A Gentle Introduction to Bilateral Filtering and its Applications

How does bilateral filter relates with other methods?

Fredo Durand (MIT CSAIL)

Slides by Pierre Kornprobst (INRIA)

Many people worked on…
edge-preserving restoration

Bilateral filter

Partial differential equations

Anisotropic diffusion

Local mode filtering

Robust statistics

Goal: Understand how does bilateral filter relates with other methods

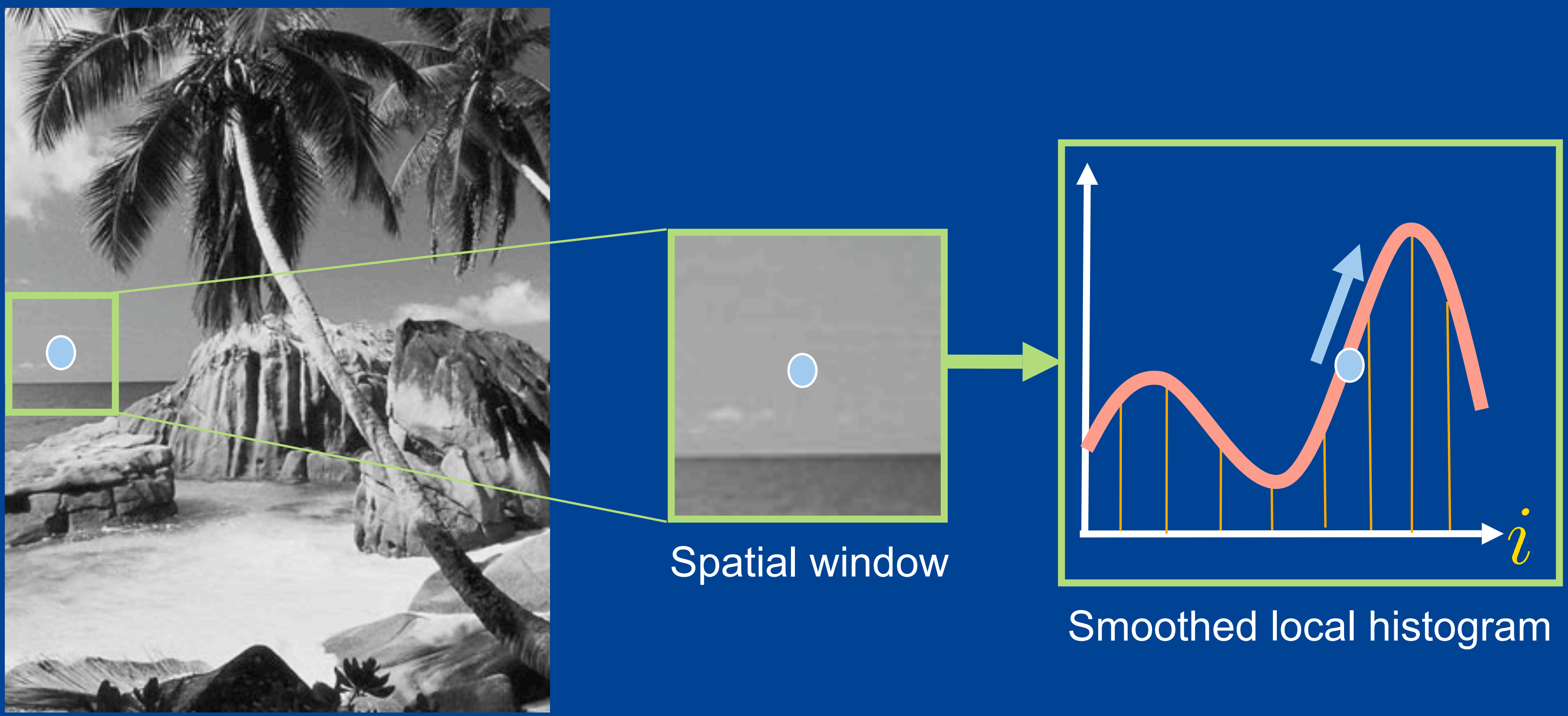
Bilateral filter

Local mode filtering

Partial differential equations

Robust statistics

Local mode filtering principle



You are going to see that BF has the same effect as local mode filtering

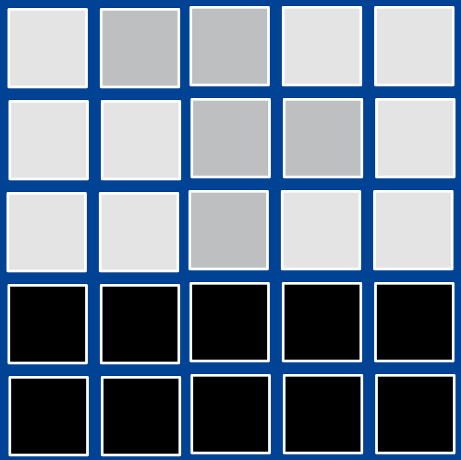
Let’s prove it!

- Define *global* histogram
- Define a *smoothed* histogram
- Define a *local* smoothed histogram
- What does it mean to look for *local modes*?
- What is the *link* with bilateral filter?

Definition of a *global* histogram

- Formal definition of histogram *H* at intensity *i*:

$$H(i) = \sum_{p \in S} \delta(I_p - i)$$

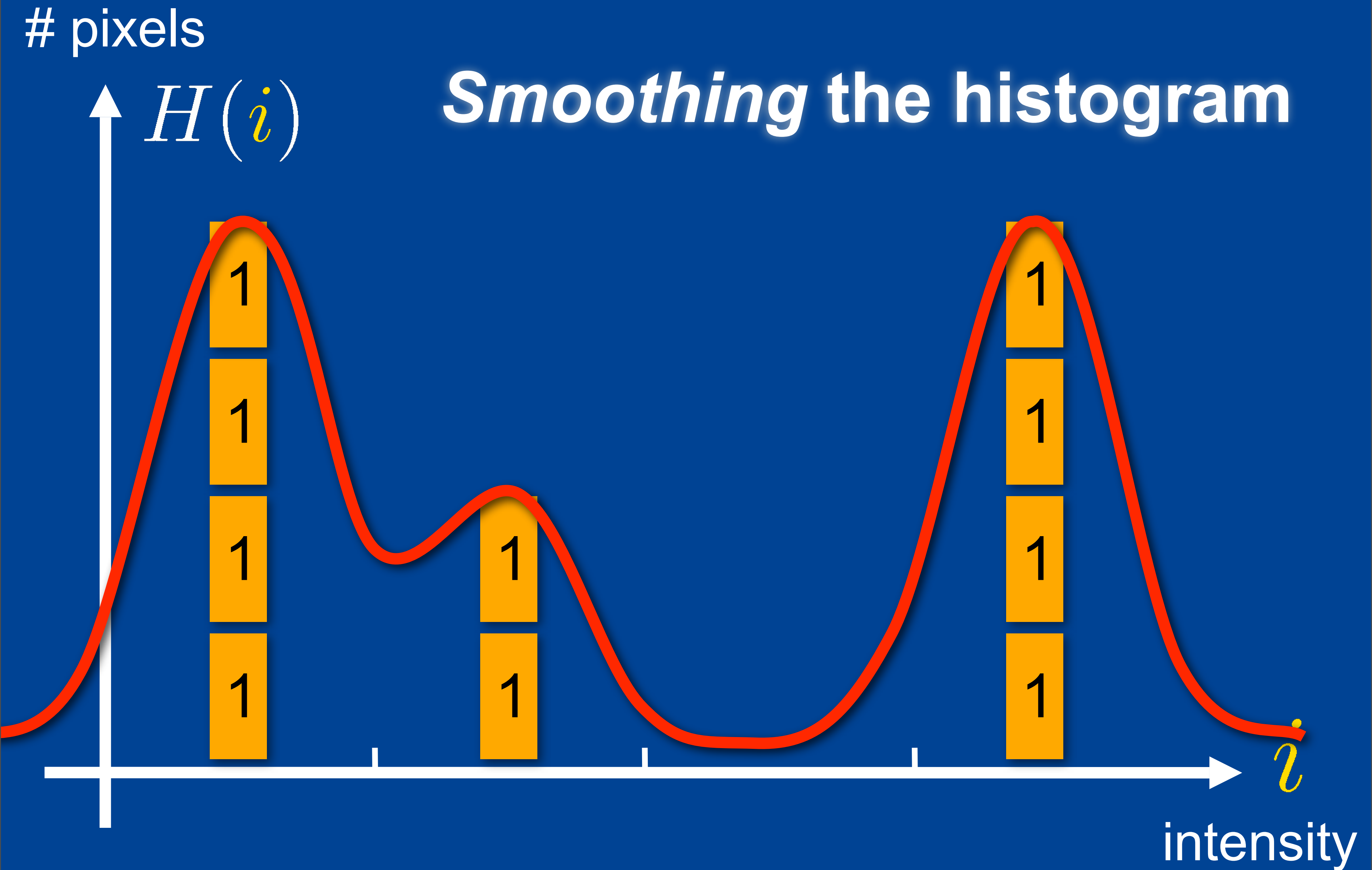


Where $\delta(.)$ is the Dirac symbol (zero everywhere except at 0)

- A sum of Dirac, « a sum of ones »

pixels
$H(i)$
1
1
1
1
1
1
1
1
1
1
i
intensity

Smoothing the histogram

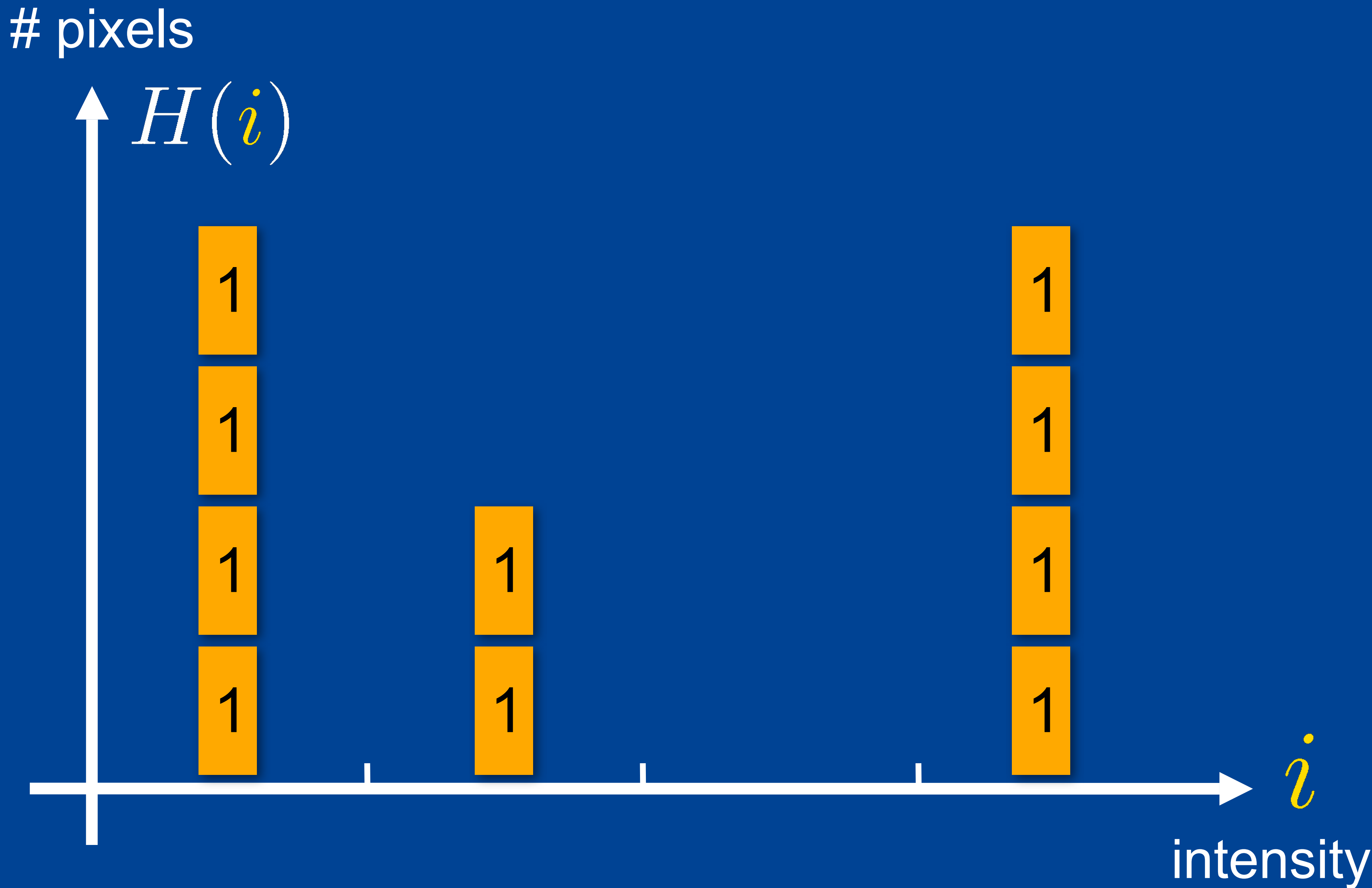


Smoothing the histogram

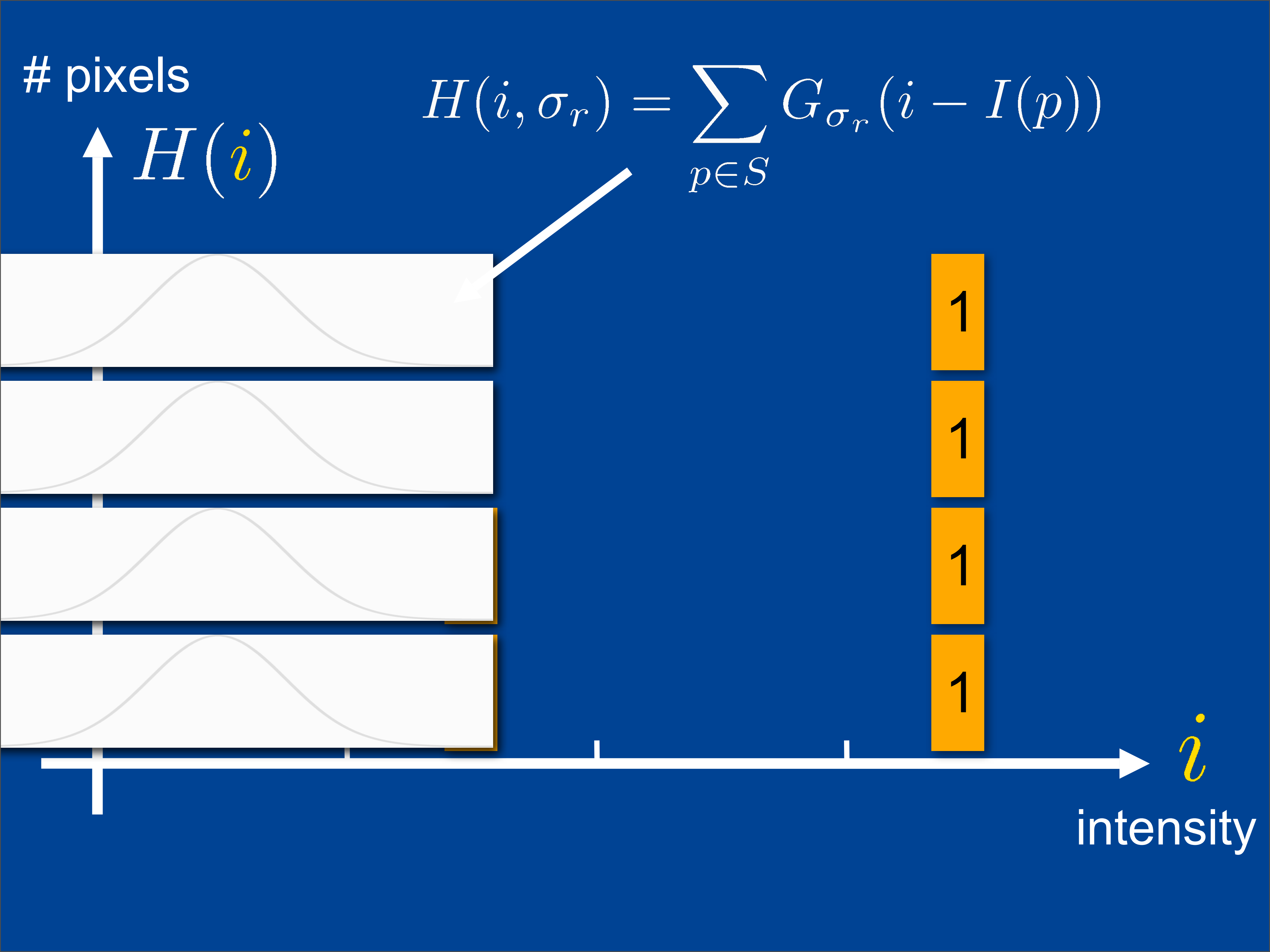
$$H \star G_{\sigma_r}(i) = \sum_{j \in \mathcal{I}} H(j) G_{\sigma_r}(i - j)$$

$$= \sum_{j \in \mathcal{I}} \sum_{p \in S} \delta(I(p) - j) G_{\sigma_r}(i - j)$$

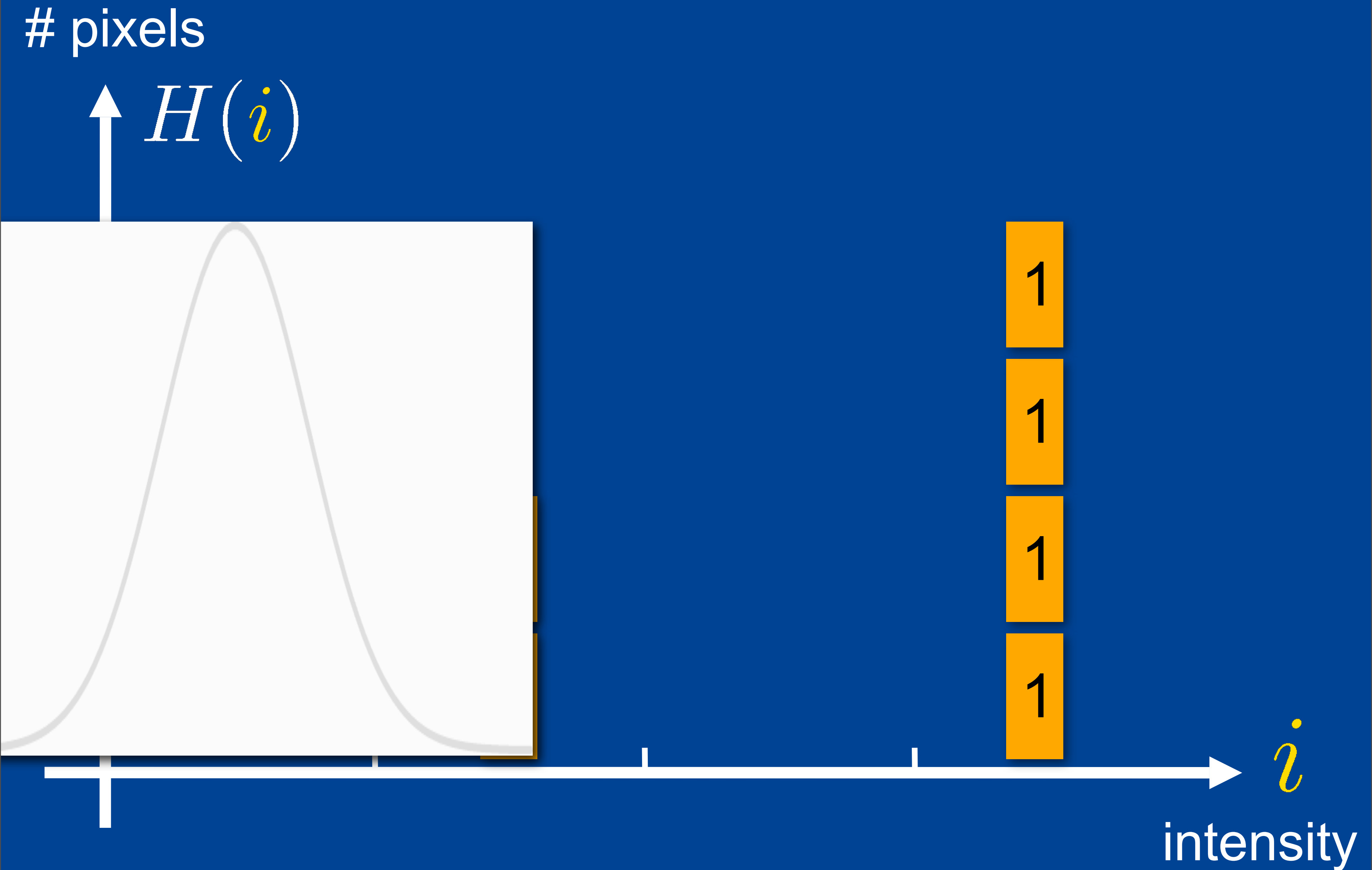
$$\delta(I(p) - j) = 0 \text{ unless } j = I(p)$$



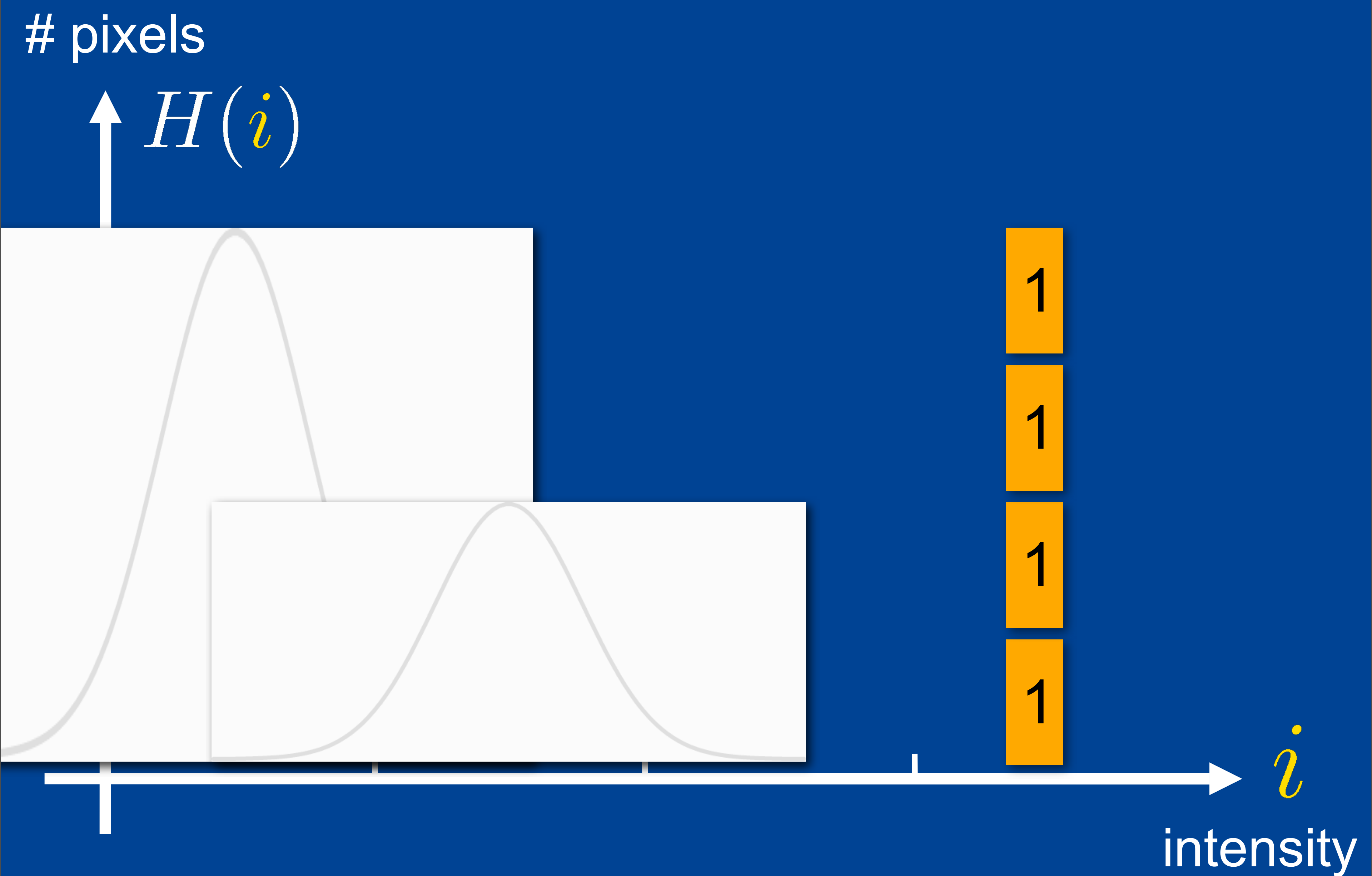
pixels
$H(i)$
1
1
1
1
1
1
1
1
1
1
i
intensity



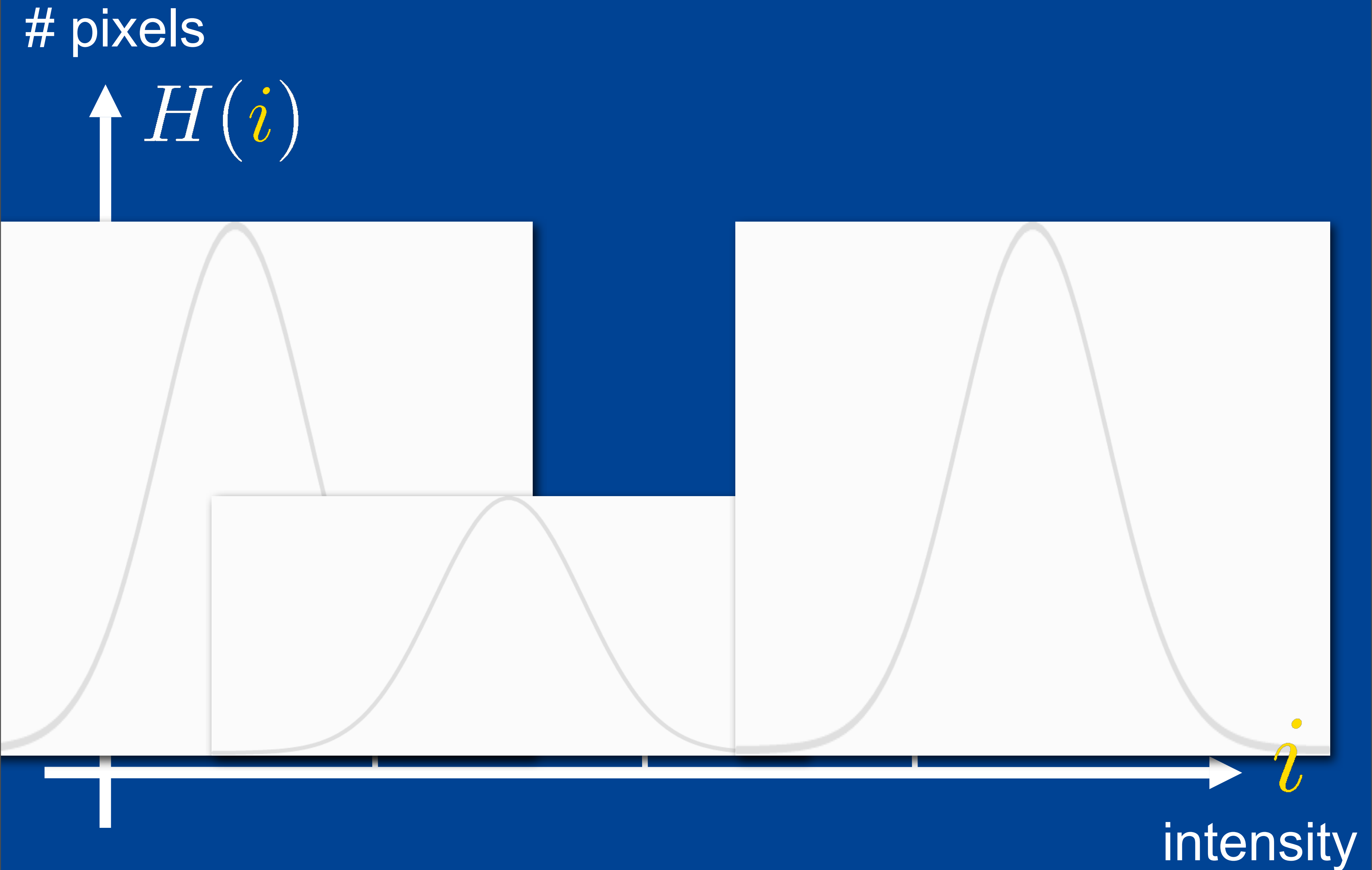
pixels
$H(i)$
$H(i,\sigma_r) = \sum_{p \in S} G_{\sigma_r}(i - I(p))$
1
1
1
1
i
intensity



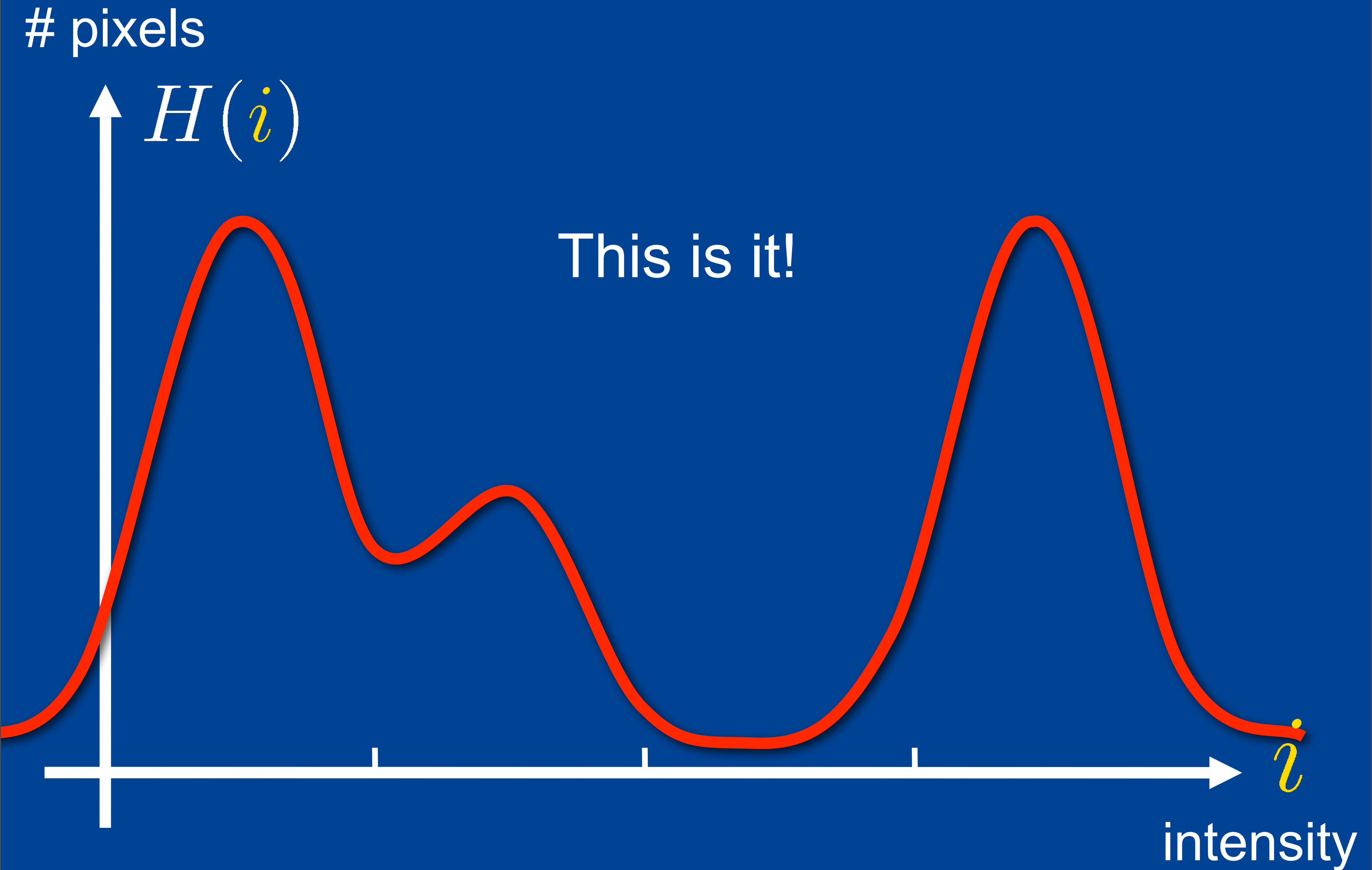
pixels
$H(i)$
1
1
1
1
i
intensity



pixels
$H(i)$
1
1
1
1
i
intensity



pixels
$H(i)$
i
intensity


// pixels
$H(i)$
This is it!
i
intensity

Definition of a *local* smoothed histogram

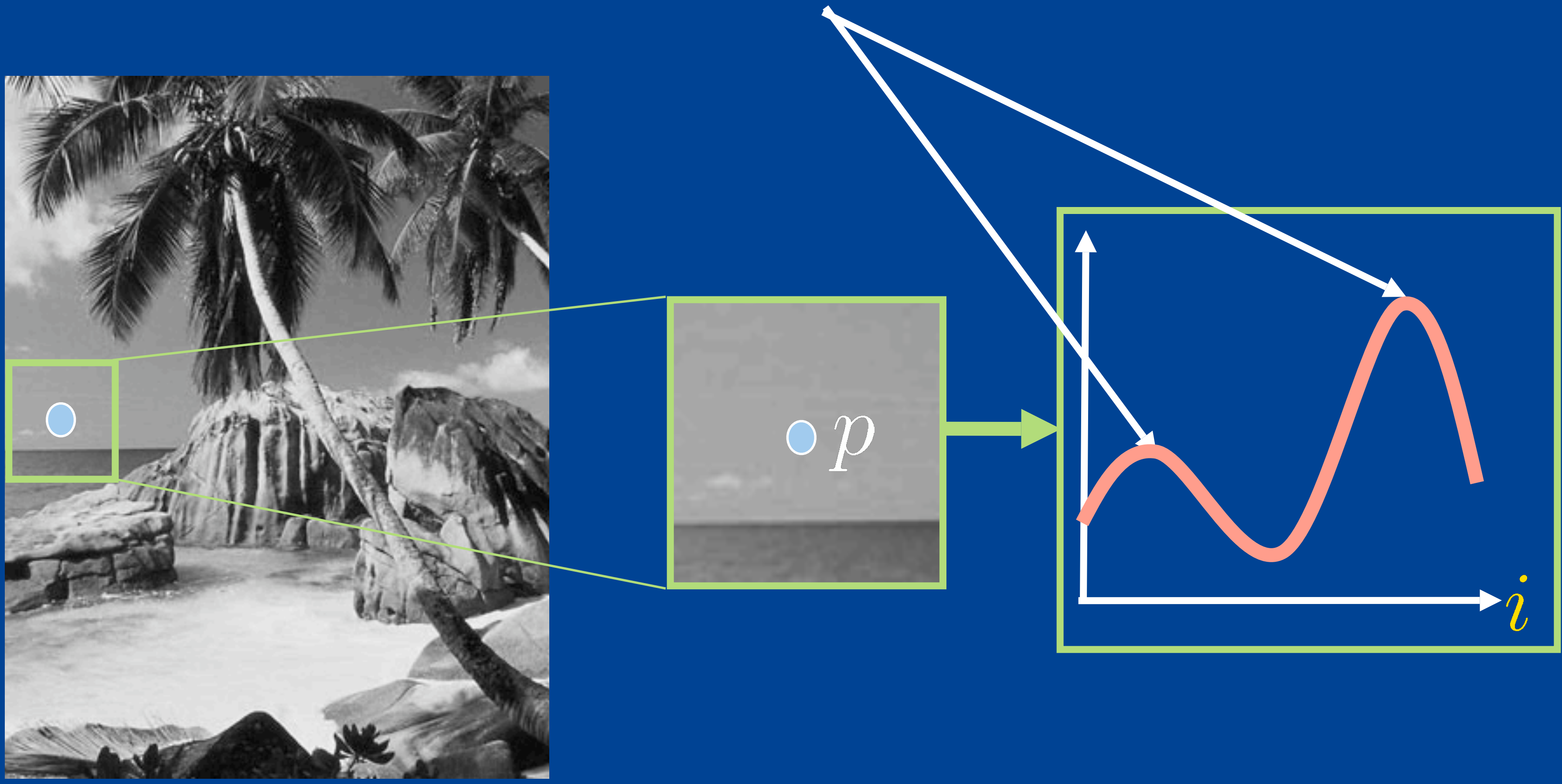
- We introduce a « smooth window »

$$H(i, p, \sigma_r, \sigma_s) = \sum_{q \in \Omega} G_{\sigma_s}(p - q) G_{\sigma_r}(i - I(q))$$

where $\begin{cases} \sigma_r = \text{Smoothing of intensities} \\ \sigma_s = \text{Spatial window} \end{cases}$

And that's the formula to have in mind!

Definition of local modes



A local mode *i* verifies $\dfrac{\partial}{\partial i} H(i, p, \sigma_r, \sigma_s) = 0?$

Local modes?

- Given

$$H(i, p, \sigma_r, \sigma_s) = \sum_{q \in \Omega} G_{\sigma_s}(p-q) G_{\sigma_r}(i - I_q)$$

- We look for

$$i \ / \ \frac{\partial}{\partial i} H(i, p, \sigma_r, \sigma_s) = 0?$$

- Result:

$$i = \frac{\sum_{q \in \Omega} G_{\sigma_s}(p-q) G_{\sigma_r}(i - I_q) I_q}{\sum_{q \in \Omega} G_{\sigma_s}(p-q) G_{\sigma_r}(i - I_q)}$$

Summary

I_p

- A **local** mode *i* verifies:

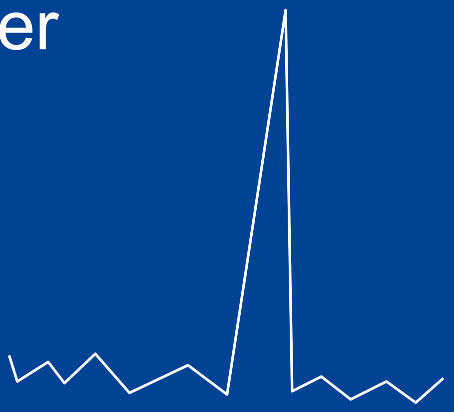
$$i = \frac{\sum_{q \in \Omega} G_{\sigma_s}(p-q) G_{\sigma_r}(i - I_q) I_q}{\sum_{q \in \Omega} G_{\sigma_s}(p-q) G_{\sigma_r}(i - I_q)}$$

- Hey! That looks like bilateral filter!!!

One iteration of the bilateral filter amounts to converge to the local mode

DIscussion

- The bilateral filter goes to a LOCAL mode, not necessarily the global mode
- Often desirable: mode closest to input pixel
- Sometimes not: impulse noise case
  - Recall the use of the median as pre-filter
  - amounts to going to the global mode



Take home message #1

Bilateral filter is equivalent to mode filtering in local histograms

[Van de Weijer, Van den Boomgaard, 2001, etc]

Goal: Understand how does bilateral filter relates with other methods

Bilateral filter

Partial differential equations

Local mode filtering

Robust statistics

Robust statistics?

- Goals: Reduce the influence of outliers

- Minimizing a cost

Error norm

$$min \sum_{p \in S} \sum_{q \in \eta_p^4} \rho(I_p - I_q)$$

- In standard robust statistics $\mathrm{I_q}$ are measured data, $\mathrm{I_p}$ is a robust average of the data

[Huber 81, Hampel 86]

Robust statistics?

- In our case: the output at a pixel should be a robust smoothing of its neighbors
- Minimizing a cost

Error norm

$$min \sum_{p \in S} \sum_{q \in \eta_p^4} \rho(I_p - I_q)$$

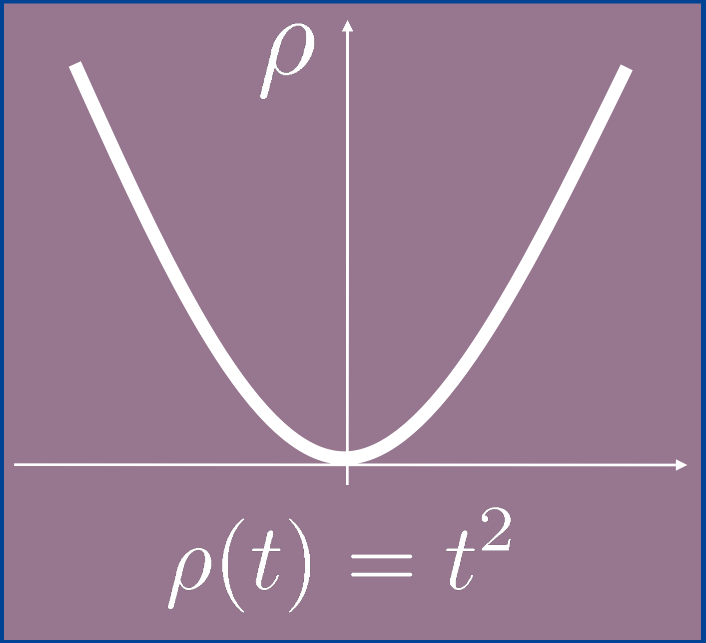
- Extended local formulation

$$min \sum_{p \in S} \sum_{q \in \eta_p} \rho(I_p - I_q) G_{\sigma_s}(p - q)$$

[Huber 81, Hampel 86]

How to choose the error norm?

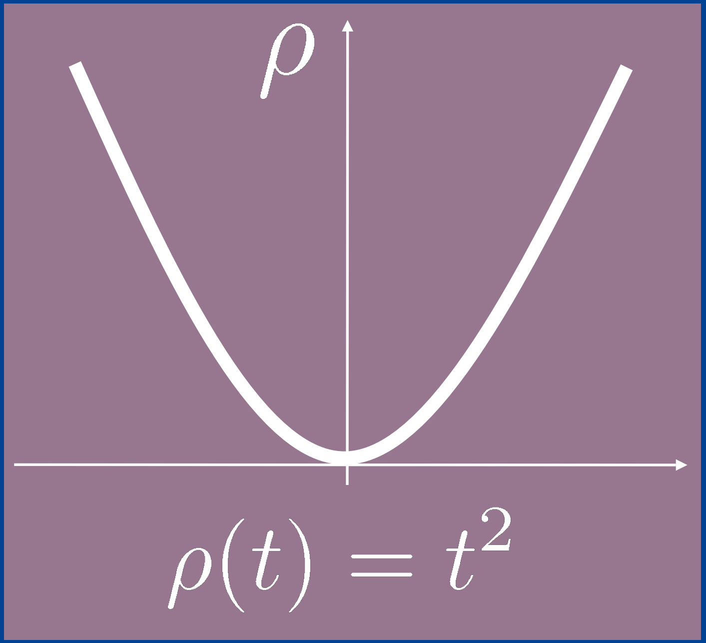
$$min \sum_{p \in \Omega} \sum_{q \in \eta_p^4} \rho(I_p - I_q)$$

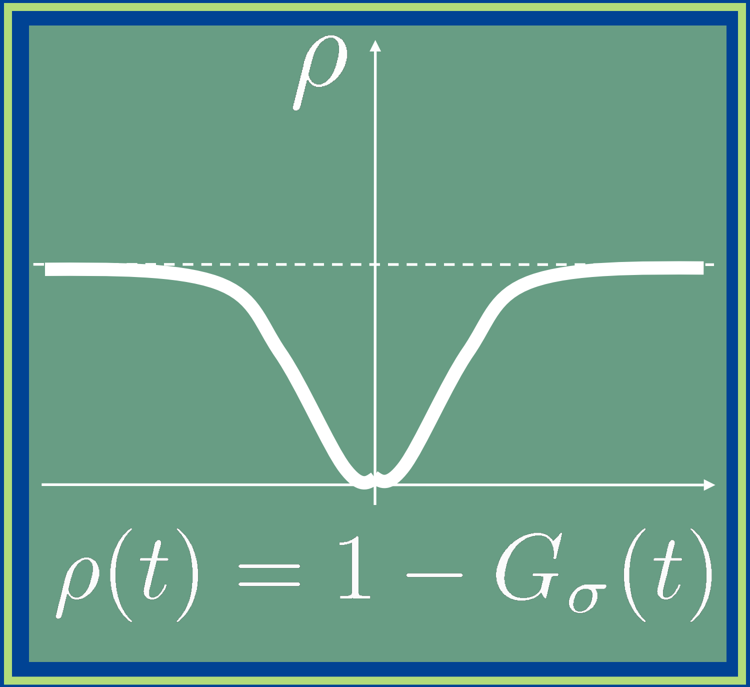


- Least square pays a big penalty for big errors
  - problem in the presence of outliers

How to choose the error norm?

$$min \sum_{p\in\Omega} \sum_{q\in\eta_p^4} \rho(I_p - I_q)$$





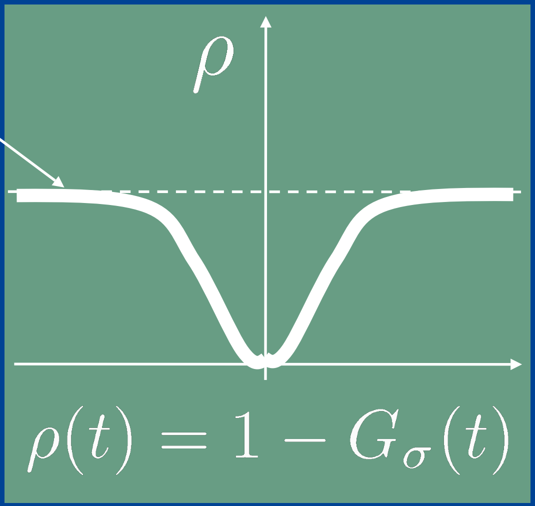
- Strong differences must not be too penalizing, otherwise, everything will be smoothed!

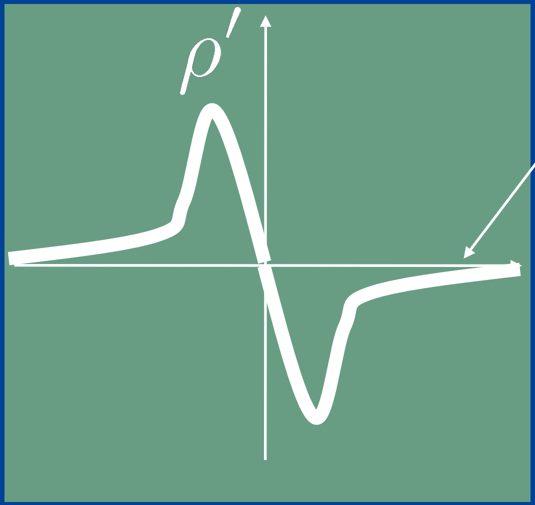
How to minimize the cost function?

- Gradient descent and iterative scheme

$$I_p^{t+1} = I_p^t + \frac{\lambda}{|\eta_p|} \sum_{q \in \eta_p} G_{\sigma_s}(q-p)\rho'(I_q^t - I_p^t)$$

Doesn't cost too much



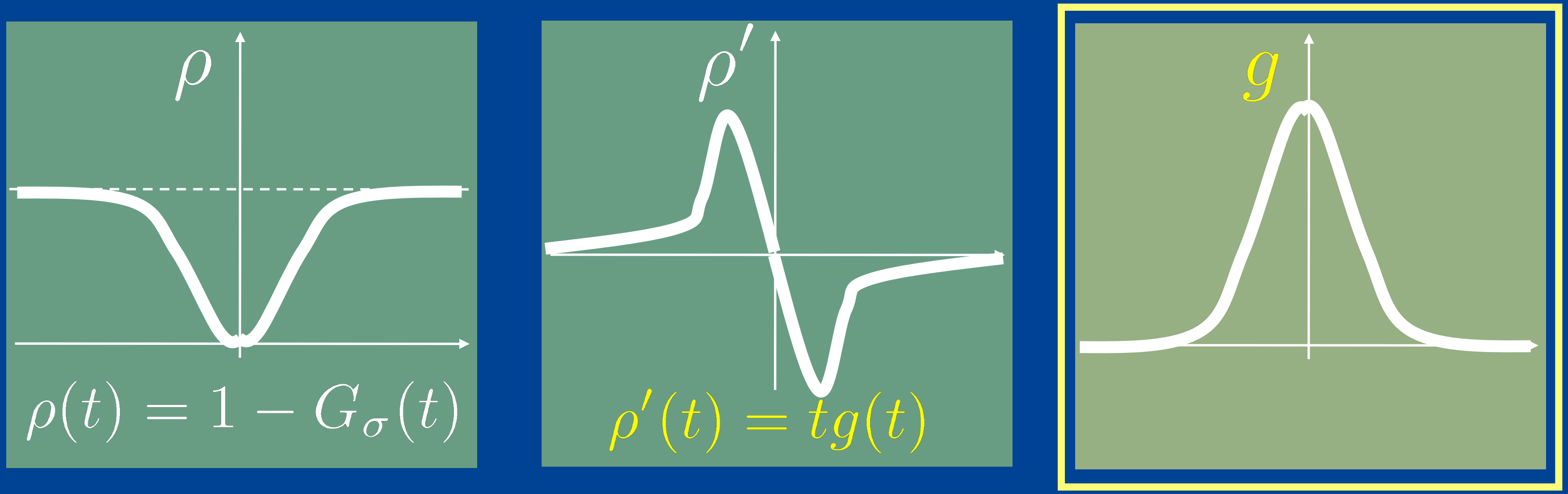


No influence In gradient

Getting closer to bilateral filter

- Rewrite introducing a new function

$$I_p^{t+1} = I_p^t + \frac{\lambda}{|\eta_p|} \sum_{q \in \eta_p} G_{\sigma_s}(q-p)\, g(I_q^t - I_p^t)(I_q^t - I_p^t)$$

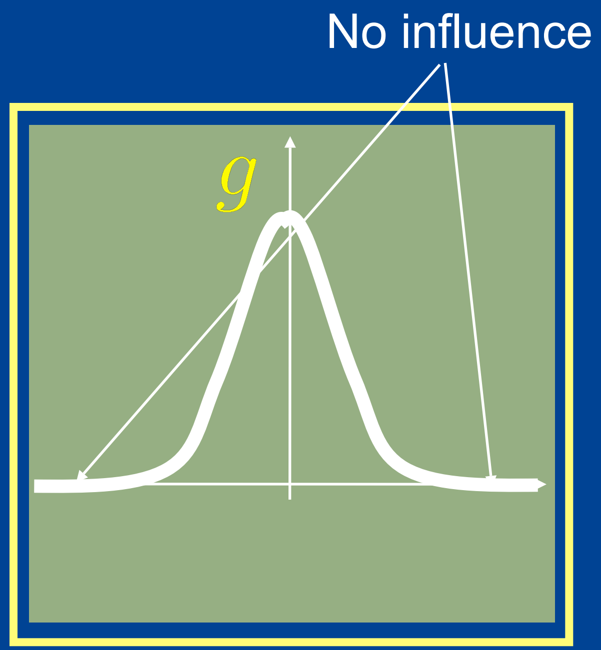


Getting closer to bilateral filter

- Rewrite introducing a new function

$$I_p^{t+1} = I_p^t + \frac{\lambda}{|\eta_p|} \sum_{q \in \eta_p} G_{\sigma_s}(q-p) g(I_q^t - I_p^t)(I_q^t - I_p^t)$$

- g has the same qualitative behavior than a Gaussian
- Now this operator reminds us about bilateral filter!



Really the same?

$$I_p^{t+1} = I_p^t + \frac{\lambda}{|\eta_p|}\sum_q G_{\sigma_s}(q-p)G_{\sigma_r}(I_q^t - I_p^t)(I_q^t - I_p^t)$$

M-estimator

[Hampel etal, 1986]: M-estimators and W_estimators are essentially equivalent and solve the same minimization problem

$$I_p^{t+1} = \frac{\sum_q G_{\sigma_s}(q-p)G_{\sigma_r}(I_q^t - I_p^t)I_q^t}{\sum_q G_{\sigma_s}(q-p)G_{\sigma_r}(I_q^t - I_p^t)}$$

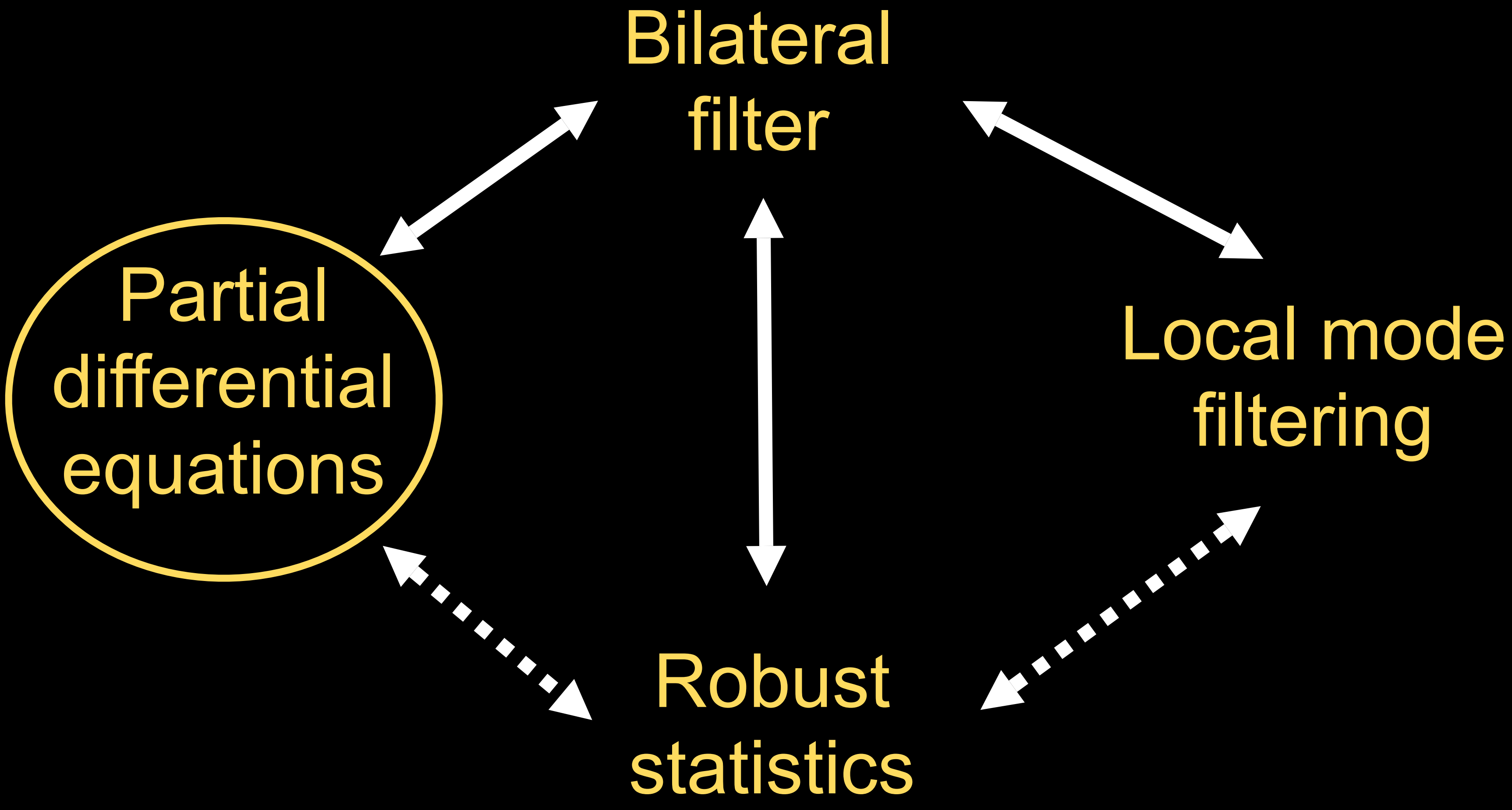
W-estimator

Take home message #2

Bilateral filter can be interpreted in term of robust statistics since it is related to a cost function!

[Durand, Dorsey, 2002, Black, Marimont, 1998, etc]

Goal: Understand how does bilateral filter relates with other methods



Disclaimer

- We will shrink the neighborhood
- This will lose some properties of the bilateral filter
- But although partial, this parallel is insightful

What do I mean by PDEs?

- Images live in a **continuous** domain
- Two kinds of formulations
  - Variational approach

$$\inf_{I} \int_{p\in\Omega} F(p, I, \nabla I)\,dp$$

  - Evolving a partial differential equation

$$\frac{\partial I}{\partial t} = G(p, I, \nabla I, H(I))$$

Recall robust statistics

$$min \sum_{p \in \Omega} \sum_{q \in \eta_p^4} \rho(I_p - I_q)$$

$$I_p^{t+1} = I_p^t + \frac{\lambda}{|\eta_p|} \sum_{q \in \eta_p} G_{\sigma_s}(q - p) g(I_q^t - I_p^t)(I_q^t - I_p^t)$$

Images are continuous

$$\inf_I \int_{p\in\Omega} \rho(\nabla I)\,dp$$

$$\frac{\partial I}{\partial t} = \mathrm{div}\left(g(\nabla I)\nabla I\right)$$

Images are discrete

$$\inf_I \sum_{p\in\Omega} \sum_{q\in\eta_p^4} \rho(I_p - I_q)$$

$$\boxed{I_p^{t+1} = I_p^t + \frac{\lambda}{|\eta_p|} \sum_{q\in\eta_s} g(\nabla I_{p,q})\nabla I_{p,q} \quad \text{with...} \nabla I_{p,q} = I_q - I_p}$$

Some technical results to establish

- Considering the Yaroslavsky Filter

$$YNF_{\sigma_s,\sigma_r} I(p) = \frac{1}{C(p)} \int_{B_{\sigma_s}(p)} G_{\sigma_r}(I(q) - I(p)) I(q) dq$$

- When

$$\sigma_s, \sigma_r \to 0 / \sigma_r = \mathcal{O}(\sigma_s^\alpha), \text{ and } \alpha = 1$$

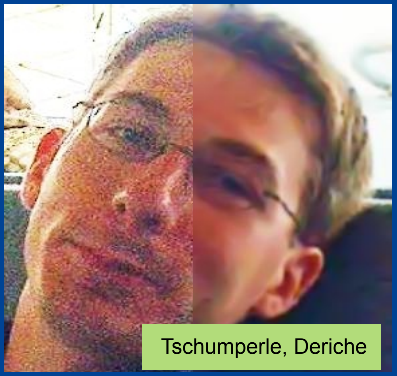
$$YNF_{\sigma_s,\sigma_r} I(p) - I(p) \approx \text{nonlinear diffusion operator}$$

(operation similar to M-estimators)

At a very local scale, the asymptotic behavior of the integral operator corresponds to a diffusion operator

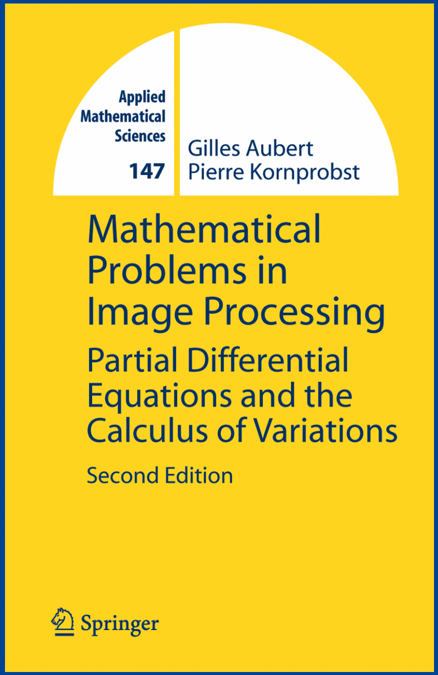
[Buades, Coll, Morel, 2005]

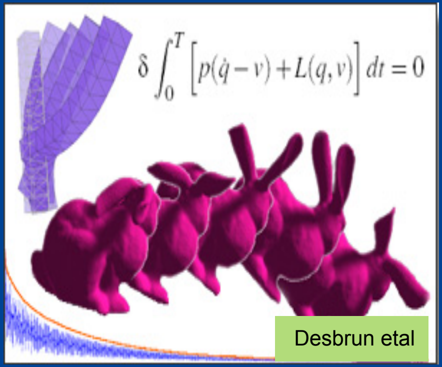
The PDE world at a glance


Tschumperle, Deriche

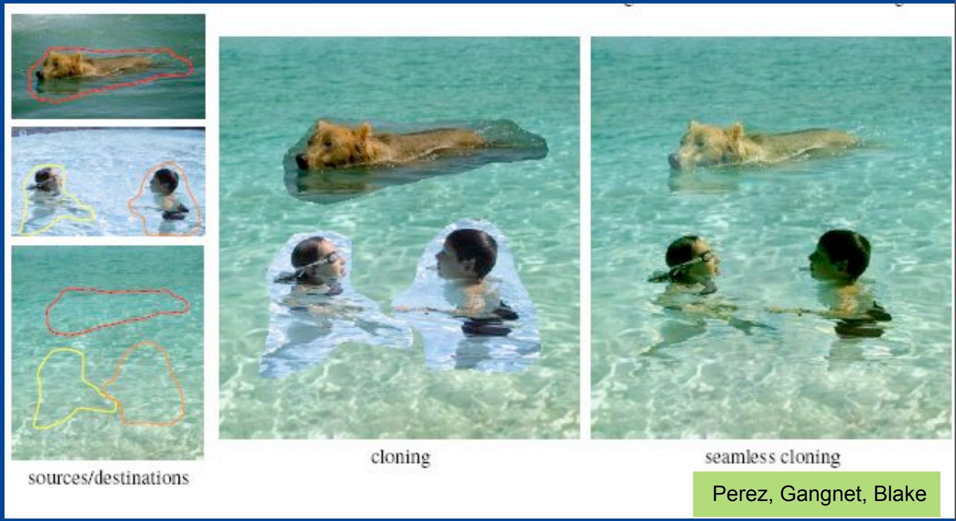

Tschumperle, Deriche


Sussman

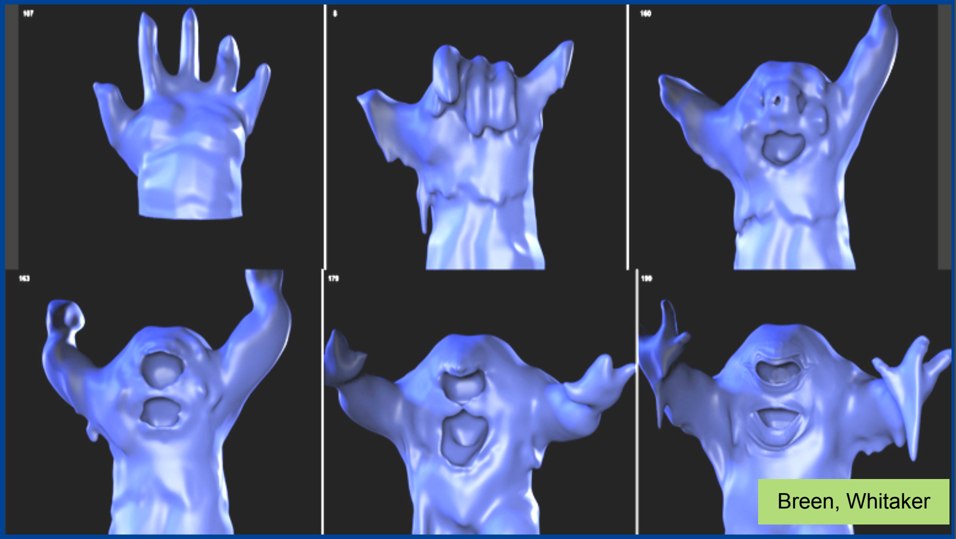




Desbrun etal



Perez, Gangnet, Blake


Breen, Whitaker

Discussion

- We shrunk the kernel
-

Take home message #3

Bilateral filter is a discretization of a particular kind of a PDE-based anisotropic diffusion.

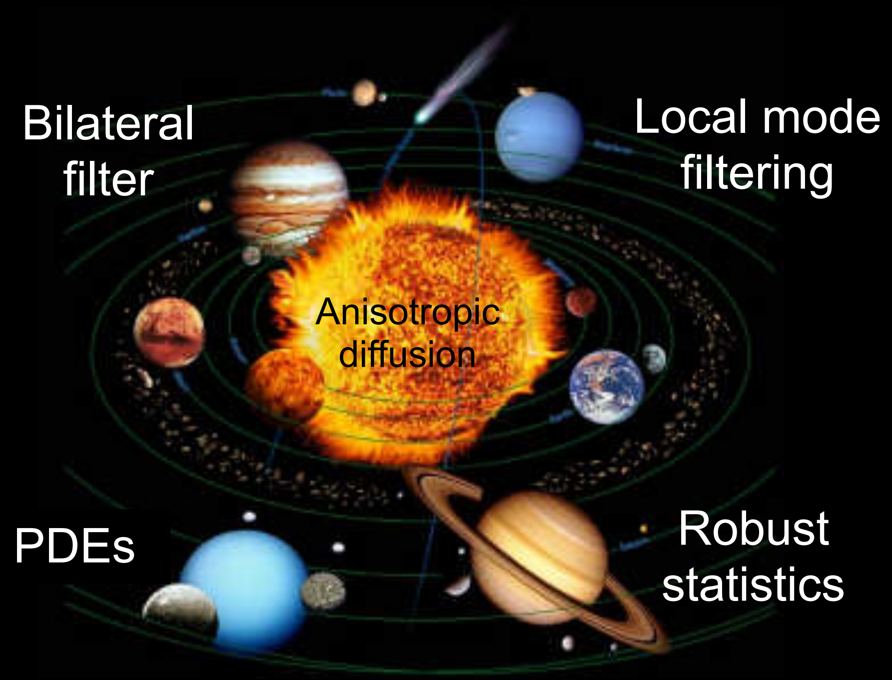
[Barash 2001, Elad 2002, Durand 2002, Buades, Coll, Morel, 2005]

Welcome to the PDE-world!

[Kornprobst 2006]

Summary

Bilateral filter is one technique for anisotropic diffusion and it makes the bridge between several frameworks. From there, you can explore news worlds!



Questions?