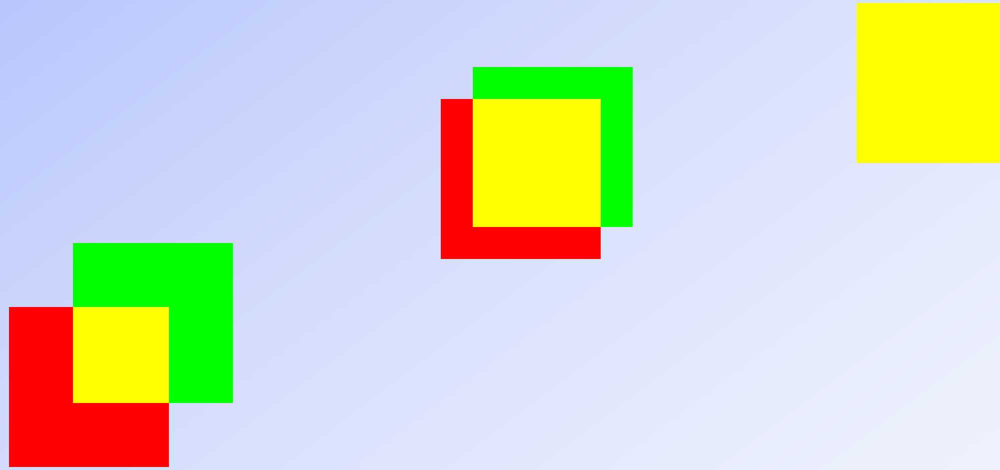
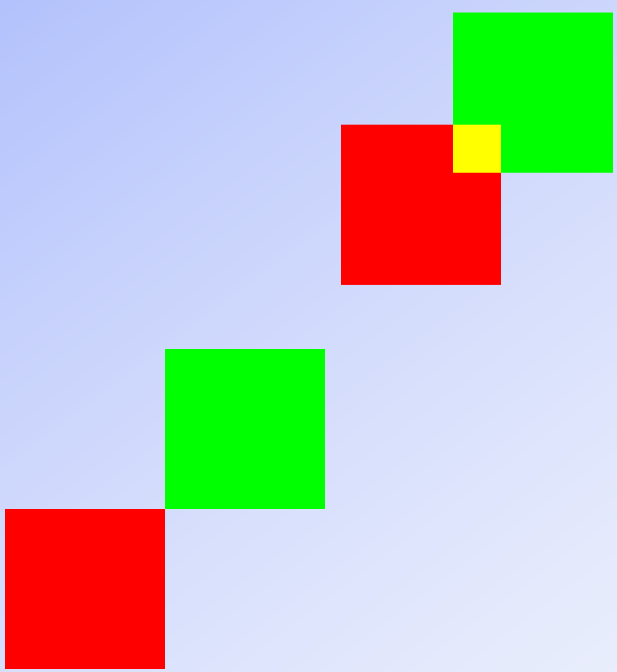
Colocalization





What is colocalization
When do you need it
How to compute (image-based / object based)
Coefficients and their limitations
Advanced methods and tests
pitfalls

What is colocalization?

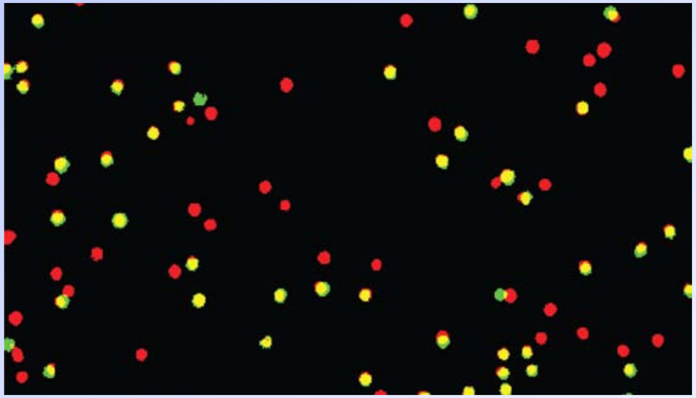
Colocalization occurs when signal of two or more channels is present in the same pixel / voxel

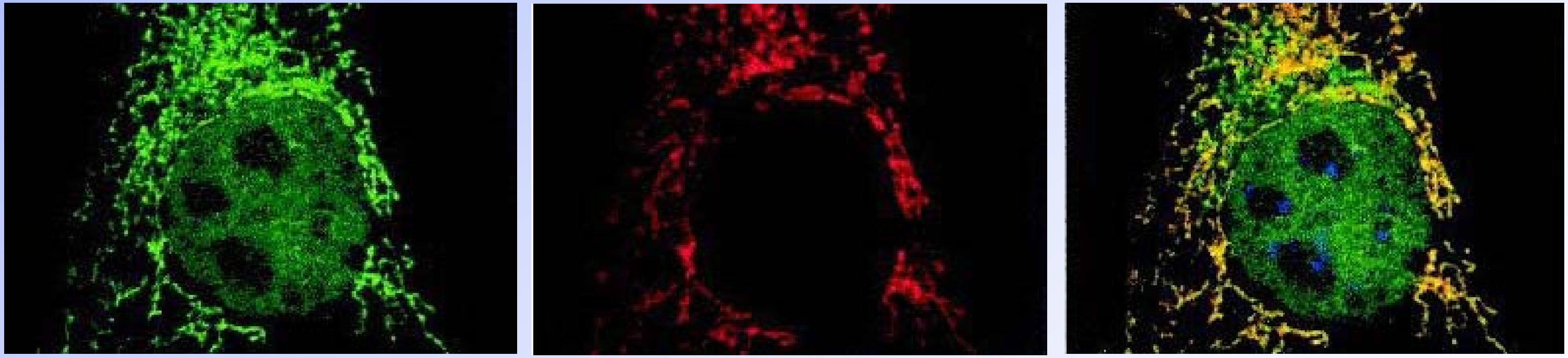
What are you looking for?

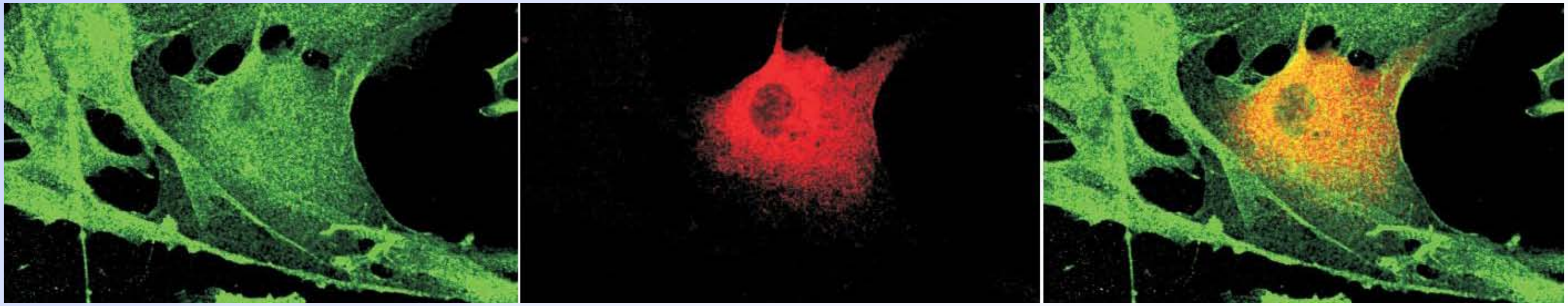
Why doing colocalization analysis?

→To show two signals coincide locally, within

- The same cell?
- The same sub-cellular compartment?
- The same resolution-limited spot?
- The exact same spatial coordinates?



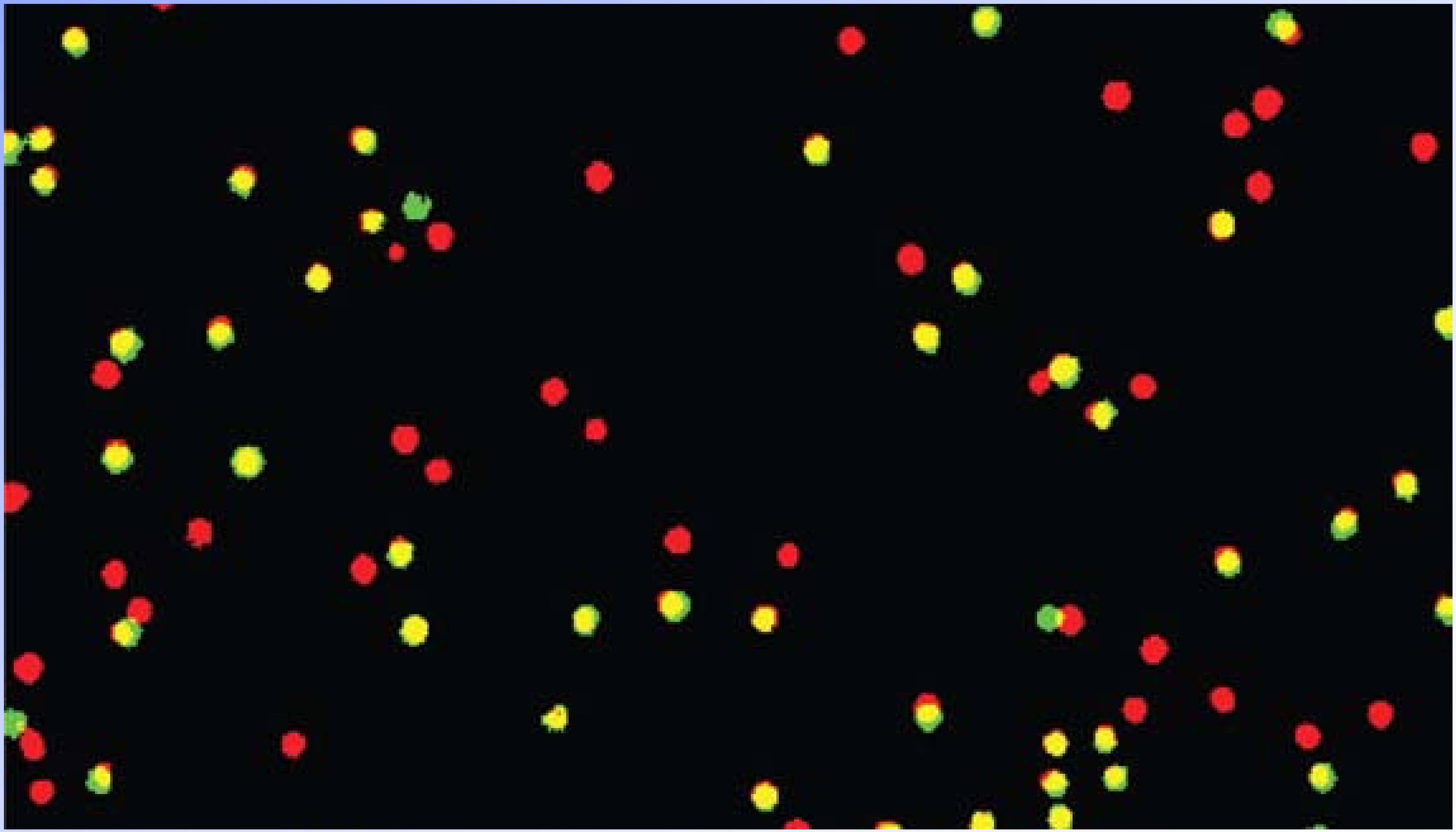




What are you looking for?

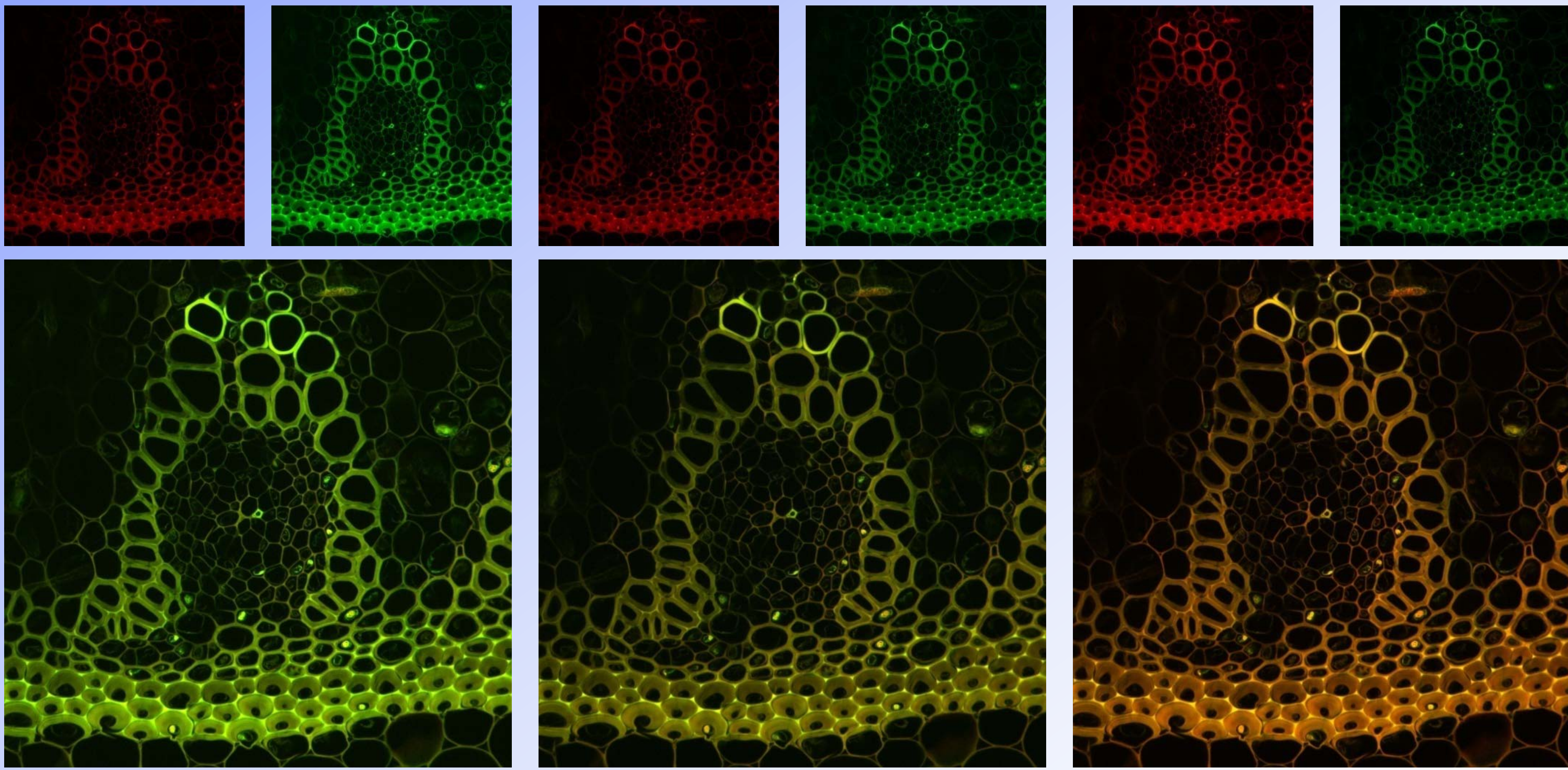
- There is no such thing as „true“ colocalization. Two molecules cannot inhabit the same space.
- Colocalization analysis exploits the resolution limit of optical microscopy
- It works only if there is actual overlap. Two spatially distinct signals that are located in the same organelle do not contribute to any colocalization coefficient.

qualitative colocalization: the yellow pixel illusion



Pixel colour is an additive superposition of the channel colours. Green plus red is only truly yellow if the intensities are equal, and these can be easily tuned, even in post-processing.

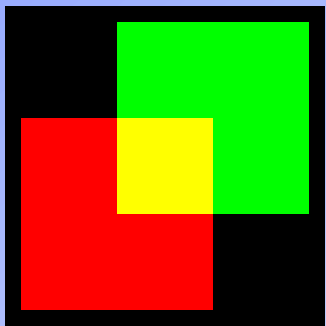
qualitative colocalization: the yellow pixel illusion



Pixel colour is an additive superposition of the channel colours. Green plus red is only truly yellow if the intensities are equal, and these can be easily tuned, even in post-processing.

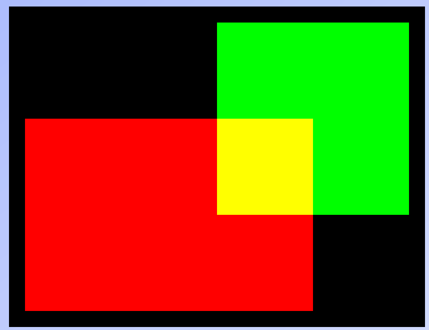
Colocalization and intuition

25% 25%



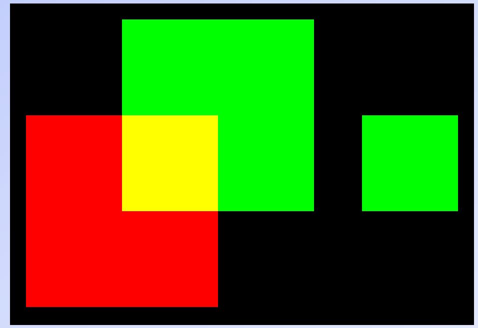
25% ?
12,5%?

17% 25%



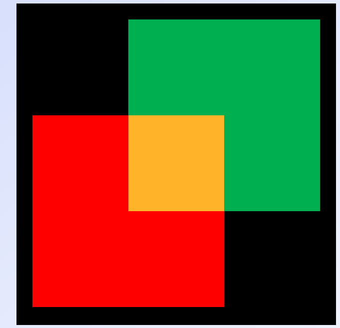
Object shape and size

25% 20%



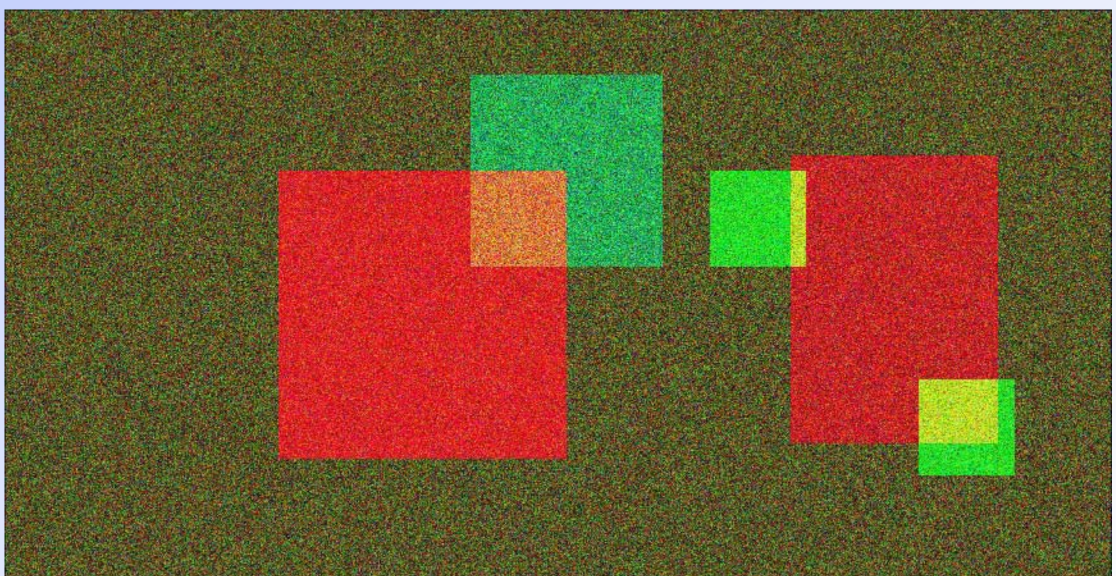
Region of interest

>25%? <25%?



Signal intensity

noise



background

Quantitative Colocalization: Coefficients

What you want:

- Automated image analysis without human bias
- sensitivity for small changes invisible to the human eye

What you get:

- Bias, if coefficients are chosen arbitrarily
- Tunable parameters that greatly influence coefficient values
- Values that need to be interpreted

What you need:

- Carefully and reasonably chosen image acquisition paramters („garbage in, garbage out“)
- A coefficient matched to the images and the scientific question
- An understanding of what the values actually mean

Coefficients:

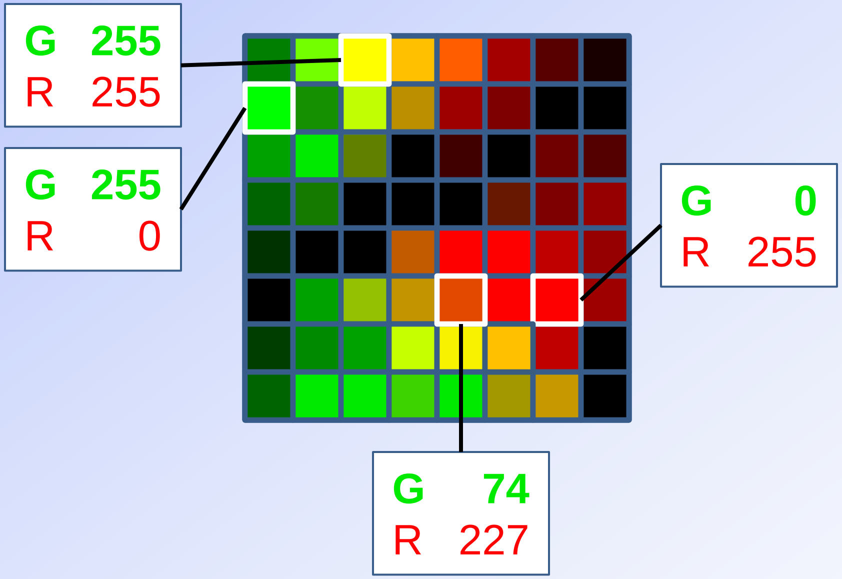
Pearson‘s coefficient Spearman‘s coefficient	Image based
Manders‘ overlap coefficient Manders‘ coefficient k1/k2 Manders coefficient M1/M2	Image based or Object based

Image-based Coefficients: **Pearson**

$$r_p = \frac{\sum((R_i - \overline{R})(G_i - \overline{G}))}{\sqrt{\sum(R_i - \overline{R})^2}\ \sqrt{\sum(G_i - \overline{G})^2}} = \frac{covariance\ (R,G)}{\sigma_R * \sigma_G}$$

0	115	255	255	255	164	88	24
0	20	193	188	158	126	0	0
0	0	193	0	164	0	112	84
0	20	0	0	0	104	150	126
0	0	0	253	255	255	192	150
0	0	193	255	227	255	255	158
0	0	0	198	248	255	192	0
0	0	0	60	0	164	200	0

Mean: 100.7

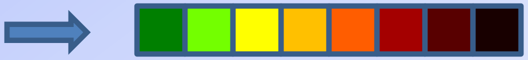


127	255	255	192	93	0	0	0
255	144	253	143	0	0	0	0
162	234	253	0	0	0	0	0
100	122	0	0	0	24	0	0
50	0	0	118	0	0	0	0
0	162	253	192	74	0	0	0
62	138	162	255	242	192	234	0
100	234	234	211	234	152	152	0

Mean: 90.9

Image-based Coefficients: **Pearson**

$$r_p = \frac{\sum((R_i - \overline{R})(G_i - \overline{G}))}{\sqrt{\sum(R_i - \overline{R})^2}\ \sqrt{\sum(G_i - \overline{G})^2}} = \frac{covariance\ (R,G)}{\sigma_R * \sigma_G}$$



$r_p =$ **0,51**

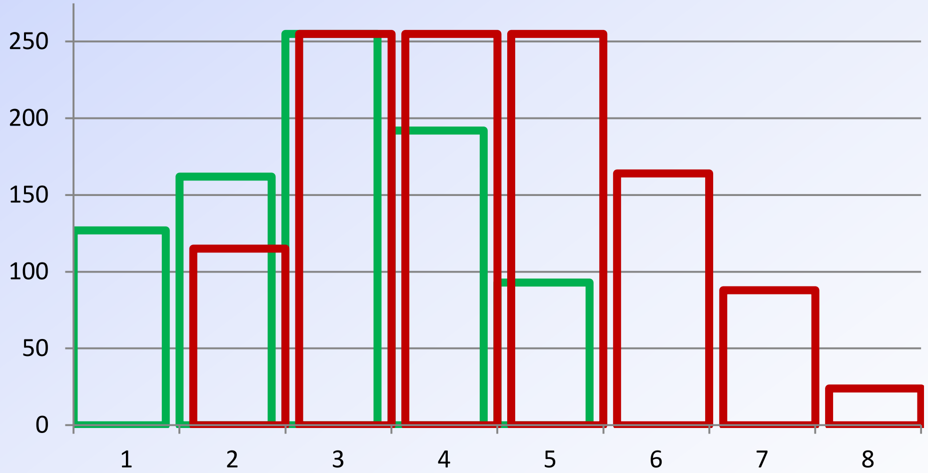
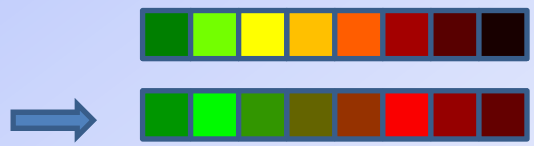


Image-based Coefficients: **Pearson**

$$r_p = \frac{\sum\left(\left(R_i - \overline{R}\right)\left(G_i - \overline{G}\right)\right)}{\sqrt{\sum\left(R_i - \overline{R}\right)^2}\ \sqrt{\sum\left(G_i - \overline{G}\right)^2}} = \frac{covariance\ (R,G)}{\sigma_R * \sigma_G}$$



$r_p =$ 0,51

$r_p =$ -0,83

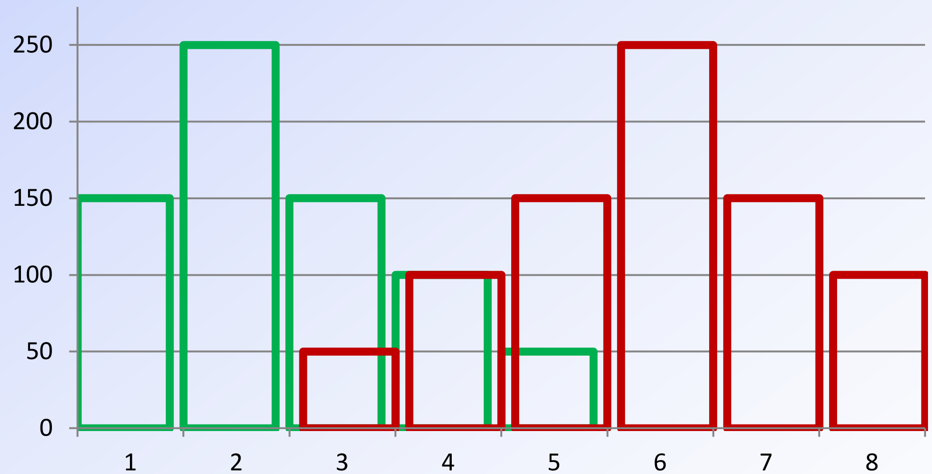
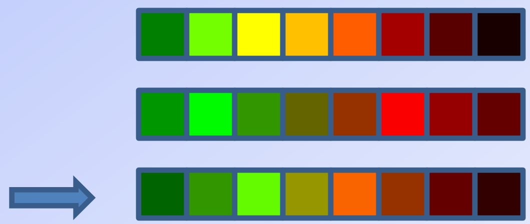


Image-based Coefficients: **Pearson**

$$r_p = \frac{\sum((R_i - \overline{R})(G_i - \overline{G}))}{\sqrt{\sum(R_i - \overline{R})^2}\ \sqrt{\sum(G_i - \overline{G})^2}} = \frac{covariance\ (R,G)}{\sigma_R * \sigma_G}$$



$r_p =$ 0,51

$r_p =$ -0,83

$r_p =$ 0

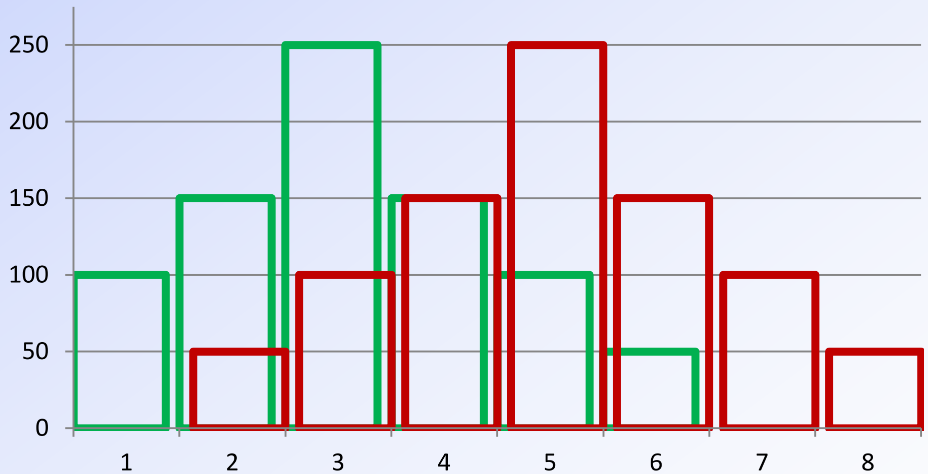


Image-based Coefficients: **Pearson**

$$r_p = \frac{\sum((R_i - \overline{R})(G_i - \overline{G}))}{\sqrt{\sum(R_i - \overline{R})^2}\ \sqrt{\sum(G_i - \overline{G})^2}} = \frac{covariance\ (R,G)}{\sigma_R * \sigma_G}$$

$r_p =$ 0,51

$r_p =$ -0,83

$r_p =$ 0

$r_p =$ 1

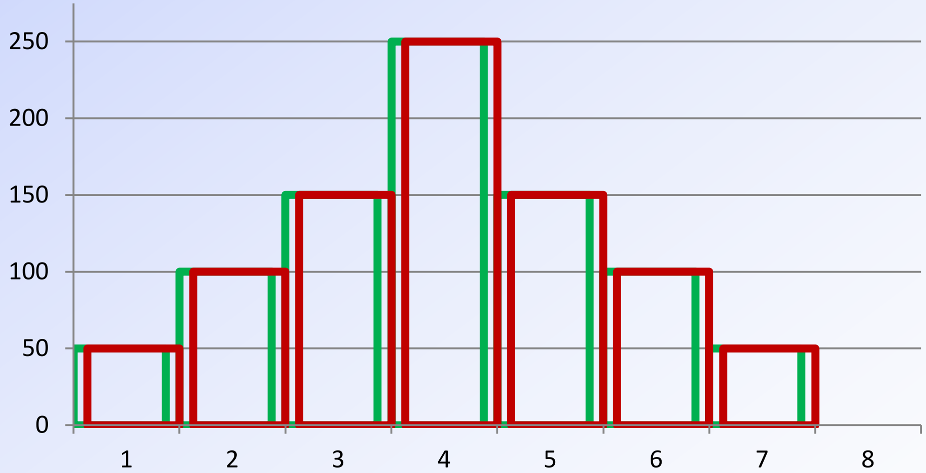


Image-based Coefficients: **Pearson**

$$r_p = \frac{\sum((R_i - \overline{R})(G_i - \overline{G}))}{\sqrt{\sum(R_i - \overline{R})^2}\ \sqrt{\sum(G_i - \overline{G})^2}} = \frac{covariance\ (R,G)}{\sigma_R * \sigma_G}$$

$r_p =$ 0,51

$r_p =$ -0,83

$r_p =$ 0

$r_p =$ 1

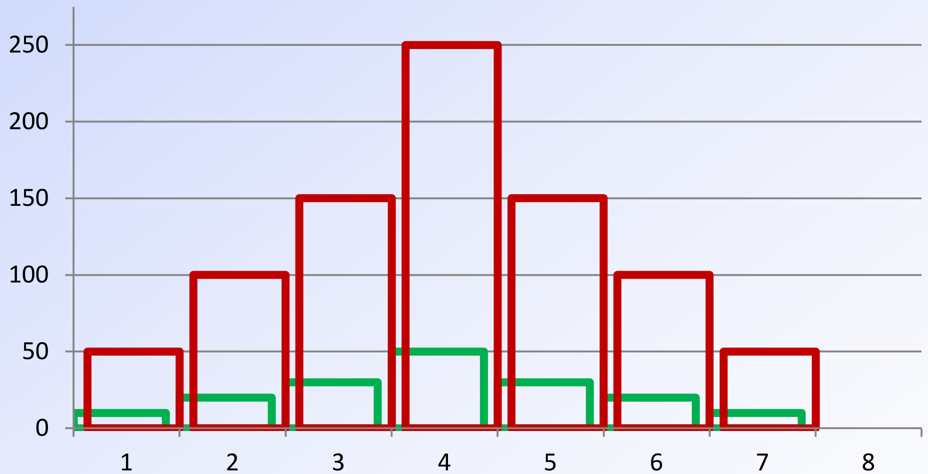


Image-based Coefficients: **Pearson**

$$r_p = \frac{\sum((R_i - \overline{R})(G_i - \overline{G}))}{\sqrt{\sum(R_i - \overline{R})^2}\ \sqrt{\sum(G_i - \overline{G})^2}} = \frac{covariance\ (R,G)}{\sigma_R * \sigma_G}$$

$r_p =$ 0,51

$r_p =$ -0,83

$r_p =$ 0

$r_p =$ 0,82

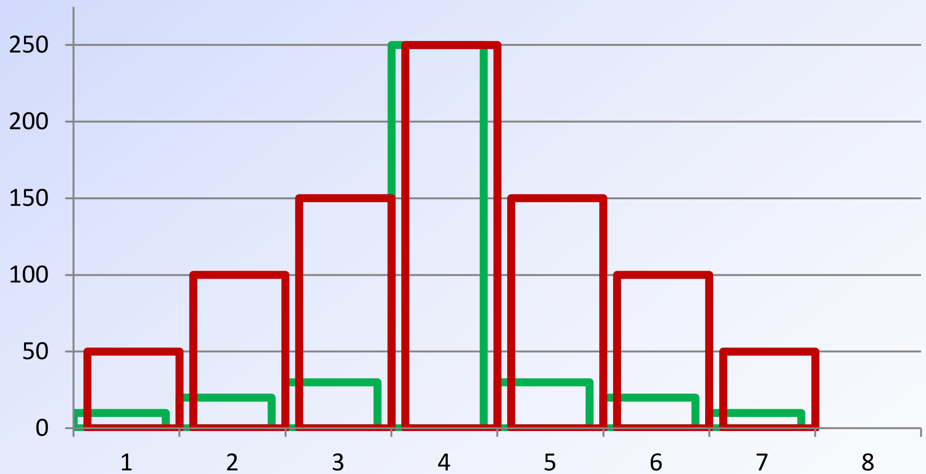


Image-based Coefficients: **Pearson**

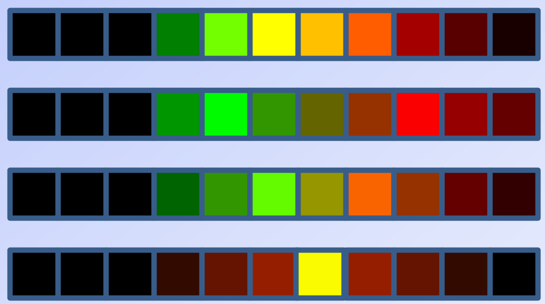
$$r_p = \frac{\sum((R_i - \overline{R})(G_i - \overline{G}))}{\sqrt{\sum(R_i - \overline{R})^2}\ \sqrt{\sum(G_i - \overline{G})^2}} = \frac{covariance\ (R,G)}{\sigma_R * \sigma_G}$$

$r_p =$ 0,66

$r_p =$ -0,35

$r_p =$ 0,37

$r_p =$ 0,80



$r_p =$ 0,51

$r_p =$ -0,83

$r_p =$ 0

$r_p =$ 0,82

Range: **-1 → 1**

Pearson's tops and flops:

- Invariant to signal offset
- Robust against background
- Invariant to unequal signal strength

- Difficult to interpret
- Sensitive to ROI size or segmentation
- Highly sensitive to detector saturation
- Ignorant of nonlinear relationships

Image-based Coefficients: **Spearman**

$$r_s = \frac{\sum((R_r - \overline{R})(G_r - \overline{G}))}{\sqrt{\sum(R_r - \overline{R})^2}\ \sqrt{\sum(G_r - \overline{G})^2}}$$

Intensity values are converted to ranks:

value	0.2	0.4	0.8	0.8	2	3	3	7	13	25
order	1	2	3	4	5	6	7	8	9	10
rank	**1**	**2**	**3,5**	**3,5**	**5**	**6,5**	**6,5**	**8**	**9**	**10**

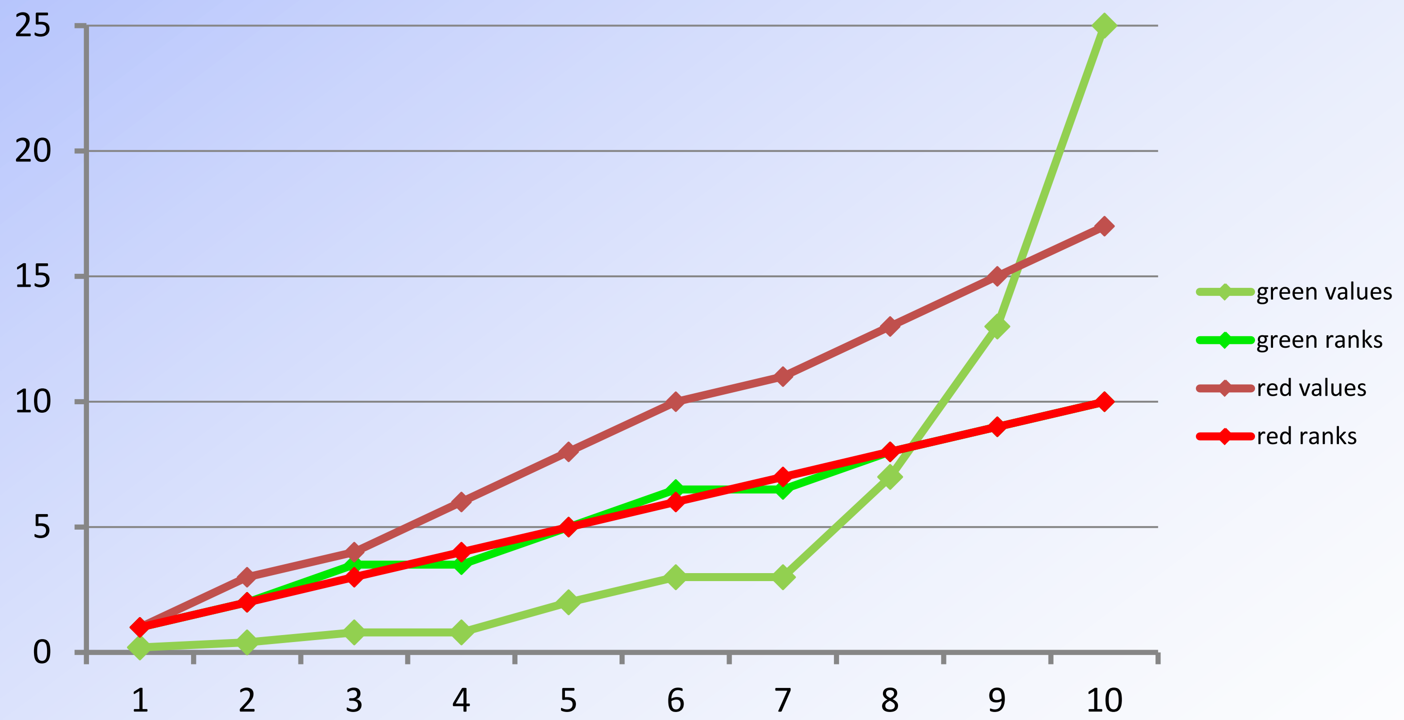


Image-based Coefficients: **Spearman**

Range: **-1 → 1**

Spearman‘s tops and flops:

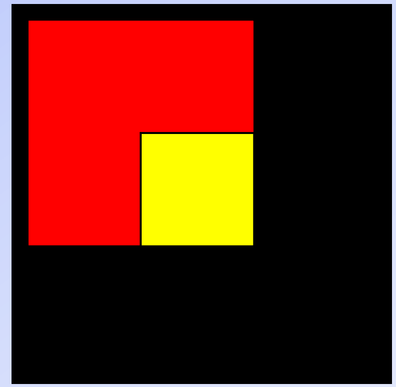
- 👍 Invariant to signal offset
- 👍 Robust against background
- 👍 Invariant to unequal signal strength
- 👍 Reflects linear + nonlinear relationships

- 👎 Difficult to interpret
- 👎 Sensitive to ROI size or segmentation
- 👎 Highly sensitive to detector saturation

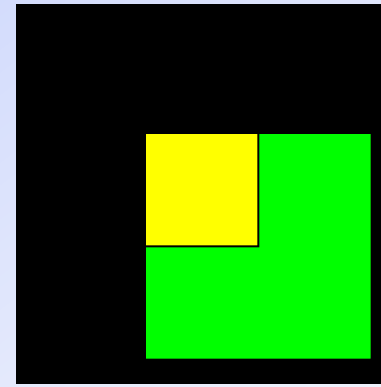
Intensity-based Coefficients: overlap coefficient

$$r_o = \frac{\sum R_i G_i}{\sqrt{\sum R_i^2 \sum G_i^2}}$$

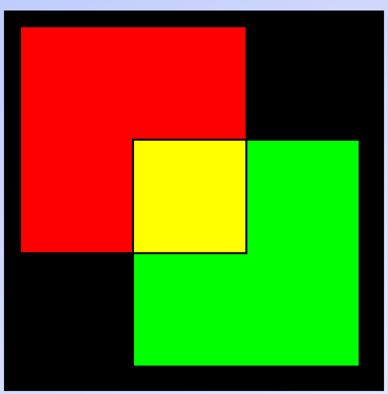
R0 = 0,5



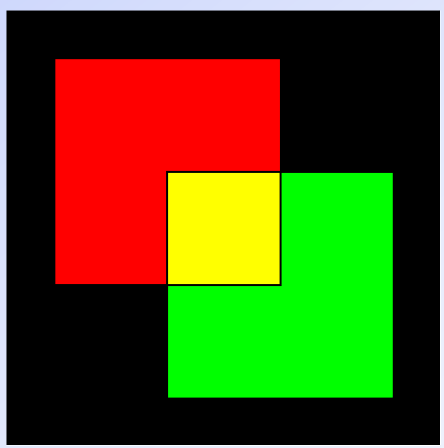
R0 = 0,5



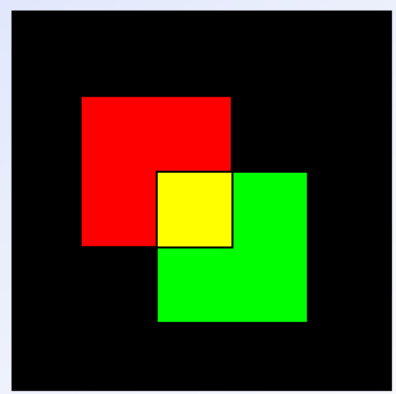
R0 = 0.25



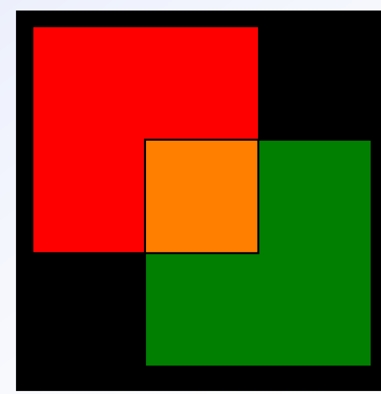
R0 = 0.25



R0 = 0.25



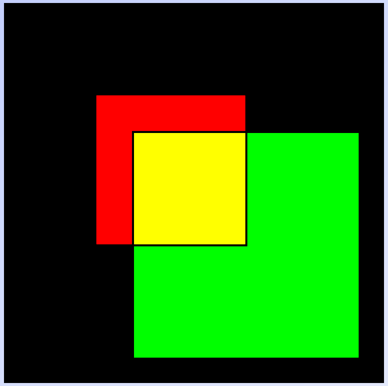
R0 = 0,25



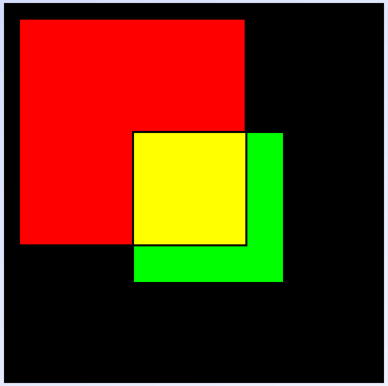
Intensity-based Coefficients: overlap coefficient

$$r_o = \frac{\sum R_i G_i}{\sqrt{\sum R_i^2 \sum G_i^2}}$$

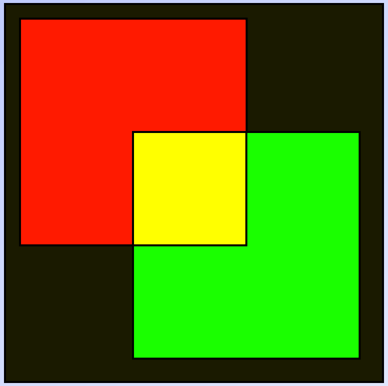
R0 = 0,38



R0 = 0,38

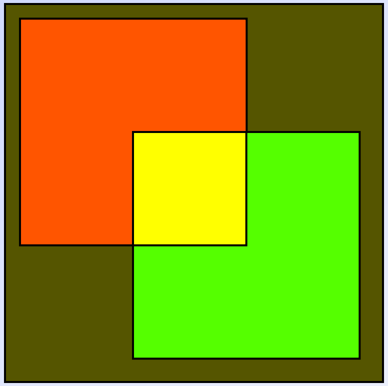


R0 = 0,41



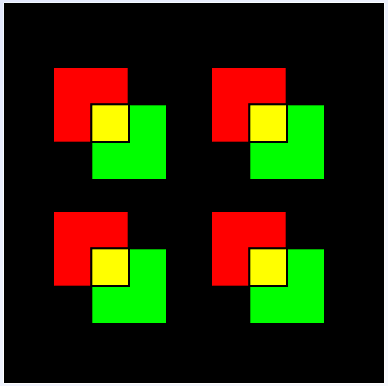
10% Background

R0 = 0,72

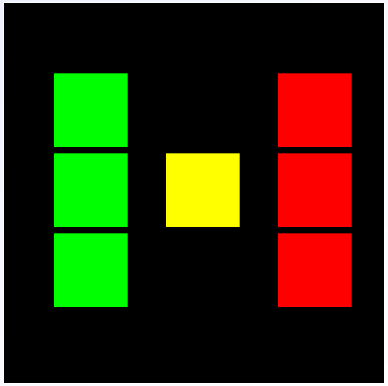


30% Background

R0 = 0,25



R0 = 0,25



Range: **0 - 1**

Overlap coefficient tops and flops:

- 👍 Invariant to ROI size
- 👍 Invariant to unequal signal strength
- 👍 Values show overlap; easier to interpret

- 👎 Highly sensitive to background or offset
- 👎 No differentiation of assymetric overlap
- 👎 Ambiguous values (different structures lead to the same values)

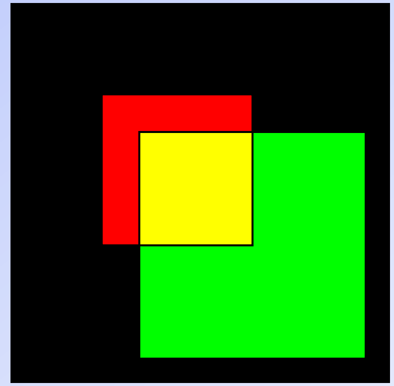
Intensity-based Coefficients: overlap, Manders

$$r_o = \frac{\sum R_i G_i}{\sqrt{\sum R_i^2 \sum G_i^2}}$$

$$k_1 = \frac{\sum R_i G_i}{\sum R_i^2}$$

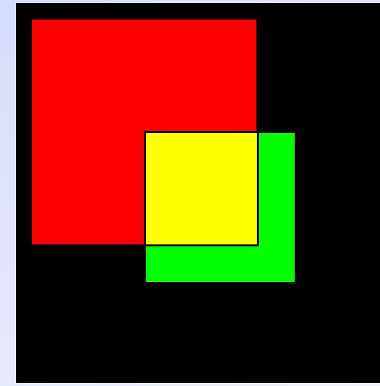
$$k_2 = \frac{\sum R_i G_i}{\sum G_i^2}$$

R0 = 0,38

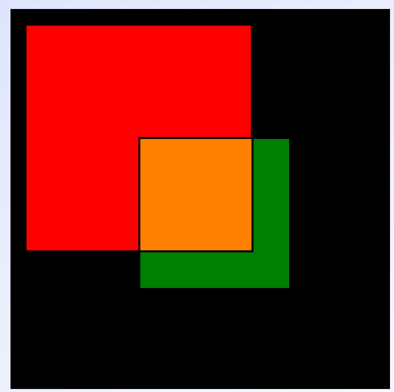


k1 = 0,75
k2 = 0,25

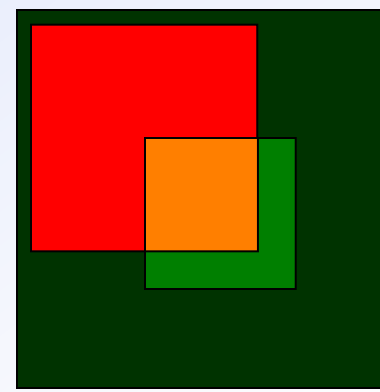
R0 = 0,38



k1 = 0,25
k2 = 0,75



k1 = 0,13
k2 = 2,25



k1 = 0,23
k2 = 1,81
20% Background

Range: **0 - 1**

Manders k1/k2 tops and flops:

- Invariant to ROI size
- Can analyze channel-specific overlap

- Highly sensitive to background or offset
- Values depend intensely on relative signal strength
- Values can be >1, very difficult to interpret

Intensity-based Coefficients: overlap, Manders

$$r_o = \frac{\sum R_i G_i}{\sqrt{\sum {R_i}^2 \sum {G_i}^2}}$$

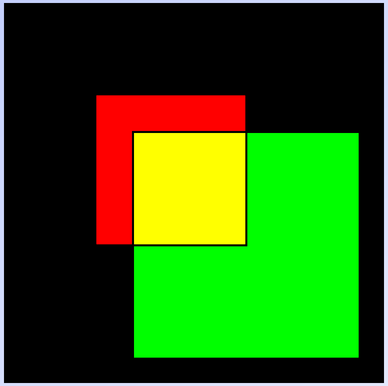
$$M_1 = \frac{\sum R_{i,coloc}}{\sum R_i}$$

$R_{i,coloc}$ = R_i at G_i >threshold

$$M_2 = \frac{\sum G_{i,coloc}}{\sum G_i}$$

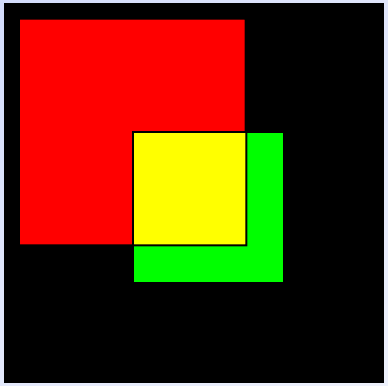
$G_{i,coloc}$ = G_i at R_i >threshold

R0 = 0,38

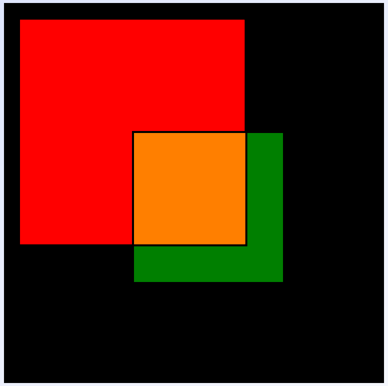


k1 = 0,74
k2 = 0,25

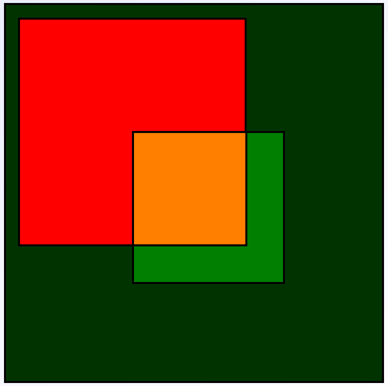
R0 = 0,38



k1 = 0,25
k2 = 0,56



k1 = 0,13
k2 = 2,25



k1 = 0,23
k2 = 1,81
20% Background

Range: **0 - 1**

Manders M1/M2 tops and flops:

- 👍 Invariant to ROI size
- 👍 Invariant to unequal signal strength
- 👍 Invariant to total signal strength
- 👍 Can analyze channel-specific overlap

- 👎 Highly sensitive to background or offset
- 👎 Dependent on segmentation method

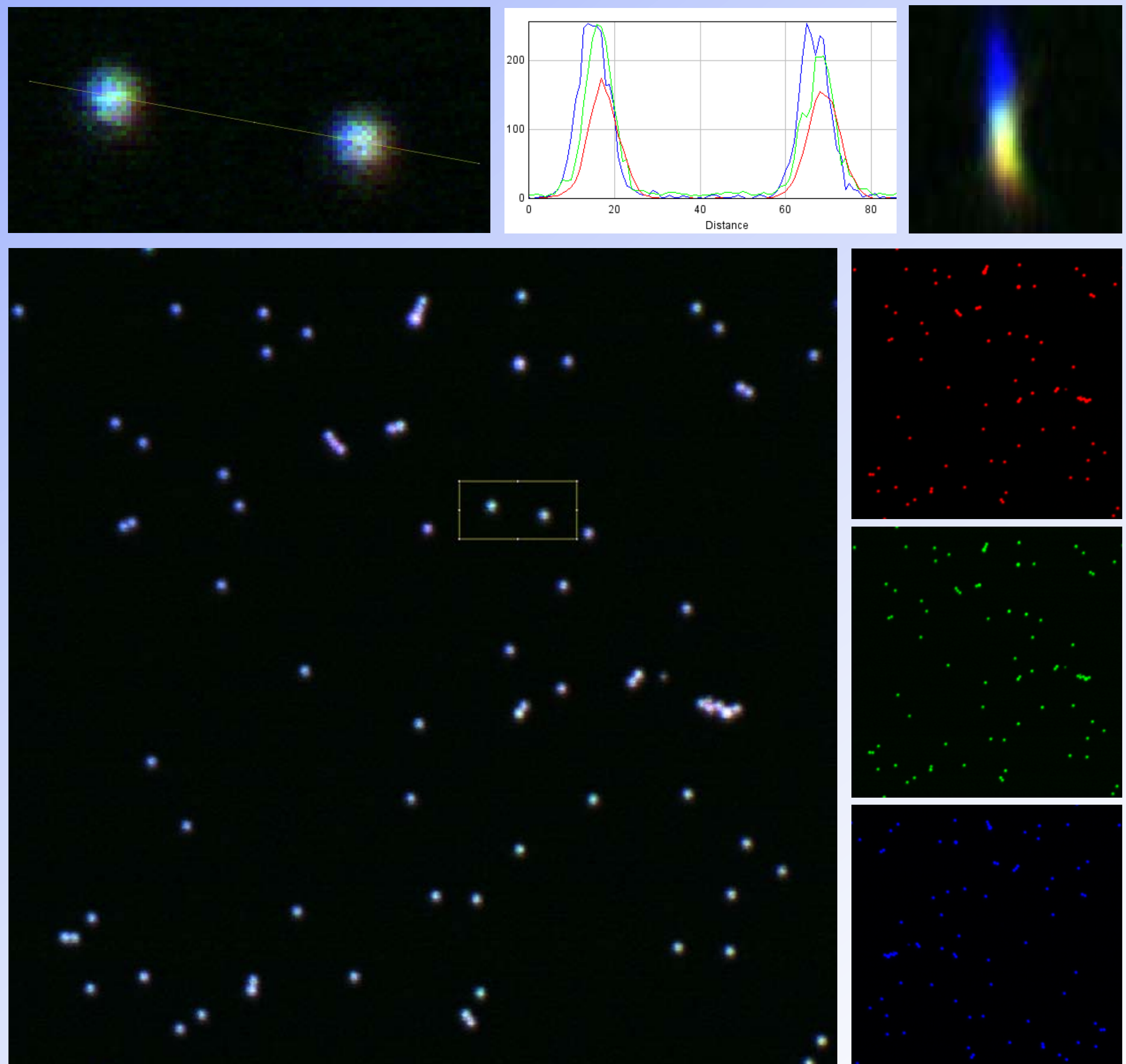
Colocalization coefficients overview

Coefficient	Main weakness / dependent on:	Use for
Pearsons	Insensitive to small but significant changes / ROI size-dependent	Samples with little area variations (e.g. monolayers)
Spearman	As above	As above, with nonlinear dependencies
Overlap coefficient	Insensitive to changes at low intensities / backround-dependent	Samples with equal intensities and large variances of covered areas
Manders k1/k2	Unlimited range of values / background- and intensity-dependent	nothing
Manders M1/M2	Will not work without threshold / background- / threshold-dependent	Everything, will give information about channels indepentently

Pitfalls

- **Crosstalk** leads to false postives in colocalization analysis
- **Chromatic aberration** leads (predominantly) to false negatives in colocalization analysis
- **Clipping** of data due to **detector saturation** leads to changes of intensity patterns, immediately influencing intensity-based coefficients such as Pearson‘s or Spearman‘s
- Errors in **background substraction** alters overlap and Manders k1/k2 coefficients
- **Segmentation** influences all colocalization coefficients!

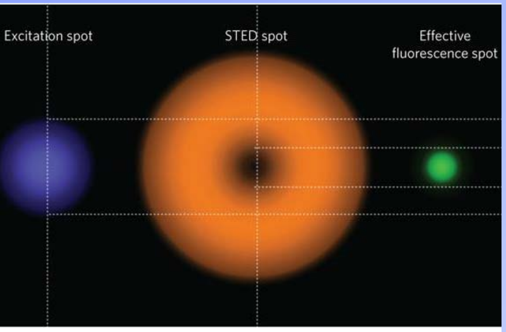
2D versus 3D



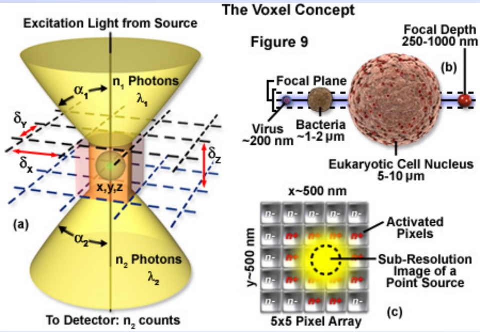
If at all possible, always record three-dimensional data for colocalization analysis!

Superresolution & other complications

STED



PALM

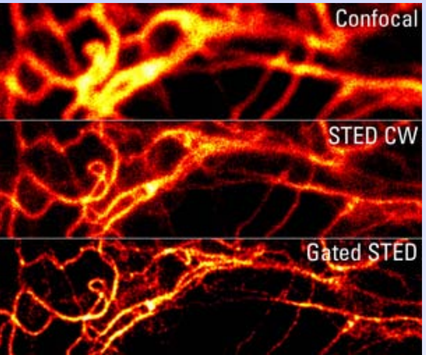


Colocalization
depends on resolution

Increasing resolving power power
leads to decreasing colocalization coefficients

As image resolution increases, „colocalizing“ structures
may be revealed to localize side by side

Colocalization as a concept may have to
be replaced by other methods such as
nearest neighbour analyses



This might also better reflect
non-overlapping localization
co-dependencies