Variational Image Analysis Using Wasserstein-Distance Based Priors

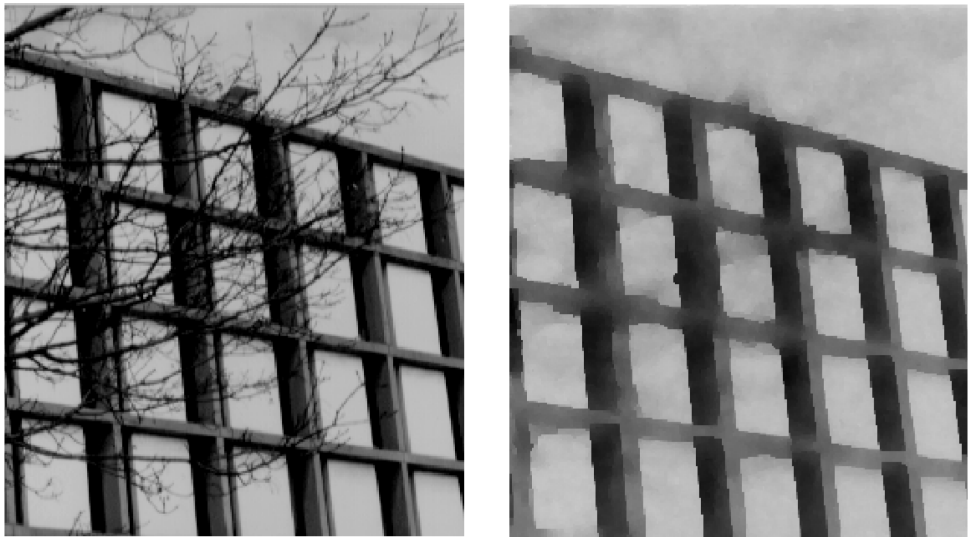
Christoph Schnörr, Paul Swoboda

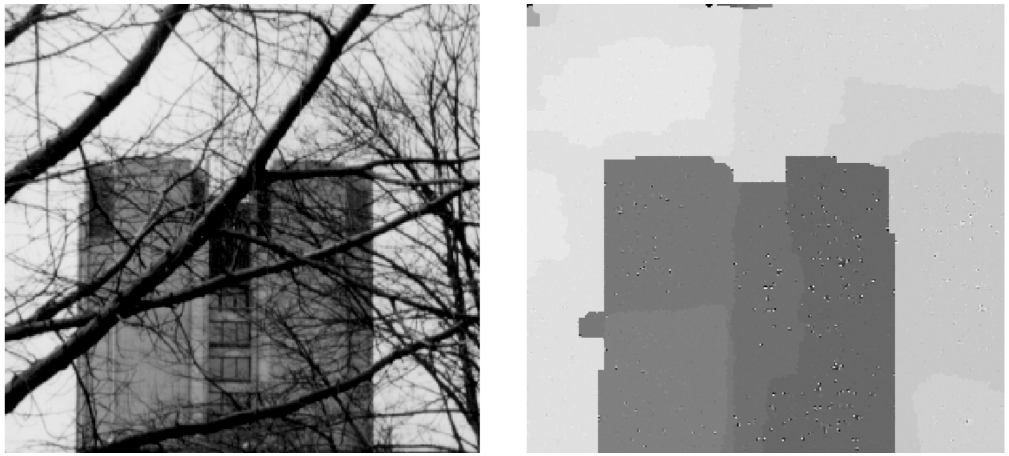
Image and Pattern Analysis Group
Heidelberg University

FGMIA
Jan 13-15, Paris

S.C. Zhu, D. Mumford: *Prior Learning and Gibbs Reaction Diffusion*

(PAMI 1997)





- generative probabilistic model
- *unsupervised* denoising by *pattern formation*

S.C. Zhu, D. Mumford: *Prior Learning and Gibbs Reaction Diffusion*

(PAMI 1997)

potentials, model parameters,
prior knowledge

filters, features

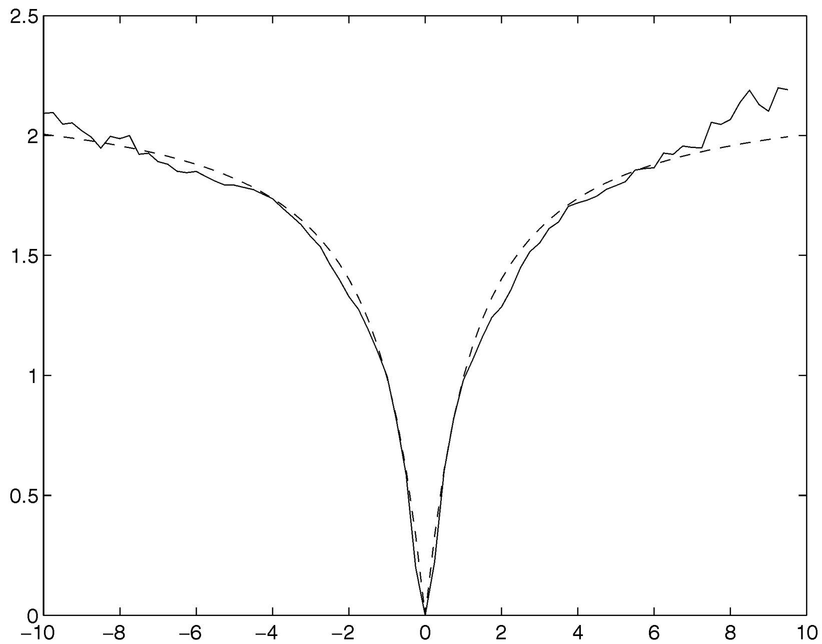
filter output statistics

$$U\left(\mathbf{I};\Lambda,S\right)=\sum_{\alpha=1}^{K}\sum_{(x,y)}\lambda^{(\alpha)}\left(\left(F^{(\alpha)}*\mathbf{I}\right)(x,y)\right)$$

functional

data

potential learned
for the Laplacian



Model Classes

Gibbs distributions

U cont. *quadratic* (Gaussian)

inference: linear system

$$p(I;\theta) \propto e^{-U(I;\theta)}$$

U *discrete* (e.g. Potts)

inference: LP relaxation, graph cuts

tractable

intractable

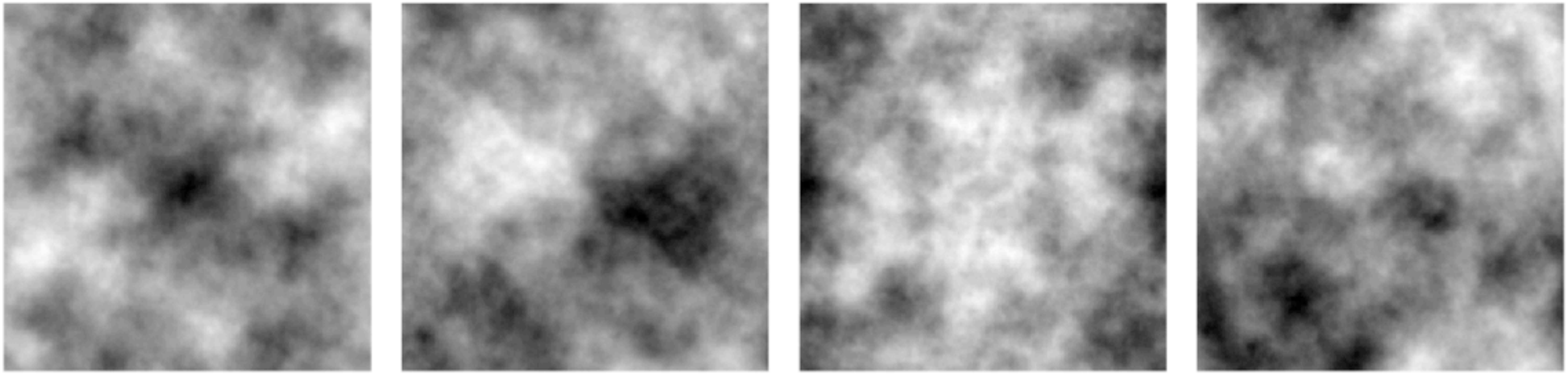
U cont. *non-convex* (reaction-diffusion)

inference: Gibbs sampler, MCMC

$$U(\mathbf{I};\Lambda,S) = \sum_{\alpha=1}^{K}\sum_{(x,y)} \lambda^{(\alpha)}\Big(\big(F^{(\alpha)} * \mathbf{I}\big)(x,y)\Big)$$

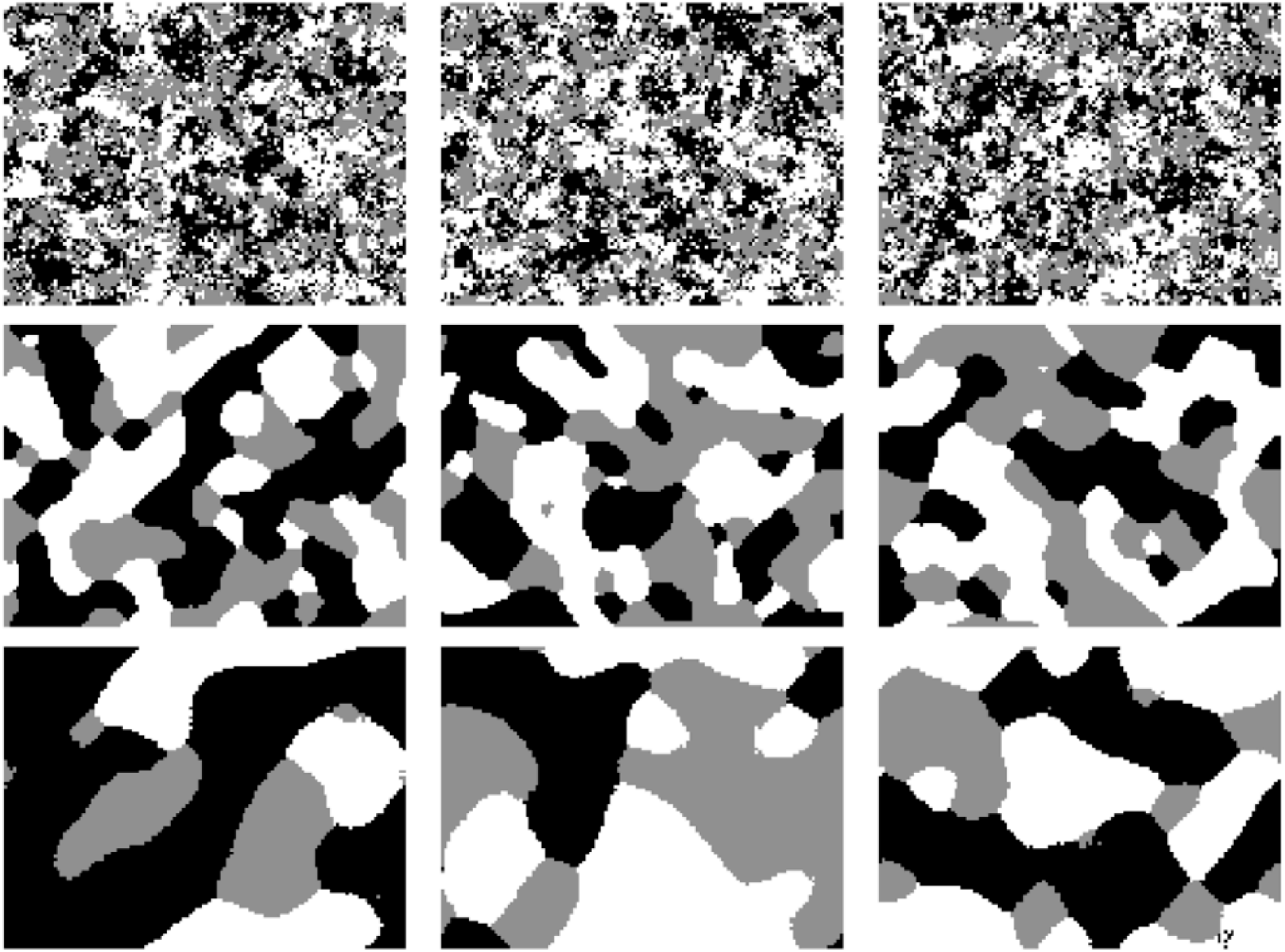
Graphical Models

Sampling a biharmonic Gaussian prior



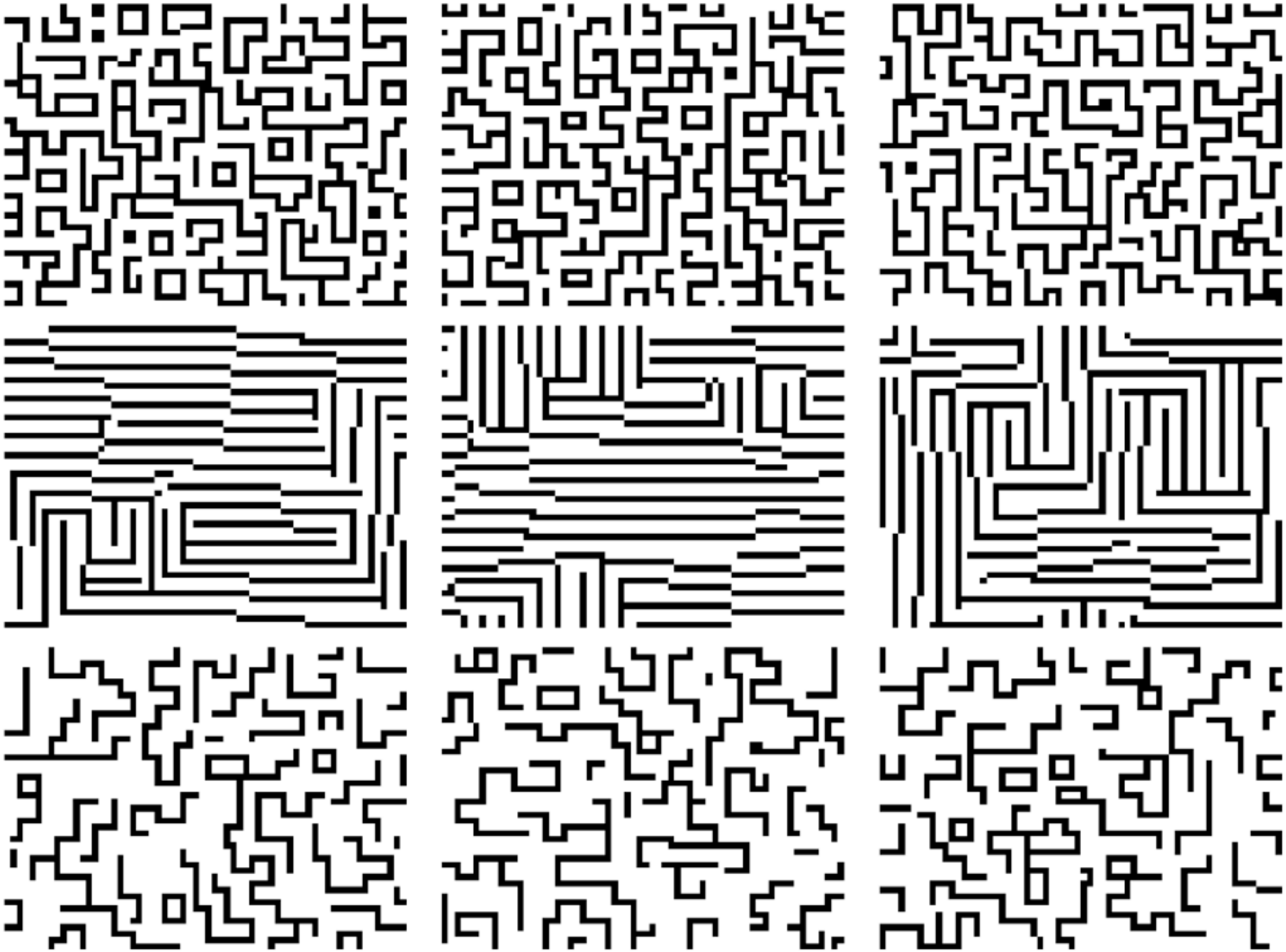
Graphical Models

Sampling a Potts prior



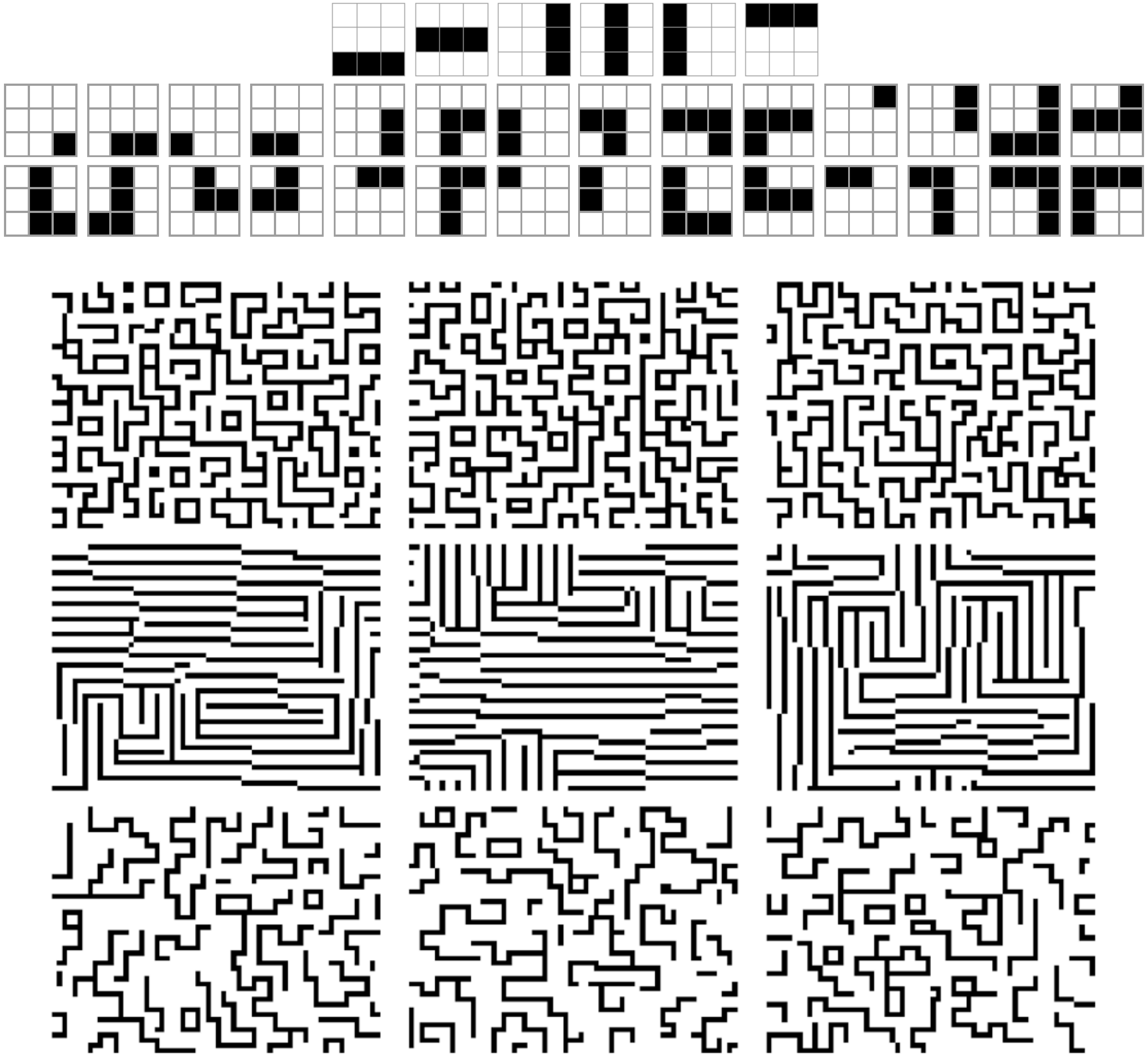
Graphical Models

Sampling a more involved discrete prior



Graphical Models

Interaction of patches - *implicit* knowledge representation



S.C. Zhu, D. Mumford: *Prior Learning and Gibbs Reaction Diffusion*

(PAMI 1997)

potentials, model parameters,
prior knowledge

filters, features

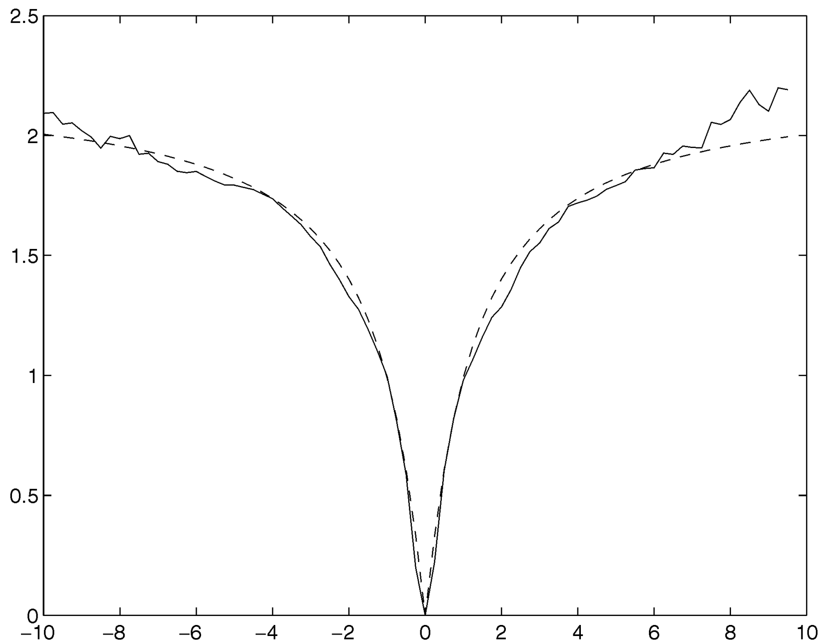
filter output statistics

$$U(\mathbf{I};\Lambda,S)=\sum_{\alpha=1}^{K}\sum_{(x,y)}\lambda^{(\alpha)}\Big(\big(F^{(\alpha)}*\mathbf{I}\big)(x,y)\Big)$$

functional

data

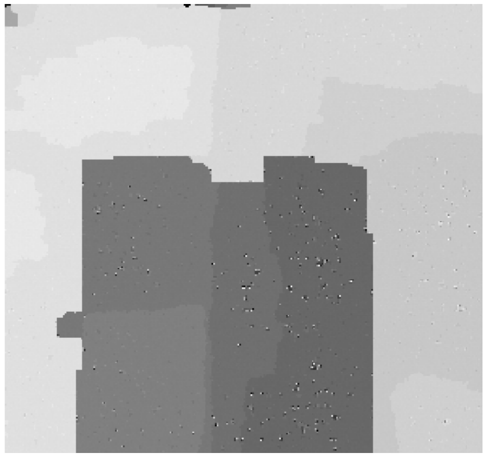
potential learned
for the Laplacian



S.C. Zhu, D. Mumford: *Prior Learning and Gibbs Reaction Diffusion*

(PAMI 1997)

- generative probabilistic model
- *unsupervised* denoising by *pattern formation*



This Talk: Focus

- *Continuous* image models
- *Explicit* knowledge: histograms

- *Tractable* Inference
- *Large-Scale* Convex Programming

- *Applications*
  Variational Restoration, Co- / Segmentation
  Unsupervised Inpainting

Outline

1. Convex Variational Restoration with *Histogram Priors*
   (Swoboda & Schnörr, *SIAM J. Imag. Sci. 2013)*

2. Convex Variational Co- / Segmentation with
   Histogram Priors (*Swoboda & Schnörr, EMMCVPR 2013*)

3. Outlook: Patch-Based Restoration and
   Unsupervised Inpainting (*in preparation*)

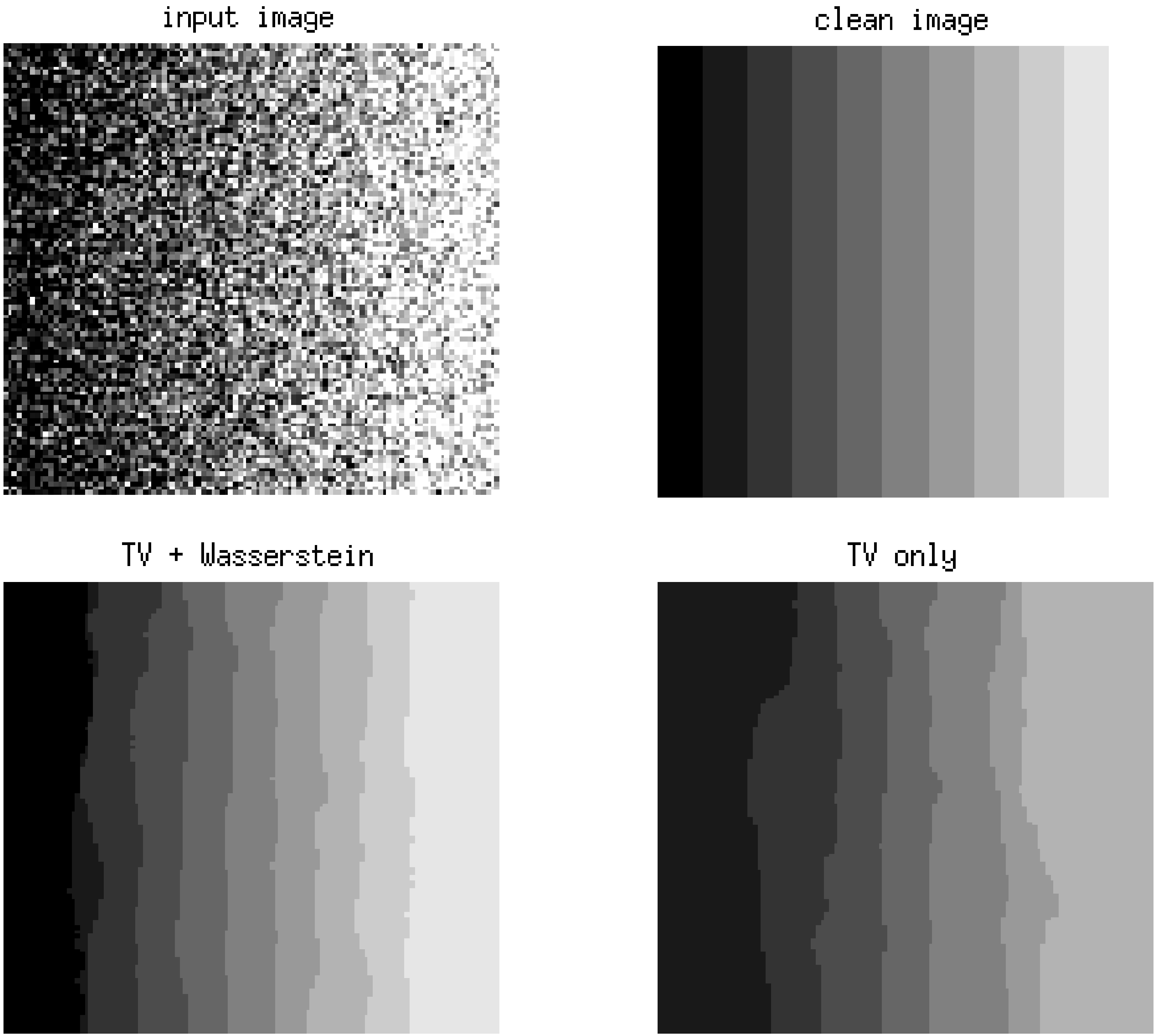
Outline

1. Convex Variational Restoration with *Histogram Priors*
   (Swoboda & Schnörr, *SIAM J. Imag. Sci. 2013)*

2. Convex Variational Co- / Segmentation with
   Histogram Priors (*Swoboda & Schnörr, EMMCVPR 2013*)

3. Outlook: Patch-Based Restoration and
   Unsupervised Inpainting (*in preparation*)

Problem, Variational Approach



Problem, Variational Approach

$$\inf_u F(u) + \lambda R(u) + \nu W(\mu^u, \mu^0)$$

data term (arbitrary)	convex spatial regularization	histogram prior

Problem, Variational Approach

$$\inf_u F(u) + \lambda R(u) + \nu W(\mu^u, \mu^0)$$

data term (arbitrary) | convex *spatial* regularization | *non-spatial* histogram prior

$$\min_{u \in \mathrm{BV}(\Omega,[0,1])} E(u) = \int_\Omega f(u(x), x) dx + \lambda \mathrm{TV}(u) + \nu W(\mu^u, \mu^0)$$

domain of u

$$W(\mu, \tilde{\mu}) = \inf_{\pi \in \Pi(\mu,\tilde{\mu})} \int_{[0,1]\times[0,1]} c(\gamma_1, \gamma_2)\; d\pi(\gamma_1, \gamma_2)$$

range of u

Contribution

- variational approach
- two convex relaxations, proof of equivalence
- convex programming approach
- numerical validation

Related Work

J. RABIN AND G. PEYRÉ, *Wasserstein regularization of imaging problem*, in Proceedings of ICIP, IEEE, Washington, DC, 2011, pp. 1541–1544.

contrast and color modification

non-convex approximation of the original energy

Convex Relaxation

based on

- duality (Wasserstein distance)
- functional lifting

alternatively

- Hoeffding-Fréchet bounds
- functional lifting

Convex Relaxation

based on

- duality (Wasserstein distance)
- functional lifting

alternatively

- Hoeffding-Fréchet bounds
- functional lifting

Duality

Wasserstein distance

$$W(\mu,\tilde{\mu}) = \inf_{\pi\in\Pi(\mu,\tilde{\mu})} \int_{[0,1]\times[0,1]} c(\gamma_1,\gamma_2)\ d\pi(\gamma_1,\gamma_2)$$

dual Kantorovich formulation

$$W(\mu^u,\mu^0) = \sup_{\substack{(\psi,\psi')\in L_1([0,1])^2 \\ \psi(\gamma_1)-\psi'(\gamma_2)\leq c(\gamma_1,\gamma_2)}} \int_0^1 \psi d\mu^u - \int_0^1 \psi' d\mu^0$$

Duality

Wasserstein distance

$$W(\mu, \tilde{\mu}) = \inf_{\pi \in \Pi(\mu, \tilde{\mu})} \int_{[0,1]\times[0,1]} c(\gamma_1, \gamma_2)\ d\pi(\gamma_1, \gamma_2)$$

dual Kantorovich formulation

$$\boxed{W(\mu^u, \mu^0)} = \sup_{\boxed{\substack{(\psi,\psi')\in L_1([0,1])^2 \\ \psi(\gamma_1)-\psi'(\gamma_2)\leq c(\gamma_1,\gamma_2)}}} \int_0^1 \psi d\mu^u - \int_0^1 \psi' d\mu^0$$

reformulation

$$E(u, \psi, \psi') = \int_\Omega f(u(x), x) dx + \lambda \mathrm{TV}(u) + \boxed{\nu \left(\int_0^1 \psi d\mu^u - \int_0^1 \psi' d\mu^0 \right)}$$

$$\inf_{u \in C} E(u) = \inf_{u \in C} \sup_{\boxed{(\psi,\psi')\in D}} E(u, \psi, \psi')$$

Reformulation

non-convexity

$$E(u,\psi,\psi') = \int_\Omega f(u(x),x)dx + \lambda \mathrm{TV}(u) + \nu \left(\int_0^1 \psi d\mu^u - \int_0^1 \psi' d\mu^0 \right)$$

$$\inf_{u\in C} E(u) = \inf_{u \in C} \sup_{(\psi,\psi')\in D} E(u,\psi,\psi')$$

$\Rightarrow$ convex relaxation

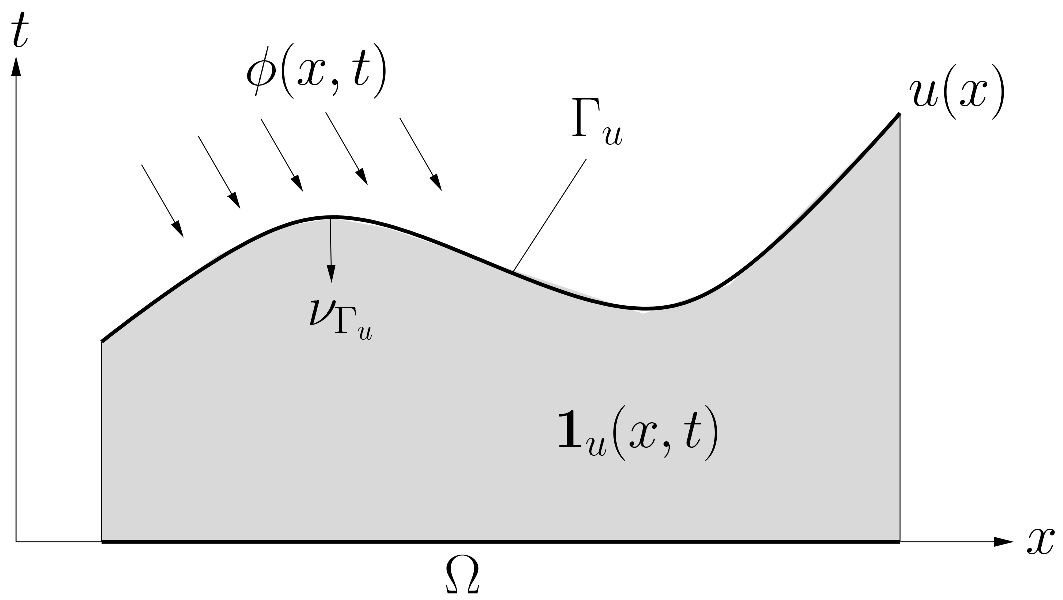
- duality (Wasserstein distance)
- functional lifting

- Hoeffding-Fréchet bounds
- functional lifting

Functional lifting

G. Alberti, G. Bouchitte, and G. Dal Maso, *The calibration method for the Mumford-Shah functional and free-discontinuity problems*, Calc. Var. Partial Differential Equations, 16 (2003), pp. 299–

T. Pock, D. Cremers, H. Bischof, and A. Chambolle, *Global solutions of variational models with convex regularization*, SIAM J. Imaging Sci., 3 (2010), pp. 1122–1145.



$$C' = \left\{ \phi \in \mathrm{BV}\left(\Omega \times \mathbb{R}, \{0,1\}\right) \ : \ \begin{array}{c} \phi(\cdot, (-\infty, 0]) \equiv 1, \ \phi(\cdot, [1, \infty)) \equiv 0, \\ D_\gamma \phi(\cdot, \gamma) \leq 0 \end{array} \right\}$$

$$-D_\gamma \phi = \mathcal{H}^2 \llcorner \mathrm{graph}(u)$$

Functional lifting

$$\mu^u(A) = \boxed{\mu^\phi(A) =} \frac{1}{|\Omega|}\int_\Omega |D_\gamma \phi(x,A)|dx = \boxed{\frac{1}{|\Omega|}\int_\Omega -D_\gamma \phi(x,A)dx}$$

$\phi \mapsto \mu^\phi$ is linear

Functional lifting

$$\mu^u(A) = \mu^\phi(A) = \frac{1}{|\Omega|}\int_\Omega |D_\gamma \phi(x,A)|dx = \frac{1}{|\Omega|}\int_\Omega -D_\gamma \phi(x,A)dx$$

$\phi \mapsto \mu^\phi$ is linear

$$E(u,\psi,\psi') = \int_\Omega f(u(x),x)dx + \lambda \mathrm{TV}(u) + \nu\left(\int_0^1 \psi d\mu^u - \int_0^1 \psi' d\mu^0\right)$$

$$\begin{aligned} E'(\phi,\psi,\psi') = & -\int_\Omega \int_0^1 f(\gamma,x) D_\gamma \phi(x,\gamma)\, dx + \lambda \int_0^1 \mathrm{TV}(\phi(\cdot,\gamma))d\gamma \\ & + \nu\left(\int_0^1 \psi d\mu^\phi - \int_0^1 \psi' d\mu^0\right). \end{aligned}$$

(coarea formula)

Functional lifting & Convex Relaxation

$$\mu^u(A) = \mu^\phi(A) = \frac{1}{|\Omega|}\int_\Omega |D_\gamma \phi(x,A)|dx = \frac{1}{|\Omega|}\int_\Omega -D_\gamma \phi(x,A)dx$$

$\phi \mapsto \mu^\phi$ is linear

$$E(u,\psi,\psi') = \int_\Omega f(u(x),x)dx + \lambda \mathrm{TV}(u) + \nu\left(\int_0^1 \psi d\mu^u - \int_0^1 \psi' d\mu^0\right)$$

$$E'(\phi,\psi,\psi') = \begin{array}{l} -\int_\Omega \int_0^1 f(\gamma,x) D_\gamma \phi(x,\gamma) dx + \lambda \int_0^1 \mathrm{TV}(\phi(\cdot,\gamma))d\gamma \\ +\nu\left(\int_0^1 \psi d\mu^\phi - \int_0^1 \psi' d\mu^0\right). \end{array}$$

(coarea formula)

Thm.: existence

$$\inf_{u\in C} \sup_{(\psi,\psi')\in D} E(u,\psi,\psi') = \inf_{\phi\in C'} \sup_{(\psi,\psi')\in D} E'(\phi,\psi,\psi')$$

relax

Reformulation

non-convexity

$$E(u,\psi,\psi') = \int_\Omega f(u(x),x)dx + \lambda \mathrm{TV}(u) + \nu \left(\int_0^1 \psi d\mu^u - \int_0^1 \psi' d\mu^0 \right)$$

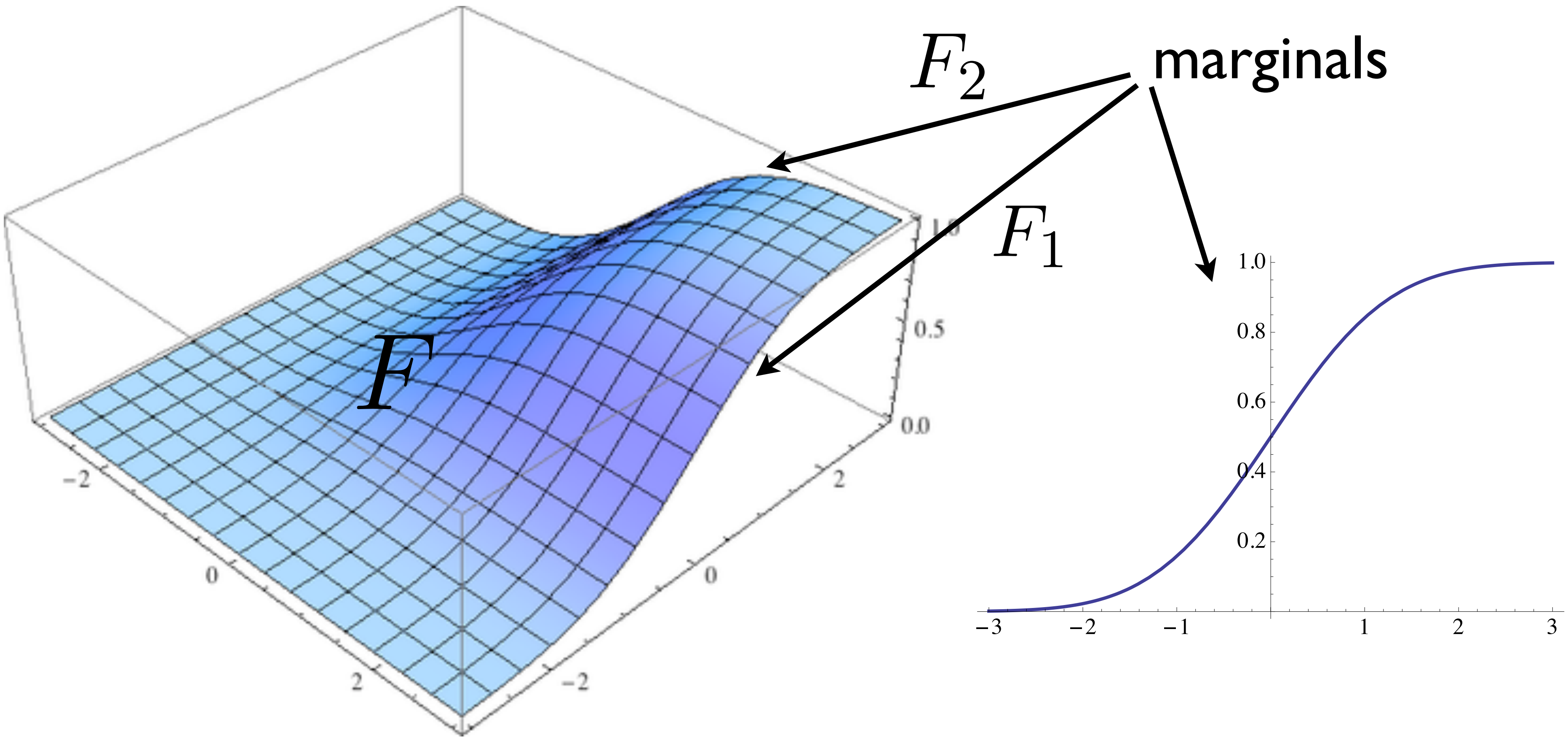
$$\inf_{u \in C} E(u) = \inf_{u \in C} \sup_{(\psi,\psi') \in D} E(u,\psi,\psi')$$

$\Rightarrow$ convex relaxation

- duality (Wasserstein distance)
- functional lifting

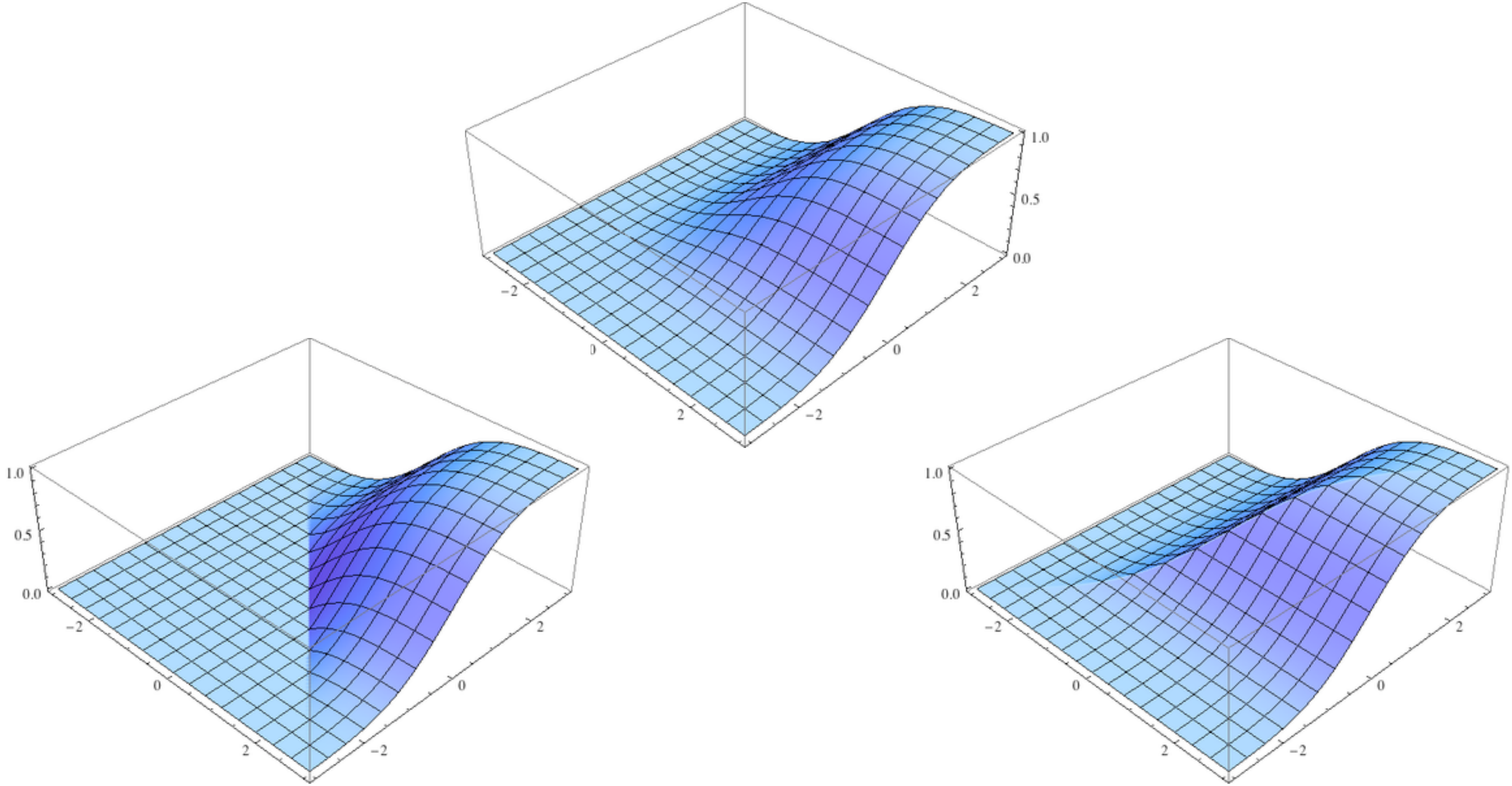
- Hoeffding-Fréchet bounds
- functional lifting

Hoeffding-Fréchet Bounds



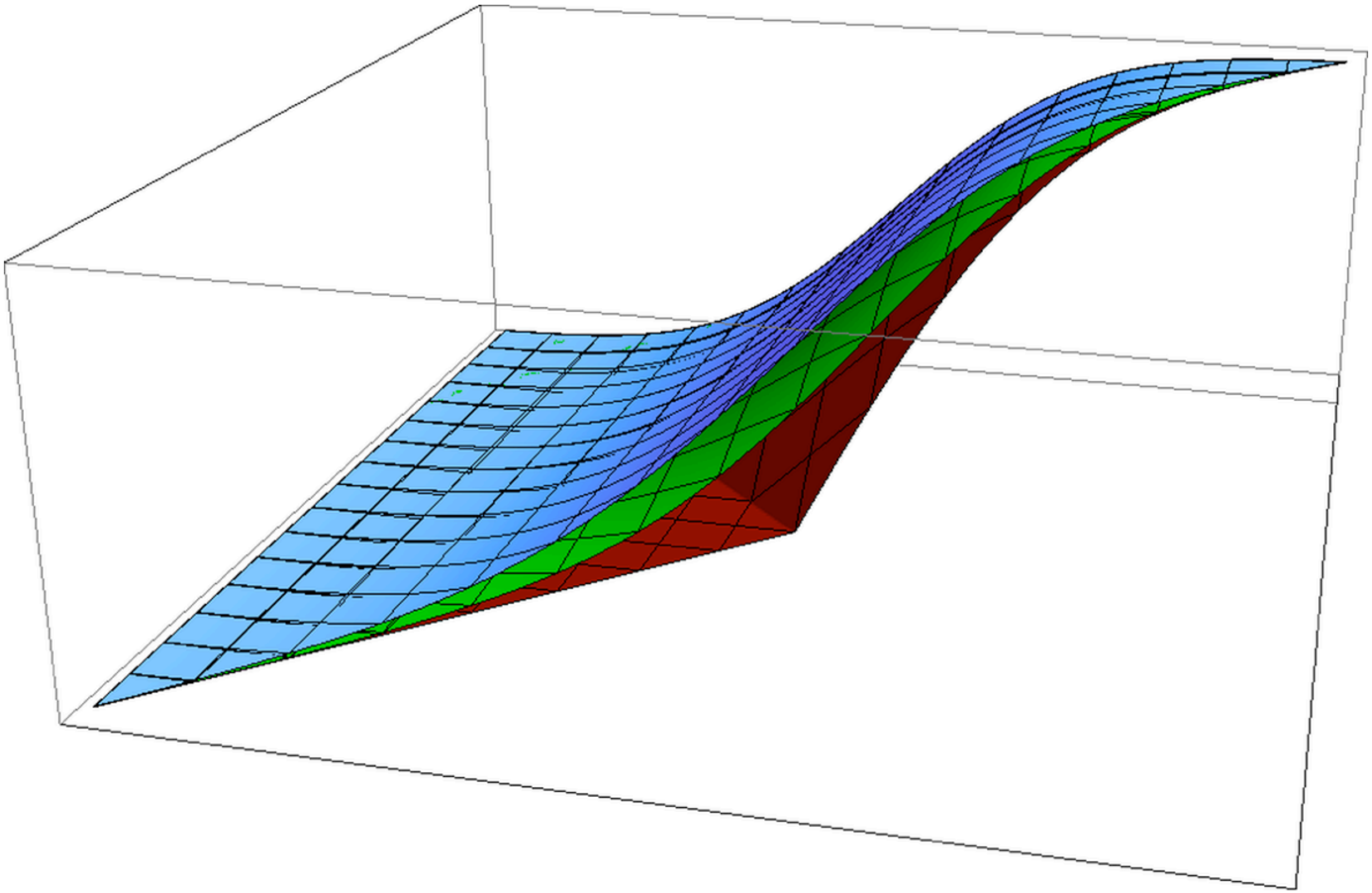
$$(F_1(\gamma_1) + F_2(\gamma_2) - 1)_+ \leq F(\gamma_1, \gamma_2) \leq \min\{F_1(\gamma_1), F_2(\gamma_2)\}$$

Hoeffding-Fréchet Bounds



$$(F_1(\gamma_1) + F_2(\gamma_2) - 1)_+ \leq F(\gamma_1, \gamma_2) \leq \min\{F_1(\gamma_1), F_2(\gamma_2)\}$$

Hoeffding-Fréchet Bounds



$$(F_1(\gamma_1) + F_2(\gamma_2) - 1)_+ \leq F(\gamma_1, \gamma_2) \leq \min\{F_1(\gamma_1), F_2(\gamma_2)\}$$

Functional Lifting & Convex Relaxation

$$
\begin{aligned}
\min_{\phi, F} \quad & \int_\Omega \int_0^1 -f(\gamma, x) D_\gamma \phi(x, \gamma) dx + \lambda \int_0^1 \mathrm{TV}(\phi(\cdot, \gamma)) d\gamma + \nu \boxed{\int_{\mathbb{R}^2} c \; dF}, \\
\text{s.t.} \quad & F_\phi(\gamma) = \frac{1}{|\Omega|} \int_\Omega -D_\gamma \phi(x, [0, \gamma]) dx, \\
& F_{\mu^0}(\gamma) = \mu^0([0, \gamma]), \\
& \boxed{F_\phi(x_1) + F_{\mu^0}(x_2) - 1 \leq F(x_1, x_2) \leq \min\{F_\phi(x_1), F_{\mu^0}(x_2)\}}, \\
& \phi \in C''.
\end{aligned}
$$

Convex Relaxations

- duality (Wasserstein distance)
- functional lifting

$\Updownarrow$ (proof: paper)

- Hoeffding-Fréchet bounds
- functional lifting

Convex Programming

Discretized problem

$$\min_{\phi \in C''_d} \boxed{\max_{(\psi,\psi') \in D_d} E'_d(\phi, \psi, \psi')}$$

based on

$$E'(\phi,\psi,\psi') = \begin{array}{l} -\int_\Omega \int_0^1 f(\gamma,x) D_\gamma \phi(x,\gamma)\, dx + \lambda \int_0^1 \mathrm{TV}(\phi(\cdot,\gamma)) d\gamma \\ + \boxed{\nu \left(\int_0^1 \psi d\mu^\phi - \int_0^1 \psi' d\mu^0 \right)}. \end{array}$$

→ generalized FB-splitting wrt. $\boxed{\ldots}$

H. Raguet, J. Fadili, and G. Peyré, *A generalized forward-backward splitting*, SIAM J. Imaging Sci., 6 (2013), pp. 1199–1226.

⇒ single *non*-standard proximal mapping only wrt. $\boxed{\ldots}$

Convex Programming

Proximal mapping with $c(\gamma_1, \gamma_2) = |\gamma_1 - \gamma_2|$

$$\overline{\phi} = \underset{\phi}{\mathrm{argmin}} \frac{1}{2} \|\phi - \tilde{\phi}\|^2 + \lambda W(\mu^{\phi}, \mu^0)$$

$$\overline{\phi}(x, \gamma) = \tilde{\phi}(x, \gamma) + \boxed{c_\gamma}$$

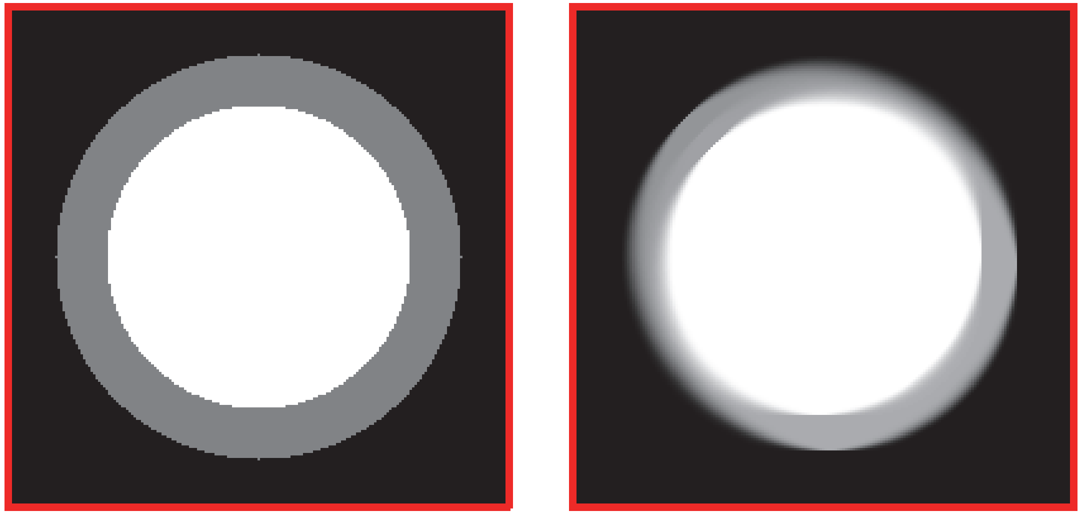
$$\boxed{c_\gamma} = \mathrm{shrink}\Big(-\frac{1}{|\Omega|} \int_\Omega \tilde{\phi}(x, \gamma) dx - F_{\mu^0}(\gamma) + 1, \frac{\lambda}{|\Omega|} \Big) + \frac{1}{|\Omega|} \int_\Omega \tilde{\phi}(x, \gamma) dx + F_{\mu^0}(\gamma) - 1$$

Numerical Validation

biased inpainting

Numerical Validation

example:
non-tight relaxation



Numerical Validation

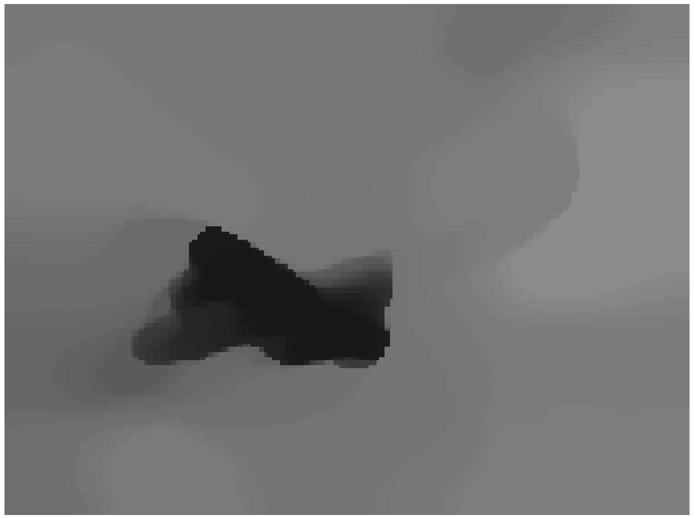
Histogram constrained denoising

contrast
better preserved

TV only

Numerical Validation

Unsupervised inpainting



TV only

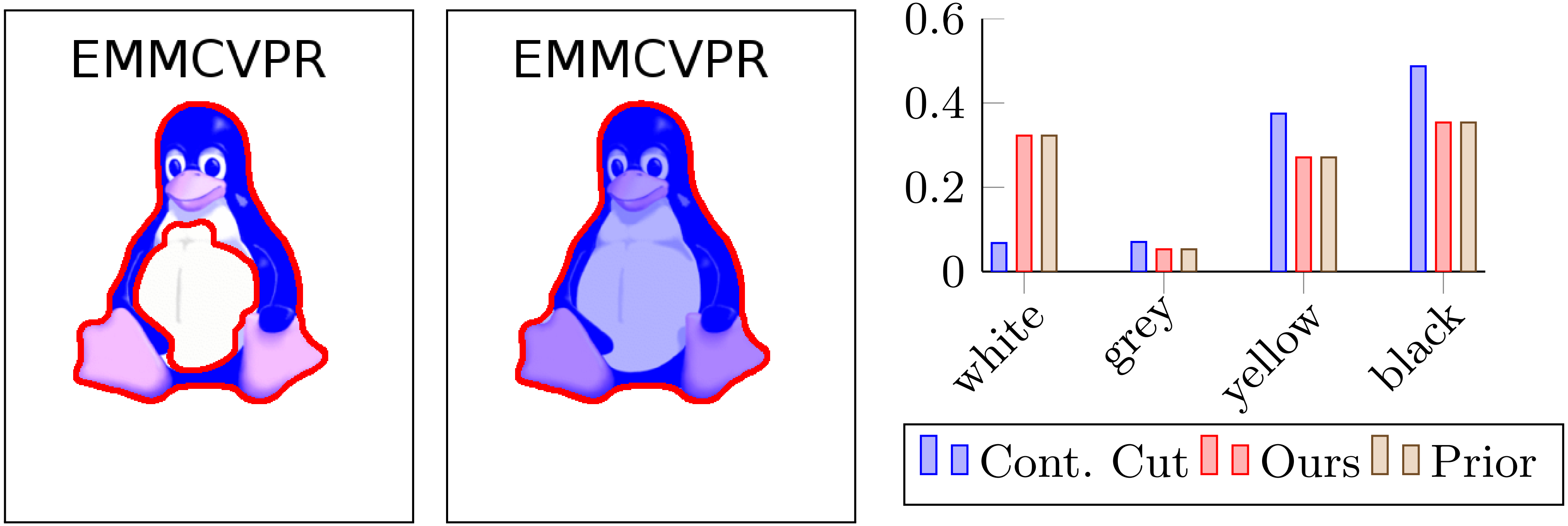
Outline

1. Convex Variational Restoration with *Histogram Priors*
   (Swoboda & Schnörr, *SIAM J. Imag. Sci. 2013)*

2. Convex Variational Co- / Segmentation with
   Histogram Priors (*Swoboda & Schnörr, EMMCVPR 2013*)

3. Outlook: Patch-Based Restoration and
   Unsupervised Inpainting (*in preparation*)

Problem, Variational Approaches



(a) Cont. Cut (b) Ours (c) Foreground histograms

Local potentials $d^i(x) = -\log(p^i(I(x)))$

\+ spatial regularization may not suffice

⇒ Histogram prior *globally depending* on the segmentation

Problem, Variational Approaches

Co-segmenting similar objects in *different* scenes
without prior knowledge (appearance)

⇒ Wasserstein distance between foreground histograms

that *globally depend* on and are inferred with the segmentation

Related Prior Work

includes

T. F. Chan, S. Esedoglu, and K. Ni. Histogram Based Segmentation Using Wasserstein Distances. In Fiorella Sgallari, Almerico Murli, and Nikos Paragios, editors, *SSVM*, volume 4485 of *Lecture Notes in Computer Science*, pages 697–708. Springer, 2007.

potentials based on *local* Wasserstein-distances

C. Rother, T. Minka, A. Blake, and V. Kolmogorov. Cosegmentation of image pairs by histogram matching - incorporating a global constraint into mrfs. In *CVPR*, pages 993–1000, Washington, DC, USA, 2006. IEEE.

S. Vicente, V. Kolmogorov, and C. Rother. Cosegmentation revisited: models and optimization. In *Proceedings of the 11th European conference on Computer vision: Part II*, ECCV'10, pages 465–479, Berlin, Heidelberg, 2010. Springer-Verlag.

less general histogram mtaching, EM-like alternating schemes, local minima issues

---

Our contribution: *single convex* variational approaches for both co-/segmentation with global histogram priors

Problem, Variational Approaches

Segmentation

global data term

histogram priors

$$E_{seg}(\Omega_1, \dots, \Omega_k) = \frac{1}{2}\sum_{i=1}^{k} \mathrm{Per}(\Omega_i, \Omega) + \sum_{i=1}^{k} W\left(\mu^I_{\Omega_i}, |\Omega_i| \cdot \mu^i\right)$$

dependency on the segmentation

Problem, Variational Approaches

Segmentation

global data term

histogram priors

$$E_{seg}(\Omega_1, \ldots, \Omega_k) = \frac{1}{2}\sum_{i=1}^{k} \mathrm{Per}(\Omega_i, \Omega) + \sum_{i=1}^{k} W\left(\mu^I_{\Omega_i}, |\Omega_i| \cdot \mu^i\right)$$

dependency on the segmentation

Co-segmentation

compare
foreground statistics

enforce decision

$$E_{coseg}(\Omega_1, \Omega_2) = \sum_{i=1}^{2} \mathrm{Per}(\Omega_i, \Omega) + W\left(\mu^{I_1}_{\Omega_1}, \mu^{I_2}_{\Omega_2}\right) + \sum_{i=1}^{2} P \cdot |\Omega \backslash \Omega_i|$$

dependency on the segmentation

Problem, Variational Approaches

Segmentation

$$E_{seg}(\Omega_1, \ldots, \Omega_k) = \frac{1}{2}\sum_{i=1}^{k} \mathrm{Per}(\Omega_i, \Omega) + \sum_{i=1}^{k} W\left(\mu^I_{\Omega_i}, |\Omega_i| \cdot \mu^i\right)$$

Co-segmentation

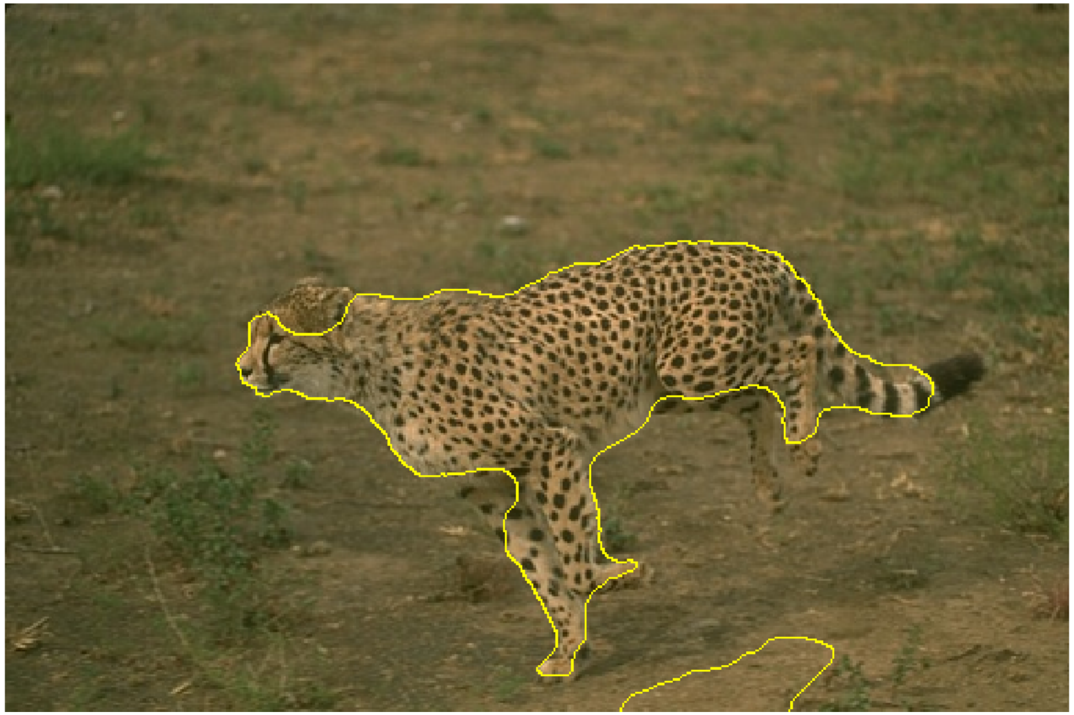
$$E_{coseg}(\Omega_1, \Omega_2) = \sum_{i=1}^{2} \mathrm{Per}(\Omega_i, \Omega) + W\left(\mu^{I_1}_{\Omega_1}, \mu^{I_2}_{\Omega_2}\right) + \sum_{i=1}^{2} P \cdot |\Omega \backslash \Omega_i|$$

Contribution

- representation via indicator functions
- convex relaxation
- problem decomposition, efficient prox-map (W-distance)
- numerical validation

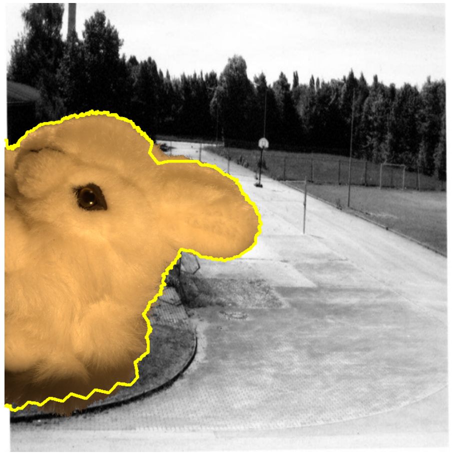
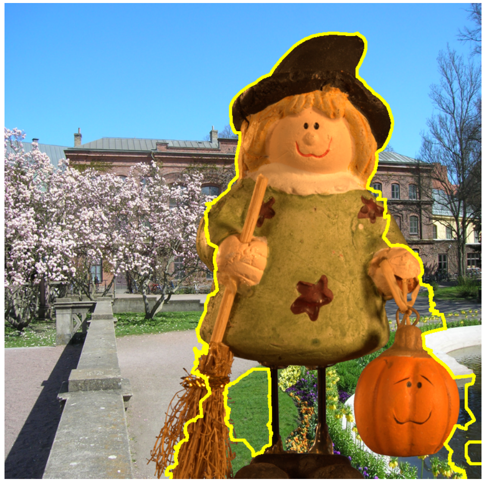
Numerical Validation

Prior: extrapolating simple color histograms



Note: this is a *global* optimum (inference).
Undesired partitions are due to features.

Numerical Validation


Outline

1. Convex Variational Restoration with *Histogram Priors*
   (*SIAM J. Imag. Sci. 2013*)

2. Convex Variational Co- / Segmentation with
   Histogram Priors

3. Outlook: Patch-Based Restoration and
   Unsupervised Inpainting

Variational Image Analysis Using Wasserstein-Distance Based Priors

Christoph Schnörr, Paul Swoboda

Image and Pattern Analysis Group
Heidelberg University

FGMIA
Jan 13-15, Paris