A Geometric Approach to Image Labeling

Freddie Åström[*,§], Stefania Petra[*,†], Bernhard Schmitzer[‡] and Christoph Schnörr[*,§]

[*]HCI, [†]MIG, [§]IPA, Heidelberg University, Germany
[‡]CEREMADE, University Paris-Dauphine, France

freddie.astroem@iwr.uni-heidelberg.de

Introduction

We introduce a smooth non-convex approach to image labeling in a novel geometric framework which complements established approaches to image labeling.

- The major underlying concept is a smooth manifold of probabilistic assignments of a pre-specified set of prior data (the "labels") to given image data.
- The Riemannian gradient flow, with respect to a corresponding objective function, evolves on the manifold and terminates, for any $\delta > 0$, within a δ-neighborhood of an unique assignment (labeling).
- Our approach is numerically implemented with sparse, highly-parallel interior-point updates that efficiently converge, largely independent from the number of labels.

Original **Input** **TRWS** 0.78 **LBP** 0.95 **Geometric** 0.99

Method

- **The Assignment Manofold.** The relative interior $\mathcal{S} = \mathring{\Delta}_{n-1}$ of the probability simplex $\Delta_{n-1} = \{p \in \mathbb{R}^n_+ \colon \langle \mathbb{1}, p\rangle = 1\}$ becomes a differentiable Riemannian manifold when endowed with the Fisher-Rao metric:

$$\langle u, v\rangle_p := \langle \frac{u}{\sqrt{p}}, \frac{v}{\sqrt{p}}\rangle, \ \forall u, v \in T_p\mathcal{S}, \quad T_p\mathcal{S} = \{v \in \mathbb{R}^n \colon \langle \mathbb{1}, v\rangle = 0\}, \ p \in \mathcal{S}.$$

Definition 1 (Assignment Manifold)**.** *The manifold of assignment matrices, called assignment manifold, is the set*

$$\mathcal{W} = \{W \in \mathbb{R}^{m\times n} \colon W_i \in \mathcal{S}, \ i \in [m]\}.$$

- **Distance Matrix, Lifting Map and Riemannian Means.** The Riemannian distance between a prior set $\mathcal{P}_\mathcal{F}$ and the data f is

$$D \in \mathbb{R}^{m\times n}, \ D_i \in \mathbb{R}^n, \quad D_{ij} = \frac{1}{\rho} d_\mathcal{F}(f_i, f_j^*), \quad \rho > 0, \quad i \in [m], \quad j \in [n].$$

Given D and W, we lift the vector field D to the manifold $\mathcal{W}$ by

$$L = L(W) := \exp_W(-U) \in \mathcal{W}, \qquad U_i = D_i - \frac{1}{n}\langle \mathbb{1}, D_i\rangle \mathbb{1}, \quad i \in [m].$$

Based on the likelihood matrix L, we define the *similarity matrix*

$$S = S(W) \in \mathcal{W}, \qquad S_i = \mathrm{mean}_\mathcal{S}\{L_j\}_{j\in\mathcal{N}_\mathcal{E}(i)}, \qquad i \in [m].$$

Definition 2 (Lifting Map (Manifolds $\mathcal{S}, \mathcal{W}$))**.** *The lifting mapping is defined by*

$$\exp \colon T\mathcal{W} \to \mathcal{W}, \qquad (W, U) \mapsto \exp_W(U) = \begin{pmatrix} \exp_{W_1}(U_1) \\ \dots \\ \exp_{W_m}(U_m) \end{pmatrix},$$

where $U_i, W_i, i \in [m]$ index the row vectors of the matrices U, W.

- **Objective Function.** The assignment matrix $W \in \mathcal{W}$ can be thought of as *posterior probabilities*,

$$W_{ij} = \Pr(f_j^* | f_i),$$

of assigning prior feature f_j^* to the observed feature f_i. A natural *objective function* to be maximized is

$$\max_{W\in\mathcal{W}} J(W), \qquad J(W) := \langle S(W), W\rangle.$$

The functional J together with the feasible set $\mathcal{W}$ formalizes the following objectives:

1. Assignments W *maximally correlate* with the feature-induced similarities $S = S(W)$
2. Assignments of prior data to observations should be done in a *spatially coherent* way.
3. Maximizers W^* should define *image labelings* in terms of rows $\overline{W}_i^* = e^{k_i} \in \{0,1\}^n, \ i, k_i \in [m]$, that are indicator vectors.

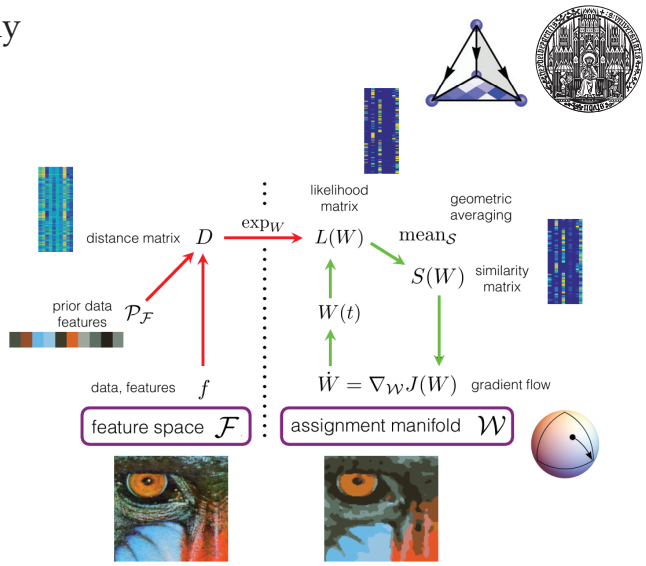
Lemma 1. *Let $\overline{\mathcal{W}}$ denote the closure of $\mathcal{W}$. We have*

$$\sup_{W\in\mathcal{W}} J(W) = m,$$

and the supremum is attained at the extreme points

$$\overline{\mathcal{W}}^* := \{\overline{W}^* \in \{0,1\}^{m\times n} \colon \overline{W}_i^* = e^{k_i}, \ i \in [m], \ k_1, \dots, k_m \in [n]\} \subset \overline{\mathcal{W}},$$

corresponding to matrices with unit vectors as row vectors.



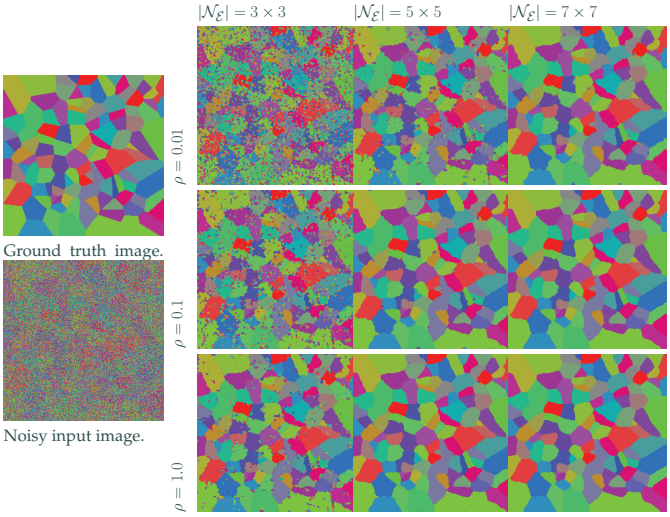
- **Optimization.** *Riemannian gradient ascent flow* on the manifold $\mathcal{W}$, with an adaptive step-size, gives in the fixed-point iteration

$$\partial_{W_{ij}} J(W) = \langle \underbrace{T^{ij}(W), W}_{\approx 0}\rangle + S_{ij}(W)$$

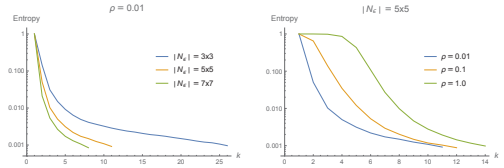
$$W_i^{(k+1)} = \frac{W_i^{(k)}(\nabla_i J(W^{(k)}))}{\langle W_i^{(k)}, \nabla_i J(W^{(k)})\rangle} \approx \frac{W_i^{(k)}(S_i(W^{(k)}))}{\langle W_i^{(k)}, S_i(W^{(k)})\rangle}, \ W_i^{(0)} = \frac{1}{n}\mathbb{1}, \ i \in [m]$$

Experiments

- **Parameter Influence on Labeling.** The labeling problem comprises of 31 color vectors forming the prior data set $\mathcal{P}_\mathcal{F}$ with the distance $d_\mathcal{F}(f^i, f^j) := \|f^i - f^j\|_1$.



- **Convergence.** **Left.** Small spatial scales necessitate to resolve more conflicting assignments through propagating information by geometric spatial averaging. **Right.** Higher selectivity leads to faster convergence.



- **Image Inpainting by Labeling.** The selectivity parameter was set to $\rho = 1$ and the neighborhood size was 3×3.

Original **Input** **TRWS** **SWAP** **BPS** **Geometric**

Further results

Åström, F.; Petra, S.; Schmitzer, B. and Schnörr, C. "Image Labeling by Assignment", arXiv, http://arxiv.org/abs/1603.05285, 2016.
Åström, F.; Petra, S.; Schmitzer, B. and Schnörr, C. "The Assignment Manifold: A Smooth Model for Image Labeling", DIFF-CVML'16, CVPR, 2016.