Medical Image Analysis and Processing

Image Noise Filtering

Total Variation

Emad Fatemizadeh

Distance/online Course: Session 10

Date: 16 March 2021, 26th Esfand 1399

Contents

- Total Variation Definition
- Calculus of variation (Variational Calculus)
- Constrained Optimization
- Total Variation Denoising (ROF Approach)
- Numerical Implementation
- Bayesian Interpretation
- One Dimensional Discrete TV
- Two Dimensional TV

Definition

› One dimensional function:

$$TV(f) = \int_a^b |f'(x)|dx \, , f \in BV(a, b)$$

› Bounded Variation (BV) functions:

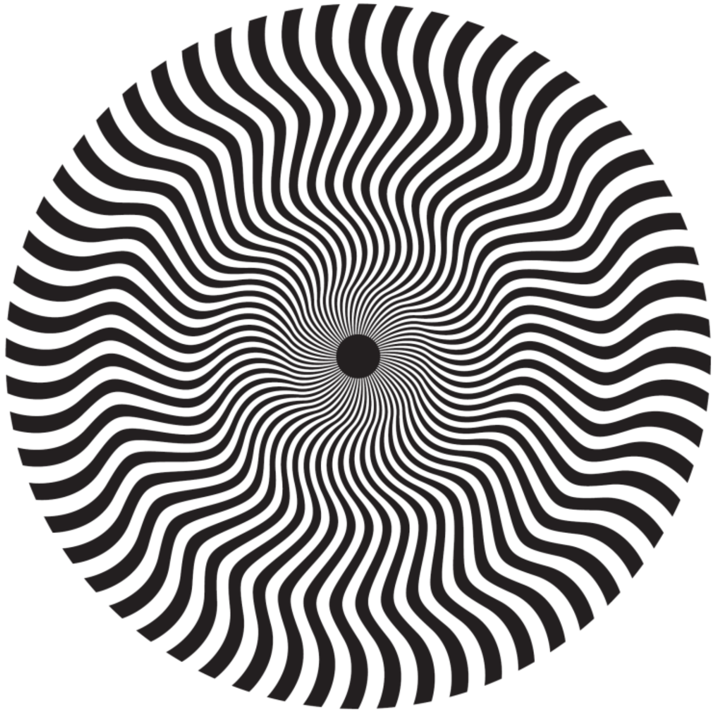
A real-valued function whose total variation is bounded (finite)

› Multivariable function:

$$TV(f) = \int_\Omega |\nabla f(x)|dx \, , \qquad \Omega \subset \mathbb{R}^n, f \in BV(\Omega)$$

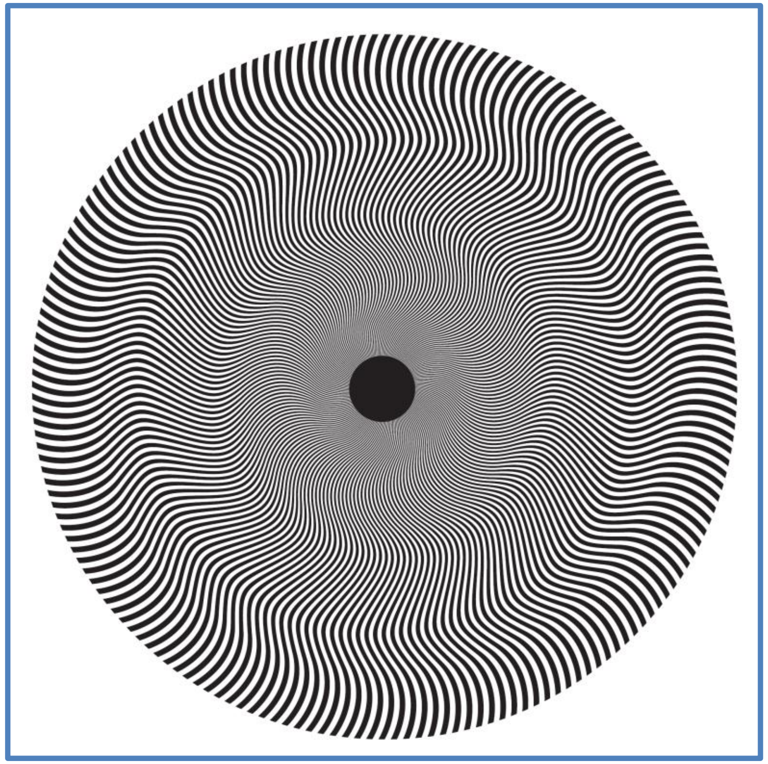
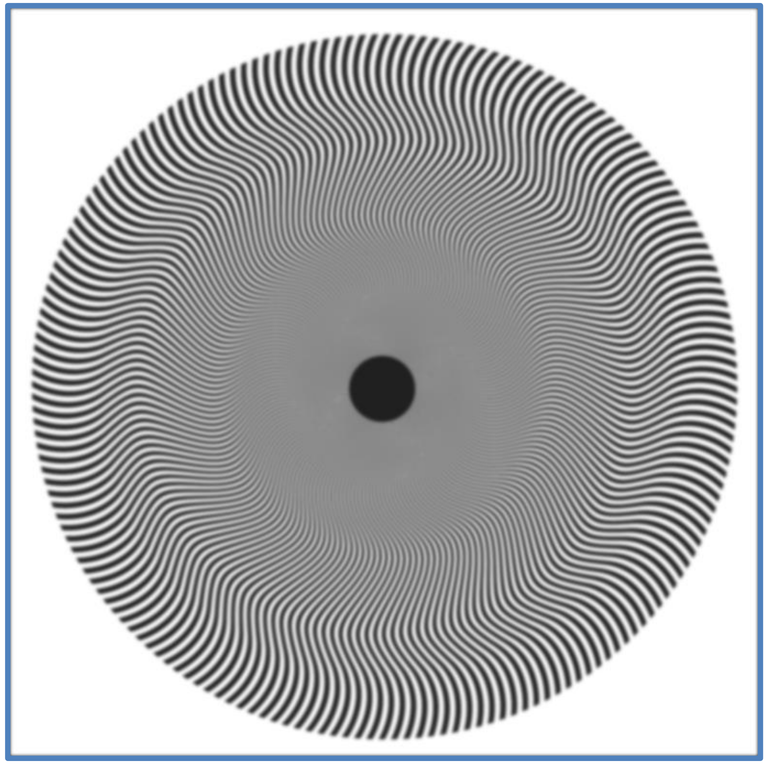
Example

- Left (0.6145), Right (0.2224), normalized to # of pixels



Example:

› Left: $\{u(x_1, x_2), \text{TV}=0.6145\}$, $\{u(x_1, x_2) * G_4(x_1, x_2), \text{TV}=0.2294\}$

Calculus of Variation

› Consider the following functional (energy function):

$$E(u) = \int_a^b L(u, u_x) dx$$

› where $u(a)$and $u(b)$ are known and $u(x) \in [a, b]$ is unknown.

› L: known as Lagrangian

› Our goal: finding extrema of $E(u)$

› Solve by *Euler-Lagrange* equation:

$$\frac{\partial L}{\partial u} - \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x} = 0$$

Calculus of Variation

- For a functional of the form:

$$E(u) = \int_a^b L(u, u_x, u_{xx})dx \Rightarrow \frac{\partial L}{\partial u} - \frac{\partial}{\partial x}\frac{\partial L}{\partial u_x} + \frac{\partial^2}{\partial x^2}\frac{\partial L}{\partial u_{xx}} = 0$$

- The derivation for the 2D problem is completely analogous:

$$E(u) = \iint_\Omega L(u, u_{x_1}, u_{x_2}, u_{x_1x_1}, u_{x_2x_2}, u_{x_1x_2})dx_1x_2$$

$$\frac{\partial L}{\partial u} - \left(\frac{\partial}{\partial x_1}\frac{\partial L}{\partial u_{x_1}} + \frac{\partial}{\partial x_2}\frac{\partial L}{\partial u_{x_2}}\right) + \left(\frac{\partial^2}{\partial x_1^2}\frac{\partial L}{\partial u_{x_1x_1}} + \frac{\partial^2}{\partial x_2^2}\frac{\partial L}{\partial u_{x_2x_2}} + \frac{\partial^2}{\partial x_1 \partial x_2}\frac{\partial L}{\partial u_{x_1x_2}}\right) = 0$$

Total Variation Denoising – ROF Approach

- Constrained optimization problem:

$$\min_{x} f(x), \qquad s.t.\ g(x) = 0$$

- Using Lagrange multiplier, solve unconstrained problem:

$$\min_{x} f(x) + \lambda g(x) \Rightarrow x = L(\lambda)$$

- Solve for λ using

$$g(x) = g\big(L(\lambda)\big) = 0 \Rightarrow \lambda = \lambda_0 \Rightarrow x = L(\lambda_0)$$

Total Variation Denoising – ROF Approach

- ROF (Rudin-Osher-Fatemi) Formulation:
- Image model:

$$v(x_1, x_2) = u(x_1, x_2) + \eta(x_1, x_2), \qquad (x_1, x_2) \subset \Omega$$

- $v(x_1, x_2)$: noisy observation
- $u(x_1, x_2)$: clean image
- $\eta(x_1, x_2)$: "zero-mean", "known variance" *i.i.d* additive noise
- *i.i.d*: independent and identically distribution

Total Variation Denoising – ROF Approach

- Motivation: Estimate clean image as smooth as possible while satisfying constraints.
- Cost function:

$$\min_{u \subset BV(\Omega)} TV(u) = \int_{\Omega} |\nabla u|$$

- Subject to:

$$\int_{\Omega} u = \int_{\Omega} v\,, \int_{\Omega} (u - v)^2 = \sigma^2 |\Omega|$$

Total Variation Denoising – ROF Approach

- Using Lagrange multiplier:

$$E(u) = \min_u \int_\Omega |\nabla u| + \frac{\lambda}{2}\left\{\int_\Omega (u-v)^2 - \sigma^2|\Omega|\right\} = \min_u \int_\Omega \left(|\nabla u| + \frac{\lambda}{2}(u-v)^2\right) - \frac{\lambda}{2}\sigma^2|\Omega|$$

- It can be shown:

$$\begin{cases} div\left(\frac{\nabla u}{|\nabla u|}\right) + \lambda(u-v) = 0 & u \in \Omega \\ \frac{\partial u}{\partial \vec{N}} = \left\langle \frac{\nabla u}{|\nabla u|} \middle| \vec{N} \right\rangle = 0 & u \in \partial\Omega \end{cases}$$

Total Variation Denoising – ROF Approach

- $div\left(\frac{\nabla u}{|\nabla u|}\right)$ is undefined and non-differentiable for $|\nabla u|$=0 (flat area), which comes from singularity of TV at zero gradients, to avoid this, $|\nabla u|$ is replaced with smooth approximation:

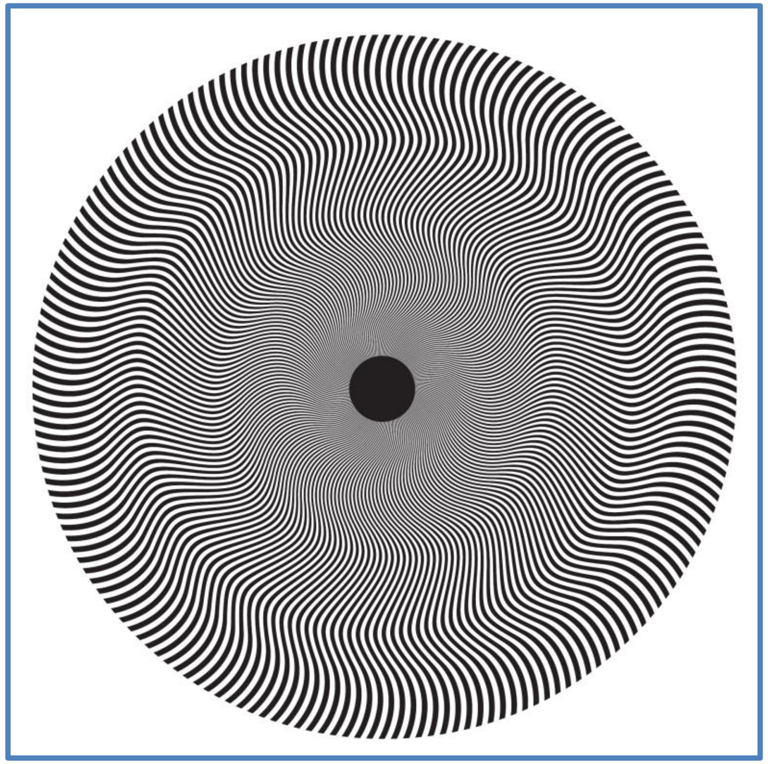
- $|\nabla u|_\beta = \sqrt{\beta^2 + |\nabla u|^2}$

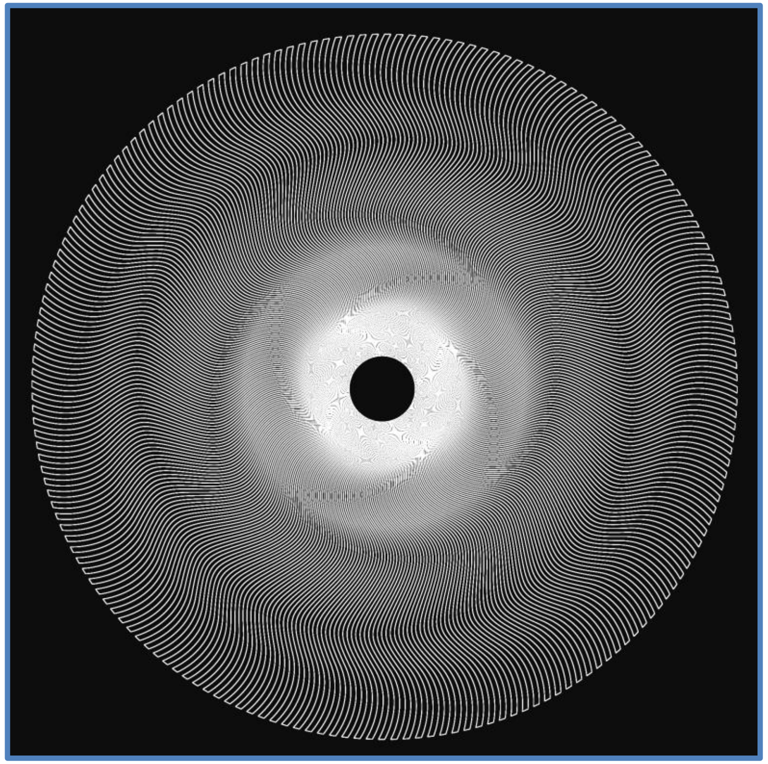
- $|\nabla u|_\alpha = |\nabla u|^\alpha \quad 1 < \alpha \leq 2$

- $|\nabla u|_\varepsilon = \begin{cases} \frac{|\nabla u|^2}{2\epsilon} + \frac{\varepsilon}{2} & |\nabla u| < \epsilon \\ |\nabla u| & |\nabla u| \geq \epsilon \end{cases}$

Total Variation - Example

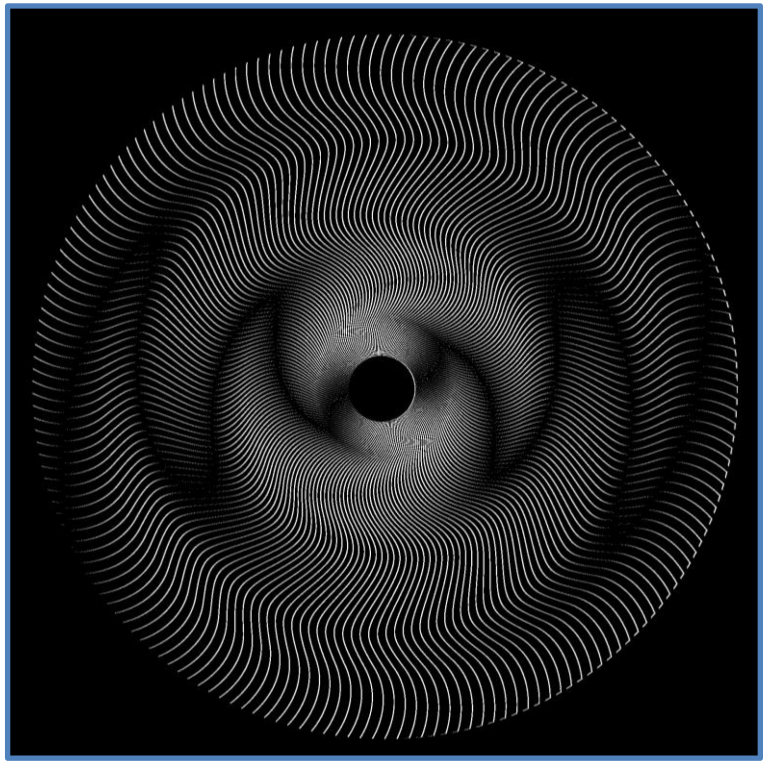
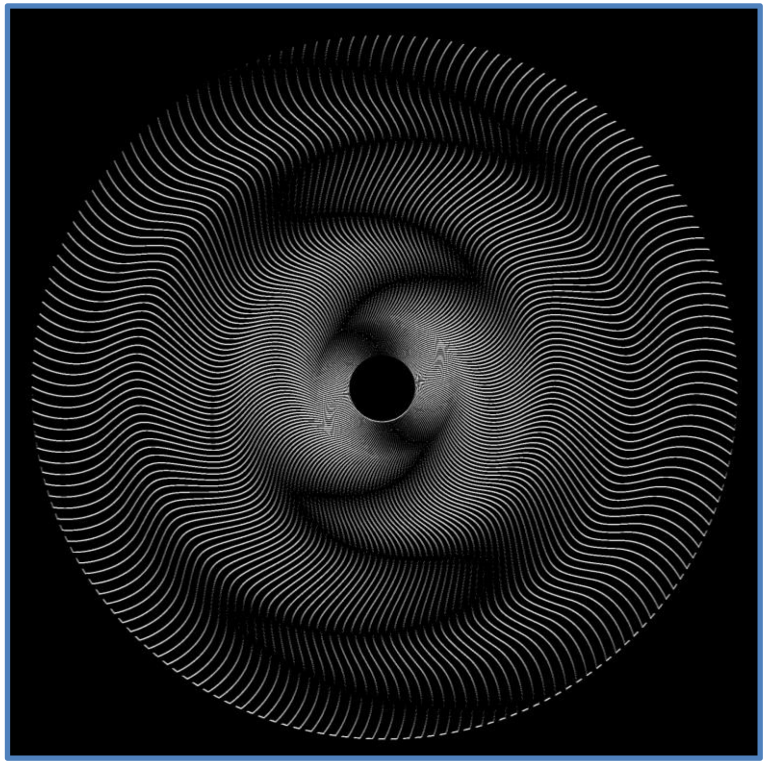
› Example: Clean Image (left), $|\nabla u|_\varepsilon$ (Right, ε=0.1)





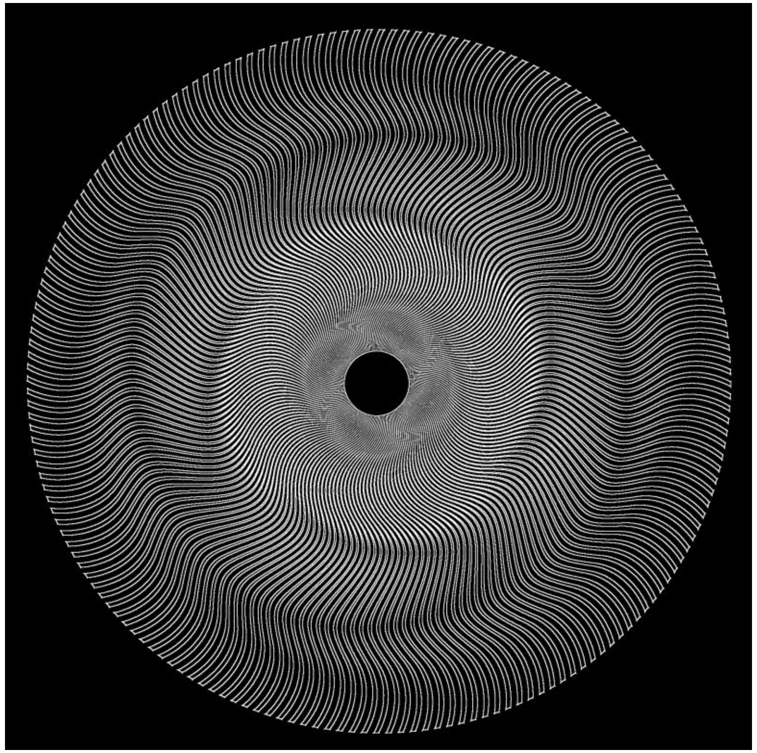
Total Variation - Example

› Example: Components of $\frac{\nabla u}{|\nabla u|_\varepsilon}$

Total Variation - Example

- Example: $div\left(\frac{\nabla u}{|\nabla u|_{\varepsilon}}\right)$



Total Variation Denoising – ROF Approach

› Most used approximation is:

$$\min_{u} \int_{\Omega} \left(\sqrt{\beta^2 + |\nabla u|^2} + \frac{\lambda}{2}(u - v)^2 \right) - \frac{\lambda}{2}\sigma^2|\Omega|$$

› or

$$\begin{cases} div\left(\dfrac{\nabla u}{\sqrt{\beta^2 + |\nabla u|^2}}\right) + \lambda(u - v) = 0 & u \in \Omega \\ \dfrac{\partial u}{\partial \vec{N}} = \left\langle \dfrac{\nabla u}{|\nabla u|} \middle| \vec{N} \right\rangle = 0 & u \in \partial\Omega \end{cases}$$

Total Variation Denoising – ROF Approach

- Using *time-marching* method:
- $u(x_1, x_2) \to u(x_1, x_2, t)$

$$\begin{cases} \dfrac{\partial u}{\partial t} = div\left(\dfrac{\nabla u}{\sqrt{\beta^2 + |\nabla u|^2}}\right) + \lambda(u - v) & u \in \Omega \\ \dfrac{\partial u}{\partial \vec{N}} = 0 & u \in \partial\Omega \\ u(x_1, x_2, t = 0) = u_0(x_1, x_2) & t = 0 \end{cases}$$

Total Variation Denoising – ROF Approach

- How to estimate λ:
  - After steady state has been reached, multiplies the PDE by $(u-v)$ and integrates by parts over Ω:

$$\int_\Omega \left(div\left(\frac{\nabla u}{\sqrt{\beta^2+|\nabla u|^2}}\right)(u-v)+\lambda(u-v)^2\right)=0$$

$$\lambda = \frac{1}{\sigma^2|\Omega|}\int_\Omega div\left(\frac{\nabla u}{\sqrt{\beta^2+|\nabla u|^2}}\right)(v-u)$$

- Iteratively find λ

Total Variation Denoising – ROF Approach

› Algorithm:

1. $\lambda \leftarrow \lambda_0, u(x_1, x_2) \leftarrow u_0(x_1, x_2)$
2. Solve time-marching PDE for steady state solution (Slide 17)
3. Estimate Lagrange multiplier, λ, (Slide 18)
4. Stop if small changes in λ, else continue from step #2

Numerical Implementation

› Total variation discretization:

$$\|u\|_{\mathrm{TV}(\Omega)} \approx \sum_{i,j} \sqrt{(\nabla_x u)^2_{i,j} + (\nabla_y u)^2_{i,j}},$$

$$\sum_{i,j} \sqrt{\epsilon^2 + (\nabla_x u)^2_{i,j} + (\nabla_y u)^2_{i,j}}.$$

Numerical Implementation

› ROF (Rudin-Osher-Fatemi) Formulation:

› There are several ways to derivatives discretization:

› Let $\nabla_x^+, \nabla_x^-, \nabla_y^+, \nabla_y^-$ denote forward and backward finite difference

- One-sided difference $\quad (\nabla_x u)^2 = (\nabla_x^+ u)^2$
- Central difference $\quad (\nabla_x u)^2 = \left((\nabla_x^+ u + \nabla_x^- u)/2\right)^2$
- Geometric average $\quad (\nabla_x u)^2 = \left((\nabla_x^+ u)^2 + (\nabla_x^- u)^2\right)/2$
- Minmod $\quad (\nabla_x u)^2 = m(\nabla_x^+ u, \nabla_x^- u)^2$
- Upwind discretization [39] $\quad (\nabla_x u)^2 = \left(\max(\nabla_x^+ u, 0)^2 + \max(\nabla_x^- u, 0)^2\right)/2$

$$m(a,b) = \left(\frac{\operatorname{sign} a + \operatorname{sign} b}{2}\right) \min(|a|, |b|)$$

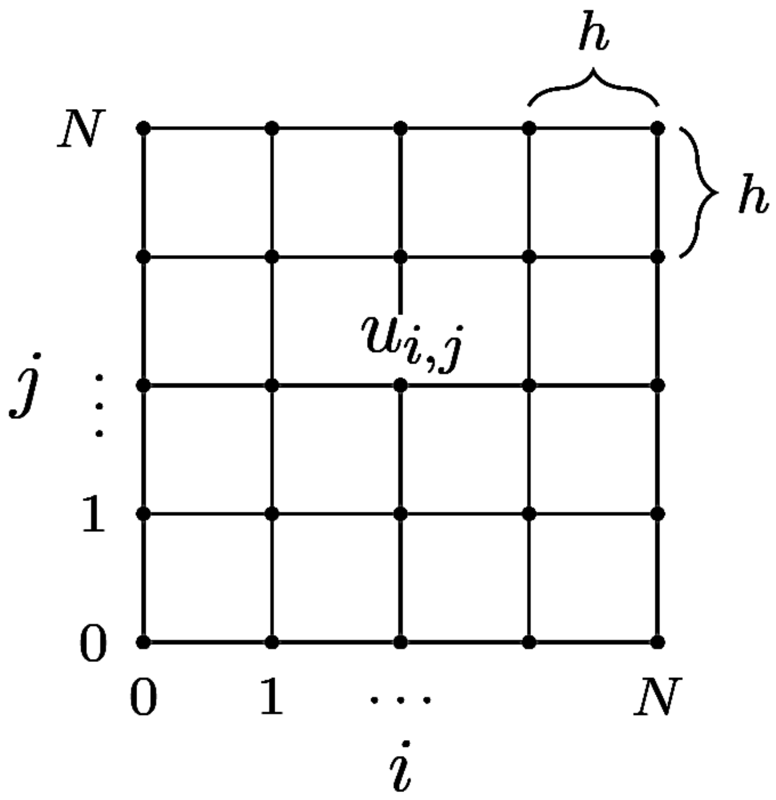
Numerical Implementation

- ROF (Rudin-Osher-Fatemi) Formulation:
- Central differences are undesirable for TV discretization because they miss thin structures, The central differences at (i, j) does not depend on $u_{i,j}$

$$\frac{\nabla_x^+ u_{i,j} + \nabla_x^- u_{i,j}}{2} = \frac{(u_{i+1,j} - u_{i,j}) + (u_{i,j} - u_{i-1,j})}{2} = \frac{u_{i+1,j} - u_{i-1,j}}{2}.$$

Numerical Implementation

- ROF (Rudin-Osher-Fatemi) Formulation:
- Sampling grid:



Numerical Implementation

- ROF (Rudin-Osher-Fatemi) Formulation:
- Time-Marching equation:

$$u_{i,j}^{n+1} = u_{i,j}^n + dt\Big[\nabla_x^-\Big(\frac{\nabla_x^+ u_{i,j}^n}{\sqrt{(\nabla_x^+ u_{i,j}^n)^2 + (m(\nabla_y^+ u_{i,j}^n, \nabla_y^- u_{i,j}^n))^2}}\Big)$$
$$+ \nabla_y^-\Big(\frac{\nabla_y^+ u_{i,j}^n}{\sqrt{(\nabla_y^+ u_{i,j}^n)^2 + (m(\nabla_x^+ u_{i,j}^n, \nabla_x^- u_{i,j}^n))^2}}\Big)\Big] + dt\lambda(f_{i,j} - u_{i,j}^n), \quad i,j = 1,\ldots,N-1,$$

- Boundary Conditions:

$$u_{0,j}^n = u_{1,j}^n, \quad u_{N,j}^n = u_{N-1,j}^n, \quad u_{i,0}^n = u_{i,1}^n, \quad u_{i,N}^n = u_{i,N-1}^n, \quad i,j = 0,\ldots,N,$$

- Initial Condition: $f(i,j)$

The End

› AnY QuEsTiOn?