Medical Image Analysis
Lecture 04

Total Variation Based Denoising

Lot of material taken from PhD thesis of Thomas Pock

Martin Urschler

Basic Problem of Computer Vision

- CV deals with inverse (often ill-posed) problems:
  - Given observed data: estimate unknown quantities!
- Image Denoising / Restoration

Observed data: Noisy Image

Unknown Quantity: Clean Image

Summary Tikhonov

$$\min\left\{E = \frac{1}{2}\int_{\Omega}|\nabla u|^2 dx + \frac{\lambda}{2}\int_{\Omega}(u-f)^2 dx\right\}$$

Example: Tikhonov Denoising

- Energy Functional
  - dependence on unknown function u (continuous domain)
- Calculus of Variations gives theorem to describe a functional at stationary points
  - Setting Functional (Gateaux) derivative to zero leads to the Euler-Lagrange PDE
  - The functional has to fulfill the Euler-Lagrange equation!

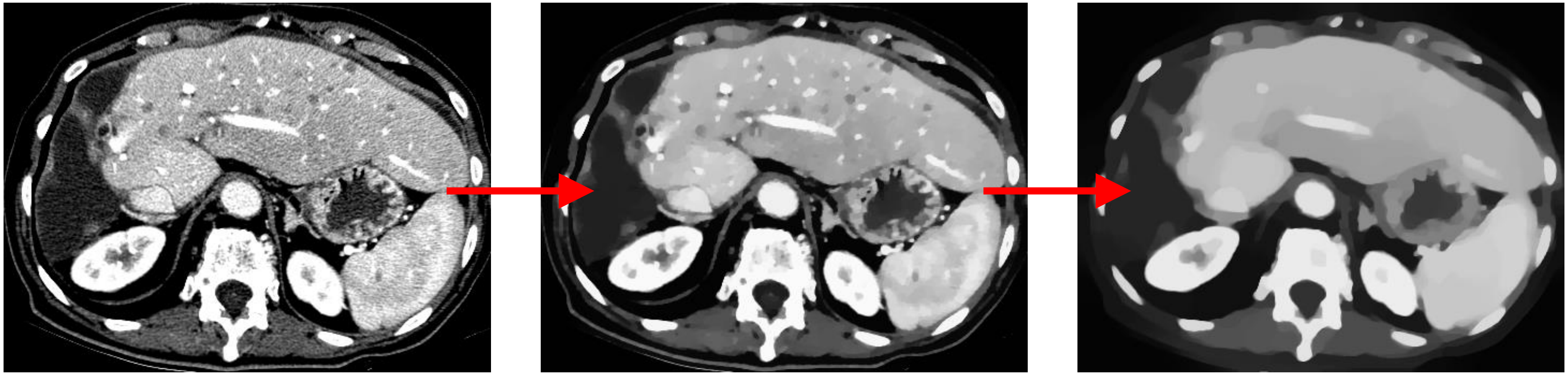
$$-\Delta u + \lambda(u-f) = 0$$

Total Variation Denoising

- Image denoising model introduced by Rudin, Osher and Fatemi in 1992 (a.k.a. ROF, TV-L2 model)

TV

$$\min_u \left\{ \int_\Omega |\nabla u| \, dx + \frac{\lambda}{2} \int_\Omega (u - f)^2 \, dx \right\}$$



$\lambda = 10$

$\lambda = 1$

Quadratic vs. Total Variation

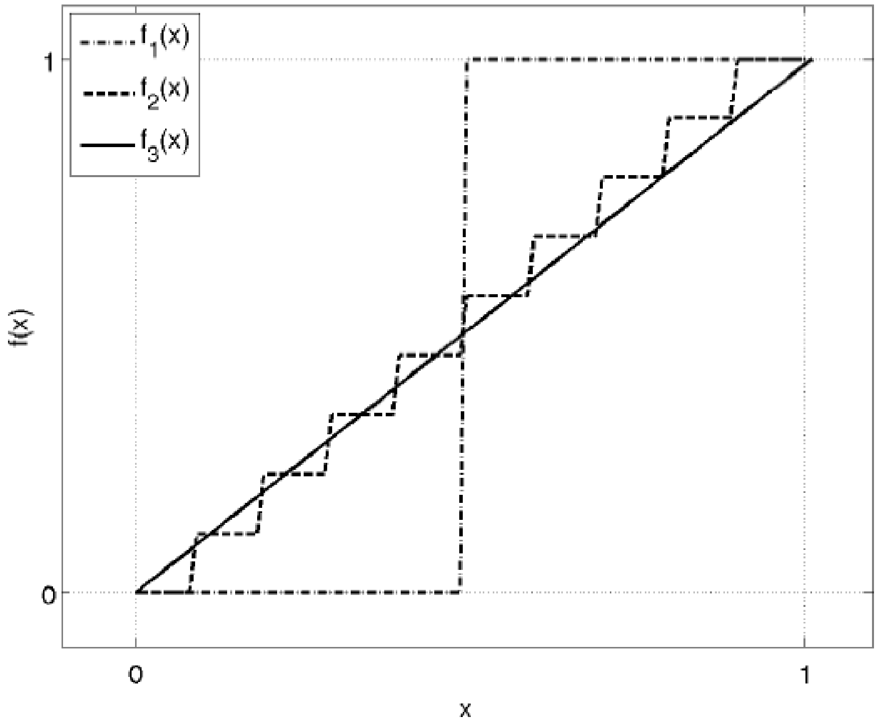


Figure 2.3: Total Variation does not see any difference between these three functions

Let's have a closer look at edges in f!

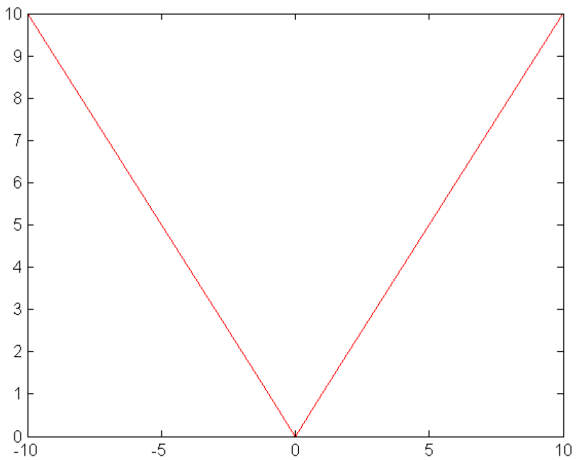
Sample f at 100 locations

Functions	Total Variation	Quadratic
$f_1(x)$	1.0	1.0
$f_2(x)$	1.0	0.11
$f_3(x)$	1.0	0.01

Minimizing quadratic norm favors f3!
Minimizing TV norm makes no distinction

Numerical Implementation - TV

- However, Total Variation (TV) model unfortunateley harder to minimize compared to quadratic!
  - Why? : Derivative undefined at zero!
  - Remember: Euler Lagrange eq. leads to derivative!
- Approaches
  - Slow Gradient Descent Methods
  - Sophisticated Primal-Dual Methods



Numerical Implementation - ROF

Minimize the following energy:

$$\min_u \left\{ \int_\Omega |\nabla u| dx + \frac{\lambda}{2} \int_\Omega (u - f)^2 dx \right\}$$

Euler-Lagrange for $J(u) = \frac{1}{p} \iint_\Omega |\nabla u|^p dxdy, \; 1 \le p < \infty,$

is

$$-\mathrm{div}\left(|\nabla u|^{p-2} \nabla u \right) = 0.$$

So: Associated Euler-Lagrange equation of our energy is:

$$-\nabla \cdot \frac{\nabla u}{|\nabla u|} + \lambda (u - f) = 0$$

Numerical Implementation - ROF

Explicit (Gradient Descent) Optimization:

$$u^{t+1} = u^t - \tau\left[-\nabla\cdot\left(\frac{\nabla u^t}{\sqrt{\left|\nabla u^t\right|^2 + \varepsilon}}\right) + \lambda(u^t - f)\right]$$

Choice of ε difficult & critical!

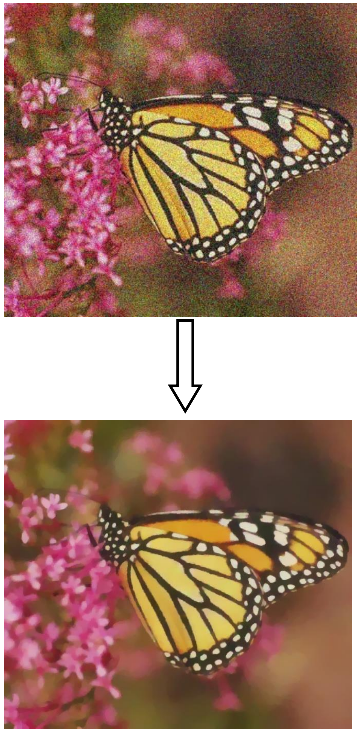
large: slow convergence, smooth over edges

small: divide by nearly zero (numerically unstable)

We call this solution: **ROF-primal**

ROF for Color Images

- Sophisticated models available combining RGB channels (e.g. vector TV)
- Simple:
  - Treat R,G,B planes separately
  - Three, uncoupled ROF steps & combine denoised RGB again



Alternative ROF Implementation

ROF energy:

$$\min_u \left\{ \int_\Omega |\nabla u| dx + \frac{\lambda}{2} \int_\Omega (u - f)^2 dx \right\}$$

Euler-Lagrange:

$$-\nabla \cdot \frac{\nabla u}{|\nabla u|} + \lambda (u - f) = 0$$

Investigate the dual formulation of the TV norm using the dual variable **p**

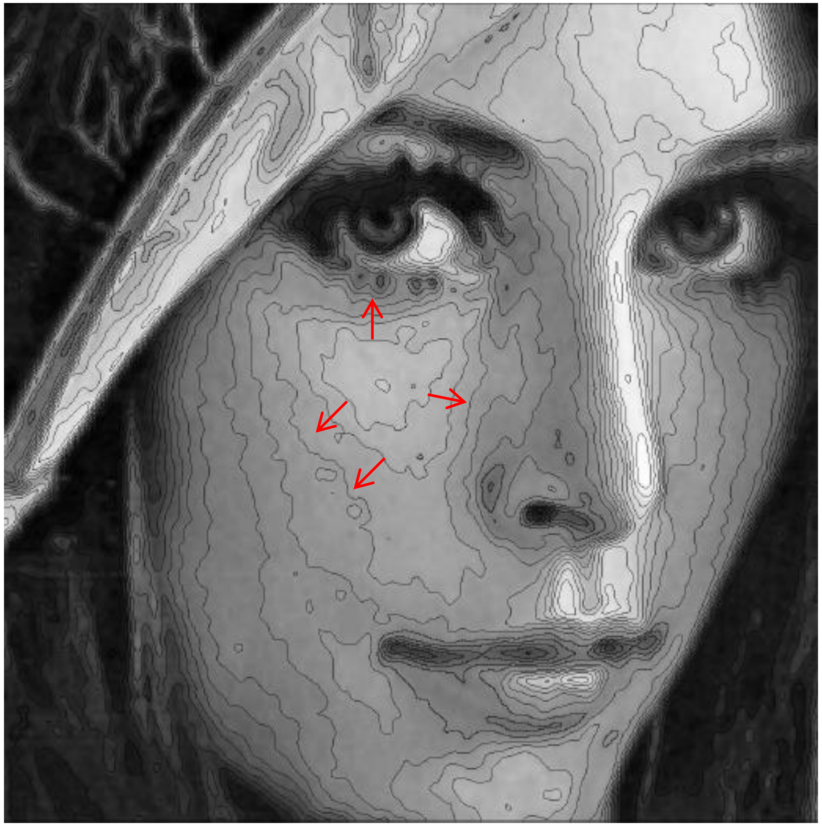
- Define

$$\mathbf{p} = \begin{cases} \dfrac{\nabla u}{|\nabla u|} & \text{if} \quad \nabla u \neq 0 \\ \text{arbitrary} & \text{if} \quad \nabla u = 0 \end{cases}$$

Alternative ROF Implementation

u: lines of same intensity (level sets)

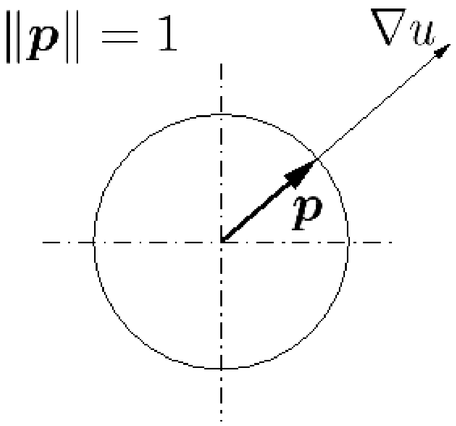
Dual representations



p: normal vector field to isointensity curves

Alternative ROF Implementation

$$|\nabla u| = \max_{\|\mathbf{p}\| \le 1} \{\mathbf{p} \cdot \nabla u\}$$



Which **p** maximizes this expression?

$$\mathbf{p} = \begin{cases} \dfrac{\nabla u}{|\nabla u|} & \text{if} \quad \nabla u \neq 0 \\ \text{arbitrary} & \text{if} \quad \nabla u = 0 \end{cases}$$

Generalization of Total Variation

Why? Scalar Product:

$$\mathbf{p} \cdot \nabla u = \langle \mathbf{p}, \nabla u \rangle = \frac{\nabla u}{|\nabla u|} \cdot \nabla u \overset{!}{=} |\nabla u|$$

The dual variable **p** is the 1-normalized image gradient ∇u

Alternative ROF Implementation

Total Variation may be expressed using dual variable:

$$TV(u) = \int_\Omega |\nabla u| dx = \max_{\|\mathbf{p}\| \le 1} \left\{ \int_\Omega \mathbf{p} \cdot \nabla u dx \right\}$$

TV norm solely defined for smooth functions u! Our image that needs reconstruction contains edges (discontinuities).

$$TV(u) := \max_{\|\mathbf{p}\| \le 1} \left\{ -\int_\Omega u \nabla \cdot \mathbf{p} dx \right\}$$

-> Alternative definition of TV norm, for discontinuous, absolutely integrable functions $u \in L^1(\Omega)$

Functions with finite TV norm are in the space of functions with bounded variations (BV)

$$BV = \left\{ u \in L^1(\Omega) : TV(u) < \infty \right\}$$

Alternative ROF Implementation

$$\max_{\mathbf{p}}\left\{\int_{\Omega}\mathbf{p}\cdot\nabla u dx : \|\mathbf{p}\|\leq 1\right\}=\int_{\Omega}|\nabla u| d\mathbf{x} \quad \overset{!}{\Leftrightarrow} \quad TV(u)=\max_{\mathbf{p}}\left\{-\int_{\Omega}u\nabla\cdot\mathbf{p}dx : \|\mathbf{p}\|\leq 1\right\}$$

Relation to the space of functions with bounded variations

Equivalence via the formula
(Blackboard, Integration
by parts & Divergence Theorem)

$$-\int_{\Omega}u\nabla\cdot\mathbf{p}dx=\int_{\Omega}\mathbf{p}\cdot\nabla u dx$$

Dual variable **p** is a differentiable vector field with compact support, where for each pixel of the domain, **p** is located in the unit disc.

Alternative ROF Implementation

$$\max_{\mathbf{p}}\left\{\int_{\Omega}\mathbf{p}\cdot\nabla u dx : \|\mathbf{p}\|\leq 1\right\} = \int_{\Omega}|\nabla u| d\mathbf{x} \quad \Longleftrightarrow \quad TV(u) = \max_{\mathbf{p}}\left\{-\int_{\Omega}u\nabla\cdot\mathbf{p}dx : \|\mathbf{p}\|\leq 1\right\}$$

Relation to the space of functions with bounded variations

Via the formula

(Blackboard, Integration

by parts & Divergence Theorem)

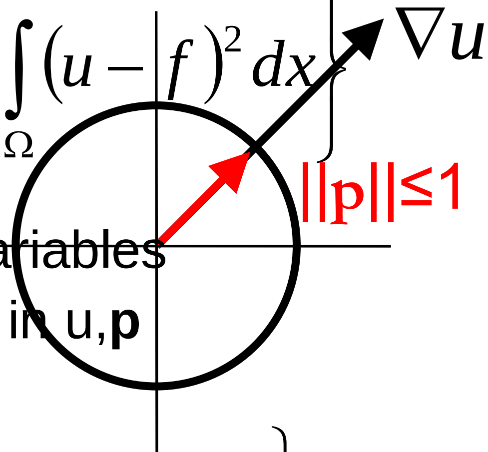
$$-\int_{\Omega}u\nabla\cdot\mathbf{p}dx = \int_{\Omega}\mathbf{p}\cdot\nabla u dx$$

Dual variable **p** is a differentiable vector field with compact support, where for each pixel of the domain, **p** is located in the unit disc.

Alternative ROF implementation

Primal-Dual Formulation of original ROF $\longrightarrow \min_u \left\{ \int_\Omega |\nabla u| dx + \frac{\lambda}{2} \int_\Omega (u - f)^2 dx \right\}$

$$\min_u \max_{\|\mathbf{p}\| \le 1} \left\{ \int_\Omega \mathbf{p} \cdot \nabla u dx + \frac{\lambda}{2} \int_\Omega (u - f)^2 dx \right\}$$



Optimization problem in 2 variables

- Alternating optimization in u,**p**

$$|\nabla u| = \max_{\|\mathbf{p}\| \le 1} (\mathbf{p} \cdot \nabla u)$$

1. $$\frac{\partial}{\partial u} \left\{ -\int_\Omega u \nabla \cdot \mathbf{p} dx + \frac{\lambda}{2} \int_\Omega (u - f)^2 dx \right\} =$$

$$-\nabla \cdot \mathbf{p} + \lambda (u - f)$$

Using $$-\int_\Omega u \nabla \cdot \mathbf{p} dx = \int_\Omega \mathbf{p} \cdot \nabla u dx$$

Alternative ROF implementation

ROF-primal-dual

(Zhu)

Optimization problem in 2 variables

1. Primal update in u

$$u^{n+1} = u^n - \tau_P\left(-\nabla\cdot\mathbf{p} + \lambda(u^n - f)\right)$$

2. Dual update in p

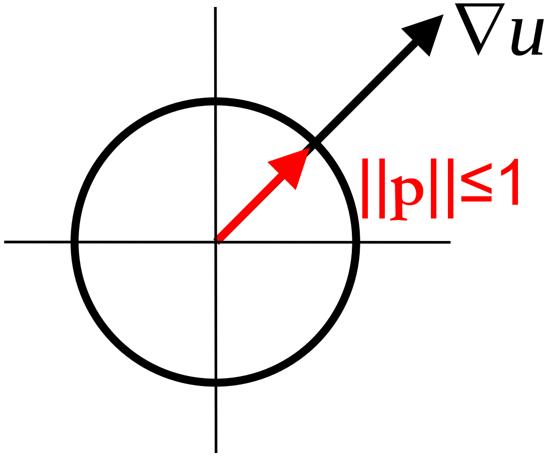
$$\frac{\partial}{\partial\mathbf{p}}\left\{\int_\Omega \mathbf{p}\cdot\nabla u\,dx + \frac{\lambda}{2}\int_\Omega (u-f)^2\,dx\right\} = \nabla u$$

$$\tilde{\mathbf{p}}^{n+1} = \mathbf{p}^n + \tau_D\nabla u$$

maximize p

$$\mathbf{p}^{n+1} = \frac{\tilde{\mathbf{p}}^{n+1}}{\max(1,\left\|\tilde{\mathbf{p}}^{n+1}\right\|)}$$

||p||≤1



$$\left|\nabla u\right| = \max_{\|\mathbf{p}\|\le 1}(\mathbf{p}\cdot\nabla u)$$

[Zhu] Zhu/Chan: An Efficient Primal-Dual Hybrid Gradient Algorithm For Total Variation Image Restoration. *UCLA CAM Report* (2008)

Primal-Dual ROF in 2D

- for n = 1:nrIterations do

  for all pixels do

Dual update $$\tilde{\mathbf{p}}^{n+1} = \mathbf{p}^n + \tau_D \nabla u^n$$

projection of p $$\mathbf{p}^{n+1} = \frac{\tilde{\mathbf{p}}^{n+1}}{\max\left(1, \left|\tilde{\mathbf{p}}^{n+1}\right|\right)}$$

  for all pixels do

Primal update $$u^{n+1} = u^n + \tau_P \left(\nabla \cdot \mathbf{p}^{n+1} - \lambda(u^n - f)\right)$$

Timesteps: $\tau_D = 0.2 + 0.08n$

$$\tau_P = \left(0.5 - \frac{5}{15-n}\right) / \tau_D$$

$$\mathbf{p} = \begin{pmatrix} p^1 \\ p^2 \end{pmatrix} \qquad \mathbf{p}^0 = 0 \qquad u^0 = f$$

Forward differences, use 0 at borders

$$\nabla u = \begin{pmatrix} u_{i+1,j} - u_{i,j} \\ u_{i,j+1} - u_{i,j} \end{pmatrix}$$

Backward differences, use p at borders

$$\nabla \cdot \mathbf{p} = (p^1_{i,j} - p^1_{i-1,j}) + (p^2_{i,j} - p^2_{i,j-1})$$

Properties of Primal-Dual ROF

- Solves a saddle point problem by alternating gradient descent and ascent
- ROF is a convex functional -> denoising solution is a global optimizer
- For more details and a more general saddle point solver in the context of TV, see [1].

[1] Chambolle/Pock: A first order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 2011

Total Variation L1 Denoising

- Gaussian Noise Assumption on Data Term
- Let‘s follow the Robust Statistics (Huber 1981) approach and go one step further
- TV – L2 model:

$$\min_u \left\{ \int_\Omega |\nabla u| dx + \frac{\lambda}{2} \int_\Omega (u - f)^2 dx \right\}$$

- Remove second quadratic term as well:

$$\min_u \left\{ \int_\Omega |\nabla u| dx + \lambda \int_\Omega |u - f| dx \right\}$$

Total Variation L1 Denoising

Minimize the following energy: $$\min_u \left\{ \int_\Omega |\nabla u| dx + \lambda \int_\Omega |u - f| dx \right\}$$

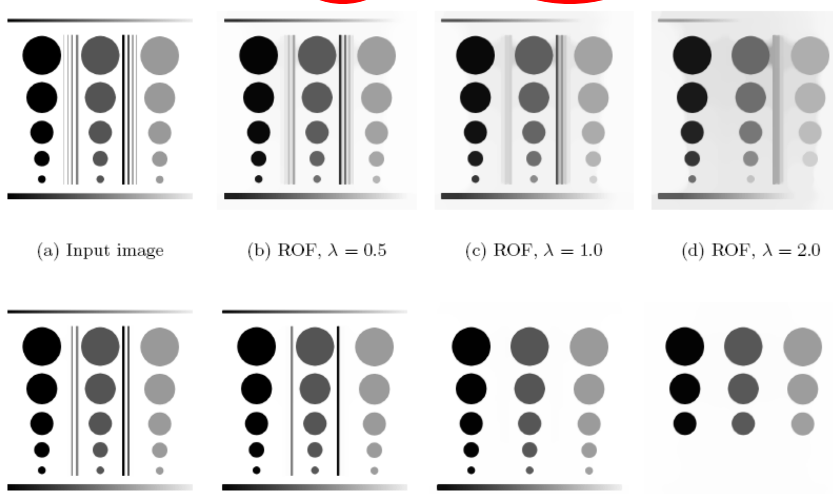
Associated Euler-Lagrange equation: $$-\nabla \cdot \frac{\nabla u}{|\nabla u|} + \lambda \frac{u - f}{|u - f|} = 0$$

Properties:

- Robust to Impulse (e.g. Salt & Pepper) Noise
- Contrast Invariance!

Possible solution:

TVL¹-primal



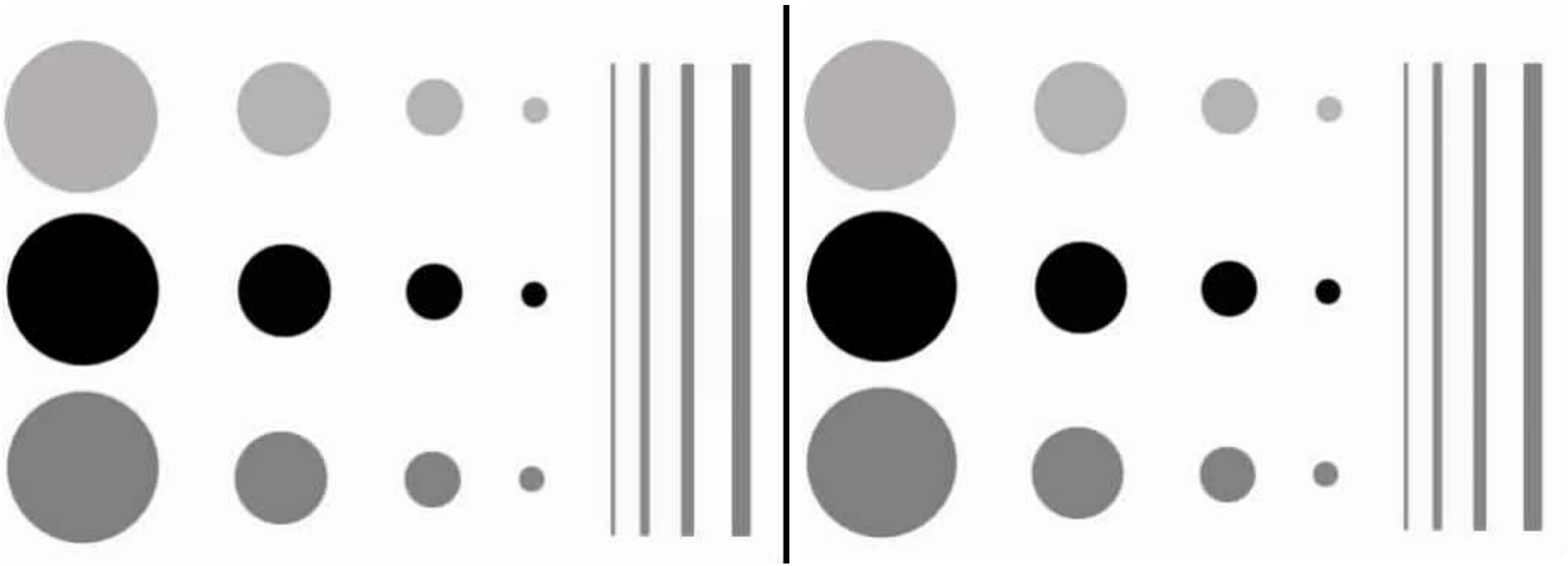
(a) Input image (b) ROF, $\lambda = 0.5$ (c) ROF, $\lambda = 1.0$ (d) ROF, $\lambda = 2.0$

(e) TV-L^1, $\lambda = 1.0$ (f) TV-L^1, $\lambda = 0.5$ (g) TV-L^1, $\lambda = 0.2$ (h) TV-L^1, $\lambda = 0.1$

Figure 2.7: Ability of the TV-L^1 model to remove structures of a certain scale.

Total Variation L1 Denoising

Influence of parameter λ



TV L1

TV L2

Contrast Invariance!

TV-L1 Implementation

$$\min_{u}\left\{\int_{\Omega}|\nabla u|dx+\lambda\int_{\Omega}|u-f|dx\right\}$$

- Approximation using auxiliary variable v (with $\theta > 0$)

$$\min_{u,v}\left\{\int_{\Omega}|\nabla u|dx+\frac{1}{2\theta}\int_{\Omega}(u-v)^2 dx+\lambda\int_{\Omega}|v-f|dx\right\}$$

- For $\theta \to 0$: Approximation approaches the original energy!

- Alternating minimization with respect to u and v

TV-L1 Implementation

$$\min_{u,v}\left\{\int_{\Omega}|\nabla u|dx+\frac{1}{2\theta}\int_{\Omega}(u-v)^2 dx+\lambda\int_{\Omega}|v-f|dx\right\}$$

- Corresponding Euler-Lagrange equations:

$$-\nabla\cdot\frac{\nabla u}{|\nabla u|}+\frac{1}{\theta}(u-v)=0$$

Solve for u $\longrightarrow$ This is the ROF model!

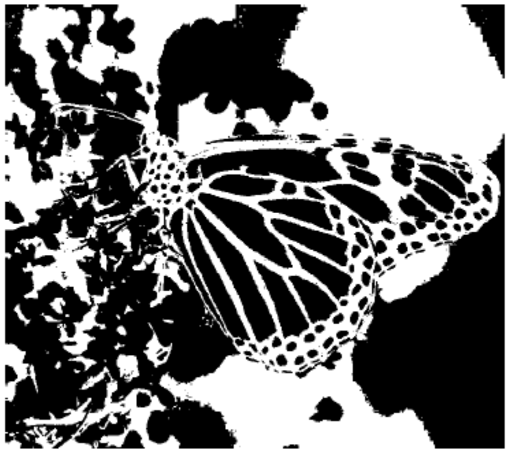
$$\frac{1}{\theta}(u-v)+\lambda\frac{v-f}{|v-f|}=0$$

Solve for v $\longrightarrow$ Simple Case Distinction Scheme!

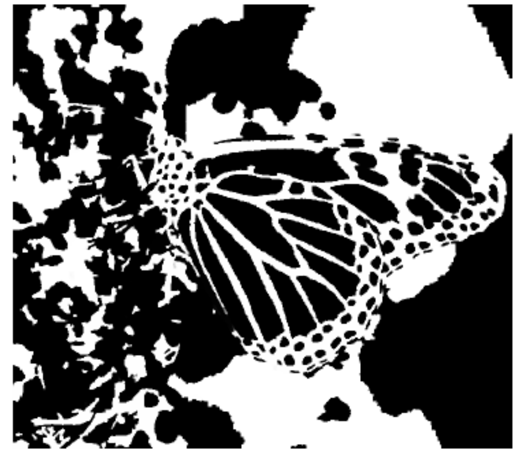
Alternating minimization!
Repeat:

1. Keep v fixed, solve for u
2. Keed u fixed, solve for v

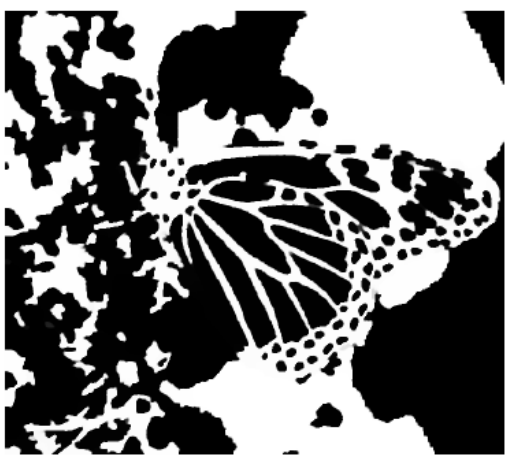
TV-L1 Interpretation – Shape Denoising



(a) Input image



(b) $\lambda = 1.0$



(c) $\lambda = 0.5$

(d) $\lambda = 0.25$

Summary

- What did we do so far?
- Edge-preserving Denoising

- Our energies are „Convex Optimization“ problems -> possess global optimum!
- These Numerical Methods are well-suited for GPU Implementation! (Nvidia CUDA, OpenCL)

END

One minute paper:

a) What did I learn today?

b) Which topics remained open/unclear?

See you next week!