A Tale That Begins with L^1-Data-Fidelity Models

Yiqiu Dong

DTU Compute, Technical University of Denmark

Scientific Day in Memory of Prof. Mila Nikolova

Mila and Me

- When did I first hear about Mila?

Mila and Me

- When did I first hear about Mila?

Journal of Mathematical Imaging and Vision 20: 99–120, 2004

A Variational Approach to Remove Outliers and Impulse Noise

MILA NIKOLOVA
CMLA UMR 8536–ENS de Cachan, 61 av. du President Wilson, 94235 Cachan Cedex, France
nikolova@cmla.ens-cachan.fr

(647 citations!)

Mila and Me

- When did I first hear about Mila?
- When did I meet Mila the first time?

Mila and Me

- When did I first hear about Mila?
- When did I meet Mila the first time?

Hong Kong Baptist University
DEPARTMENT OF MATHEMATICS

- **Tuesday, 6th March, 2007**

Title: CMIV Lecture Series: Optimization for Image Processing (Lecture 3)
Speaker: Prof. Mila Nikolova, CMLA ENS de Cachan, France
Time/Place: 14:30 - 16:30
FSC 1217

Mila and Me

- When did I first hear about Mila?
- When did I meet Mila the first time?
- Inspired by Mila's work on nonsmooth data-fidelity models, I finished my PhD project on "impulse noise removal in image processing".

Nonsmooth Data-Fidelity Models

Consider the variational model

$$\mathcal{F}(u, z) = \Psi(u, z) + \alpha\Phi(u).$$

- $\Psi(u, z) = \sum_{i=1}^{q} \psi_i(a_i^T u - z_i)$ is a data-fidelity term, where $A \in \mathbb{R}^{q\times p}$ shows the forward system whose rows are a_i^T for $i = 1, \cdots, q$ and ψ is even, convex, nonsmooth and increasing on $\mathbb{R}_+$.
- $\Phi(u) = \sum_{i=1}^{r} \phi(\|G_i u\|)$ is a regularization term, where G_i are linear operators and ϕ is smooth on $\mathbb{R} \setminus \{0\}$, and increasing on $\mathbb{R}_+$.
- $\alpha > 0$ is the regularization parameter.
- The expected solution u minimises the regularized cost function $\mathcal{F}$.

Nonsmooth Data-Fidelity Models

Theorem. For $J \subset \{1, \cdots, q\}$, let $\{a_i, i \in J\}$ be linearly independent. Set

$$K_J = \{w \in \mathbb{R}^p : a_i^T w = 0, \forall i \in J\}.$$

Let $\hat{u}$ be such that $\mathcal{F}$ has a strict local minimum over $\hat{u} + K_J \cup K_J^{\perp}$. Then there is an open subset $O_J \subset \mathbb{R}^q$ and a continuous function $\mathcal{U} : O_J \to \mathbb{R}^p$ such that for any $z \in O_J$, $\mathcal{F}$ has a strict (local) minimum at $\hat{u} := \mathcal{U}(z)$ satisfying

$$\begin{aligned} a_i^T \hat{u} &= z_i, \quad \forall i \in J, \\ a_i^T \hat{u} &\neq z_i, \quad \forall i \in J^c. \end{aligned}$$

Crucial consequence:

$$P(a_i^T \hat{U} - Z = 0) = P(Z \in O_j) > 0.$$

[Ref.] Nikolova 2002; Nikolova 2004.

Impulse Noise

$$z_i = \begin{cases} \bar{u}_i, & \text{with probability } 1-r, \\ n_i, & \text{with probability } r, \end{cases}$$

- $\bar{u} \in \mathbb{R}^q$ denotes the clean image
- $z \in \mathbb{R}^q$ is the degraded image corrupted by impulse noise
- r represents the corruption rate by impulse noise
- Only part of pixels are corrupted
- Noise is independent of the image

Two Main Type

- *Salt-and-pepper noise:*

$$z_i = \begin{cases} \bar{u}_i, & \text{with probability } 1-r, \\ n_{\max}, & \text{with probability } \frac{r}{2}, \\ n_{\min}, & \text{with probability } \frac{r}{2}, \end{cases}$$

  where $n_{\max}$ and $n_{\min}$ are the maximum and minimum of the gray-level range.

- *Random-valued impulse noise:*

$$z_i = \begin{cases} \bar{u}_i, & \text{with probability } 1-r, \\ n_i, & \text{with probability } r, \end{cases}$$

  where n comes from a uniformly distributed random variable with values in $[n_{\min}, n_{\max}]$.

L^1-Data-Fidelity Model to Remove Impulse Noise

$$\mathcal{F}_z^1(u) = \sum_{i=1}^{q} |u_i - z_i| + \beta \sum_{i=1}^{q} \sum_{j \in \mathcal{N}_i} \phi(|u_i - u_j|)$$

where $\mathcal{N}_i$ is the neighborhood of the pixel i.

Corollary Let ϕ be convex. Then, the function $\mathcal{F}_z^1$ reaches its minimizer at $\hat{u}$ if, and only if

$$i \in J \Longrightarrow \left| \sum_{j \in \mathcal{N}_i} \phi'(z_i - \hat{u}_j) \right| \leq \frac{1}{\beta}$$

$$i \in J^c \Longrightarrow \sum_{j \in \mathcal{N}_i} \phi'(\hat{u}_i - \hat{u}_j) = \frac{\sigma_i}{\beta}, \qquad \sigma_i = \mathsf{sign}\left(\sum_{j \in \mathcal{N}_i} \phi'(z_i - \hat{u}_j) \right)$$

where $J = \{i : \hat{u}_i = z_i\}$.

Two-Phase Methods

- **Noise Detection.** Identify image pixels contaminated by noise by suitable noise detector. The possible noise candidate set is $\mathcal{I}$, and the noise-free candidate set is $\mathcal{U} = \mathcal{A} \setminus \mathcal{I}$
- **Noise Removal.** Based upon the information on the location of noise-free pixels, images are denoised by solving the variational model

$$\mathcal{F}_z^1|_{\mathcal{I}}(u) = \sum_{i \in \mathcal{I}} \left[|u_i - z_i| + \beta \sum_{j \in \mathcal{N}_i} \phi(|u_i - u_j|) \right]$$

[Ref.] Chan, Ho, Nikolova 2005; Chan, Hu, Nikolova 2004; Cai, Chan, Nikolova 2008; Cai, Chan, Nikolova 2010.

Noise Detector

- **Salt-and-Pepper Noise.** The adaptive median filter
- **Random-Valued Impulse Noise.** The adaptive center-weighted median (ACWM) filter

- The usage of noise detector is due to the fact that some part of the pixels in images corrupted by impulse noise and the rest are noise-free
- The capability of the two-phase method is mainly limited by the accuracy of the noise detector in the first phase.

More Two-Phase Methods

- **Noise Detection.** The statistic "Rank-Ordered Logarithmic Difference" was proposed to detect random-valued impulse noise, especially for high noise level. It is defined as

$$ROLD_m(z_i) = \sum_{k=1}^{m} R_k(z_i)$$

where R_k denotes the $k$the smallest value in $\{D_j : j \in \mathcal{N}_i\}$ with $D_j(z_i) = 1 + \max\{\log_2 |z_j - z_i|, -5\}/5$. The possible noise candidate set is $\mathcal{I}$, and the noise-free candidate set is $\mathcal{U} = \mathcal{A} \setminus \mathcal{I}$

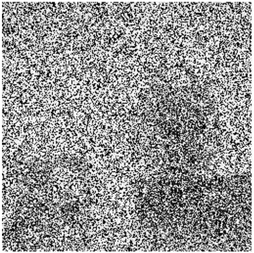
- **Image Restoration.** Based upon the information on the location of noise-free pixels, images are deblurred and denoised simultaneously by solving the variational model

$$\mathcal{F}_z^2(u) = \sum_{i \in \mathcal{U}} |a_i^T u - z_i| + \beta \sum_{i \in \mathcal{A}} \sum_{j \in \mathcal{N}_i} \phi(|u_i - u_j|)$$

[Ref.] Chan, Dong, Xu 2007; Chan, Dong, Hintermüller 2010.

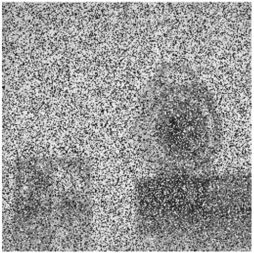
Example

Original image



70% SP Noise

Restored result



60% RV Noise

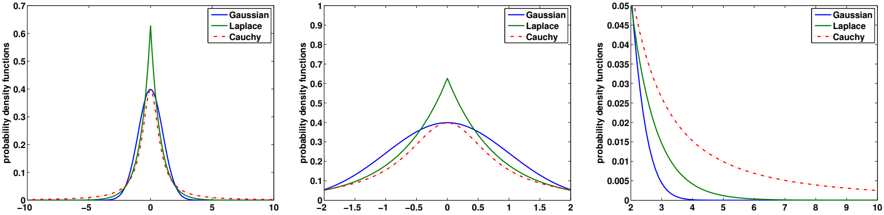
Restored result

Cauchy Noise

- Cauchy noise often arises in echo of radar, low-frequency atmospheric noises and underwater acoustic signals
- Cauchy noise is additive, bell-shaped, and very impulsive
- Cauchy distribution has probability density function:

$$p_\gamma(x) = \frac{\gamma}{\pi((x-\mu)^2 + \gamma^2)},$$

where μ is the location parameter, and $\gamma > 0$ is the scale parameter



Variational Model

According to Bayes' Law and using the maximum a posteriori (MAP) estimator, we obtain a variational model for Cauchy noise removal

$$\min_{u \in BV(\Omega)} \frac{\lambda}{2} \int_\Omega \log\left(\gamma^2 + (Ku - f)^2\right)\, dx + \int_\Omega |Du|.$$

- $BV(\Omega)$ denotes the space of functions of bounded variation, i.e., $BV(\Omega) = \{u \in L^1(\Omega) : \int_\Omega |Du| < +\infty\}$
- $\int_\Omega |Du| = \sup\left\{ \int_\Omega u \,\mathrm{div}\vec{v} dx : \vec{v} \in (C_0^\infty(\Omega))^2, \|\vec{v}\|_\infty \leq 1\right\}$
- $\lambda > 0$ is the regularization parameter
- The model is NON-CONVEX

[Ref.] Sciacchitano, Dong, Zeng 2015; Mei, Dong, Huang, Yin 2018; Laus, Pierre, Steidl 2018.

Properties for Nonconvex Model

$$\min_{u \in BV(\Omega)} \frac{\lambda}{2} \int_\Omega \log\left(\gamma^2 + (Ku - f)^2\right) \, dx + \int_\Omega |Du|$$

- **(Existence)** Assume $f \in L^2(\Omega)$, $K \in \mathcal{L}(L^1(\Omega), L^2(\Omega))$, and $K\mathbf{1} \neq 0$; then the model has at least one solution in $BV(\Omega)$. In the denoising case with $f \in L^\infty(\Omega)$, the solutions satisfy $\inf_\Omega f \leq u \leq \sup_\Omega f$.
- **(Uniqueness)** Assume that $f \in L^2(\Omega)$ and K is injective. Then, the model has a unique solution u^* in $\Omega_U := \{u \in BV(\Omega) : f(x) - \gamma < (Ku)(x) < f(x) + \gamma$ for all $x \in \Omega\}$.
- **(Minimum-maximum principle)** Let $f_1, f_2 \in L^\infty(\Omega)$ and $\gamma \geq 1$ with $a_1 = \inf_\Omega f_1$, $a_2 = \inf_\Omega f_2$ and $b_2 = \sup_\Omega f_2$. Assume that $f_1 < f_2$. Then, denoting with u_1 (resp., u_2) a solution of the denoising problem for $f = f_1$ (resp., $f = f_2$), we have $u_1 \leq u_2$, if $(b_2 - a_1)(b_2 - a_2) < 1$.

Convex Variational Model

$$\min_{u\in BV(\Omega)} \frac{\lambda}{2}\left(\int_\Omega \log\left(\gamma^2+(Ku-f)^2\right)dx+\mu\|Ku-u_0\|_2^2\right)+\int_\Omega |Du|$$

- $\lambda > 0$ and $\mu > 0$ are regularization parameters.
- u_0 is the median filter result.
- If $8\mu\gamma^2 \geq 1$ and K is injective, the model is strictly convex.
- **(Existence & Uniqueness)** Assume $f \in L^2(\Omega)$, $K \in \mathcal{L}(L^2(\Omega))$, and $K\mathbf{1} \neq 0$; then the model admits a solution. If $8\mu\gamma^2 \geq 1$ and K is injective, then the solution is unique.
- In the denoising case with $f \in L^\infty(\Omega)$ and $8\mu\gamma^2 \geq 1$, the unique solutions satisfy

$$\min\{\inf_\Omega f, \inf_\Omega u_0\} \leq u \leq \max\{\sup_\Omega f, \sup_\Omega u_0\}.$$

Example

(PSNR=25.68 dB)

(PSNR=26.69 dB)

(PSNR=26.74 dB)

(PSNR=27.14 dB)

Noisy images — Convex model — Nonconvex model

Thank you!