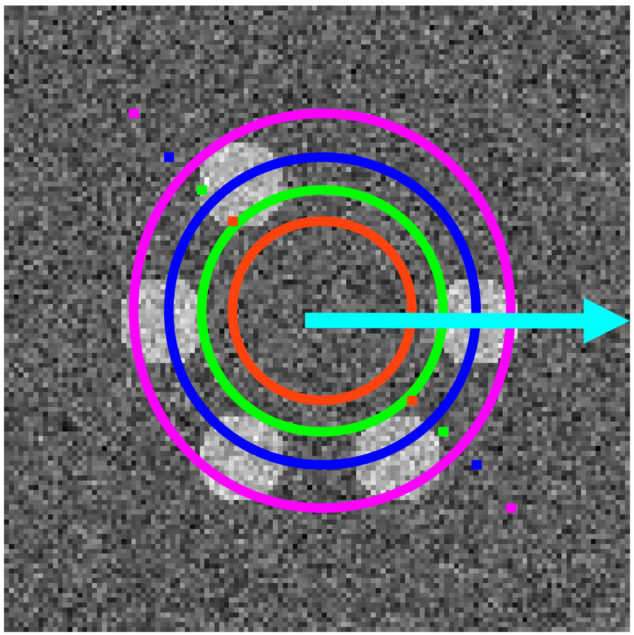


Image 1

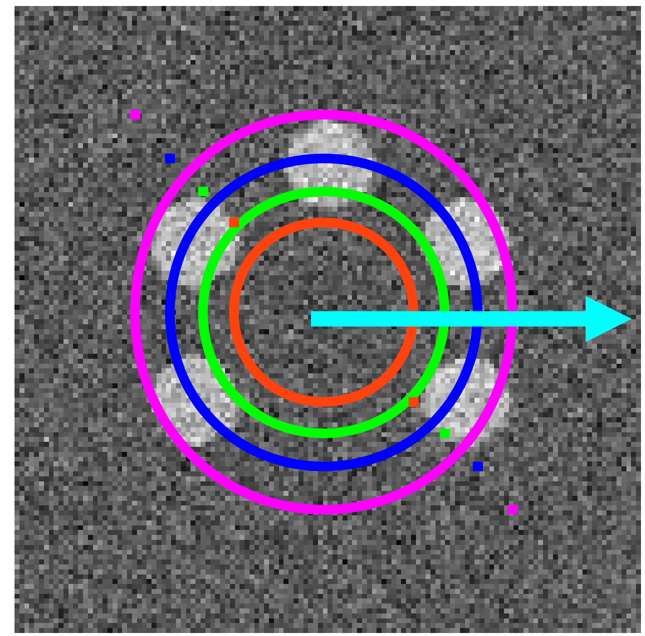
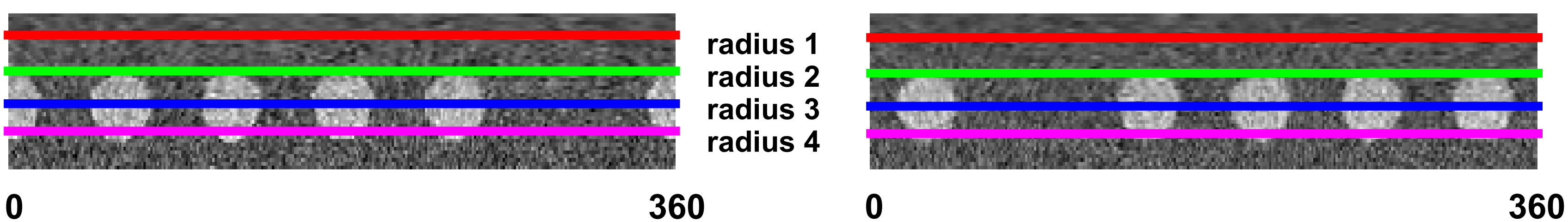
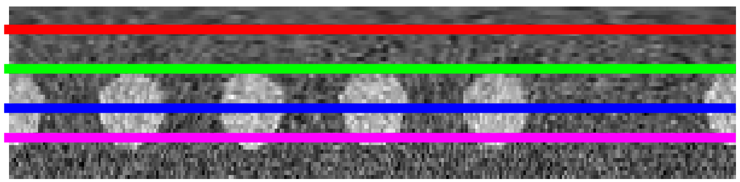


Image 2

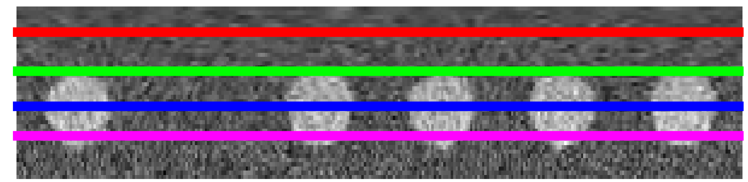


Polar representation

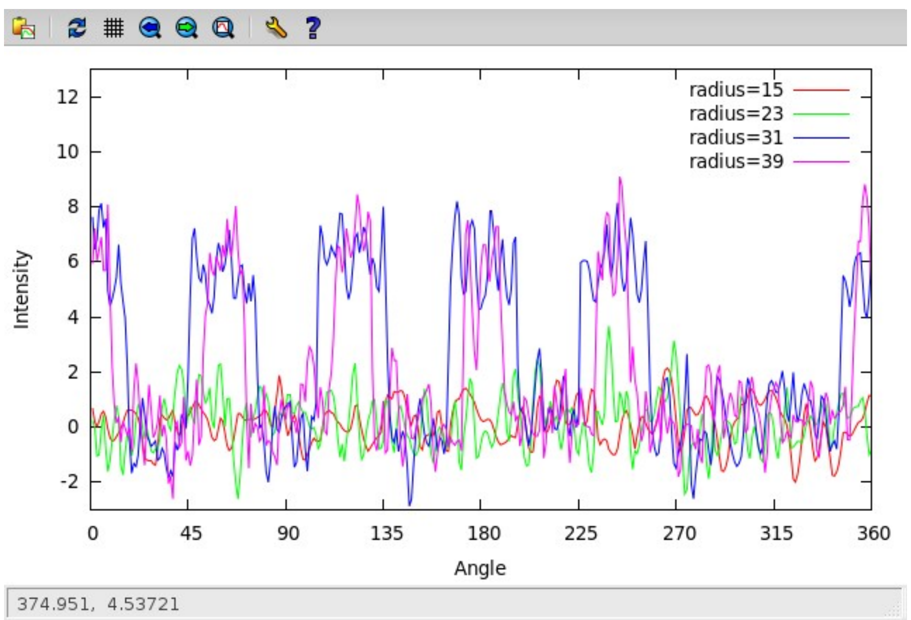
Orientation alignment

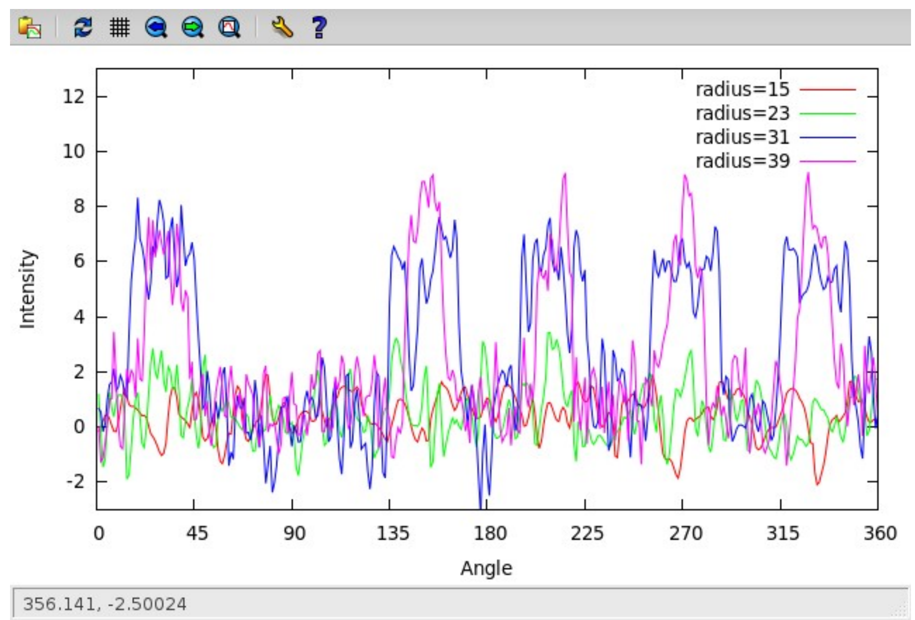


radius 1
radius 2
radius 3
radius 4



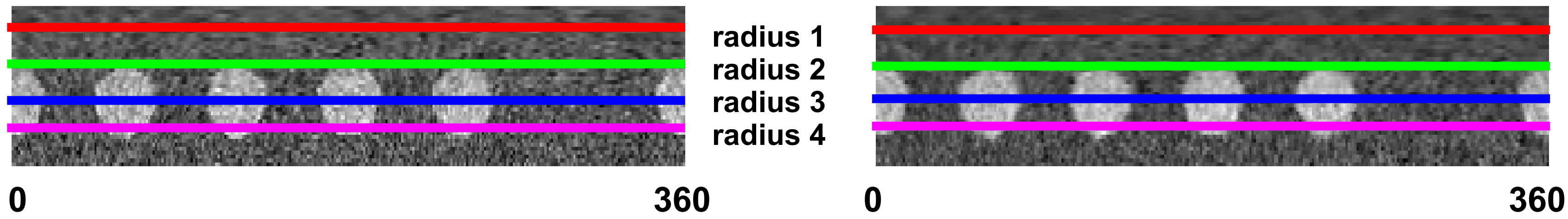
0 360 0 360

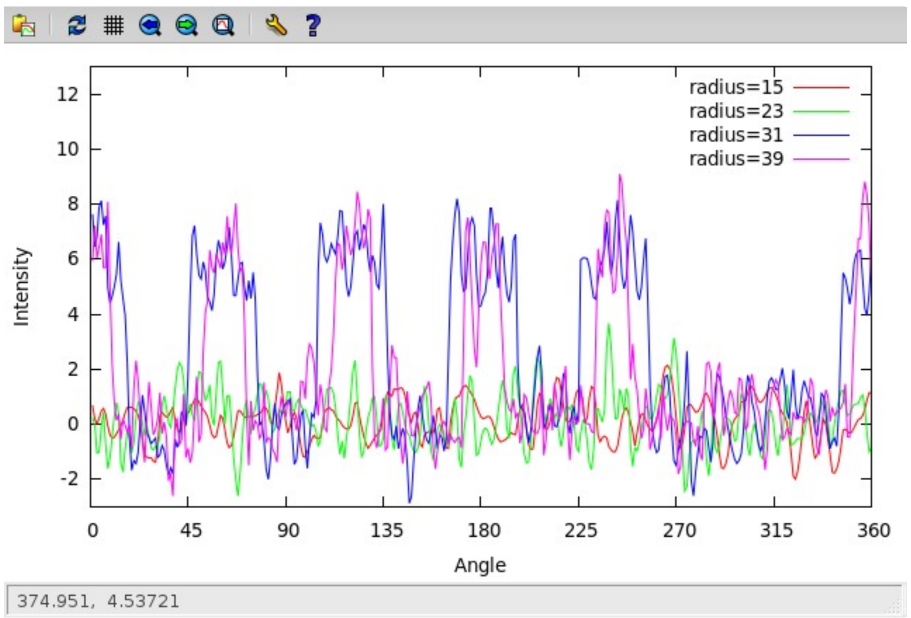


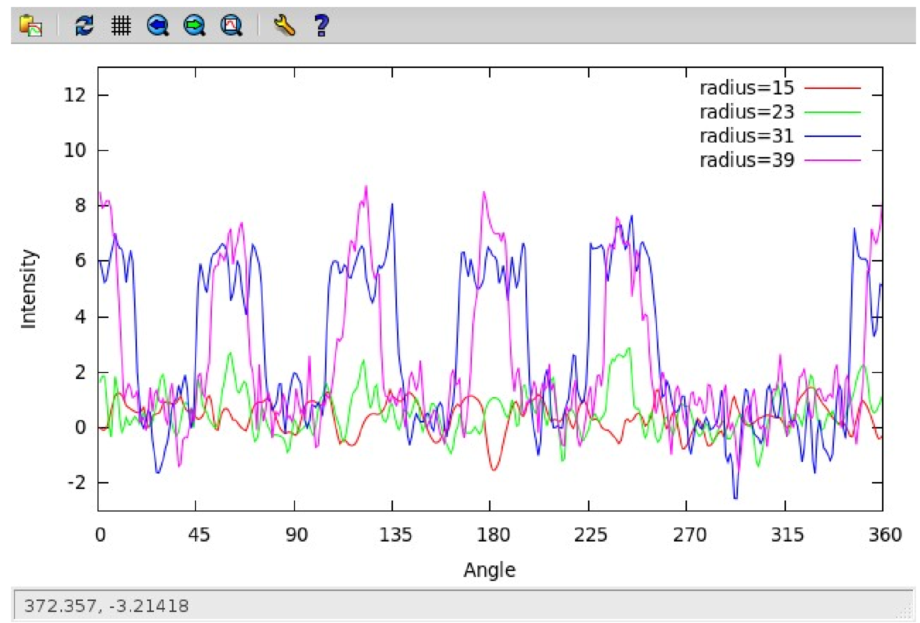


CEITEC

Orientation alignment: After rotation







Which do you perform first?
Translational or orientation alignment?

Translational and orientation alignment are interdependent

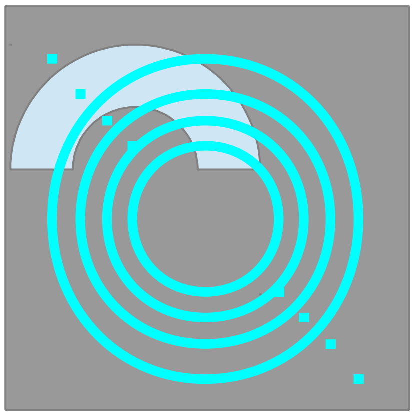


Image 1

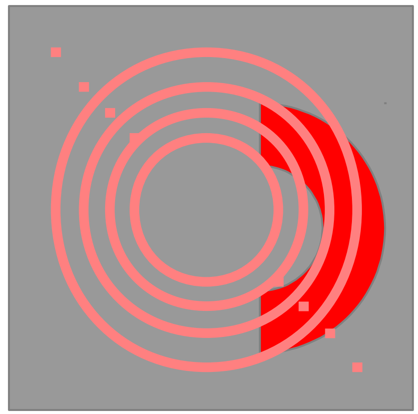
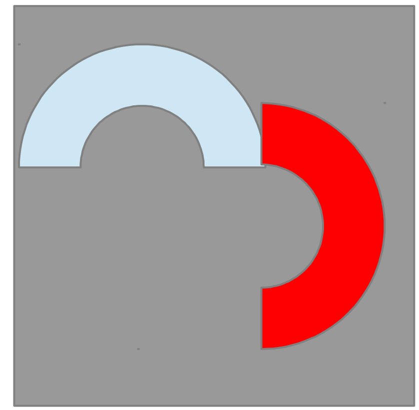


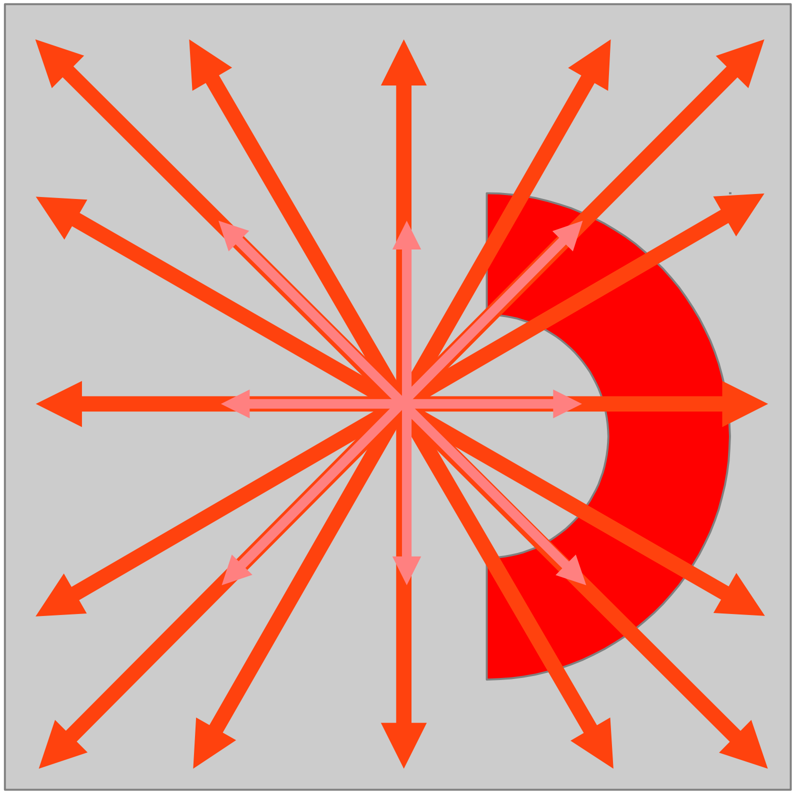
Image 2



Superimposed

SOLUTION: You try a set of reasonable shifts, and perform separate orientation alignments for each.

Translational and orientation alignment are interdependent



Set of all shifts of up to 1 pixel

Set of all new shifts of up to 2 pixels

Shifts of (0, +/-1, +/-2) pixels results in 25 orientation searches.

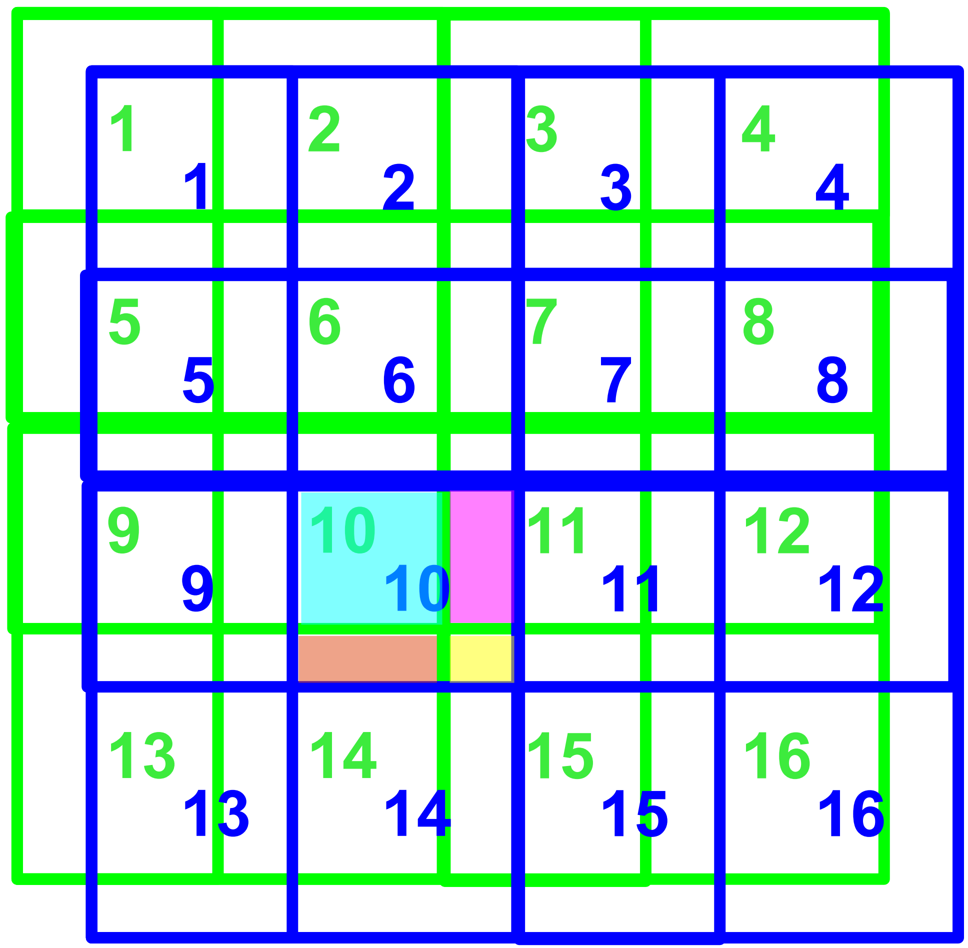
Outline

Image analysis II

- Fourier transforms revisited
  - Ducks and other animals
- Analogy to the Ewald sphere
- Aliasing
- Alignment
- Interpolation
- Multivariate data analysis

How to apply the best shift and rotation?

Shifts

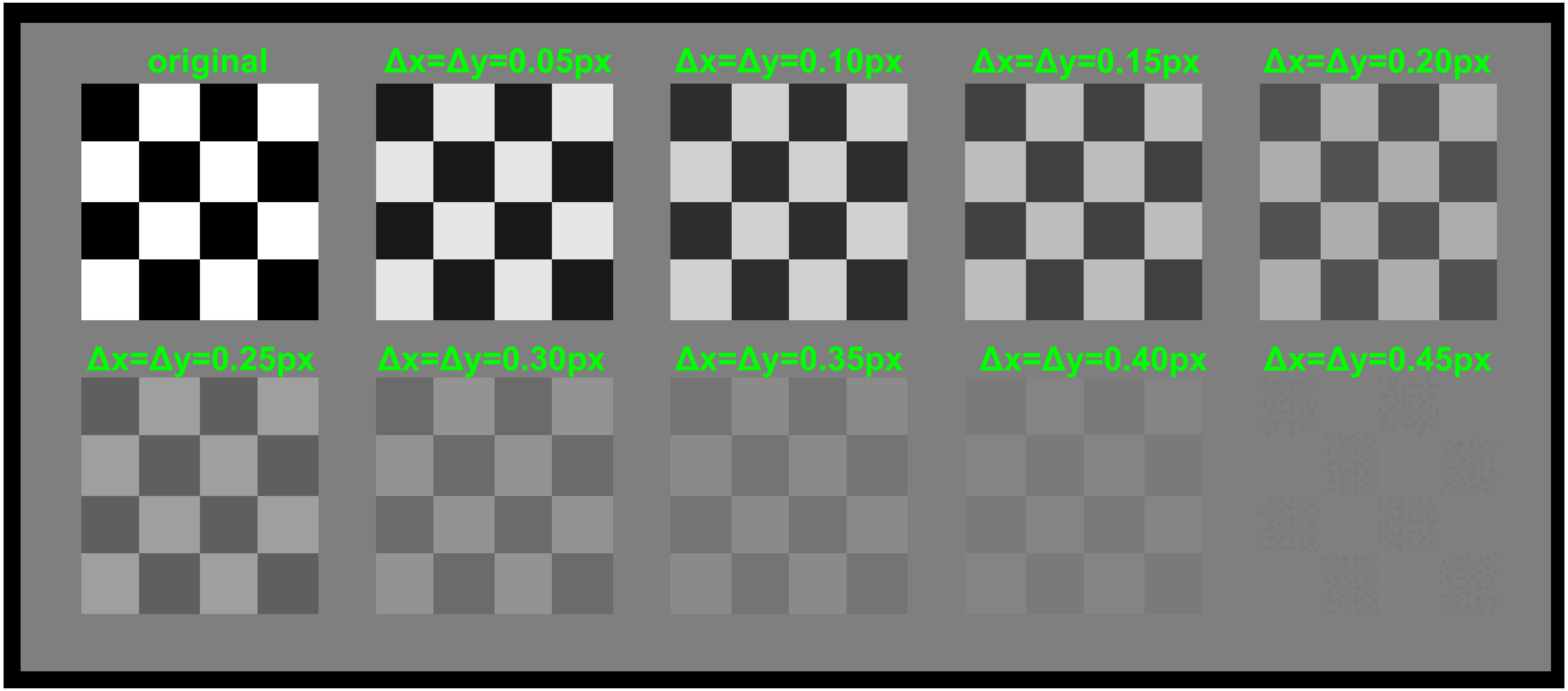


Suppose we shift the image in *x* & *y*.

The new pixels will be weighted averages of the old pixels.

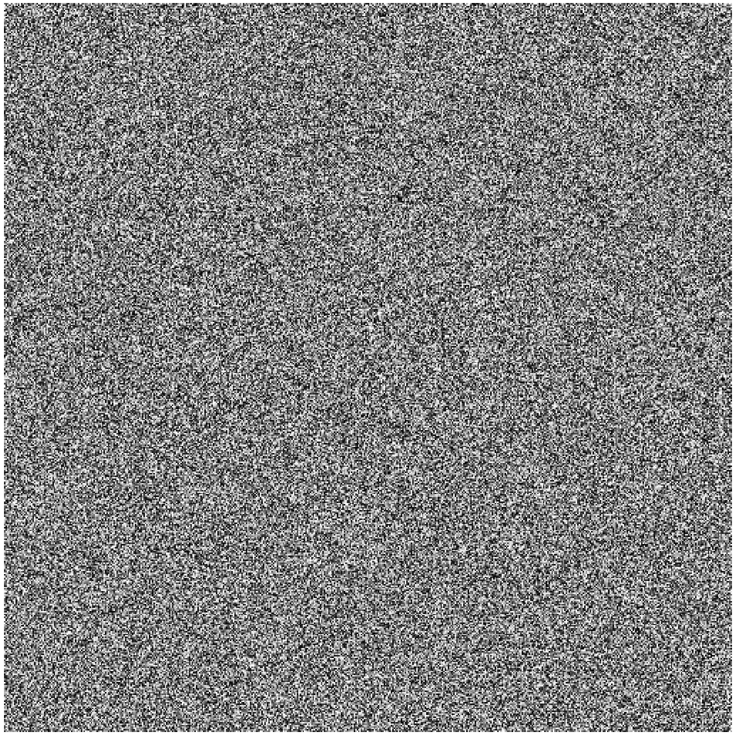
The more the mix the pixels, the worse the result will be.

Effect of shifts

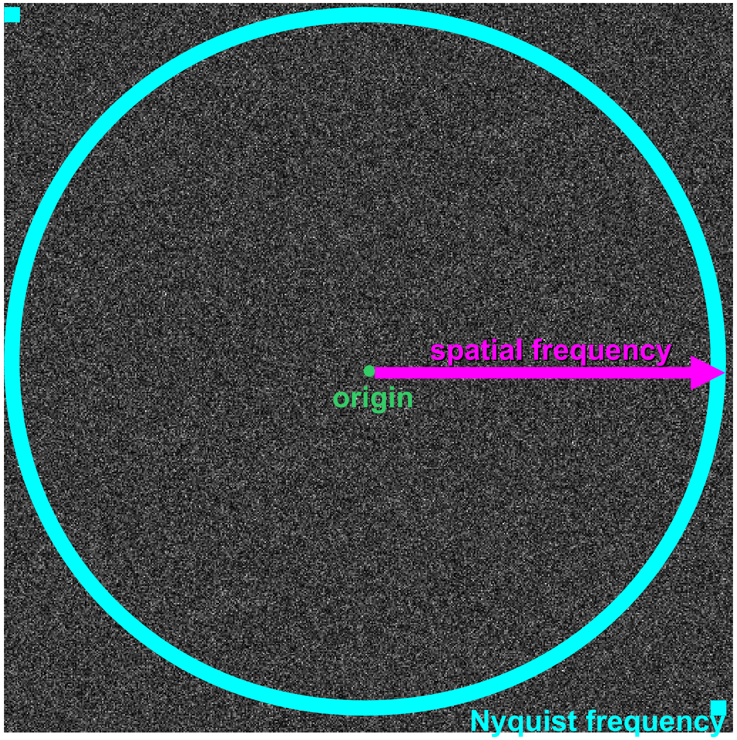


Two more properties of Fourier transforms: Noise

- The Fourier transform of noise is noise
- “White” noise is evenly distributed in Fourier space
    - “White” means that each pixel is independent



White noise

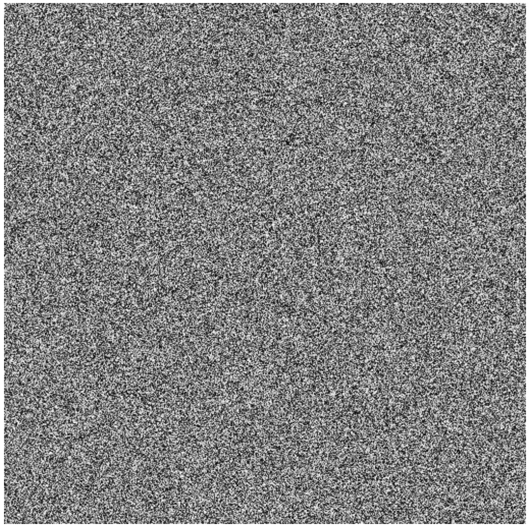


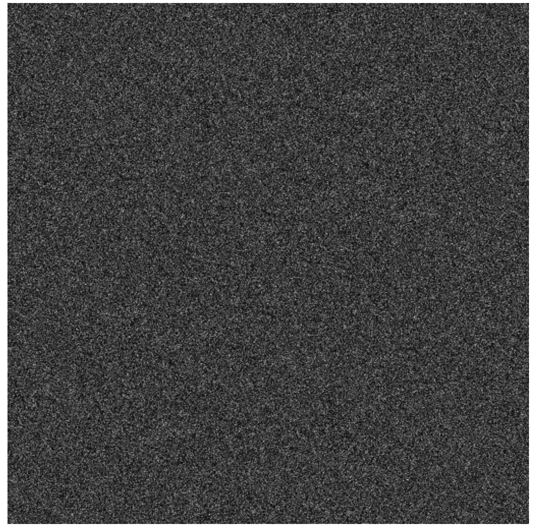
Power spectrum

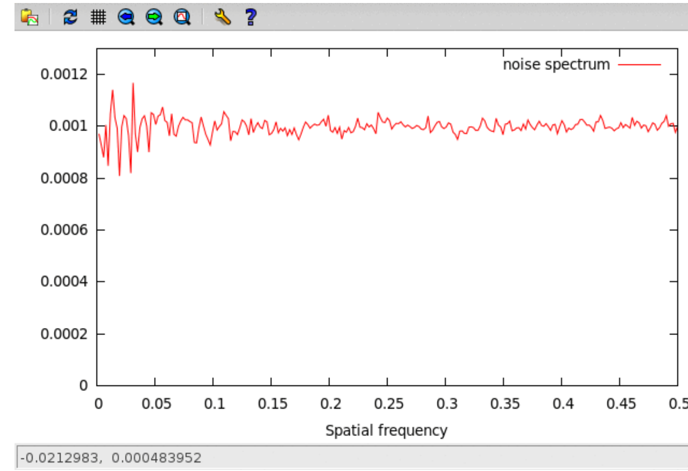
CEITEC

Effects of interpolation are resolution-dependent

Original





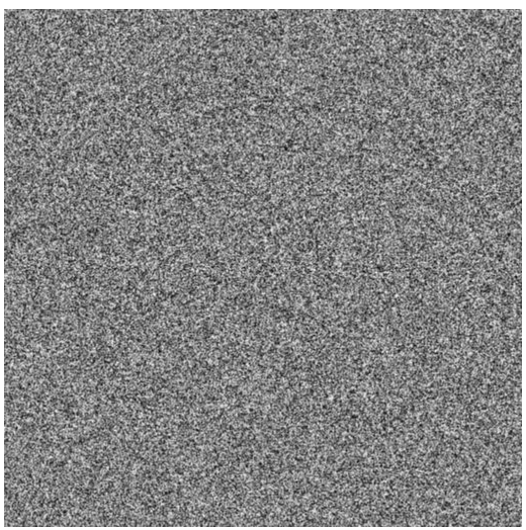


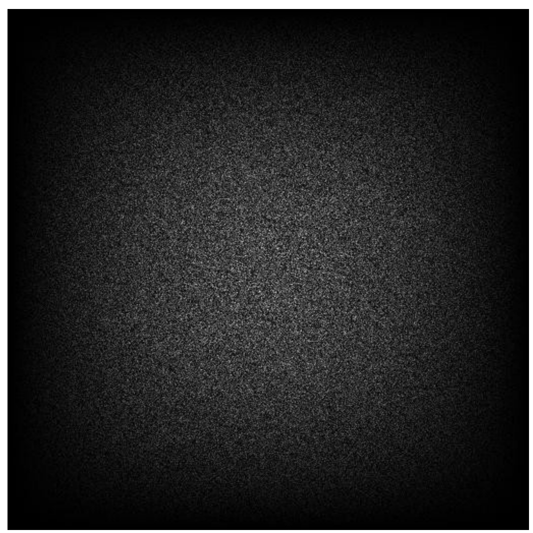
Image

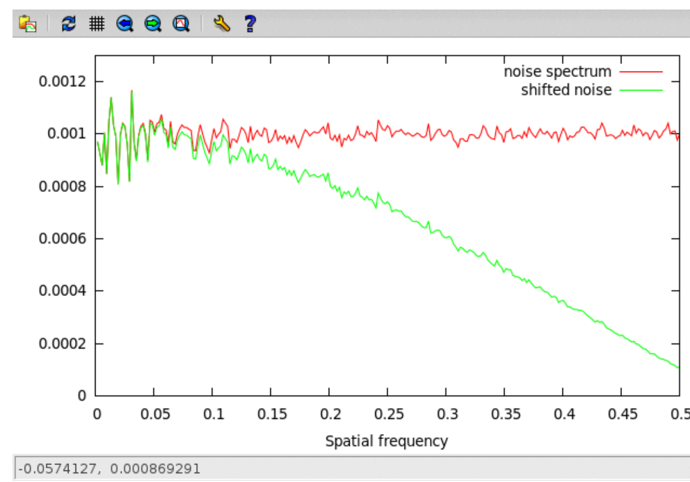
Power spectrum

Profile

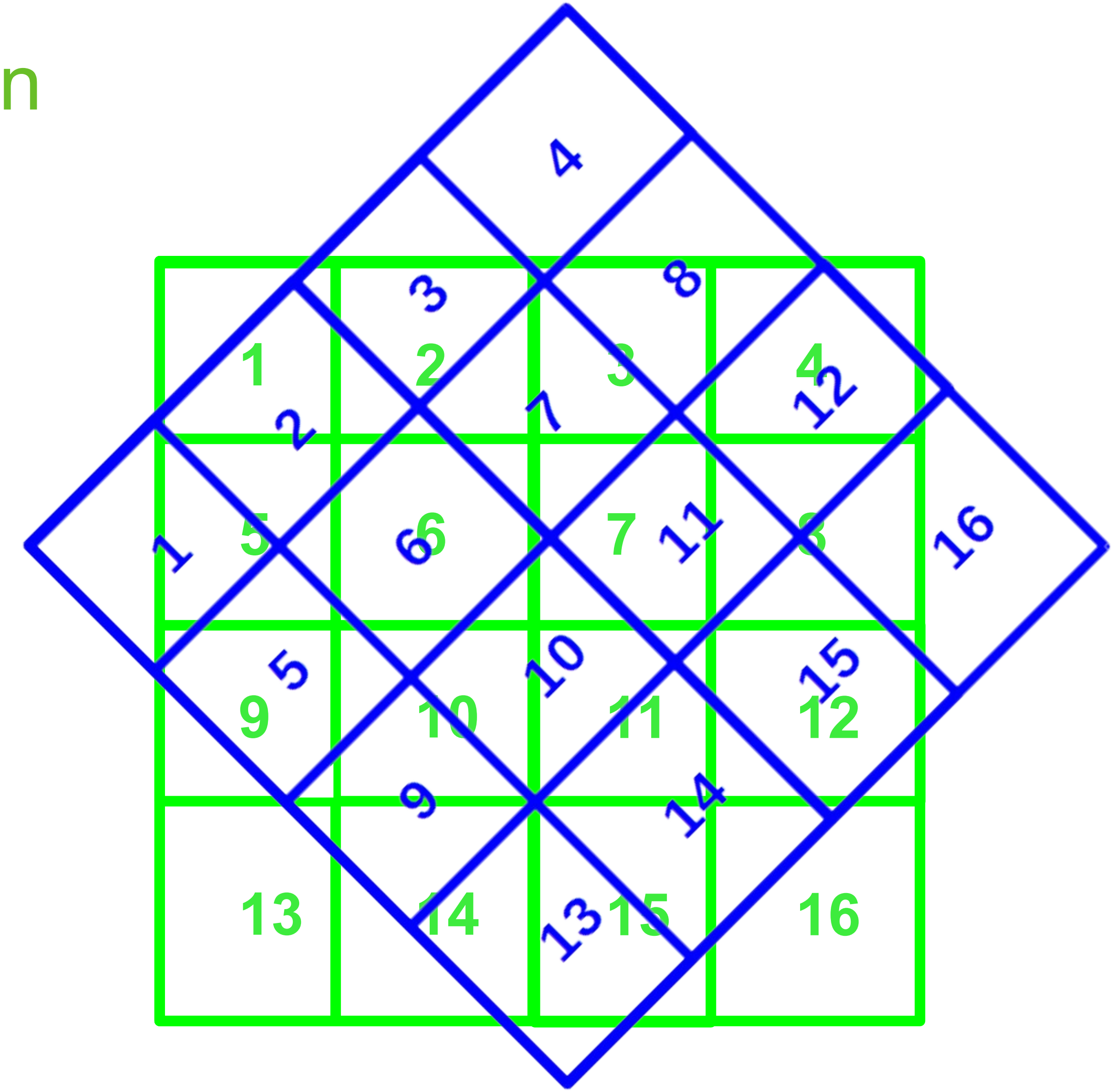
Shifted by (0.5,0.5) px







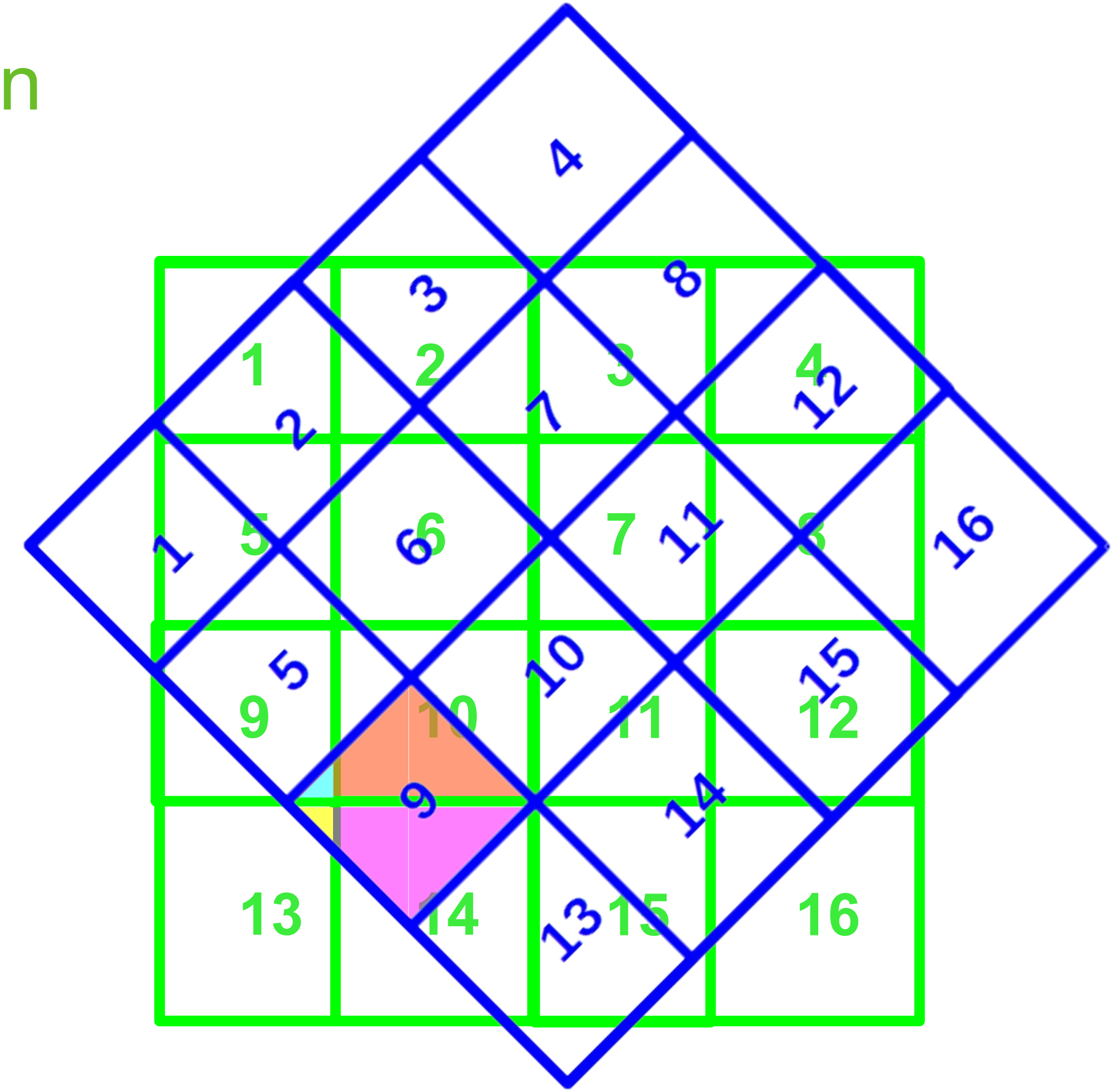
Rotation



Suppose we rotate the image.

The new pixels will be weighted averages of the old pixels.

Rotation



Suppose we rotate the image.

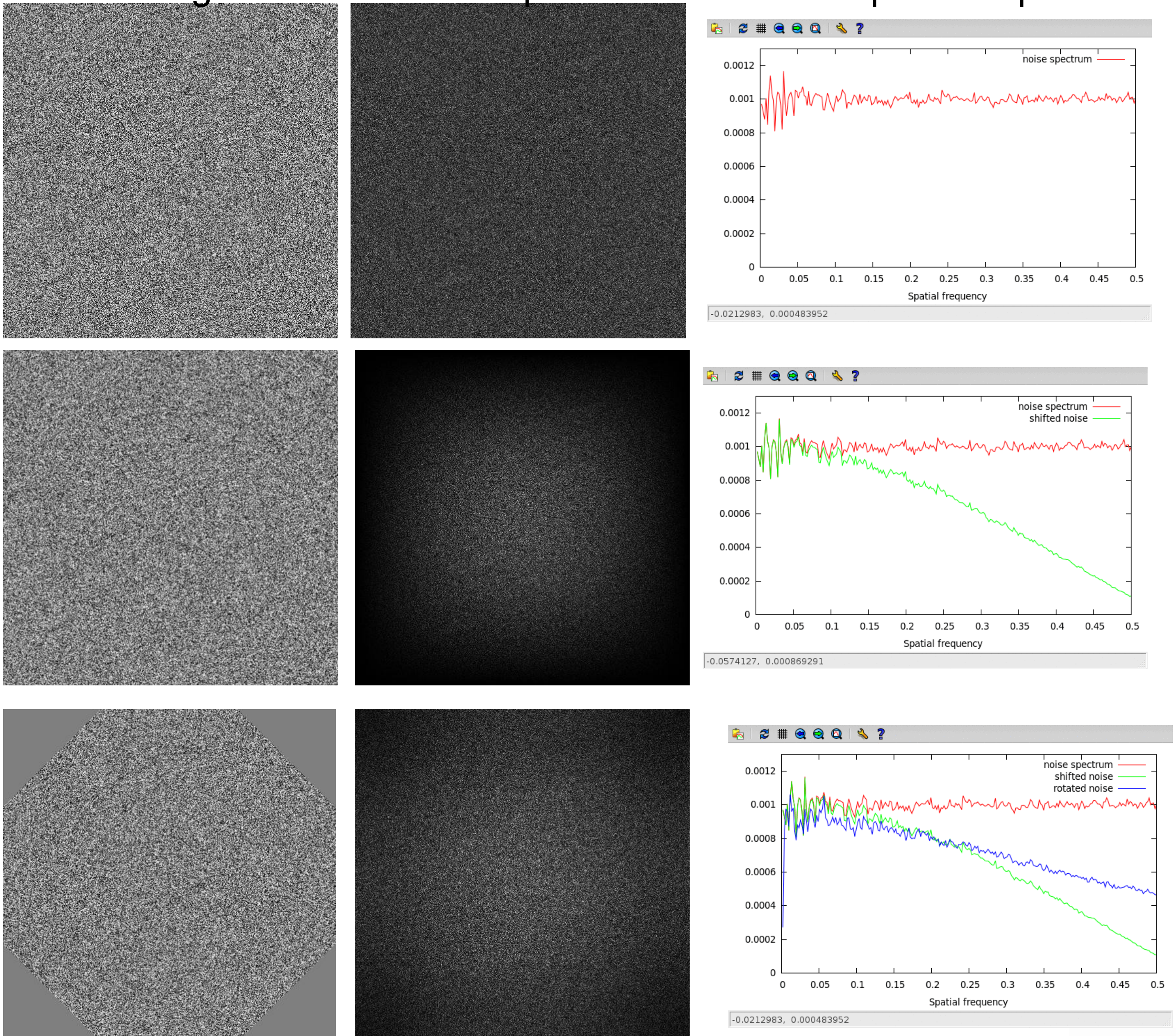
New pixel #9 will be a weighted sum of old pixels 9, 10, 13, and 14.

Image | Power spectrum | Power spectrum profile

Original

Shifted by (0.5,0.5) px

Rotated by 45°



The degradation of the images means that we should minimize the number of interpolations.

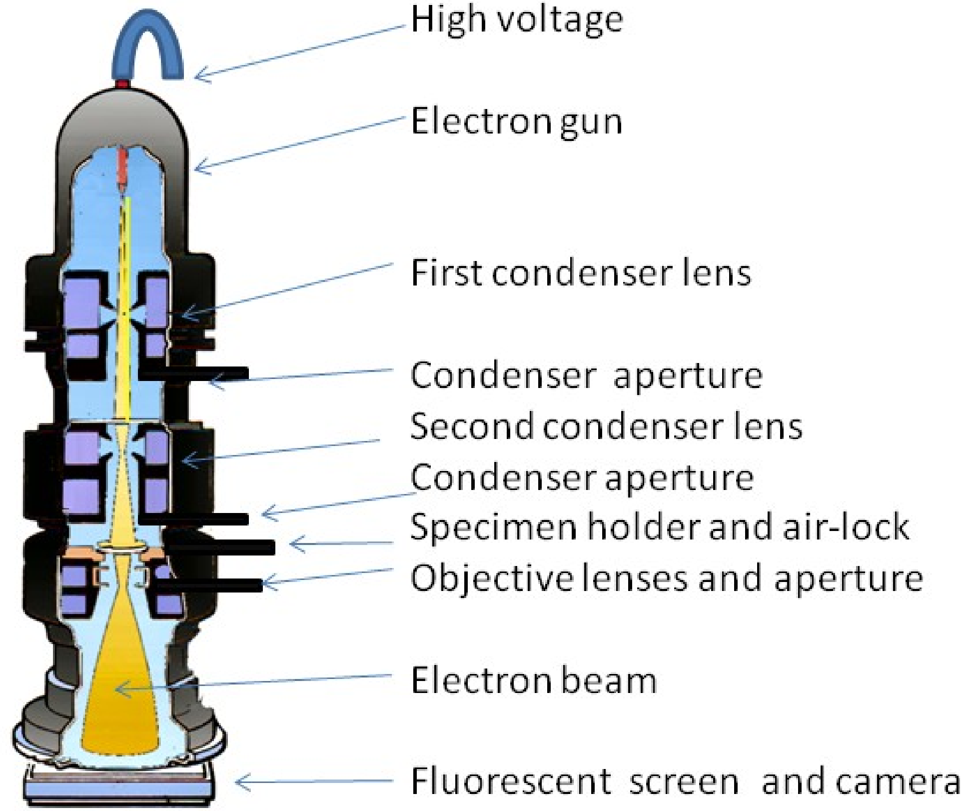
From two weeks ago...

Typical magnification: 50,000X
Typical detector element: 15µm
(pixel size on the camera scale)

Pixel size on the specimen scale:
15×10^{-6} m/px / 50000 =
3.0×10^{-10} m/px = **3.0 Å/px**

In other words,
the best resolution we can achieve (or, the finest oscillation we can detect) at 3.0 Å/px is **6.0 Å**.

It will be worse due to interpolation, so to be safe, a pixel should be 3X smaller than your target resolution.

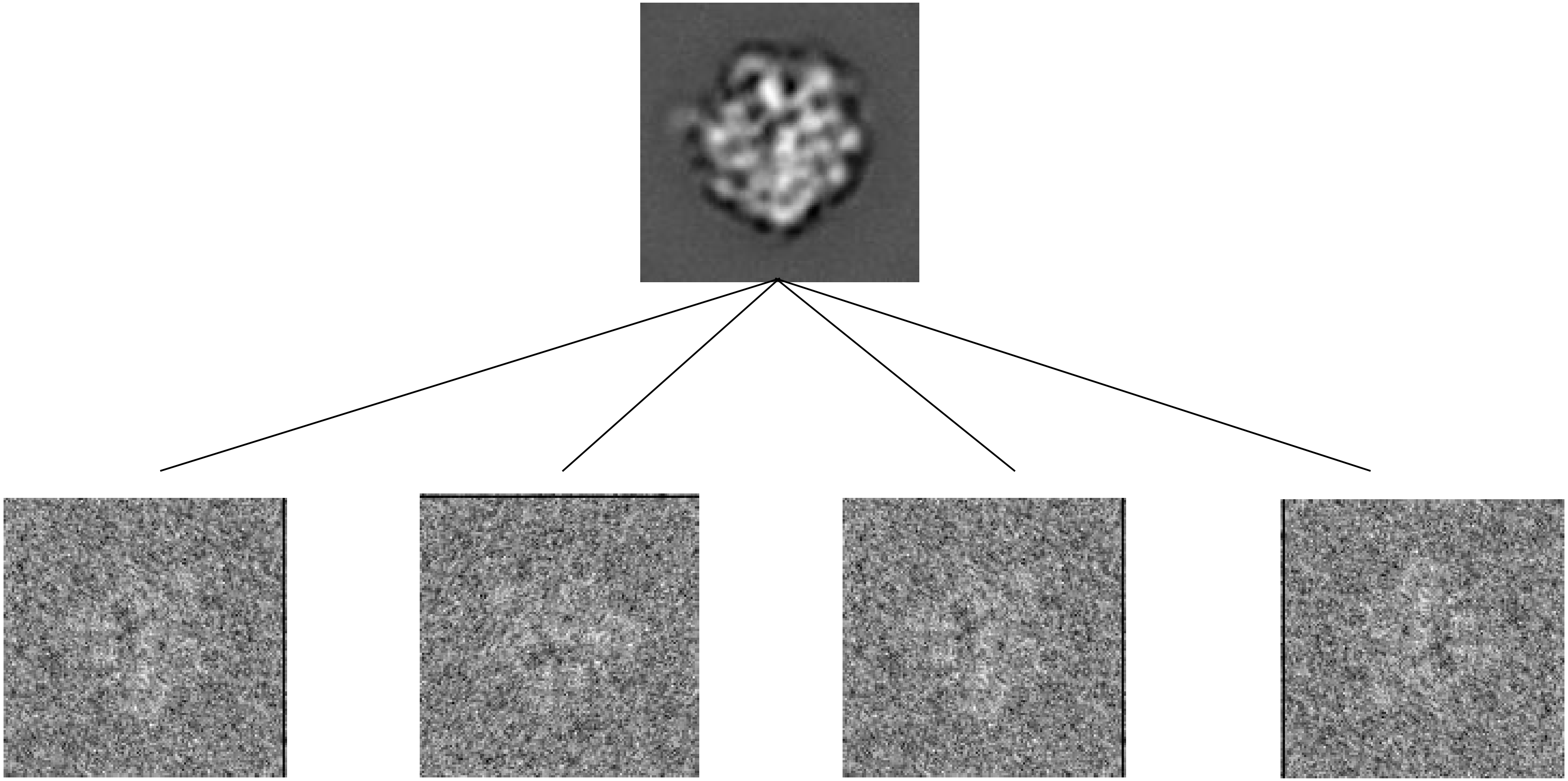


Transmission Electron Microscope

http://www.en.wikipedia.org

Different alignment strategies

Reference-based alignment

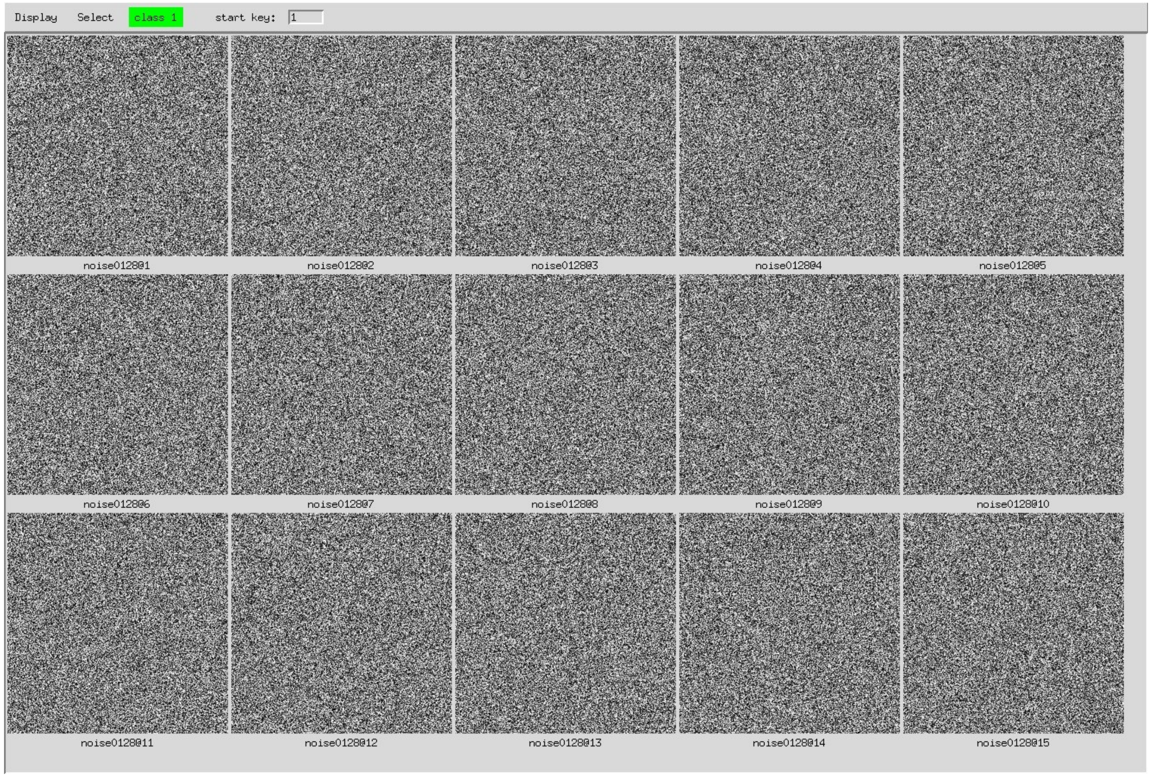


There's a problem with reference-based alignment:

Model bias

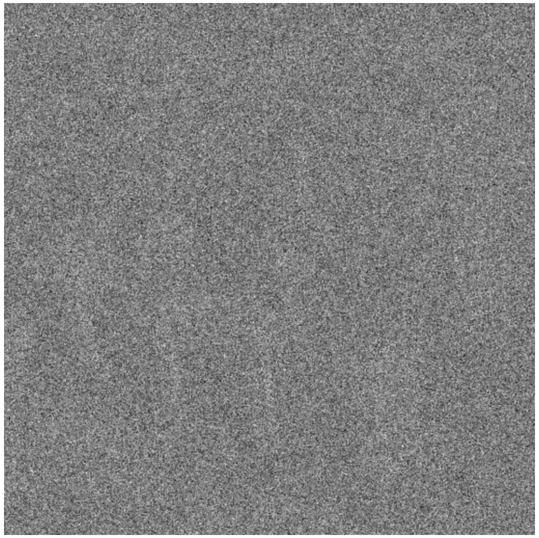
Model bias

Reference

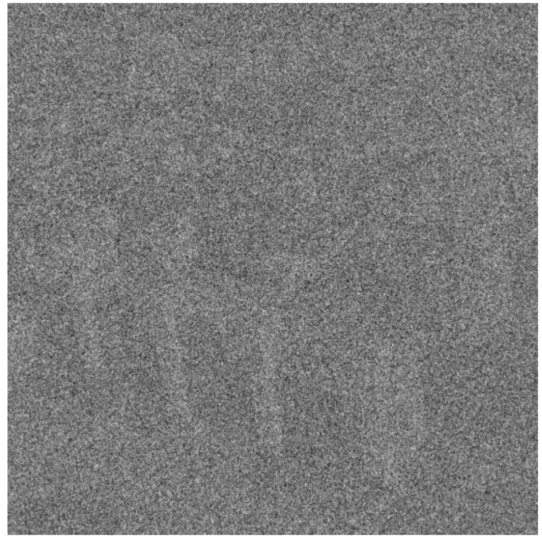


Images of pure noise

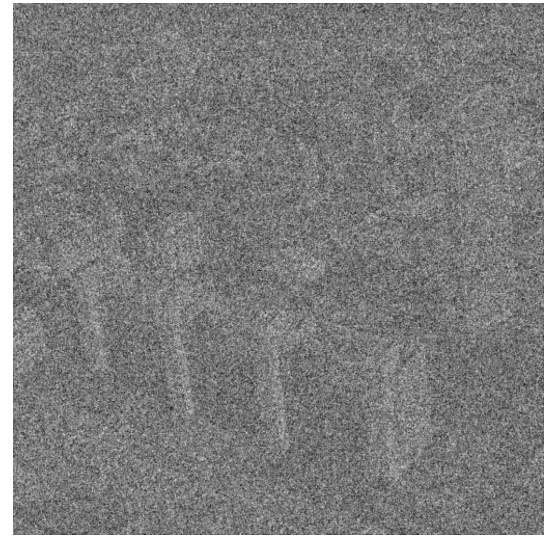
Averages of images of pure noise



N = 128

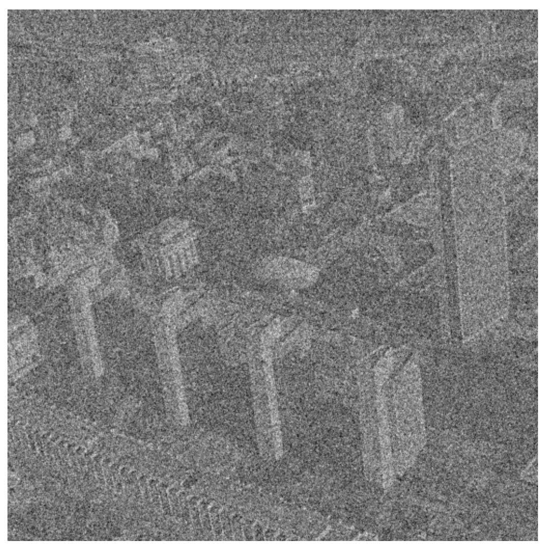


N = 256



N = 512

N = 1024

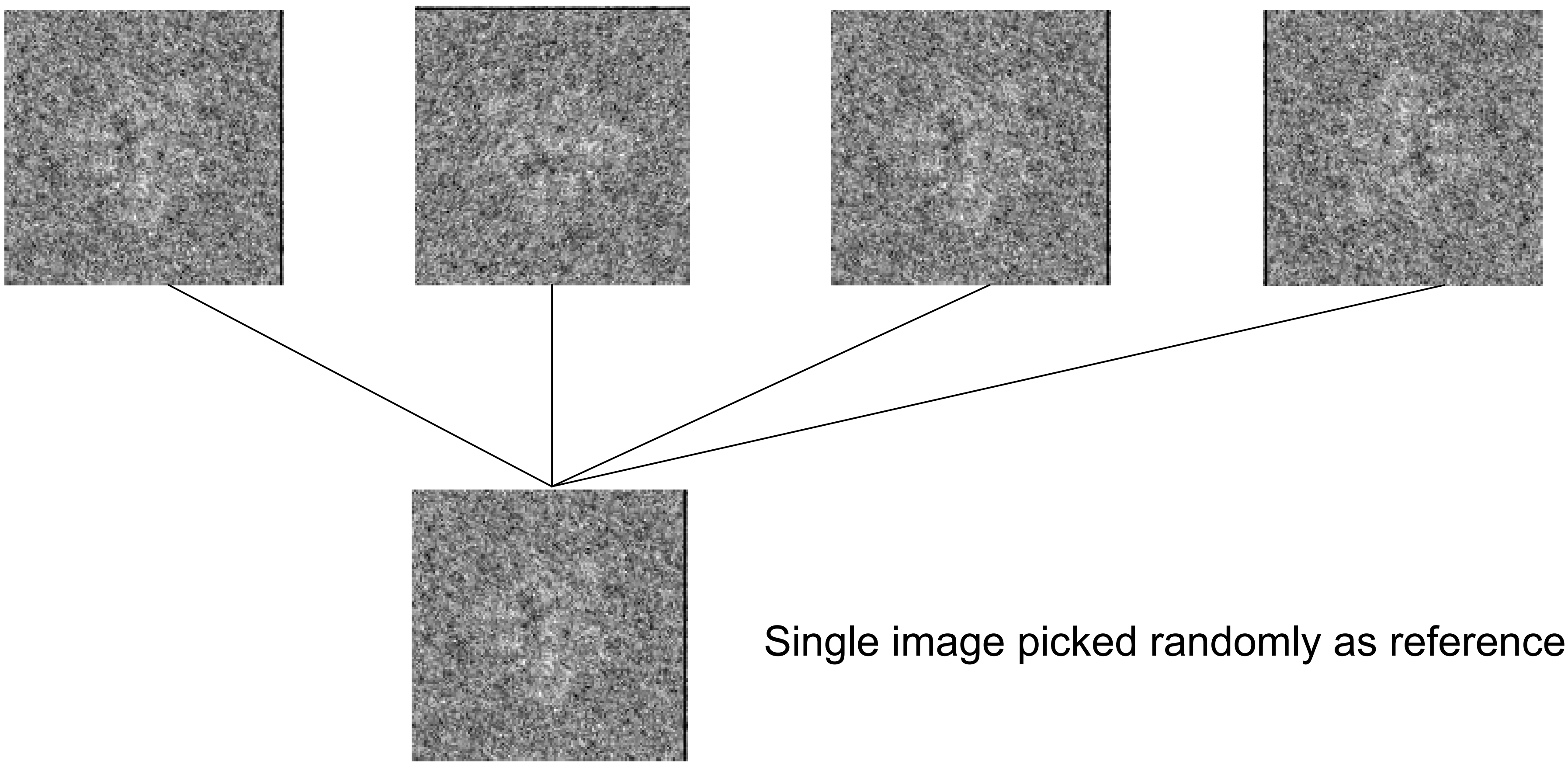


N = 2048

original

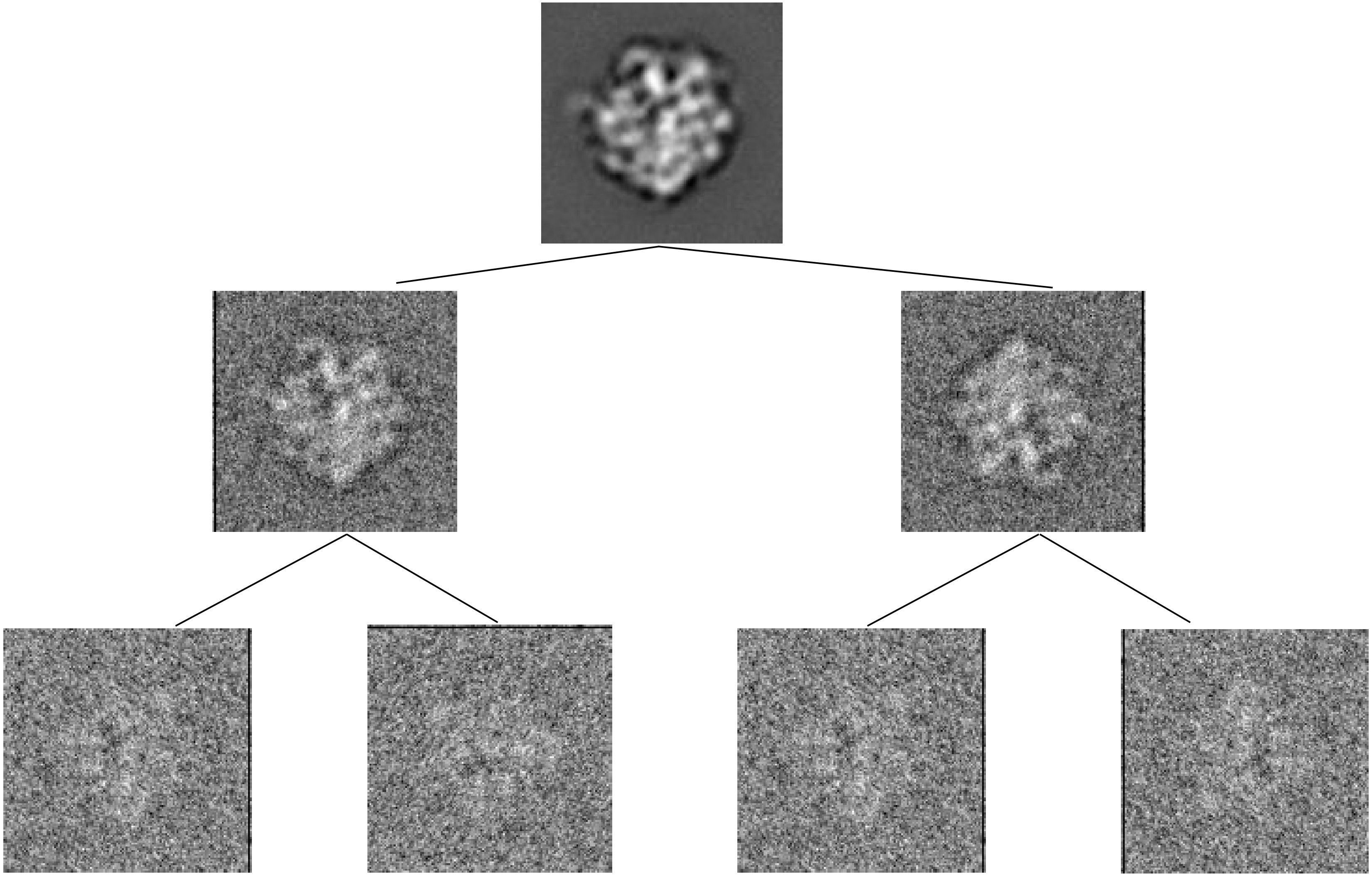
There are reference-free alignment schemes

Reference-free alignment (SPIDER command AP SR)



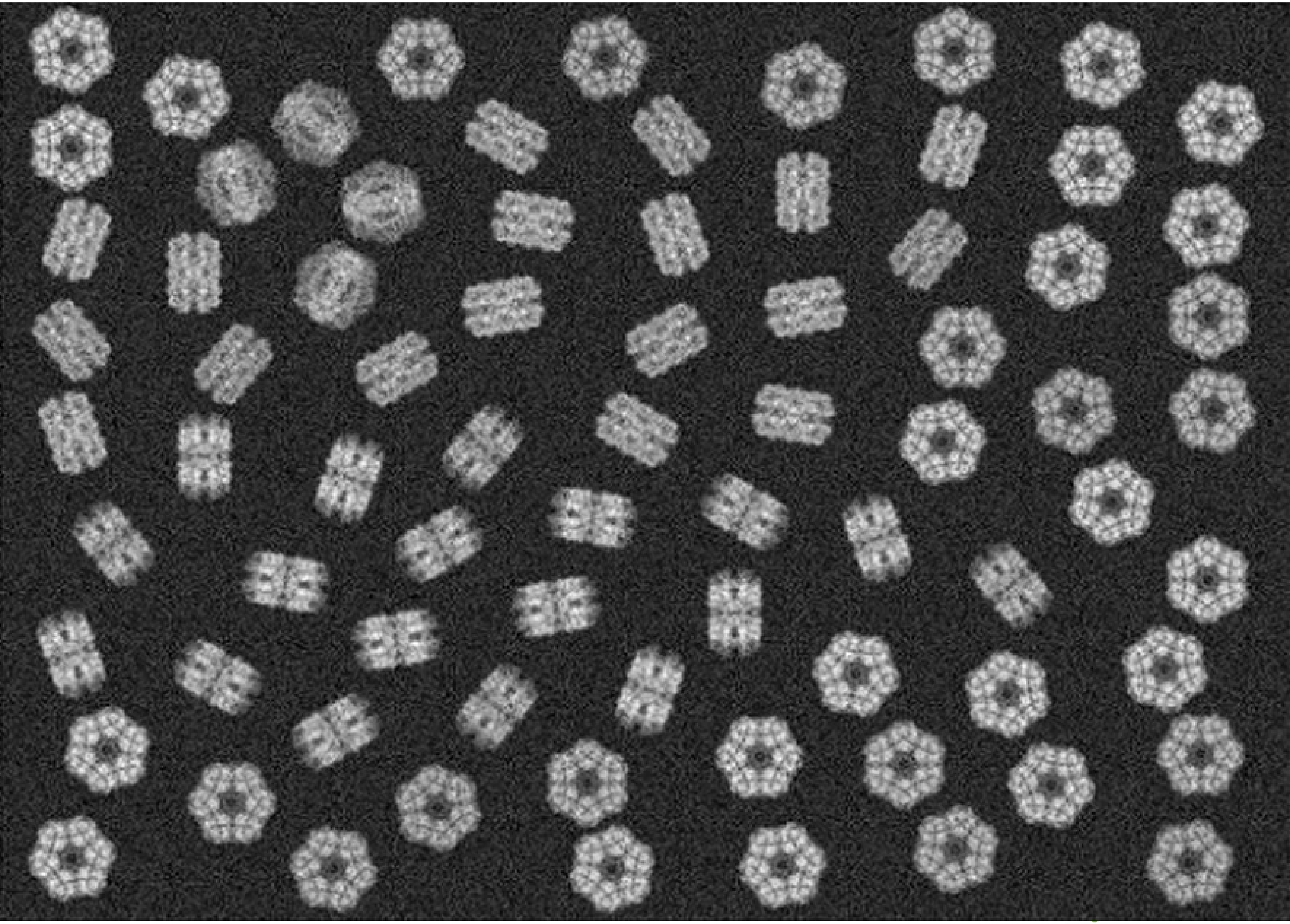
Disadvantage: Alignment depends on the choice of random seed.

Pyramidal/pairwise alignment



Marco... Carrascosa (1996) Ultramicroscopy

You have aligned images,
but they don't all look the same.

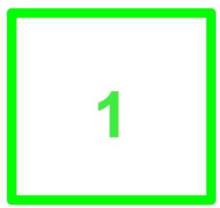


Outline

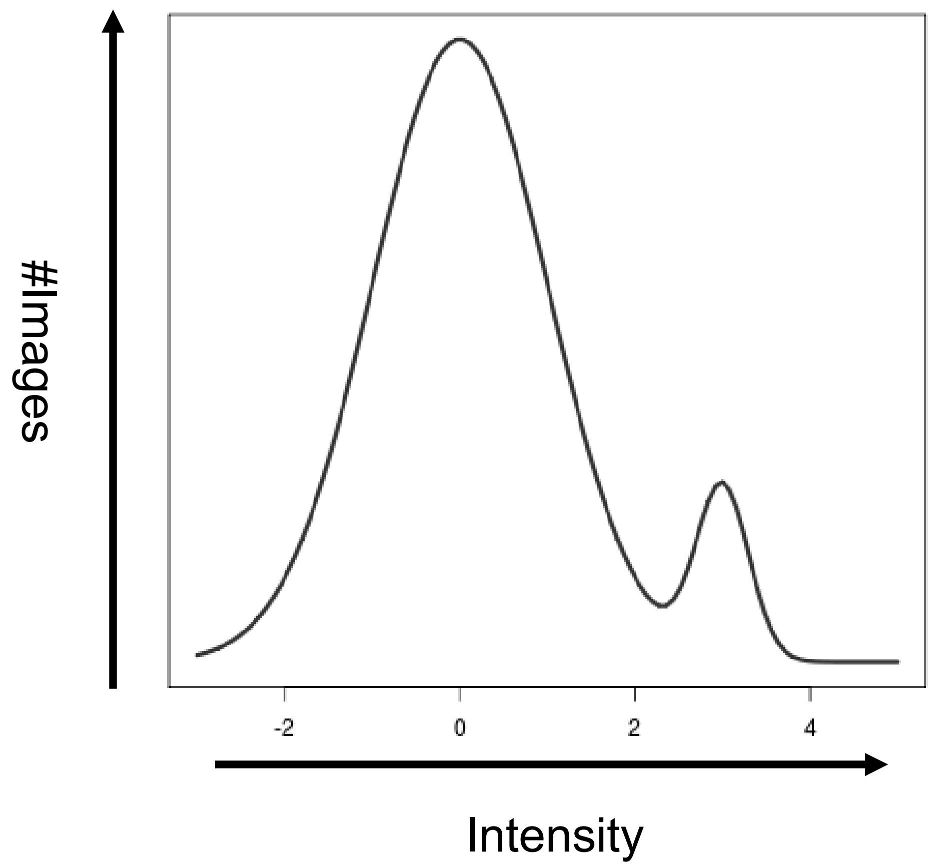
Image analysis II

- Fourier transforms revisited
    - Ducks and other animals
- Analogy to the Ewald sphere
- Aliasing
- Alignment
- Interpolation
- Multivariate data analysis

A one-pixel image

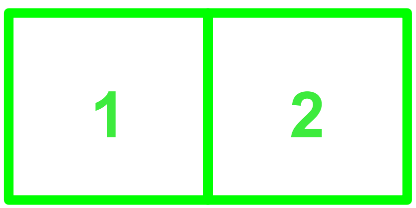


1-pixel image

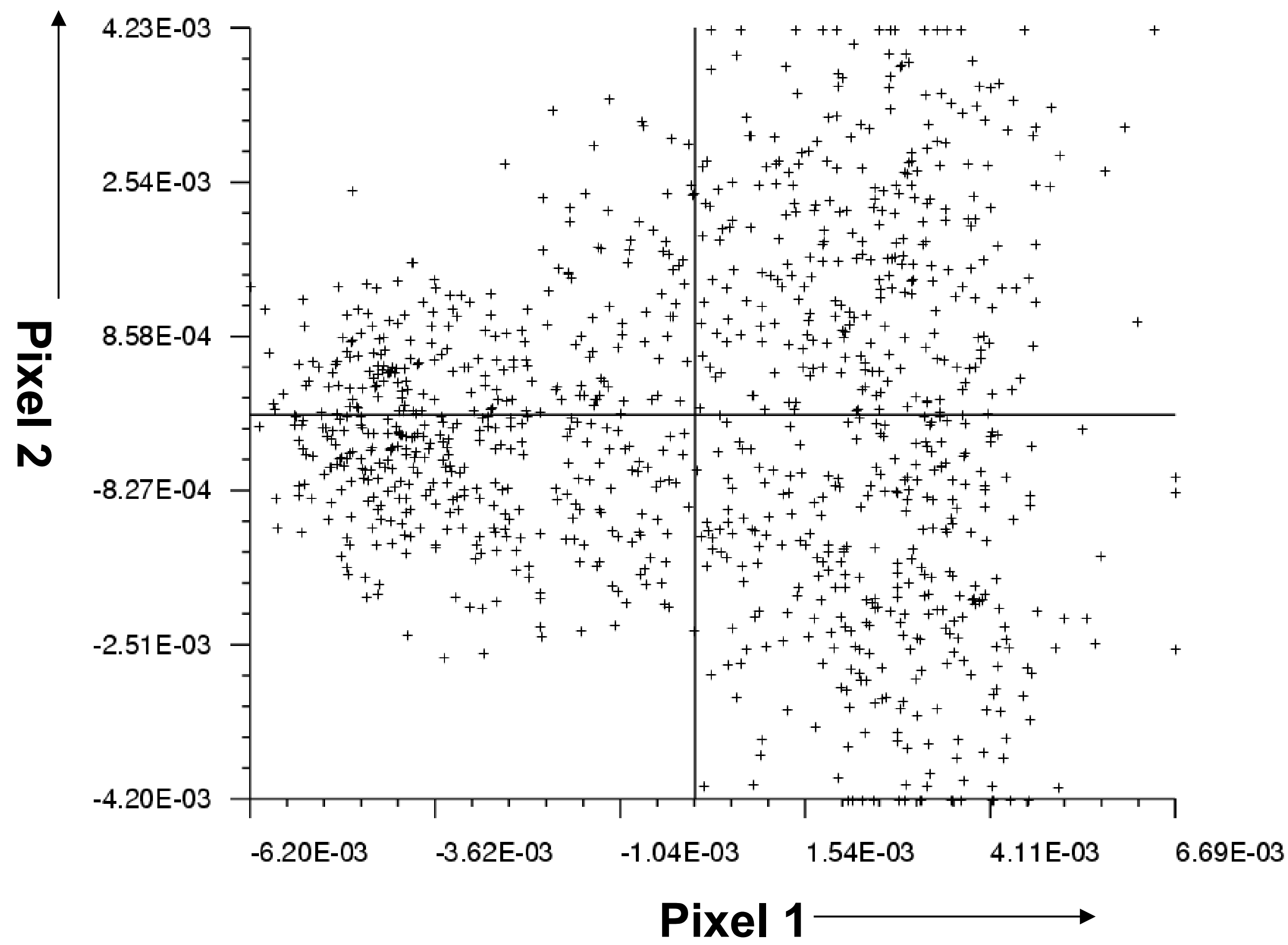


http://isomorphism.es

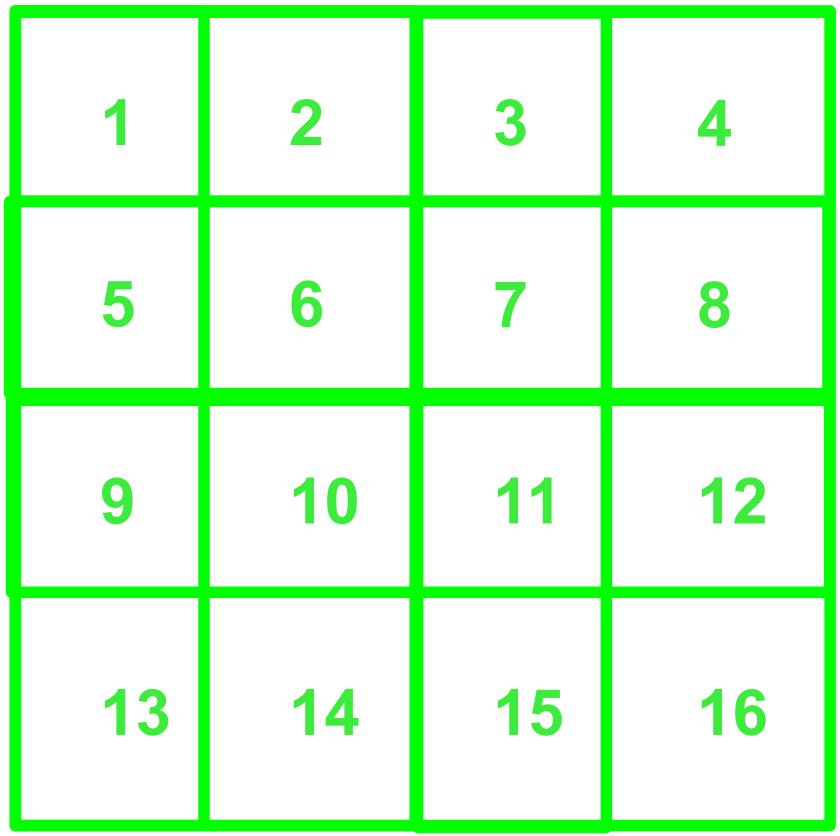
A two-pixel image



2-pixel image

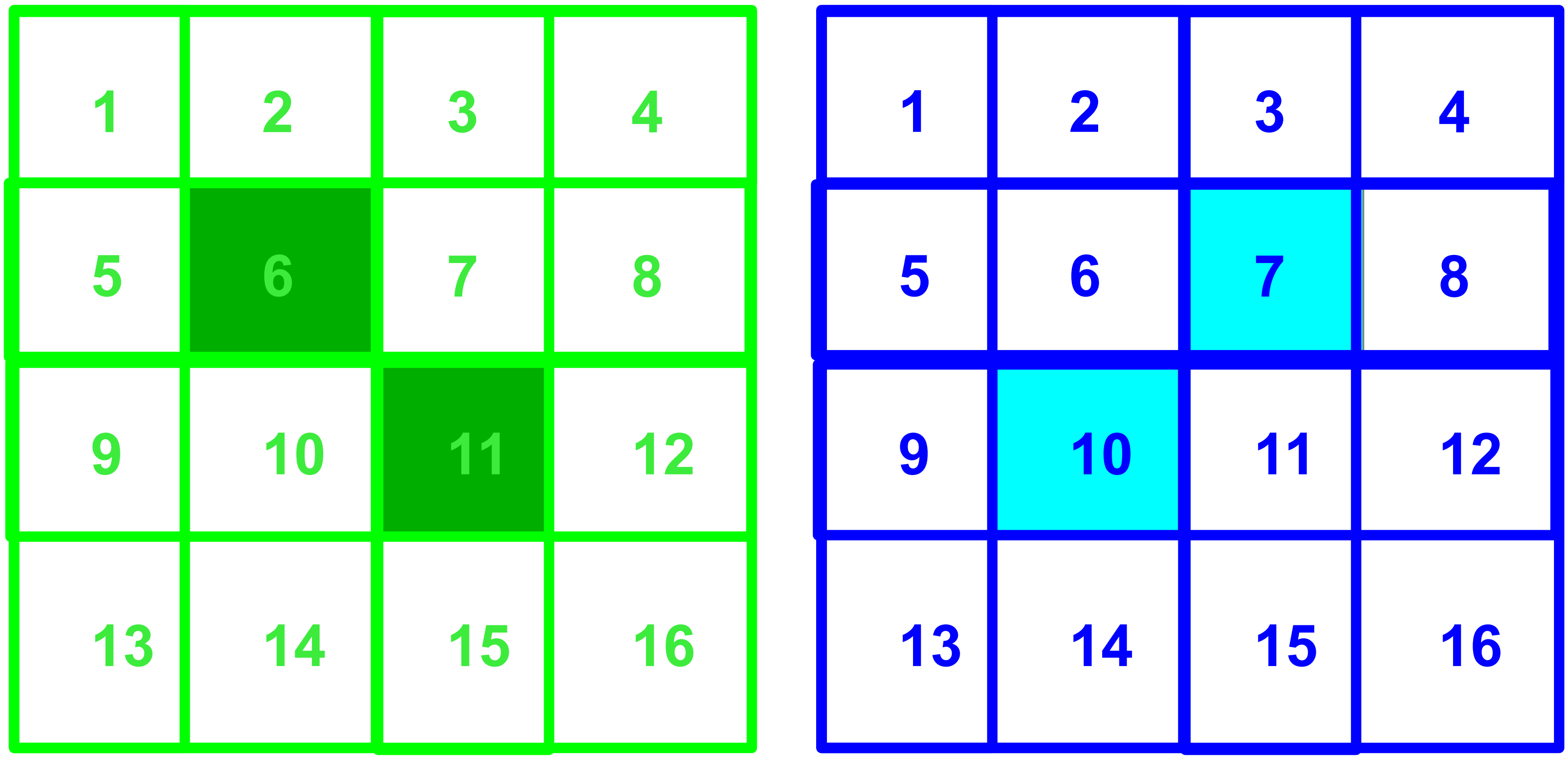


A 16-pixel image



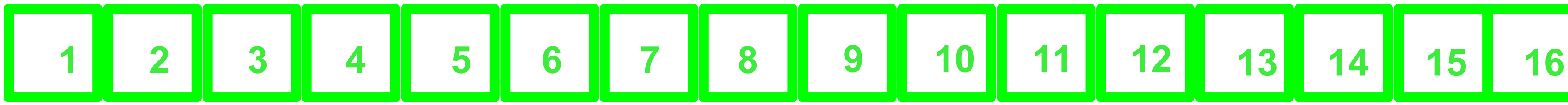
Now, we have a 16-dimensional problem.

Multivariate data analysis (MDA), or Multivariate statistical analysis (MSA)



Suppose pixel 6 coincided with pixel 11,
And pixel 7 coincided with pixel 10.
Then, we're back to two variables, and a 2D problem.

Multivariate data analysis (MDA), or Multivariate statistical analysis (MSA)



Our 16-pixel image can be reorganized into a 16-coordinate vector.

Covariance of measurements *x* and *y*:

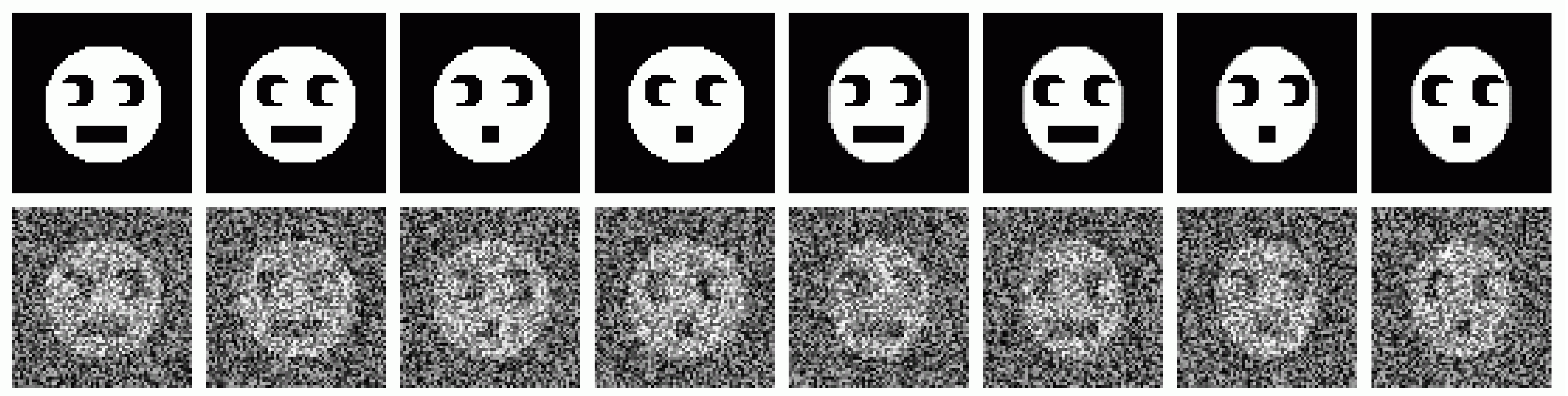
$\langle xy \rangle - \langle x \rangle \langle y \rangle$,

where $\langle x \rangle$ is the mean of *x*.

A high covariance is a measure of the correlation between two variables.

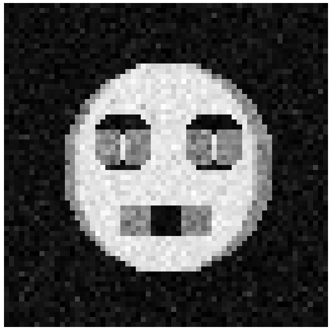
MDA: An example

8 classes of faces, 64x64 pixels



With noise added

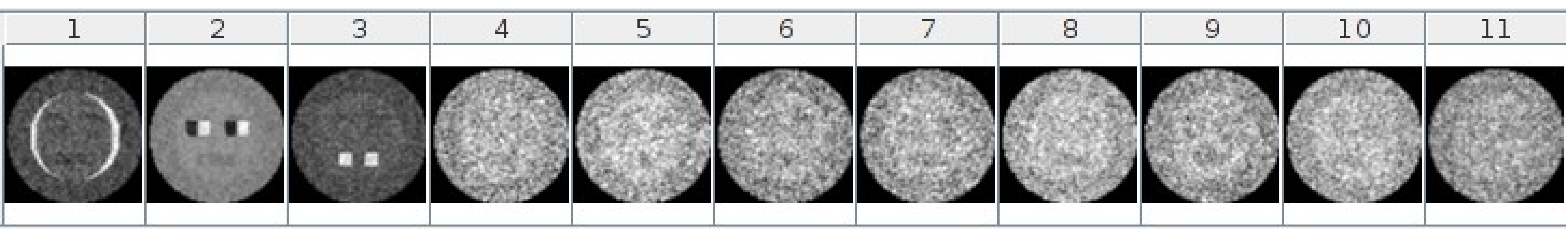
Average:



From http://spider.wadsworth.org/spider_doc/spider/docs/techs/classification/tutorial.html

Principal component analysis (PCA) or Correspondence analysis (CA)

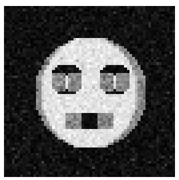
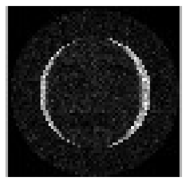
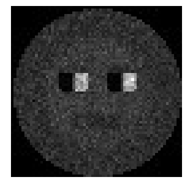
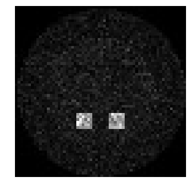
- For a 4096-pixel image, we will have a 4096x4096 covariance matrix.
- Row-reduction of the covariance matrix gives us “eigenvectors.”
  - The eigenvectors describe correlated variations in the data.
  - The eigenvectors have 4096 elements and can be converted back into images, called “eigenimages.”
  - The first eigenvectors will account for the most variation. The later eigenvectors may only describe noise.
  - Linear combinations of these images will give us approximations of the classes that make up the data.



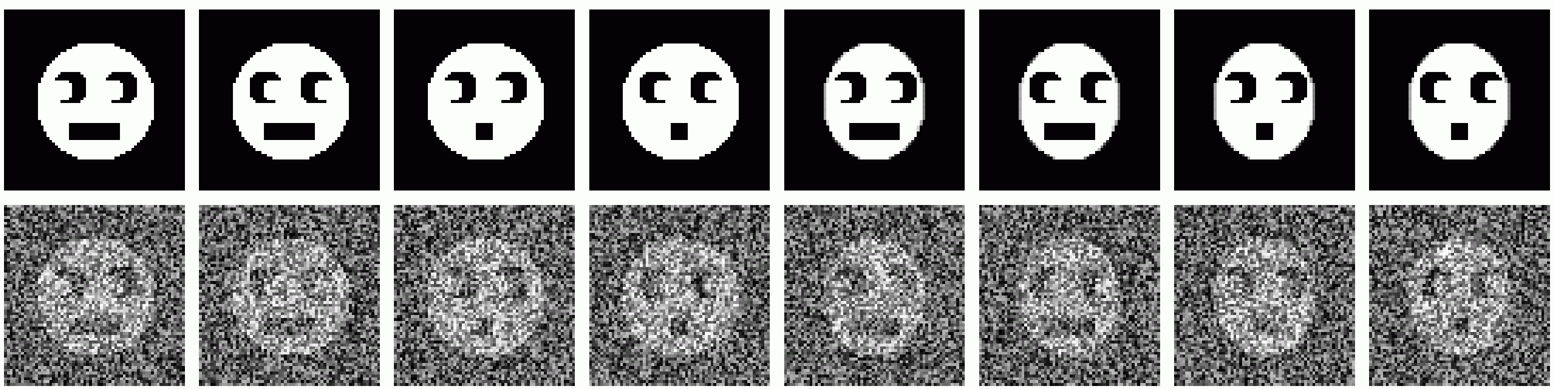
Eigenimages

Reconstituted images

Linear combinations of these images will give us approximations of the classes that make up the data.

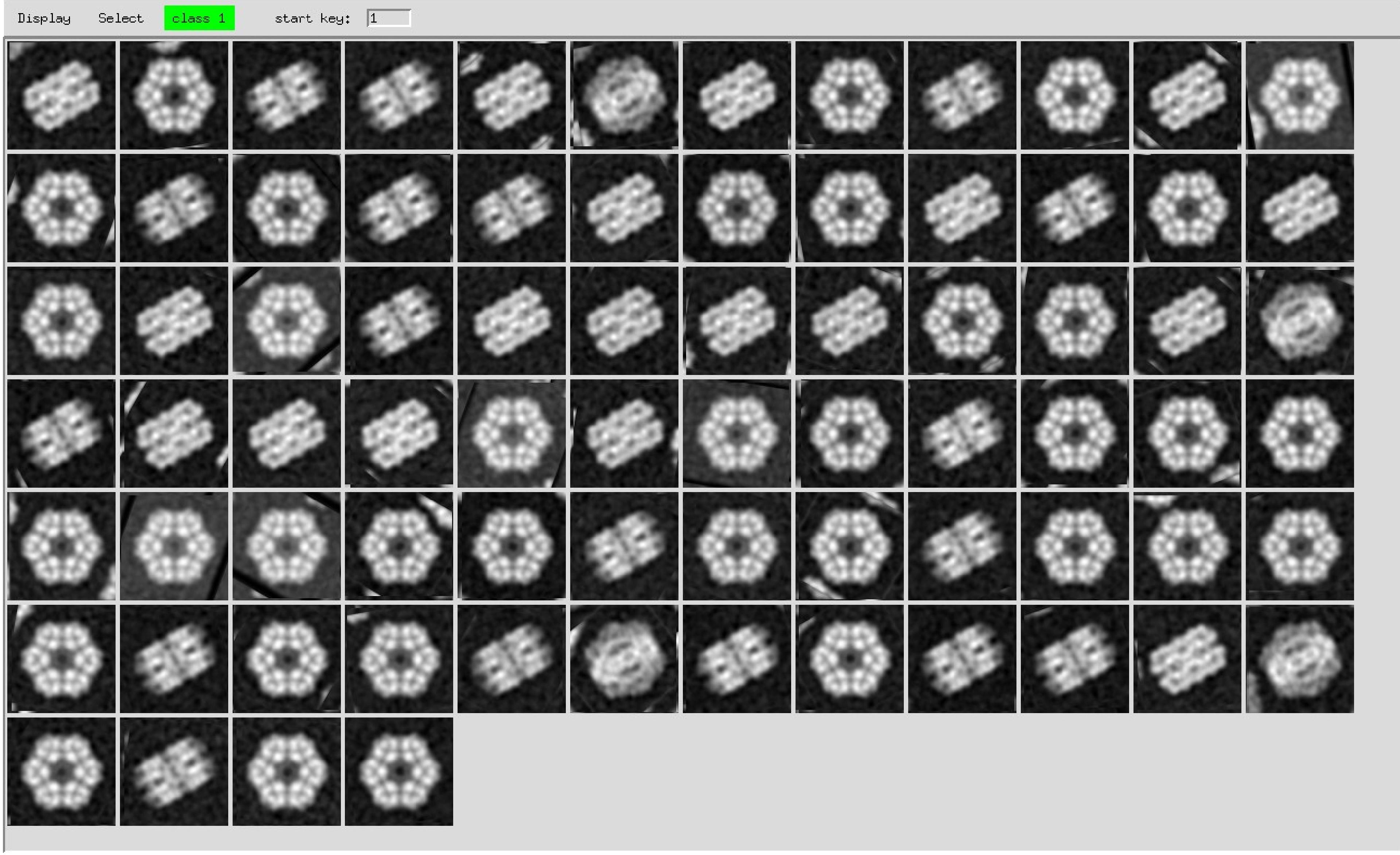
c_0  $+ c_1$  $+ c_2$  $+ c_3$  $+ \ldots$

Average | Eigenimage #1 | Eigenimage #2 | Eigenimage #3



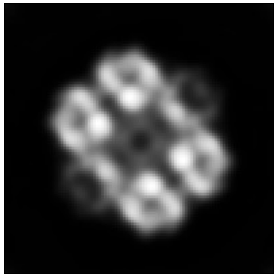
A reminder of what our original images looked like

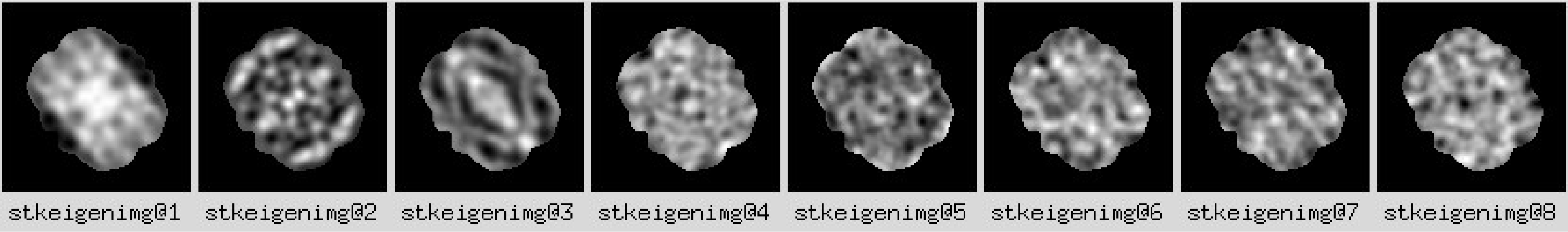
Another example: worm hemoglobin



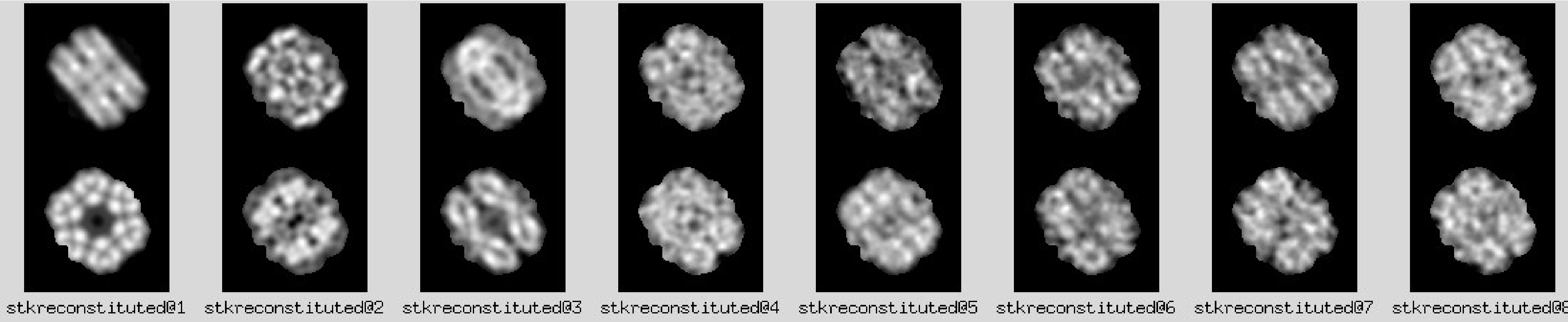
Phantom images of worm hemoglobin

PCA of worm hemoglobin

Average: 



$+c_0$ $+c_1$ $+c_2$ $+c_3$ $+c_4$ $+c_5$



$-c_0$ $-c_1$ $-c_2$ $-c_3$ $-c_4$ $-c_5$

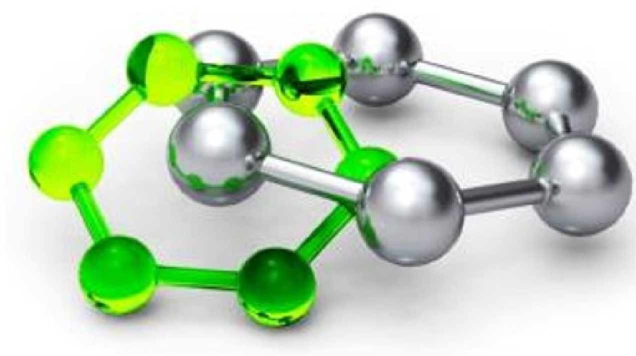
Next week:
Classification & 3D Reconstruction

Thank you for your attention

CEITEC

Central European Institute of Technology
Masaryk University
Kamenice 753/5
625 00 Brno, Czech Republic

www.ceitec.muni.cz | info@ceitec.muni.cz



UNIVERSITAS MASARYKIANA BRUNENSIS

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND
INVESTING IN YOUR FUTURE

2007–13
OP Research and Development for Innovation