Non-smooth regularization and imaging

Models and optimization algorithms

Jingwei Liang

DAMTP, University of Cambridge

Outline

Outline

Regularization

Regularization

- A key tool in the solution of inverse problems.
- Introducing prior knowledge, making the approximation of ill-posed (pseudo-)inverses feasible.

Three approaches to motivate regularzation

- Variational regularization — roadmap of the course.
- MAP (maximum a-posteriori probability) — Bayesian perspective.
- Application driven — this lecture.

Outline

From Gaussian denoising to heat equation

Image observation

$$f = \mathring{u} + \varepsilon,$$

where ε is zero mean white Gaussian noise.

Original image

Noised image

Gaussian denoising

$$u = G_\sigma \star f,$$

where G_σ is the Gaussian kernel $G_\sigma(x) = \frac{1}{2\pi\sigma^2}\exp(-\frac{\|x\|^2}{\sigma^2})$.

Noised image | $\sigma = 1$ | $\sigma = 3$ | $\sigma = 10$

From Gaussian denoising to heat equation

Heat equation

$$\begin{aligned} \frac{\partial}{\partial t} u(x,t) &= \Delta u(x,t), \quad t > 0, \\ u(x,0) &= f, \end{aligned}$$

with appropriate boundary condition. Solution: $u(x,t) = G_{\sqrt{2t}} \star f$.

Noised image | $t = 1/2$ | $t = 4.5$ | $t = 50$

From Gaussian denoising to heat equation

Non-linear diffusion

Linear diffusion

- PDE appear as a natural way to denoise.
- Linear PDE (or convolution) does not preserve edges.
- Heat equation

$$\begin{aligned}\Delta u(x,t) &= \mathrm{div}\big(\nabla u(x,t)\big) \\ &= -\nabla^T\big(\nabla u(x,t)\big).\end{aligned}$$

Goal

- Diffusion along the edges.
- No diffusion cross the edges.

Non-linear diffusion

Non-linear diffusion [Perona and Malik '90]

$$\frac{\partial}{\partial t} u(x,t) = \mathrm{div}\big(\phi(\|\nabla u(x,t)\|)\nabla u(x,t)\big), \quad t > 0,$$
$$u(x,0) = f,$$

Choices of scalar function ϕ

$$\phi(x) = \frac{1}{\sqrt{1+x}} \quad \text{or} \quad \exp(-x).$$

Non-linear diffusion

TV flow [Rudin, Osher & Fatemi '92]

$$\frac{\partial}{\partial t}u(x,t) = \mathrm{div}\Big(\frac{1}{\|\nabla u\|}\nabla u\Big), \quad t > 0,$$

$$u(x,0) = f.$$

ϵ-regularized TV flow

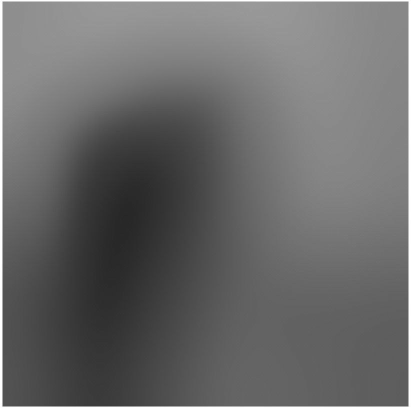
$$\frac{\partial}{\partial t}u(x,t) = \mathrm{div}\Big(\frac{1}{\sqrt{\|\nabla u\|^2 + \epsilon^2}}\nabla u\Big).$$


non-linear diffusion


tv flow

Non-linear diffusion

- Non-linear PDE (anisotropic diffusion) can preserve the edges of image.
- As it still is diffusion, the image becomes a constant as $t \to +\infty$.



Gradient flow

$$\begin{aligned}\frac{\partial}{\partial t}u(x,t) &= -\nabla E(u), \quad t > 0, \\ u(x,0) &= f.\end{aligned}$$

Consider $E(u) = \frac{1}{2}e(\|\nabla u\|^2)$, then

$$\nabla E(u) = \nabla^T\big(e'(\|\nabla u\|^2)\nabla u\big).$$

- Heat equation
$$e'(\|\nabla u\|^2) = 1.$$
- Non-linear diffusion
$$e'(\|\nabla u\|^2) = \frac{1}{\sqrt{1+\|\nabla u\|^2}}.$$
- TV flow
$$e'(\|\nabla u\|^2) = \frac{1}{\|\nabla u\|_1}.$$

Optimization problem

$$\min_{u} \; E(u).$$

Gradient descent: $u_0 = f$,

$$u_{k+1} = u_k - \gamma \nabla E(u_k).$$

Trivial solution: $c \in [0, 1]$

$$u(x) = c, \quad x \in \Omega.$$

Problem: the initial condition f is discarded...

Tradeoff: fidelity and diffusion

$$\min_u E(u) + \frac{1}{2\mu}\int_\Omega \|u - f\|^2 \mathrm{d}x.$$

- μ provides a balance between diffusion and fidelity.
- The value of μ depends on noise level.
- The quadratic fidelity term accounts for noise with *bounded L^2-norm*.

Gradient flow

$$\begin{aligned} \frac{\partial}{\partial t}u(x,t) &= -\nabla E(u) - \frac{1}{\mu}(u - f), \quad t > 0, \\ u(x,0) &= f. \end{aligned}$$

non-linear diffusion

tv flow

Outline

From diffusion to regularization

Previously

- To preserve edges — from isotropic diffusion to anisotropic diffusion.

Namely, there is certain *prior information* of u that we want to keep.

Regularization

- Promoting prior information to the solution.
- Making ill-posed problem “solvable” or preventing over-fitting.

Application

- Signal/image processing, compressed sensing, inverse problems
- Data science, machine learning
- Statstics
- ...

Examples: Tikhonov regularization

From now on: `discrete` setting...

Tikhonov regularization [Tikhonov '63]

$$\|\Gamma u\|^2$$

- Γ is some properly chosen linear operator.

Isotropic heat diffusion

$$\min_u \|\nabla u\|^2 + \frac{1}{2\mu}\|f - u\|^2.$$

Total variation [Rudin, Osher & Fatemi '92]

$$\|\nabla u\|_1$$

- Higher-order TV...

Original image

Horizontal gradient

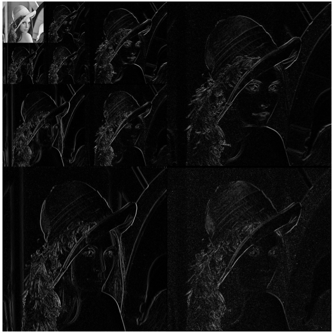
Vertical gradient

Wavelet decomposition [Morlet, Meyer, Mallat, Daubechies and many others]

Family of functions $\{\psi_{j,k} : j, k \in \mathbb{Z}\}$

$$\psi_{j,k}(\cdot) = 2^{j/2}\psi(2^j \cdot -k).$$

Original image

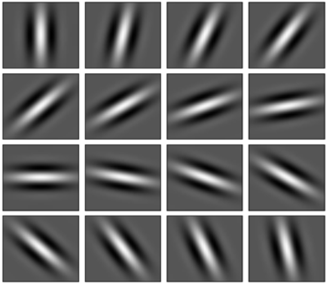


Wavelet coefficients

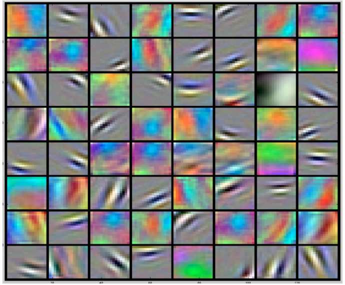
Dictionary [Mairal et al ’09]

Family of filters $\psi_j : j = \{1, ..., m\}$

$$\min_{\mathcal{D}, r} \frac{1}{2}\|u - \mathcal{D}r\|^2 + \mu\|r\|_0 + \iota_{\mathcal{C}}(\mathcal{D}).$$



Gabor filters



Redundant dictionary

Other examples

- ℓ_1-norm, $\ell_{1,2}$-norm, ℓ_0-pseudo norm, ℓ_p-norm, $p \in [0, 1]$
- Fourier transform, discrete cosine transform
- Curvelet, shearlet...
- Matrix: nuclear norm, rank function
- Constraints: non-negativity, simplex, box constraint
- Physics laws
- ...

How to use them...

Outline

Mathematical formulation

$$f = u + \varepsilon,$$

where

- u is the true image which is **piecewise constant** — total variation.
- ε is additive noise.

TV+L^p denoising [Chambolle & Pock '11]

$$\min_u \frac{1}{2}\|f - u\|_p^p + \mu\|\nabla u\|_1.$$

- μ provides the balance between fidelity and regularization.
- p depends on noise model:
  - Additive white Gaussian noise: $p = 2$;
  - Sparse noise (salt and pepper noise): $p = 1$.

Image denoising


Original image


Noised image


Denoised image

Mathematical formulation

$$f = \mathcal{F}u + \varepsilon,$$

where

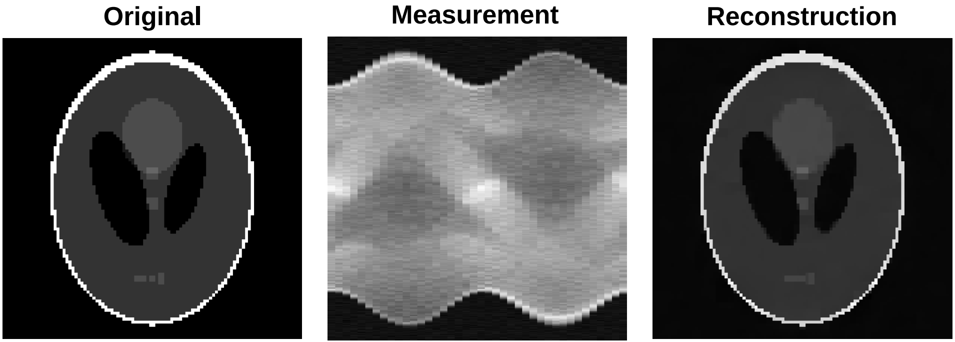
- u is the true image which is **piecewise constant** — total variation.
- $\mathcal{F}$ is partial Fourier transform.
- ε is additive noise.

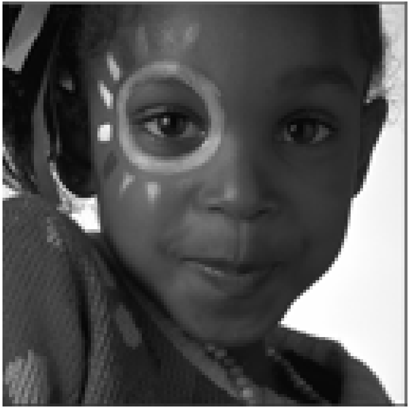
Wavelet frame MRI reconstruction

$$\min_u \ \frac{1}{2}\|f - \mathcal{F}u\|_p^p + \mu\|\mathcal{W}u\|_1.$$

- μ provides the balance between fidelity and regularization.

Medical imaging





Original image

Blurred image

Mathematical formulation

$$f = h \star u + \varepsilon,$$

where

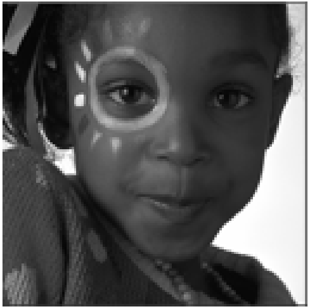
- u is the true image which is **piecewise constant** — total variation.
- h is the blur kernel which is **non-negative and smooth** and $\sum_i h_i = 1$ — Tikhonov regularization.
- ε is additive white Gaussian noise.

Blind image deblur [Driggs et al '20]

$$\min_{u,h} \frac{1}{2}\|f - h \star u\|^2 + \mu\|\nabla u\|_1 + \nu\|h\|^2 + \iota_{\mathcal{H}}(h).$$

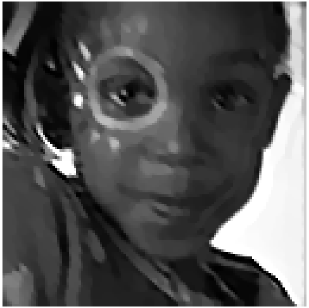
- The problem is **non-convex**.

Blind image deblur



Original image

Blurred image



Deblurred image

Optical flow constraint:

$$u_0(x) = u_1(x + v) + w.$$

Taylor expansion for v

$$u_0(x) \approx u_1(x + v_0) + \nabla u_1(x + v_0) \cdot (v - v_0) + w,$$

for some chosen v_0.

Denote

$$\rho(w, v) = u_0(x) - \big(u_1(x + v_0) + \nabla u_1(x + v_0) \cdot (v - v_0) + w\big).$$

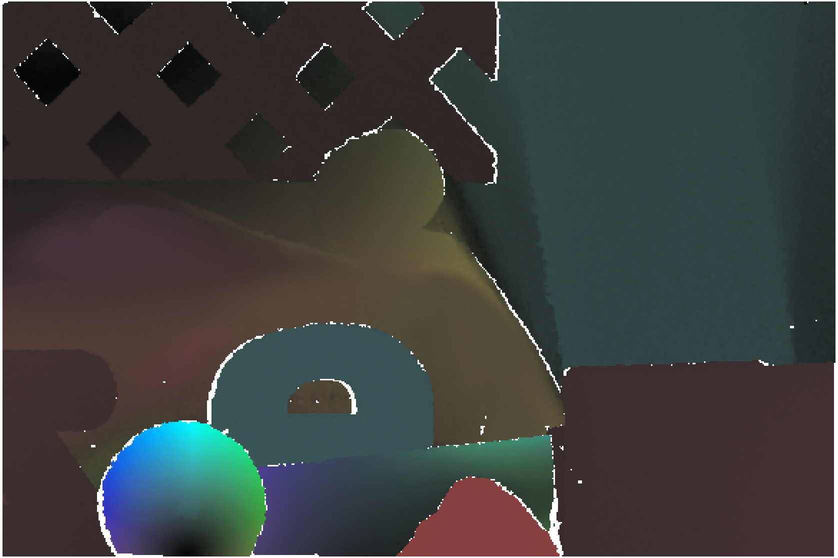
TV+L^1 Optical flow [Chambolle & Pock '11]

$$\min_{v,w} \|\rho(v, w)\|_1 + \mu\big(\|\nabla v\|_1 + \nu\|\nabla w\|_1\big).$$

Optical flow

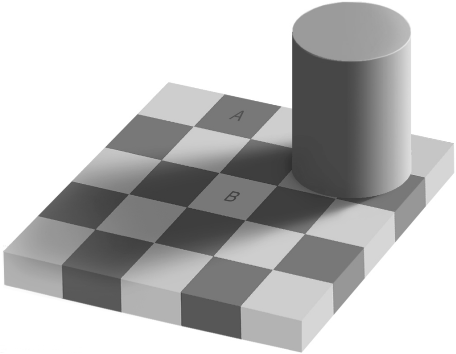
Optical flow

Optical flow

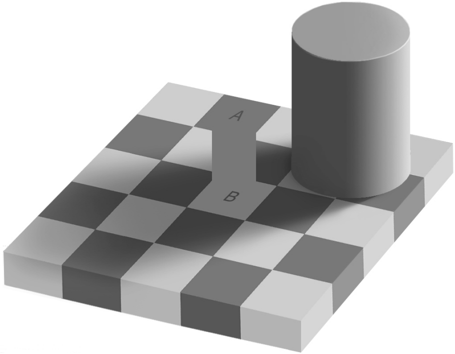


Optical flow

Optical flow



Checkboard illusion



Illusion free

Mathematical formulation

$$f = u + s + \varepsilon,$$

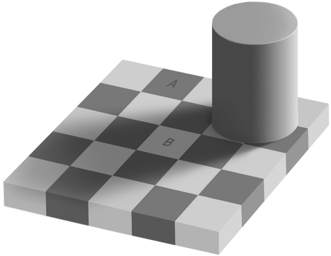
where

- u is the true image which is **piecewise constant** — total variation.
- s is the shadow which is **piecewise linear** — 2nd-order total variation.
- ε is additive white Gaussian noise.

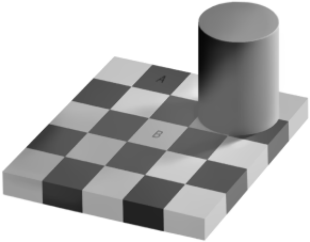
TV+L^1 Retinex [Liang & Zhang '15]

$$\min_{u,s} \frac{1}{2}\|f - u - s\|^2 + \mu\big(\|\nabla u\| + \nu\|\nabla^2 s\|\big).$$

- ν provides the balance between $\|\nabla u\|$ and $\|\nabla^2 s\|$.
- μ provides the balance between fidelity and regularization.

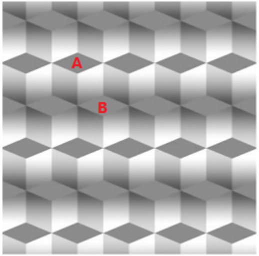


Checkboard

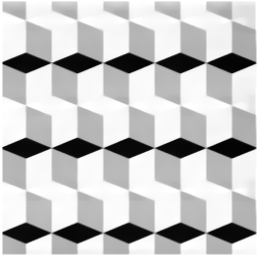


Estimated u

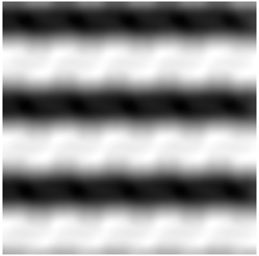
Estimated s



Illusion



Estimated u



Estimated s

Mathematical formulation

$$f = l + s + \varepsilon,$$

where

- l is the background which is **low rank** — nuclear norm.
- h is the foreground which is **sparse** — ℓ_1-norm.
- ε is additive white Gaussian noise.

Principal component pursuit [Candès et al '11]

$$\min_{l,s} \frac{1}{2}\|f - l - s\|^2 + \mu\|l\|_* + \nu\|s\|_1.$$

Video decomposition

Take-away messages

- Regularization is widely used in fields including inverse problems, imaging, statistics and other related areas.
- Mathematical models end up with certain optimization problem.
- Challenges of solving the optimization problem: non-smoothness, composition, dimensionality and non-convexity.
- First-order non-smooth optimization methods — to be introduced in the later stage of the course.

- Regularization is widely used in fields including inverse problems, imaging, statistics and other related areas.
- Mathematical models end up with certain optimization problem.
- Challenges of solving the optimization problem: non-smoothness, composition, dimensionality and non-convexity.
- First-order non-smooth optimization methods — to be introduced in the later stage of the course.

Thank you very much!

https://jliang993.github.io/