Texture segmentation based on fractal attributes

Barbara Pascal
(bpascal-fr.github.io)

Joint work with Patrice Abry, Nelly Pustelnik, Valérie Vidal, and Samuel Vaiter

July 4th 2023

Harmonic and Multifractal Analyses: from Mathematics to Quantitative Neuroscience

Centre de Recherches Mathématiques (CRM), Montréal, Canada

Textures

Textures

Textures

Textures

Crucial to describe real-world images

Textured image segmentation

Textured image segmentation

Goal: obtain a partition of the image into K homogeneous textures

$$\Omega = \Omega_1 \bigsqcup \ldots \bigsqcup \Omega_K$$

Textured image segmentation

Goal: obtain a partition of the image into K **homogeneous textures**

$$\Omega = \Omega_1 \bigsqcup \ldots \bigsqcup \Omega_K$$

Piecewise monofractal model

Piecewise monofractal model

Fractal attributes

- variance σ^2 *amplitude of variations*

Piecewise monofractal model

Fractal attributes

- variance σ^2 *amplitude of variations*
- local regularity h *scale invariance*

Piecewise monofractal model

Fractal attributes

- variance σ^2 *amplitude of variations*
- local regularity h *scale invariance*

$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$

Fractal attributes

- variance σ^2 *amplitude of variations*
- local regularity h *scale invariance*

$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$

$h(x) \equiv h_1 = 0.9$

$h(x) \equiv h_2 = 0.3$

Fractal attributes

- variance σ^2 *amplitude of variations*
- local regularity h *scale invariance*

$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$

$h(x) \equiv h_1 = 0.9$

$h(x) \equiv h_2 = 0.3$

Segmentation

- σ^2 and h piecewise constant

(h_2, σ_2^2)

(h_1, σ_1^2)

(h_1, σ_1^2)

Fractal attributes

- variance σ^2 *amplitude of variations*
- local regularity h *scale invariance*

$$|f(x) - f(y)| \leq \sigma(x)|x - y|^{h(x)}$$

$h(x) \equiv h_1 = 0.9$

$h(x) \equiv h_2 = 0.3$

Segmentation

- σ^2 and h piecewise constant
- region Ω_k characterized by (σ_k^2, h_k)

(h_2, σ_2^2)

(h_1, σ_1^2)

(h_1, σ_1^2)

Textured image

Multiscale analysis

Textured image

Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

Multiscale analysis

Textured image

Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

Scale

$a = 2^1$

$a = 2^2$

...

$a = 2^5$

Multiscale analysis

Textured image

Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

Scale

$a = 2^1$

$a = 2^2$

$\dots$

$a = 2^5$

Proposition (Jaffard, 2004), (Wendt, 2008)

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a) \underbrace{\boldsymbol{h}}_{\text{regularity}} + \underbrace{\boldsymbol{v}}_{\substack{\propto \log(\boldsymbol{\sigma}^2) \\ \text{(variance)}}}$$

Multiscale analysis

Textured image

Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

Scale

$a = 2^1$ $a = 2^2$ $a = 2^5$

Proposition (Jaffard, 2004), (Wendt, 2008)

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a)\,\underbrace{\boldsymbol{h}}_{\text{regularity}} + \underbrace{\boldsymbol{v}}_{\substack{\propto\log(\boldsymbol{\sigma}^2)\\ \text{(variance)}}}$$

Multiscale analysis

Textured image

Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

Scale

$a = 2^1$ $a = 2^2$ $a = 2^5$

Proposition (Jaffard, 2004), (Wendt, 2008)

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a)\,\underbrace{\boldsymbol{h}}_{\text{regularity}} + \underbrace{\boldsymbol{v}}_{\substack{\propto\log(\boldsymbol{\sigma}^2)\\ \text{(variance)}}}$$

Multiscale analysis

Textured image

Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

Scale

$a = 2^1$ $a = 2^2$ $a = 2^5$

Proposition (Jaffard, 2004), (Wendt, 2008)

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a)\,\underbrace{\boldsymbol{h}}_{\text{regularity}} + \underbrace{\boldsymbol{v}}_{\substack{\propto\log(\boldsymbol{\sigma}^2)\\ \text{(variance)}}}$$

Multiscale analysis

Textured image

Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

Scale

$a = 2^1$ $a = 2^2$ $a = 2^5$

Proposition (Jaffard, 2004), (Wendt, 2008)

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a)\,\underbrace{\boldsymbol{h}}_{\text{regularity}} + \underbrace{\boldsymbol{v}}_{\substack{\propto\log(\boldsymbol{\sigma}^2)\\ \text{(variance)}}}$$

Multiscale analysis

Textured image

Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

Scale

$a = 2^1$

$a = 2^2$

$\dots$

$a = 2^5$

Proposition (Jaffard, 2004), (Wendt, 2008)

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a)\,\underbrace{\boldsymbol{h}}_{\text{regularity}} + \underbrace{\boldsymbol{v}}_{\substack{\propto\log(\boldsymbol{\sigma}^2)\\ \text{(variance)}}}$$

Multiscale analysis

Textured image

Local maximum of wavelet coefficients: $\mathcal{L}_{a,\cdot}$

Scale

$a = 2^1$ $a = 2^2$ $a = 2^5$

Proposition (Jaffard, 2004), (Wendt, 2008)

$$\log\left(\mathcal{L}_{a,\cdot}\right) \underset{a\to 0}{\simeq} \log(a)\,\underbrace{\boldsymbol{h}}_{\text{regularity}} + \underbrace{\boldsymbol{v}}_{\substack{\propto\log(\boldsymbol{\sigma}^2)\\ \text{(variance)}}}$$

Direct punctual estimation

Linear regression $\log\left(\mathcal{L}_{a,\cdot}\right) \simeq \log(a) \underset{\text{regularity}}{\boldsymbol{h}} + \underset{\propto \log(\boldsymbol{\sigma}^2)}{\boldsymbol{v}}$

Textured image

Direct punctual estimation

Linear regression $\log\left(\mathcal{L}_{a,\cdot}\right) \simeq \log(a) \underset{\text{regularity}}{\boldsymbol{h}} + \underset{\propto \log(\boldsymbol{\sigma}^2)}{\boldsymbol{v}}$

$$\left(\widehat{\boldsymbol{h}}^{\mathrm{LR}}, \widehat{\boldsymbol{v}}^{\mathrm{LR}}\right) = \underset{\boldsymbol{h},\boldsymbol{v}}{\mathrm{argmin}} \sum_{a=a_{\min}}^{a_{\max}} \left\|\log\left(\mathcal{L}_{a,\cdot}\right) - \log(a)\boldsymbol{h} - \boldsymbol{v}\right\|^2$$

Textured image

Direct punctual estimation

Linear regression $\log\left(\mathcal{L}_{a,\cdot}\right) \simeq \log(a) \underset{\text{regularity}}{\boldsymbol{h}} + \underset{\propto \log(\boldsymbol{\sigma}^2)}{\boldsymbol{v}}$

$$\left(\widehat{\boldsymbol{h}}^{\mathrm{LR}}, \widehat{\boldsymbol{v}}^{\mathrm{LR}}\right) = \underset{\boldsymbol{h},\boldsymbol{v}}{\mathrm{argmin}} \sum_{a=a_{\min}}^{a_{\max}} \left\|\log\left(\mathcal{L}_{a,\cdot}\right) - \log(a)\boldsymbol{h} - \boldsymbol{v}\right\|^2$$

Textured image

Local regularity $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Local power $\widehat{\boldsymbol{v}}^{\mathrm{LR}}$

Linear regression $\underbrace{\mathbb{E}\log\left(\mathcal{L}_{a,\cdot}\right)}_{\text{expected value}} = \log(a)\underbrace{\bar{\boldsymbol{h}}}_{\text{regularity}} + \underbrace{\bar{\boldsymbol{v}}}_{\propto\log(\boldsymbol{\sigma}^2)}$

$$\left(\widehat{\boldsymbol{h}}^{\mathrm{LR}}, \widehat{\boldsymbol{v}}^{\mathrm{LR}}\right) = \underset{\boldsymbol{h},\boldsymbol{v}}{\mathrm{argmin}} \sum_{a=a_{\min}}^{a_{\max}} \left\|\log\left(\mathcal{L}_{a,\cdot}\right) - \log(a)\boldsymbol{h} - \boldsymbol{v}\right\|^2$$

Textured image

Local regularity $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Local power $\widehat{\boldsymbol{v}}^{\mathrm{LR}}$

$\longrightarrow$ large estimation variance

A posteriori regularization

Filter smoothing (linear)

$$\left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \widehat{\boldsymbol{h}}^{\mathrm{LR}}$$

Linear regression $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Smoothing

A posteriori regularization

Filter smoothing (linear)

$$\left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \widehat{\boldsymbol{h}}^{\mathrm{LR}}$$

ROF denoising (nonlinear)

$$\underset{\boldsymbol{h}}{\operatorname{argmin}} \; \|\boldsymbol{h} - \widehat{\boldsymbol{h}}^{\mathrm{LR}}\|^2 + \lambda \|\mathbf{D}\boldsymbol{h}\|_{2,1}$$

Linear regression $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Smoothing

ROF

$\longrightarrow$ cumulative estimation variance and regularization bias

Functionals with either free or co-localized contours

$$\sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}}$$

$\rightarrow$ fidelity to the log-linear model

Functionals with either free or co-localized contours

$$\sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

$\rightarrow$ fidelity to the log-linear model

$\rightarrow$ favors piecewise constancy

Functionals with either free or co-localized contours

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\substack{\textbf{Least-Squares} \\ \rightarrow \text{ fidelity to the log-linear model}}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\substack{\textbf{Total Variation} \\ \rightarrow \text{ favors piecewise constancy}}}$$

Functionals with either free or co-localized contours

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

$\rightarrow$ fidelity to the log-linear model

$\rightarrow$ favors piecewise constancy

Finite differences $\mathbf{D}_1\boldsymbol{x}$ (horizontal), $\mathbf{D}_2\boldsymbol{x}$ (vertical) in each pixel

Functionals with either free or co-localized contours

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\substack{\textbf{Least-Squares} \\ \rightarrow \text{ fidelity to the log-linear model}}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\substack{\textbf{Total Variation} \\ \rightarrow \text{ favors piecewise constancy}}}$$

Finite differences $\mathbf{D}\boldsymbol{x} = [\mathbf{D}_1\boldsymbol{x}, \mathbf{D}_2\boldsymbol{x}]$

<u>Free:</u> $\boldsymbol{h}$, $\boldsymbol{v}$ are **independently** piecewise constant

$$\mathcal{Q}_{\mathsf{F}}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha) = \alpha\|\mathbf{D}\boldsymbol{h}\|_{2,1} + \|\mathbf{D}\boldsymbol{v}\|_{2,1}$$

Functionals with either free or co-localized contours

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\substack{\textbf{Least-Squares} \\ \rightarrow \text{ fidelity to the log-linear model}}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\substack{\textbf{Total Variation} \\ \rightarrow \text{ favors piecewise constancy}}}$$

Finite differences $\mathbf{D}\boldsymbol{x} = [\mathbf{D}_1\boldsymbol{x}, \mathbf{D}_2\boldsymbol{x}]$

<u>Free:</u> $\boldsymbol{h}$, $\boldsymbol{v}$ are **independently** piecewise constant

$$\mathcal{Q}_{\mathrm{F}}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha) = \alpha\|\mathbf{D}\boldsymbol{h}\|_{2,1} + \|\mathbf{D}\boldsymbol{v}\|_{2,1}$$

<u>Co-localized:</u> $\boldsymbol{h}$, $\boldsymbol{v}$ are **concomitantly** piecewise constant

$$\mathcal{Q}_{\mathrm{C}}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha) = \|[\alpha\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}]\|_{2,1}$$

Functionals minimization

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

Functionals minimization

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

- gradient descent $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$ $\qquad \boldsymbol{x} = (\boldsymbol{h}, \boldsymbol{v})$

Functionals minimization

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

nonsmooth

- gradient descent $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$ $\qquad \boldsymbol{x} = (\boldsymbol{h}, \boldsymbol{v})$
- implicit subgradient descent: proximal point algorithm

$$\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \boldsymbol{u}^n, \;\; \boldsymbol{u}^n \in \partial\varphi(\boldsymbol{x}^{n+1}) \;\Leftrightarrow\; \boldsymbol{x}^{n+1} = \mathrm{prox}_{\tau\varphi}(\boldsymbol{x}^n)$$

Functionals minimization

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

nonsmooth

- gradient descent $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$ $\qquad \boldsymbol{x} = (\boldsymbol{h}, \boldsymbol{v})$
- implicit subgradient descent: proximal point algorithm

$$\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \boldsymbol{u}^n, \;\; \boldsymbol{u}^n \in \partial\varphi(\boldsymbol{x}^{n+1}) \;\Leftrightarrow\; \boldsymbol{x}^{n+1} = \mathrm{prox}_{\tau\varphi}(\boldsymbol{x}^n)$$

- splitting proximal algorithm (Chambolle, 2011)

$$\boldsymbol{y}^{n+1} = \mathrm{prox}_{\sigma(\lambda\mathcal{Q})^*}\left(\boldsymbol{y}^n + \sigma \mathbf{D}\bar{\boldsymbol{x}}^n\right)$$

$$\boldsymbol{x}^{n+1} = \mathrm{prox}_{\tau\|\boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}\cdot\|_2^2}\left(\boldsymbol{x}^n - \tau \mathbf{D}^\top \boldsymbol{y}^{n+1}\right), \qquad \boldsymbol{\Phi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$$

$$\bar{\boldsymbol{x}}^{n+1} = 2\boldsymbol{x}^{n+1} - \boldsymbol{x}^n$$

Functionals minimization

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

nonsmooth

- gradient descent $\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \nabla \varphi(\boldsymbol{x}^n)$ $\qquad \boldsymbol{x} = (\boldsymbol{h}, \boldsymbol{v})$
- implicit subgradient descent: proximal point algorithm

$$\boldsymbol{x}^{n+1} = \boldsymbol{x}^n - \tau \boldsymbol{u}^n, \;\; \boldsymbol{u}^n \in \partial\varphi(\boldsymbol{x}^{n+1}) \;\Leftrightarrow\; \boldsymbol{x}^{n+1} = \mathrm{prox}_{\tau\varphi}(\boldsymbol{x}^n)$$

- splitting proximal algorithm (Chambolle, 2011)

$$\boldsymbol{y}^{n+1} = \mathrm{prox}_{\sigma(\lambda\mathcal{Q})^*}\left(\boldsymbol{y}^n + \sigma \mathbf{D}\bar{\boldsymbol{x}}^n\right)$$

$$\boldsymbol{x}^{n+1} = \mathrm{prox}_{\tau\|\boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}\cdot\|_2^2}\left(\boldsymbol{x}^n - \tau \mathbf{D}^\top \boldsymbol{y}^{n+1}\right), \qquad \boldsymbol{\Phi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$$

$$\bar{\boldsymbol{x}}^{n+1} = 2\boldsymbol{x}^{n+1} - \boldsymbol{x}^n$$

Accelerated algorithm based on strong-convexity

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

nonsmooth

Primal-dual algorithm (Chambolle, 2011)

δ: duality gap, $\delta(\boldsymbol{x}^n, \boldsymbol{y}^n) \underset{n\to\infty}{\longrightarrow} 0$

Accelerated algorithm based on strong-convexity

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

nonsmooth

Primal-dual algorithm (Chambolle, 2011)

δ: duality gap, $\delta(\boldsymbol{x}^n, \boldsymbol{y}^n) \xrightarrow[n\to\infty]{} 0$

Convexity properties

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

μ-strongly convex

nonsmooth

Convexity properties

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_a \underbrace{\|\log \mathcal{L}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

μ-strongly convex

nonsmooth

Strong-convexity

- φ μ-strongly convex iff $\varphi - \frac{\mu}{2}\|\cdot\|^2$ convex

✓ strictly convex
✗ non strongly convex

✓ strictly convex
✓ 1-strongly convex

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

$\boldsymbol{\mu}$-strongly convex

nonsmooth

Strong-convexity

- φ $\boldsymbol{\mu}$-strongly convex iff $\varphi - \dfrac{\boldsymbol{\mu}}{2}\|\cdot\|^2$ convex
- φ $\mathcal{C}^2$ with Hessian matrix $\boldsymbol{H}\varphi \succeq 0 \implies \boldsymbol{\mu} = \min \mathrm{Sp}(\boldsymbol{H}\varphi)$

Convexity properties

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \underbrace{\sum_a \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

$\boldsymbol{\mu}$-strongly convex

nonsmooth

Strong-convexity

- φ $\boldsymbol{\mu}$-strongly convex iff $\varphi - \dfrac{\boldsymbol{\mu}}{2}\|\cdot\|^2$ convex
- φ $\mathcal{C}^2$ with Hessian matrix $\boldsymbol{H}\varphi \succeq 0 \implies \boldsymbol{\mu} = \min \mathrm{Sp}(\boldsymbol{H}\varphi)$

Proposition (Pascal, 2019)

$\sum_a \|\log \boldsymbol{\mathcal{L}} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2$ is $\boldsymbol{\mu}$-strongly convex.

$a_{\min} = 2^1$, $a_{\max}$	2^2	2^3	2^4	2^5	2^6
$\boldsymbol{\mu} = \min \mathrm{Sp}\left(2\boldsymbol{\Phi}^\top\boldsymbol{\Phi}\right)$	0.29	**0.72**	1.20	1.69	2.20

Accelerated algorithm based on strong-convexity

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_a \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

μ-strongly convex

nonsmooth

Accelerated Primal-dual algorithm *(Chambolle, 2011)*

for $n = 0, 1, \ldots$ $\qquad \boldsymbol{x} = (\boldsymbol{h}, \boldsymbol{v})$

$$\boldsymbol{y}^{n+1} = \operatorname{prox}_{\sigma_n(\lambda\mathcal{Q})^*}\left(\boldsymbol{y}^n + \sigma_n \mathbf{D}\bar{\boldsymbol{x}}^n\right)$$

$$\boldsymbol{x}^{n+1} = \operatorname{prox}_{\tau_n\|\boldsymbol{\mathcal{L}} - \boldsymbol{\Phi}\cdot\|_2^2}\left(\boldsymbol{x}^n - \tau_n \mathbf{D}^\top \boldsymbol{y}^{n+1}\right)$$

$$\theta_n = \sqrt{1 + 2\mu\tau_n}, \quad \tau_{n+1} = \tau_n/\theta_n, \quad \sigma_{n+1} = \theta_n\sigma_n$$

$$\bar{\boldsymbol{x}}^{n+1} = \boldsymbol{x}^{n+1} + \theta_n^{-1}\left(\boldsymbol{x}^{n+1} - \boldsymbol{x}^n\right)$$

Accelerated algorithm based on strong-convexity

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

μ-strongly convex

nonsmooth

Accelerated Primal-dual algorithm *(Chambolle, 2011)*

δ: duality gap, $\delta(\boldsymbol{x}^n, \boldsymbol{y}^n) \xrightarrow[n\to\infty]{} 0$

Segmentation *via* iterated thresholding

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

Textured image

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Segmentation *via* iterated thresholding

$$\underset{\boldsymbol{h},\boldsymbol{v}}{\text{minimize}} \quad \sum_{a} \underbrace{\|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2}_{\textbf{Least-Squares}} \quad + \quad \lambda \underbrace{\mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)}_{\textbf{Total Variation}}$$

Textured image

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

Co-localized contours $\widehat{\boldsymbol{h}}^{\mathrm{C}}$

Threshold estimate† $T\widehat{\boldsymbol{h}}^{\mathrm{C}}$

†(Cai, 2013)

Threshold-ROF on $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

(Nafornita, 2014), (Pustelnik, 2016)

$$\underset{\boldsymbol{h}}{\operatorname{argmin}} \; \|\boldsymbol{h} - \widehat{\boldsymbol{h}}^{\mathrm{LR}}\|^2 + \lambda \|\mathbf{D}\boldsymbol{h}\|_{2,1}$$

Lin. reg. ROF Threshold

Only based on regularity $\boldsymbol{h}$.

State-of-the-art methods for texture segmentation

Threshold-ROF on $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

(Nafornita, 2014), (Pustelnik, 2016)

$$\underset{\boldsymbol{h}}{\operatorname{argmin}} \; \|\boldsymbol{h} - \widehat{\boldsymbol{h}}^{\mathrm{LR}}\|^2 + \lambda \|\mathbf{D}\boldsymbol{h}\|_{2,1}$$

Lin. reg. ROF Threshold

Only based on regularity $\boldsymbol{h}$.

Factorization based segmentation†

(Yuan, 2015)

(i) local histograms

(ii) matrix factorization

(a) (b)

Fig. 2. Scatterplot of features in subspace. (a) Scatterplot of features projected onto the 3-d subspace. (b) Scatterplot after removing features with high edgeness.

†https://sites.google.com/site/factorizationsegmentation/

Multiphase flow through porous media

Laboratoire de Physique, ENS Lyon, V. Vidal, T. Busser, (M. Serres, IFPEN)

Solid
foam

Multiphase flow through porous media

Laboratoire de Physique, ENS Lyon, V. Vidal, T. Busser, (M. Serres, IFPEN)

Multiphase flow through porous media

Laboratoire de Physique, ENS Lyon, V. Vidal, T. Busser, (M. Serres, IFPEN)

→ 1600 × 1100 pixels
→ video: ∼ 1000 images
→ phase diagram: ∼ 10 flow rates

Low activity: $Q_{\mathrm{G}} = 300\mathrm{mL/min}$ - $Q_{\mathrm{L}} = 300\mathrm{mL/min}$

Low activity: $Q_{\mathrm{G}} = 300\mathrm{mL/min}$ - $Q_{\mathrm{L}} = 300\mathrm{mL/min}$

Liquid: $h_{\mathrm{L}} = 0.4$ $\quad \sigma^2_{\mathrm{dark}} = 10^{-2}$

Gas: $h_{\mathrm{G}} = 0.9$ $\quad \sigma^2_{\mathrm{dark}} = 10^{-2}$ (dark bubbles)
$\sigma^2_{\mathrm{clear}} = 10^{-1}$ (clear bubbles).

Transition: $Q_{\mathrm{G}} = 400\mathrm{mL/min}$ - $Q_{\mathrm{L}} = 700\mathrm{mL/min}$

Liquid: $h_{\mathrm{L}} = 0.4$ $\quad \sigma^2_{\mathrm{dark}} = 10^{-2}$

Gas: $h_{\mathrm{G}} = 0.9$ $\quad \sigma^2_{\mathrm{dark}} = 10^{-2}$ (dark bubbles)
$\sigma^2_{\mathrm{clear}} = 10^{-1}$ (clear bubbles).

High activity: $Q_{\mathrm{G}} = 1200\mathrm{mL/min}$ - $Q_{\mathrm{L}} = 300\mathrm{mL/min}$

Liquid: $h_{\mathrm{L}} = 0.4$ $\quad \sigma^2_{\mathrm{dark}} = 10^{-2}$

Gas: $h_{\mathrm{G}} = 0.9$ $\quad \left| \begin{array}{ll} \sigma^2_{\mathrm{dark}} = 10^{-2} & \text{(dark bubbles)} \\ \sigma^2_{\mathrm{clear}} = 10^{-1} & \text{(clear bubbles).} \end{array} \right.$

High activity: $Q_{\mathrm{G}} = 1200\mathrm{mL/min}$ - $Q_{\mathrm{L}} = 300\mathrm{mL/min}$

Liquid: $h_{\mathrm{L}} = 0.4$ $\quad \sigma^2_{\mathrm{dark}} = 10^{-2}$

Gas: $h_{\mathrm{G}} = 0.9$ $\quad \sigma^2_{\mathrm{dark}} = 10^{-2}$ (dark bubbles)
$\sigma^2_{\mathrm{clear}} = 10^{-1}$ (clear bubbles).

Regularization parameters selection

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

Regularization parameters selection

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

$(\lambda, \alpha) = (0, 0)$

Regularization parameters selection

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

$(\lambda, \alpha) = (0, 0)$

Co-localized contours estimate $\widehat{\boldsymbol{h}}^{\mathrm{C}}$

$(\lambda, \alpha) = (0.5, 0.5)$

too small

Regularization parameters selection

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$

$(\lambda, \alpha) = (0, 0)$

Co-localized contours estimate $\widehat{\boldsymbol{h}}^{\mathrm{C}}$

$(\lambda, \alpha) = (0.5, 0.5)$

too small

$(\lambda, \alpha) = (500, 500)$

too large

Regularization parameters selection

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

Lin. reg. $\widehat{\boldsymbol{h}}^{\mathrm{LR}}$ Co-localized contours estimate $\widehat{\boldsymbol{h}}^{\mathrm{C}}$

$(\lambda, \alpha) = (0, 0)$ $(\lambda, \alpha) = (0.5, 0.5)$ $(\lambda^\dagger, \alpha^\dagger) = (11.5, 0.8)$ $(\lambda, \alpha) = (500, 500)$

too small optimal too large

What *optimal* means? How to determine $\lambda^\dagger$ and $\alpha^\dagger$?

Parameter tuning (Grid search)

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliary*

Parameter tuning (Grid search)

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliary*

$\bar{\boldsymbol{h}}$: *true* regularity

$$\mathcal{R}(\lambda, \alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha) - \bar{\boldsymbol{h}}\right\|^2$$

Parameter tuning (Grid search)

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliary*

$\bar{\boldsymbol{h}}$: *true* regularity

$$\mathcal{R}(\lambda, \alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha) - \bar{\boldsymbol{h}}\right\|^2$$

Parameter tuning (Grid search)

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliary*

$\bar{\boldsymbol{h}}$: *true* regularity

$$\mathcal{R}(\lambda, \alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha) - \bar{\boldsymbol{h}}\right\|^2$$

Parameter tuning (Grid search)

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliary*

$\bar{\boldsymbol{h}}$: *true* regularity

$$\mathcal{R}(\lambda, \alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha) - \bar{\boldsymbol{h}}\right\|^2$$

$\bar{\boldsymbol{h}}$: unknown!

?

Parameter tuning (Grid search)

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

$\boldsymbol{h}$: *discriminant*, $\boldsymbol{v}$: *auxiliary*

$\bar{\boldsymbol{h}}$: *true* regularity

$$\mathcal{R}(\lambda, \alpha) = \left\|\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha) - \bar{\boldsymbol{h}}\right\|^2$$

$\bar{\boldsymbol{h}}$: unknown!

?

Stein Unbiased Risk Estimate (SURE)

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: truth and $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: truth and $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Parametric estimator $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

Ex. $$\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linear)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \, \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(nonlinear)} \end{cases}$$

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: truth and $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Parametric estimator $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

$$\textbf{Ex.} \quad \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linear)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \, \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(nonlinear)} \end{cases}$$

Quadratic error $R(\lambda) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 \stackrel{?}{=} \mathbb{E}_{\boldsymbol{\zeta}} \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ unknown

Stein Unbiased Risk Estimate (Principe)

Observations $\boldsymbol{y} = \bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P$, $\bar{\boldsymbol{x}}$: truth and $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$

Parametric estimator $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$

$$\textbf{Ex.} \quad \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) = \begin{cases} \left(\mathbf{I} + \lambda \mathbf{D}^\top \mathbf{D}\right)^{-1} \boldsymbol{y} & \text{(linear)} \\ \underset{\boldsymbol{x}}{\operatorname{argmin}} \, \|\boldsymbol{y} - \boldsymbol{x}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{x}) & \text{(nonlinear)} \end{cases}$$

Quadratic error $R(\lambda) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \bar{\boldsymbol{x}}\|^2 \stackrel{?}{=} \mathbb{E}_{\boldsymbol{\zeta}} \widehat{R}(\boldsymbol{y}; \lambda)$ $\bar{\boldsymbol{x}}$ unknown

Theorem (Stein, 1981)

Let $(\boldsymbol{y}; \lambda) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)$ an estimator of $\bar{\boldsymbol{x}}$

- weakly differentiable w.r.t. $\boldsymbol{y}$,
- such that $\boldsymbol{\zeta} \mapsto \langle \widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \boldsymbol{\zeta}; \lambda), \boldsymbol{\zeta} \rangle$ is integrable w.r.t. $\mathcal{N}(\mathbf{0}, \rho^2 \mathbf{I})$.

$$\widehat{R}(\boldsymbol{y}; \lambda) \triangleq \|\widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda) - \boldsymbol{y}\|^2 + 2\rho^2 \operatorname{tr}\left(\partial_{\boldsymbol{y}} \widehat{\boldsymbol{x}}(\boldsymbol{y}; \lambda)\right) - \rho^2 P$$
$$\Longrightarrow R(\lambda) = \mathbb{E}_{\boldsymbol{\zeta}}[\widehat{R}(\boldsymbol{y}; \lambda)].$$

Generalized *Stein Unbiased Risk Estimate*

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P$, $\bar{\boldsymbol{x}} \in \mathbb{R}^N$, $\boldsymbol{\Phi} : \mathbb{R}^{P \times N}$ and $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$

E.g. the estimators $\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha)$ with free or co-localized contours

$\log \boldsymbol{\mathcal{L}} = \boldsymbol{\Phi}(\bar{\boldsymbol{h}}, \bar{\boldsymbol{v}}) + \boldsymbol{\zeta}$ $\qquad$ $\boldsymbol{\zeta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\mathcal{S}})$ $\qquad$ $\mathcal{R} = \|\widehat{\boldsymbol{h}} - \bar{\boldsymbol{h}}\|^2$

$\boldsymbol{\Phi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$

$\boldsymbol{\Pi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto (\boldsymbol{h}, \boldsymbol{0})$

Projected estimation error $\quad R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2$

Observations $\boldsymbol{y} = \boldsymbol{\Phi}\bar{\boldsymbol{x}} + \boldsymbol{\zeta} \in \mathbb{R}^P$, $\bar{\boldsymbol{x}} \in \mathbb{R}^N$, $\boldsymbol{\Phi} : \mathbb{R}^{P \times N}$ and $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$

E.g. the estimators $\widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha)$ with free or co-localized contours

$\log \boldsymbol{\mathcal{L}} = \boldsymbol{\Phi}(\bar{\boldsymbol{h}}, \bar{\boldsymbol{v}}) + \boldsymbol{\zeta}$ $\qquad$ $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$ $\qquad$ $\mathcal{R} = \|\widehat{\boldsymbol{h}} - \bar{\boldsymbol{h}}\|^2$

$\boldsymbol{\Phi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto \{\log(a)\boldsymbol{h} + \boldsymbol{v}\}_a$

$\boldsymbol{\Pi} : (\boldsymbol{h}, \boldsymbol{v}) \mapsto (\boldsymbol{h}, \mathbf{0})$

Projected estimation error $\quad R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) \triangleq \mathbb{E}_{\boldsymbol{\zeta}} \|\boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda}) - \boldsymbol{\Pi}\bar{\boldsymbol{x}}\|^2$

Theorem (Pascal, 2020)

Let $(\boldsymbol{y}; \boldsymbol{\Lambda}) \mapsto \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})$ an estimator of $\bar{\boldsymbol{x}}$

- weakly differentiable w.r.t. $\boldsymbol{y}$,
- such that $\boldsymbol{\zeta} \mapsto \langle \boldsymbol{\Pi}\widehat{\boldsymbol{x}}(\bar{\boldsymbol{x}} + \boldsymbol{\zeta}; \lambda), \mathbf{A}\boldsymbol{\zeta} \rangle$ is integrable w.r.t. $\mathcal{N}(\mathbf{0}, \boldsymbol{\mathcal{S}})$.

$$\widehat{R}(\boldsymbol{\Lambda}) \triangleq \|\mathbf{A}(\boldsymbol{\Phi}\widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda}) - \boldsymbol{y})\|^2 + 2\mathrm{tr}\left(\boldsymbol{\mathcal{S}}\mathbf{A}^\top \boldsymbol{\Pi} \partial_{\boldsymbol{y}} \widehat{\boldsymbol{x}}(\boldsymbol{y}; \boldsymbol{\Lambda})\right) - \mathrm{tr}\left(\mathbf{A}\boldsymbol{\mathcal{S}}\mathbf{A}^\top\right)$$

$$\Longrightarrow R_{\boldsymbol{\Pi}}(\boldsymbol{\Lambda}) = \mathbb{E}_{\boldsymbol{\zeta}}[\widehat{R}(\boldsymbol{\Lambda})].$$

Parameter tuning (Grid search)

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

$\bar{\boldsymbol{h}}$: *true* regularity

$$\mathcal{R}(\lambda, \alpha) = \left\| \widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha) - \bar{\boldsymbol{h}} \right\|^2$$

$\bar{\boldsymbol{h}}$: unknown!

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha | \boldsymbol{\mathcal{S}})$$

Parameter tuning (Grid search)

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

$\bar{\boldsymbol{h}}$: *true* regularity

$$\mathcal{R}(\lambda, \alpha) = \left\| \widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha) - \bar{\boldsymbol{h}} \right\|^2$$

$\bar{\boldsymbol{h}}$: unknown!

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha | \boldsymbol{\mathcal{S}})$$

Parameter tuning (Grid search)

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

$\bar{\boldsymbol{h}}$: *true* regularity

$$\mathcal{R}(\lambda, \alpha) = \left\| \widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha) - \bar{\boldsymbol{h}} \right\|^2$$

$\bar{\boldsymbol{h}}$: unknown!

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha | \boldsymbol{\mathcal{S}})$$

Parameter tuning (Grid search)

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

$\bar{\boldsymbol{h}}$: *true* regularity

$$\mathcal{R}(\lambda, \alpha) = \left\| \widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha) - \bar{\boldsymbol{h}} \right\|^2$$

$\bar{\boldsymbol{h}}$: unknown!

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha | \boldsymbol{\mathcal{S}})$$

Parameter tuning (Automatic selection)

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h}, \boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,.} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

$\bar{\boldsymbol{h}}$: *true* regularity

$$\mathcal{R}(\lambda, \alpha) = \left\| \widehat{\boldsymbol{h}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha) - \bar{\boldsymbol{h}} \right\|^2$$

$\bar{\boldsymbol{h}}$: unknown!

$$\widehat{R}_{\nu,\boldsymbol{\varepsilon}}(\boldsymbol{\mathcal{L}}; \lambda, \alpha | \boldsymbol{\mathcal{S}})$$

Automated selection of regularization parameters

$$\left(\widehat{\boldsymbol{h}}, \widehat{\boldsymbol{v}}\right)(\boldsymbol{\mathcal{L}}; \lambda, \alpha) = \underset{\boldsymbol{h},\boldsymbol{v}}{\operatorname{argmin}} \sum_{a} \|\log \boldsymbol{\mathcal{L}}_{a,\cdot} - \log(a)\boldsymbol{h} - \boldsymbol{v}\|^2 + \lambda \mathcal{Q}(\mathbf{D}\boldsymbol{h}, \mathbf{D}\boldsymbol{v}; \alpha)$$

Example

$\widehat{\boldsymbol{h}}^{\mathsf{F}}(\boldsymbol{\mathcal{L}}; \lambda^{\dagger}, \alpha^{\dagger})$ (grid)

$\widehat{\boldsymbol{h}}^{\mathsf{F}}(\boldsymbol{\mathcal{L}}; \widehat{\lambda}^{\dagger}, \widehat{\alpha}^{\dagger})$ (grid)

$\widehat{\boldsymbol{h}}^{\mathsf{F}}(\boldsymbol{\mathcal{L}}; \widehat{\lambda}^{\mathsf{qN}}, \widehat{\alpha}^{\mathsf{qN}})$ (quasi-Newton)

225 calls of the estimator over the grid v.s. **40** for quasi-Newton

Take home messages

- Fractal texture model based on local *regularity* and *variance*
  - appropriate for real-world texture characterization
  - complementary attributes able to finely discriminate

Take home messages

- Fractal texture model based on local *regularity* and *variance*
  - appropriate for real-world texture characterization
  - complementary attributes able to finely discriminate
- Simultaneous estimation and regularization
  - significant decrease of the estimation error
  - accurate and regular contours thanks to co-localized penalization

Take home messages

- Fractal texture model based on local *regularity* and *variance*
  - appropriate for real-world texture characterization
  - complementary attributes able to finely discriminate
- Simultaneous estimation and regularization
  - significant decrease of the estimation error
  - accurate and regular contours thanks to co-localized penalization
- Fast algorithms for automated tuning of hyperparameters
  - possibility to manage huge amount of data
  - amenable to process data corrupted by *correlated* Gaussian noise
  - ensured objectivity and reproducibility

Take home messages

- Fractal texture model based on local *regularity* and *variance*
  - appropriate for real-world texture characterization
  - complementary attributes able to finely discriminate
- Simultaneous estimation and regularization
  - significant decrease of the estimation error
  - accurate and regular contours thanks to co-localized penalization
- Fast algorithms for automated tuning of hyperparameters
  - possibility to manage huge amount of data
  - amenable to process data corrupted by *correlated* Gaussian noise
  - ensured objectivity and reproducibility

MATLAB toolbox available: **github.com/bpascal-fr/gsugar**