Image Processing and related PDEs
Lecture 4: Image segmentation

Yves van Gennip

School of Mathematical Sciences, University of Nottingham

Minicourse on Image Processing and related PDEs
University of Oxford

7 September, 2016

What's brewing?

- Mumford-Shah variational segmentation
- Chan-Vese variational segmentation
- Graph based variational segmentation

Mumford-Shah functional

We will consider $\Omega \subset \mathbb{R}^2$ to be open and bounded.

Mumford, Shah, 1989, Chan, Shen §7.4, Aubert, Kornprobst §4.2

Let $f \in L^2(\Omega)$ and $\alpha, \beta > 0$. Then we minimise

$$(u^*, \Gamma^*) \in \operatorname*{arg\,min}_{(u,\Gamma)} F(u, \Gamma),$$

where

$$F(u, \Gamma) := \alpha \mathcal{H}^1(\Gamma) + \beta \int_{\Omega \setminus \Gamma} |\nabla u|^2 + \int_{\Omega \setminus \Gamma} (u - f)^2.$$

- Admissible $u \in W^{1,2}(\Omega \setminus \Gamma) \cap L^\infty(\Omega)$
- Admissible $\Gamma \subset \Omega$ closed, with $\mathcal{H}^1(\Omega) < \infty$
- Existence proof problem! The function $A \mapsto \mathcal{H}^1(\partial A)$ is not lower semicontinuous w.r.t. any compact topology (Aubert, Kornprobst §4.2.2).

Mumford-Shah reformulated

Reformulation, De Giorgi, Carriero, Leaci, 1989

$$G(u) := \alpha\mathcal{H}^1(S_u) + \beta\int_{\Omega\setminus S_u}|\nabla u|^2 + \int_{\Omega\setminus S_u}(u-f)^2,$$

where S_u is the *jump set* of u.

- If $u \in BV(\Omega)$, then Du, the distributional derivative of u, can be decomposed into the sum of an absolutely continuous (w.r.t. the Lebesgue measure) part and a singular part.
- The latter in turn can be decomposed into the sum of a part with support on the jump set of u and a remainder part, which is called the *Cantor part*.
- There exist a nonconstant BV function which is continuous and has zero distributional derivative almost everywhere and only has a Cantor part: the Cantor-Vitali function (Ambrosio, Fusco, Pallara Examples 1.67, 3.34).

Special functions of bounded variation

- Even worse, the class of $BV(\Omega)$ functions with zero absolutely continuous part is dense in $L^2(\Omega)$ (Ambrosio, 1998).
- If $\{v\}_n$ is a sequence of such functions $L^2(\Omega)$-converging to f, then for all n,

$$0 \leq \inf_{u \in BV(\Omega)} G(u) \leq G(v_n) = \int_\Omega (v_n - f)^2.$$

  Letting $n \to \infty$, we find $\inf_{u \in BV(\Omega)} G(u) = 0$, hence G has no minimiser in general.
- Therefore we restrict G to *special functions of bounded variation*. A function u is in $SBV(\Omega)$ if it is in $BV(\Omega)$ and its distributional derivative has no Cantor part.
- To be precise: $G : SBV(\Omega) \cap L^\infty(\Omega) \to \mathbb{R}$
- *SBV* has the right kind of compactness properties and $\mathcal{H}^1$ is lower semicontinuous w.r.t. to the right topology to make an existence proof work (details in Ambrosio, Fusco, Pallara §4, Aubert, Kornprobst §4.2).

Back to the original functional

- Finally we connect back to the original Mumford-Shah functional F.
- If u^* is a minimiser of G over $SBV(\Omega)$, then it can be shown that $(u^*, \Omega \cap \overline{S_{u^*}})$ is a minimiser of F, over $\left[W^{1,2}(\Omega \setminus \Gamma) \cap L^\infty(\Omega)\right] \times \left\{\Gamma \subset \Omega : \Gamma \text{ closed and } \mathcal{H}^1(\Omega) < \infty\right\}$.
- Details in De Giorgi, Carriero, Leaci, 1989, Aubert, Kornprobst §4.2.

Limit for $\beta \to \infty$

Formally[1], if we send $\beta \to \infty$ in Mumford-Shah, we are left with minimising

$$G_\infty(u, \Gamma) := \alpha \mathcal{H}^1(\Gamma) + \int_{\Omega \setminus \Gamma} (u - f)^2$$

over a restricted admissible set, which requires that the distributional derivative of u is zero on $\Omega \setminus \Gamma$.
If we (uniquely, up to labelling) decompose $\Omega \setminus \Gamma = \bigcup_i \Omega_i$ into connected components, then for every i, $u|_{\Omega_i}$ must be constant. In fact

$$u|_{\Omega_i} = \frac{1}{|\Omega_i|} \int_{\Omega_i} f$$

is the optimal choice for the constants.
The Chan-Vese method (Chan, Vese, 2001) builds on this idea.

[1] There are ways of making this precise.

Chan-Vese method

Chan, Vese, 2001

$$H(c_1, c_2, \Gamma) := \alpha \mathcal{H}^1(\Gamma) + \nu |A| + \lambda_1 \int_A |c_1 - f|^2 + \lambda_2 \int_A |c_2 - f|^2,$$

where A is the region inside the closed curve Γ.
Minimisation is over constants $c_1, c_2 \in \mathbb{R}$ and closed curves Γ that are the boundary of open, bounded sets $A \subset \mathbb{R}^2$.

- Minimisers exist (Chan, Vese, 2001, gives some references).
- A reformulation in terms of level sets (Osher, Setian, 1988) is useful. This allows us to minimise over functions u, instead of sets, such that $\Gamma = \{x \in \Omega : u(x) = 0\}$.

Examples

Show example(s) on Image Processing On Line:

`http://www.ipol.im/`

Graph based image segmentation

(Olive Oil by Tétine)

(Calatroni, vG, Schönlieb, Rowland, Flenner, 2016)

Variational model for phase separation

Continuum Ginzburg-Landau functional

Let $W(x) = x^2(x-1)^2$ and $\varepsilon > 0$. Minimise

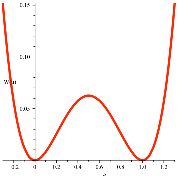
$$u^* \in \underset{u \in W^{1,2}(\Omega)}{\arg\min} \; GL(u) + \dots,$$

where

$$GL(u) := \frac{\varepsilon}{2} \int_\Omega |\nabla u|^2 + \frac{1}{\varepsilon} \int_\Omega W(u).$$

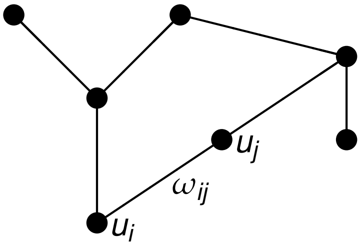
The ... indicate that, in order to avoid trivial minimisers, we need to add extra terms to the functional (e.g. a fidelity term) or add extra constraints to the set of admissible functions (e.g. a mass constraint).

Double well potential W:



New setting: graphs

Data points (pixels) are represented by nodes in an undirected graph $G = (V, E)$. Similarity is encoded in edge weights ω_{ij}. 'Object' and 'background' are distinguished by node labels $u_i : V \to \mathbb{R}$.



For simplicity, cluster into two groups: $\{u = 0\}$ and $\{u = 1\}$.

Phase separation on graphs

Ginzburg-Landau functional + fidelity to a priori known labels

$$\min_{u:V\to\mathbb{R}} \tfrac{1}{2}\sum_{i,j}\omega_{ij}(u_i-u_j)^2+\tfrac{1}{\varepsilon}\sum_i W(u_i)+\tfrac{1}{2}\sum_i \lambda_i(u_i-(u_{\text{known}})_i)^2$$

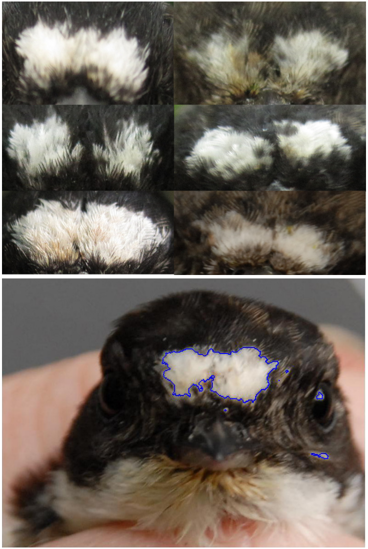
$$\lambda_i = \begin{cases} 1 & i \text{ in known} \\ 0 & \text{otherwise} \end{cases}$$

(method from Bertozzi, Flenner, 2012)

Gradient flow: Allen-Cahn equation + fidelity

$$\begin{aligned} u_i &= -\sum_j \omega_{ij}(u_i-u_j)-\frac{1}{\varepsilon}W'(u_i)-\lambda_i(u_i-(u_{\text{known}})_i) \\ &= -(\Delta u)_i-\frac{1}{\varepsilon}W'(u_i)-\lambda_i(u_i-(u_{\text{known}})_i) \end{aligned}$$

Ginzburg-Landau segmentation



(Calatroni, vG, Schönlieb, Rowland, Flenner, 2016)

- Vertices: the pixels
- Edge weight: $\omega_{ij} = e^{-\|x_i - x_j\|^2/\tau}$ where τ is a scale parameter and
- x_i is the feature vector of pixel i
- Constraint: fidelity to data in original image

Translate PDEs to graphs

$\mathcal{V}$: Functions on the nodes $u, v : V \to \mathbb{R}$

$\mathcal{E}$: Functions on the edges $\varphi, \phi : E \to \mathbb{R}$

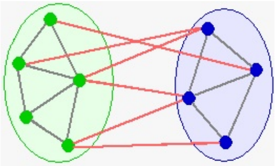
Build (choose!) a differential structure on the graph (r is a parameter):

- Node degree $d_i = \sum_j \omega_{ij}$
- $\langle u, v \rangle_{\mathcal{V}} = \sum_i d_i^r u_i v_i$
- $\langle \varphi, \phi \rangle_{\mathcal{E}} = \frac{1}{2} \sum_{i,j} \omega_{ij} \phi_{ij} \varphi_{ij}$
- Gradient: $(\nabla u)_{ij} = \begin{cases} (u_j - u_i) & \text{if } i \sim j, \\ 0 & \text{else} \end{cases}$
- Divergence: $(\text{div}\, \varphi)_i = d_i^{-r} \sum_j \omega_{ij} \varphi_{ji}$
- Laplacian: $(\Delta u)_i = (\text{div}\, \nabla u)_i = d_i^{-r} \sum_j \omega_{ij} (u_i - u_j)$

  $r = 0$: unnormalized Laplacian; $r = 1$: random walk Laplacian; symmetric normalized Laplacian: not in this framework

Dirichlet energy, total variation

- Dirichlet energy: $\frac{1}{2}\|\nabla u\|_{\mathcal{E}}^2 = \frac{1}{4}\sum_{ij}\omega_{ij}(u_i - u_j)^2$
  The Laplacian is the 1st variation of the Dirichlet energy.
- (Anisotropic) total variation:
  $TV_a(u) = \max\{\langle \operatorname{div}\varphi, u\rangle_{\mathcal{V}} : \max_{i,j}\ |\varphi_{ij}| \leq 1\} = \frac{1}{2}\sum_{ij}\omega_{ij}|u_i - u_j|$
- ...is the graph cut objective functional, if $u = \chi_S$ (indicator function of node set S).



Γ-convergence and its relevance to minimisation

If $\mathrm{GL}_\varepsilon \xrightarrow{\Gamma} \mathrm{GL}_0$ and a compactness property holds, then:

If u_ε minimizes GL_ε and $u_\varepsilon \to u_0$, then u_0 minimizes GL_0

Definitions:

Γ-limit of sequence of functionals

A sequence $\{F_n\}$ Γ-converges to F_0 as $n \to \infty$ if, for all $u \in \mathrm{Dom}(F)$

1. $\forall u_n \to u \; \liminf_{n\to\infty} F_n(u_n) \geq F(u)$ and
2. $\exists u_n \to u \; \limsup_{n\to\infty} F_n(u_n) \leq F(u)$.

Compactness property

Compactness: $F_n(u_n) < C \Rightarrow \{u_n\}$ has a convergent subsequence.

Γ-convergence on graphs

For a fixed general graph with edge weights ω_{ij} independent of ε

$$\sum_{ij} \omega_{ij}(u_i - u_j)^2 + \frac{1}{\varepsilon}\sum_i W(u_i) \xrightarrow[\varepsilon\to 0]{\Gamma} \begin{cases} \sum_{i,j}\omega_{ij}|u_i - u_j|, & \text{if } u = \chi_S \text{ for some } S,\\ +\infty, & \text{otherwise.}\end{cases}$$

(vG, Bertozzi, 2012)

This is analogous to a famous result in the continuum setting (Modica, Mortola, 1977, Modica 1987, Modica 1987):

$$GL(u) \xrightarrow[\varepsilon\to 0]{\Gamma} \begin{cases} \int_\Omega |\nabla u|, & \text{if } u = \chi_S \text{ for some set } S \text{ of finite perimeter,}\\ +\infty, & \text{otherwise.}\end{cases}$$

Bibliography: image analysis specific references

- L. Calatroni, Y. van Gennip, C.-B. Schönlieb, H.M. Rowland, A. Flenner, *Graph Clustering, Variational Image Segmentation Methods and Hough Transform Scale Detection for Object Measurement in Images*, J. Math. Imaging Vis, 2016, Springer Online First DOI 10.1007/s10851-016-0678-0
- T.F. Chan, J. Shen, *Image Processing and Analysis*, SIAM, 2005
- T.F. Chan, L.A. Vese, *Active Contours Without Edges*, IEEE Transactions on Image Processing, 10(2), 2001, 26–277
- E. De Giorgi, M. Carriero, A. Leaci, *Existence theorem for a minimum problem with free discontinuity set*, Archive for Rational Mechanics and Analysis 108(4), 1989, 195–218.
- D. Mumford, J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm Pure Appl. Math, 42, 1989, 577–685
- C.-B. Schönlieb, *Image Processing — Variational and PDE Methods*, online lecture notes, 2013/14, http://www.damtp.cam.ac.uk/user/cbs31/Teaching_files/PDEImageLectureNotesLent2014.pdf

Bibliography: PDEs on graphs references

- A.L. Bertozzi, A. Flenner, *Diffuse interface models on graphs for classification of high dimensional data*, Multiscale Modeling & Simulation 10(3), 2012, 1090–1118
- Y. van Gennip, A.L. Bertozzi, Γ*-convergence of graph Ginzburg-Landau functionals*, Advances in Differential Equations, 17(11–12), 2012, 1115–1180
- Y. van Gennip, N. Guillen, B. Osting, A.L. Bertozzi, *Mean Curvature, Threshold Dynamics, and Phase Field Theory on Finite Graphs*, Milan Journal of Mathematics, 82(1), 2014, 3–65

Bibliography: general PDE and analysis references

- L. Ambrosio, *Free discontinuity problems and special functions with bounded variation*, Progress in Mathematics, 168, 1998, 15–35
- E. De Giorgi, M. Carriero, A. Leaci, *Existence theorem for a minimum problem with free discontinuity set*, Archive for Rational Mechanics and Analysis 108(4), 1989, 195–218.
- L. Modica, *The gradient theory of phase transitions and the minimal interface criterion*, Arch. Rational Mech. Anal., 98, 1987, 123–142
- L.Modica, *Gradient theory of phase transitions with boundary contact energy*, Ann. Inst. Henri Poincaré, 4, 1987, 487–512
- L. Modica, S. Mortola, *Un esempio di Γ-convergenza*, Bollettino U.M.I., 5, 1977, 285–299
- S. Osher, J.A. Sethian, *Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations*, Journal of Computational Physics 79(1), 1988, 12–49