Non Convex Minimization using Convex Relaxation
Some Hints to Formulate Equivalent Convex Energies

Mila Nikolova (CMLA, ENS Cachan, CNRS, France)

SIAM Imaging Conference (IS14) Hong Kong
Minitutorial: May 13, 2014

Outline

1. Energy minimization methods

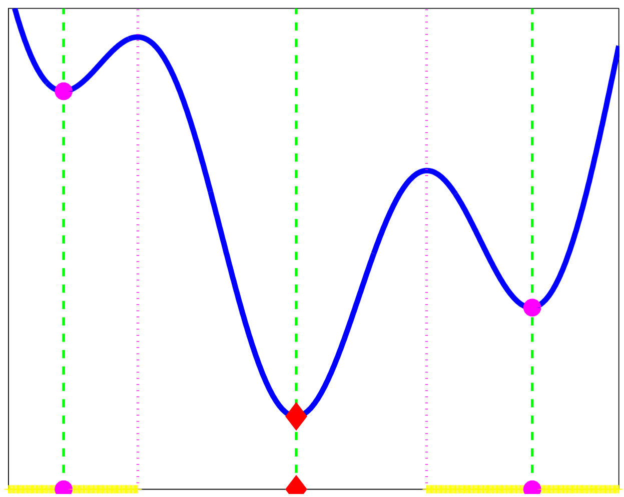
In many imaging problems the sought-after image $\widehat{u} : \Omega \to \mathbb{R}^k$ is defined by

$$\widehat{\boldsymbol{u}} = \arg\min_{\boldsymbol{u}} \boldsymbol{\mathcal{E}}(\boldsymbol{u}) \quad \text{for} \quad \mathcal{E}(u) := \Psi(u, f) + \lambda\Phi(u) + \imath_S(u) \quad \lambda > 0$$

f given image, Ψ data fidelity, Φ regularization, S set of constraints, $\imath$ indicator function ($\imath_S(u) = 0$ if $u \in S$ and $\imath_S(u) = +\infty$ otherwise)

- Often $u \mapsto \mathcal{E}(u)$ is nonconvex

Algorithms easily get trapped in local minima



How to find a global minimizer? Many algorithms, usually suboptimal.

Some famous nonconvex problems for labeling and segmentation

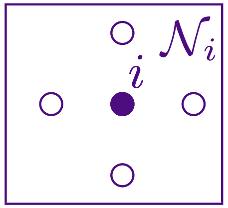
Potts model [Potts 52] (ℓ_0 semi-norm applied to differences):

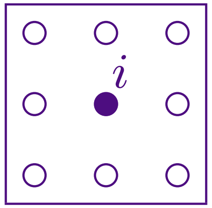
$$\mathcal{E}(u) = \Psi(u, f) + \lambda \sum_{i,j} \phi(u[i] - u[j]) \quad \phi(t) := \begin{cases} 0 & \text{if} \quad t = 0 \\ 1 & \text{if} \quad t \neq 0 \end{cases}$$

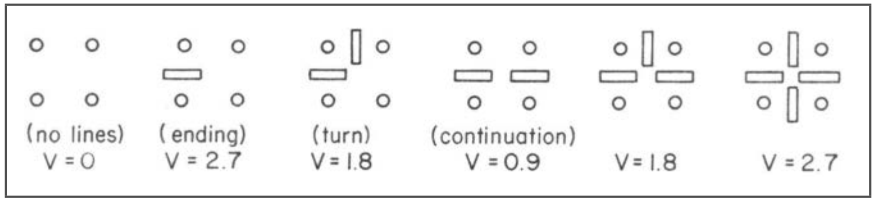
Line process in Markov random field priors [Geman, Geman 84]: $(\hat{u}, \hat{\ell}) = \arg\min_{u,\ell} \mathcal{F}(u, \ell)$

$$\mathcal{F}(u, \ell) = \|A(u) - f\|_2^2 + \lambda \sum_i \Big(\sum_{j \in \mathcal{N}_i} \varphi(u[i] - u[j])(1 - \ell_{i,j}) + \sum_{(k,n) \in \mathcal{N}_{i,j}} \mathrm{V}(\ell_{i,j}, \ell_{k,n}) \Big)$$

$[\ell_{i,j} = 0 \;\Leftrightarrow\;$ no edge$]$, $\quad [\ell_{i,j} = 1 \;\Leftrightarrow\;$ edge between i and $j]$







M.-S. functional [Mumford, Shah 89]:

$$\mathcal{F}(u, L) = \int_\Omega (u - v)^2 dx + \lambda \left(\int_{\Omega \setminus L} \|\nabla u\|^2 dx + \alpha \mid L \mid \right) \qquad \mid L \mid = \text{length}(L)$$

GEMAN AND GEMAN: STOCHASTIC RELAXATION, GIBBS DISTRIBUTIONS, AND BAYESIAN RESTORATION

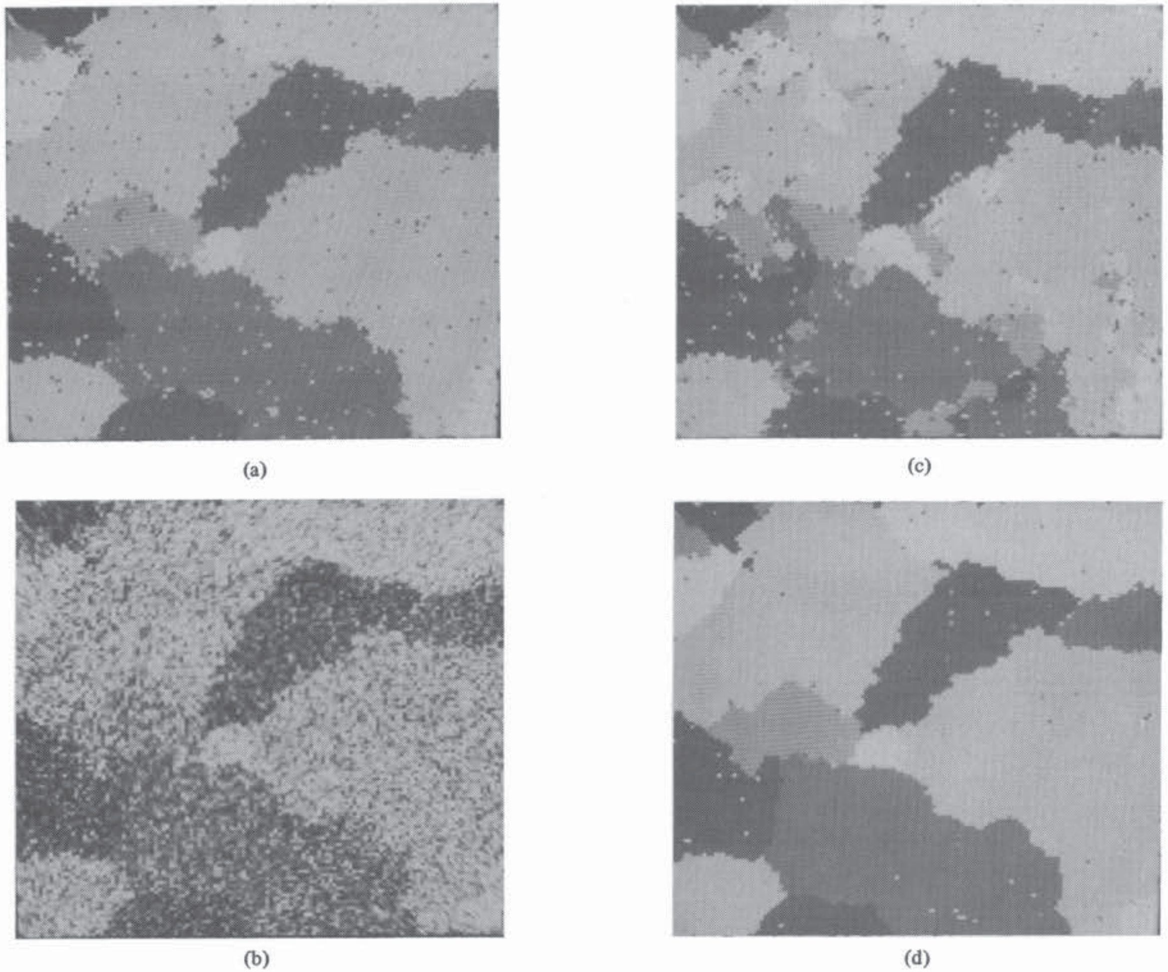


Fig. 2. (a) Original image: Sample from MRF. (b) Degraded image: Additive noise. (c) Restoration: 25 iterations. (d) Restoration: 300 iterations.

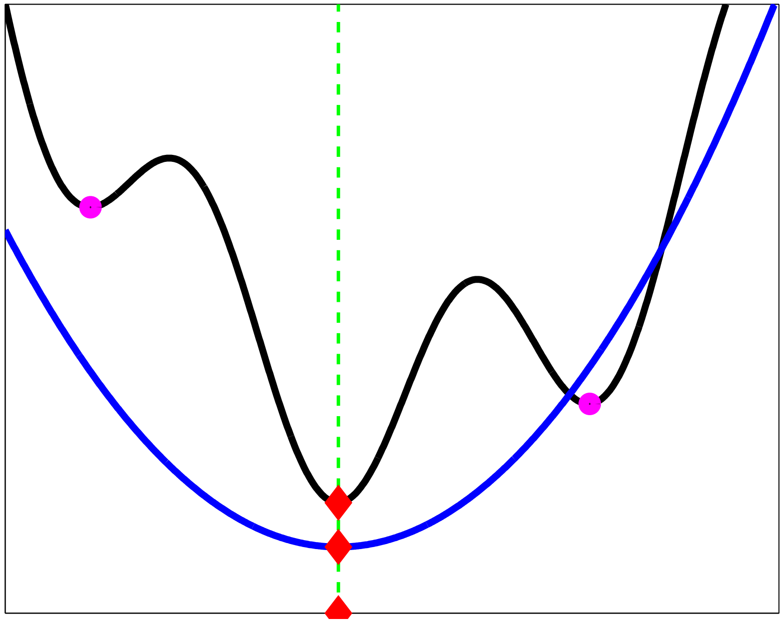
Image credits: S. Geman and D. Geman 1984. Restoration with 5 labels using Gibbs sampler

"We make an analogy between images and statistical mechanics systems. Pixel gray levels and the presence and orientation of edges are viewed as states of atoms or molecules in a lattice-like physical system. The assignment of an energy function in the physical system determines its Gibbs distribution. Because of the Gibbs distribution, Markov random field (MRF) equivalence, this assignment also determines an MRF image model." [S. Geman, D. Geman 84]

A perfect bypass: Find another functional $\mathcal{F} : \Omega \to \mathbb{R}$, easy to minimize, such that

$$\arg\min_u \mathcal{F}(u) \subseteq \arg\min_u \mathcal{E}(u)$$

e.g., $\mathcal{F}$ is convex and coercive.



- Subtle and case-dependent.
- We are in the inception phase...

Finding a globally optimal solution to a hard problem by conceiving **another** problem having the same set of optimal solutions and easy to solve has haunted researchers for a long time.

- The Weiszfeld algorithm: E. Weiszfeld, Sur le point pour lequel la somme des distances de n points données est minimum," Tôhoku Mathematical Journal, vol. 43, pp. 355–386, 1937.

  The word **algorithm** was unknown to most mathematicians by 1937.

  The Weiszfeld algorithm has extensively been used (e.g., in economics) when computers were available.

- G. Dantzig, R. Fulkerson and S. Johnson, "Solution of a large-scale traveling-salesman problem", Operations Research, vol. 2, pp. 393–410, 1954

- R. E. Gomory, "Outline of an algorithm for integer solutions to linear programs" Bull. Amer. Math. Soc., 64(5), pp. 217–301, 1958.

(Tight) convex relaxation is only one somehow "secured" way to tackle hard minimization problems. This talk focuses on convex relaxations for imaging applications.

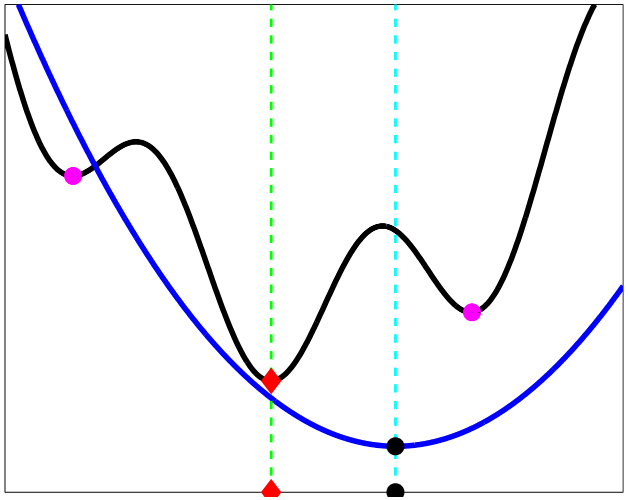
– Discrete setting – MRF – geometry of images may be difficult to handle.

– Continuous setting – in general more accurate approximations can be derived.

Experimental comparison of discrete and continuous shape optimization – [Klodt et al, 2008]

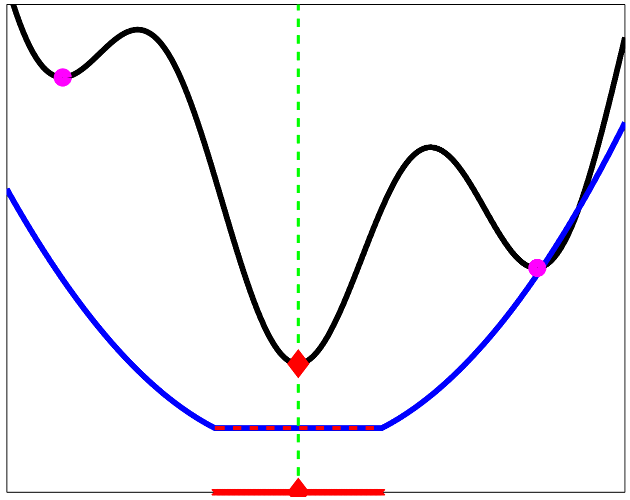
Applications in imaging: image restoration, image segmentation, disparity estimation of stereo images, depth map estimation, optical flow estimation, (multi) labeling problems, among many others.

Loose convex relaxation



No way to get $\widehat{u}$

Often in practice



How to get $\widehat{u}$?

- In practice

$$\boldsymbol{\arg\min_{u}\ \mathcal{E}(u) \subseteq \arg\min_{u}\ \mathcal{F}(u)}$$

- Convex relaxation is tight in each of the cases
  - $\arg\min\limits_{u}\ \mathcal{E}(u) \supseteq \arg\min\limits_{u}\ \mathcal{F}(u)$
  - we know how to reach $\widehat{u} \in \arg\min\limits_{u}\ \mathcal{E}(u)$ from $\widetilde{u} \in \arg\min\limits_{u}\ \mathcal{F}(u)$

We will explain how several successful convex relaxations have been obtained.

We will exhibit some limits of the approach.

Notation

- Image domain and derivatives
  - $\Omega \subset \mathbb{R}^2$ continuous setting, Du is the (distributional) derivative of u;
  - $\Omega = h\{1, \cdots, M\} \times h\{1, \cdots, N\}$ grid with step h, Du is a set difference operators

$x = (x_1, x_2) \in \Omega$

- $\boldsymbol{\{u > t\} := \{x \in \Omega \ : \ u(x) > t\}}$ the super-levels of u
- $\Sigma \subset \Omega$ (in general non connected) $\partial\Sigma$ is its boundary in Ω and $\mathrm{Per}(\Sigma)$ its perimeter
- $\mathbb{1}_\Sigma(x) = \begin{cases} 1 & \text{if} \quad x \in \Sigma \\ 0 & \text{otherwise} \end{cases}$ the characteristic function of Σ
- $\imath_\Sigma(x) = \begin{cases} 0 & \text{if} \quad x \in \Sigma \\ +\infty & \text{otherwise} \end{cases}$ the indicator function of Σ
- $\mathrm{supp}\,(u) := \{x \in \Omega \ : \ u(x) \neq 0\}$
- $BV(\Omega)$ – the set of all functions of bounded variation defined on Ω

Useful formulas

◇ $u \in BV(\Omega)$

- Coarea formula $\mathrm{TV}(u) = \int \|Du\|\,dx = \int_{-\infty}^{+\infty} \mathrm{Per}(\{x \,:\, u(x) > t\})\,dt$ (coa)

  $\mathrm{Per}(\Sigma) = \mathrm{TV}(\mathbb{1}_\Sigma)$ (per)

- Layer-cake formulas
  - $u(x) = \int_{-\infty}^{+\infty} \mathbb{1}_{\{x \,:\, u(x)>t\}}(x)\,dt$ (cake)
  - $\|u-f\|_1 = \int_{-\infty}^{+\infty} \big|\{x \,:\, u(x) > t\} \bigtriangleup \{x \,:\, f(x) > t\}\big|\,dt$ (cake1)

    $\bigtriangleup$ symmetric difference [T.Chan, Esedoglu 05], [T. Chan, Esedoglu, Nikolova 06]

◇ V is a normed vector space, V^* its dual and $F : V \to \mathbb{R}$ is proper

- The convex conjugate of F is $F^*(v) := \sup_{u \in V} \big\{\langle u, v\rangle - F(u)\big\} \quad v \in V^*$ (cc)

2. Simple Convex Binary Labeling / Restoration [T. Chan, Esedoglu, Nikolova 06]

Given a binary input image $f = 1\!\!1_\Sigma$, we are looking for a binary $\widehat{u}(x) = 1\!\!1_{\widehat{\Sigma}}(x)$

Constraint : $u(x) = 1\!\!1_\Sigma(x)$ [Vese, Osher 02]

$$\mathcal{E}(u) = \|u - 1\!\!1_\Sigma\|_2^2 + \lambda \mathrm{TV}(u) + \imath_S(u)$$

$S := \{u = 1\!\!1_E \;:\; E \subset \mathbb{R}^2, \quad E \text{ bounded}\}$ (the binary images)

$\mathcal{E}$ is nonconvex because of the constraint S $\Rightarrow$ Nonconvex (intuitive) minimization:

- Level set method [Osher, Sethian 88] $E = \{x \in \mathbb{R}^2 : \phi(x) > 0\} \Rightarrow \partial E = \{x \in \mathbb{R}^2 : \phi(x) = 0\}$
  Then $\mathcal{E}$ is equivalent to

$$\mathcal{E}_1(\phi) = \|H(\phi) - 1\!\!1_\Sigma\|_2^2 + \lambda \int_{\mathbb{R}^2} |\nabla H(\phi(x))| \, dx$$

$H : \mathbb{R} \to \mathbb{R}$ the Heaviside function $H(t) = \begin{cases} 1 & \text{if} \quad t \geq 0 \\ 0 & \text{if} \quad t < 0 \end{cases}$

Computation gets stuck in local minima

$\boldsymbol{L_1 - TV}$ energy: $\boldsymbol{\mathcal{F}(u) = \|u - f\|_1 + \lambda \mathrm{TV}(u)} \quad \boldsymbol{f(x) = 1\!\!1_\Sigma(x), \ \Sigma} \subset \mathbb{R}^2$ bounded

$\mathcal{F}$ is coercive and non-strictly convex $\Rightarrow \arg\min \mathcal{F}$ is nonempty, closed and convex

By (coa) and (cake1)

$$\mathcal{F}(u) = \int_{-\infty}^{+\infty} \big|\{u > t\} \triangle \{f > t\}\big| + \lambda \mathrm{Per}\big(\{u > t\}\big)\, dt = \int_{-\infty}^{+\infty} \big|\{u > t\} \triangle \Sigma\big| + \lambda \mathrm{Per}\big(\{u > t\}\big)\, dt$$

$E \subset \mathbb{R}^2$ bounded $\quad \Rightarrow \|1\!\!1_E - 1\!\!1_\Sigma\|_2^2 = \|1\!\!1_E - 1\!\!1_\Sigma\|_1 \quad \Rightarrow \quad \mathcal{E}(1\!\!1_E) = \mathcal{F}(1\!\!1_E)$

Geometrical nonconvex problem: $\quad \mathcal{E}_1(E) = |E \triangle \Sigma| + \lambda \mathrm{Per}(E) \equiv \mathcal{E}(1\!\!1_E)$ (geo)

There exists $\quad \widehat{\Sigma} \in \arg \min\limits_{E \subset \mathbb{R}^2} \mathcal{E}_1(E)$

For $\widetilde{u} \in \arg \min\limits_{u \in \mathbb{R}^2} \mathcal{F}(u)$ set $\widetilde{\Sigma}(\gamma) = \{\widetilde{u} > \gamma\}$ for a.e. $\gamma \in [0, 1]$

$\mathcal{F}(1\!\!1_{\widetilde{\Sigma}(\gamma)}) \geq \mathcal{E}(1\!\!1_{\widehat{\Sigma}}) = \mathcal{F}(\widetilde{u}) \quad \Rightarrow \quad \widehat{u} := 1\!\!1_{\widehat{\Sigma}} \in \arg\min\limits_{u} \mathcal{F}(u)$

Further, $\mathcal{F}(1\!\!1_{\widetilde{\Sigma}(\gamma)}) = \mathcal{F}(\widetilde{u})$ for a.e. $\gamma \in [0, 1]$. Therefore

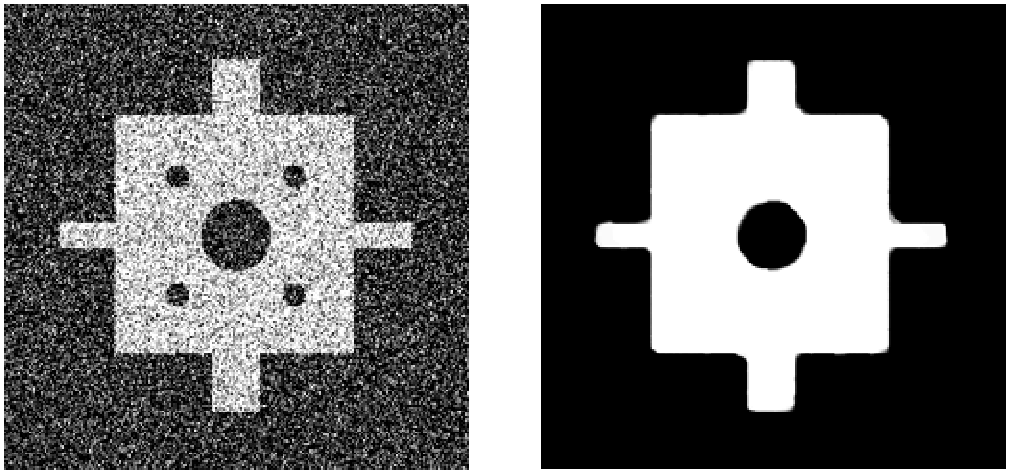
(i) $\widehat{u} = 1\!\!1_{\widehat{\Sigma}}$ is a global minimizer of $\mathcal{E} \quad \Rightarrow \quad \widehat{u} \in \arg \min\limits_{u \in \mathbb{R}^2} \mathcal{F}(u)$;

(ii) $\widetilde{u} \in \arg \min\limits_{u \in \mathbb{R}^2} \mathcal{F}(u) \quad \Rightarrow \quad \widehat{u} := 1\!\!1_{\widehat{\Sigma}} \in \arg\min\limits_{u \in S} \mathcal{E}(u)$, $\widehat{\Sigma} := \{\widetilde{u} > \gamma\}$ for a.e. $\gamma \in [0, 1]$.

For a.e. $\lambda > 0$, $\mathcal{F}$ has a unique minimizer $\widehat{u}$ which is binary by (i) [T. Chan, Esedoglu 05]

- In practice one finds a binary minimizer of $\mathcal{F}$
- If $f = \mathbb{1}_\Sigma$ is noisy, the noise is in the shape $\partial\Sigma$

  Restoring $\widehat{u}$ = denoising = 0-1 segmentation = shape optimization
- The crux: L_1 data fidelity [Alliney 92], [Nikolova 02], [T. Chan, Esedoglu 05] $\Rightarrow$ (cake1)



Data Restored

3. MS for two phase segmentation: The Chan-Vese (CV) model

[T. Chan, Vese 2001]

$$\mathrm{MS}(\Sigma, c_1, c_2) = \int_\Sigma (c_1 - f)^2\,dx + \int_{\Omega\setminus\Sigma} (c_2 - f)^2\,dx + \lambda \mathrm{Per}(\Sigma;\Omega) \quad \text{for} \quad \Omega \subset \mathbb{R}^2 \quad \text{bounded}$$

One should solve $\min_{c_1, c_2 \in \mathbb{R}, \Sigma\subset\Omega} \mathrm{MS}(\Sigma, c_1, c_2) \quad \text{for} \quad f : \Omega \to \mathbb{R}^2.$

For $c_1 = 1$, $c_2 = 0$ and $f = 1\!\!1_\Sigma$ this amounts to $\mathcal{E}_1(E)$ in (geo)

For the optimal $\widehat{\Sigma}$ one has $\hat{c}_1 = \frac{1}{|\widehat{\Sigma}|}\int_{\widehat{\Sigma}} f\,dx$ and $\hat{c}_2 = \frac{1}{|\Omega\setminus\widehat{\Sigma}|}\int_{\Omega\setminus\widehat{\Sigma}} f\,dx$

Two-step iterative algorithms to approximate the solution [T. Chan, Vese 2001]

(a) Solve $\min_\phi \int_\Omega H(\phi)(c_1 - f)^2 + (1 - H(\phi))(c_2 - f)^2 + \lambda\|DH(\phi)\|$

(b) Update c_1 and c_2

Step (a) solves for c_1 and c_2 fixed the nonconvex problem

$$\mathcal{E}(\Sigma) = \int_\Sigma (c_1 - f)^2\,dx + \int_{\Omega\setminus\Sigma} (c_2 - f)^2\,dx + \lambda\,\mathrm{Per}(\Sigma;\Omega)$$

Alternative for step (a): Variational approximation + Γ convergence [Modica, Mortola 77]

$\mathcal{E}_\varepsilon(u) = \int_{\mathbb{R}^2} u^2(c_1 - f)^2 + (1-u)^2(c_2 - f)^2 + \lambda\left(\varepsilon\|Du\|^2 + \frac{1}{\varepsilon}W(u)\right)\,dx$

W double-well potential, $W(0) = W(1) = 0,\ W(u) > 0$ else. E.g., $W(u) = u^2(1 - u^2)$

W forces $\widehat{u}$ to be a characteristic function when $\varepsilon \searrow 0$.

Finding a global minimizer of $\mathcal{E}$ using a convex $\mathcal{F}$ [T. Chan, Esedoglu, Nikolova 06]

For $0 \le u \le 1$ one shows that for a constant K independent of u

$$\int_\Sigma (c_1 - f)^2\,dx + \int_{\Omega\setminus\Sigma} (c_2 - f)^2\,dx = \int_\Omega \big((c_1 - f)^2 - (c_2 - f)^2\big)u\,dx + K$$

and using (coa) one has $\mathcal{E}(\Sigma) = \mathcal{F}(\mathbb{1}_\Sigma) + K$ where

$$\mathcal{F}(u) := \int_\Omega \big((c_1 - f)^2 - (c_2 - f)^2\big)u\,dx + \lambda\|Du\| + \imath_S(u)\,dx \quad \text{for} \quad S := \{u \in \Omega \,:\, u(x) \in [0,1]\}$$

$\mathcal{F}$ – nonstrictly convex and constrained $\quad\Rightarrow\quad \arg\min\limits_u \mathcal{F}(u) \neq \varnothing$ – convex and compact

To summarize:

(i) $\widehat{\Sigma}$ is a global minimizer of $\mathcal{E}$ $\quad\Rightarrow\quad \widehat{u} = \mathbb{1}_{\widehat{\Sigma}} \in \arg\min\limits_{u\in S} \mathcal{F}$;

(ii) $\widetilde{u} \in \arg\min\limits_{u\in S} \mathcal{F}(u) \quad\Rightarrow\quad \widehat{\Sigma} := \{\widetilde{u} > \gamma\} \in \arg\min\limits_{\Sigma\subset\mathbb{R}^2} \mathcal{E}(\Sigma)$ for a.e. $\gamma \in [0,1]$.

$\mathcal{F}$ provides a tight relaxation of $\mathcal{E}$

Convex non tight relaxation for the **full** CV model: [Brown, T. Chan, Bresson 12]

4. Nonconvex data fidelity with convex regularization

[Pock, Cremers, Bischof, Chambolle SIIMS 10], [Pock et al, 08]

$u : \Omega \to \Gamma$ (bounded), $\Omega \subset \mathbb{R}^2$. Continuous energy $\mathcal{E}$:

$$\mathcal{E}(u) := \int_\Omega g(x, u(x)) + \lambda h(\|Du\|)\, dx$$

Data term based on Cartesian currents: depends on the whole graph $(x, u(x))$. Nonconvex in general. The regularization is convex, one-homogeneous w.r.t. $\|Du\|$.

Approach: embed the minimization of $\mathcal{E}$ in a higher dimensional space [Chambolle 01]

Similar approach in discrete setting: [Ishikawa, Geiger 03] with numerical intricacies.

Using (cake) and the fact that $|\partial_t \mathbb{1}_{\{u>t\}}(x)| = \delta(u(x) - t) = +\infty$ if $u(x) = t$ and $= 0$ otherwise, one can find a global minimizer of $\mathcal{E}$ by minimizing

$$\mathcal{E}_1(\mathbb{1}_{\{u>t\}}) = \int_{\Omega\times\Gamma} g(x,t)|\partial_t \mathbb{1}_{\{u>t\}}(x)| + \lambda h(\|D_x \mathbb{1}_{\{u>t\}}\|)\, dx\, dt$$

$\mathcal{E}_1$ is convex w. r. t. $\mathbb{1}_{\{u>t\}}$ but $\mathbb{1}_{\{u>t\}} : [\Omega \times \Gamma] \to \{0, 1\}$ is discontinuous.

Relaxation: replace $1\!\!1_{\{\cdot\}}$ by $\phi \in S$ where

$$S := \left\{ \phi \in BV(\Omega \times \mathbb{R}; [0,1]) : \lim_{t \to -\infty} \phi(x,t) = 1, \lim_{t \to +\infty} \phi(x,t) = 0 \right\}$$

$\mathcal{F}$ below is convex and constrained:

$$\mathcal{F}(\phi) = \int_{\Omega \times \mathbb{R}} g(x,t)|\partial_t \phi(x,t)| + \lambda h(\|D_x \phi(x,t)\|) + \imath_S(\phi)\, dx\, dt$$

Facts:

- $\mathcal{F}$ obeys the generalized coarea formula $\mathcal{F}(\phi) = \displaystyle\int_{-\infty}^{+\infty} \mathcal{F}(1\!\!1_{\{\phi > t\}}) dt$
- $\widehat{\phi} \in \arg\min_{\phi} \mathcal{F}(\phi) \quad \Rightarrow \quad \mathcal{F}(\widehat{\phi}) = \displaystyle\int_0^1 \mathcal{F}(1\!\!1_{\{\widehat{\phi} > t\}}) dt$
- for a.e. $\gamma \in [0,1)$, $1\!\!1_{\{\widehat{\phi} > \gamma\}} \in \arg\min_{\phi} \mathcal{F}(\phi)$

Therefore

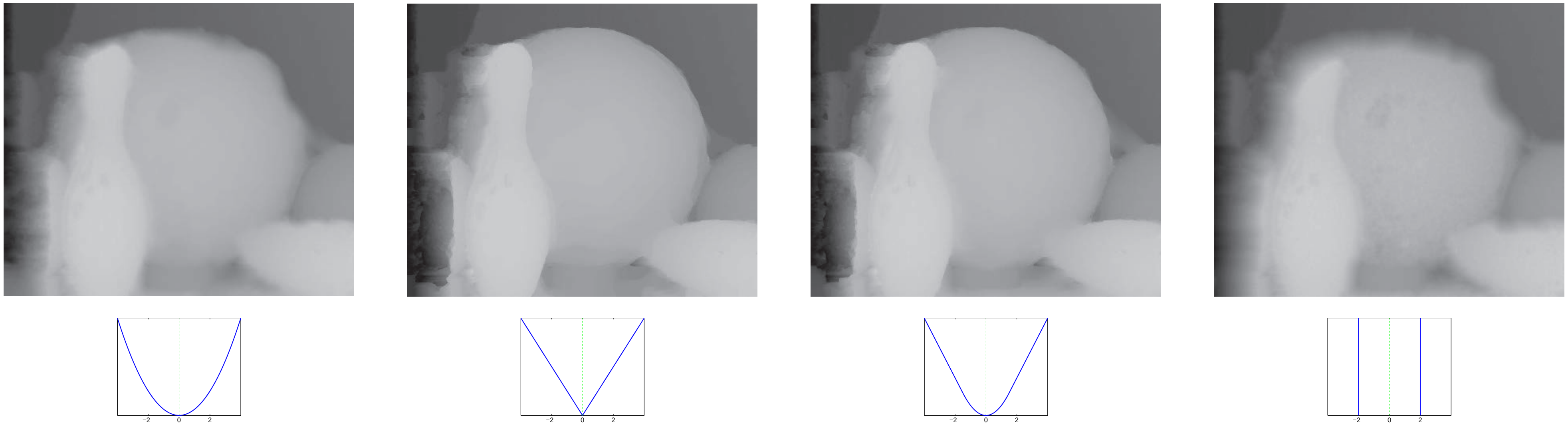
(i) $\widehat{\phi} \in \arg\min_{\phi} \mathcal{F}(\phi) \quad \Rightarrow \quad 1\!\!1_{\{\widehat{\phi} > \gamma\}}$ for a.e. $\gamma \in [0,1]$ is a global minimizer of $\mathcal{E}_1$;

(ii) From $1\!\!1_{\{\widehat{\phi} > \gamma\}}$ a global minimizer $\widehat{u}$ of $\mathcal{E}$ is found.

Disparity estimation

(a) Left input image (b) Right input image (c) True disparity

Figure 7. *Rectified stereo image pair and the ground truth disparity. Light gray pixels indicate structures near to the camera, and black pixels correspond to unknown disparity values.*



quadratic TV Huber Lipscitz

Image credits to the authors: Pock, Cremers, Bischof, Chambolle 2010

5. Minimal Partitions [Chambolle, Cremers, Pock, SIIMS 12]

$u : \Omega \to \mathbb{R}, \quad \Omega \subset \mathbb{R}^2$ (open) Extension to $\mathbb{R}^3$ is also considered

Goal: partition Ω into (at most) k regions whose total perimeter is minimal using external data.

This amounts to partition Ω into ideal soap films [Brakke 95].

$$\begin{aligned} \mathcal{E}(\{\Sigma_i\}_{i=1}^k) &= \sum_{i=1}^k \left(\int_{\Sigma_i} g_i(x) + \frac{1}{2}\mathrm{Per}(\Sigma_i;\Omega) + \imath_{S^\Sigma}(x)\, dx \right) \\ S^\Sigma &= \left\{ \{E_i\}_{i=1}^k \subset \mathbb{R}^2 \ : \ E_i \cap E_j = \varnothing \text{ if } i \neq j \ \text{ and } \ \cup_{i=1}^k E_i = \Omega \right\} \end{aligned}$$

$g_i : \Omega \to \mathbb{R}_+$, $1 \leq i \leq k$ are given external potentials (e.g., extracted from input data). Set

$$\chi_i(x) := \mathbb{1}_{\Sigma_i}(x) \quad i \in \{1,\dots,k\} \quad \chi := (\chi_1,\dots,\chi_k) \in \mathbb{R}^{d\times k} \quad \text{and} \quad g := (g_1,\dots,g_k)$$

By (coa) and (per), the interfacial energy reads as

$$\Phi(\chi) := \frac{1}{2}\sum_{i=1}^k \int_\Omega \|D\chi_i\| + \imath_{S^0}(\chi) \quad \text{for} \quad S^0 = \left\{ \chi \in BV(\Omega;\{0,1\}^k) \ : \ \sum_{i=1}^k \chi_i = 1 \text{ a.e. in } \Omega \right\}$$

Minimizing $\mathcal{E}$ is equivalent to minimize: $\mathcal{E}_1(\chi) = \int_\Omega \chi(x)\cdot g(x) + \Phi(\chi)(x)\, dx$

$\mathcal{E}_1$ is convex but S^0 is discrete. It is known that $\mathcal{E}_1$ has global minimizers.

A straightforward convex relaxation is to replace $\chi \in S^0$ by $v \in S$

$$S := \left\{ v \in BV(\Omega; [0,1]^k) \ : \ \sum_{i=1}^{k} v_i(x) = 1 \text{ a.e. in } \Omega \right\}$$

and to minimize the convex $\mathcal{F}_1(v) = \int_\Omega v(x) \cdot g(x) + \frac{1}{2} \sum_{i=1}^k \int_\Omega \|Dv_i\| + \imath_S(v)\, dx$

e.g. [Zach et al, 08], [Bae, Yuan, Tai 11]

This relaxation is not tight except for $k = 2$

The goal is to conceive a convex relaxation of $\mathcal{E}_1$ as tight as possible.

Construction of a "local" convex envelope $\widetilde{\Phi}$ of Φ

Let $\mathcal{E}_1^*$ be the convex conjugate of $\mathcal{E}_1$, see (cc). Then $\mathcal{E}_1^{**}$ is the convex envelope of $\mathcal{E}_1$, so $\arg\min_v \mathcal{E}_1 \subset \arg\min_v \mathcal{E}_1^{**}$. Its domain is $\operatorname{dom} \mathcal{E}_1^{**} = S$, but $\mathcal{E}_1^{**}$ is seldom computable.

One looks for the largest non-negative, even, convex envelope of Φ of the form $\widetilde{\Phi}(v) = \int_\Omega h(v, Dv) \quad \text{for} \quad h : \Omega \times \mathbb{R}^{k\times d} \to \mathbb{R}_+$ satisfying

$$\widetilde{\Phi}(v) \leq \Phi(v) \ \forall\, v \in L^2(\Omega; \mathbb{R}^k) \quad \text{and} \quad \widetilde{\Phi}(v) = \Phi(v) \ \forall\, v \in S^0$$

Result: $\widetilde{\Phi}(v) = \displaystyle\int_\Omega h(Dv) + \imath_S(v)\, dx$ where

$$h(p) = \sup_{q \in K} q \cdot p \quad \text{for} \quad K = \left\{ q = (q_1, \ldots, q_k)^T \in \mathbb{R}^{k\times d} \ : \ \|q_i - q_j\| \leq 1 \ \ \forall\, i < j \right\}$$

This estimate $\widetilde{\Phi}$ of Φ is nearly optimal.

The convex partition problem

$$\mathcal{F}(v) = \int_\Omega v(x) \cdot g(x) + \widetilde{\Phi}(v)\, dx$$

Let $\widehat{v} \in \arg\min \mathcal{F}(v)$. Cases:

1. $\widehat{v} = (\widehat{v}_1, \cdots, \widehat{v}_k) \in S^0 \Rightarrow \left\{ \widehat{\Sigma}_i := \mathrm{supp}\ (\widehat{v}_i) \right\}_{i=1}^{k}$ is a global minimizer of $\mathcal{E}$;
2. $\widehat{v} \in S \setminus S^0$ and $\widehat{v}$ is a convex combination of several $\widehat{w}_i \in \arg\min \mathcal{E}_1(v)$ then for each i, $\widehat{w}_i \in \arg\min \mathcal{F}(v)$ and $\mathcal{F}(\widehat{v}) = \min_v \mathcal{E}_1(v)$. For $k \geq 3$ a binarization may be used (see, e.g. [Lellmann Schnörr 11]) or a slight perturbation of g.
3. $\widehat{v} \in S \setminus S^0$ and $\widehat{v}$ is not a convex combination of some global minimizers of $\mathcal{E}_1$. Then $\mathcal{F}(\widehat{v}) \leq \min_v \mathcal{E}_1(v)$.

Case 1 occurs much more often than cases 2 and 3.

Minimization of $\mathcal{F}$ by primal-dual ArrowHurwicz-type algorithm.

For less tight relaxations such as $\mathcal{F}_1$ case 1 is less frequent.

The 3 cases

Figure 7. *Completion of four regions.*

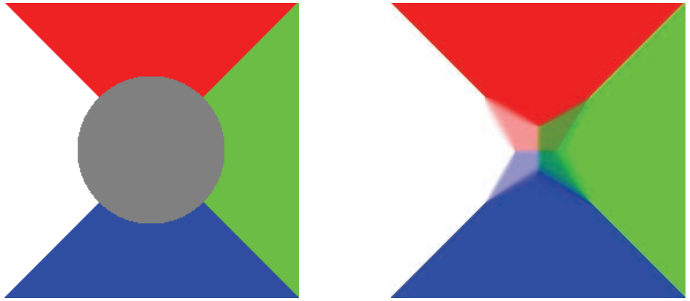


Figure 8. *Completion of four regions: in case of nonuniqueness, the method may find a combination of the solutions.*

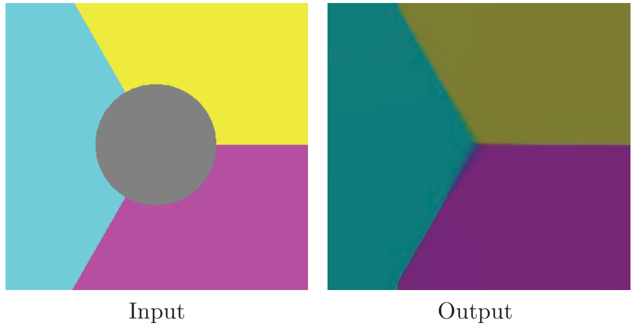


Figure 6. *Example of a nonbinary solution.*

Cases 1 and 2

$$g_1 = (1,0,0),\ g_2 = (0,1,0),\ g_3 = (0,0,1),\ g_4 = (1,1,1)$$

Case 3

$$(g_1, g_2, g_3)$$

Image credits to the authors: Chambolle, Cremers, Pock, SIIMS 12

Comparison with other methods

[Chambolle, Cremers, Pock 12]

[Zach,Gallup,Frahm,Niethammer 08]

[Lellmann,Kappes,Yuan,Becker,Schnörr 09]

Image credits to the authors: Chambolle, Cremers, Pock, SIIMS 12

References

E. Bae, J. Yuan, and X.-C. Tai, Global minimization for continuous multiphase partitioning problems using a dual approach, Int. J. Comput. Vis., 92 (2011), pp. 112-129.

K. A. Brakke, Soap films and covering spaces, J. Geom. Anal., 5 (1995), pp. 445514.

E. Brown, T. Chan, X. Bresson: Completely convex formulation of the Chan-Vese image segmentation model. International Journal of Computer Vision 98, 103121 (2012)

A. Chambolle, Convex representation for lower semicontinuous envelopes of functionals in L1, J. Convex. Anal., 1 (2001), pp. 149-170.

A. Chambolle, D. Cremers, and T. Pock, A Convex Approach to Minimal Partitions, SIAM Journal on Imaging Sciences, 5(4) 2012, pp. 1113-1158

T. F. Chan, S. Esedoglu, and M. Nikolova, Algorithms for finding global minimizers of image segmentation and denoising models, SIAM J. Appl. Math., 66 (2006), pp. 16321648.

S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Trans. on Pattern Analysis and Machine Intelligence, PAMI-6 (1984), pp. 721-741.

L. Hammer, P. Hansen, and B. Simeone, Roof duality, complementation and persistency in quadratic 0-1 optimization, Math. Programming, 28 (1984), pp. 121–155.

H. Ishikawa and D. Geiger, Segmentation by grouping junctions, in Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR), 1998, pp. 125131.

M. Klodt, T. Schoenemann, K. Kolev, M. Schikora, and D. Cremers, An experimental comparison of discrete and continuous shape optimization methods, in Proceedings of the 10th European Conference on Computer Vision, 2008, pp. 332-345.

J. Lellmann, J. Kappes, J. Yuan, F. Becker, and C. Schnörr, Convex multi-class image labeling by simplex-constrained total variation, in Scale Space and Variational Methods in Computer Vision, Lecture Notes in Comput. Sci. 5567, Springer, Berlin, 2009, pp. 150-162.

J. Lellmann and C. Schnörr, Continuous multiclass labeling approaches and algorithms, SIAM J. Imaging Sci., 4 (2011), pp. 1049-1096.

N. Papadakis, J.-F. Aujol, V. Caselles, and R. Yildizoğlu, High-dimension multi-label problems: convex or non convex relaxation? SIAM Journal on Imaging Sciences, 6(4), 2013, pp. 2603–2639

T. Pock, T. Schoenemann, G. Graber, H. Bischof, and D. Cremers, A convex formulation of continuous multi-label problems, in European Conference on Computer Vision (ECCV), Lecture Notes in Comput. Sci. 5304, Springer-Verlag, Berlin, Heidelberg, 2008, pp. 792-805.

T. Pock, A. Chambolle, D. Cremers, and H. Bischof, A convex relaxation approach for computing minimal partitions, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009, pp. 810-817.

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

R. B. Potts, Some generalized order-disorder transformations, Proc. Cambridge Philos. Soc., 48

(1952), pp. 106109.

S. Roy and I. J. Cox, A maximum-flow formulation of the N-camera stereo correspondence problem, in Proc. of the IEEE Int. Conf. on Computer Vision (ICCV), 1998, pp. 492-502.

C. Zach, D. Gallup, J. M. Frahm, and M. Niethammer, Fast global labeling for real-time stereo using multiple plane sweeps, in Vision, Modeling, and Visualization 2008, IOS Press, Amsterdam, The Netherlands, 2008, pp. 243252.