Joint reconstruction-segmentation on graphs

Jeremy Budd [1] Yves van Gennip [2] Jonas Latz [3] Simone Parisotto [4] Carola-Bibiane Schönlieb [5]

[1]Universität Bonn [2]Technische Universiteit Delft [3]Heriot-Watt University [4]siHealth Photonics [5]University of Cambridge

The problem

The task of **image segmentation** concerns locating the key parts of an image $x^* : Y \to \mathbb{R}^\ell$. However, in practice images are typically observed indirectly. That is, for some **forward model** $\mathcal{T}$ and noise e, we have **observations**:

$$y = \mathcal{T}(x^*) + e. \tag{1}$$

Thus, our segmentation task is a **reconstruction-segmentation** task. Given y and an already reconstructed and segmented **reference image** $x_d : Z \to \mathbb{R}^\ell$ with **a priori segmentation** $f : Z \to \{0, 1\}$, we seek to reconstruct and segment x^*. *(N.B. Our focus is primarily on the accuracy of the segmentation.)*

Background

Following (Tikhonov, 1963), a major approach to inverting (1) is via **variational methods**. Traditionally, reconstruction-segmentation was performed **sequentially**: first reconstruct, then segment that reconstruction. On the other extreme are **end-to-end** methods, which learn a map from observations to a segmentation. The method of **joint reconstruction-segmentation** lies between these extremes: we reconstruct and segment simultaneously, using each to guide the other. For a detailed overview of these methods, see (Adler *et al.*, 2018) and (Corona *et al.*, 2019). In this work, we incorporate the powerful **graph-based segmentation** methods pioneered by (Bertozzi, Flenner, 2012) into this joint reconstruction-segmentation framework.

Analysis on graphs, and the graph Ginzburg–Landau energy

- A **graph** G is a (finite) set of **vertices** V linked by **edges** E.
- We assign each edge (i, j) a **weight** ω_{ij} and each vertex a **degree** $d_i := \sum_j \omega_{ij}$.
- Define spaces $\mathcal{V} := \{u : V \to \mathbb{R}\}$ and $\mathcal{E} := \{\varphi : E \to \mathbb{R}\}$ with inner products

$$\langle u, v\rangle_{\mathcal{V}} := \sum_{i\in V} d_i u_i v_i \qquad \text{and} \qquad \langle \varphi, \phi\rangle_{\mathcal{E}} := \frac{1}{2}\sum_{i,j\in V} \omega_{ij}\varphi_{ij}\phi_{ij}.$$

- Define graph variants of the **gradient** and **Laplacian**:

$$(\nabla u)_{ij} := \begin{cases} u_j - u_i, & (i, j) \in E \\ 0, & \text{otherwise} \end{cases} \qquad \text{and} \qquad (\Delta u)_i := d_i^{-1}\sum_{j\in V} \omega_{ij}(u_i - u_j).$$

- We define a graph analogue of the **Ginzburg–Landau functional**, for W a double-well potential

$$\mathrm{GL}_{\varepsilon,\mu,f}(u, \omega) = \frac{1}{2}||\nabla u||^2_{\mathcal{E}} + \frac{1}{\varepsilon}\langle W \circ u, \mathbf{1}\rangle_{\mathcal{V}} + \frac{1}{2}\left|\left|\mu^{\frac{1}{2}} \odot (u - f)\right|\right|^2_{\mathcal{V}}.$$

- The ODE for the gradient flow of $\mathrm{GL}_{\varepsilon,\mu,f}$ with respect to u is the **Allen–Cahn equation**

$$\frac{du}{dt} = -\Delta u - \frac{1}{\varepsilon}W' \circ u - \mu \odot (u - f).$$

An iterative scheme for joint reconstruction-segmentation

We will model our reconstruction-segmentation task as the following variational problem:

$$\min_{x\in\mathbb{R}^{N\times\ell},u\in\mathcal{V}} \mathcal{R}(x) + \alpha||\mathcal{T}(x) - y||^2_F + \beta\,\mathrm{GL}_{\varepsilon,\mu,f}(u, \Omega(x)) \tag{2}$$

where $\mathcal{R}$ is a convex **regulariser**. The first two terms in the objective functional are a standard Tikhonov reconstruction energy, and the final Ginzburg–Landau term is the segmentation energy.

This joint problem is a lot to solve all at once, so we will use the following alternating iterative scheme to approach solutions (where $\alpha, \beta, \eta_n, \nu_n$ are parameters):

$$x_{n+1} = \operatorname*{argmin}_{x\in\mathbb{R}^{N\times\ell}} \mathcal{R}(x) + \alpha||\mathcal{T}(x) - y||^2_F + \beta\,\mathrm{GL}_{\varepsilon,\mu,f}(u_n, \Omega(x)) + \eta_n||x - x_n||^2_F, \tag{3a}$$

$$u_{n+1} = \operatorname*{argmin}_{u\in\mathcal{V}} \beta\,\mathrm{GL}_{\varepsilon,\mu,f}(u, \Omega(x_{n+1})) + \nu_n||u|_Y - u_n|_Y||^2_{\mathcal{V}}. \tag{3b}$$

We can understand this scheme intuitively as iterating the following steps:

I. Given the current segmentation, update the reconstruction using the segmentation energy as an extra regulariser and the previous reconstruction as a momentum term.

II. Given the current reconstruction, update the segmentation using the previous segmentation of the image to be reconstructed as a momentum term.

Convergence analysis

We analyse the convergence of (3) to (2) using the theory of (Attouch *et al.*, 2010). If we assume that $\mathcal{R}$ is **sub-analytic**, then it follows that the joint energy in (2) has the **Kurdyka–Łojasiewicz (KŁ) property**. It therefore follows that, if $\mathcal{T}$ is **coercive** and the η_n, ν_n are bounded both above and away from zero:

- The joint energy monotonically decreases with each iteration.
- For all feasible (u_0, x_0), (u_n, x_n) converges to a critical point of the joint energy.
- If (u_0, x_0) is near a global minimum, then (u_n, x_n) converges to a global minimum.

Furthermore, there are provable convergence rates if the **KŁ exponent** of the joint energy is known.

N.B. To prove these results, we had to restrict the feasibility set so that $u|_Z = f$.

Conclusions

- We formulated joint reconstruction-segmentation, with graph-based segmentation, as a variational problem.
- We devised an iterative scheme for this problem, and developed algorithms for computing this scheme.
- We proved that this iterative scheme converges to critical points of the joint energy.
- We tested this scheme for a deblurring/denoising-segmentation task, with very promising results.

References

- A. N. Tikhonov, *Solution of incorrectly formulated problems and the regularization method*, Soviet Math. Dokl., 4 (1963), pp. 1035–1038.
- A. L. Bertozzi and A. Flenner, *Diffuse interface models on graphs for analysis of high dimensional data*, Multiscale Modeling and Simulation 10 (2012), no. 3, 1090–1118.
- J. Adler, S. Lunz, O. Verdier, C.-B. Schönlieb, and O. Öktem, *Task adapted reconstruction for inverse problems*, 2018, `https://arxiv.org/abs/1809.00948`
- V. Corona, M. Benning, M. J. Ehrhardt, L. F. Gladden, R. Mair, A. Reci, A. J. Sederman, S. Reichelt, and C.-B. Schönlieb, *Enhancing joint reconstruction and segmentation with non convex Bregman iteration*, Inverse Problems, 35 (2019), p. 055001.
- J. Budd, Y. van Gennip, and J. Latz, *Classification and image processing with a semi-discrete scheme for fidelity forced Allen–Cahn on graphs*, GAMM Mitteilungen, 44 (2021), pp. 1–43.
- H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, *Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka–Łojasiewicz inequality*, Mathematics of Operations Research, 35 (2010), pp. 438–457.

Turning an image into a graph

Let $V := Y \cup Z$ be the set of pixels in our images. Then our combined image (x, x_d) is a function from V to $\mathbb{R}^\ell$. To each pixel $i \in V$ we assign a **feature vector** $z_i \in \mathbb{R}^q$, via some (linear) feature map. These feature vectors will encode the "key information" about the pixel. Then we build our graph by defining the weight on edge $(i, j) \in V^2$ according to the similarity of the feature vectors z_i and z_j. For example:

$$\text{if } i \neq j \qquad \omega_{ij} = e^{-||z_i - z_j||^2_2/\sigma^2}, \qquad \text{if } i = j \qquad \omega_{ij} = 0.$$

We summarise all of this as $\omega = \Omega(x)$, where Ω is the **"image-to-graph" map**.

Solving (3a)

It is highly computationally challenging to solve (3a), because of the Ginzburg–Landau term. We therefore **linearise** that term. Let $\mathcal{G}(x) := \mathrm{GL}_{\varepsilon,\mu,f}(u_n, \Omega(x))$. Then linearising $\mathcal{G}$ around x_n, (3a) becomes:

$$\operatorname*{argmin}_x \mathcal{R}(x) + \alpha||\mathcal{T}(x) - y||^2 + \beta\langle x, \nabla_x\mathcal{G}(x_n)\rangle + \eta_n||x - x_n||^2$$

or equivalently,

$$\operatorname*{argmin}_x \mathcal{R}(x) + \alpha||\mathcal{T}(x) - y||^2 + \eta_n||x - \tilde{x}_n||^2 \tag{4}$$

where $\tilde{x}_n := x_n - \frac{1}{2}\beta\eta_n^{-1}\nabla_x\mathcal{G}(x_n)$ is a segmentation-driven adjustment of the previous reconstruction.

We can solve (4) by e.g. primal-dual methods, so it remains to compute $\nabla_x\mathcal{G}(x_n)$. For the choice of Ω above, this can be reduced to computing matrix-vector products with $\Omega(x)$. That matrix is too large to compute these products directly, but they can be approximated using the **Nyström extension**.

Solving (3b)

Because μ and f are zero on Y, we can rewrite the objective function in (3b) as:

$$\beta\,\mathrm{GL}_{\varepsilon,\mu',f'}(u, \Omega(x_{n+1}))$$

where $\mu' := \mu + 2\nu_n\beta^{-1}\chi_Y$ and $f' := f + u_n \odot \chi_Y$. We minimise this in u by numerically solving the Allen–Cahn equation, using the **SDIE scheme** described in (Budd, van Gennip, Latz, 2021).

Deblurring/denoising-segmentation results

Figure 1. Observations y. The y is obtained by applying a 25 pixel motion blur, then adding Gaussian noise with mean 0 and standard deviation 0.7. The PSNR of this relative to the original image is 6.55 dB. The image is 480×640 pixels.

Figure 2. Reference image x_d. f (not shown) is a hand-drawn label of the cows in this image. The image is 480×640 pixels.

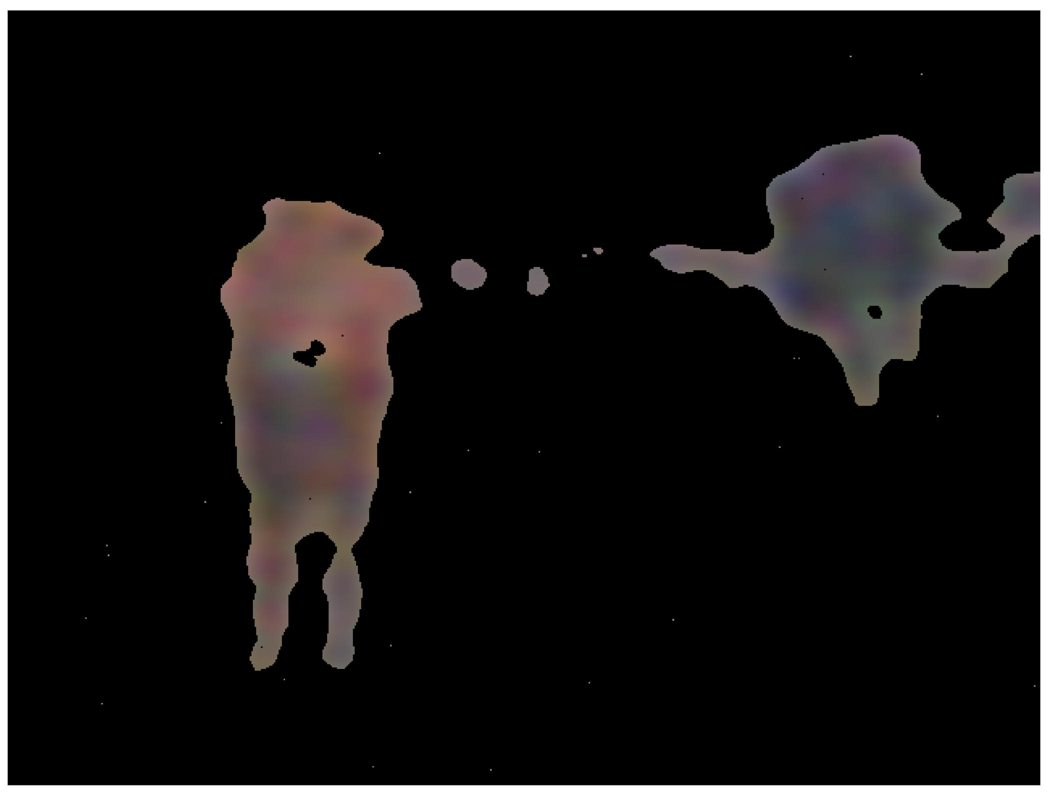

Figure 3. x_8 masked with the best segmentation u_8. The pixel accuracy of the segmentation is 95.53%, with Dice score 0.8237. The PSNR of the reconstruction was 17.93 dB. The computation time (on a basic laptop) to compute this reconstruction-segmentation was 158.56 seconds.