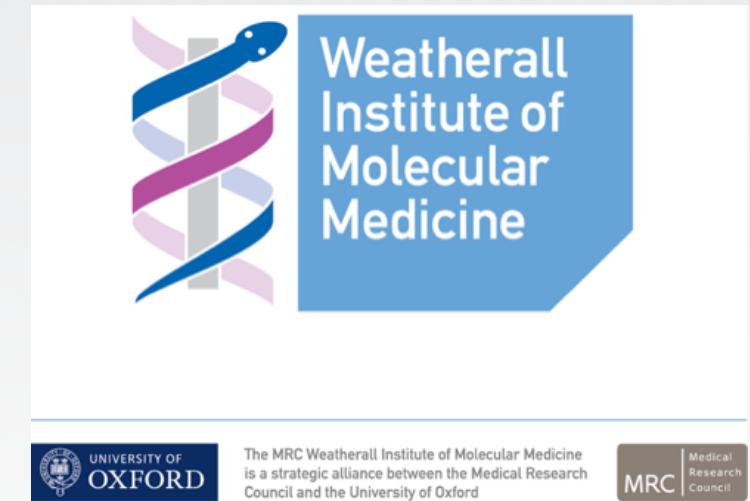


## Correlation in Image Analysis

by Dominic Waithe  
Weatherall Institute of Molecular Medicine

- The dot product.
- Understanding correlation
- Practical colocalisation.
- Registration introduction.



The MRC Weatherall Institute of Molecular Medicine is a strategic alliance between the Medical Research Council and the University of Oxford



Source:

# The dot product of two vectors (algebraic)

in  $\mathbf{R}^{12}$

$a$	$b$
5,	4,
6,	5,
8,	6,
6,	5,
5,	4,
3,	3,
2,	5,
3,	6,
5,	5,
7,	4,
9,	3,
7,	4,

$a$	$\bullet$	$b$
(5	$\times$	4) +
(6	$\times$	5) +
(8	$\times$	6) +
(6	$\times$	5) +
(5	$\times$	4) +
(3	$\times$	3) +
(2	$\times$	5) +
(3	$\times$	6) +
(5	$\times$	5) +
(7	$\times$	4) +
(9	$\times$	3) +
(7	$\times$	4) +

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^n a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Scalar/dot product

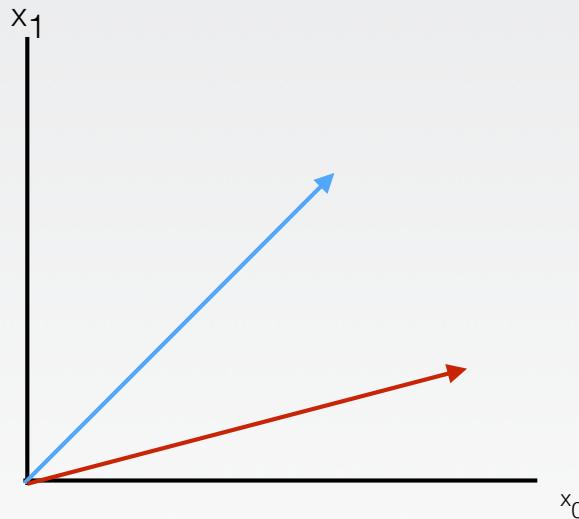
$$= 293$$

Source:

# A vector can be represented on axis.

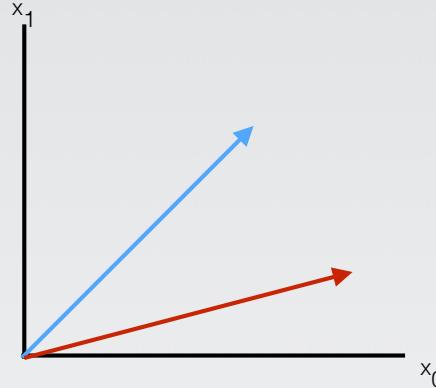
	a	b
$x_0$	30	40
$x_1$	30	20

in  $\mathbb{R}^2$



Source:

# What does the dot product mean?

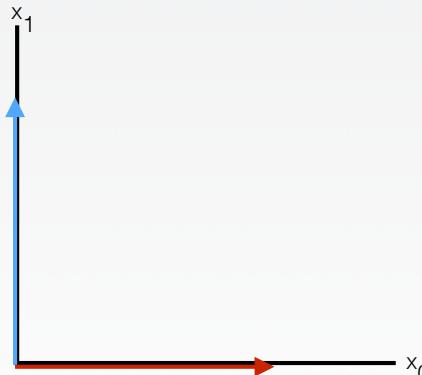


a	b
30	40
30	20

$$\begin{aligned} a \bullet b &= (30 \cdot 40) + (30 \cdot 20) \\ &= 1800 \end{aligned}$$

In Euclidean space a vector has magnitude (length<sup>2</sup>) and direction.

The dot product of the two vectors represent the magnitude within a common dimensional space.



a	b
0	60
60	0

$$\begin{aligned} a \bullet b &= (0 \cdot 60) + (60 \cdot 0) \\ &= 0 \end{aligned}$$

Source:

# The dot product of two vectors (algebraic)

in  $\mathbf{R}^{12}$

$a$	$b$
5,	4,
6,	5,
8,	6,
6,	5,
5,	4,
3,	3,
2,	5,
3,	6,
5,	5,
7,	4,
9,	3,
7,	4,

$a$	$\bullet$	$b$
(5	$\times$	4) +
(6	$\times$	5) +
(8	$\times$	6) +
(6	$\times$	5) +
(5	$\times$	4) +
(3	$\times$	3) +
(2	$\times$	5) +
(3	$\times$	6) +
(5	$\times$	5) +
(7	$\times$	4) +
(9	$\times$	3) +
(7	$\times$	4) +

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^n a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Scalar/dot product

$$= 293$$

Source:

# Motivation for understanding dot product

colocalisation

Fluorescence correlation spectroscopy

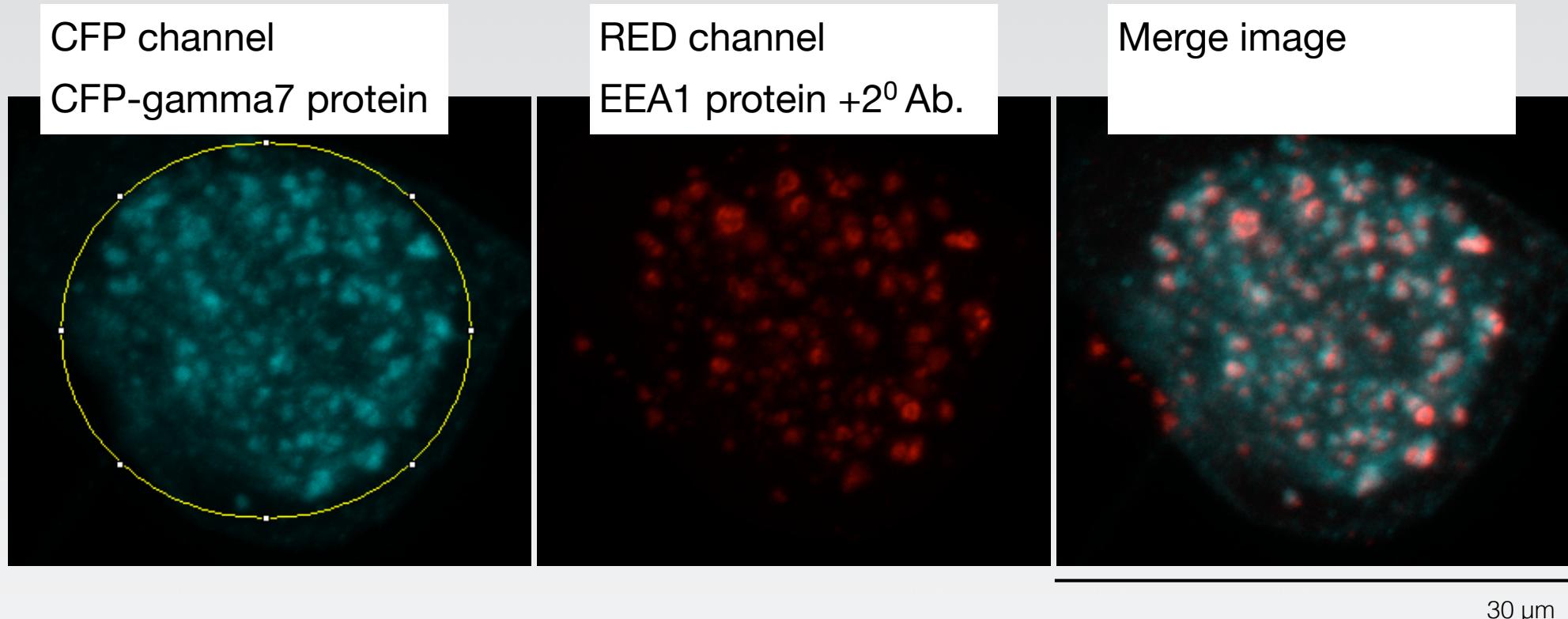
Registration

Template-matching (grayscale)

Convolution

At the core of a lot of  
techniques

# colocalisation

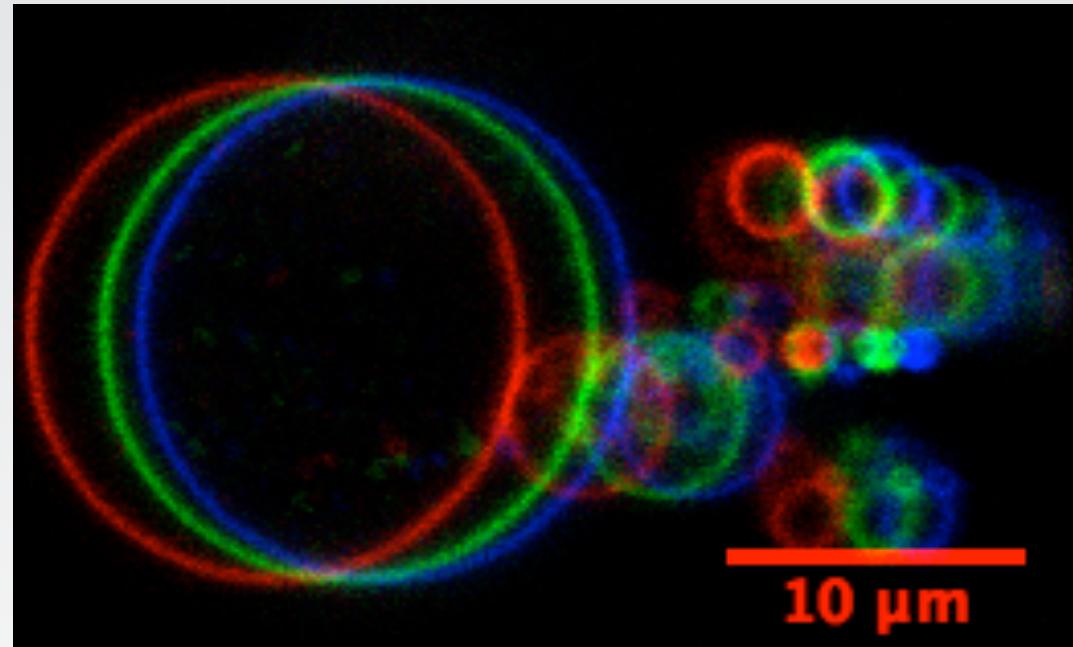
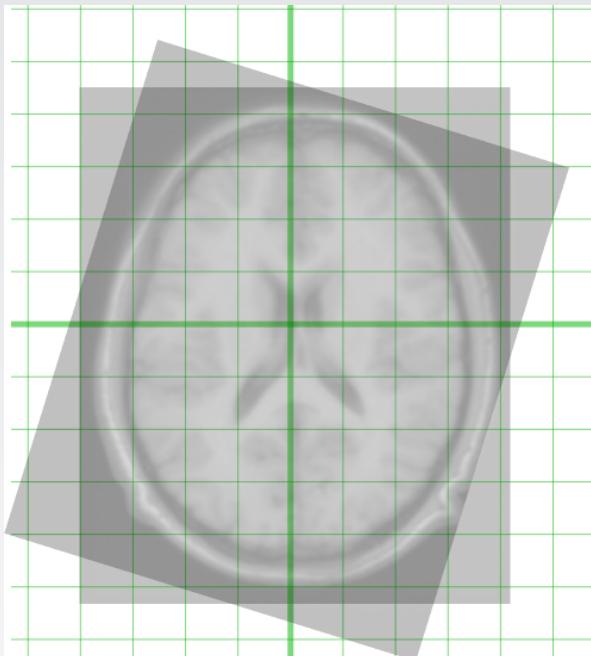


Want to find a measure of similarity between the expression patterns of the two proteins within this cell. For this we use correlation (colocalisation) to compare the intensity distribution in the different channels and measure the similarity of the two distributions.

Source:

# Registration in imaging

We want to align images by analysing the similarity of the different images (or channels) in different spatial locations (often using correlation as our similarity measure).



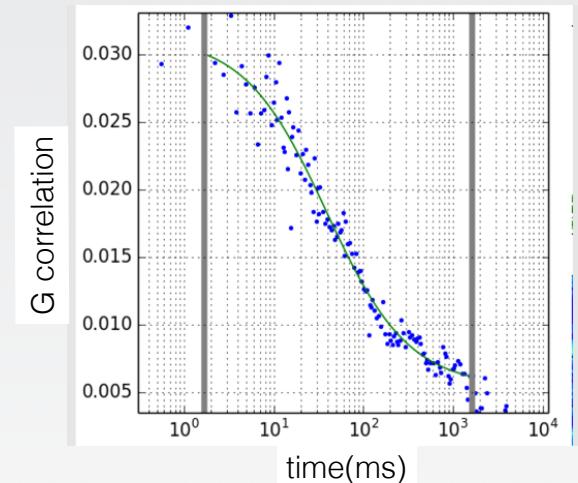
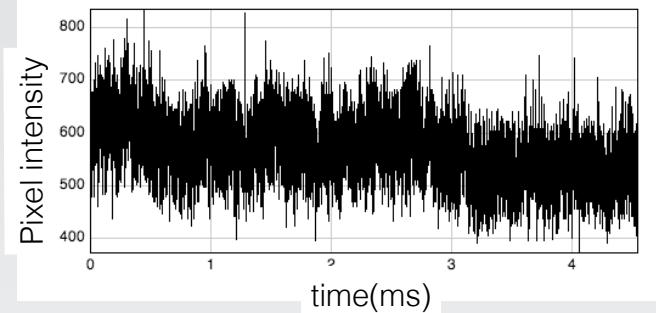
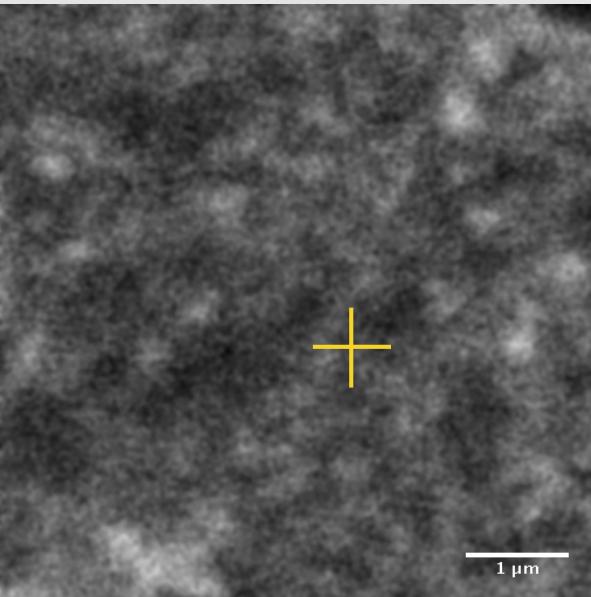
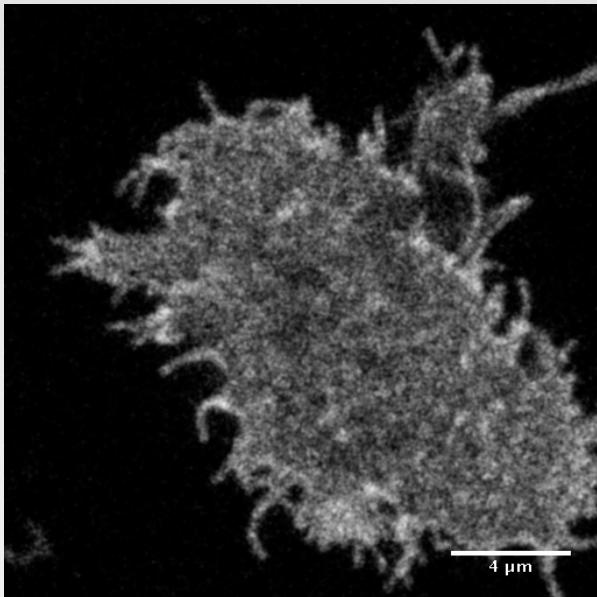
Spatial shifts in medical sciences (e.g. different imaging modalities, different patients)  
Spatial shifts in microscopy (eg. chromatic aberration, drift during experimentation).

Source: [http://en.wikipedia.org/wiki/Image\\_registration#mediaviewer/File:Mni-autoreg\\_03-registered.png](http://en.wikipedia.org/wiki/Image_registration#mediaviewer/File:Mni-autoreg_03-registered.png)

# Temporal correlation



e.g. Fluorescence Correlation Spectroscopy



$$R(\tau) = \frac{E[(X_t - \mu)(X_{t+\tau} - \mu)]}{\sigma^2},$$

where 'mu' is the mean of Xt and sigma^2 is your standard deviation.

By fitting the correlation data we are able to calculate the rate at which the protein of interest is diffusing in the specimen.

Source: Jorge Bernardino de la serna, Christian Eggeling

# Grayscale template matching



*Template image*



*Input image*



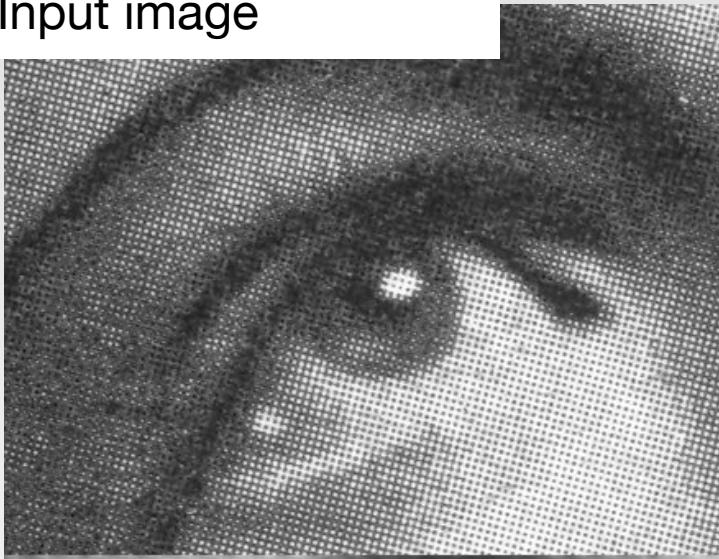
*Results of multi-angle matching*

Grayscale template matching uses correlation to establish location of matches of a template and an image. Matches represent areas of highest correlation.

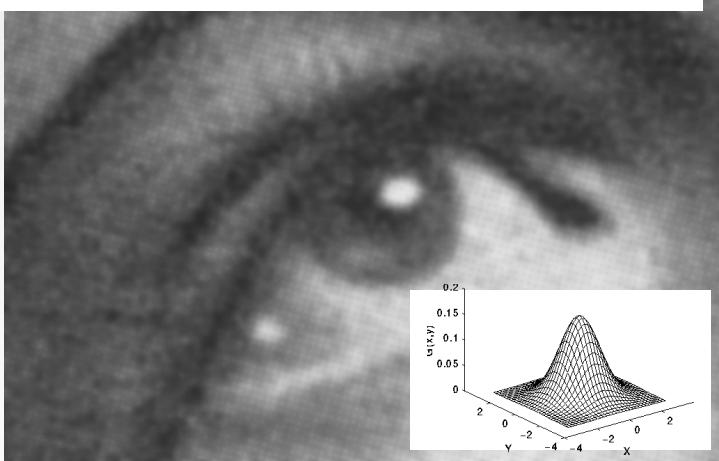
Source: [http://docs.adaptive-vision.com/current/studio/machine\\_vision\\_guide/TemplateMatching.html](http://docs.adaptive-vision.com/current/studio/machine_vision_guide/TemplateMatching.html)

# Dot product and convolution

Input image



gaussian smoothed

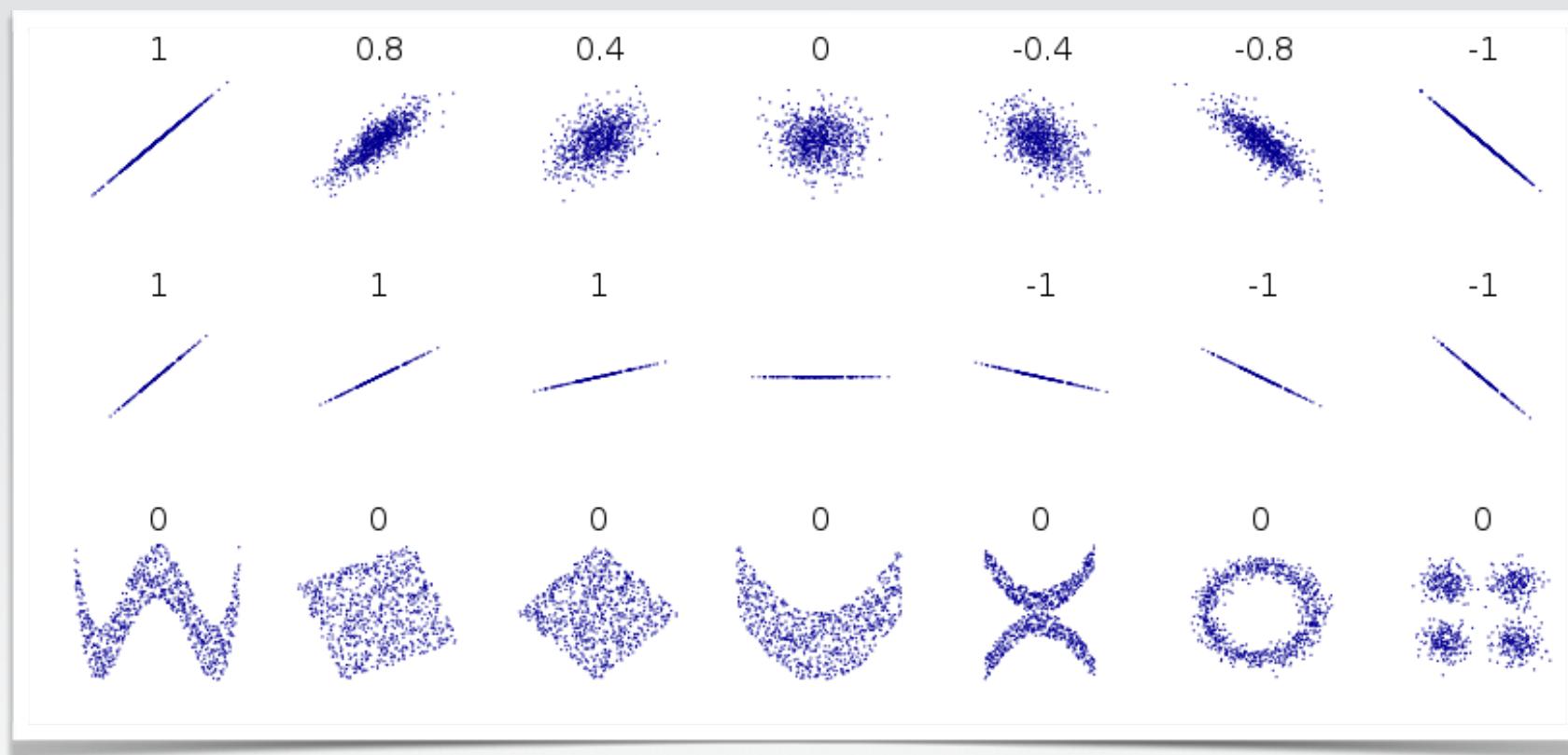


Discrete deconvolution and correlation are almost the same thing\*.

The output of a filter convolution represents the similarity function at each pixel of the filter and the input image.

Source: <http://en.wikipedia.org/wiki/Convolution>

\*If the filter is symmetrical.



## Understanding correlation

Source: [https://en.wikipedia.org/wiki/Correlation\\_and\\_dependence](https://en.wikipedia.org/wiki/Correlation_and_dependence)

# Why can't we always use dot product.

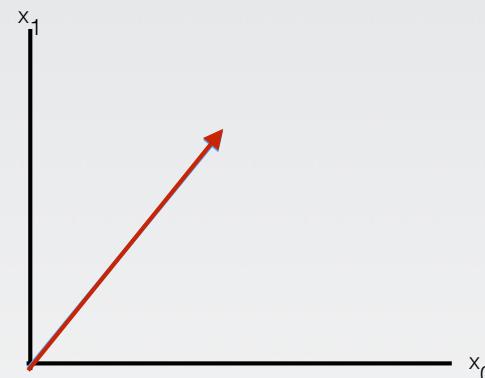
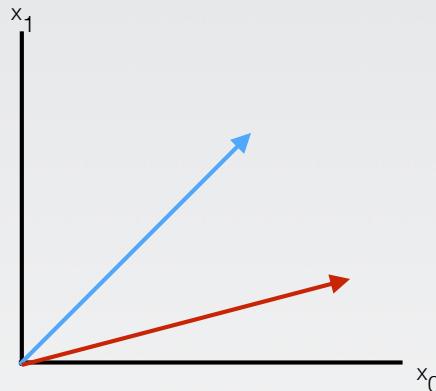
in  $\mathbb{R}^2$

a	b
30	40
30	20

$$\begin{aligned} a \cdot b &= (30 \cdot 40) + (30 \cdot 20) \\ &= 1800 \end{aligned}$$

a	b
30	30
30	30

$$\begin{aligned} a \cdot b &= (30 \cdot 30) + (30 \cdot 30) = 1800 \end{aligned}$$



In the second example the magnitude is the same as in the first?

The magnitude of the dot product alone is not an accurate method to establish the similarity of two vectors as the vectors can differ in length.

A better measure of similarity is the angle between the two vectors.

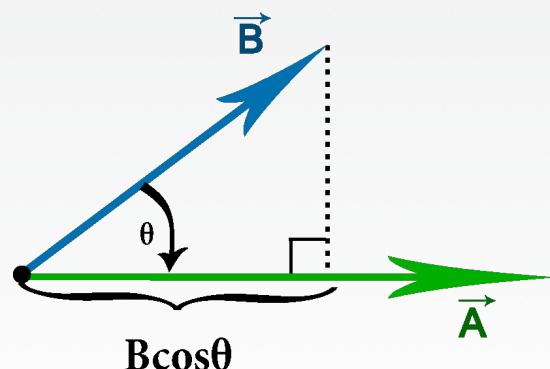
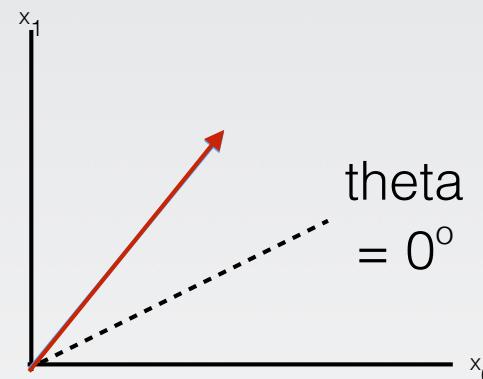
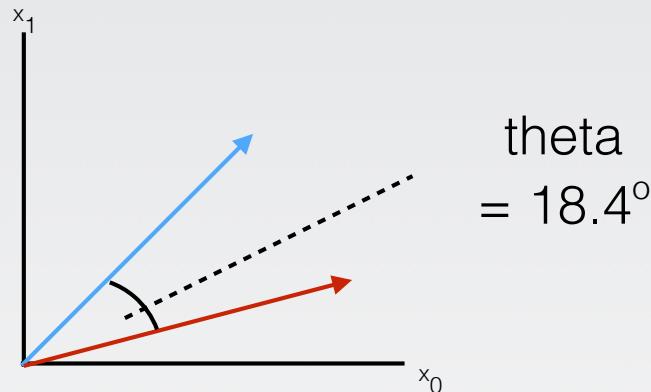
Source:

# Geometric dot product

in  $\mathbb{R}^2$

$$\begin{array}{|c|c|} \hline a & b \\ \hline 30 & 40 \\ \hline 30 & 20 \\ \hline \end{array}$$
$$a \cdot b = (30 \cdot 40) + (30 \cdot 20) = 1800$$

$$\begin{array}{|c|c|} \hline a & b \\ \hline 30 & 30 \\ \hline 30 & 30 \\ \hline \end{array}$$
$$a \cdot b = (30 \cdot 30) + (30 \cdot 30) = 1800$$



We use the relationship between the algebraic and geometric dot product

$$\mathbf{A} \cdot \mathbf{B} = \|\mathbf{A}\| \|\mathbf{B}\| \cos \theta,$$

$$\|\mathbf{A}\| = \sqrt{\mathbf{A} \cdot \mathbf{A}}$$

# Pearson's product-moment correlation test.

$$\mathbf{A} \cdot \mathbf{B} = \|\mathbf{A}\| \|\mathbf{B}\| \cos \theta,$$

Pearson's equation:

$$r = \frac{\sum (R_i - \bar{R}) \times (G_i - \bar{G})}{\sqrt{\sum (R_i - \bar{R})^2 \times \sum (G_i - \bar{G})^2}}$$

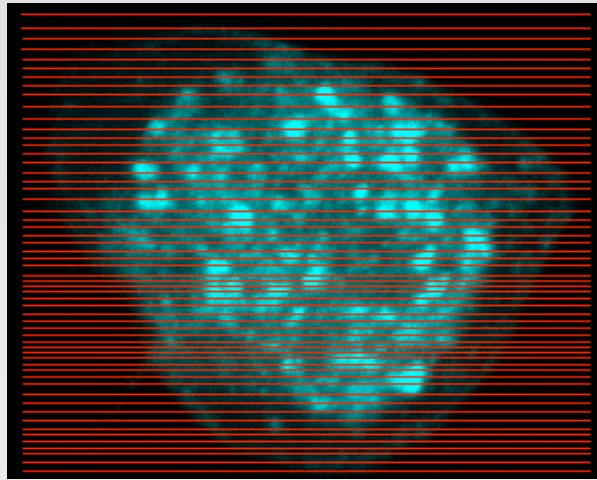
centred vector

r is the  $\cos(\theta)$

magnitude of our centred vectors

Source: [http://en.wikipedia.org/wiki/Correlation\\_coefficient](http://en.wikipedia.org/wiki/Correlation_coefficient)

# Image (2d array) to list (1d)



=

[ 93], [ 23], [ 23], [155], [155], [155]	155
[ 107], [198], [198], [140], [140], [140]	140
[ 121], [ 11], [ 11], [ 7], [ 7], [ 7]	7
[ 135], [235], [235], [198], [198], [198]	198
[ 149], [114], [114], [213], [213], [213]	213
[ 163], [187], [187], [ 9], [ 9], [ 9]	9
[ 8], [ 80], [ 80], [150], [150], [150]	150
[ 22], [187], [187], [ 20], [ 20], [ 20]	20
[ 16], [165], [165], [111], [111], [111]	111
[ 158], [ 15], [ 15], [ 34], [ 34], [ 34]	34
[ 200], [120], [120], [ 69], [ 69], [ 69]	69
	155
	140
	7
	198
	213
	9
	150
	20
	111
	34
	69
	etc
	etc

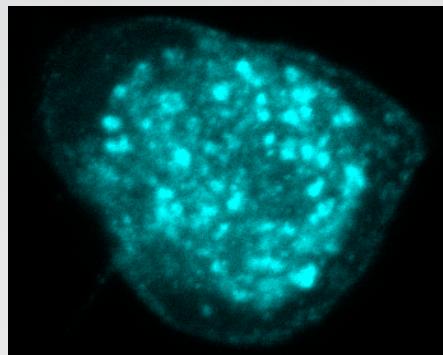
=

Yes. We take our image and represent it as a very long list of pixel intensities.

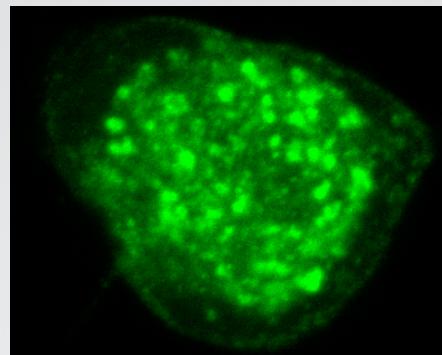
In python we can do this by reshaping our image array (e.g. `im.reshape(-1)`).

Source:

# Correlation of two images to find fluorescence correlation



VS

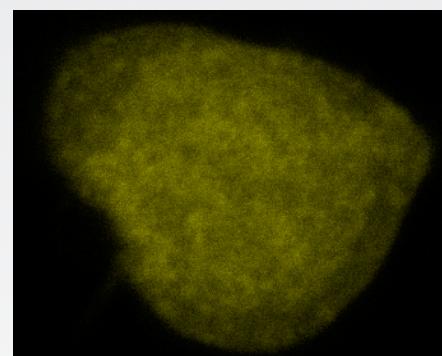


correlation

0.8



VS



dot product total:

0.2

Images which have similar ‘spatial’ distribution of pixel values will be highly-correlated. We can use this to establish colocalisation.

# Pearson's product-moment correlation test.

Pearson's equation:

$$r = \frac{\sum (R_i - \bar{R}) \times (G_i - \bar{G})}{\sqrt{\sum (R_i - \bar{R})^2 \times \sum (G_i - \bar{G})^2}}$$

if  $r$  is 1.0 means correlation  
if  $r$  is close to '0.0' no correlation.  
if  $r$  is -1.0 it means anti-correlation.

$R$  refers to one channel,  $G$  refers to Green channel.  $G$  or  $R$  with a bar refers to mean intensity in that channel. 'i' refers to each pixel in image. Sigma (big E) refers to sum. So sum of all pixels minus their mean.



VS  
high

dot product total:



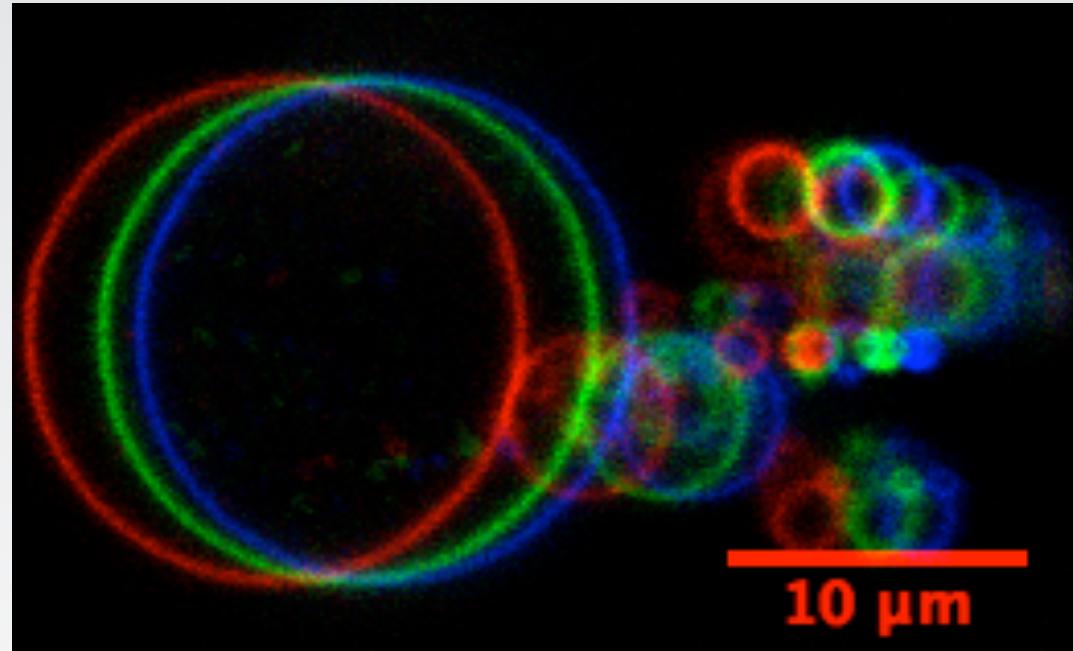
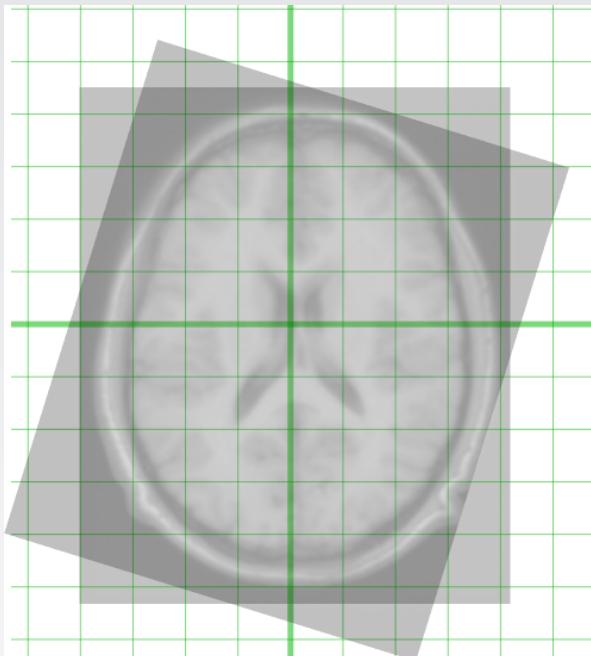
VS  
low

Dimensionless and normalised comparison. Can be used on any two images as long as they are the same spatial size and don't have too many black pixels

Source: [http://en.wikipedia.org/wiki/Correlation\\_coefficient](http://en.wikipedia.org/wiki/Correlation_coefficient)

# Registration in imaging

We want to align images by analysing the similarity of the different images (or channels) in different spatial locations (often using correlation as our similarity measure).



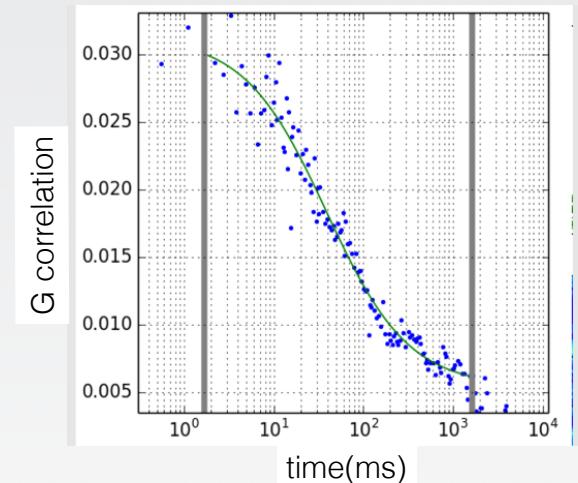
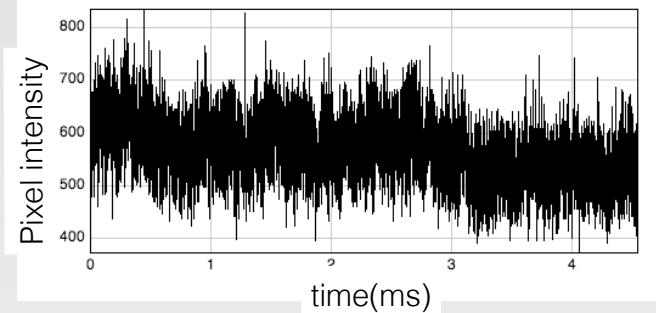
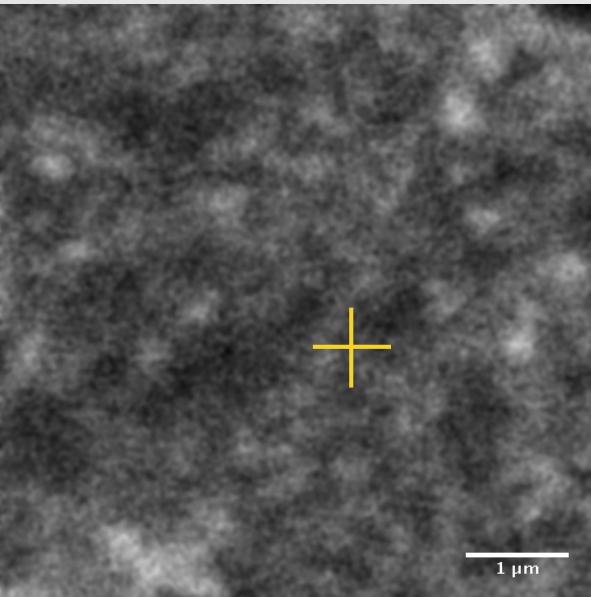
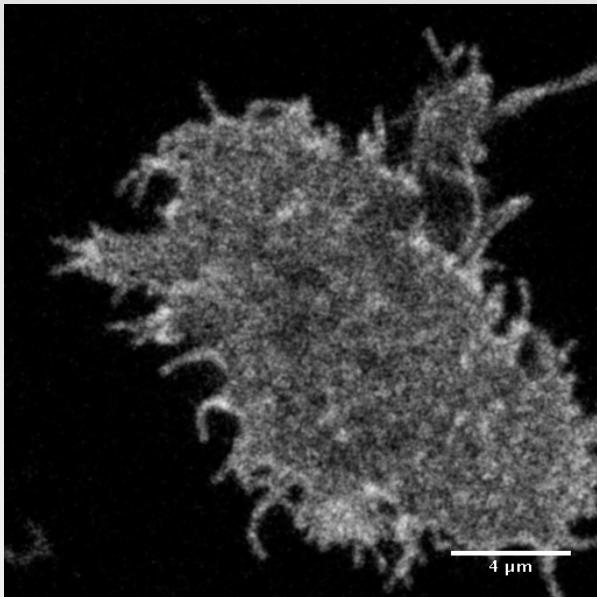
Spatial shifts in medical sciences (e.g. different imaging modalities, different patients)  
Spatial shifts in microscopy (eg. chromatic aberration, drift during experimentation).

Source: [http://en.wikipedia.org/wiki/Image\\_registration#mediaviewer/File:Mni-autoreg\\_03-registered.png](http://en.wikipedia.org/wiki/Image_registration#mediaviewer/File:Mni-autoreg_03-registered.png)

# Temporal correlation



e.g. Fluorescence Correlation Spectroscopy



$$R(\tau) = \frac{E[(X_t - \mu)(X_{t+\tau} - \mu)]}{\sigma^2},$$

where 'mu' is the mean of Xt and sigma^2 is your standard deviation.

By fitting the correlation data we are able to calculate the rate at which the protein of interest is diffusing in the specimen.

Source: Jorge Bernardino de la serna, Christian Eggeling

# Grayscale template matching



*Template image*



*Input image*



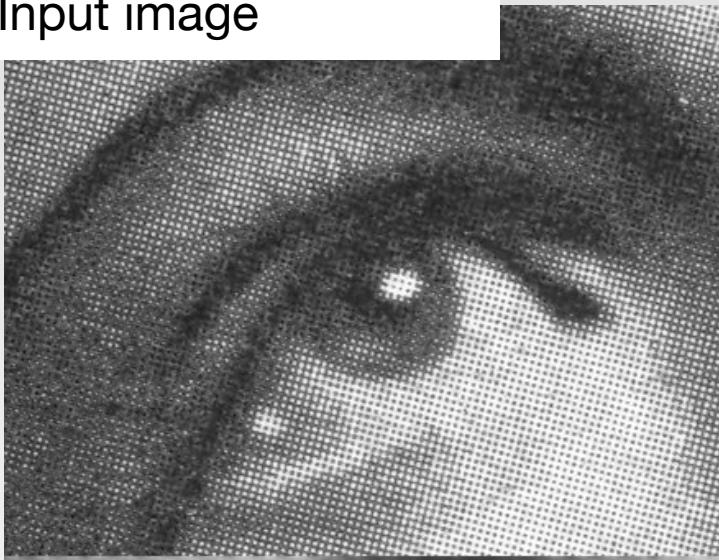
*Results of multi-angle matching*

Grayscale template matching uses correlation to establish location of matches of a template and an image. Matches represent areas of highest correlation.

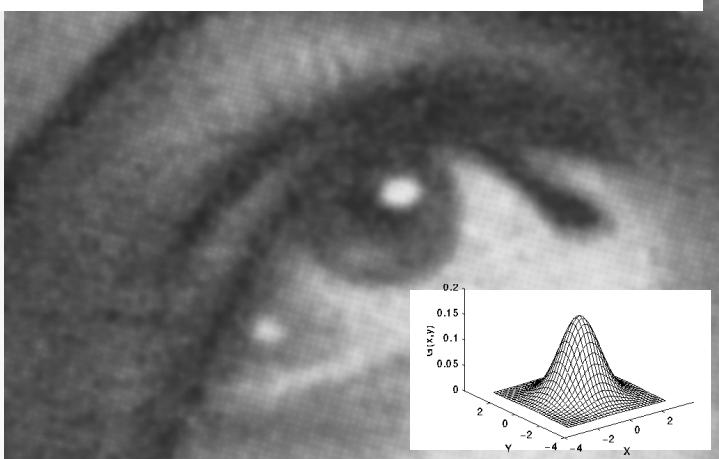
Source: [http://docs.adaptive-vision.com/current/studio/machine\\_vision\\_guide/TemplateMatching.html](http://docs.adaptive-vision.com/current/studio/machine_vision_guide/TemplateMatching.html)

# Dot product and convolution

Input image



gaussian smoothed

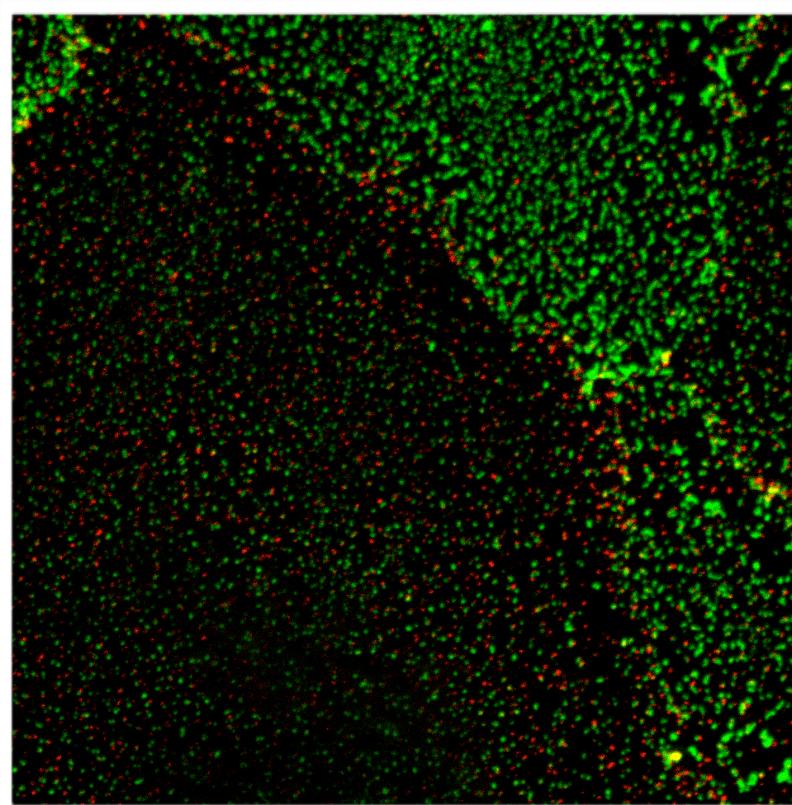


Discrete convolution and correlation are almost the same thing\*.

The output of a filter convolution represents the similarity function at each pixel of the filter and the input image.

Source: <http://en.wikipedia.org/wiki/Convolution>

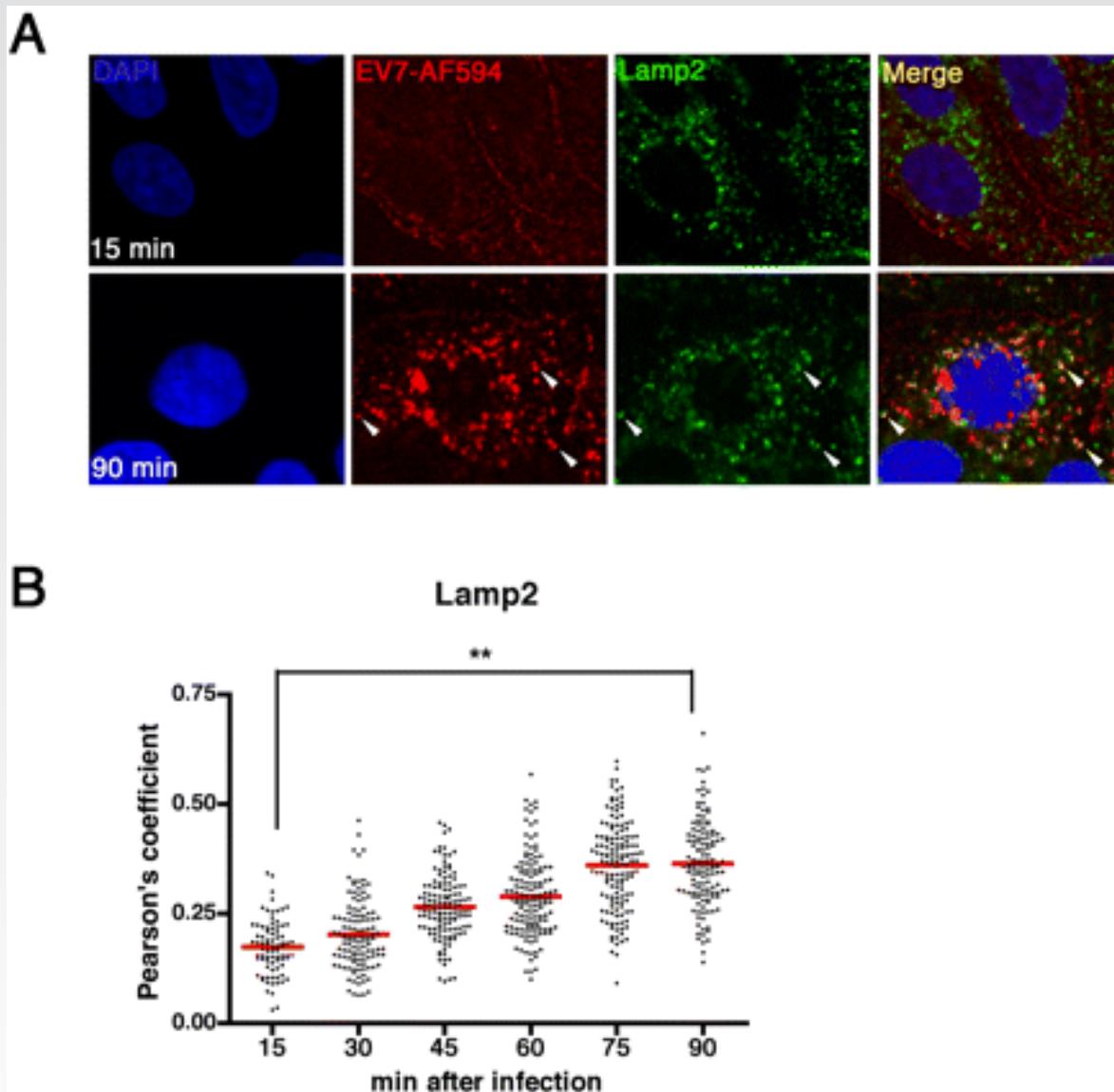
\*If the filter is symmetrical.



## Practical colocalisation

Source: Tess Stanley

# A good example of pearson's test for colocalisation



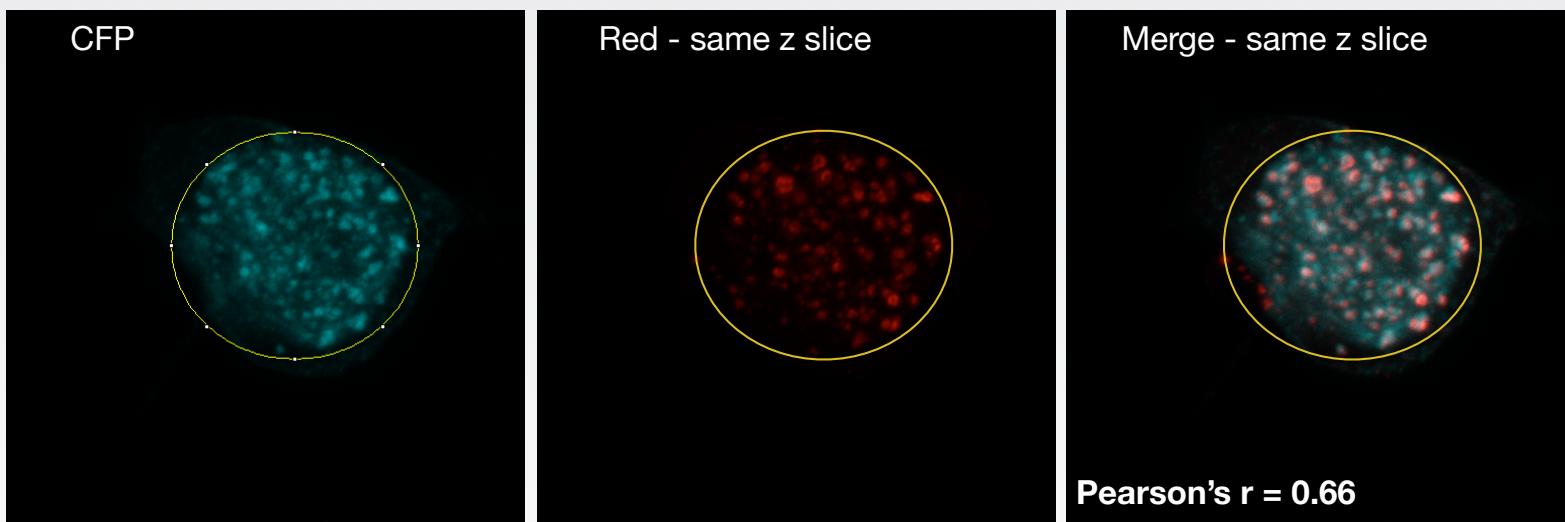
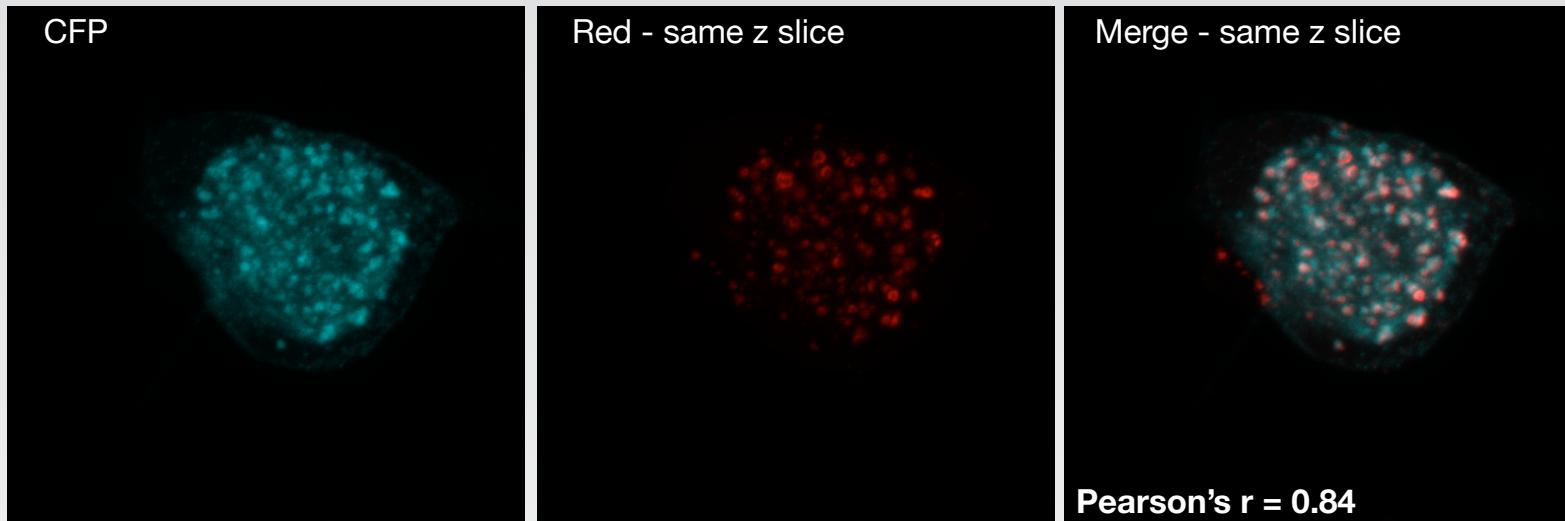
Comparing EV7 staining with Lamp2 over time.

Source: <http://mbio.asm.org/content/3/2/e00304-11/F5.expansion.html>

MRC Weatherall Institute of Molecular Medicine - Wolfson Centre for Imaging - Nano Group-CDT image analysis

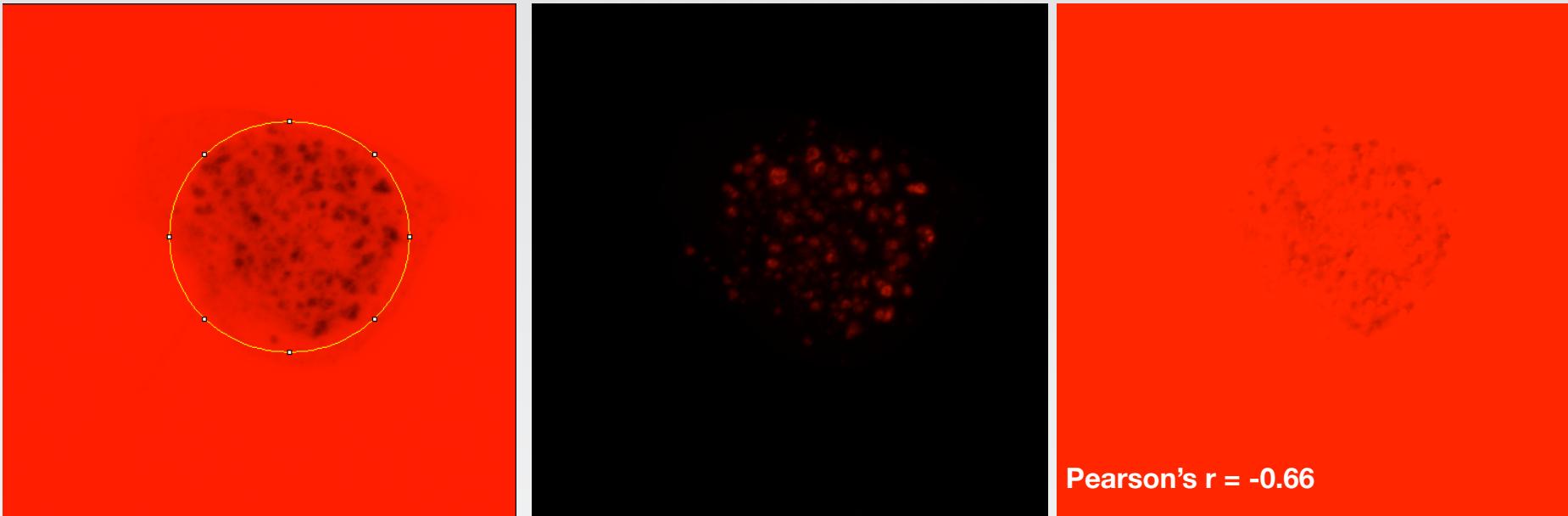
24 Waite 2015

# P's test is sensitive zero pixels and saturation



- Pearson's test doesn't ignore '0' pixels and noise within calculation.
- Coloc 2 plugin does warn you however: The ratio between zero-zero pixels and other pixels is larger 0.37. Maybe you should use a ROI.

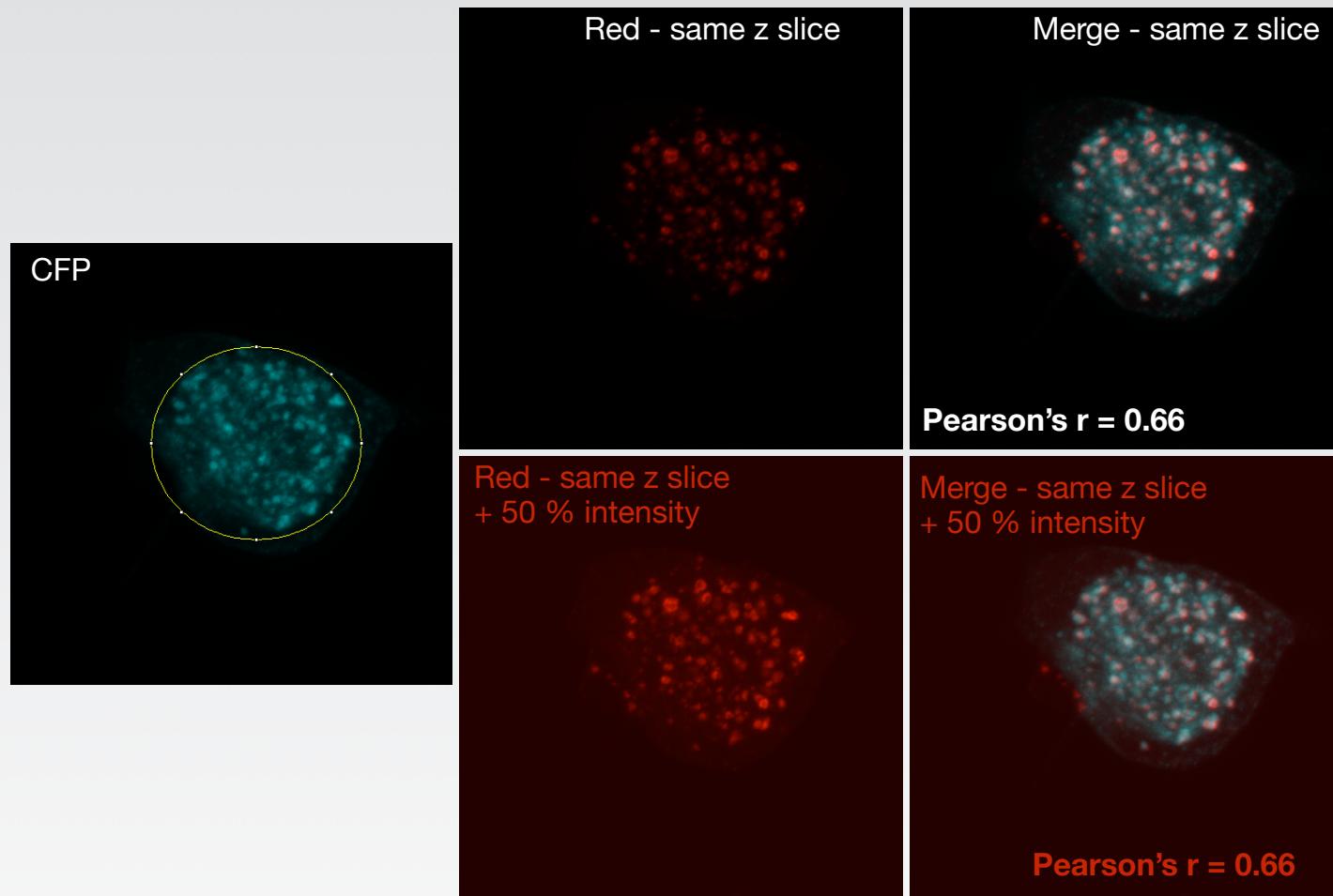
# Pearson's test and anti-correlation



- Pearson's test can also be used to establish when something is negatively colocalised. Can be shown by taking inverse of input image.
- For when something is being actively excluded from an area.

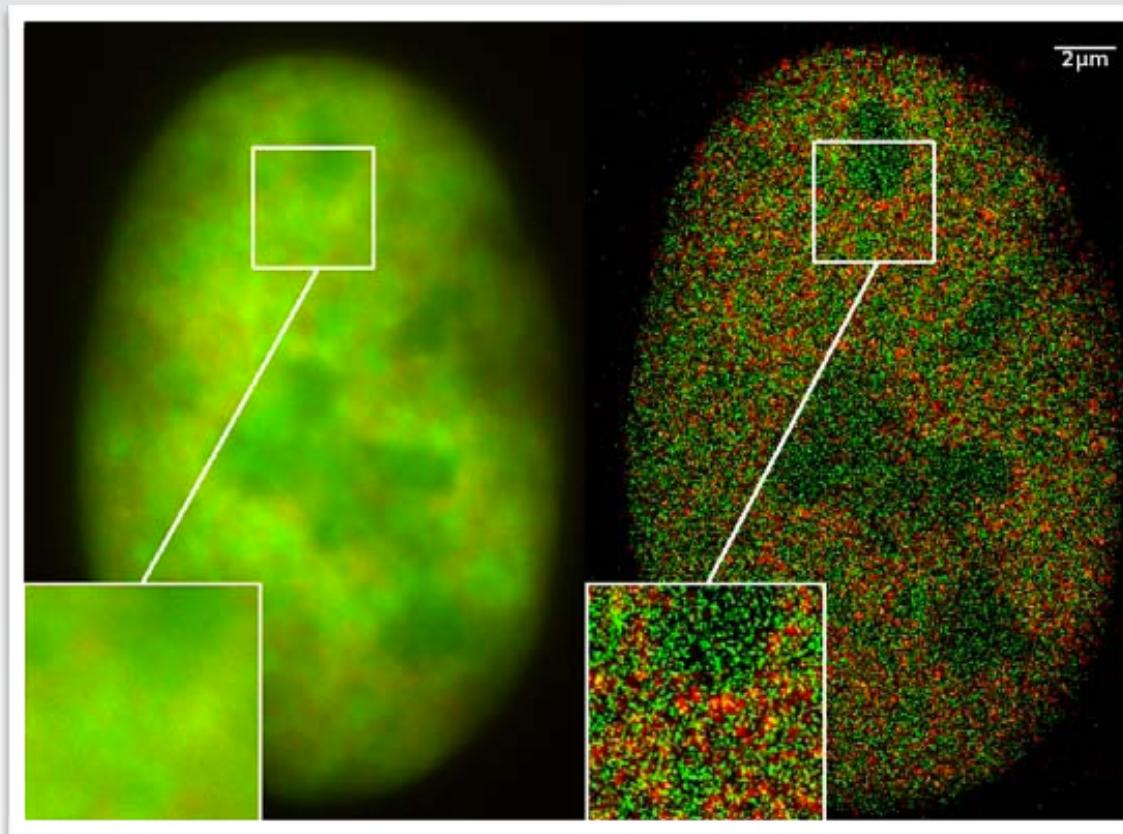
Source:

# Pearson's test is insensitive to global intensity



- Pearson's test is (within reason) insensitive to linear changes in intensity.
- This is good, it looks at trends rather than absolute values.
- This means expression variation between cells does not ruin experiment

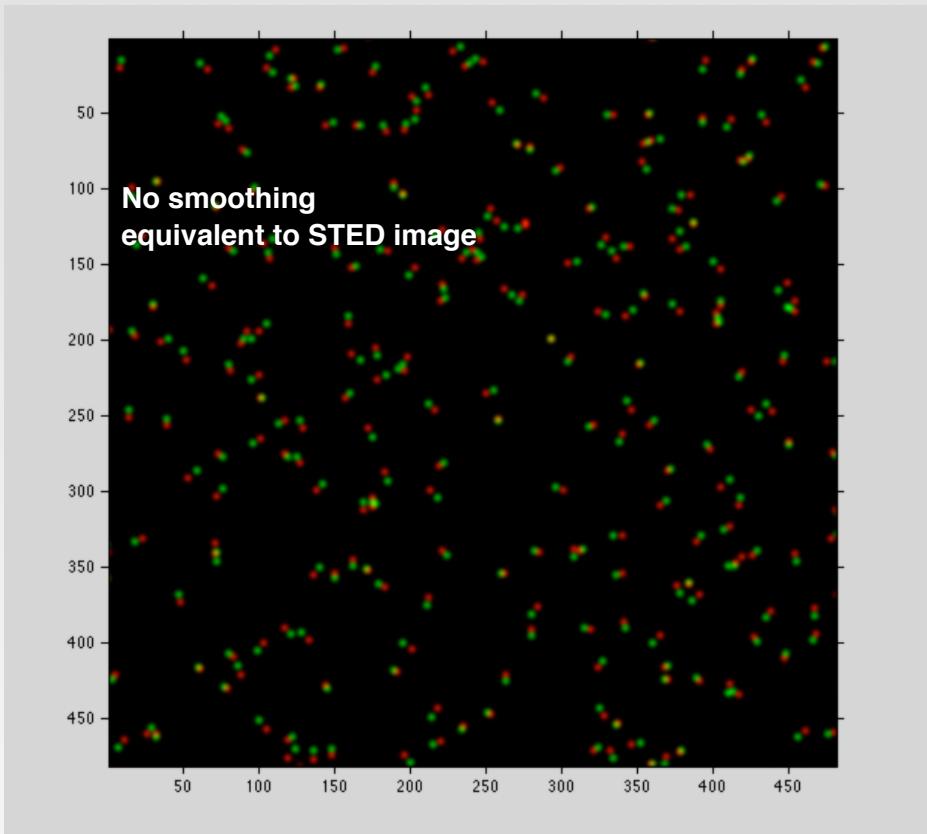
Source:



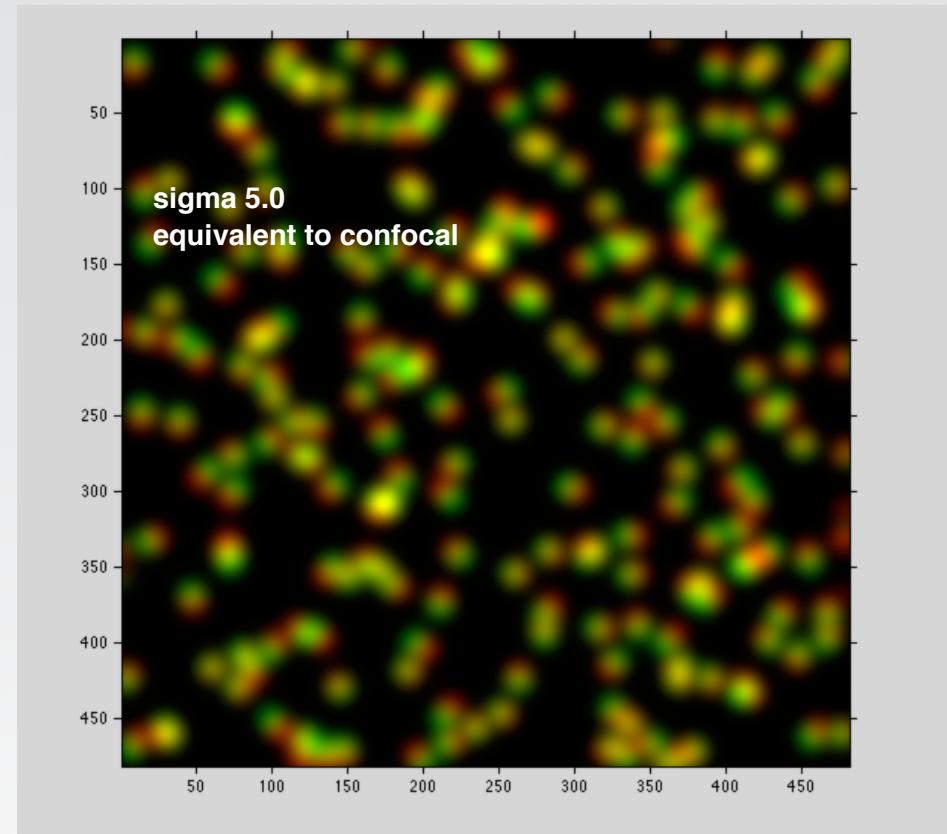
## Super-resolution and colocalisation

Source: [http://en.wikipedia.org/wiki/Super-resolution\\_microscopy](http://en.wikipedia.org/wiki/Super-resolution_microscopy)

# localisation and colocalisation



$r = 0.23$



$r = 0.84$

Source:

# Pearson's test with saturation and clipping

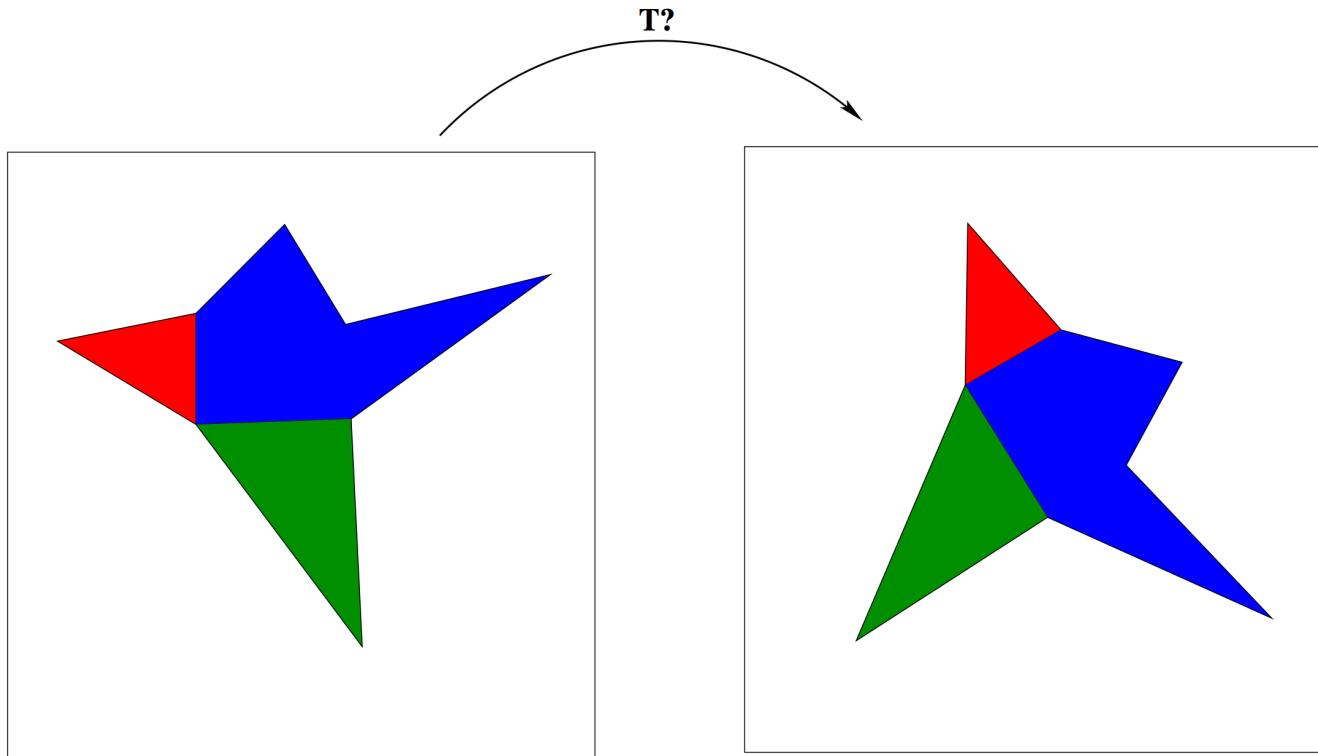
- The Pearson's test doesn't work well if there are many saturated or clipped ('0') pixels in the image. It works best when fluorescence is distributed in both channels in every pixel (at least to some extent).
- If you have many clipped or saturated pixels then you will find it harder to distinguish trends in your data.



Source: <https://en.wikipedia.org/wiki/Apple>

# Pixel-based registration

## Example

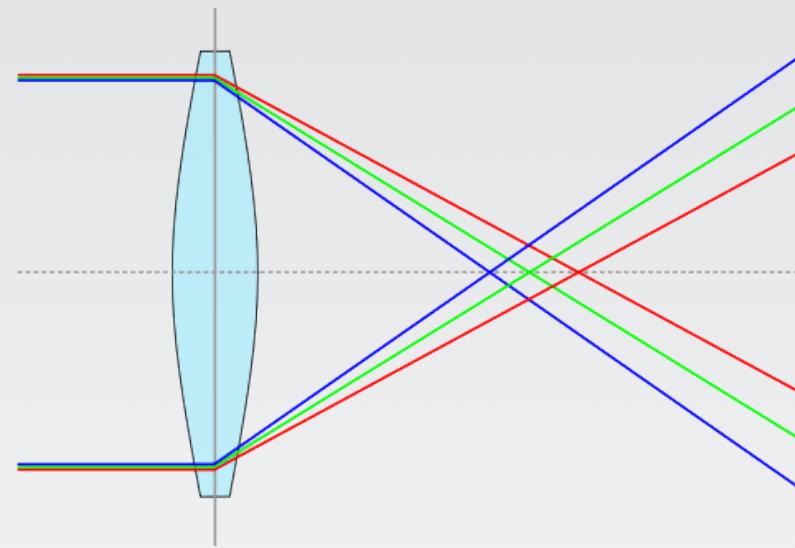


We want to register/align these two stars so that they overlap perfectly. To do this we want to find  $T$ , which is the transformation function.

Source: Sébastien Ourselin, UCL

# Image Restoration

## Optics:

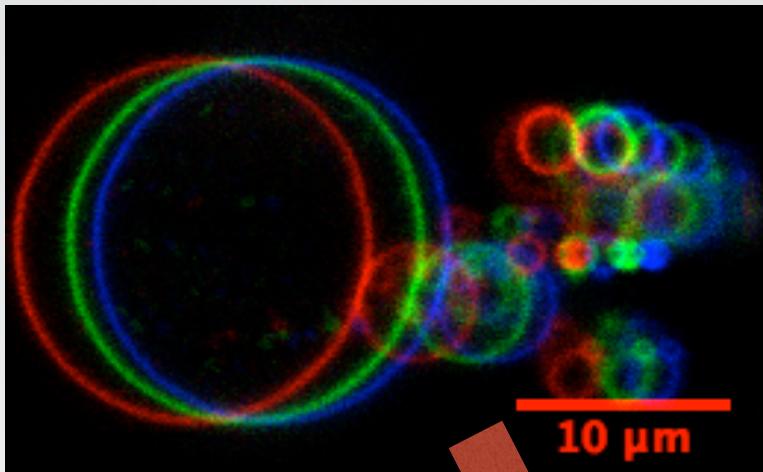


- A artefact of all optical lens is chromatic aberration.
- Refractive index of lens is wavelength dependent.
- More obvious for high-resolution imaging.
- Can be corrected for using careful bead calibration + other.

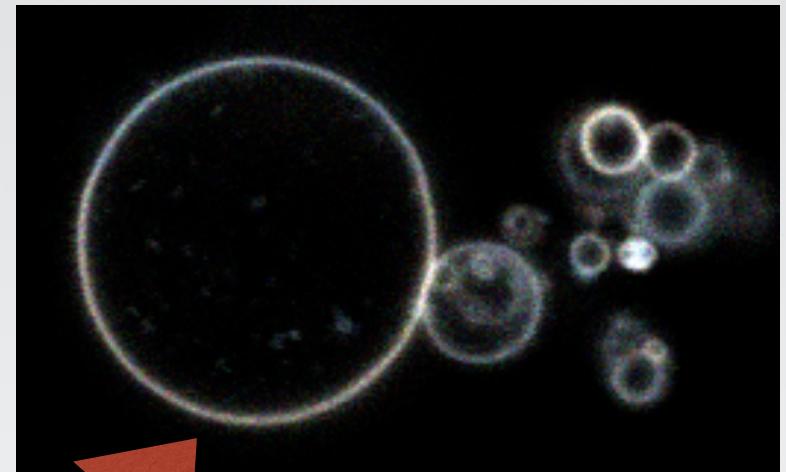
Source: [http://commons.wikimedia.org/wiki/File:Chromatic\\_aberration\\_convex.svg](http://commons.wikimedia.org/wiki/File:Chromatic_aberration_convex.svg)

# Fiji color alignment

Artificially Unaligned

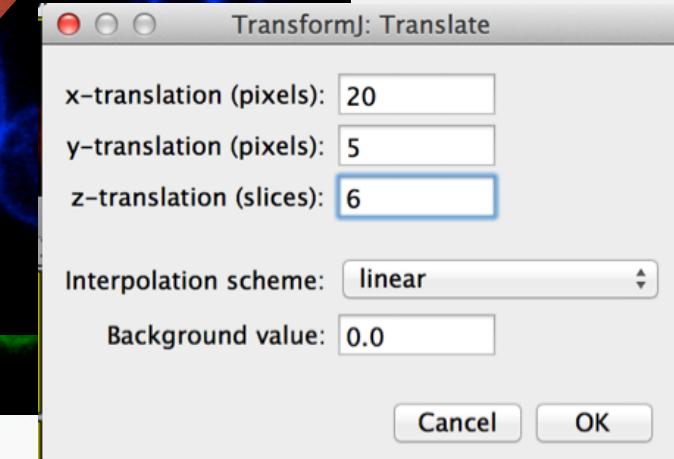
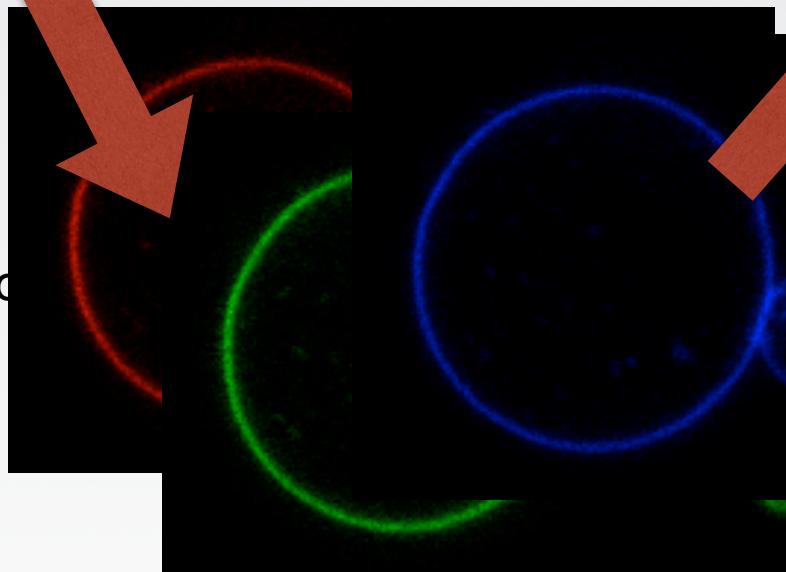


Aligned



Possible in Fiji through

Can be written as



Source: Image, Erdinc Sezgin.

# Pixel based registration

## Correlation Coefficient

- Image normalisation

$$\tilde{I}(x_k) = \frac{I(x_k) - \bar{I}}{\sigma(I)}$$

with  $\bar{I} = \frac{1}{n} \sum_{k=1}^n I(x_k)$  and  $\sigma(I) = \sqrt{\frac{1}{n} \sum_{k=1}^n [I(x_k) - \bar{I}]^2}$

- We want a measure which is invariant by linear transformation of the intensity

$$J = aI + b$$

$$CC = \frac{Cov(I, J)}{\sqrt{Var(I) Var(J)}}$$

This is the same as the Pearson's test  
but theoretically 'I' and 'J' are 2d not 1d

Source:

# Colocalisation Summary

- The dot-product is at the heart of many image-processing algorithms.
- If used correctly, Pearson's test is ingenious method for establishing colocalisation characteristics.
- Poor staining results in unconvincing colocalisation.
- Other tests are available, but must be fully understood.
- Briefly introduced registration.

A guided tour into subcellular colocalization analysis in light microscopy

S Bolte, FP Cordelieres - Journal of microscopy, 2006 - Wiley Online Library

Quantifying colocalization by correlation: The Pearson correlation coefficient is superior to the Mander's overlap coefficient

J Adler, I Parmryd - Cytometry Part A, 2010 - Wiley Online Library

$$(A \bullet B)/c = ||A|| ||B|| \cos\theta \text{ where } c = ||B||/||A||$$

$$(A \bullet B)/c = A \bullet A \cos\theta$$

IF  $||A|| = ||B||$  then  $c = 1$ . and then

$$(A \bullet B) = A \bullet A \cos\theta$$

therefore B can be anything and the dot product is proportional to the angle.

As long as  $||A|| = ||B||$

in autocorrelation  $||A|| = ||B||$  so why do we bother with the rest?, just take dot product.

Source: