Optimization Models for Inverse Problems in Image Processing

Alireza Hosseini

University of Tehran

April 2021

Outline

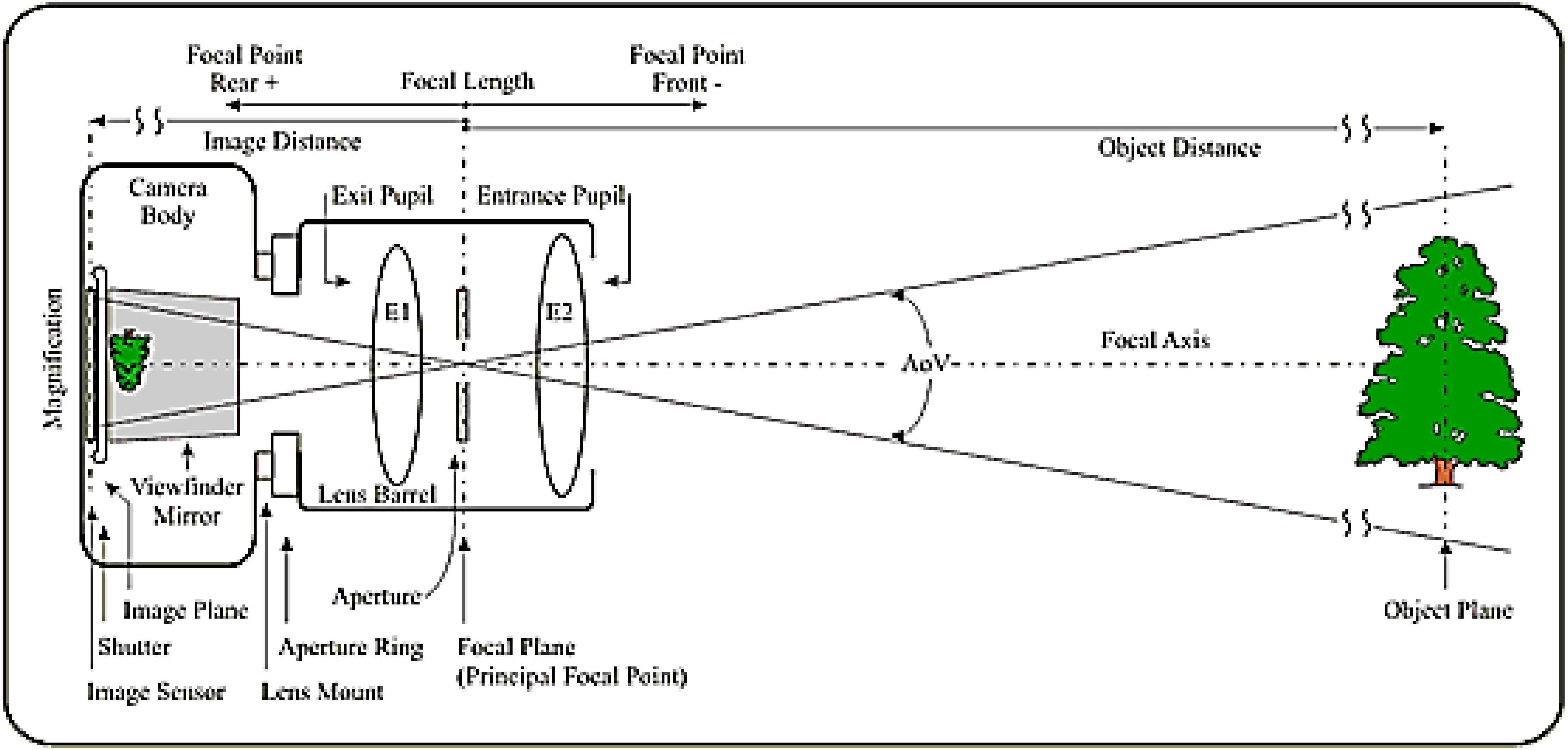
- What are images?
- Compressed sensing MRI.
- Distributional derivatives.
- Variational optimization models for image problems.
- Discretization of variational models.
- Primal-dual algorithm to solve MRI restoration problems.
- About SOAL Optimizer Competition 2021.

What are images?
(Bredies et al. 2018)

What are images; images can be produced in many different ways

- **Photography:** Photography produces two dimensional images by projecting a scene of the real world through some optics onto a two-dimensional image plane. The optics are focused onto some plane, called the focal plane, and objects appear more blurred the farther they are from the focal plane. Hence, photos usually have both sharp and blurred regions.

Photography

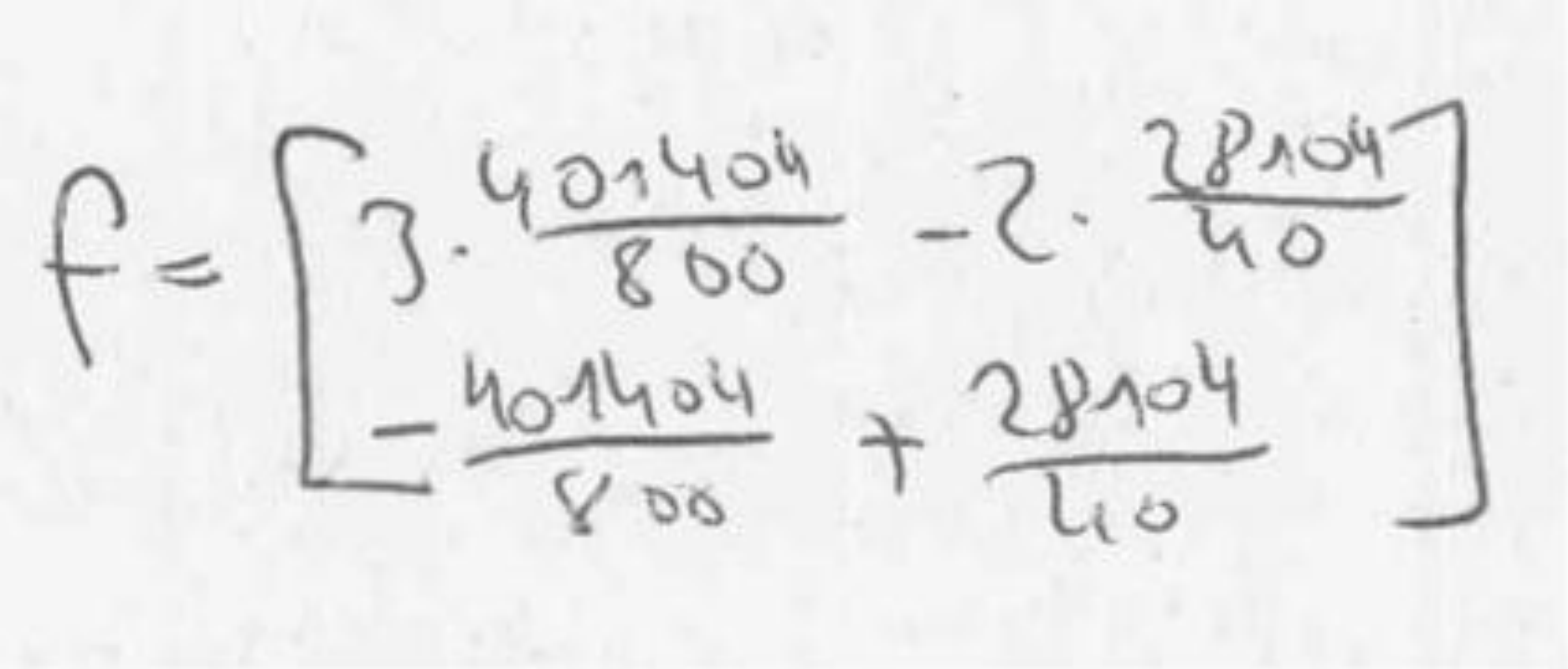


Scans

To digitize photos one may use a scanner. The scanner illuminates the

photo row by row and measures the brightness or color along the lines. Usually this does not result in some additional blur. However, a scanner operates at some resolution, which results in a reduction of information. Moreover, the scanning process may result in some additional artifacts. Older scans are often pale and may contain some contamination. The correction of such errors is an important problem in image processing.

Scans: An example



Mathematical definition of an image

Ω: Image domain, F: Color space;

$$u: \Omega \rightarrow F \text{ (Image).}$$

Discrete and Continuous Images:

Discrete d-dimensional image:

$$\Omega = \{1, \cdots, N_1\} \times \cdots \times \{1, \cdots, N_d\}.$$

Continuous d-dimensional image:

$\Omega \subset \Re^d$ or specifically

$$\Omega = [0, a_1] \times \cdots \times [0, a_d].$$

Different kinds of color spaces

Black-and-white (binary) images: $F = \{0,1\}$.

Grayscale images with discrete color space with k-bit depth:

$$F = \{0, \cdots, 2^k - 1\}.$$

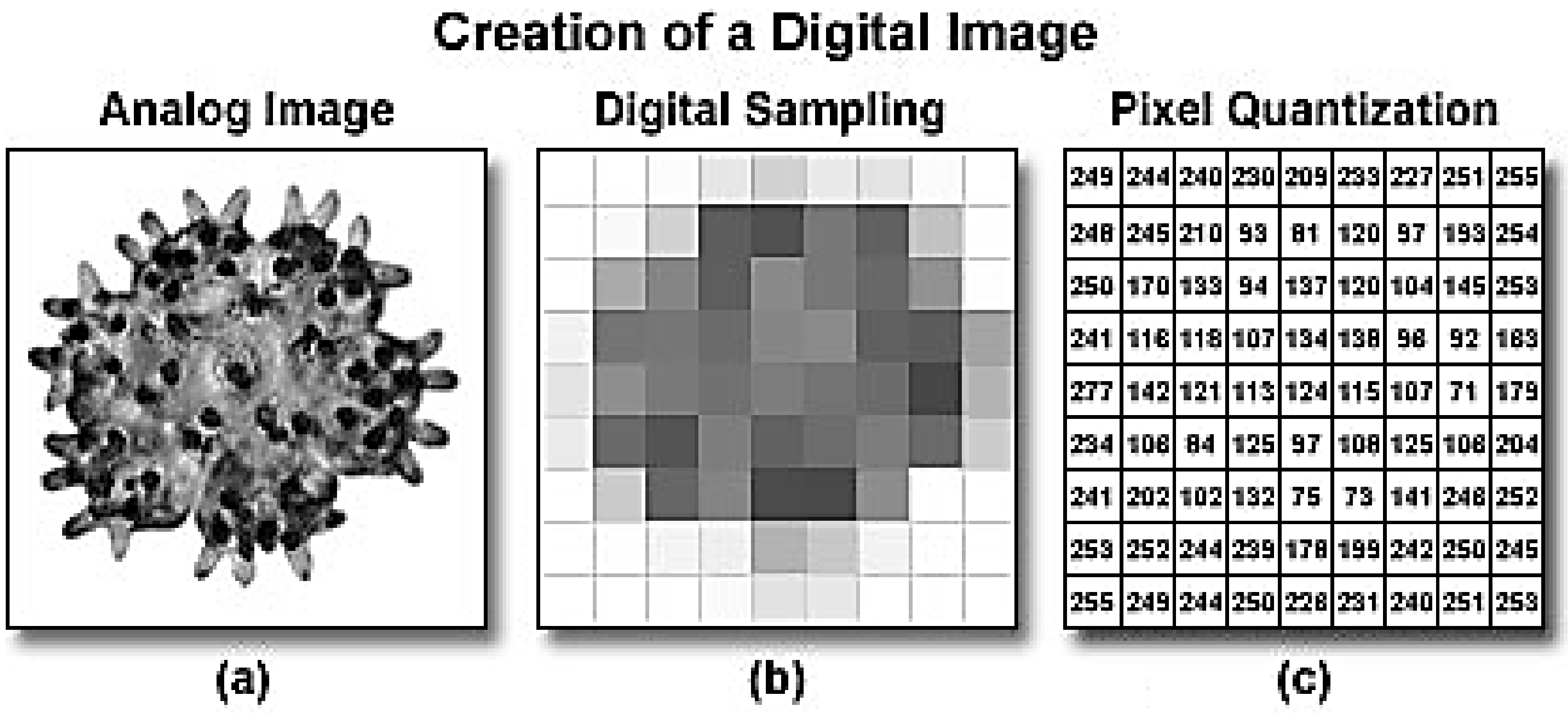
Color images with k-bit depth for each of 3 color channels:

$$F = \{0, \cdots, 2^k - 1\}^3.$$

Images with continuous gray values: $F = [0,1]$ or $F = \mathfrak{R}$.

Images with continuous colors: $F = [0,1]^3$ or $F = \mathfrak{R}^3$.

A discrete image

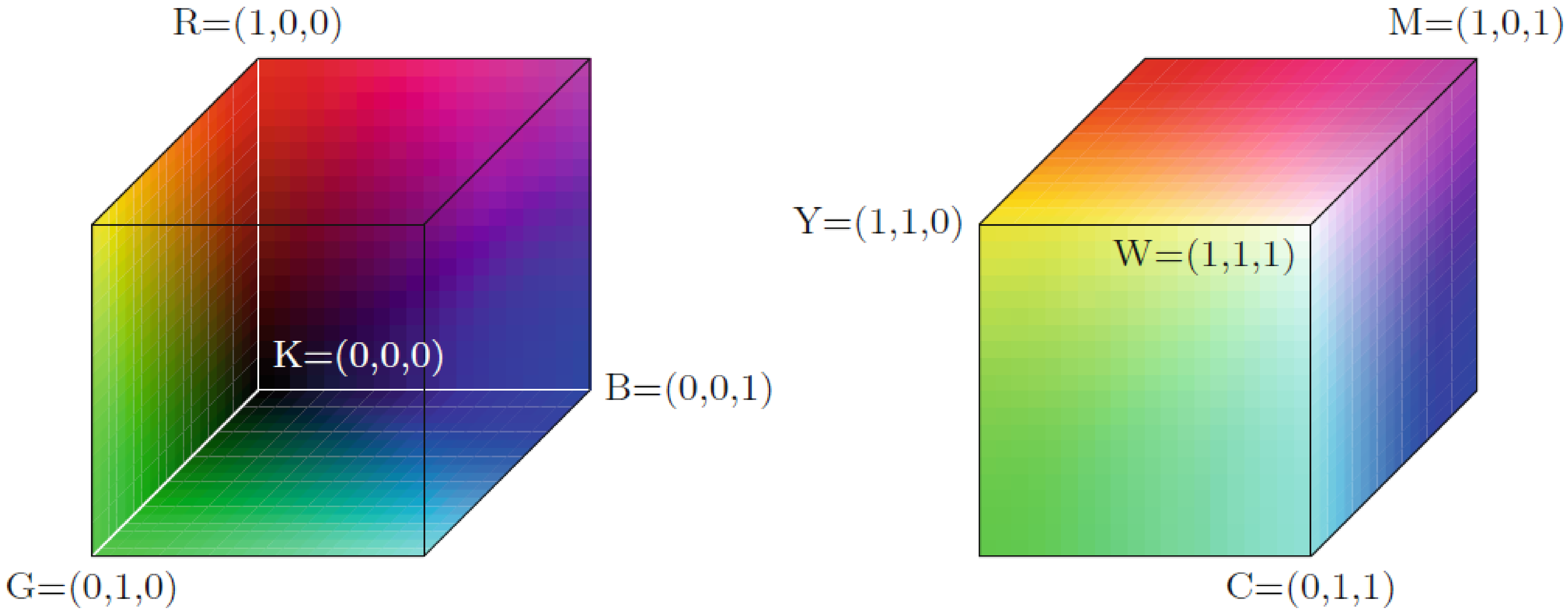


Encoding colors in different color channels

- How to measure distances in the color space?
- The perception of color is very complex and also subjective.
- Color Channels:

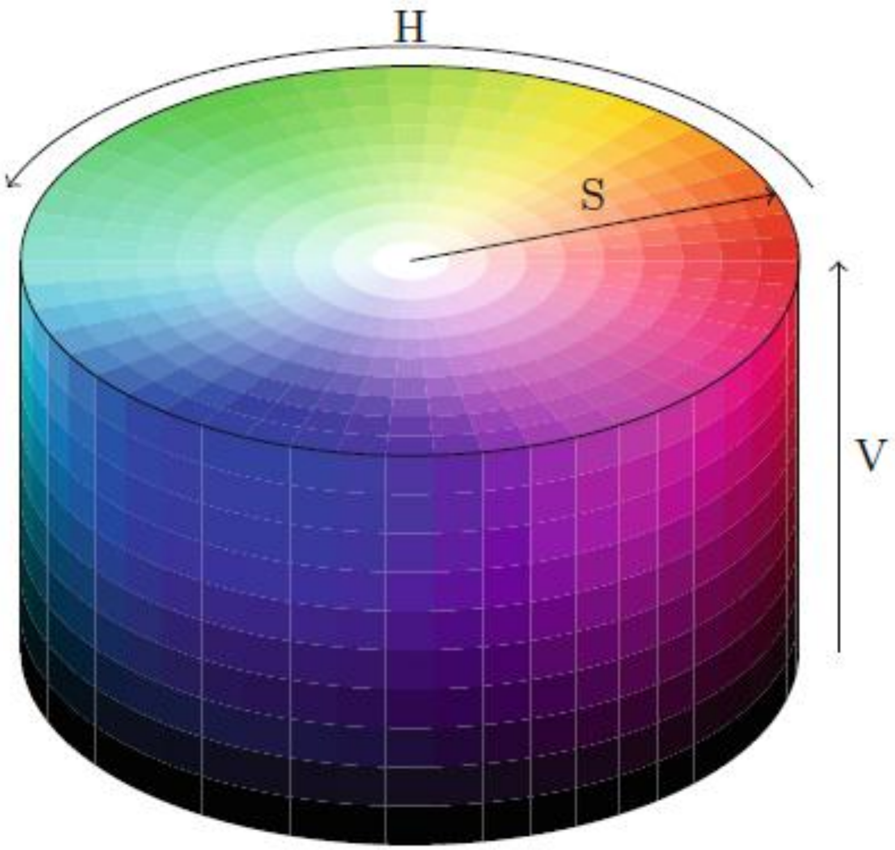
RGB space: $(R, G, B) \in [0,1]^3$

HSV space: $(H, S, V) \in [0,360) \times [0,100] \times [0,100]$.



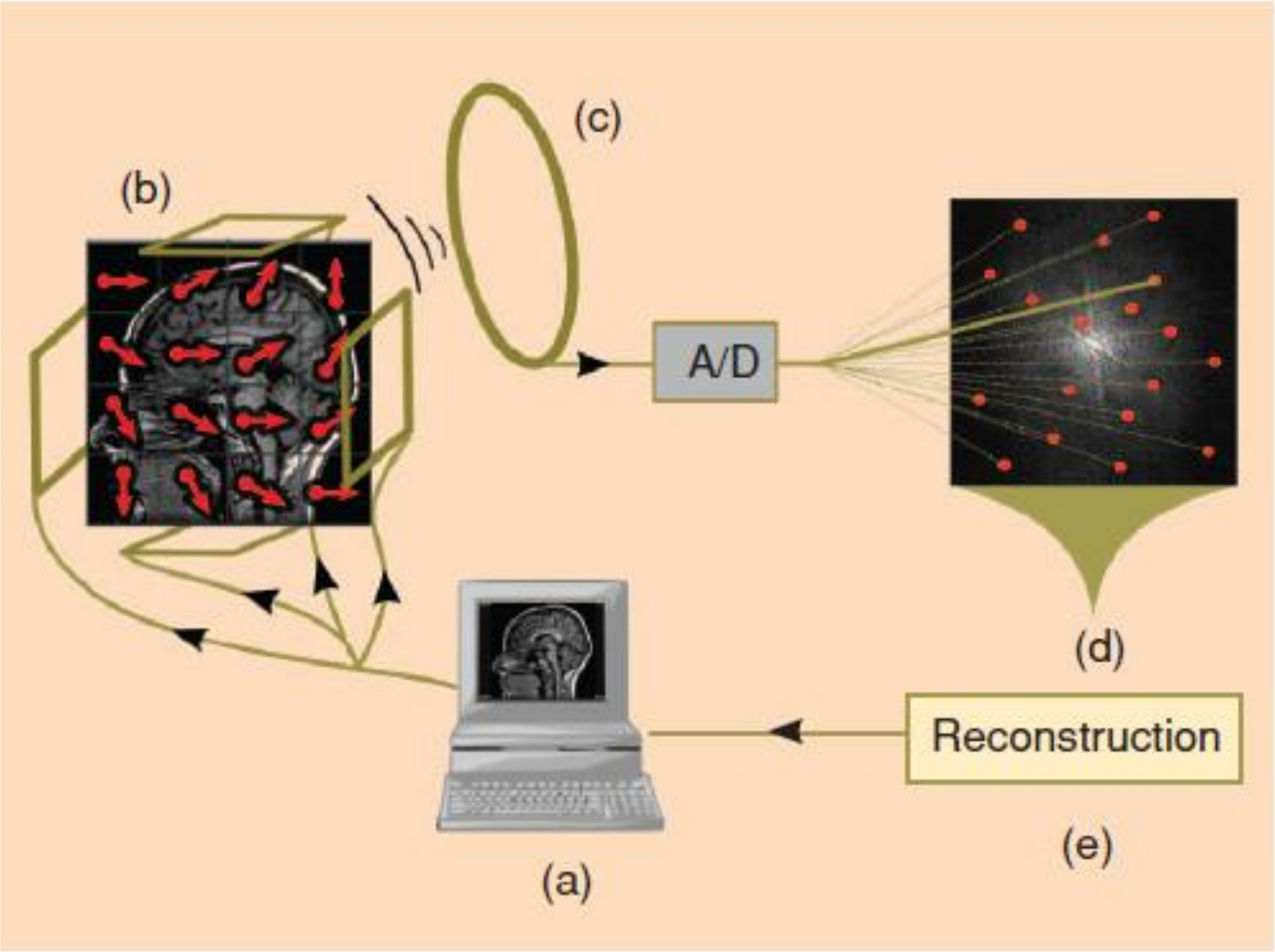
RGB space, visualized as a cube

HSV space, visualized as a cylinder

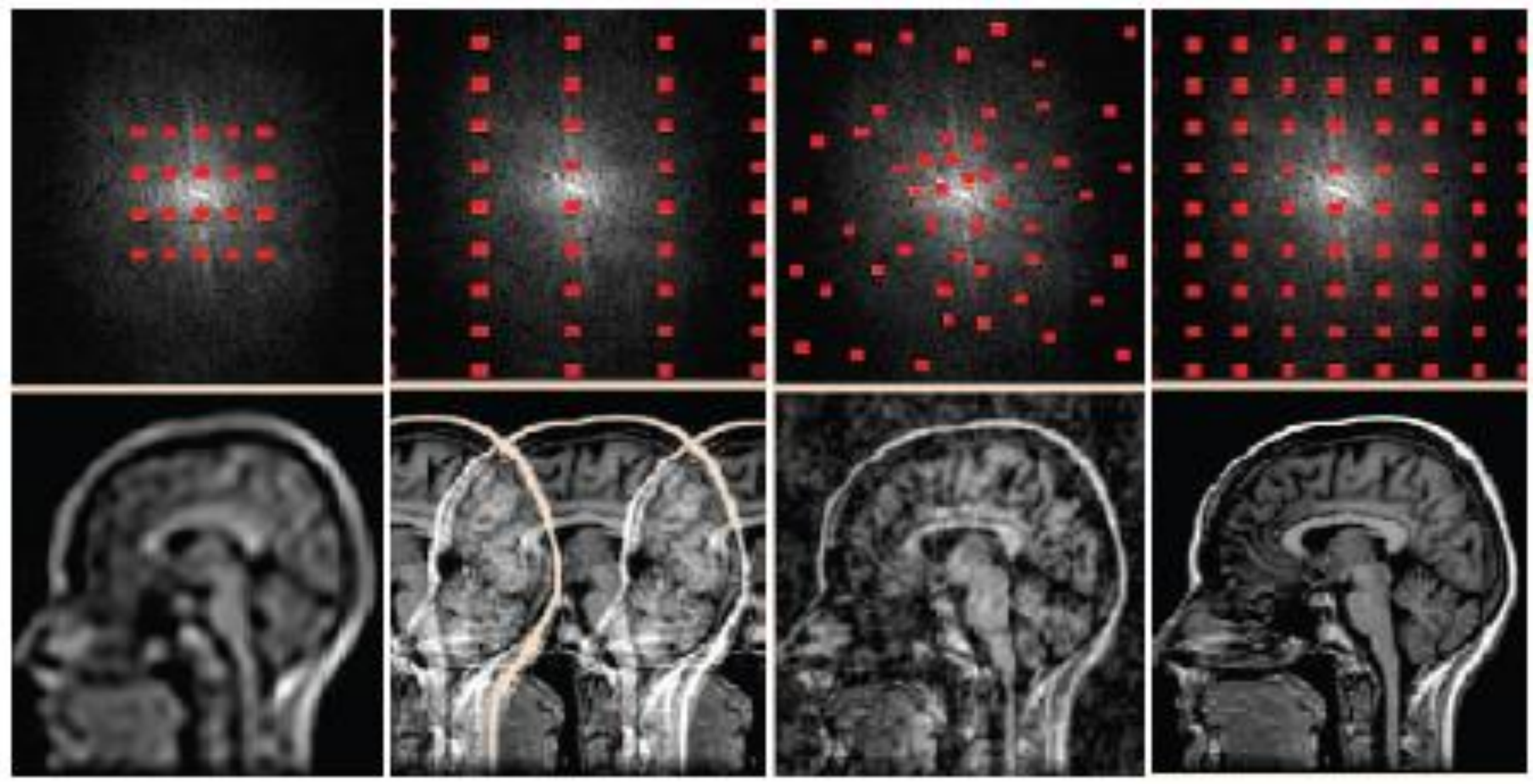


Compressed sensing MRI (Lustig et al. 2008)

MRI as a compressed sensing system.



Spatial encoding



Direct problem

- $x \in R^n$: The clean image in the spatial domain;
- F ∈ $C^{n \times n}$: Fourier Transform;
- $M \in diag(n, n), M(i,i) = 0 \; or \; 1, i = 1, \cdots, n$: Sampling Mask;
- MFx: Sampled Fourier transform of the clean image;
- Direct process: $y = MFx + v$, $v \in R^{n \times n}$ (noise), $y \in R^{n \times n}$(observed data).

Inverse problem

- Problem: How to restore x?

Consider the following linear system

$$\mathrm{y} = MFx \ (\mathrm{y, K \ and \ F \ are \ given}).$$

This system may have infinitely many solutions. Which solution should be chosen?

This system may be unsolvable.

How to insert the unknown noise to the problem?

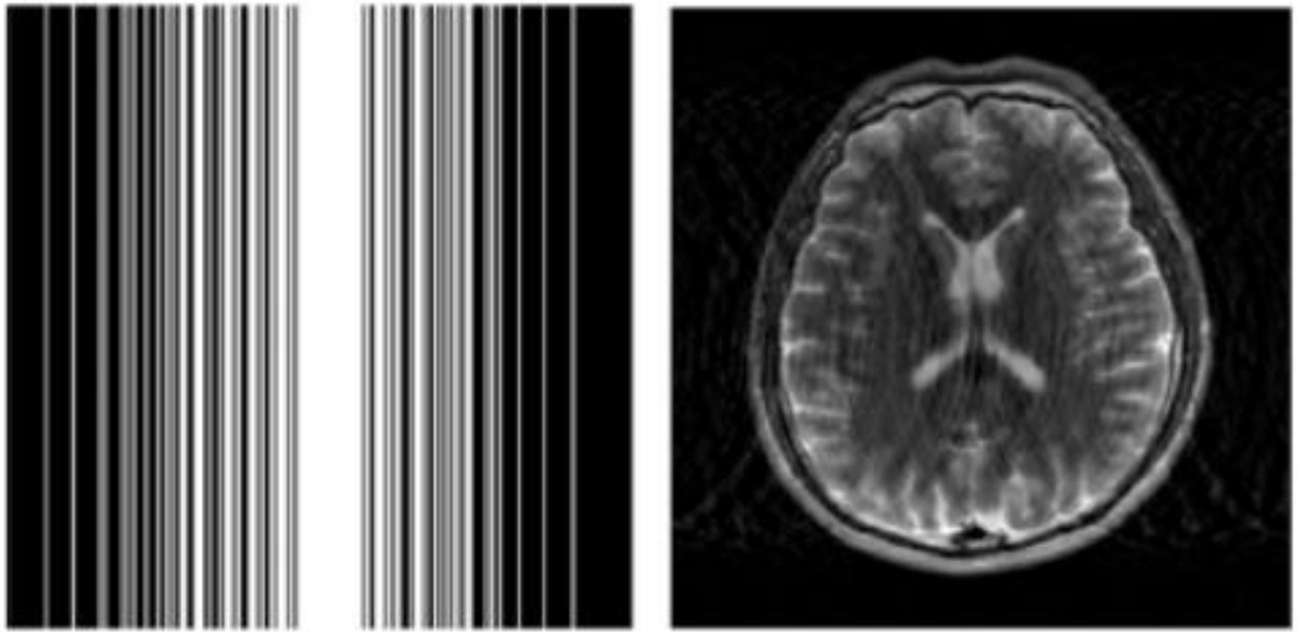
Zero filled solution

$$y = M\mathsf{F}x + v, b = y - v \rightarrow b = M\mathsf{F}x \rightarrow$$

$$x = \mathsf{F}^{-1}b \text{ is a solution.}$$

Definition (zero filled solution): $ZF = \mathsf{F}^{-1}y.$

Sampling pattern (a), zero filled solution (b)



(a) (b)

Compressed MRI

$$\min ||\Psi x||_1$$
$$s.t. ||y - M\mathrm{F}x||_2^2 \leq \epsilon$$

Ψ is a linear operator such that Ψx is a sparse representation of the vector x.

Compressed sensing problem: regularization

Compressed sensing problem:

$$\min ||\Psi x||_1 + \lambda\, R(x),$$
$$\text{s.t.}\ \ ||y - M\mathsf{F}x||_2^2 \le \epsilon.$$

R, is a regularizer. The equivalent form is

$$\min \frac{1}{2}||y - M\mathsf{F}x||_2^2 + \alpha R(x) + \beta||\Psi x||_1,$$

where, α, β are some scalars.

Variational models in the continuous setting

$$g, u: \Omega \to \Re \ or \ [0,1], \Omega \subset \Re^2 \ or \ [0,1]^2,$$

$A: L^2(\Omega) \to L^2(\Omega)$ is a bounded linear operator.

Assume g is an image containing some artifacts (noise, blur,…) and A is an approximation of a linear operator, operating on a clean image u, and turns it in to g . Consider the following optimization problem to restore clean image u:

$$\min_{u \in L^2(\Omega)} \lambda R(u(x)) + \frac{1}{2} \int_\Omega |Au(x) - g(x)|^2 \, dx.$$

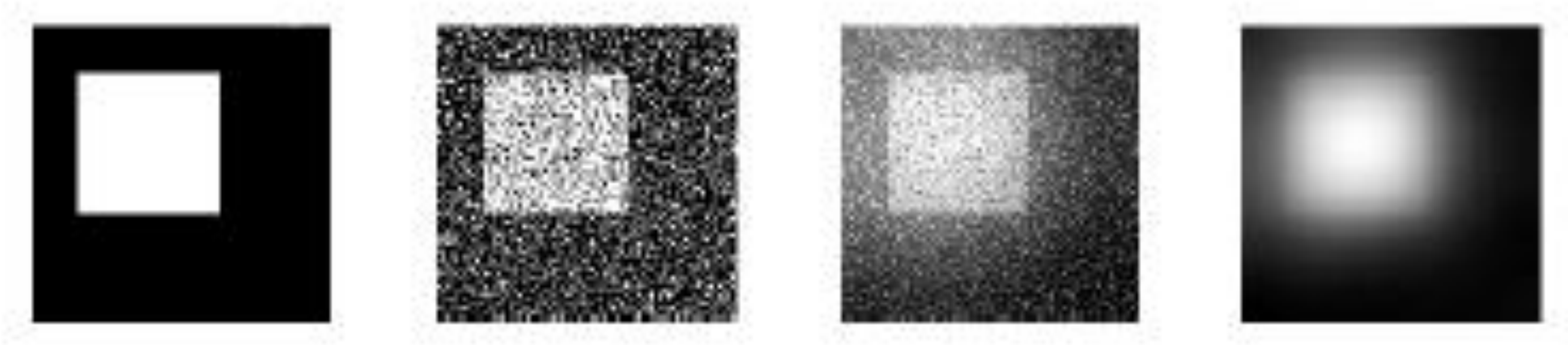
How to choose regularization function R?

- Standard Tychonov regularization functions:

$$\Omega \subset \Re^N, u \in L^2(\Omega), R(u) = \frac{1}{2}\int_\Omega \mathrm{u}^2 dx \text{ or } \frac{1}{2}\int_\Omega |\nabla u|^2 dx,$$

$$\nabla u(x) = \left(\frac{\partial u(x)}{\partial x_1}, \cdots, \frac{\partial u(x)}{\partial x_N}\right), |\,.\,| = ||.||_2.$$

Wrong Choice (clean, noisy and restored images)



Distributional derivatives (Reddy 1998)

Distribution

Assume $\Omega \subset \Re^n$.

Every linear, continuous functional

$$F : D(\Omega) \to \Re,$$

is a distribution on the domain Ω, that is $F \in D(\Omega)'$, where $D(\Omega) = C_c^\infty(\Omega)$.

Special case: $u \in L^1_{loc}(\Omega) \to (u, \phi) = \int_\Omega u\phi dx$.

Classic Green Theorem

Assume $\Omega \subset \Re^n$, $u, v \in C^1(\overline{\Omega}), \nu_i$ is ith component of the outward unit normal ν to the sufficiently smooth boundary Γ of the domain, then

$$\int_\Omega u \frac{\partial v}{\partial x_i} dx = \int_\Gamma u v \nu_i ds - \int_\Omega v \frac{\partial u}{\partial x_i} dx .$$

General case: $u, v \in C^m(\overline{\Omega}), |\alpha| = m$:

$$\int_\Omega u D^\alpha v dx = \int_\Gamma h(u, v) ds + (-1)^{|\alpha|} \int_\Omega v D^\alpha u dx .$$

Derivatives of distributions

$$(D^\alpha u, \phi) = (-1)^{|\alpha|}(u, D^\alpha \phi), \phi \in D(\Omega).$$

Definition (weak derivative): If $u \in L^1_{loc}(\Omega)$ and there exists a $w \in L^1_{loc}(\Omega)$ such that

$$(w, \phi) = (-1)^{|\alpha|}(u, D^\alpha \phi), \phi \in D(\Omega),$$

then $D^\alpha u = w$ is called the αth weak derivative of u.

Example

The ramp function $R(\boldsymbol{x})$ on $\Omega = (-1,1) \times (-1,1) \subset \mathbb{R}^2$ is defined by

$$R(\boldsymbol{x}) = \begin{cases} xy & \text{if} \quad x \geq 0, \;\; y \geq 0, \\ 0 & \text{if} \quad x < 0 \text{ or } y < 0. \end{cases}$$

The generalized derivative $D^{(1,0)}R = \partial R / \partial x$ is found from

$$\begin{aligned} \left\langle \frac{\partial R}{\partial x}, \phi \right\rangle &= -\left\langle R, \frac{\partial \phi}{\partial x} \right\rangle = -\int_{-1}^{1}\int_{-1}^{1} R(\boldsymbol{x}) \frac{\partial \phi}{\partial x}\, dx dy \\ &\qquad (R \text{ is locally integrable}) \\ &= -\int_{0}^{1}\int_{0}^{1} xy \frac{\partial \phi}{\partial x}\, dx dy = \int_{0}^{1}\int_{0}^{1} y\phi \; dx dy \end{aligned}$$

$$
\begin{aligned}
\left\langle \frac{\partial^2 R}{\partial x \partial y}, \phi \right\rangle &= (-1)^2 \left\langle R, \frac{\partial^2 \phi}{\partial x \partial y} \right\rangle = \int_0^1 \int_0^1 xy \frac{\partial^2 \phi}{\partial x\, \partial y}\, dx dy \\
&= \int_0^1 \int_0^1 \phi\, dx dy \quad \text{(applying Green's theorem twice)} \\
&= \int_\Omega H(\boldsymbol{x}) \phi(x)\, dx dy,
\end{aligned}
$$

where H is the two-dimensional step function:

$$
H(\boldsymbol{x}) = \begin{cases} 1 & \text{if} \quad x \geq 0,\ y \geq 0, \\ 0 & \text{if} \quad x < 0 \text{ or } y < 0. \end{cases}
$$

Hence

$$
\left\langle \frac{\partial^2 R}{\partial x \partial y}, \phi \right\rangle = \langle H, \phi \rangle \text{ so that } D^{(1,1)} R = H.
$$

Definition of Sobolev spaces

$$W^{m,p}(\Omega) = \{u:\ D^{\alpha}u \in L^{p}(\Omega); \forall\alpha, |\alpha| \leq m\}.$$

Variational optimization models for image problems (Chambolle et al. 2010)

Images on L^1_{loc}

Assume $u: \Omega \to [0,1]$,

$$\frac{\partial u}{\partial x}: D(\Omega) \to \Re, \qquad \left(\frac{\partial u}{\partial x}, \phi\right) = -\left(u, \frac{\partial \phi}{\partial x}\right), \phi \in D(\Omega),$$

$$\frac{\partial u}{\partial y}: D(\Omega) \to \Re, \left(\frac{\partial u}{\partial y}, \phi\right) = -\left(u, \frac{\partial \phi}{\partial y}\right), \phi \in D(\Omega),$$

$$\nabla u : D(\Omega) \times D(\Omega) \to \Re, \left(\nabla u, \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}\right) = \left(\frac{\partial u}{\partial x}, \phi_1\right) + \left(\frac{\partial u}{\partial y}, \phi_2\right), \phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \in D(\Omega)^2.$$

A motivation for definition of TV (Total Variation)

$$(\nabla u, \phi) = -\left(u, \frac{\partial \phi_1}{\partial x}\right) - \left(u, \frac{\partial \phi_2}{\partial y}\right) = -\left(u, \frac{\partial \phi_1}{\partial x} + \frac{\partial \phi_2}{\partial y}\right) = -(\mathrm{u}, \mathrm{div}\, \phi).$$

From the definition of operator norm, we get

$$||\nabla u|| = \sup\{|(\nabla u, \phi)|, \phi \in D(\Omega)^2, |\phi| \leq 1\},$$

where

$$|\phi| = \sup\left\{\sqrt{\phi_1(x,y)^2 + \phi_2(x,y)^2}, (x,y) \in \Omega\right\}.$$

Total Variation: Definition

$u \in L^1_{loc}(\Omega), \Omega \subset \Re^N$

$$J(u) = \sup\{-\int_\Omega u \,\mathrm{div}\, \phi dx : \phi \in D(\Omega), |\phi(x)| \leq 1, \forall x \in \Omega\}.$$

A function is said to have bounded variation whenever $J(u) < +\infty$.

- Definition: Functions with bounded variation ($BV(\Omega)$):

$$BV(\Omega) = \{u \in L^1(\Omega) : J(u) < \infty\}.$$

- $||u||_{BV(\Omega)} = ||u||_{L^1(\Omega)} + J(u)$ and $BV(\Omega)$ is a Banach space.

Equivalent definition for the smooth cases $W^{1,1}(\Omega)$ and $C^1(\Omega)$

Assume $u \in C^1(\Omega)\ or\ u \in W^{1,1}(\Omega)$, then

$$-\int_\Omega u\mathrm{div}\phi\, dx = \int_\Omega \phi \nabla u\, dx\,,$$

and the sup over all ϕ with $|\phi| \leq 1$ is

$$-\int_\Omega u\mathrm{div}\phi\, dx = \int_\Omega |\nabla u| dx\,.$$

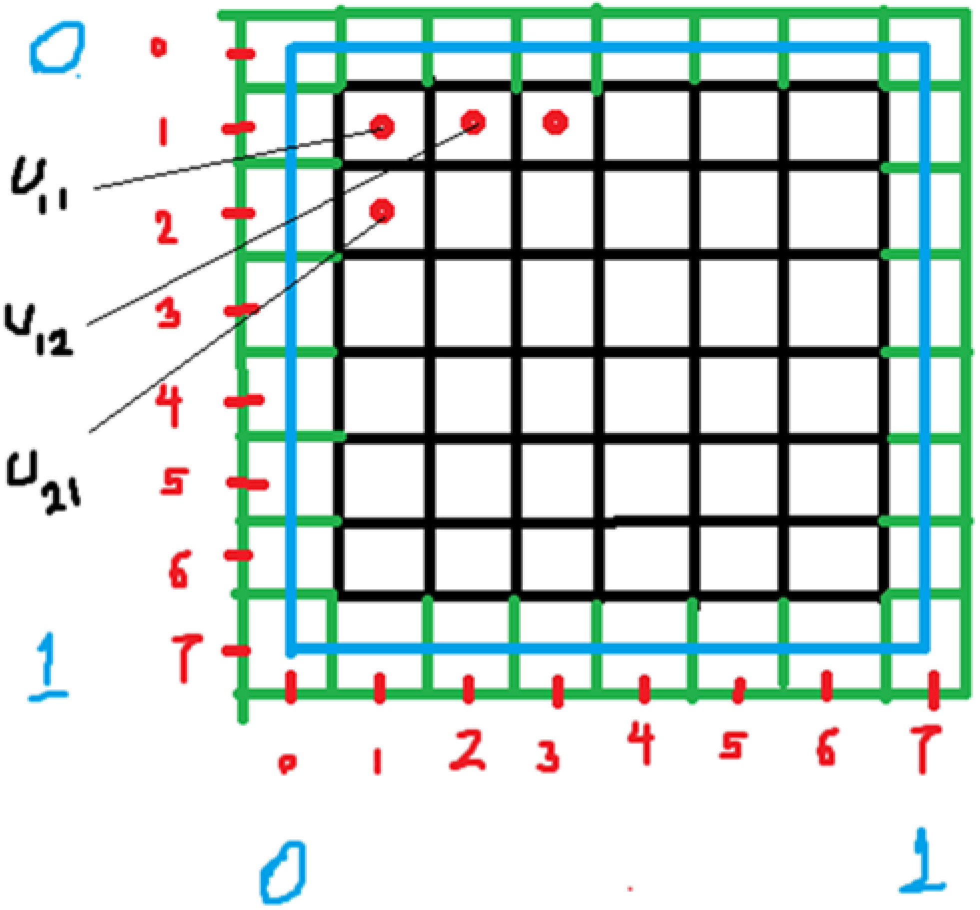
Optimization problem

$$\min_{u} \mathcal{E}(u) \; := J(u) + ||u - g||^2_{L^2(\Omega)}.$$

- J and $\mathcal{E}$ are convex.
- This problem has a unique solution ($\mathcal{E}$ is strictly convex).

Discretization of variational models (Chambolle et al. 2010)

Discrete TV



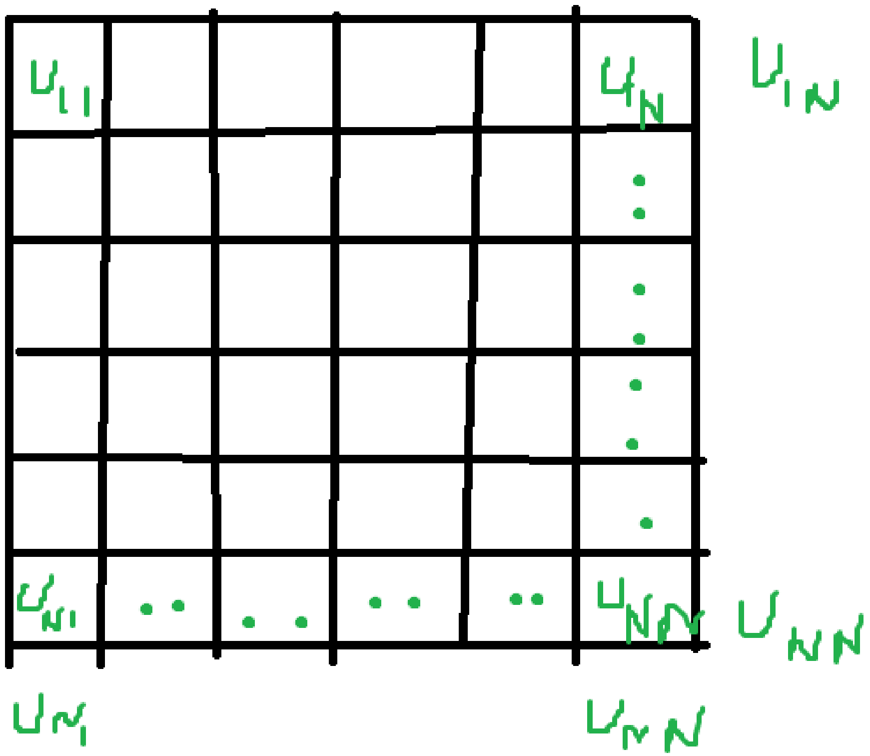
$$N = 6, u\left(\frac{i}{N+1}, \frac{j}{N+1}\right) = u(i,j),$$

$$i,j = 0, \cdots, N+1$$

$$D_x^+ u(i,j) = u(i+1,j) - u(i,j),$$

$$D_y^+ u(i,j) = u(i,j+1) - u(i,j).$$

Symmetric boundary conditions for u



Discrete TV

$$h = \frac{1}{N+1},$$

$$TV_h = \frac{1}{N+1} \sum_{i=1}^{N} \sum_{j=1}^{N} \sqrt{D_x^+ u(i,j) + D_y^+ u(i,j)} \quad .$$

Discrete total variation is correct approximation of total variation functional

Proposition *Let* $\Omega = (0,1)^2$, $p \in [1, +\infty)$, *and* $G : L^p(\Omega) \to \mathbb{R}$ *a continuous functional such that* $\lim_{c\to\infty} G(c+u) = +\infty$ *for any* $u \in L^p(\Omega)$ *(this coerciveness assumption is just to ensure the existence of a solution to the problem, and other situations could be considered). Let* $h = 1/N > 0$ *and* $u^h = (u_{i,j})_{1\le i,j\le N}$, *identified with* $u^h(x) = \sum_{i,j} u_{i,j}\, \chi_{((i-1)h,ih)\times((j-1)h,jh)}(x)$, *be the solution of*

$$\min_{u^h} TV_h(u^h) \; + \; G(u^h)\,.$$

Then, there exists $u \in L^p(\Omega)$ *such that some subsequence* $u^{h_k} \to u$ *as* $k \to \infty$ *in* $L^1(\Omega)$, *and* u *is a minimizer in* $L^p(\Omega)$ *of*

$$J(u) \; + \; G(u)\,.$$

ROF optimization model

$$\min_{u \in \Re^{n \times n}} \lambda ||\nabla u||_{2,1} + \frac{1}{2} ||Au - g||_2,$$

$$p(i,j) = \left(p_1(i,j), p_2(i,j)\right), ||p||_{2,1} = \sum_{i,j=1}^{n} \sqrt{p_1(i,j)^2 + p_2(i,j)^2},$$

$$\nabla u(i,j) = \left(D_x^+ u(i,j), D_y^+ u(i,j)\right),$$

If g is a noisy image, then put $A = I$.

ROF model for MRI restoration problem

$$\min \mathrm{E(u)} = \frac{1}{2}||z - M\mathrm{F}u||_2^2 + \alpha||\nabla u||_{2,1} + \beta||\Psi u||_1.$$

A General Form:

$$\min_{u \in \Re^n} F(Au) + G(u),$$

F and G are suitable convex functions.

Primal-dual algorithm to solve MRI problems
(Esser 2009)

Primal-Dual Algorithm (Preliminaries)

Adjoint of a linear operator:

$A: \Re^n \to \Re^m$ is a linear operator :

$$\forall x \in \Re^n, \forall y \in \Re^m, (Ax, y) = (x, A^* y).$$

Conjugate of a convex nonlinear operator:

$F: \Re^n \to \Re$ is a suitable convex function:

$$F^*: \Re^n \to \Re: \quad F^*(y) = \sup_x \{(y, x) - F(x)\}.$$

Proximal operator:

$$\mathrm{Prox}_{\sigma F}(y) = \mathrm{argmin}_x\, \sigma F(x) + \left|\left|x - y\right|\right|_2^2/2.$$

Primal-dual problems

Primal Problem:

$$\min_{x \in \Re^n} F(Ax) + G(x),$$

Dual Problem:

$$\max_{y} -\left(G^*(-A^*y) - F^*(y)\right).$$

Primal-dual algorithm

- Initialization: choose $x_0 = \overline{x_0},\ \ y_0$ arbitrary;
- For $n \geq 0$, until the primal value is not equal to the dual value, do:
- $y^{n+1} = \mathrm{Prox}_{\sigma F^*}(y^n + \sigma A\, \overline{x^n})$,
- $x^{n+1} = \mathrm{Prox}_{\tau G}(x^n - \tau A^* y^{n+1})$,
- $\overline{x^{n+1}} = 2x^{n+1} - x^n$.

Theorem: Under some weak assumptions, convergence will be guaranteed if $\tau\sigma L^2 < 1$ $(L = ||A||)$.

Primal-dual algorithm for MRI restoration

$$\min \mathrm{E(u)} = \frac{1}{2}||z - M\mathsf{F}u||_2^2 + \alpha||\nabla u||_{2,1} + \beta||\Psi u||_1$$

$F(y_1, y_2, y_3) = \frac{1}{2}||z - y_1||_2^2 + \alpha||y_2||_{2,1} + \beta||y_3||_1, G(u) = 0,$

$A = (M\mathsf{F}, \nabla, \Psi) \quad , y = (y_1, y_2, y_3)^t \rightarrow F(Au) = E(u),$

$A^* = (F^* M^t, -\mathrm{div}, \Psi^*)^t,$

$$\begin{aligned} p = (p_1, p_2), \mathrm{div}\, p(i,j) &= D_x^- p_1(i,j) + D_y^- p_2(i,j) \\ &= (p_1(i,j) - p_1(i-1,j)) + (p_2(i,j) - p_2(i,j-1)). \end{aligned}$$

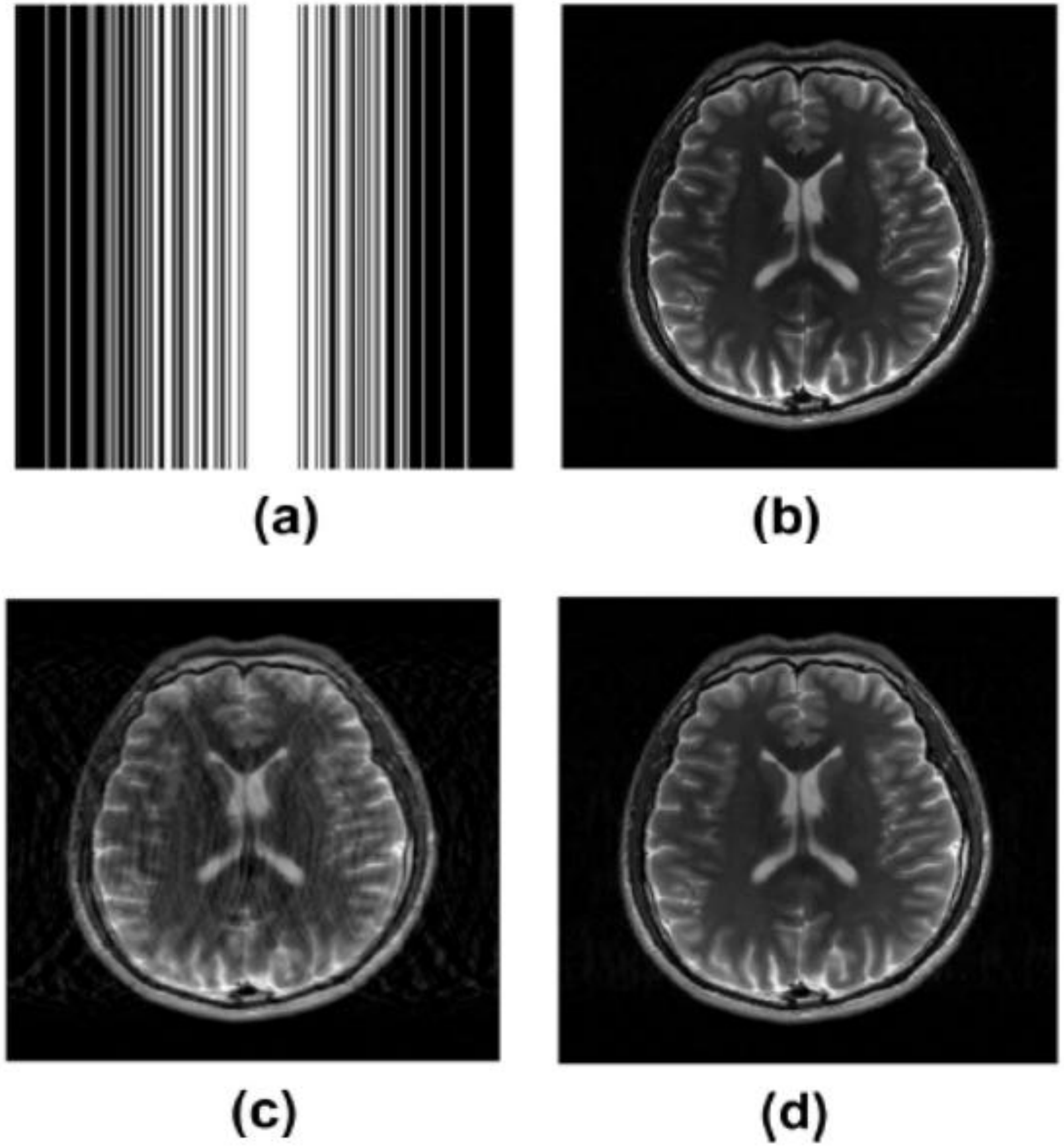
Primal-dual algorithm for MRI restoration

$$F^*(v) = F^*(v_1, v_2, v_3) = (z, v_1) + \frac{1}{2}||v_1||_2^2 + I_{\{||v_2||_{2,\infty} \leq \alpha, ||v_3||_\infty \leq \beta\}}(v),$$

$$\mathrm{Prox}_{\sigma F^*}(w_1, w_2, w_3) =$$

$$\left(\frac{w_1 - \sigma z}{2}, \begin{cases} w_2, ||w_2||_{2,\infty} \leq \alpha \\ \dfrac{\alpha w_2}{||w_2||_{2,\infty}}, \text{otherwise} \end{cases}, \begin{cases} w_3, ||w_3||_\infty \leq \beta \\ \dfrac{\alpha w_3}{||w_3||_\infty}, \text{otherwise} \end{cases} \right).$$

Sampling pattern (a), clean image (b), zero filled solution (c), restored image, using variational model (d).



About SOAL Optimizer Competition (Selesnick et al. 2017)

Non-separable regularization

$$\min ||x||_0$$
$$\text{s.t. } \text{Ax} = \text{b}$$

A suggestion: Using non-separable regularization:

$$\min \psi(x) + \lambda ||Ax - b||_2^2,$$

$$\psi(x) = \|x\|_1 - \frac{\lambda}{\gamma\|B\|_1^2} S\left(\frac{\gamma\|B\|_1}{\lambda} Bx\right),$$

$$B^{\mathsf{T}}B \preccurlyeq A^{\mathsf{T}}A, \qquad 0 < \gamma \leqslant 1 \cdot$$

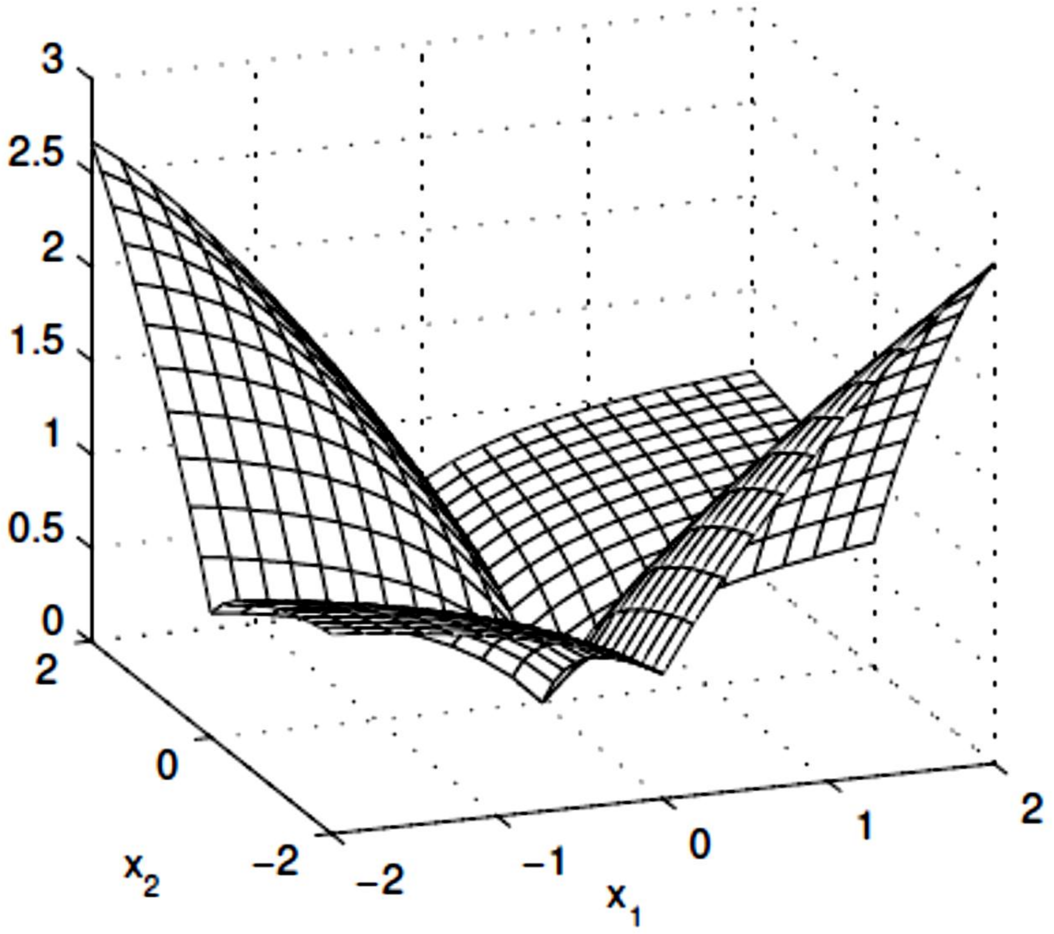
Various kindes of S

Log	$\phi(t) = \log(1 + \lvert t \rvert)$
Rat	$\phi(t) = \dfrac{\lvert t \rvert}{1 + \lvert t \rvert / 2}$
Atan	$\phi(t) = \frac{2}{\sqrt{3}}\left(\tan^{-1}\left(\frac{1+2\lvert t \rvert}{\sqrt{3}}\right) - \frac{\pi}{6}\right)$
Exp	$\phi(t) = 1 - e^{-\lvert t \rvert}$
MC	$\phi(t) = \begin{cases} \lvert t \rvert - \frac{1}{2}t^2, & \lvert t \rvert \leqslant 1 \\ \frac{1}{2}, & \lvert t \rvert \geqslant 1 \end{cases}$

$$s(t) = \lvert t \rvert - \phi(t).$$

$$S(x) = \sum_n s(x_n).$$

Penalty function ψ



References

[1] Bredies, K., Lorenz, D.: Mathematical Image Processing, Springer Nature Switzerland AG (2018).

[2] Reddy, B. D.: Introductory Functional Analysis, Springer (1998).

[3] Chambolle, A., Caselles, V., Cremers, D., Novaga, M., Pock, T.: An introduction to total variation for image analysis. In: Theoretical Foundations and Numerical Methods for Sparse Recovery, Radon Series on Computational and Applied Mathematics, vol. 9, pp. 263–340. Walter de Gruyter, Berlin (2010).

References

[4] Esser, E.: Applications of lagrangian-based alternating direction methods and connections to split Bregman. CAM Reports 09-31, UCLA, Center for Applied Math. (2009).

[5] Selesnick, I., Farshchian, M.: Sparse signal approximation via non-separable Regularization, IEEE Transactions on Signal Processing, vol. 65, pp. 2561-2575. (2017).