Total Variation in Image Analysis (The Homo Erectus Stage?)

François Lauze

[1] Department of Computer Science
University of Copenhagen

Hólar Summer School on Sparse Coding, August 2010

Outline

Outline

- In mathematics: the Plateau problem of minimal surfaces, i.e. surfaces of minimal area with a given boundary

- In image analysis: denoising, image reconstruction, segmentation...
- An ubiquitous prior for many image processing tasks.

- In mathematics: the Plateau problem of minimal surfaces, i.e. surfaces of minimal area with a given boundary

- In image analysis: denoising, image reconstruction, segmentation...
- An ubiquitous prior for many image processing tasks.

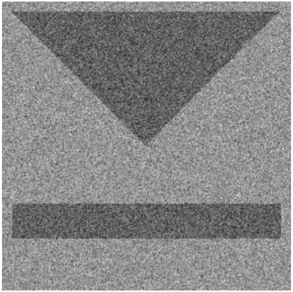
- In mathematics: the Plateau problem of minimal surfaces, i.e. surfaces of minimal area with a given boundary

- In image analysis: denoising, image reconstruction, segmentation...
- An ubiquitous prior for many image processing tasks.

Outline

Denoising

Determine an unknown image from a noisy observation.



Methods

All methods based on some statistical inference.

- Fourier/Wavelets
- Markov Random Fields
- Variational and Partial Differential Equations methods
- ...

We focus on variational and PDE methods.

A simple corruption model

- A digital image u of size $N \times M$ pixels, corrupted by Gaussian white noise of variance σ^2
- write it as observed image $u_0 = u + \eta$, $\|u - u_0\|^2 = \sum_{ij}(u_{ij} - u_{0ij})^2 = NM\sigma^2$ (noise variance = σ^2), $\sum_{ij} u_{ij} = \sum_{ij} u_{0ij}$ (zero mean noise).
- could add a blur degradation $u_0 = Ku + \eta$ for instance, so to have $\|Ku - u_0\|^2 = NM\sigma^2$.

A simple corruption model

- A digital image u of size $N \times M$ pixels, corrupted by Gaussian white noise of variance σ^2
- write it as observed image $u_0 = u + \eta$, $\|u - u_0\|^2 = \sum_{ij}(u_{ij} - u_{0ij})^2 = NM\sigma^2$ (noise variance = σ^2), $\sum_{ij} u_{ij} = \sum_{ij} u_{0ij}$ (zero mean noise).
- could add a blur degradation $u_0 = Ku + \eta$ for instance, so to have $\|Ku - u_0\|^2 = NM\sigma^2$.

A simple corruption model

- A digital image u of size $N \times M$ pixels, corrupted by Gaussian white noise of variance σ^2
- write it as observed image $u_0 = u + \eta$, $\|u - u_0\|^2 = \sum_{ij}(u_{ij} - u_{0ij})^2 = NM\sigma^2$ (noise variance = σ^2), $\sum_{ij} u_{ij} = \sum_{ij} u_{0ij}$ (zero mean noise).
- could add a blur degradation $u_0 = Ku + \eta$ for instance, so to have $\|Ku - u_0\|^2 = NM\sigma^2$.

A simple corruption model

- A digital image u of size $N \times M$ pixels, corrupted by Gaussian white noise of variance σ^2
- write it as observed image $u_0 = u + \eta$, $\|u - u_0\|^2 = \sum_{ij}(u_{ij} - u_{0ij})^2 = NM\sigma^2$ (noise variance = σ^2), $\sum_{ij} u_{ij} = \sum_{ij} u_{0ij}$ (zero mean noise).
- could add a blur degradation $u_0 = Ku + \eta$ for instance, so to have $\|Ku - u_0\|^2 = NM\sigma^2$.

A simple corruption model

- A digital image u of size $N \times M$ pixels, corrupted by Gaussian white noise of variance σ^2
- write it as observed image $u_0 = u + \eta$, $\|u - u_0\|^2 = \sum_{ij}(u_{ij} - u_{0ij})^2 = NM\sigma^2$ (noise variance = σ^2), $\sum_{ij} u_{ij} = \sum_{ij} u_{0ij}$ (zero mean noise).
- could add a blur degradation $u_0 = Ku + \eta$ for instance, so to have $\|Ku - u_0\|^2 = NM\sigma^2$.

Recovery

- The problem: Find u such that

$$\|u - u_0\|^2 = NM\sigma^2, \quad \sum_{ij} u_{ij} = \sum_{ij} u_{0ij} \tag{1}$$

  is not well-posed. Many solutions possible.

- In order to recover u, extra information is needed, e.g. in the form of a prior on u.
- For images, smoothness priors often used.
- Let Ru a digital gradient of u, Then find smoothest u that satisfy constraints (1), the smoothest meaning with smallest

$$T(u) = \|Ru\| = \sqrt{\sum_{ij} |Ru|_{ij}^2}.$$

Recovery

- The problem: Find u such that
$$\|u - u_0\|^2 = NM\sigma^2, \quad \sum_{ij} u_{ij} = \sum_{ij} u_{0ij} \tag{1}$$
is not well-posed. Many solutions possible.
- In order to recover u, extra information is needed, e.g. in the form of a prior on u.
- For images, smoothness priors often used.
- Let Ru a digital gradient of u, Then find smoothest u that satisfy constraints (1), the smoothest meaning with smallest
$$T(u) = \|Ru\| = \sqrt{\sum_{ij} |Ru|_{ij}^2}.$$

Recovery

- The problem: Find u such that

$$\|u - u_0\|^2 = NM\sigma^2, \quad \sum_{ij} u_{ij} = \sum_{ij} u_{0ij} \tag{1}$$

  is not well-posed. Many solutions possible.

- In order to recover u, extra information is needed, e.g. in the form of a prior on u.

- For images, smoothness priors often used.

- Let Ru a digital gradient of u, Then find smoothest u that satisfy constraints (1), the smoothest meaning with smallest

$$T(u) = \|Ru\| = \sqrt{\sum_{ij} |Ru|^2_{ij}}.$$

Recovery

- The problem: Find u such that

$$\|u - u_0\|^2 = NM\sigma^2, \quad \sum_{ij} u_{ij} = \sum_{ij} u_{0ij} \tag{1}$$

  is not well-posed. Many solutions possible.

- In order to recover u, extra information is needed, e.g. in the form of a prior on u.
- For images, smoothness priors often used.
- Let Ru a digital gradient of u, Then find smoothest u that satisfy constraints (1), the smoothest meaning with smallest

$$T(u) = \|Ru\| = \sqrt{\sum_{ij} |Ru|_{ij}^2}.$$

Recovery

- The problem: Find u such that

$$\|u - u_0\|^2 = NM\sigma^2, \quad \sum_{ij} u_{ij} = \sum_{ij} u_{0ij} \tag{1}$$

  is not well-posed. Many solutions possible.

- In order to recover u, extra information is needed, e.g. in the form of a prior on u.
- For images, smoothness priors often used.
- Let Ru a digital gradient of u, Then find smoothest u that satisfy constraints (1), the smoothest meaning with smallest

$$T(u) = \|Ru\| = \sqrt{\sum_{ij} |Ru|_{ij}^2}.$$

Outline

Tikhonov regularization

- It can be show that this is equivalent to minimize

$$E(u) = ||Ku - u_0||^2 + \lambda ||Ru||^2$$

for a $\lambda = \lambda(\sigma)$ (Wahba?).

- $E(u)$ minimizaton can be derived from a Maximum a Posteriori formulation

$$\underset{u}{\text{Arg.max}} p(u|u_0) = \frac{p(u_0|u)p(u)}{p(u_0)}$$

- Rewriting in a continuous setting:

$$E(u) = \int_\Omega (Ku - u_o)^2 \, dx + \lambda \int_\Omega |\nabla u|^2 \, dx$$

Tikhonov regularization

- It can be show that this is equivalent to minimize

$$E(u) = \|Ku - u_0\|^2 + \lambda \|Ru\|^2$$

for a $\lambda = \lambda(\sigma)$ (Wahba?).

- $E(u)$ minimizaton can be derived from a Maximum a Posteriori formulation

$$\underset{u}{\text{Arg.max}} p(u|u_0) = \frac{p(u_0|u)p(u)}{p(u_0)}$$

- Rewriting in a continuous setting:

$$E(u) = \int_\Omega (Ku - u_o)^2 \, dx + \lambda \int_\Omega |\nabla u|^2 \, dx$$

Tikhonov regularization

- It can be show that this is equivalent to minimize

$$E(u) = \|Ku - u_0\|^2 + \lambda \|Ru\|^2$$

for a $\lambda = \lambda(\sigma)$ (Wahba?).

- $E(u)$ minimizaton can be derived from a Maximum a Posteriori formulation

$$\text{Arg.max}_u \, p(u|u_0) = \frac{p(u_0|u)p(u)}{p(u_0)}$$

- Rewriting in a continuous setting:

$$E(u) = \int_\Omega (Ku - u_o)^2 \, dx + \lambda \int_\Omega |\nabla u|^2 \, dx$$

Tikhonov regularization

- It can be show that this is equivalent to minimize

$$E(u) = \|Ku - u_0\|^2 + \lambda \|Ru\|^2$$

for a $\lambda = \lambda(\sigma)$ (Wahba?).

- $E(u)$ minimizaton can be derived from a Maximum a Posteriori formulation

$$\underset{u}{\text{Arg.max}} p(u|u_0) = \frac{p(u_0|u)p(u)}{p(u_0)}$$

- Rewriting in a continuous setting:

$$E(u) = \int_\Omega (Ku - u_o)^2 \, dx + \lambda \int_\Omega |\nabla u|^2 \, dx$$

Tikhonov regularization

- It can be show that this is equivalent to minimize

$$E(u) = \|Ku - u_0\|^2 + \lambda\|Ru\|^2$$

for a $\lambda = \lambda(\sigma)$ (Wahba?).

- $E(u)$ minimizaton can be derived from a Maximum a Posteriori formulation

$$\underset{u}{\text{Arg.max}} p(u|u_0) = \frac{p(u_0|u)p(u)}{p(u_0)}$$

- Rewriting in a continuous setting:

$$E(u) = \int_\Omega (Ku - u_o)^2 \, dx + \lambda \int_\Omega |\nabla u|^2 \, dx$$

How to solve?

- Solution satisfies the Euler-Lagrange equation for E:

$$K^*\left(Ku - u_0\right) - \lambda \Delta u = 0.$$

  (K^* is the adjoint of K)
- A linear equation, easy to implement, and many fast solvers exit, but...

How to solve?

- Solution satisfies the Euler-Lagrange equation for E:

$$K^* (Ku - u_0) - \lambda \Delta u = 0.$$

(K^* is the adjoint of K)

- A linear equation, easy to implement, and many fast solvers exit, but...

How to solve?

- Solution satisfies the Euler-Lagrange equation for E:

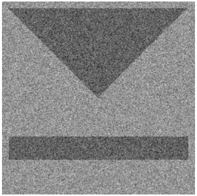
$$K^* (Ku - u_0) - \lambda \Delta u = 0.$$

  (K^* is the adjoint of K)
- A linear equation, easy to implement, and many fast solvers exit, but...

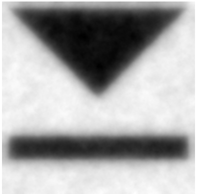
Tikhonov example

Denoising example, $K = Id$.

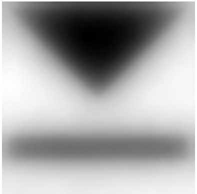
Original



$\lambda = 50$



$\lambda = 500$

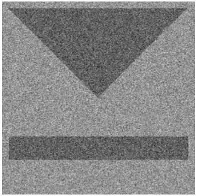


- Not good: images contain edges but Tikhonov blur them. Why?
- The term $\int_\Omega (u - u_0)^2 \, dx$: not guilty!
- Then it must be $\int_\Omega |\nabla u|^2 \, dx$. Derivatives and step edges do not go too well together?

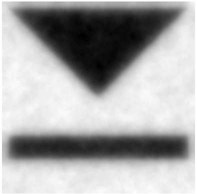
Tikhonov example

Denoising example, $K = Id$.

Original



$\lambda = 50$



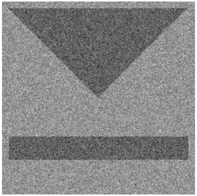
$\lambda = 500$

- Not good: images contain edges but Tikhonov blur them. Why?
- The term $\int_\Omega (u - u_0)^2\, dx$: not guilty!
- Then it must be $\int_\Omega |\nabla u|^2\, dx$. Derivatives and step edges do not go too well together?

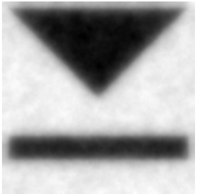
Tikhonov example

Denoising example, $K = Id$.

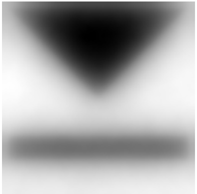
Original



$\lambda = 50$



$\lambda = 500$

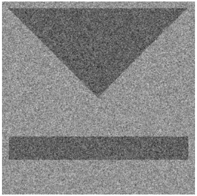


- Not good: images contain edges but Tikhonov blur them. Why?
- The term $\int_\Omega (u - u_0)^2 \, dx$: not guilty!
- Then it must be $\int_\Omega |\nabla u|^2 \, dx$. Derivatives and step edges do not go too well together?

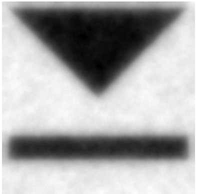
Tikhonov example

Denoising example, $K = Id$.

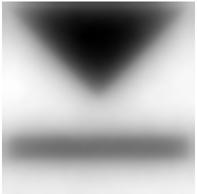
Original



$\lambda = 50$



$\lambda = 500$



- Not good: images contain edges but Tikhonov blur them. Why?
- The term $\int_\Omega (u - u_0)^2 \, dx$: not guilty!
- Then it must be $\int_\Omega |\nabla u|^2 \, dx$. Derivatives and step edges do not go too well together?

Outline

- Set $\Omega = [-1, 1]$, a a real number and u the step-edge function

$$u(x) = \begin{cases} 0 & x \leq 0 \\ a & x > 0 \end{cases}$$

- Not differentiable at 0, but forget about it and try to compute

$$\int_{-1}^{1} |u'(x)|^2 \, dx.$$

- Around 0 "approximate" $u'(x)$ by

$$\frac{u(h) - u(-h)}{2h}, \quad h > 0, \text{small}$$

- Set $\Omega = [-1, 1]$, a a real number and u the step-edge function

$$u(x) = \begin{cases} 0 & x \leq 0 \\ a & x > 0 \end{cases}$$

- Not differentiable at 0, but forget about it and try to compute

$$\int_{-1}^{1} |u'(x)|^2 \, dx.$$

- Around 0 "approximate" $u'(x)$ by

$$\frac{u(h) - u(-h)}{2h}, \quad h > 0, \text{small}$$

- Set $\Omega = [-1, 1]$, a a real number and u the step-edge function

$$u(x) = \begin{cases} 0 & x \leq 0 \\ a & x > 0 \end{cases}$$

- Not differentiable at 0, but forget about it and try to compute

$$\int_{-1}^{1} |u'(x)|^2 \, dx.$$

- Around 0 “approximate” $u'(x)$ by

$$\frac{u(h) - u(-h)}{2h}, \quad h > 0, \text{small}$$

- with this finite difference approximation

$$u'(x) \approx \frac{a}{2h}, \quad x \in [-h, h]$$

- then

$$\begin{aligned}\int_{-1}^{1} |u'(x)|^2\,dx &= \int_{-1}^{-h} |u'(x)|^2\,dx + \int_{-h}^{h} |u'(x)|^2\,dx + \int_{h}^{1} |u'(x)|^2\,dx \\ &= 0 + 2h \times \left(\frac{a}{2h}\right)^2 + 0 \\ &= \frac{a^2}{2h} \to \infty, \quad h \to 0\end{aligned}$$

- So a step-edge has “infinite energy”. It cannot minimizes Tikhonov.
- What went “wrong”: the square:

- with this finite difference approximation

$$u'(x) \approx \frac{a}{2h}, \quad x \in [-h, h]$$

- then

$$\begin{aligned} \int_{-1}^{1} |u'(x)|^2 \, dx &= \int_{-1}^{-h} |u'(x)|^2 \, dx + \int_{-h}^{h} |u'(x)|^2 \, dx + \int_{h}^{1} |u'(x)|^2 \, dx \\ &= 0 + 2h \times \left(\frac{a}{2h}\right)^2 + 0 \\ &= \frac{a^2}{2h} \to \infty, \quad h \to 0 \end{aligned}$$

- So a step-edge has “infinite energy”. It cannot minimizes Tikhonov.
- What went “wrong”: the square:

- with this finite difference approximation

$$u'(x) \approx \frac{a}{2h}, \quad x \in [-h, h]$$

- then

$$\begin{aligned}\int_{-1}^{1} |u'(x)|^2\, dx &= \int_{-1}^{-h} |u'(x)|^2\, dx + \int_{-h}^{h} |u'(x)|^2\, dx + \int_{h}^{1} |u'(x)|^2\, dx \\ &= 0 + 2h \times \left(\frac{a}{2h}\right)^2 + 0 \\ &= \frac{a^2}{2h} \to \infty, \quad h \to 0\end{aligned}$$

- So a step-edge has “infinite energy”. It cannot minimizes Tikhonov.
- What went “wrong”: the square:

- with this finite difference approximation

$$u'(x) \approx \frac{a}{2h}, \quad x \in [-h, h]$$

- then

$$\begin{aligned} \int_{-1}^{1} |u'(x)|^2 \, dx &= \int_{-1}^{-h} |u'(x)|^2 \, dx + \int_{-h}^{h} |u'(x)|^2 \, dx + \int_{h}^{1} |u'(x)|^2 \, dx \\ &= 0 + 2h \times \left(\frac{a}{2h}\right)^2 + 0 \\ &= \frac{a^2}{2h} \to \infty, \quad h \to 0 \end{aligned}$$

- So a step-edge has “infinite energy”. It cannot minimizes Tikhonov.
- What went “wrong”: the square:

- with this finite difference approximation

$$u'(x) \approx \frac{a}{2h}, \quad x \in [-h, h]$$

- then

$$\begin{aligned}\int_{-1}^{1} |u'(x)|^2\,dx &= \int_{-1}^{-h} |u'(x)|^2\,dx + \int_{-h}^{h} |u'(x)|^2\,dx + \int_{h}^{1} |u'(x)|^2\,dx \\ &= 0 + 2h \times \left(\frac{a}{2h}\right)^2 + 0 \\ &= \frac{a^2}{2h} \to \infty, \quad h \to 0\end{aligned}$$

- So a step-edge has “infinite energy”. It cannot minimizes Tikhonov.
- What went “wrong”: the square:

- Replace the square in the previous computation by $p > 0$ and redo:
- Then

$$\begin{aligned}\int_{-1}^{1} |u'(x)|^p\,dx &= \int_{-1}^{-h} |u'(x)|^p\,dx + \int_{-h}^{h} |u'(x)|^p\,dx + \int_{h}^{1} |u'(x)|^p\,dx \\ &= 0 + 2h \times \left|\frac{a}{2h}\right|^p + 0 \\ &= |a|^p (2h)^{1-p} < \infty \quad \text{when } p \le 1\end{aligned}$$

- When $p \le 1$ this is finite! Edges can survive here!
- Quite ugly when $p < 1$ (but not uninteresting)
- When $p = 1$, this is the Total Variation of u.

- Replace the square in the previous computation by $p > 0$ and redo:
- Then

$$\begin{aligned}\int_{-1}^{1} |u'(x)|^p\,dx &= \int_{-1}^{-h} |u'(x)|^p\,dx + \int_{-h}^{h} |u'(x)|^p\,dx + \int_{h}^{1} |u'(x)|^p\,dx \\ &= 0 + 2h \times \left|\frac{a}{2h}\right|^p + 0 \\ &= |a|^p (2h)^{1-p} < \infty \quad \text{when } p \leq 1\end{aligned}$$

- When $p \leq 1$ this is finite! Edges can survive here!
- Quite ugly when $p < 1$ (but not uninteresting)
- When $p = 1$, this is the Total Variation of u.

- Replace the square in the previous computation by $p > 0$ and redo:
- Then

$$\begin{aligned}\int_{-1}^{1} |u'(x)|^p\,dx &= \int_{-1}^{-h} |u'(x)|^p\,dx + \int_{-h}^{h} |u'(x)|^p\,dx + \int_{h}^{1} |u'(x)|^p\,dx \\ &= 0 + 2h \times \left|\frac{a}{2h}\right|^p + 0 \\ &= |a|^p (2h)^{1-p} < \infty \quad \text{when } p \le 1\end{aligned}$$

- When $p \le 1$ this is finite! Edges can survive here!
- Quite ugly when $p < 1$ (but not uninteresting)
- When $p = 1$, this is the Total Variation of u.

- Replace the square in the previous computation by $p > 0$ and redo:
- Then

$$\begin{aligned}\int_{-1}^{1} |u'(x)|^p \, dx &= \int_{-1}^{-h} |u'(x)|^p \, dx + \int_{-h}^{h} |u'(x)|^p \, dx + \int_{h}^{1} |u'(x)|^p \, dx \\ &= 0 + 2h \times \left|\frac{a}{2h}\right|^p + 0 \\ &= |a|^p (2h)^{1-p} < \infty \quad \text{when } p \leq 1\end{aligned}$$

- When $p \leq 1$ this is finite! Edges can survive here!
- Quite ugly when $p < 1$ (but not uninteresting)
- When $p = 1$, this is the Total Variation of u.

- Replace the square in the previous computation by $p > 0$ and redo:
- Then

$$\begin{aligned}\int_{-1}^{1} |u'(x)|^p \, dx &= \int_{-1}^{-h} |u'(x)|^p \, dx + \int_{-h}^{h} |u'(x)|^p \, dx + \int_{h}^{1} |u'(x)|^p \, dx \\ &= 0 + 2h \times \left|\frac{a}{2h}\right|^p + 0 \\ &= |a|^p (2h)^{1-p} < \infty \quad \text{when } p \leq 1\end{aligned}$$

- When $p \leq 1$ this is finite! Edges can survive here!
- Quite ugly when $p < 1$ (but not uninteresting)
- When $p = 1$, this is the Total Variation of u.

- Replace the square in the previous computation by $p > 0$ and redo:
- Then

$$\begin{aligned}\int_{-1}^{1} |u'(x)|^p\,dx &= \int_{-1}^{-h} |u'(x)|^p\,dx + \int_{-h}^{h} |u'(x)|^p\,dx + \int_{h}^{1} |u'(x)|^p\,dx \\ &= 0 + 2h \times \left|\frac{a}{2h}\right|^p + 0 \\ &= |a|^p (2h)^{1-p} < \infty \quad \text{when } p \leq 1\end{aligned}$$

- When $p \leq 1$ this is finite! Edges can survive here!
- Quite ugly when $p < 1$ (but not uninteresting)
- When $p = 1$, this is the Total Variation of u.

Outline

- Let $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}$. Define total variation as

$$J(u) = \int_\Omega |\nabla u| \, dx, \quad |\nabla u| = \sqrt{\sum_{i=1}^n u_{x_i}^2}.$$

- When $J(u)$ is finite, one says that u has bounded variations and the space of function of bounded variations on Ω is denoted $BV(\Omega)$.

- Let $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}$. Define total variation as

$$J(u) = \int_\Omega |\nabla u| \, dx, \quad |\nabla u| = \sqrt{\sum_{i=1}^{n} u_{x_i}^2}.$$

- When $J(u)$ is finite, one says that u has bounded variations and the space of function of bounded variations on Ω is denoted $BV(\Omega)$.

- Expected: when minimizing $J(u)$ with other constraints, edges are less penalized that with Tikhonov.
- Indeed edges are “naturally present” in bounded variation functions. In fact: functions of bounded variations can be decomposed in
  - smooth parts, ∇u well defined,
  - Jump discontinuities (our edges)
  - something else (Cantor part) which can be nasty...
- The functions that do not possess this nasty part form a subspace of $BV(\Omega)$ called $SBV(\Omega)$, The Special functions of Bounded Variation, (used for instance when studying Mumford-Shah functional)

- Expected: when minimizing $J(u)$ with other constraints, edges are less penalized that with Tikhonov.
- Indeed edges are "naturally present" in bounded variation functions. In fact: functions of bounded variations can be decomposed in
  - smooth parts, ∇u well defined,
  - Jump discontinuities (our edges)
  - something else (Cantor part) which can be nasty...
- The functions that do not possess this nasty part form a subspace of $BV(\Omega)$ called $SBV(\Omega)$, The Special functions of Bounded Variation, (used for instance when studying Mumford-Shah functional)

- Expected: when minimizing $J(u)$ with other constraints, edges are less penalized that with Tikhonov.
- Indeed edges are “naturally present” in bounded variation functions. In fact: functions of bounded variations can be decomposed in

  1. smooth parts, ∇u well defined,
  2. Jump discontinuities (our edges)
  3. something else (Cantor part) which can be nasty...

- The functions that do not possess this nasty part form a subspace of $BV(\Omega)$ called $SBV(\Omega)$, The Special functions of Bounded Variation, (used for instance when studying Mumford-Shah functional)

- Expected: when minimizing $J(u)$ with other constraints, edges are less penalized that with Tikhonov.
- Indeed edges are “naturally present” in bounded variation functions. In fact: functions of bounded variations can be decomposed in

  1. smooth parts, ∇u well defined,
  2. Jump discontinuities (our edges)
  3. something else (Cantor part) which can be nasty...

- The functions that do not possess this nasty part form a subspace of $BV(\Omega)$ called $SBV(\Omega)$, The Special functions of Bounded Variation, (used for instance when studying Mumford-Shah functional)

- Expected: when minimizing $J(u)$ with other constraints, edges are less penalized that with Tikhonov.
- Indeed edges are "naturally present" in bounded variation functions. In fact: functions of bounded variations can be decomposed in

  1. smooth parts, ∇u well defined,
  2. Jump discontinuities (our edges)
  3. something else (Cantor part) which can be nasty...

- The functions that do not possess this nasty part form a subspace of $BV(\Omega)$ called $SBV(\Omega)$, The Special functions of Bounded Variation, (used for instance when studying Mumford-Shah functional)

- Expected: when minimizing $J(u)$ with other constraints, edges are less penalized that with Tikhonov.
- Indeed edges are “naturally present” in bounded variation functions. In fact: functions of bounded variations can be decomposed in

    1. smooth parts, ∇u well defined,
    2. Jump discontinuities (our edges)
    3. something else (Cantor part) which can be nasty...

- The functions that do not possess this nasty part form a subspace of $BV(\Omega)$ called $SBV(\Omega)$, The Special functions of Bounded Variation, (used for instance when studying Mumford-Shah functional)

- Expected: when minimizing $J(u)$ with other constraints, edges are less penalized that with Tikhonov.
- Indeed edges are "naturally present" in bounded variation functions. In fact: functions of bounded variations can be decomposed in

  1. smooth parts, ∇u well defined,
  2. Jump discontinuities (our edges)
  3. something else (Cantor part) which can be nasty...

- The functions that do not possess this nasty part form a subspace of $BV(\Omega)$ called $SBV(\Omega)$, The Special functions of Bounded Variation, (used for instance when studying Mumford-Shah functional)

- Expected: when minimizing $J(u)$ with other constraints, edges are less penalized that with Tikhonov.
- Indeed edges are “naturally present” in bounded variation functions. In fact: functions of bounded variations can be decomposed in

  1. smooth parts, ∇u well defined,
  2. Jump discontinuities (our edges)
  3. something else (Cantor part) which can be nasty...

- The functions that do not possess this nasty part form a subspace of $BV(\Omega)$ called $SBV(\Omega)$, The Special functions of Bounded Variation, (used for instance when studying Mumford-Shah functional)

- Expected: when minimizing $J(u)$ with other constraints, edges are less penalized that with Tikhonov.
- Indeed edges are “naturally present” in bounded variation functions. In fact: functions of bounded variations can be decomposed in

  1. smooth parts, ∇u well defined,
  2. Jump discontinuities (our edges)
  3. something else (Cantor part) which can be nasty...

- The functions that do not possess this nasty part form a subspace of $BV(\Omega)$ called $SBV(\Omega)$, The Special functions of Bounded Variation, (used for instance when studying Mumford-Shah functional)

- Expected: when minimizing $J(u)$ with other constraints, edges are less penalized that with Tikhonov.
- Indeed edges are “naturally present” in bounded variation functions. In fact: functions of bounded variations can be decomposed in

  1. smooth parts, ∇u well defined,
  2. Jump discontinuities (our edges)
  3. something else (Cantor part) which can be nasty...

- The functions that do not possess this nasty part form a subspace of $BV(\Omega)$ called $SBV(\Omega)$, The Special functions of Bounded Variation, (used for instance when studying Mumford-Shah functional)

Outline

ROF Denoising

- State the denoising problem as minimizing $J(u)$ under the constraints

$$\int_\Omega u\,dx = \int_\Omega u_o\,dx, \quad \int_\Omega (u-u_0)^2\,dx = |\Omega|\sigma^2 \quad (|\Omega| = \text{area/volume of } \Omega)$$

- Solve via Lagrange multipliers.

ROF Denoising

- State the denoising problem as minimizing $J(u)$ under the constraints

$$\int_\Omega u\,dx = \int_\Omega u_o\,dx, \quad \int_\Omega (u-u_0)^2\,dx = |\Omega|\sigma^2 \quad (|\Omega| = \text{area/volume of } \Omega)$$

- Solve via Lagrange multipliers.

ROF Denoising

- State the denoising problem as minimizing $J(u)$ under the constraints

$$\int_\Omega u\,dx = \int_\Omega u_o\,dx, \quad \int_\Omega (u-u_0)^2\,dx = |\Omega|\sigma^2 \quad (|\Omega| = \text{area/volume of } \Omega)$$

- Solve via Lagrange multipliers.

TV-denoising

- Chambolle-Lions: there exists λ such the solution minimizes

$$E_{TV}(u) = \frac{1}{2}\int_\Omega (Ku - u_0)^2\,dx + \lambda \int_\Omega |\nabla u|\,dx$$

- Euler-Lagrange equation:

$$K^*(Ku - u_0) - \lambda \text{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0.$$

- The term $\text{div}\left(\frac{\nabla u}{|\nabla u|}\right)$ is highly non linear. Problems especially when $|\nabla u| = 0$.
- In fact $\frac{\nabla u/}{|\nabla u|}(x)$ is the unit normal of the level line of u at x and $\text{div}\left(\frac{\nabla u}{|\nabla u|}\right)$ is the (mean)curvature of the level line: not defined when the level line is singular or does not exist!

TV-denoising

- Chambolle-Lions: there exists λ such the solution minimizes

$$E_{TV}(u) = \frac{1}{2}\int_\Omega (Ku - u_0)^2\, dx + \lambda \int_\Omega |\nabla u|\, dx$$

- Euler-Lagrange equation:

$$K^*(Ku - u_0) - \lambda \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0.$$

- The term $\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)$ is highly non linear. Problems especially when $|\nabla u| = 0$.
- In fact $\frac{\nabla u/}{|\nabla u|}(x)$ is the unit normal of the level line of u at x and $\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)$ is the (mean)curvature of the level line: not defined when the level line is singular or does not exist!

TV-denoising

- Chambolle-Lions: there exists λ such the solution minimizes

$$E_{TV}(u) = \frac{1}{2}\int_\Omega (Ku - u_0)^2 \, dx + \lambda \int_\Omega |\nabla u| \, dx$$

- Euler-Lagrange equation:

$$K^*(Ku - u_0) - \lambda \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0.$$

- The term div $\left(\frac{\nabla u}{|\nabla u|}\right)$ is highly non linear. Problems especially when $|\nabla u| = 0$.
- In fact $\frac{\nabla u /}{|\nabla u|}(x)$ is the unit normal of the level line of u at x and div $\left(\frac{\nabla u}{|\nabla u|}\right)$ is the (mean)curvature of the level line: not defined when the level line is singular or does not exist!

TV-denoising

- Chambolle-Lions: there exists λ such the solution minimizes

$$E_{TV}(u) = \frac{1}{2}\int_\Omega (Ku - u_0)^2\,dx + \lambda \int_\Omega |\nabla u|\,dx$$

- Euler-Lagrange equation:

$$K^*(Ku - u_0) - \lambda \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0.$$

- The term $\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)$ is highly non linear. Problems especially when $|\nabla u| = 0$.
- In fact $\frac{\nabla u/}{|\nabla u|}(x)$ is the unit normal of the level line of u at x and $\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)$ is the (mean)curvature of the level line: not defined when the level line is singular or does not exist!

TV-denoising

- Chambolle-Lions: there exists λ such the solution minimizes

$$E_{TV}(u) = \frac{1}{2}\int_\Omega (Ku - u_0)^2\,dx + \lambda\int_\Omega |\nabla u|\,dx$$

- Euler-Lagrange equation:

$$K^*(Ku - u_0) - \lambda \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0.$$

- The term $\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)$ is highly non linear. Problems especially when $|\nabla u| = 0$.
- In fact $\frac{\nabla u/}{|\nabla u|}(x)$ is the unit normal of the level line of u at x and $\mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)$ is the (mean)curvature of the level line: not defined when the level line is singular or does not exist!

Acar-Vogel

- Replace it by regularized version

$$|\nabla u|_\beta = \sqrt{|\nabla u|^2 + \beta}, \quad \beta > 0$$

- Acar - Vogel show that

$$\lim_{\beta \to 0} \left(J_\beta(u) = \int_\Omega |\nabla u|_\beta \, dx \right) = J(u).$$

- Replace energy by

$$E'(u) = \int_\Omega (Ku - u_0)^2 \, dx + \lambda J_\beta(u)$$

- Euler-Lagrange equation:

$$K^*(Ku - u_0) - \lambda \mathrm{div} \left(\frac{\nabla u}{|\nabla u|_\beta} \right) = 0$$

The null denominator problem disappears.

Acar-Vogel

- Replace it by regularized version

$$|\nabla u|_\beta = \sqrt{|\nabla u|^2 + \beta}, \quad \beta > 0$$

- Acar - Vogel show that

$$\lim_{\beta \to 0} \left(J_\beta(u) = \int_\Omega |\nabla u|_\beta \, dx \right) = J(u).$$

- Replace energy by

$$E'(u) = \int_\Omega (Ku - u_0)^2 \, dx + \lambda J_\beta(u)$$

- Euler-Lagrange equation:

$$K^*(Ku - u_0) - \lambda \mathrm{div} \left(\frac{\nabla u}{|\nabla u|_\beta} \right) = 0$$

The null denominator problem disappears.

Acar-Vogel

- Replace it by regularized version

$$|\nabla u|_\beta = \sqrt{|\nabla u|^2 + \beta}, \quad \beta > 0$$

- Acar - Vogel show that

$$\lim_{\beta \to 0} \left(J_\beta(u) = \int_\Omega |\nabla u|_\beta \, dx \right) = J(u).$$

- Replace energy by

$$E'(u) = \int_\Omega (Ku - u_0)^2 \, dx + \lambda J_\beta(u)$$

- Euler-Lagrange equation:

$$K^*(Ku - u_0) - \lambda \text{div} \left(\frac{\nabla u}{|\nabla u|_\beta} \right) = 0$$

The null denominator problem disappears.

Acar-Vogel

- Replace it by regularized version

$$|\nabla u|_\beta = \sqrt{|\nabla u|^2 + \beta}, \quad \beta > 0$$

- Acar - Vogel show that

$$\lim_{\beta \to 0} \left(J_\beta(u) = \int_\Omega |\nabla u|_\beta \, dx \right) = J(u).$$

- Replace energy by

$$E'(u) = \int_\Omega (Ku - u_0)^2 \, dx + \lambda J_\beta(u)$$

- Euler-Lagrange equation:

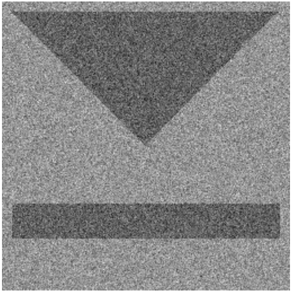
$$K^*(Ku - u_0) - \lambda \text{div} \left(\frac{\nabla u}{|\nabla u|_\beta} \right) = 0$$

The null denominator problem disappears.

Example

Implementation by finite differences, fixed-point strategy, linearization.

Original

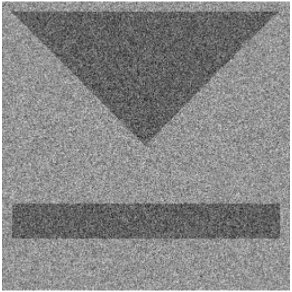


$\lambda = 1.5$, $\beta = 10^{-4}$

Example

Implementation by finite differences, fixed-point strategy, linearization.

Original



$\lambda = 1.5$, $\beta = 10^{-4}$

Outline

- Filling u in the subset $H \subset \Omega$ where data is missing, denoise known data
- Inpainting energy (Chan & Shen):

$$E_{ITV}(u) = \frac{1}{2}\int_{\Omega \setminus H} (u - u_0)^2\, dx + \lambda \int_{\Omega} |\nabla u|\, dx$$

- Euler-Lagrange Equation:

$$(u - u_0)\chi - \lambda \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0.$$

($\chi(x) = 1$ is $x \notin H$, 0 otherwise).

- Very similar to denoising. Can use the same approximation/implementation.

- Filling u in the subset $H \subset \Omega$ where data is missing, denoise known data
- Inpainting energy (Chan & Shen):

$$E_{ITV}(u) = \frac{1}{2}\int_{\Omega\setminus H}(u-u_0)^2\,dx + \lambda\int_\Omega |\nabla u|\,dx$$

- Euler-Lagrange Equation:

$$(u-u_0)\chi - \lambda \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0.$$

($\chi(x) = 1$ is $x \notin H$, 0 otherwise).
- Very similar to denoising. Can use the same approximation/implementation.

- Filling u in the subset $H \subset \Omega$ where data is missing, denoise known data
- Inpainting energy (Chan & Shen):

$$E_{ITV}(u) = \frac{1}{2}\int_{\Omega \setminus H} (u - u_0)^2 \, dx + \lambda \int_{\Omega} |\nabla u| \, dx$$

- Euler-Lagrange Equation:

$$(u - u_0)\chi - \lambda \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0.$$

($\chi(x) = 1$ is $x \notin H$, 0 otherwise).

- Very similar to denoising. Can use the same approximation/implementation.

- Filling u in the subset $H \subset \Omega$ where data is missing, denoise known data
- Inpainting energy (Chan & Shen):

$$E_{ITV}(u) = \frac{1}{2}\int_{\Omega\setminus H}(u-u_0)^2\,dx + \lambda\int_{\Omega}|\nabla u|\,dx$$

- Euler-Lagrange Equation:

$$(u-u_0)\chi - \lambda \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0.$$

($\chi(x) = 1$ is $x \notin H$, 0 otherwise).

- Very similar to denoising. Can use the same approximation/implementation.

- Filling u in the subset $H \subset \Omega$ where data is missing, denoise known data
- Inpainting energy (Chan & Shen):

$$E_{ITV}(u) = \frac{1}{2}\int_{\Omega \setminus H} (u - u_0)^2 \, dx + \lambda \int_{\Omega} |\nabla u| \, dx$$

- Euler-Lagrange Equation:

$$(u - u_0)\chi - \lambda \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0.$$

  ($\chi(x) = 1$ is $x \notin H$, 0 otherwise).

- Very similar to denoising. Can use the same approximation/implementation.

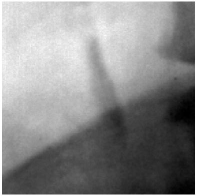
Degraded

Inpainted

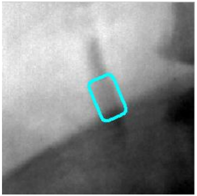
Segmention

Inpainting - driven segmention (Lauze, Nielsen 2008, IJCV)

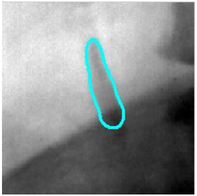
Aortic calcifiction



Detection



Segmention



Outline

- With definition of total variation as

$$J(u) = \int_\Omega |\nabla u| \, dx$$

u ***must have (weak) derivatives***.

- But we just saw that the computation is possible for a step-edge $u(x) = 0$, $x < 0$, $u(x) = a$, $x > 0$:

$$\int_{-1}^{1} |u'(x)| \, dx = |a|$$

- Can we avoid the use of derivatives of u?

- With definition of total variation as

$$J(u) = \int_{\Omega} |\nabla u| \, dx$$

u ***must have (weak) derivatives***.

- But we just saw that the computation is possible for a step-edge $u(x) = 0$, $x < 0$, $u(x) = a$, $x > 0$:

$$\int_{-1}^{1} |u'(x)| \, dx = |a|$$

- Can we avoid the use of derivatives of u?

- With definition of total variation as

$$J(u) = \int_\Omega |\nabla u| \, dx$$

u ***must have (weak) derivatives***.

- But we just saw that the computation is possible for a step-edge $u(x) = 0$, $x < 0$, $u(x) = a$, $x > 0$:

$$\int_{-1}^{1} |u'(x)| \, dx = |a|$$

- Can we avoid the use of derivatives of u?

- With definition of total variation as

$$J(u) = \int_{\Omega} |\nabla u| \, dx$$

  u ***must have (weak) derivatives***.

- But we just saw that the computation is possible for a step-edge $u(x) = 0$, $x < 0$, $u(x) = a$, $x > 0$:

$$\int_{-1}^{1} |u'(x)| \, dx = |a|$$

- Can we avoid the use of derivatives of u?

- Assume first that ∇u exists.

$$|\nabla u| = \nabla u \cdot \frac{\nabla u}{|\nabla u|}$$

(except when $\nabla u = 0$) and $\frac{\nabla u}{|\nabla u|}$ is the normal to the level lines of u, it has everywhere norm 1.

- Let V the set of vector fields $v(x)$ on Ω with $|v(x)| \leq 1$. I claim

$$J(u) = \sup_{v \in V} \int_\Omega \nabla u(x) \cdot v(x)\, dx$$

(consequence of Cauchy-Schwarz inequality).

- Assume first that ∇u exists.

$$|\nabla u| = \nabla u \cdot \frac{\nabla u}{|\nabla u|}$$

(except when $\nabla u = 0$) and $\frac{\nabla u}{|\nabla u|}$ is the normal to the level lines of u, it has everywhere norm 1.

- Let V the set of vector fields $v(x)$ on Ω with $|v(x)| \leq 1$. I claim

$$J(u) = \sup_{v \in V} \int_\Omega \nabla u(x) \cdot v(x)\, dx$$

(consequence of Cauchy-Schwarz inequality).

- Assume first that ∇u exists.

$$|\nabla u| = \nabla u \cdot \frac{\nabla u}{|\nabla u|}$$

(except when $\nabla u = 0$) and $\frac{\nabla u}{|\nabla u|}$ is the normal to the level lines of u, it has everywhere norm 1.

- Let V the set of vector fields $v(x)$ on Ω with $|v(x)| \leq 1$. I claim

$$J(u) = \sup_{v \in V} \int_\Omega \nabla u(x) \cdot v(x)\, dx$$

(consequence of Cauchy-Schwarz inequality).

- Restrict to the set W of such v's that are differentiable and vanishing at $\partial\Omega$, the boundary of Ω Then

$$J(u) = \sup_{v \in W} \int_\Omega \nabla u(x) \cdot v(x)\, dx$$

- But then I can use Divergence theorem: $H \subset D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}$ differentiable function, $g = (g^1, \dots, g^n) : D \to \mathbb{R}^n$ differentiable vector field and $\operatorname{div} g = \sum_{i=1}^n g^i_{x_i}$,

$$\int_H \nabla f \cdot g\, dx = -\int_H f \operatorname{div} g\, dx + \int_{\partial H} f g \cdot n(s)\, ds$$

  with $n(s)$ exterior normal field to ∂H.

- Apply it to J(u) above:

$$J(u) = \sup_{v \in W} \left(-\int_\Omega u(x) \operatorname{div} v(x)\, dx \right)$$

- The gradient has disappeared from u! This is the classical definition of total variation.
- Note that when $\nabla u(x) \neq 0$, optimal $v(x) = (\nabla u / |\nabla| u)(x)$ and $\operatorname{div} v(x)$ is the mean curvature of the level set of u at x. Geometry is there!

- Restrict to the set W of such v's that are differentiable and vanishing at $\partial\Omega$, the boundary of Ω Then

$$J(u) = \sup_{v \in W} \int_\Omega \nabla u(x) \cdot v(x)\, dx$$

- But then I can use Divergence theorem: $H \subset D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}$ differentiable function, $g = (g^1, \dots, g^n) : D \to \mathbb{R}^n$ differentiable vector field and $\operatorname{div} g = \sum_{i=1}^n g^i_{x_i}$,

$$\int_H \nabla f \cdot g\, dx = -\int_H f \operatorname{div} g\, dx + \int_{\partial H} f g \cdot n(s)\, ds$$

with $n(s)$ exterior normal field to ∂H.

- Apply it to J(u) above:

$$J(u) = \sup_{v \in W} \left(-\int_\Omega u(x) \operatorname{div} v(x)\, dx \right)$$

- The gradient has disappeared from u! This is the classical definition of total variation.
- Note that when $\nabla u(x) \neq 0$, optimal $v(x) = (\nabla u / |\nabla| u)(x)$ and $\operatorname{div} v(x)$ is the mean curvature of the level set of u at x. Geometry is there!

- Restrict to the set W of such v's that are differentiable and vanishing at $\partial\Omega$, the boundary of Ω Then

$$J(u) = \sup_{v \in W} \int_\Omega \nabla u(x) \cdot v(x)\, dx$$

- But then I can use Divergence theorem: $H \subset D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}$ differentiable function, $g = (g^1, \dots, g^n) : D \to \mathbb{R}^n$ differentiable vector field and $\operatorname{div} g = \sum_{i=1}^n g^i_{x_i}$,

$$\int_H \nabla f \cdot g\, dx = -\int_H f \operatorname{div} g\, dx + \int_{\partial H} f g \cdot n(s)\, ds$$

with $n(s)$ exterior normal field to ∂H.

- Apply it to J(u) above:

$$J(u) = \sup_{v \in W} \left(-\int_\Omega u(x) \operatorname{div} v(x)\, dx \right)$$

- The gradient has disappeared from u! This is the classical definition of total variation.
- Note that when $\nabla u(x) \neq 0$, optimal $v(x) = (\nabla u / |\nabla| u)(x)$ and $\operatorname{div} v(x)$ is the mean curvature of the level set of u at x. Geometry is there!

- Restrict to the set W of such v's that are differentiable and vanishing at $\partial\Omega$, the boundary of Ω Then

$$J(u) = \sup_{v \in W} \int_\Omega \nabla u(x) \cdot v(x)\, dx$$

- But then I can use Divergence theorem: $H \subset D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}$ differentiable function, $g = (g^1, \dots, g^n) : D \to \mathbb{R}^n$ differentiable vector field and $\operatorname{div} g = \sum_{i=1}^n g^i_{x_i}$,

$$\int_H \nabla f \cdot g\, dx = -\int_H f \operatorname{div} g\, dx + \int_{\partial H} fg \cdot n(s)\, ds$$

with $n(s)$ exterior normal field to ∂H.

- Apply it to J(u) above:

$$J(u) = \sup_{v \in W} \left(-\int_\Omega u(x) \operatorname{div} v(x)\, dx \right)$$

- The gradient has disappeared from u! This is the classical definition of total variation.
- Note that when $\nabla u(x) \neq 0$, optimal $v(x) = (\nabla u / |\nabla| u)(x)$ and $\operatorname{div} v(x)$ is the mean curvature of the level set of u at x. Geometry is there!

- Restrict to the set W of such v's that are differentiable and vanishing at $\partial\Omega$, the boundary of Ω Then

$$J(u) = \sup_{v \in W} \int_\Omega \nabla u(x) \cdot v(x)\, dx$$

- But then I can use Divergence theorem: $H \subset D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}$ differentiable function, $g = (g^1, \dots, g^n) : D \to \mathbb{R}^n$ differentiable vector field and $\operatorname{div} g = \sum_{i=1}^n g^i_{x_i}$,

$$\int_H \nabla f \cdot g\, dx = -\int_H f \operatorname{div} g\, dx + \int_{\partial H} fg \cdot n(s)\, ds$$

  with $n(s)$ exterior normal field to ∂H.
- Apply it to J(u) above:

$$J(u) = \sup_{v \in W} \left(-\int_\Omega u(x) \operatorname{div} v(x)\, dx \right)$$

- The gradient has disappeared from u! This is the classical definition of total variation.
- Note that when $\nabla u(x) \neq 0$, optimal $v(x) = (\nabla u / |\nabla| u)(x)$ and $\operatorname{div} v(x)$ is the mean curvature of the level set of u at x. Geometry is there!

- Restrict to the set W of such v's that are differentiable and vanishing at $\partial\Omega$, the boundary of Ω Then

$$J(u) = \sup_{v \in W} \int_\Omega \nabla u(x) \cdot v(x)\, dx$$

- But then I can use Divergence theorem: $H \subset D \subset \mathbb{R}^n$, $f : D \to \mathbb{R}$ differentiable function, $g = (g^1, \dots, g^n) : D \to \mathbb{R}^n$ differentiable vector field and $\operatorname{div} g = \sum_{i=1}^n g^i_{x_i}$,

$$\int_H \nabla f \cdot g\, dx = -\int_H f \operatorname{div} g\, dx + \int_{\partial H} f g \cdot n(s)\, ds$$

  with $n(s)$ exterior normal field to ∂H.

- Apply it to J(u) above:

$$J(u) = \sup_{v \in W} \left(-\int_\Omega u(x) \operatorname{div} v(x)\, dx \right)$$

- The gradient has disappeared from u! This is the classical definition of total variation.
- Note that when $\nabla u(x) \neq 0$, optimal $v(x) = (\nabla u / |\nabla| u)(x)$ and $\operatorname{div} v(x)$ is the mean curvature of the level set of u at x. Geometry is there!

Outline

Step-edge

- u the step-edge function defined in previous slides. We compute $J(u)$ with the new definition.
- here $W = \{\phi : [-1, 1] \to \mathbb{R}$ differentiable, $\phi(-1) = \phi(1) = 0, |\phi(x)| \leq 1\}$,

$$J(u) = \sup_{\phi \in W} \int_{-1}^{1} u(x)\phi'(x)\,dx$$

- we compute

$$\begin{aligned}\int_{-1}^{1} u(x)\phi'(x)\,dx &= a \int_{0}^{1} \phi'(x)\,dx \\ &= a\left(\phi(1) - \phi(0)\right) \\ &= -a\phi(0)\end{aligned}$$

- As $-1 \leq \phi(0) \leq 1$, the maximum is $|a|$.

Step-edge

- u the step-edge function defined in previous slides. We compute $J(u)$ with the new definition.
- here $W = \{\phi : [-1,1] \to \mathbb{R}$ differentiable, $\phi(-1) = \phi(1) = 0, |\phi(x)| \leq 1\}$,

$$J(u) = \sup_{\phi \in W} \int_{-1}^{1} u(x)\phi'(x)\,dx$$

- we compute

$$\begin{aligned}\int_{-1}^{1} u(x)\phi'(x)\,dx &= a\int_{0}^{1} \phi'(x)\,dx \\ &= a\left(\phi(1) - \phi(0)\right) \\ &= -a\phi(0)\end{aligned}$$

- As $-1 \leq \phi(0) \leq 1$, the maximum is $|a|$.

Step-edge

- u the step-edge function defined in previous slides. We compute $J(u)$ with the new definition.
- here $W = \{\phi : [-1,1] \to \mathbb{R}$ differentiable, $\phi(-1) = \phi(1) = 0, |\phi(x)| \leq 1\}$,

$$J(u) = \sup_{\phi \in W} \int_{-1}^{1} u(x)\phi'(x)\,dx$$

- we compute

$$\begin{aligned}\int_{-1}^{1} u(x)\phi'(x)\,dx &= a\int_{0}^{1} \phi'(x)\,dx \\ &= a\left(\phi(1) - \phi(0)\right) \\ &= -a\phi(0)\end{aligned}$$

- As $-1 \leq \phi(0) \leq 1$, the maximum is $|a|$.

Step-edge

- u the step-edge function defined in previous slides. We compute $J(u)$ with the new definition.
- here $W = \{\phi : [-1, 1] \to \mathbb{R}$ differentiable, $\phi(-1) = \phi(1) = 0, |\phi(x)| \leq 1\}$,

$$J(u) = \sup_{\phi \in W} \int_{-1}^{1} u(x)\phi'(x)\,dx$$

- we compute

$$\begin{aligned}\int_{-1}^{1} u(x)\phi'(x)\,dx &= a\int_{0}^{1} \phi'(x)\,dx \\ &= a\left(\phi(1) - \phi(0)\right) \\ &= -a\phi(0)\end{aligned}$$

- As $-1 \leq \phi(0) \leq 1$, the maximum is $|a|$.

Step-edge

- u the step-edge function defined in previous slides. We compute $J(u)$ with the new definition.
- here $W = \{\phi : [-1,1] \to \mathbb{R} \text{ differentiable}, \phi(-1) = \phi(1) = 0, |\phi(x)| \leq 1\}$,

$$J(u) = \sup_{\phi \in W} \int_{-1}^{1} u(x)\phi'(x)\,dx$$

- we compute

$$\begin{aligned}\int_{-1}^{1} u(x)\phi'(x)\,dx &= a\int_{0}^{1} \phi'(x)\,dx \\ &= a\left(\phi(1) - \phi(0)\right) \\ &= -a\phi(0)\end{aligned}$$

- As $-1 \leq \phi(0) \leq 1$, the maximum is $|a|$.

Step-edge

- u the step-edge function defined in previous slides. We compute $J(u)$ with the new definition.
- here $W = \{\phi : [-1,1] \to \mathbb{R} \text{ differentiable}, \phi(-1) = \phi(1) = 0, |\phi(x)| \leq 1\}$,

$$J(u) = \sup_{\phi \in W} \int_{-1}^{1} u(x)\phi'(x)\,dx$$

- we compute

$$\begin{aligned}\int_{-1}^{1} u(x)\phi'(x)\,dx &= a\int_{0}^{1} \phi'(x)\,dx \\ &= a\left(\phi(1) - \phi(0)\right) \\ &= -a\phi(0)\end{aligned}$$

- As $-1 \leq \phi(0) \leq 1$, the maximum is $|a|$.

2D example

- B open set with regular boundary curve $partialB$, Ω large enough to contain B and χ_B the characteristic function of B

$$\chi_B(x) = \begin{cases} 1 & x \in B \\ 0 & x \notin B \end{cases}$$

- For $v \in W$, by the divergence theorem on B and its boundary ∂B

$$\begin{aligned} \int_\Omega \chi(x) \operatorname{div} v(x)\, dx &= \int_B \operatorname{div} v(x)\, dx \\ &= -\int_{\partial B} v(s) \cdot n(s)\, ds \end{aligned}$$

($n(s)$ is the exterior normal to ∂B)

- This integral is maximized when $v = -n$: length of ∂B perimeter of B.

2D example

- B open set with regular boundary curve *partialB*, Ω large enough to contain B and χ_B the characteristic function of B

$$\chi_B(x) = \begin{cases} 1 & x \in B \\ 0 & x \notin B \end{cases}$$

- For $v \in W$, by the divergence theorem on B and its boundary ∂B

$$\int_\Omega \chi(x) \mathrm{div}\, v(x)\, dx = \int_B \mathrm{div}\, v(x)\, dx$$
$$= -\int_{\partial B} v(s) \cdot n(s)\, ds$$

  ($n(s)$ is the exterior normal to ∂B)
- This integral is maximized when $v = -n$: length of ∂B perimeter of B.

2D example

- B open set with regular boundary curve *partialB*, Ω large enough to contain B and χ_B the characteristic function of B

$$\chi_B(x) = \begin{cases} 1 & x \in B \\ 0 & x \notin B \end{cases}$$

- For $v \in W$, by the divergence theorem on B and its boundary ∂B

$$\begin{aligned} \int_\Omega \chi(x) \mathrm{div}\, v(x)\, dx &= \int_B \mathrm{div}\, v(x)\, dx \\ &= -\int_{\partial B} v(s) \cdot n(s)\, ds \end{aligned}$$

($n(s)$ is the exterior normal to ∂B)

- This integral is maximized when $v = -n$: length of ∂B perimeter of B.

2D example

- B open set with regular boundary curve *partialB*, Ω large enough to contain B and χ_B the characteristic function of B

$$\chi_B(x) = \begin{cases} 1 & x \in B \\ 0 & x \notin B \end{cases}$$

- For $v \in W$, by the divergence theorem on B and its boundary ∂B

$$\begin{aligned} \int_\Omega \chi(x) \mathrm{div}\, v(x)\, dx &= \int_B \mathrm{div}\, v(x)\, dx \\ &= -\int_{\partial B} v(s) \cdot n(s)\, ds \end{aligned}$$

  ($n(s)$ is the exterior normal to ∂B)
- This integral is maximized when $v = -n$: length of ∂B perimeter of B.

Sets of finite perimeter

- Let $H \subset \Omega$. If its characteristic function χ_H satisfies

$$J(\chi_H) < \infty$$

  H is called set of finite perimeter (and $Per_\Omega(H) := J(\chi_H)$ is its perimeter)

- This is used for instance in the Chan and Vese algorithm.
- If $J(u) < \infty$ and $H_t = \{x \in \Omega,\ u(x) < t\}$ the lower t-level set of u,

$$J(u) = \int_{-\infty}^{+\infty} J(\chi_{H_t})\, dt \quad \text{Coarea formula}$$

Sets of finite perimeter

- Let $H \subset \Omega$. If its characteristic function χ_H satisfies

$$J(\chi_H) < \infty$$

  H is called set of finite perimeter (and $Per_\Omega(H) := J(\chi_H)$ is its perimeter)

- This is used for instance in the Chan and Vese algorithm.

- If $J(u) < \infty$ and $H_t = \{x \in \Omega,\ u(x) < t\}$ the lower t-level set of u,

$$J(u) = \int_{-\infty}^{+\infty} J(\chi_{H_t})\, dt \quad \text{Coarea formula}$$

Sets of finite perimeter

- Let $H \subset \Omega$. If its characteristic function χ_H satisfies

$$J(\chi_H) < \infty$$

  H is called set of finite perimeter (and $Per_\Omega(H) := J(\chi_H)$ is its perimeter)

- This is used for instance in the Chan and Vese algorithm.
- If $J(u) < \infty$ and $H_t = \{x \in \Omega,\ u(x) < t\}$ the lower t-level set of u,

$$J(u) = \int_{-\infty}^{+\infty} J(\chi_{H_t})\, dt \quad \text{Coarea formula}$$

Outline

Chambolle algorithm

- Let $K \in L^2(\Omega)$ the closure of the set $\{\operatorname{div} v,\ v \in C_0^1(\Omega)^2,\ |v(x)| \leq 1\}$ i.e. the image of W by div.
- Then
$$J(u) = \sup_{\phi \in K} \left(\int_\Omega u\,\phi\,dx = \langle u, \phi \rangle_{L^2(\Omega)} \right)$$
- Solution of the denoising problem arg.min $\int_\Omega (u - u_0)^2 + \lambda J(u)$ given by
$$u = u_0 - \pi_{\lambda K}(u_0)$$
with $\pi_{\lambda K}$ orthogonal projection onto the convex set λK (Chambolle).
- Needs a bit of convex analysis to show that: subdifferentials and subgradients, Fenchel transforms, indicators/characteristic functions and elementary results on them

Chambolle algorithm

- Let $K \in L^2(\Omega)$ the closure of the set $\{\mathrm{div}\, v,\ v \in C_0^1(\Omega)^2,\ |v(x)| \leq 1\}$ i.e. the image of W by div.
- Then
$$J(u) = \sup_{\phi \in K} \left(\int_\Omega u\, \phi\, dx = \langle u, \phi \rangle_{L^2(\Omega)} \right)$$
- Solution of the denoising problem arg.min $\int_\Omega (u - u_0)^2 + \lambda J(u)$ given by
$$u = u_0 - \pi_{\lambda K}(u_0)$$
with $\pi_{\lambda K}$ orthogonal projection onto the convex set λK (Chambolle).
- Needs a bit of convex analysis to show that: subdifferentials and subgradients, Fenchel transforms, indicators/characteristic functions and elementary results on them

Chambolle algorithm

- Let $K \in L^2(\Omega)$ the closure of the set $\{\text{div}\, v,\ v \in C_0^1(\Omega)^2,\ |v(x)| \leq 1\}$ i.e. the image of W by div.
- Then

$$J(u) = \sup_{\phi \in K} \left(\int_\Omega u\, \phi\, dx = \langle u, \phi \rangle_{L^2(\Omega)} \right)$$

- Solution of the denoising problem arg.min $\int_\Omega (u - u_0)^2 + \lambda J(u)$ given by

$$u = u_0 - \pi_{\lambda K}(u_0)$$

  with $\pi_{\lambda K}$ orthogonal projection onto the convex set λK (Chambolle).
- Needs a bit of convex analysis to show that: subdifferentials and subgradients, Fenchel transforms, indicators/characteristic functions and elementary results on them

Chambolle algorithm

- Let $K \in L^2(\Omega)$ the closure of the set $\{\mathrm{div}\, v,\ v \in C_0^1(\Omega)^2,\ |v(x)| \leq 1\}$ i.e. the image of W by div.
- Then

$$J(u) = \sup_{\phi \in K} \left(\int_\Omega u\, \phi\, dx = \langle u, \phi \rangle_{L^2(\Omega)} \right)$$

- Solution of the denoising problem arg.min $\int_\Omega (u - u_0)^2 + \lambda J(u)$ given by

$$u = u_0 - \pi_{\lambda K}(u_0)$$

  with $\pi_{\lambda K}$ orthogonal projection onto the convex set λK (Chambolle).
- Needs a bit of convex analysis to show that: subdifferentials and subgradients, Fenchel transforms, indicators/characteristic functions and elementary results on them

Chambolle algorithm

- Let $K \in L^2(\Omega)$ the closure of the set $\{\mathrm{div}\, v,\ v \in C_0^1(\Omega)^2,\ |v(x)| \leq 1\}$ i.e. the image of W by div.
- Then
$$J(u) = \sup_{\phi \in K} \left(\int_\Omega u\, \phi\, dx = \langle u, \phi \rangle_{L^2(\Omega)} \right)$$
- Solution of the denoising problem arg.min $\int_\Omega (u - u_0)^2 + \lambda J(u)$ given by
$$u = u_0 - \pi_{\lambda K}(u_0)$$
with $\pi_{\lambda K}$ orthogonal projection onto the convex set λK (Chambolle).
- Needs a bit of convex analysis to show that: subdifferentials and subgradients, Fenchel transforms, indicators/characteristic functions and elementary results on them

Chambolle algorithm

- Let $K \in L^2(\Omega)$ the closure of the set $\{\mathrm{div}\, v,\ v \in C_0^1(\Omega)^2,\ |v(x)| \leq 1\}$ i.e. the image of W by div.
- Then

$$J(u) = \sup_{\phi \in K} \left(\int_\Omega u\,\phi\,dx = \langle u, \phi \rangle_{L^2(\Omega)} \right)$$

- Solution of the denoising problem arg.min $\int_\Omega (u - u_0)^2 + \lambda J(u)$ given by

$$u = u_0 - \pi_{\lambda K}(u_0)$$

  with $\pi_{\lambda K}$ orthogonal projection onto the convex set λK (Chambolle).
- Needs a bit of convex analysis to show that: subdifferentials and subgradients, Fenchel transforms, indicators/characteristic functions and elementary results on them

Fenchel Transform

- X Hilbert space, $f : X \to \mathbb{R}$ convex, proper. Fenchel transform of F:

$$F^*(v) = \sup_{u \in X} \left(\langle u, v \rangle_X - F(u) \right)$$

- Geometric meaning: take u^* such that $F^*(u^*) < +\infty$: the affine function

$$a(u) = \langle u, u^* \rangle - F^*(u^*)$$

  is tangent to F and $a(0) = -F^*(u^*)$.

Fenchel Transform

- X Hilbert space, $f : X \to \mathbb{R}$ convex, proper. Fenchel transform of F:

$$F^*(v) = \sup_{u \in X} \left(\langle u, v \rangle_X - F(u) \right)$$

- Geometric meaning: take u^* such that $F^*(u^*) < +\infty$: the affine function

$$a(u) = \langle u, u^* \rangle - F^*(u^*)$$

is tangent to F and $a(0) = -F^*(u^*)$.

Fenchel Transform

- X Hilbert space, $f : X \to \mathbb{R}$ convex, proper. Fenchel transform of F:

$$F^*(v) = \sup_{u \in X} \left(\langle u, v \rangle_X - F(u) \right)$$

- Geometric meaning: take u^* such that $F^*(u^*) < +\infty$: the affine function

$$a(u) = \langle u, u^* \rangle - F^*(u^*)$$

is tangent to F and $a(0) = -F^*(u^*)$.

Fenchel transform

- Interesting properties:
  - Convex
  - If Φ is the transform of F and $\lambda > 0$, then the transform of $u \mapsto \lambda F(\lambda^{-1}(u))$ is $\lambda\Phi$.
  - If F 1-homogeneous, i.e. $F(\lambda u) = \lambda F(u)$ then $F^*(u)$ only take values 0 and $+\infty$ as the property above implies $F^* = \lambda F^*$, $\lambda > 0$.
  - In that case, the set where $F^* = 0$ i a closed convex set of X, $F^* = \delta_C$, the indicator function of C,

$$\delta_C(x) = \begin{cases} 0 & , x \in C \\ +\infty & , x \notin C \end{cases}$$

  - For $x \in \mathbb{R} \mapsto |x|$, $C = [-1, 1]$
  - For $J(u)$, $C = K$

Fenchel transform

- Interesting properties:
  - Convex
  - if Φ is the transform of F and $\lambda > 0$, then the transform of $u \mapsto \lambda F(\lambda^{-1}(u)$ is $\lambda\Phi$.
  - if F 1-homogeneous, i.e. $F(\lambda u) = \lambda F(u)$ then $F^*(u)$ only take values 0 and $+\infty$ as the property above implies $F^* = \lambda F^*$, $\lambda > 0$.
  - In that case, the set where $F^* = 0$ i a closed convex set of X, $F^* = \delta_C$, the indicator function of C,

$$\delta_C(x) = \begin{cases} 0 & , x \in C \\ +\infty & , x \notin C \end{cases}$$

  - For $x \in \mathbb{R} \mapsto |x|$, $C = [-1, 1]$
  - For $J(u)$, $C = K$.

Fenchel transform

- Interesting properties:
  - Convex
  - if Φ is the transform of F and $\lambda > 0$, then the transform of $u \mapsto \lambda F(\lambda^{-1}(u)$ is $\lambda\Phi$.
  - if F 1-homogeneous, i.e. $F(\lambda u) = \lambda F(u)$ then $F^*(u)$ only take values 0 and $+\infty$ as the property above implies $F^* = \lambda F^*$, $\lambda > 0$.
  - In that case, the set where $F^* = 0$ i a closed convex set of X, $F^* = \delta_C$, the indicator function of C,

$$\delta_C(x) = \begin{cases} 0 & , x \in C \\ +\infty & , x \notin C \end{cases}$$

  - For $x \in \mathbb{R} \mapsto |x|$, $C = [-1, 1]$
  - For $J(u)$, $C = K$.

Fenchel transform

- Interesting properties:
  - Convex
  - if Φ is the transform of F and $\lambda > 0$, then the transform of $u \mapsto \lambda F(\lambda^{-1}(u)$ is $\lambda\Phi$.
  - if F 1-homogeneous, i.e. $F(\lambda u) = \lambda F(u)$ then $F^*(u)$ only take values 0 and $+\infty$ as the property above implies $F^* = \lambda F^*$, $\lambda > 0$.
  - In that case, the set where $F^* = 0$ i a closed convex set of X, $F^* = \delta_C$, the indicator function of C,

$$\delta_C(x) = \begin{cases} 0 & , x \in C \\ +\infty & , x \notin C \end{cases}$$

  - For $x \in \mathbb{R} \mapsto |x|$, $C = [-1, 1]$
  - For $J(u)$, $C = K$.

Fenchel transform

- Interesting properties:
  - Convex
  - if Φ is the transform of F and $\lambda > 0$, then the transform of $u \mapsto \lambda F(\lambda^{-1}(u)$ is $\lambda\Phi$.
  - if F 1-homogeneous, i.e. $F(\lambda u) = \lambda F(u)$ then $F^*(u)$ only take values 0 and $+\infty$ as the property above implies $F^* = \lambda F^*$, $\lambda > 0$.
  - In that case, the set where $F^* = 0$ i a closed convex set of X, $F^* = \delta_C$, the indicator function of C,

$$\delta_C(x) = \begin{cases} 0 & , x \in C \\ +\infty & , x \notin C \end{cases}$$

  - For $x \in \mathbb{R} \mapsto |x|$, $C = [-1, 1]$
  - For $J(u)$, $C = K$.

Fenchel transform

- Interesting properties:
  - Convex
  - if Φ is the transform of F and $\lambda > 0$, then the transform of $u \mapsto \lambda F(\lambda^{-1}(u)$ is $\lambda\Phi$.
  - if F 1-homogeneous, i.e. $F(\lambda u) = \lambda F(u)$ then $F^*(u)$ only take values 0 and $+\infty$ as the property above implies $F^* = \lambda F^*$, $\lambda > 0$.
  - In that case, the set where $F^* = 0$ i a closed convex set of X, $F^* = \delta_C$, the indicator function of C,

$$\delta_C(x) = \begin{cases} 0 & , x \in C \\ +\infty & , x \notin C \end{cases}$$

  - For $x \in \mathbb{R} \mapsto |x|$, $C = [-1, 1]$
  - For $J(u)$, $C = K$.

Fenchel transform

- Interesting properties:
  - Convex
  - if Φ is the transform of F and $\lambda > 0$, then the transform of $u \mapsto \lambda F(\lambda^{-1}(u)$ is $\lambda\Phi$.
  - if F 1-homogeneous, i.e. $F(\lambda u) = \lambda F(u)$ then $F^*(u)$ only take values 0 and $+\infty$ as the property above implies $F^* = \lambda F^*$, $\lambda > 0$.
  - In that case, the set where $F^* = 0$ i a closed convex set of X, $F^* = \delta_C$, the indicator function of C,

$$\delta_C(x) = \begin{cases} 0 & , x \in C \\ +\infty & , x \notin C \end{cases}$$

  - For $x \in \mathbb{R} \mapsto |x|$, $C = [-1, 1]$
  - For $J(u)$, $C = K$.

Fenchel transform

- Interesting properties:
  - Convex
  - if Φ is the transform of F and $\lambda > 0$, then the transform of $u \mapsto \lambda F(\lambda^{-1}(u)$ is $\lambda\Phi$.
  - if F 1-homogeneous, i.e. $F(\lambda u) = \lambda F(u)$ then $F^*(u)$ only take values 0 and $+\infty$ as the property above implies $F^* = \lambda F^*$, $\lambda > 0$.
  - In that case, the set where $F^* = 0$ i a closed convex set of X, $F^* = \delta_C$, the indicator function of C,

$$\delta_C(x) = \begin{cases} 0 & , x \in C \\ +\infty & , x \notin C \end{cases}$$

  - For $x \in \mathbb{R} \mapsto |x|$, $C = [-1, 1]$
  - For $J(u)$, $C = K$.

Subdifferentials

- subdifferential of F at u: $\partial F(u) = \{v \in X, F(w) - F(u) \geq \langle w - u, v \rangle, \forall w \in X\}$. $v \in \partial F(u)$ is a subgradient of F at u.
- Three fundamental (and easy) properties:
  - $0 \in \partial F(u)$ iff u global minimizer of F
  - $u^* \in \partial F(u) \Leftrightarrow F(u) + F^*(u^*) = \langle u, u^* \rangle$
  - Duality: $u^* \in \partial F(u) \Leftrightarrow u \in \partial F^*(u)$
- The duality above allows to transform optimization of homogeneous functions into domain constraints!

Subdifferentials

- subdifferential of F at u: $\partial F(u) = \{v \in X, F(w) - F(u) \geq \langle w - u, v\rangle, \forall w \in X\}$. $v \in \partial F(u)$ is a subgradient of F at u.
- Three fundamental (and easy) properties:
  - $0 \in \partial F(u)$ iff u global minimizer of F
  - $u^* \in \partial F(u) \Leftrightarrow F(u) + F^*(u^*) = \langle u, u^* \rangle$
  - Duality: $u^* \in \partial F(u) \Leftrightarrow u \in \partial F^*(u)$
- The duality above allows to transform optimization of homogeneous functions into domain constraints!

Subdifferentials

- subdifferential of F at u: $\partial F(u) = \{v \in X, F(w) - F(u) \geq \langle w - u, v \rangle, \forall w \in X\}$. $v \in \partial F(u)$ is a subgradient of F at u.
- Three fundamental (and easy) properties:
  - $0 \in \partial F(u)$ iff u global minimizer of F
  - $u^* \in \partial F(u) \Leftrightarrow F(u) + F^*(u^*) = \langle u, u^* \rangle$
  - Duality: $u^* \in \partial F(u) \Leftrightarrow u \in \partial F^*(u)$
- The duality above allows to transform optimization of homogeneous functions into domain constraints!

Subdifferentials

- subdifferential of F at u: $\partial F(u) = \{v \in X, F(w) - F(u) \geq \langle w - u, v\rangle, \forall w \in X\}$. $v \in \partial F(u)$ is a subgradient of F at u.
- Three fundamental (and easy) properties:
  - $0 \in \partial F(u)$ iff u global minimizer of F
  - $u^* \in \partial F(u) \Leftrightarrow F(u) + F^*(u^*) = \langle u, u^*\rangle$
  - Duality: $u^* \in \partial F(u) \Leftrightarrow u \in \partial F^*(u)$
- The duality above allows to transform optimization of homogeneous functions into domain constraints!

Subdifferentials

- subdifferential of F at u: $\partial F(u) = \{v \in X, F(w) - F(u) \geq \langle w - u, v \rangle, \forall w \in X\}$. $v \in \partial F(u)$ is a subgradient of F at u.
- Three fundamental (and easy) properties:
  - $0 \in \partial F(u)$ iff u global minimizer of F
  - $u^* \in \partial F(u) \Leftrightarrow F(u) + F^*(u^*) = \langle u, u^* \rangle$
  - Duality: $u^* \in \partial F(u) \Leftrightarrow u \in \partial F^*(u)$
- The duality above allows to transform optimization of homogeneous functions into domain constraints!

Subdifferentials

- subdifferential of F at u: $\partial F(u) = \{v \in X, F(w) - F(u) \geq \langle w - u, v \rangle, \forall w \in X\}$. $v \in \partial F(u)$ is a subgradient of F at u.
- Three fundamental (and easy) properties:
  - $0 \in \partial F(u)$ iff u global minimizer of F
  - $u^* \in \partial F(u) \Leftrightarrow F(u) + F^*(u^*) = \langle u, u^* \rangle$
  - Duality: $u^* \in \partial F(u) \Leftrightarrow u \in \partial F^*(u)$
- The duality above allows to transform optimization of homogeneous functions into domain constraints!

Subdifferentials

- subdifferential of F at u: $\partial F(u) = \{v \in X, F(w) - F(u) \geq \langle w - u, v \rangle, \forall w \in X\}$. $v \in \partial F(u)$ is a subgradient of F at u.
- Three fundamental (and easy) properties:
  - $0 \in \partial F(u)$ iff u global minimizer of F
  - $u^* \in \partial F(u) \Leftrightarrow F(u) + F^*(u^*) = \langle u, u^* \rangle$
  - Duality: $u^* \in \partial F(u) \Leftrightarrow u \in \partial F^*(u)$
- The duality above allows to transform optimization of homogeneous functions into domain constraints!

TV-denoising

- To minimize:

$$\frac{1}{2}\|u - u_0\|^2_{L^2(\Omega} + \lambda J(u)$$

- optimality condition:

$$0 \in u - u_0 + \lambda \partial J(u) \Leftrightarrow \frac{u_0 - u}{\lambda} \in \partial J(u)$$

- Duality

$$\frac{u_0}{\lambda} \in \frac{u_0 - u}{\lambda} + \frac{1}{\lambda}\partial J^*(\frac{u_0 - u}{\lambda})$$

- Set $w = \frac{u_0 - u}{\lambda}$: w satisfies

$$0 \in w - \frac{u_0}{\lambda} + \frac{1}{\lambda}\partial J^*(w)$$

- This is the subdifferential of the convex function

$$\frac{1}{2}\|w - u_0/\lambda\|^2 + \frac{1}{\lambda}J^*(w)$$

- But $J^*(w) = \delta_K(w)$: we get $w = \pi_K(g\lambda)$.

TV-denoising

- To minimize:

$$\frac{1}{2}\|u-u_0\|^2_{L^2(\Omega}+\lambda J(u)$$

- optimality condition:

$$0 \in u-u_0+\lambda\partial J(u) \Leftrightarrow \frac{u_0-u}{\lambda} \in \partial J(u)$$

- Duality

$$\frac{u_0}{\lambda} \in \frac{u_0-u}{\lambda}+\frac{1}{\lambda}\partial J^*(\frac{u_0-u}{\lambda})$$

- Set $w=\frac{u_0-u}{\lambda}$: w satisfies

$$0 \in w-\frac{u_0}{\lambda}+\frac{1}{\lambda}\partial J^*(w)$$

- This is the subdifferential of the convex function

$$\frac{1}{2}\|w-u_0/\lambda\|^2+\frac{1}{\lambda}J^*(w)$$

- But $J^*(w)=\delta_K(w)$: we get $w=\pi_K(g\lambda)$.

TV-denoising

- To minimize:

$$\frac{1}{2}\|u - u_0\|^2_{L^2(\Omega} + \lambda J(u)$$

- optimality condition:

$$0 \in u - u_0 + \lambda \partial J(u) \Leftrightarrow \frac{u_0 - u}{\lambda} \in \partial J(u)$$

- Duality

$$\frac{u_0}{\lambda} \in \frac{u_0 - u}{\lambda} + \frac{1}{\lambda}\partial J^*(\frac{u_0 - u}{\lambda})$$

- Set $w = \frac{u_0 - u}{\lambda}$: w satisfies

$$0 \in w - \frac{u_0}{\lambda} + \frac{1}{\lambda}\partial J^*(w)$$

- This is the subdifferential of the convex function

$$\frac{1}{2}\|w - u_0/\lambda\|^2 + \frac{1}{\lambda}J^*(w)$$

- But $J^*(w) = \delta_K(w)$: we get $w = \pi_K(g\lambda)$.

TV-denoising

- To minimize:

$$\frac{1}{2}\|u - u_0\|^2_{L^2(\Omega} + \lambda J(u)$$

- optimality condition:

$$0 \in u - u_0 + \lambda \partial J(u) \Leftrightarrow \frac{u_0 - u}{\lambda} \in \partial J(u)$$

- Duality

$$\frac{u_0}{\lambda} \in \frac{u_0 - u}{\lambda} + \frac{1}{\lambda}\partial J^*\left(\frac{u_0 - u}{\lambda}\right)$$

- Set $w = \frac{u_0 - u}{\lambda}$: w satisfies

$$0 \in w - \frac{u_0}{\lambda} + \frac{1}{\lambda}\partial J^*(w)$$

- This is the subdifferential of the convex function

$$\frac{1}{2}\|w - u_0/\lambda\|^2 + \frac{1}{\lambda}J^*(w)$$

- But $J^*(w) = \delta_K(w)$: we get $w = \pi_K(g\lambda)$.

TV-denoising

- To minimize:

$$\frac{1}{2}\|u-u_0\|^2_{L^2(\Omega} + \lambda J(u)$$

- optimality condition:

$$0 \in u - u_0 + \lambda \partial J(u) \Leftrightarrow \frac{u_0-u}{\lambda} \in \partial J(u)$$

- Duality

$$\frac{u_0}{\lambda} \in \frac{u_0-u}{\lambda} + \frac{1}{\lambda}\partial J^*(\frac{u_0-u}{\lambda})$$

- Set $w = \frac{u_0-u}{\lambda}$: w satisfies

$$0 \in w - \frac{u_0}{\lambda} + \frac{1}{\lambda}\partial J^*(w)$$

- This is the subdifferential of the convex function

$$\frac{1}{2}\|w - u_0/\lambda\|^2 + \frac{1}{\lambda}J^*(w)$$

- But $J^*(w) = \delta_K(w)$: we get $w = \pi_K(g\lambda)$.

TV-denoising

- To minimize:

$$\frac{1}{2}\|u-u_0\|^2_{L^2(\Omega} + \lambda J(u)$$

- optimality condition:

$$0 \in u - u_0 + \lambda \partial J(u) \Leftrightarrow \frac{u_0 - u}{\lambda} \in \partial J(u)$$

- Duality

$$\frac{u_0}{\lambda} \in \frac{u_0 - u}{\lambda} + \frac{1}{\lambda}\partial J^*(\frac{u_0 - u}{\lambda})$$

- Set $w = \frac{u_0 - u}{\lambda}$: w satisfies

$$0 \in w - \frac{u_0}{\lambda} + \frac{1}{\lambda}\partial J^*(w)$$

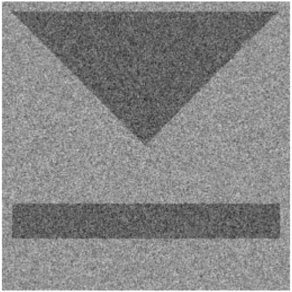
- This is the subdifferential of the convex function

$$\frac{1}{2}\|w - u_0/\lambda\|^2 + \frac{1}{\lambda}J^*(w)$$

- But $J^*(w) = \delta_K(w)$: we get $w = \pi_K(g\lambda)$.

Example

The usual original



Denoised by projection

Outline

Camerman Example

Solution of denoising energy present numerically stair-casing effect (Nikolova)

Original

xc

$\lambda = 100$

$\lambda = 500$

The gradient becomes “sparse”.

Bibliography

- Tikhonov, A. N.; Arsenin, V. Y. 1977. Solution of Ill-posed Problems.
- Wahba, G, 1990. Spline Models for Observational Data.
- Rudin, L.; Osher, S.; Fatemi, E. 1992. Nonlinear Total Variation Based Noise Removal Algorithms.
- Chambolle, A. 2004. An algorithm for Total Variation Minimization and Applications.
- Nikolova, M. 2004. Weakly Constrained Minimization: Application to the Estimation of Images and Signals Involving Constant Regions

The End