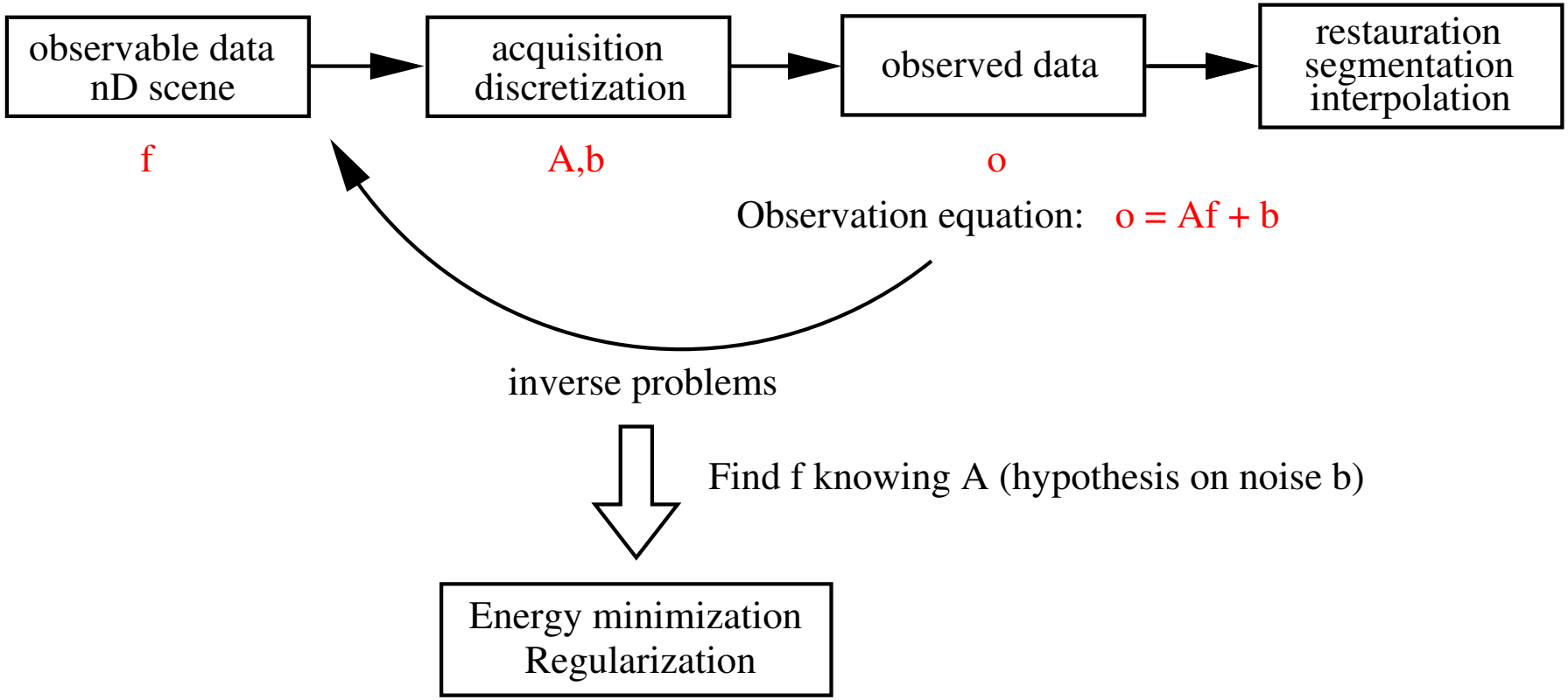
Regularization in Image Processing

Sébastien Bougleux

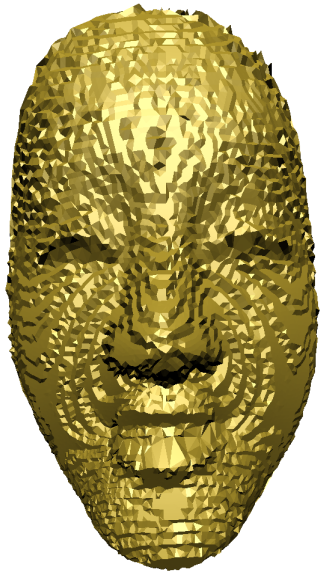
Université de Caen Basse-Normandie
GREYC - CNRS UMR 6072 - Image Team

October 12, 2008

Variational Models: Motivations



Variational Models: Motivations



Problem

Given:

- a weighted graph (V, W),
- a fonction $f^0 : V \to \mathbb{R}^m$ such that $f^0 = Af + n$.

Approximate $f : V \to \mathbb{R}^m$ such that it is sufficiently regular and sufficiently close to f^0

Continuous methods

Continuous variational methods:

- hypothesis : $f^0 : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ continuous, dfined on a continuous domain Ω,
- minimization of an energy of the form $E(f, f^0, \lambda) = E_{regu}(f) + \lambda E_{approx}(f, f^0)$,

$$\text{example : } \min_{f:\Omega \rightarrow \mathbb{R}} \left\{ \int_\Omega | \bigtriangledown f|^p + \lambda \|f - f^0\|_2^2 \right\}$$

- dicretization of the solution

Principal problems:

- definition of differential operators for spaces of complex topology
- data defined on spaces of high dimension

Solution: digitalization of continuous methods

- $f^0 : V \rightarrow \mathbb{R}$ discrete, defined on the vertices of (V, W)
- remplace $\int_\Omega$ by $\sum_V$, definition of discrete differential operators

Outline

1. Preliminary Definitions
2. Operators and Diffusion
3. Non-Adaptive Regularization
4. Adaptive Regularization

Outline

1. Preliminary Definitions
2. Operators and Diffusion
3. Non-Adaptive Regularization
4. Adaptive Regularization

Graphs

Graph (V, W)

- V a set of vertices with $|V| = n$
- $W = (w_{ij})_{i,j=1,\cdots,n}$ a similarity matrix with $w_{ij} \geq 0$
- (V, W) undirected if W is symmetric ($w_{ij} = w_{ji}$), and directed otherwise
- $E \subset V \times V$ a set of edges $e_{ij} = (i, j)$
- $N_i = \{j \in V \ : \ (i, j) \in E\}$ neighborhood of $i \in V$

Degree of a vertex

- directed graph: outgoing and ingoing sum of weights

$$d_i^{\text{out}} \stackrel{\text{def.}}{=} \frac{1}{n} \sum_{j \in V} w_{ij}, \qquad d_i^{\text{in}} \stackrel{\text{def.}}{=} \frac{1}{n} \sum_{j \in V} w_{ji}, \qquad \text{with } \underbrace{d_i^{\text{in}} + d_i^{\text{out}} > 0, \ \forall i \in V}_{\text{connected graph}}$$

- undirected graph: $d_i^{\text{out}} \equiv d_i^{\text{in}}$

$$d_i \stackrel{\text{def.}}{=} \frac{1}{n} \sum_{j \in V} w_{ij}$$

Functions on Graphs

Functions on vertices

- function space $\mathbb{R}^V$ with inner product

$$\langle f, g \rangle_V \stackrel{\text{def.}}{=} \frac{1}{n} \sum_{i \in V} f_i g_i \varphi_i$$

with $f, g : V \to \mathbb{R}$

and $\varphi_i \stackrel{\text{def.}}{=} \varphi_{\text{in}}(d_i^{\text{in}}) + \varphi_{\text{out}}(d_i^{\text{out}})$

- Hilbert space
$\mathcal{H}(V, \varphi) \stackrel{\text{def.}}{=} (\mathbb{R}^V, \langle \cdot, \cdot \rangle_V)$

Functions on edges

- function space $\mathbb{R}^E$ with inner product

$$\langle F, G \rangle_E \stackrel{\text{def.}}{=} \frac{1}{2n^2} \sum_{i \in V} \sum_{j \in V} F_{ij} G_{ij} \phi_{ij}$$

with $F, G : E \to \mathbb{R}$

and $\phi_{ij} \stackrel{\text{def.}}{=} \phi(w_{ij})$

- Hilbert space
$\mathcal{H}(E, \phi) \stackrel{\text{def.}}{=} (\mathbb{R}^E, \langle \cdot, \cdot \rangle_E)$

2D Images and Graphs

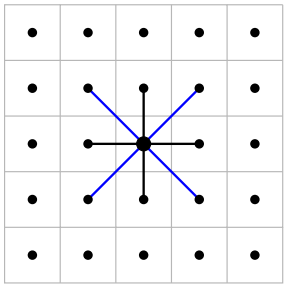
Image $f : \Omega \subset \mathbb{Z}^2 \rightarrow \mathbb{R}$ becomes $f : V \rightarrow \mathbb{R}$
Each pixel $p_i = (x_i, y_i) \in \Omega$ becomes a vertex $i \in V$

Adjacency grid graphs: natural representation

- 4-adjacency grid graph
- 8-adjacency grid graph

$\Rightarrow$ undirected graphs

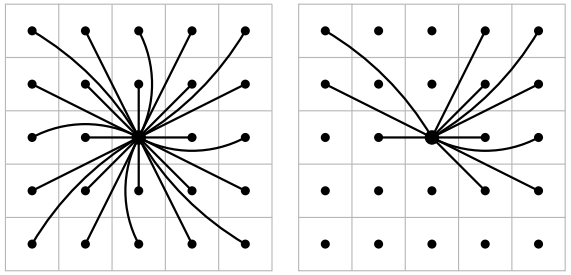
Neighborhoods



$w_{ij} = 1$, $w_{ij} = \frac{1}{\sqrt{2}}$

Neighborhood graphs

- ϵ-neighborhood graph: $j \in N_i \Leftrightarrow \text{dist}(i,j) \leq \epsilon$
  - $\text{dist}(i,j) = \|p_i - p_j\|_2$
  - $\text{dist}(i,j) =$ Chebyshev distance (patches)
  - $\text{dist}(i,j) = |f_i - f_j|$
  - ...
- k-nearest neighbors graph (directed)
- ...



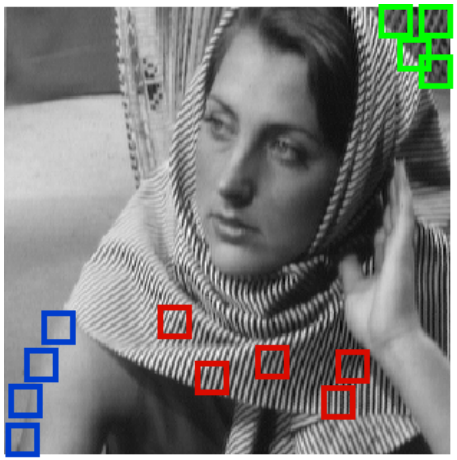
$w_{ij} = \frac{1}{\text{dist}(i,j)}$

Weight Functions

Image $f : \Omega \subset \mathbb{Z}^2 \rightarrow \mathbb{R}$ becomes $f : V \rightarrow \mathbb{R}$

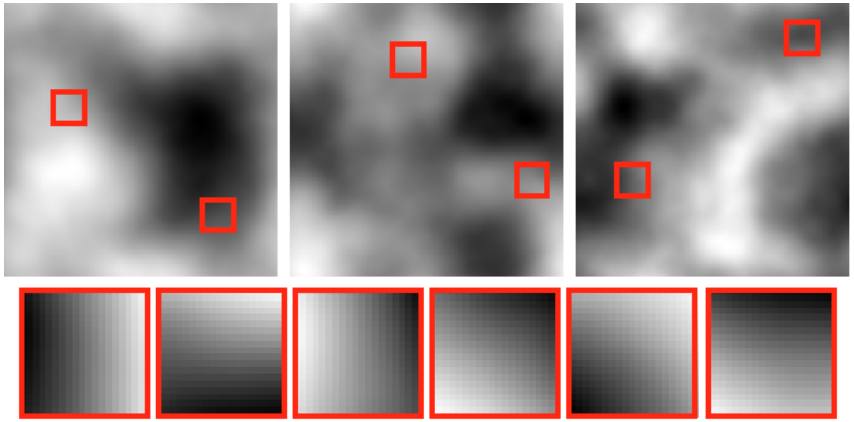
Whatever the structure of the graph, weights are generally computed using:

- $w_{ij} = \frac{1}{\|f_i - f_j\|_2^{\alpha} + \epsilon}$, $\alpha > 0$
- $w_{ij} = \exp\left(-\frac{\|f_i - f_j\|_2^2}{\sigma^2}\right)$
- $w_{ij} = \exp\left(-\frac{\|p_i - p_j\|_2^2}{\sigma_\Omega^2}\right) \exp\left(-\frac{\|f_i - f_j\|_2^2}{\sigma^2}\right)$
- $w_{ij} = \exp\left(-\frac{\|\text{patch}_i - \text{patch}_j\|_{2,a}^2}{h^2}\right)$

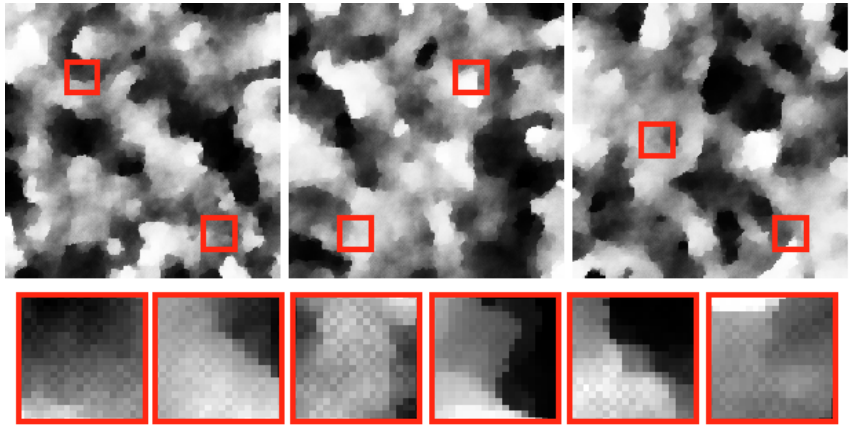


similar patches

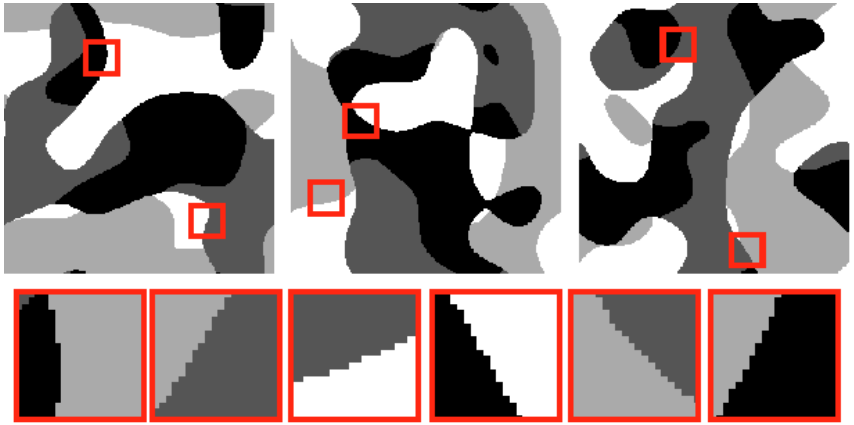
Patch-Based Geometry of Images



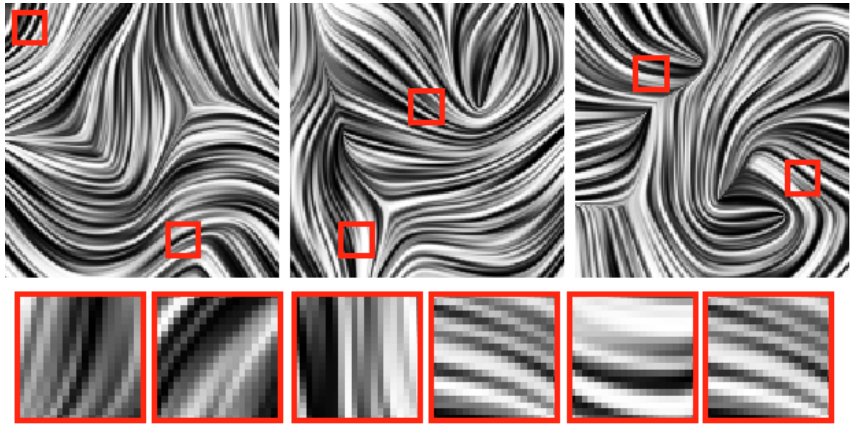
uniformly smooth images
patches: linear gradient of intensity



discontinuous images with bounded variations
patches: finite length level sets



cartoon images
patches: binary straight edges



locally parallel textures
patches: directional oscillations

Outline

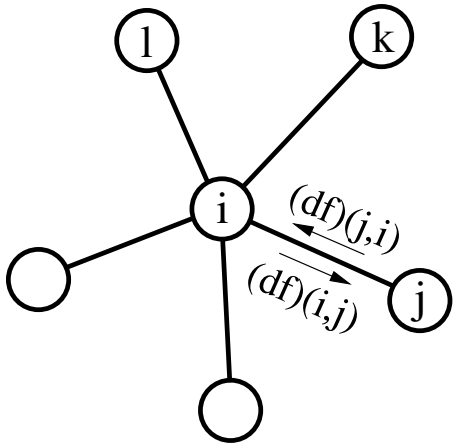
1. Preliminary Definitions
2. Operators and Diffusion
3. Non-Adaptive Regularization
4. Adaptive Regularization

Difference Operator and its Adjoint

Difference operator $d : \mathcal{H}(V, \varphi) \rightarrow \mathcal{H}(E, \phi)$

$$(df)_{ij} \stackrel{\text{def.}}{=} \gamma_{ij} \left(\alpha_j f_j - \alpha_i f_i\right)$$

- with $\gamma : V \times V \rightarrow \mathbb{R}_+$
- *example*: $\gamma_{ij} = \sqrt{w_{ij}}$ and $\alpha_i = \sqrt{1/d_i}$
- bounded, linear, antisymmetric
- $(\nabla f)_i = ((df)_{ij})_{j \in V}$



Adjoint $d^* : \mathcal{H}(E, \phi) \rightarrow \mathcal{H}(V, \varphi)$ of difference d

$$\langle df, G \rangle_E \stackrel{\text{def.}}{=} \langle f, d^* G \rangle_V$$

$$(d^* G)(i) = \underbrace{\frac{\alpha_i}{2n\varphi_i} \sum_{j \in V} \gamma_{ji} G_{ji} \phi_{ji}}_{\text{outgoing flow}} - \underbrace{\frac{\alpha_i}{2n\varphi_i} \sum_{j \in V} \gamma_{ij} G_{ij} \phi_{ij}}_{\text{ingoing flow}}$$

Laplace Operator

General Laplacian $\Delta : \mathcal{H}(V,\varphi) \to \mathcal{H}(V,\varphi)$

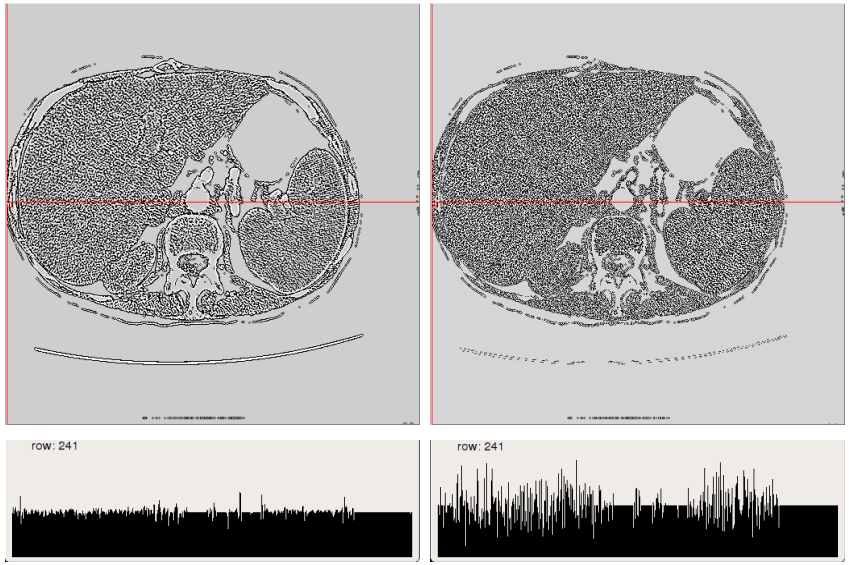
$$\Delta \stackrel{\text{def.}}{=} d^* d$$

$$(\Delta f)_i = \frac{\alpha_i}{2n\varphi_i} \sum_{j \in V} \left(\gamma_{ij}^2 \phi_{ij} + \gamma_{ji}^2 \phi_{ji} \right) (\alpha_i f_i - \alpha_j f_j)$$

$\Rightarrow$ linear

Undirected graph

$\gamma_{ij} = \gamma_{ji}$



unweighted graph | weighted graph

Combinatorial Laplacian

$$(\Delta^{(c)} f)_i = \delta_i f_i - \frac{1}{n} \sum_{j \in V} w_{ij} f_j$$

$$\Delta^{(c)} f = (D - W) f$$

with $\alpha_i = 1$, $\gamma_{ij} = \sqrt{w_{ij}}$, $\varphi_i = 1$, $\phi_{ij} = 1$

Random walk Laplacian

$$(\Delta^{(r)} f)_i = f_i - \frac{1}{n\delta_i} \sum_{j \in V} w_{ij} f_j$$

$$\Delta^{(r)} f = (I - D^{-1} W) f$$

with $\alpha_i = 1$, $\gamma_{ij} = \sqrt{w_{ij}}$, $\varphi_i = \delta_i$, $\phi_{ij} = 1$

Normalized Laplacian

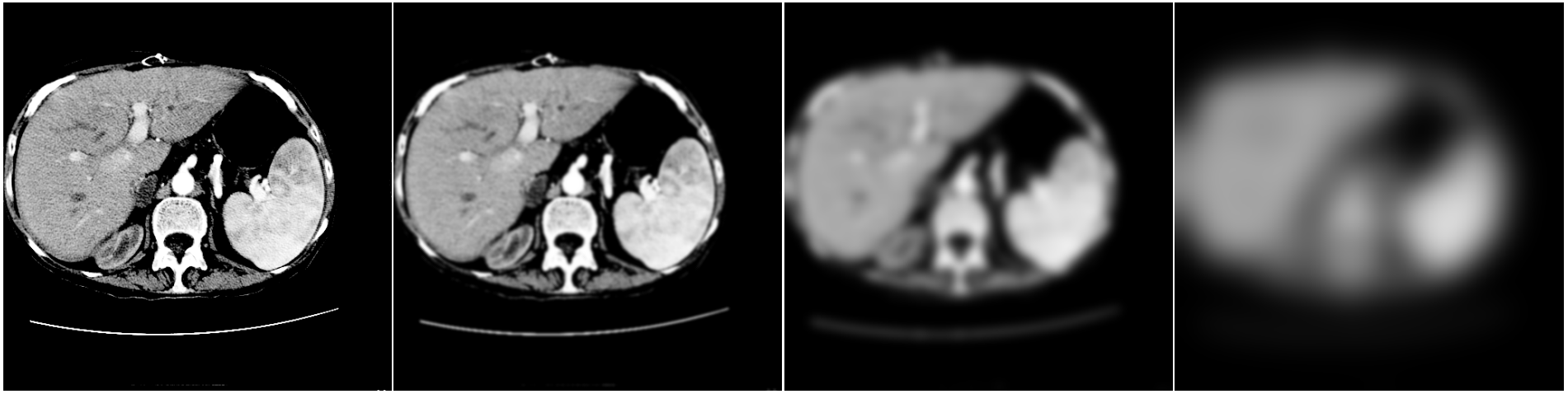
$$(\Delta^{(n)} f)_i = f_i - \frac{1}{n\sqrt{\delta_i}} \sum_{j \in V} \frac{w_{ij}}{\sqrt{\delta_j}} f_j$$

$$\Delta^{(n)} f = (I - D^{-\frac{1}{2}} W D^{-\frac{1}{2}}) f$$

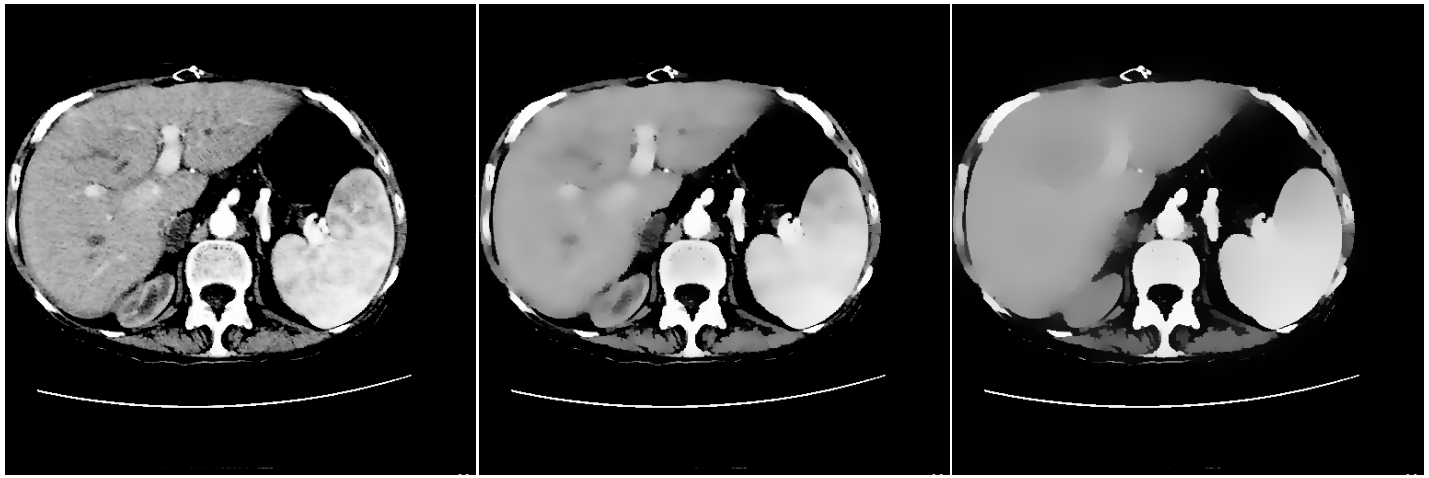
with $\alpha_i = 1/\sqrt{\delta_i}$, $\gamma_{ij} = \sqrt{w_{ij}}$, $\varphi_i = 1$, $\phi_{ij} = 1$

Laplacian Diffusion

$\partial_t f_i = (\Delta f)_i, \ \forall i \in V$ with combinatorial Laplacian and 4-adjacency grid graph



unweighted graph



weighted graph

Curvature Operator

Curvature operator $\kappa : \mathcal{H}(V,\varphi) \to \mathcal{H}(V,\varphi)$

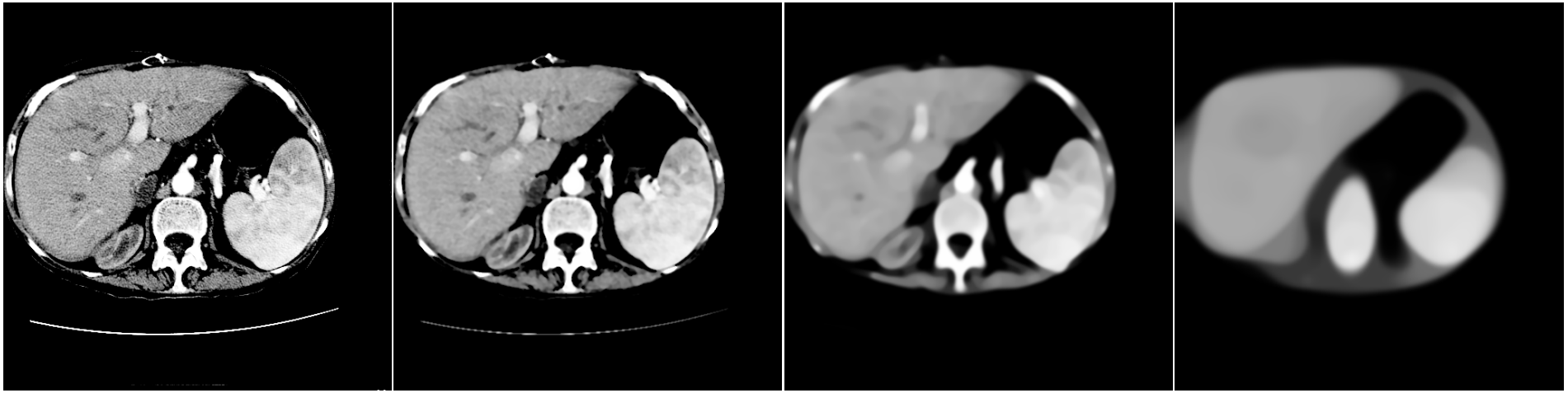
$$\kappa f \stackrel{\text{def.}}{=} d^* \left(\frac{df}{\|df\|_2} \right)$$

$$(\kappa f)_i = \frac{\alpha_i}{2n\varphi_i} \sum_{j \in V} \left(\frac{\gamma_{ij}^2 \phi_{ij}}{\|(\nabla f)_i\|_2} + \frac{\gamma_{ji}^2 \phi_{ji}}{\|(\nabla f)_j\|_2} \right) (\alpha_i f_i - \alpha_j f_j)$$

$\Rightarrow$ nonlinear

Diffusion Based on Curvature

$\partial_t f_i = (\kappa f)_i, \ \forall i \in V$ with combinatorial curvature and 4-adjacency grid graph



unweighted graph

weighted graph

p-Laplacian Operator

Graph p-Laplacian $\Delta_p : \mathcal{H}(V,\varphi) \to \mathcal{H}(V,\varphi)$, $p \in (0,+\infty)$

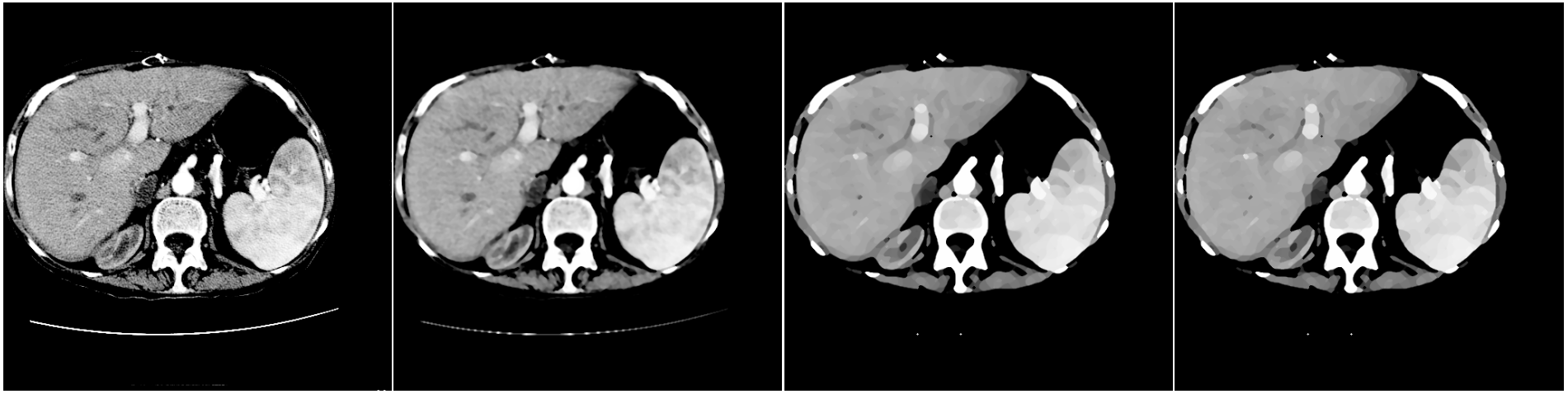
$$\Delta_p f \stackrel{\text{def.}}{=} d^* \left(\frac{df}{\|df\|_2^{2-p}} \right)$$

$$(\Delta_p f)_i = \frac{\alpha_i}{2n\varphi_i} \sum_{j \in V} \left(\frac{\gamma_{ij}^2 \phi_{ij}}{\|(\nabla f)_i\|_2^{2-p}} + \frac{\gamma_{ji}^2 \phi_{ji}}{\|(\nabla f)_j\|_2^{2-p}} \right) (\alpha_i f_i - \alpha_j f_j)$$

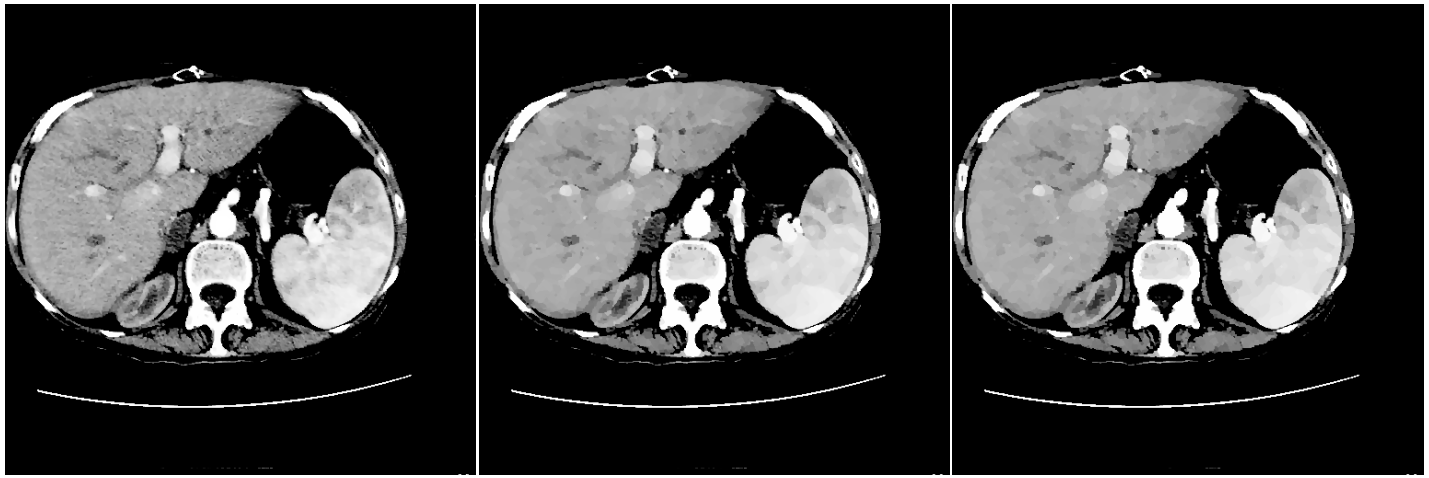
$\Rightarrow$ nonlinear for $p \neq 2$
$\Rightarrow$ extends Laplace and curvature operators

p-Laplacian Diffusion

$\partial_t f_i = (\Delta_p f)_i,\ \forall i \in V$ with combinatorial curvature, 4-adjacency grid graph and $p = 0.7$



unweighted graph



weighted graph

p-Laplacian Operator Based on Differences

Graph p-Laplacian $\tilde{\Delta}_p : \mathcal{H}(V,\varphi) \to \mathcal{H}(V,\varphi)$, $p \in (0,+\infty)$

$$\tilde{\Delta}_p f \stackrel{\text{def.}}{=} d^* \left(\frac{df}{|df|^{2-p}} \right)$$

$$(\tilde{\Delta}_p f)_i = \frac{\alpha_i}{2n\varphi_i} \sum_{j \in V} \left(\frac{\gamma_{ij}^2 \phi_{ij}}{|(df)_{ij}|^{2-p}} + \frac{\gamma_{ji}^2 \phi_{ji}}{|(df)_{ji}|^{2-p}} \right) (\alpha_i f_i - \alpha_j f_j)$$

$\Rightarrow$ nonlinear for $p \neq 2$
$\Rightarrow$ extends Laplace operator with differences

Outline

1. Preliminary Definitions
2. Operators and Diffusion
3. Non-Adaptive Regularization
4. Adaptive Regularization

Discrete Regularization on graphs

Considered discrete variational problem

Given a graph (V, W) and a function $f^0 : V \to \mathbb{R}$, find a function $f : V \to \mathbb{R}$ which is sufficiently regular on G, and sufficiently close to f^0:

$$\underset{f:V\to\mathbb{R}}{\text{argmin}} \ \{ \ E(f, f^0, \lambda, p) = \underbrace{\frac{1}{2p}\sum_{i\in V} \|(\nabla f)_i\|_2^p}_{\text{regularization term} J(f)} + \underbrace{\frac{\lambda}{2}\sum_{i\in V}(f_i - f_i^0)^2}_{\text{fitting term } \|f-f^0\|_2^2} \ \}$$

Special cases

- $p = 2$: Tikhonov regularization (for example [Chung, 97])
- $p = 1$: fitted TV regularization (on unweighted graphs [Osher and Shen, 00])

Existance of the minimizer

- $p \geq 1$: unique solution (E is strickly convex)
- $p < 1$: the global minimizer may not exist (E is non-convex)

Discrete Regularization on graphs

Considered discrete variational problem

Given a graph (V, W) and a function $f^0 : V \to \mathbb{R}$, find a function $f : V \to \mathbb{R}$ which is sufficiently regular on G, and sufficiently close to f^0:

$$\underset{f:V\to\mathbb{R}}{\text{argmin}} \ \{ \ E(f, f^0, \lambda, p) = \underbrace{\frac{1}{2p}\sum_{i\in V} \|(\nabla f)_i\|_2^p}_{\text{regularization term} J(f)} + \underbrace{\frac{\lambda}{2}\sum_{i\in V}(f_i - f_i^0)^2}_{\text{fitting term } \|f-f^0\|_2^2} \ \}$$

Regularization functional

- nonlocal expression:

$$J(f) = \frac{1}{2p}\sum_{i\in V}\left(\sum_{j\in V} w_{ij}(f_j - f_i)^2\right)^{\frac{p}{2}}$$

- continous analogue ([Kinderman et al., 05] [Gilboa and Osher, 07] pour $p = 2$ et $p = 1$):

$$J(f) = \frac{1}{p}\int_\Omega \left(\int_\Omega w(x,y)(f(y) - f(x))^2 dy\right)^{\frac{p}{2}} dx, \quad f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$$

Discrete Regularization on graphs

Considered discrete variational problem

Given a graph (V, W) and a function $f^0 : V \to \mathbb{R}$, find a function $f : V \to \mathbb{R}$ which is sufficiently regular on G, and sufficiently close to f^0:

$$\underset{f:V\to\mathbb{R}}{\text{argmin}} \ \{ \ E(f, f^0, \lambda, p) = \underbrace{\frac{1}{2p}\sum_{i\in V} \|(\nabla f)_i\|_2^p}_{\text{regularization term} J(f)} + \underbrace{\frac{\lambda}{2}\sum_{i\in V}(f_i - f_i^0)^2}_{\text{fitting term } \|f-f^0\|_2^2} \ \}$$

Solution: system of equations based on the p-Laplacian

Euler-Lagrange equation

$$\frac{\partial J(f)}{\partial f_i} + \lambda(f_i - f_i^0) = 0, \quad \forall i \in V$$

Property: $\frac{\partial J(f)}{\partial f_i} = (\Delta_p f)_i \ \forall i \in V$

Discrete Diffusion Process

Considered discrete variational problem

Given a graph (V, W) and a function $f^0 : V \to \mathbb{R}$, find a function $f : V \to \mathbb{R}$ which is sufficiently regular on G, and sufficiently close to f^0:

$$\underset{f:V\to\mathbb{R}}{\text{argmin}} \ \{ \ E(f, f^0, \lambda, p) = \underbrace{\frac{1}{2p}\sum_{i\in V} \|(\nabla f)_i\|_2^p}_{\text{regularization term} J(f)} + \underbrace{\frac{\lambda}{2}\sum_{i\in V}(f_i - f_i^0)^2}_{\text{fitting term } \|f-f^0\|_2^2} \ \}$$

Solution: discrete diffusion process

$$\begin{cases} f^{(0)} = f^0 \\ f_i^{(t+1)} = \dfrac{\lambda f_i^0 + \sum_{j\in V} \gamma_w^p(f^{(t)}, i, j) f_j^{(t)}}{\lambda + \sum_{j\in V} \gamma_w^p(f^{(t)}, i, j)}, \quad \forall j \in V. \end{cases}$$

where $\gamma_w^p(f, i, j) = w_{ij}(\|(\nabla f)_i\|_2^{p-2} + \|(\nabla f)_j\|_2^{p-2})$

Unified Framework

Depending on the choice of:

- degree of regularity p
- fitting parameter λ
- structure of the graph

Different processes using the discrete diffusion:

- Spectral analysis on graphs and Laplacien smoothing ($p = 2$, $\lambda = 0$)
- Image processing:
  - Regular grid:
    - *digital TV filter* ($p = 1$, $w = 1$) [Osher and Shen, 00]
    - Bilateral filter ($p = 2$, $\lambda = 0$, 1 itration) [Tomasi et Manduchi, 98]
    - NL-means filter ($p = 2$, $\lambda = 0$, 1 itration) [Buades et al., 05]
  - Abstract representation of images
- Mesh processing
- Non-organized data processing

Behavior: image processing

Local regularization on the 4-adjacency graph

digital filters [Osher and Shen, 2000]

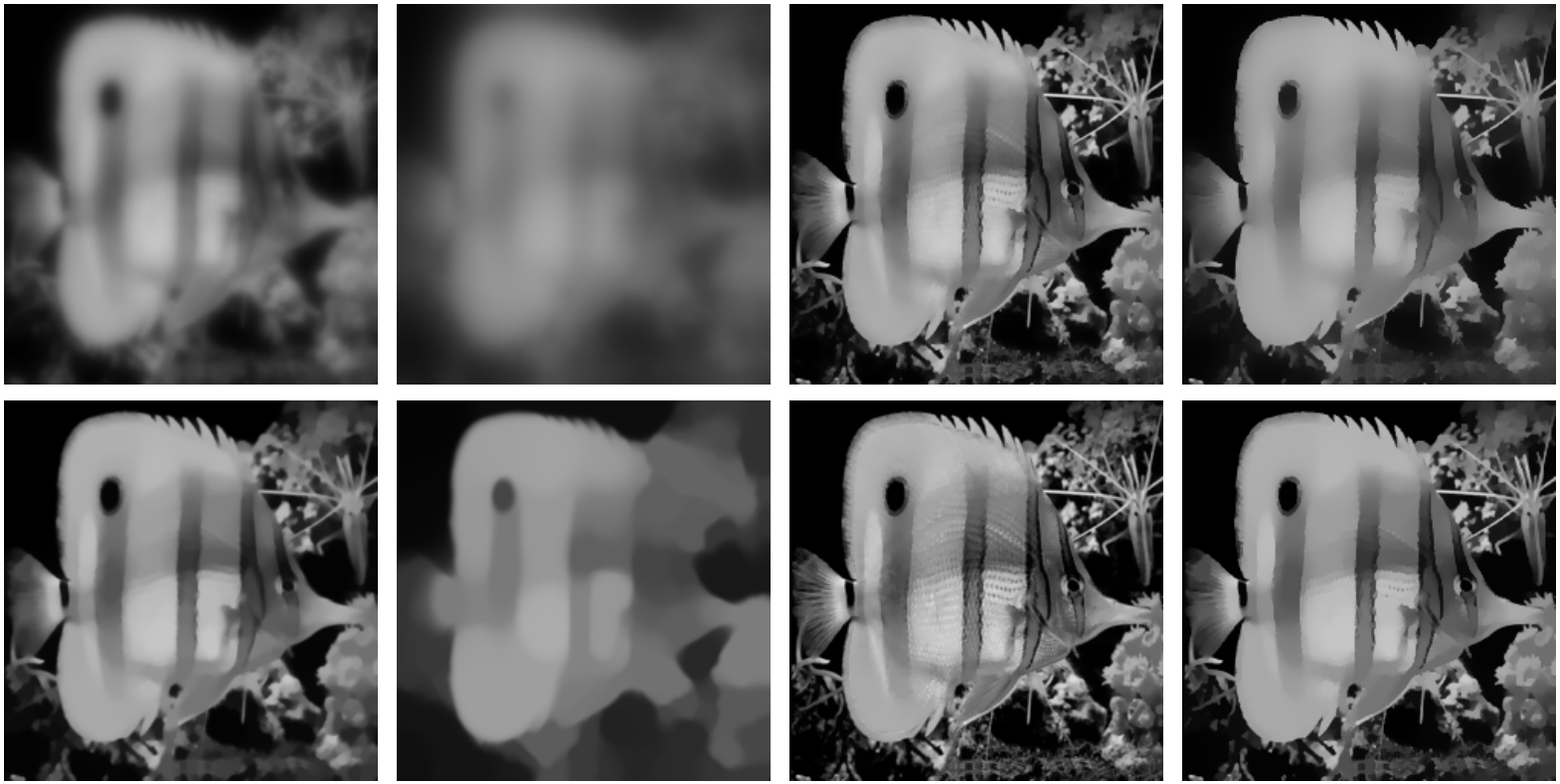
$w = 1$

[Bougleux et al, SSVM 07]

$w_{ij} = \exp\left(-\frac{\|f_i^0 - f_i^0\|_2^2}{\sigma^2}\right)$

$p = 2$

$p = 1$

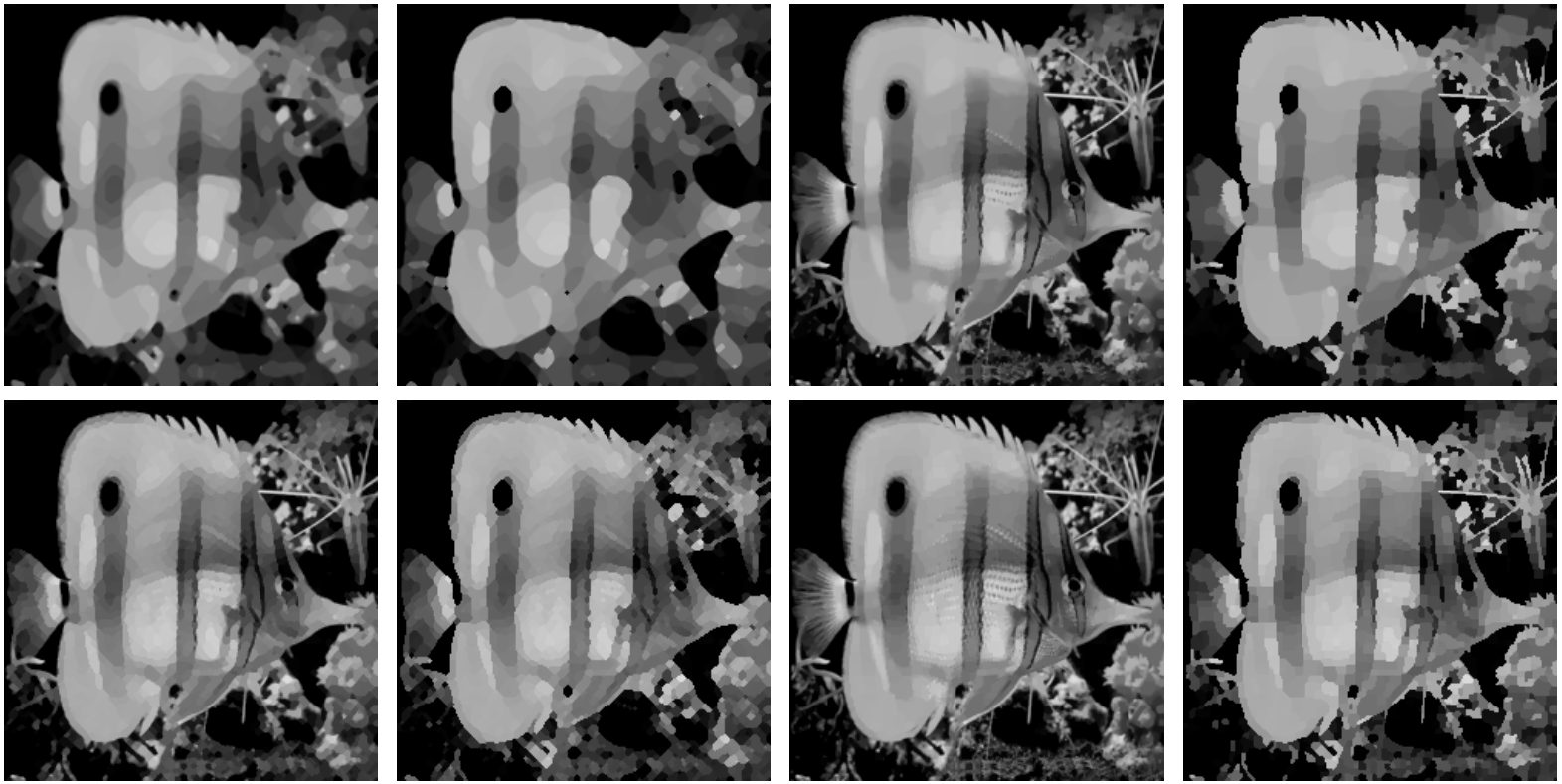


$\lambda = 0.05$ $\lambda = 0.005$ $\lambda = 0.05$ $\lambda = 0.005$

Behavior: image processing

Local regularization on the 4-adjacency graph

$$w_{ij} = \exp\left(-\frac{\|f_i^0 - f_j^0\|_2^2}{\sigma^2}\right) \qquad w_{ij} = \exp\left(-\frac{\|f_i^0 - f_i^0\|_2^2}{\sigma^2}\right)$$



$p = 0.7$

$p = 0.1$

$\lambda = 0.005$ $\lambda = 0$ $\lambda = 0.005$ $\lambda = 0$

Behavior: image processing

Nonlocal regularization on regular neighborhood graphs

- particular case: $p = 2$, $\lambda = 0$ and 1 iteration $\rightarrow$ NL-means filter [Buades et al., 05]
- extension for $p \neq 2$, $\lambda > 0$ and several iterations

NL-means weight:

$$w_{ij} = \exp\left(-\frac{\|\text{patch}_i(f^0), \text{patch}_j(f^0)\|_{2,a}^2}{h^2}\right).$$

$p = 2$ $p = 1$ $p = 0.7$ $p = 0.1$

Behavior: image processing

Nonlocal regularization on regular neighborhood graphs

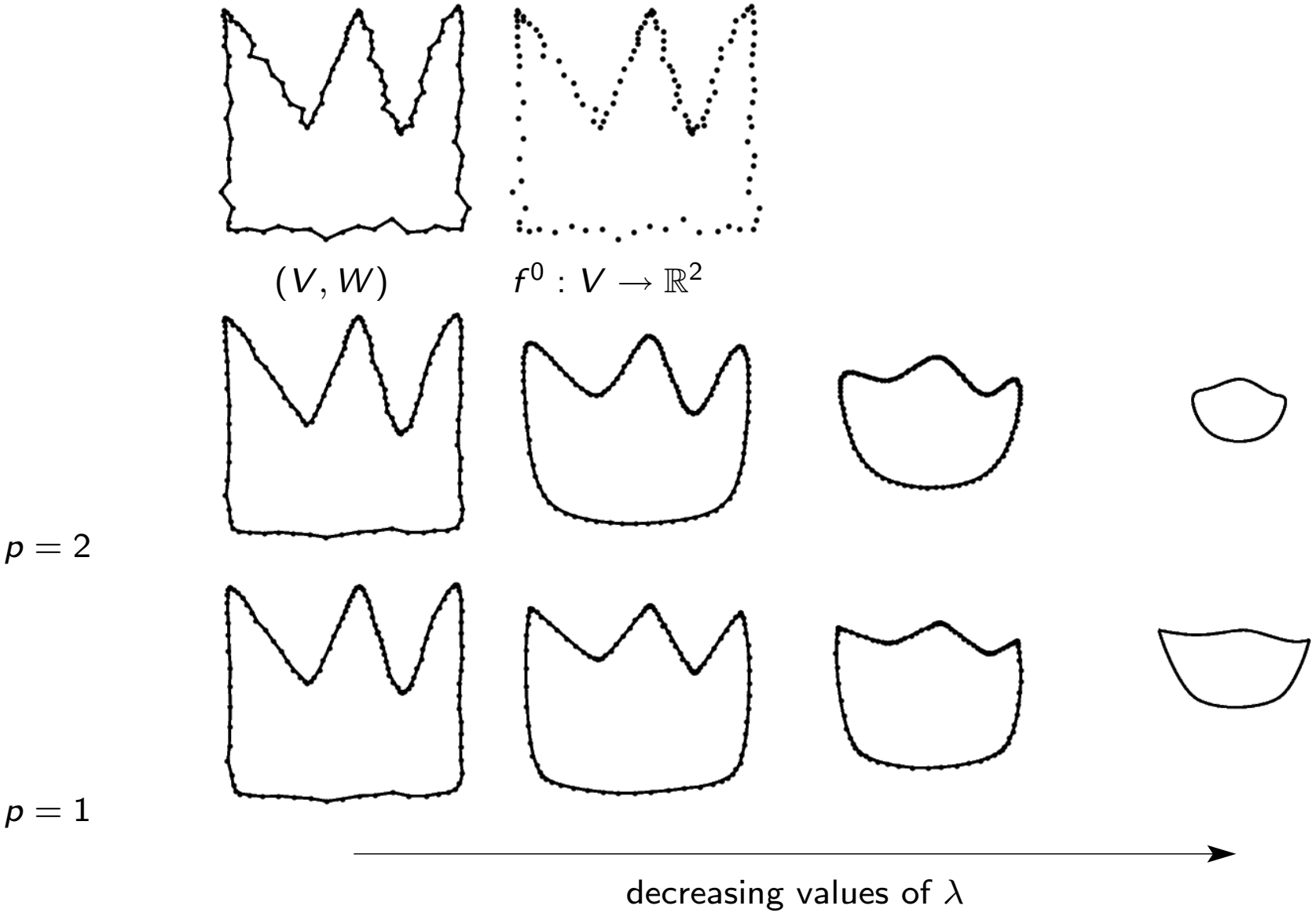
- particular case: $p = 2$, $\lambda = 0$ and 1 iteration $\rightarrow$ NL-means filter [Buades et al., 05]
- extension for $p \neq 2$, $\lambda > 0$ and several iterations

NL-means weight:

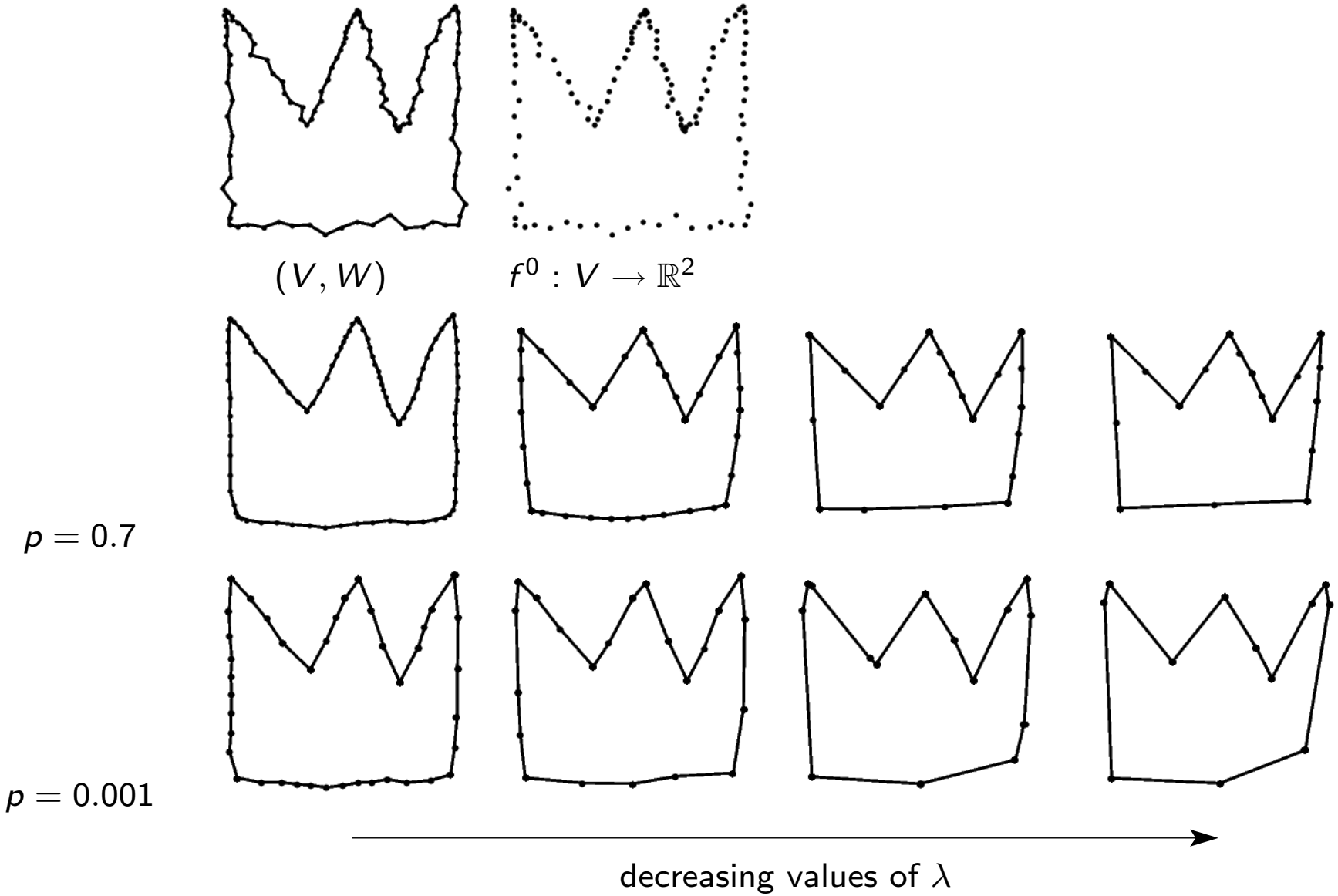
$$w_{ij} = \exp\left(-\frac{\|\text{patch}_i(f^0), \text{patch}_j(f^0)\|_{2,a}^2}{h^2}\right).$$

$p = 2$ $p = 1$ $p = 0.7$ $p = 0.1$

Behavior: polygonal curves

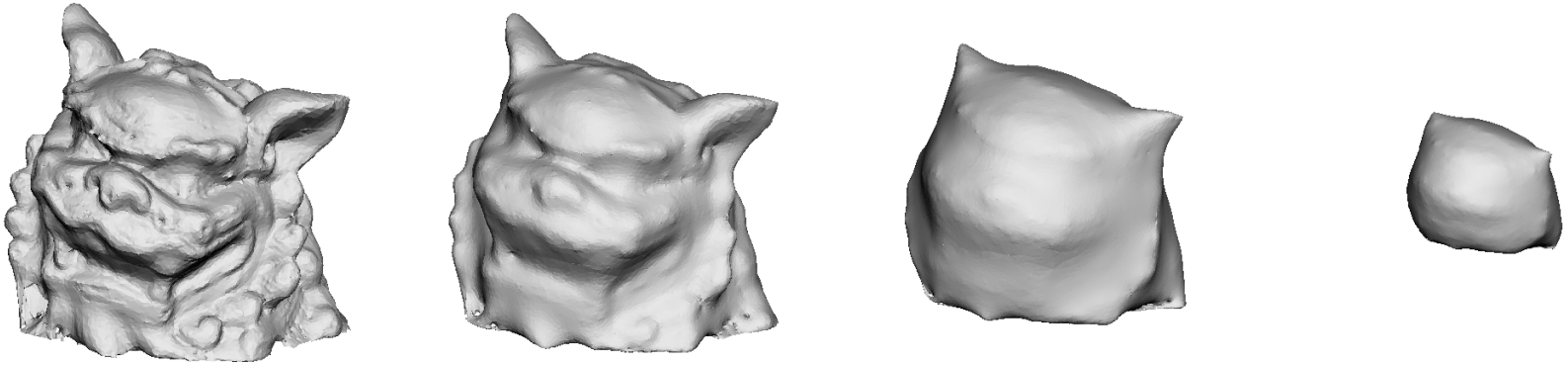


Behavior: polygonal curves



Behavior: polygonal surfaces

(V, W)
$f^0 : V \to \mathbb{R}^3$

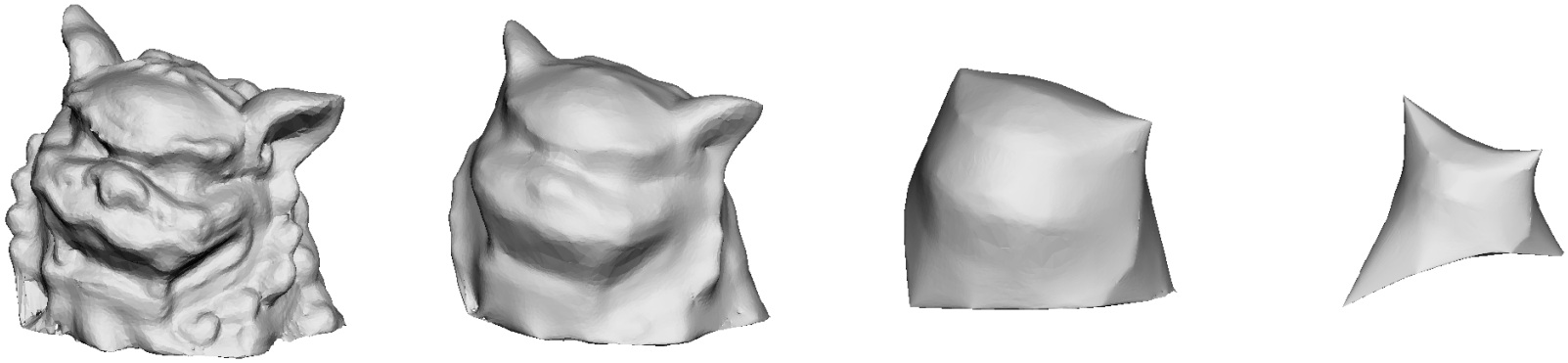


$p = 2$

decreasing values of λ

Behavior: polygonal surfaces

(V, W)
$f^0 : V \to \mathbb{R}^3$

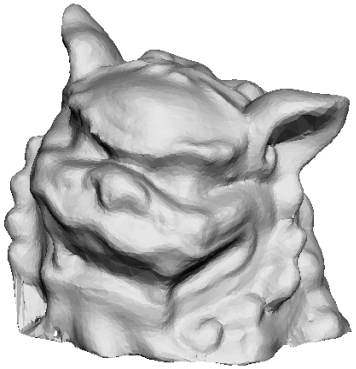


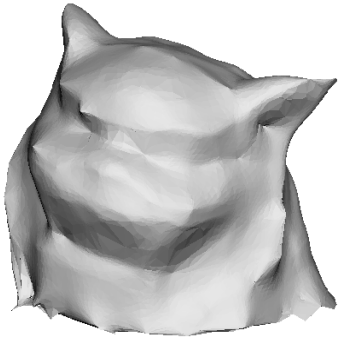
$p = 1$

decreasing values of λ

Behavior: polygonal surfaces

(V, W)
$f^0 : V \rightarrow \mathbb{R}^3$



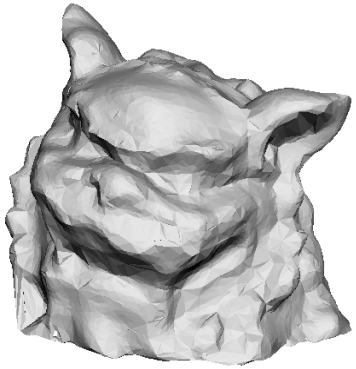


$p = 0.7$

decreasing values of λ

Behavior: polygonal surfaces

(V, W)
$f^0 : V \rightarrow \mathbb{R}^3$

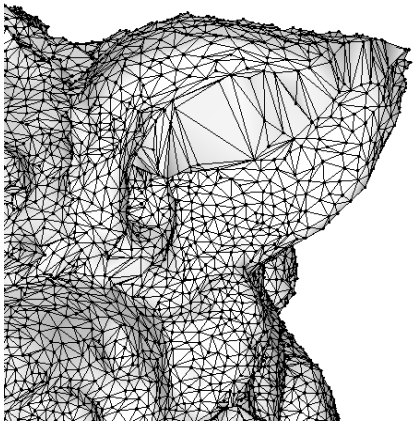


$p = 0.1$

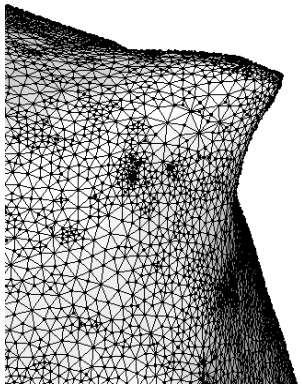
decreasing values of λ

Behavior: polygonal surfaces

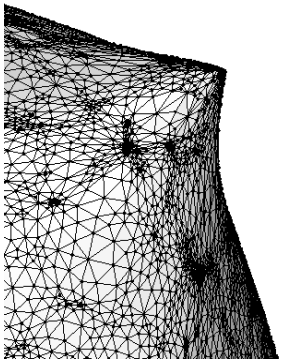
(V, W)
$f^0 : V \to \mathbb{R}^3$



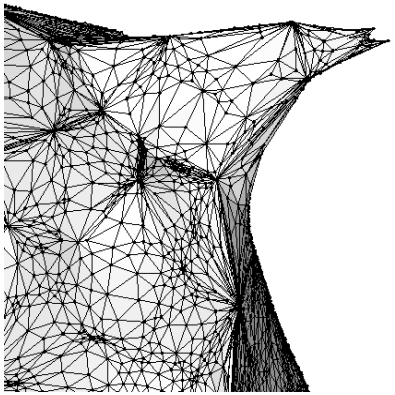
zoom



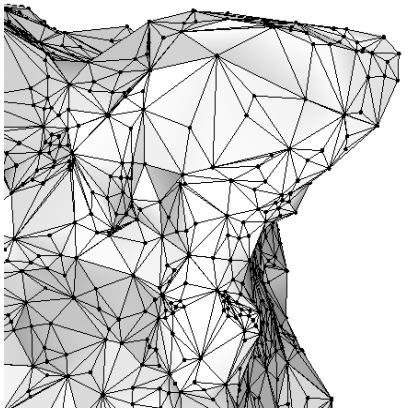
$p = 2$



$p = 1$



$p = 0.7$



$p = 0.1$

Outline

1. Preliminary Definitions
2. Operators and Diffusion
3. Non-Adaptive Regularization
4. Adaptive Regularization

Adaptive Model

Variational Problem

Find both solution $f^* : V \longrightarrow \mathbb{R}$ and adapted weights w^*

$$(f^*, w^*) = \underset{(f,w)}{\text{argmin}} \, \frac{1}{2} \|f^0 - Af\|_2^2 + \lambda J_w(f)$$

where

$$J_w(f) = \frac{1}{2p} \sum_{i \in V} \left(\sum_{j \in V} w_{ij}(f)(f_j - f_i)^2 \right)^{\frac{p}{2}}$$

Iterative Minimization Algorithm for $p = 2$

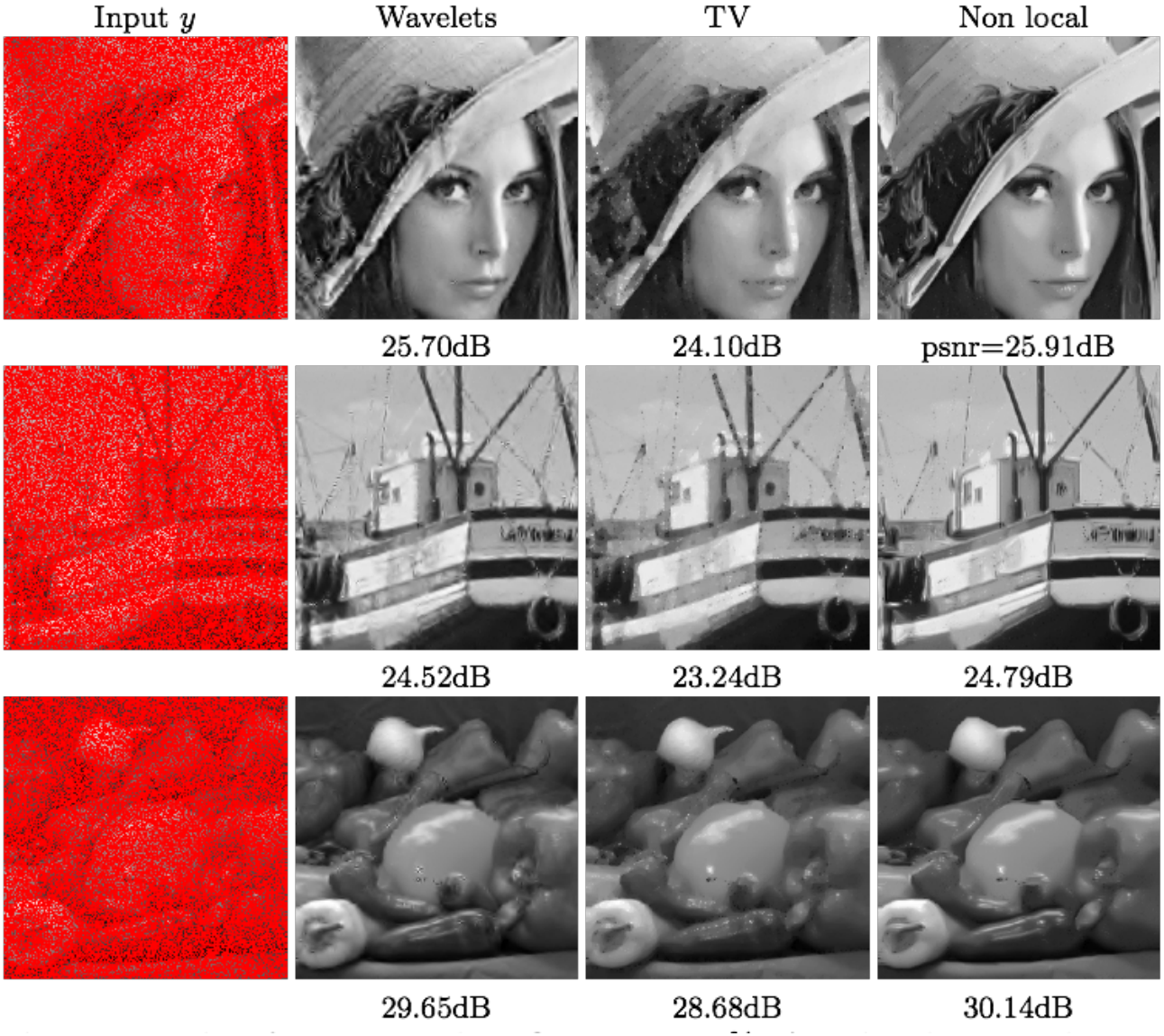
- Step 1: w fixed, find f by a gradient descent

$$f^* \leftarrow f + \tau A(Af - f^0) - \tau \lambda \Delta f$$

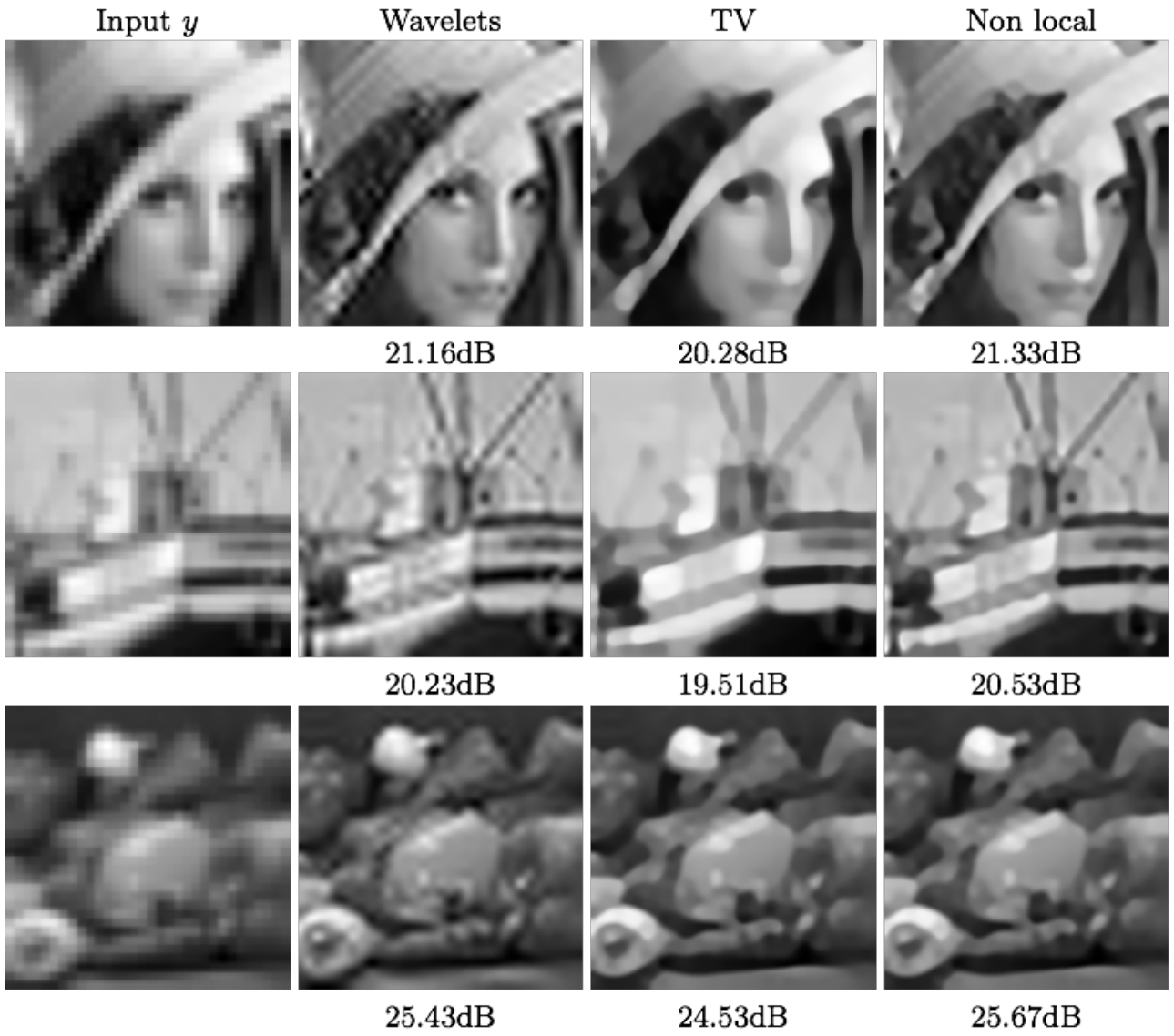
- Step 2: f^* fixed, estimate w^* and go back to step 1

For non-smooth functionals ($p = 1$), replace gradient descent by proximal iterations
See [Peyré, Bougleux and Cohen, ECCV 2008]

Inpainting Results



Super-resolution Results



Conclusion

Open problem: find both solution f^*, adpated weights w^* and adapted graph topology

Regularization in Image Processing

Sébastien Bougleux

Université de Caen Basse-Normandie
GREYC - CNRS UMR 6072 - Image Team

October 12, 2008