Convex Optimization for Multi-Class Image Labeling

Jan Lellmann and Christoph Schnörr
Image and Pattern Analysis Group, HCI, University of Heidelberg
e-mail: {lellmann, schnoerr}@math.uni-heidelberg.de

Overview

- Continuous Multi-Class Image Labeling
- Relaxation yields linear data term
- Use weighted variant of total variation on vector fields as regularizer
- Allows for non-trivial interaction potentials, can be optimized quickly
- Models Euclidean distance interaction potentials exactly, non-Euclidan distances can be approximated
- Globally optimized using Nesterov's first-order approach with explicit optimality bounds

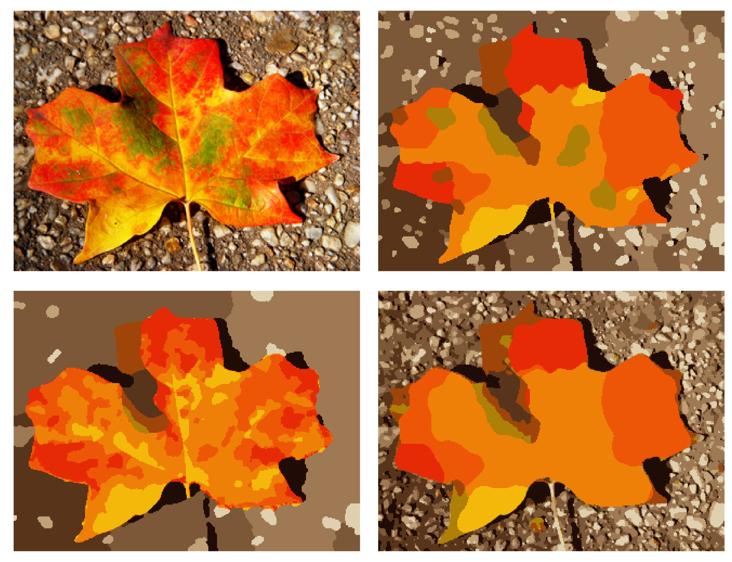
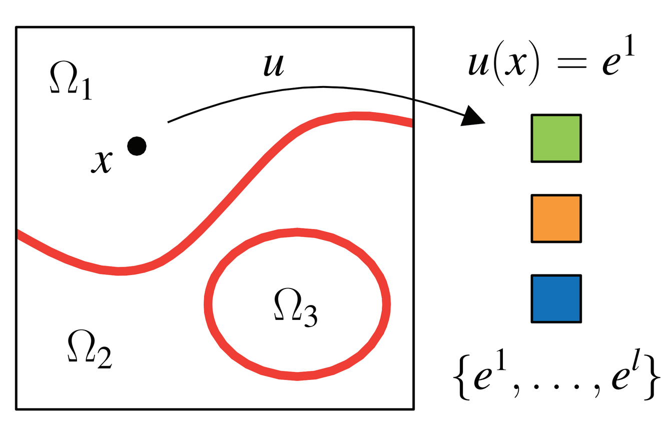


Figure: Application to color segmentation. Parametrization of the regularizer allows for selective suppression of background or foreground structure.

Continuous Problem Formulation

Problem

For each pixel $x \in \Omega \subseteq \mathbb{R}^d$, find a class label $\ell(x) \in \{1, \ldots, l\}$ according to *local data fidelity* and *regularization term*
⇒ Combinatorial problem.



Relaxation

Identify the i-th label with e^i and relax to the unit simplex in $\mathbb{R}^l$, find

$$\inf_{u \in \mathcal{C}} \int_\Omega \langle u(x), s(x) \rangle dx + J(u) \,, \qquad (1)$$

where $\mathcal{C} := \{u : \Omega \to \mathbb{R}^l | u_i(x) \geqslant 0, \sum_{i=1}^l u_i(x) = 1\}$ and J penalizes label changes.
⇒ *Convex* problem for any data term. Formulating J means trading off generality vs. simplicity of computation.

Total Variation Based Regularizer

Approach

Fix an interface potential $d : \{1, \ldots, l\}^2 \to \mathbb{R}$ and require regularization according to boundary length with weight depending on labels i and j of adjoining regions,

$$J(e^i 1_S + e^j 1_{S^c}) = d(i,j) Per(S) \qquad (2)$$

for any set $S \subset \Omega$ with finite perimeter. If J convex, positively homogeneous and $J(u) = 0$ for constant u, then d must be a metric.

Euclidean Distances

Idea: Use linear modification of total variation on vectors,

$$J(u) := \mathrm{TV}_A(u) := \int_\Omega \|D(Au)\|_F \, dx \,, \qquad (3)$$

here $\|D(\cdot)\|_F$ Frobenius norm of the Jacobian and $A \in \mathbb{R}^{k \times l}$. Then

$$\mathrm{TV}_A(e^i 1_S + e^j 1_{S^c}) = \|a^i - a^j\| \operatorname{Per}(S) \,. \qquad (4)$$

If $d(i,j) = \|a^i - a^j\|$, i.e. d is an Euclidean distance, J is exact for hard labeling.

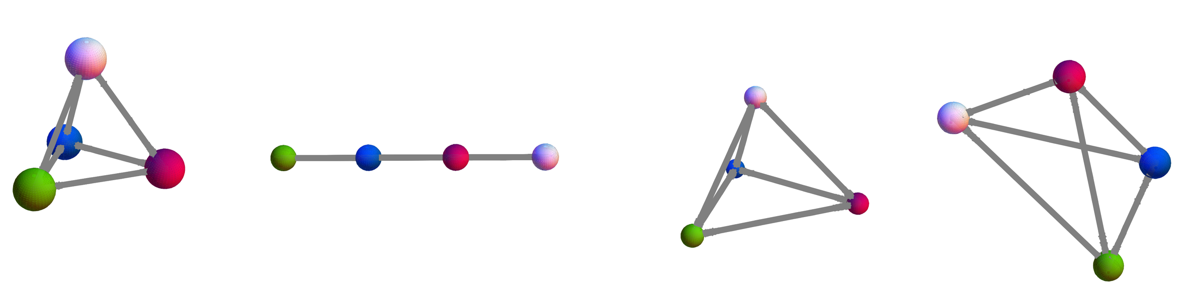


Figure: Euclidean embeddings for several distances into $\mathbb{R}^3$: Potts distance, Linear distance, approximated truncated label distance, distance for foreground-background separation

Approximation of Non-Euclidean Distances

In case d is non-Euclidean, e.g. $d(i,j) = \min(1, |i-j|)$: Set $D_{ij} = d(i,j)^2$ and compute Euclidean approximation by minimizing

$$\|E - D\|_M \qquad (5)$$

over all Euclidean distance matrices E by solving a convex semidefinite program.
Example: truncated linear distance, absolute error bound $\varepsilon_E = 0.145$:

$$\begin{pmatrix} 0 & 1 & 2 & 2 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 2 & 2 & 1 & 0 \end{pmatrix} \text{ vs. } \begin{pmatrix} 0 & 1.15 & 1.92 & 2.08 \\ 1.15 & 0 & 1.15 & 1.92 \\ 1.92 & 1.15 & 0 & 1.15 \\ 2.08 & 1.92 & 1.15 & 0 \end{pmatrix}$$

Discretized Problem

Use forward differences and support function representation of discrete total variation to get a *bilinear saddle point problem*:

$$\min_{u \in \mathcal{C}} \max_{v \in \mathcal{D}} \langle u, s \rangle + \langle Lu, v \rangle - \langle b, v \rangle \,, \qquad (6)$$

where L discretization of gradient and

$$\mathcal{C} := \{u \in \mathbb{R}^{n \times l} | u_{i,\cdot} \in \Delta_l, i = 1, \ldots, n\} \,, \qquad (7)$$

$$\mathcal{D} := \prod_{x \in \Omega} \{p = (p^1, \ldots, p^l) \in \mathbb{R}^{d \times l} \,|\, \|p\|_F \leqslant 1\} \subseteq \mathbb{R}^{n \times d \times l} \,. \qquad (8)$$

Existence and strong duality follow from boundedness of $\mathcal{C}, \mathcal{D}$. Projections onto $\mathcal{C}, \mathcal{D}$ can be computed exactly in a finite number of steps.

Optimization

Solve non-smooth problem using Nesterov's approach [1]

- First-order method: exploit sparsity
- Combines controlled smoothing with accumulated gradients
- Requires only evaluations of L and projections onto $\mathcal{C}, \mathcal{D}$
- $O(1/n)$ convergence, compared to $O(1/\sqrt{n})$ for subgradient methods
- *Explicit* suboptimality bound for given number of iterations, $O(\varepsilon^{-1} n \sqrt{d} \|A\|)$ to find ε-optimal solution
- Fully automatic, parameter–free

Algorithm 1 Convex Multi-Class Labeling

Input: $c_1 \in \mathcal{C}$, $c_2 \in \mathcal{D}$ and $r_1, r_2 \in \mathbb{R}$ s.t. $\mathcal{C} \subseteq \mathcal{B}_{r_1}(c_1)$ and $\mathcal{D} \subseteq \mathcal{B}_{r_2}(c_2)$; $x^{(0)} \in \mathcal{C}$; $\mathbb{R} \ni C \geqslant \|L\|$, $N \in \mathbb{N}$.
Output: $u^{(N)} \in \mathcal{C}$, $v^{(N)} \in \mathcal{D}$.
Let $\mu \leftarrow \frac{2\|L\|}{N+1} \frac{r_1}{r_2}$.
Set $G^{(-1)} = 0$, $v^{(-1)} = 0$.
for $k = 0, \ldots, N$ **do**
  $V \leftarrow \Pi_{\mathcal{D}}\left(c_2 + \frac{1}{\mu}\left(Lx^{(k)} - b\right)\right)$.
  $v^{(k)} \leftarrow v^{(k-1)} + 2\frac{(k+1)}{(N+1)(N+2)} V$.
  $G \leftarrow s + L^\top V$.
  $G^{(k)} \leftarrow G^{(k-1)} + \frac{k+1}{2} G$.
  $u^{(k)} \leftarrow \Pi_{\mathcal{C}}\left(x^{(k)} - \frac{\mu}{\|L\|^2} G\right)$.
  $z^{(k)} \leftarrow \Pi_{\mathcal{C}}\left(c_1 - \frac{\mu}{\|L\|^2} G^{(k)}\right)$.
  $x^{(k+1)} \leftarrow \frac{2}{k+3} z^{(k)} + \left(1 - \frac{2}{k+3}\right) u^{(k)}$.
end for

Experiments

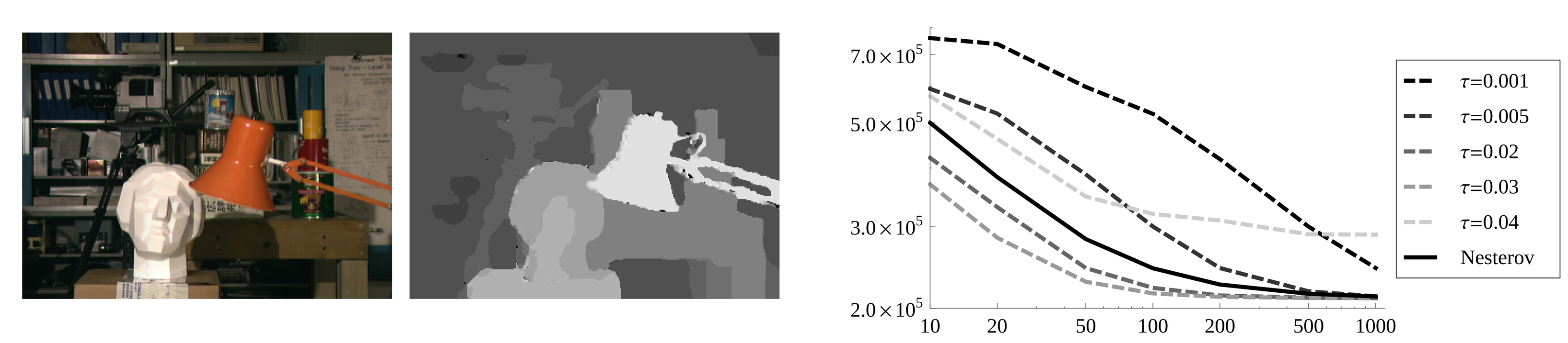


Figure: Convergence for stereo disparity estimation with non-Euclidean distance and 16 disparities: Objective vs. number of iterations N of our method (solid) compared to the Arrow-Hurwicz method from [2] (dashed) for various step sizes.

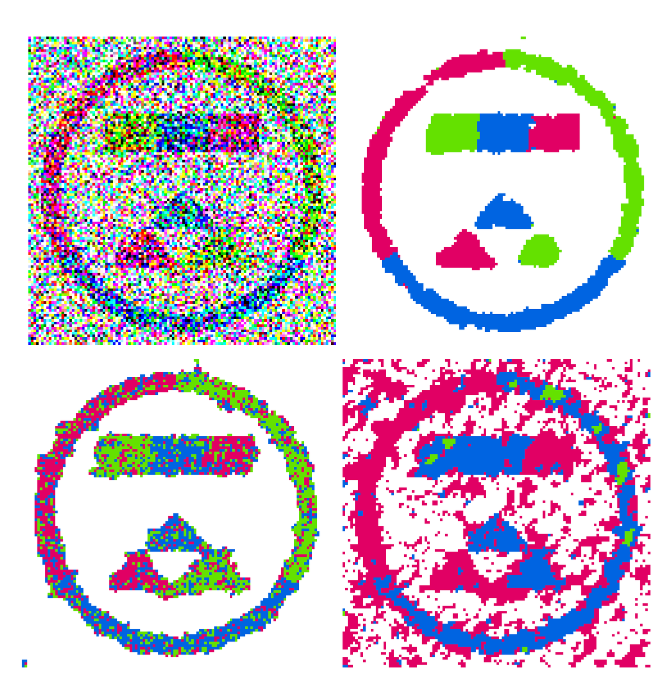


Figure: Four-class color segmentation with varying distance: Input, proposed method with Potts, fg-bg, and linear distance. The fg-bg distance clearly segments the three foreground classes from the white background but allows for a large variance within the foreground. The linear distance corresponds to a degenerate embedding and results in a strongly suboptimal discrete solution.

Figure: Region filling properties: α-β swap generates block artifacts. The relaxation yields a non-binary solution; after binarization the result is in accordance with the expected solution.

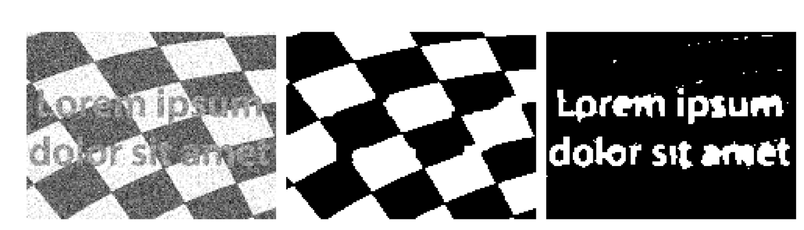


Figure: Simultaneous segmentation and background reconstruction: Noisy image; background and foreground reconstructed using a non-uniform distance.

References

Y. Nesterov.
Smooth minimization of non-smooth functions.
Math. Prog., 103(1):127–152, 2004.

A. Chambolle, D. Cremers, and T. Pock.
A convex approach for computing minimal partitions.
Technical Report 649, Ecole Polytechnique CMAP, 2008.

J. Lellmann et al.
Convex Multi-Class Image Labeling by Simplex-Constrained Total Variation.
In: *Scale Space and Variational Methods*, LNCS 5567:150–162, 2009.