MA3111: Mathematical Image Processing
Variational Image Segmentation

Suh-Yuh Yang (楊肅煜)

Department of Mathematics, National Central University
Jhongli District, Taoyuan City 320317, Taiwan

First version: June 09, 2021/Last updated: July 2, 2025

Outline of "variational image segmentation"

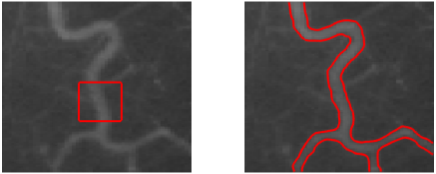
In this lecture, we will give a brief introduction to the topics:

- *The energy-based models for image segmentation: the Mumford-Shah model and the Chan-Vese model based on the level set formulation.*
- *An efficient iterative thresholding method for model implementation.*

The material of this lecture is mainly based on

- P. Getreuer, Chan-Vese segmentation, *Image Processing On Line,* 2 (2012), pp. 214-224.
- D. Wang, H. Li, X. Wei, X.-P. Wang (王筱平), An efficient iterative thresholding method for image segmentation, *Journal of Computational Physics,* 350 (2017), pp. 657-667.

Image segmentation in medical imaging



f & initialization C *segmented image*

In what follows, Ω denotes an open bounded subset in $\mathbb{R}^2$ and $f : \overline{\Omega} \to \mathbb{R}$ denotes the given grayscale image to be segmented.

Mumford-Shah model (CPAM 1989)

Mumford-Shah model: it finds a piecewise smooth function u and a curve set $\mathcal{C}$, which separates the image domain into disjoint regions, minimizing the energy functional:

$$\min_{u,\mathcal{C}} \left(\mu|\mathcal{C}| + \lambda \int_{\Omega} (f(\mathbf{x}) - u(\mathbf{x}))^2 \, d\mathbf{x} + \int_{\Omega\setminus\mathcal{C}} |\nabla u(\mathbf{x})|^2 \, d\mathbf{x} \right),$$

where $|\mathcal{C}|$ denotes the total length of the curves in $\mathcal{C}$.

- The first term plays the regularization role, which ensures the target objects can tightly be wrapped by $\mathcal{C}$.
- The second term is the data fidelity term, which forces u to be close to the input image f.
- The third term is the smoothing term, which forces the target function u to be piecewise smooth within each of the regions separated by the curves in $\mathcal{C}$.
- $\mu > 0$, $\lambda > 0$ *are tuning parameters to modulate these three terms.*

Simplified Mumford-Shah model

- *The non-convexity of energy functional in the Mumford-Shah model* makes the minimization problem difficult to analyze and the computational cost is much considerable.
- The piecewise smooth model suffers for its *sensitivity to the initialization of $\mathcal{C}$.*
- **Simplified Mumford-Shah model:** it finds *a piecewise constant function u* and a curve set $\mathcal{C}$ to minimize the energy functional:

$$\min_{u,\mathcal{C}} \left(\mu \left| \mathcal{C} \right| + \int_\Omega \left(f(\boldsymbol{x}) - u(\boldsymbol{x}) \right)^2 d\boldsymbol{x} \right).$$

  Note that u is constant on each connected component of $\Omega \setminus \mathcal{C}$. *The minimization problem is still non-convex.*

Chan (陳繁昌)-Vese two-phase model

In 1999, Chan and Vese proposed a two-phase segmentation model based on the level set formulation ("active contours without edges", LNCS 1999):

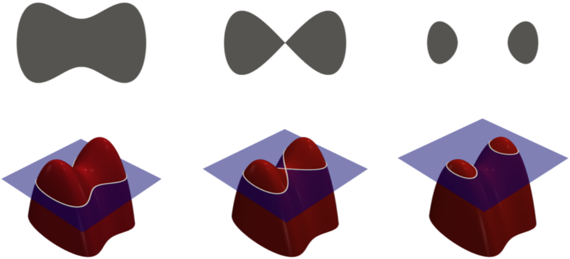
$$\min_{c_1,c_2,\mathcal{C}} \left(\mu|\mathcal{C}| + \nu|\Omega_{\text{in}}| + \lambda_1 \int_{\Omega_{\text{in}}} (f(\boldsymbol{x}) - c_1)^2 \, d\boldsymbol{x} + \lambda_2 \int_{\Omega_{\text{out}}} (f(\boldsymbol{x}) - c_2)^2 \, d\boldsymbol{x} \right).$$

- Ω_{in} represents the region enclosed by and contains the curves in $\mathcal{C}$ with area $|\Omega_{\text{in}}|$, and $\Omega_{\text{out}} := \Omega \setminus \Omega_{\text{in}}$.
- $\mu > 0$, $\nu \geq 0$, $\lambda_1 > 0$, and $\lambda_2 > 0$ are tuning parameters (actually, one of them can be fixed as 1).
- Chan-Vese model finds a piecewise constant function u and a curve set $\mathcal{C}$ to minimize the energy functional, where u has only two constant values,

$$u(\boldsymbol{x}) = \begin{cases} c_1, & \boldsymbol{x} \text{ is inside } \mathcal{C}, \\ c_2, & \boldsymbol{x} \text{ is outside } \mathcal{C}. \end{cases}$$

Topological changes of $\mathcal{C}$

To solve the minimization problem of Chan-Vese model, we evolve $\mathcal{C}$ and find c_1, c_2 to minimize the energy functional. However, it is generally hard to handle *topological changes* of the curves in $\mathcal{C}$.



(quoted from wikipedia)

Level set function

Therefore, we represent $\mathcal{C}$ implicitly by the zero level contour of a level set function $\phi : \overline{\Omega} \to \mathbb{R}$, i.e.,

$$\mathcal{C} = \{\boldsymbol{x} \in \overline{\Omega} : \ \phi(\boldsymbol{x}) = 0\}.$$

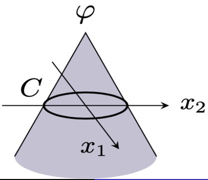
The zero level contour $\mathcal{C}$ partitions the image domain into two disjoint regions Ω_{in} and Ω_{out} such that

$$\phi(\boldsymbol{x}) \geq 0 \ \text{ for } \ \boldsymbol{x} \in \Omega_{\text{in}} \quad \text{and} \quad \phi(\boldsymbol{x}) < 0 \ \text{ for } \ \boldsymbol{x} \in \Omega_{\text{out}}.$$

For example, given $r > 0$, we define a level set function, which is a *signed distance function*,

$$\phi(\boldsymbol{x}) = \phi(x, y) = r - \sqrt{x^2 + y^2},$$

whose zero level contour is the circle of radius $r > 0$.



Chan-Vese model

- Let H denote the Heaviside function and δ the Dirac delta function. Then

$$H(s) = \begin{cases} 1 & s \geq 0, \\ 0 & s < 0, \end{cases} \qquad \text{and} \qquad \frac{d}{ds}H(s) = \delta(s).$$

- In terms of H, δ, and the level set function ϕ, the Chan-Vese model has the form

$$\begin{aligned} \min_{c1,\, c2,\, \phi} \Big(\mu \int_\Omega \delta(\phi(\boldsymbol{x})) |\nabla \phi(\boldsymbol{x})| \, d\boldsymbol{x} + \nu \int_\Omega H(\phi(\boldsymbol{x})) \, d\boldsymbol{x} \\ + \lambda_1 \int_\Omega (f(\boldsymbol{x}) - c_1)^2 H(\phi(\boldsymbol{x})) \, d\boldsymbol{x} \\ + \lambda_2 \int_\Omega (f(\boldsymbol{x}) - c_2)^2 (1 - H(\phi(\boldsymbol{x}))) \, d\boldsymbol{x} \Big). \end{aligned}$$

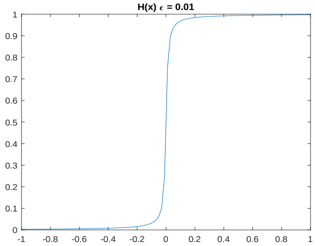
Original formulation:

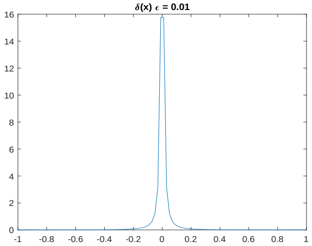
$$\min_{c_1, c_2, \mathcal{C}} \left(\mu |\mathcal{C}| + \nu |\Omega_{\text{in}}| + \lambda_1 \int_{\Omega_{\text{in}}} (f(\boldsymbol{x}) - c_1)^2 + \lambda_2 \int_{\Omega_{\text{out}}} (f(\boldsymbol{x}) - c_2)^2 \right).$$

The regularized Heaviside and delta functions

The Heaviside function H and the Dirac delta function δ can be approximately regularized as follows: for a sufficiently small $\epsilon > 0$,

$$
\begin{aligned}
H_\epsilon(t) &:= \frac{1}{2}\left(1 + \frac{2}{\pi}\tan^{-1}\left(\frac{t}{\epsilon}\right)\right), \\
\delta_\epsilon(t) &:= \frac{d}{dt}H_\epsilon(t) = \frac{\epsilon}{\pi(\epsilon^2 + t^2)}, \\
\int_{-\infty}^{\infty} \delta_\epsilon(t)dt &= \int_{-\infty}^{\infty} \frac{\epsilon}{\pi(\epsilon^2 + t^2)}dt = \cdots = 1.
\end{aligned}
$$





Total length of $\mathcal{C}$

The first term of the energy functional is the length of $\mathcal{C}$, which can be expressed as the total variation of $H(\phi)$,

$$\begin{aligned} |\mathcal{C}| &= \int_\Omega \delta(\phi(\boldsymbol{x}))|\nabla\phi(\boldsymbol{x})|d\boldsymbol{x} = \int_\Omega \left|\frac{dH}{d\phi}(\phi(\boldsymbol{x}))\right||\nabla\phi(\boldsymbol{x})|d\boldsymbol{x} \\ &= \int_\Omega |\nabla H(\phi(\boldsymbol{x}))|d\boldsymbol{x}. \end{aligned}$$

A heuristic argument to prove $|\mathcal{C}| = \int_\Omega \delta(\phi(\boldsymbol{x}))|\nabla\phi(\boldsymbol{x})|d\boldsymbol{x}$:

Suppose that the level set function ϕ is a signed distance function, i.e.,

$$\phi(\boldsymbol{x}) = \begin{cases} d(\boldsymbol{x},\mathcal{C}) & \text{if } \boldsymbol{x} \in \Omega_{\text{in}}, \\ -d(\boldsymbol{x},\mathcal{C}) & \text{if } \boldsymbol{x} \in \Omega_{\text{out}}. \end{cases}$$

Then $\phi(\boldsymbol{x})$ is differentiable almost everywhere and $|\nabla\phi(\boldsymbol{x})| = 1$ for $\boldsymbol{x} \in \overline{\Omega}$ a.e. The contour $\mathcal{C}$ can be parametrized in arc length s, $\boldsymbol{z}(s) = (x(s), y(s))$ for $0 \le s \le L := |\mathcal{C}|$. Let $N \gg 1$ be a large integer. We approximate the Dirac δ-function by

$$\delta_N(t) := \begin{cases} N, & |t| \le \frac{1}{2N}, \\ 0, & \text{otherwise.} \end{cases}$$

Total length of $\mathcal{C}$: a heuristic argument (cont'd)

Let B_N be the narrow band defined by

$$B_N := \{\boldsymbol{x} \in \overline{\Omega} : \ |\phi(\boldsymbol{x})| \leq 1/(2N)\}.$$

Then we have

$$\int_{\Omega} \delta(\phi(\boldsymbol{x}))|\nabla\phi(\boldsymbol{x})|d\boldsymbol{x} \approx N \int_{B_N} |\nabla\phi(\boldsymbol{x})|d\boldsymbol{x}.$$

The "centerline" of this band B_N is the curve $\mathcal{C} = \{\boldsymbol{x} \in \overline{\Omega} : \ \phi(\boldsymbol{x}) = 0\}$. Consider a point $\boldsymbol{p} = \boldsymbol{z}(s) \in \mathcal{C}$. Then the tangent vector and the normal vector are $\boldsymbol{z}'(s) = (x'(s), y'(s))$ and $\nabla\phi(\boldsymbol{z}(s))$, respectively. Starting at $\boldsymbol{p}$ in the direction $\nabla\phi(\boldsymbol{p})$, we reach the boundary of B_N when we have traversed the length $h > 0$ such that $|\nabla\phi(\boldsymbol{p})|h = \frac{1}{2N}$. It follows that near $\boldsymbol{p} = \boldsymbol{z}(s)$ the width $\rho(s)$ of this band is approximately given by

$$\rho(s) = 2h = \frac{1}{N|\nabla\phi(\boldsymbol{z}(s))|} = \frac{1}{N}.$$

Therefore we have

$$\int_{\Omega} \delta(\phi(\boldsymbol{x}))|\nabla\phi(\boldsymbol{x})|d\boldsymbol{x} \approx N \int_{B_N} |\nabla\phi(\boldsymbol{x})|d\boldsymbol{x} \approx N \int_0^L \rho(s)ds = L = |\mathcal{C}|.$$

An alternating iterative scheme

The minimization is solved by *an alternating iterative scheme*, i.e., alternatingly updating c_1, c_2 and ϕ.

(S1) Fixed ϕ, the optimal values of c_1 and c_2 are the region averages,

$$c_1 = \frac{\int_\Omega f(\boldsymbol{x}) H(\phi(\boldsymbol{x}))\, d\boldsymbol{x}}{\int_\Omega H(\phi(\boldsymbol{x}))\, d\boldsymbol{x}}, \quad c_2 = \frac{\int_\Omega f(\boldsymbol{x}) \big(1 - H(\phi(\boldsymbol{x}))\big)\, d\boldsymbol{x}}{\int_\Omega \big(1 - H(\phi(\boldsymbol{x}))\big)\, d\boldsymbol{x}}.$$

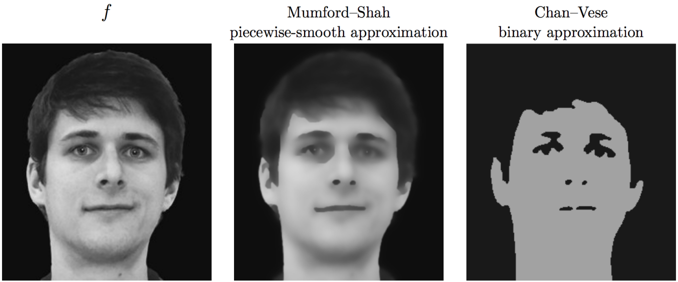
(S2) Fixed c_1, c_2, we solve the initial-boundary value problem (IBVP) to reach a steady-state:

$$\frac{\partial \phi}{\partial t} = \delta_\epsilon(\phi)\left(\mu \nabla \cdot \frac{\nabla \phi}{|\nabla \phi|} - \nu - \lambda_1 (f - c_1)^2 + \lambda_2 (f - c_2)^2\right), \quad \text{for } t > 0, \boldsymbol{x} \in \Omega,$$

$$\phi(0, \boldsymbol{x}) = \phi_0(\boldsymbol{x}), \boldsymbol{x} \in \Omega,$$

$$\frac{\partial \phi}{\partial \boldsymbol{n}} = 0 \text{ on } \partial\Omega, t \geq 0.$$

Example: Mumford-Shah vs. Chan-Vese



P. Getreuer, Chan-Vese segmentation,
Image Processing On Line, 2 (2012), pp. 214-224.

Numerical implementation

- Assume that the image domain $\overline{\Omega}$ is the unit square $[0,1] \times [0,1]$.
- Let $\Omega_D := \{(x_i, y_j) |\ i,j = 0,1,\cdots,M\}$ be the set of grid points of a uniform partition of $\overline{\Omega}$ with size $h = 1/M$.
- Then $x_i = ih$ and $y_j = jh$, $i,j = 0,1,\cdots,M$. Let $\phi_{i,j}(t)$ be the spatial difference approximation to $\phi(t, x_i, y_j)$.
- Let $t_n = n\Delta t$, $n \geq 0$, and $\Delta t > 0$ be the time step, and let $\phi^n_{i,j}$ be the full difference approximation to $\phi(t_n, x_i, y_j)$.

Discrete differential operators and BC

- *Define the discrete differential operators: for* $1 \le i,j \le M-1$,

$$
\begin{aligned}
\nabla_x^+ \phi_{i,j} &= \frac{\phi_{i+1,j} - \phi_{i,j}}{h}, \text{ (forward difference)} \\
\nabla_x^- \phi_{i,j} &= \frac{\phi_{i,j} - \phi_{i-1,j}}{h}, \text{ (backward difference)} \\
\nabla_y^+ \phi_{i,j} &= \frac{\phi_{i,j+1} - \phi_{i,j}}{h}, \text{ (forward difference)} \\
\nabla_y^- \phi_{i,j} &= \frac{\phi_{i,j} - \phi_{i,j-1}}{h}, \text{ (backward difference)} \\
\nabla_x^0 \phi_{i,j} &:= \left(\frac{\nabla_x^+ + \nabla_x^-}{2}\right)\phi_{i,j}, \quad \nabla_y^0 \phi_{i,j} := \left(\frac{\nabla_y^+ + \nabla_y^-}{2}\right)\phi_{i,j}. \\
& \quad \text{(central differences)}
\end{aligned}
$$

- *Discretize the homogeneous Neumann BC:* $\dfrac{\partial \phi}{\partial \boldsymbol{n}} = 0$ *on* $\partial\Omega$

$$
\phi_{0,j} = \phi_{1,j}, \quad \phi_{M,j} = \phi_{M-1,j}, \quad \phi_{i,0} = \phi_{i,1}, \quad \phi_{i,M} = \phi_{i,M-1}.
$$

Finite difference discretization: spatial variables

Performing the spatial discretization [Getreuer-2012], we have

$$
\begin{aligned}
\frac{\partial \phi_{i,j}}{\partial t} &= \delta_\epsilon(\phi_{i,j})\Bigg\{\mu\Bigg(\nabla_x^- \frac{\nabla_x^+\phi_{i,j}}{\sqrt{\eta^2 + (\nabla_x^+\phi_{i,j})^2 + (\nabla_y^0\phi_{i,j})^2}} \\
&\qquad + \nabla_y^- \frac{\nabla_y^+\phi_{i,j}}{\sqrt{\eta^2 + (\nabla_x^0\phi_{i,j})^2 + (\nabla_y^+\phi_{i,j})^2}}\Bigg) \\
&\qquad - \nu - \lambda_1(f_{i,j} - c_1)^2 + \lambda_2(f_{i,j} - c_2)^2\Bigg\},
\end{aligned}
$$

where $i, j = 1, 2, \cdots, M-1$.

The purpose of small positive parameter η in the denominators prevents division by zero.

Spatial discretization

Define

$$
\begin{aligned}
A_{i,j} &= \frac{\mu}{\sqrt{\eta^2 + (\nabla_x^+ \phi_{i,j})^2 + (\nabla_y^0 \phi_{i,j})^2}}, \\
B_{i,j} &= \frac{\mu}{\sqrt{\eta^2 + (\nabla_x^0 \phi_{i,j})^2 + (\nabla_y^+ \phi_{i,j})^2}}.
\end{aligned}
$$

Using the fact $\nabla_x^+ \phi_{i,j} = \frac{\phi_{i+1,j} - \phi_{i,j}}{h}$, $\nabla_y^+ \phi_{i,j} = \frac{\phi_{i,j+1} - \phi_{i,j}}{h}$ and taking the backward difference at $A_{i,j}(\phi_{i+1,j} - \phi_{i,j})$ and $B_{i,j}(\phi_{i,j+1} - \phi_{i,j})$, then the discretization can be written as

$$
\begin{aligned}
\frac{\partial \phi_{i,j}}{\partial t} = \delta_\epsilon(\phi_{i,j}) \Big\{ & \frac{1}{h^2}\Big(A_{i,j}(\phi_{i+1,j} - \phi_{i,j}) - A_{i-1,j}(\phi_{i,j} - \phi_{i-1,j}) \Big) \\
& + \frac{1}{h^2}\Big(B_{i,j}(\phi_{i,j+1} - \phi_{i,j}) - B_{i,j-1}(\phi_{i,j} - \phi_{i,j-1}) \Big) \\
& - \nu - \lambda_1 (f_{i,j} - c_1)^2 + \lambda_2 (f_{i,j} - c_2)^2 \Big\}.
\end{aligned}
$$

Temporal discretization

Define

$$\widetilde{A}_{i,j} = \frac{1}{h^2}A_{i,j}, \quad \widetilde{A}_{i-1,j} = \frac{1}{h^2}A_{i,j},$$
$$\widetilde{B}_{i,j} = \frac{1}{h^2}B_{i,j}, \quad \widetilde{B}_{i,j-1} = \frac{1}{h^2}B_{i,j-1}.$$

Time is discretized with a semi-implicit Gauss-Seidel method, values $\phi_{i,j}$, $\phi_{i-1,j}$, $\phi_{i,j-1}$ are evaluated at time t_{n+1} and all others at time t_n.

$$\begin{aligned}\frac{\phi_{i,j}^{n+1} - \phi_{i,j}^{n}}{\Delta t} \quad = \quad & \delta_\epsilon(\phi_{i,j}^n)\Big\{\widetilde{A}_{i,j}\phi_{i+1,j}^n + \widetilde{A}_{i-1,j}\phi_{i-1,j}^{n+1} + \widetilde{B}_{i,j}\phi_{i,j+1}^n + \widetilde{B}_{i,j-1}\phi_{i,j-1}^{n+1} \\ & -\left(\widetilde{A}_{i,j} + \widetilde{A}_{i-1,j} + \widetilde{B}_{i,j} + \widetilde{B}_{i,j-1}\right)\phi_{i,j}^{n+1} \\ & -\nu - \lambda_1(f_{i,j} - c_1)^2 + \lambda_2(f_{i,j} - c_2)^2\Big\}.\end{aligned}$$

Gauss-Seidel scheme

This allows ϕ at time t_{n+1} to be solved by one Gauss-Seidel *sweep from left to right, bottom to top:*

$$\begin{aligned}
\phi_{i,j}^{n+1} &= \Big\{ \phi_{i,j}^{n} + \Delta t \delta_\epsilon(\phi_{i,j}^{n}) \Big(\widetilde{A}_{i,j}\phi_{i+1,j}^{n} + \widetilde{A}_{i-1,j}\phi_{i-1,j}^{n+1} + \widetilde{B}_{i,j}\phi_{i,j+1}^{n} \\
&\qquad + \widetilde{B}_{i,j-1}\phi_{i,j-1}^{n+1} - \nu - \lambda_1 (f_{i,j} - c_1)^2 + \lambda_2 (f_{i,j} - c_2)^2 \Big) \Big\} \\
&\qquad \times \Big\{ 1 + \Delta t \delta_\epsilon(\phi_{i,j}) \Big(\widetilde{A}_{i,j} + \widetilde{A}_{i-1,j} + \widetilde{B}_{i,j} + \widetilde{B}_{i,j-1} \Big) \Big\}^{-1},
\end{aligned}$$

where

$$\begin{aligned}
\widetilde{A}_{i,j} &= \frac{\mu}{h^2 \sqrt{\eta^2 + \Big((\phi_{i+1,j}^{n} - \phi_{i,j}^{n})/h \Big)^2 + \Big((\phi_{i,j+1}^{n} - \phi_{i,j-1}^{n+1})/(2h) \Big)^2}}, \\
\widetilde{B}_{i,j} &= \frac{\mu}{h^2 \sqrt{\eta^2 + \Big((\phi_{i+1,j}^{n} - \phi_{i-1,j}^{n+1})/(2h) \Big)^2 + \Big((\phi_{i,j}^{n} - \phi_{i+1,j}^{n})/h \Big)^2}}.
\end{aligned}$$

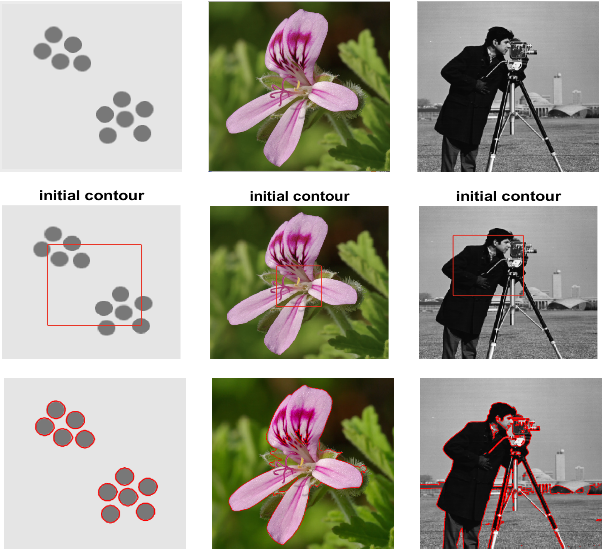
Gauss-Seidel scheme

We can rewrite $\widetilde{A}_{i,j}$ and $\widetilde{B}_{i,j}$ as follows:

$$\begin{aligned}
\widetilde{A}_{i,j} &= \frac{\mu}{h^2\sqrt{\eta^2 + \left((\phi^n_{i+1,j} - \phi^n_{i,j})/h\right)^2 + \left((\phi^n_{i,j+1} - \phi^{n+1}_{i,j-1})/(2h)\right)^2}}, \\
&= \frac{(\mu/h)}{\sqrt{(h\eta)^2 + (\phi^n_{i+1,j} - \phi^n_{i,j})^2 + \left((\phi^n_{i,j+1} - \phi^{n+1}_{i,j-1})/2\right)^2}}, \\
\widetilde{B}_{i,j} &= \frac{\mu}{h^2\sqrt{\eta^2 + \left((\phi^n_{i+1,j} - \phi^{n+1}_{i-1,j})/(2h)\right)^2 + \left((\phi^n_{i,j} - \phi^n_{i+1,j})/h\right)^2}} \\
&= \frac{(\mu/h)}{\sqrt{(h\eta)^2 + \left((\phi^n_{i+1,j} - \phi^{n+1}_{i-1,j})/2\right)^2 + (\phi^n_{i,j} - \phi^n_{i+1,j})^2}}.
\end{aligned}$$

In numerical implementation, we take $(h\eta) = 10^{-8}$.

Numerical experiments



An iterative thresholding scheme

- Most image segmentation models incorporate *the level set formulation* for solving the associated minimization problems. It generally results in initial-boundary value problems for PDEs.
- We are going to employ an *iterative thresholding (IT) scheme* for multi-phase image segmentation based on the Chan-Vese model.
- In the IT scheme, total length of $\mathcal{C}$ is approximated by a non-local multi-phase energy constructed based on *convolution of the heat kernel with the characteristic functions of regions.*
- The IT scheme is divided into two steps. *It works by alternating a thresholding step with an averaging step.*

The approximate Chan-Vese functional

Let $f : \overline{\Omega} \to \mathbb{R}$ be the given grayscale image to be segmented.

- Suppose f approximately takes n distinct constants $c_1, \cdots, c_n$ in the disjoint regions $\Omega_1, \cdots, \Omega_n$ *(n-phase partition)* with boundaries $\mathcal{C}_1, \cdots, \mathcal{C}_n$, respectively, that separate Ω.

  Let $\mathcal{C} = \cup_{i=1}^{n} \mathcal{C}_i$. Then $\Omega \setminus \mathcal{C} = \cup_{i=1}^{n} \Omega_i$.

- Let χ_i be the characteristic function of the desirable region Ω_i,

$$\chi_i(\boldsymbol{x}) = \begin{cases} 1 & \boldsymbol{x} \in \Omega_i, \\ 0 & \text{otherwise}, \end{cases} \qquad \text{and} \qquad \sum_{i=1}^{n} \chi_i = 1 \text{ in } \Omega \setminus \mathcal{C}.$$

- Let $\boldsymbol{\chi} = (\chi_1, \chi_2, \cdots, \chi_n)$. We define the set $\mathcal{S}$ of the characteristic vector functions by

$$\mathcal{S} = \left\{ \boldsymbol{\chi} \in (BV(\Omega))^n : \chi_i(\boldsymbol{x}) \in \{0, 1\}, \sum_{i=1}^{n} \chi_i(\boldsymbol{x}) = 1 \ \forall \boldsymbol{x} \in \Omega \setminus \mathcal{C} \right\},$$

  where $BV(\Omega)$ is the usual bounded variation space.

The approximate Chan-Vese functional (cont'd)

In [WLWW-JCP2017], the authors considered the following model:

$$\min_{\{\Omega_i\},\{c_i\}} \sum_{i=1}^{n} \Big(\lambda |\mathcal{C}_i| + \int_{\Omega_i} (f(\boldsymbol{x}) - c_i)^2 \, d\boldsymbol{x} \Big).$$

Let $\boldsymbol{c} := (c_1, c_2, \cdots, c_n)$. Then we look for χ^* and $\boldsymbol{c}^*$ such that

$$(\chi^*, \boldsymbol{c}^*) = \operatorname*{arg\,min}_{\chi \in \mathcal{S}, \boldsymbol{c} \in \mathbb{R}^n} \sum_{i=1}^{n} \Big(\lambda |\mathcal{C}_i| + \int_{\Omega} \chi_i(\boldsymbol{x}) g_i(\boldsymbol{x}) \, d\boldsymbol{x} \Big),$$

where

$$g_i(\boldsymbol{x}) := (f(\boldsymbol{x}) - c_i)^2.$$

The length of $\mathcal{C}_i$

Let $0 < \tau \ll 1$. Define the heat kernel G_τ by

$$G_\tau(\boldsymbol{x}) := \frac{1}{4\pi\tau} \exp\Big(-\frac{\|\boldsymbol{x}\|_2^2}{4\tau}\Big).$$

Then the length of $\mathcal{C}_i \cap \mathcal{C}_j$ can be approximated by (see CPAM-2015)

$$|\mathcal{C}_i \cap \mathcal{C}_j| \approx \sqrt{\frac{\pi}{\tau}} \int_\Omega \chi_i(\boldsymbol{x}) G_\tau(\boldsymbol{x}) * \chi_j(\boldsymbol{x})\, d\boldsymbol{x},$$

where $*$ represents the convolution operation, and therefore

$$|\mathcal{C}_i| \approx \sum_{j=1, j\neq i}^{n} \sqrt{\frac{\pi}{\tau}} \int_\Omega \chi_i(\boldsymbol{x}) G_\tau(\boldsymbol{x}) * \chi_j(\boldsymbol{x})\, d\boldsymbol{x}.$$

S. Esedoḡlu and F. Otto, Threshold dynamics for networks with arbitrary surface tensions, *Communications on Pure and Applied Mathematics*, 68 (2015), pp. 808-864.

The approximate energy functional and ICT scheme

The total energy functional can be approximated by

$$\mathcal{E}_\tau(\boldsymbol{\chi}, \boldsymbol{c}) = \sum_{i=1}^{n} \Big(\lambda \sum_{j=1, j\neq i}^{n} \sqrt{\frac{\pi}{\tau}} \int_\Omega \chi_i(\boldsymbol{x}) G_\tau(\boldsymbol{x}) * \chi_j(\boldsymbol{x})\, d\boldsymbol{x} + \int_\Omega \chi_i(\boldsymbol{x}) g_i(\boldsymbol{x})\, d\boldsymbol{x} \Big),$$

and our goal is to solve the following minimization problem:

$$(\boldsymbol{\chi}^*, \boldsymbol{c}^*) = \underset{\boldsymbol{\chi}\in\mathcal{S}, \boldsymbol{c}\in\mathbb{R}^n}{\arg\min}\ \mathcal{E}_\tau(\boldsymbol{\chi}, \boldsymbol{c}).$$

The minimization problem can be solved by the ICT scheme, i.e., alternatively updating $\boldsymbol{\chi}$ and $\boldsymbol{c}$. Suppose that we have the k-th iterations for $k \geq 0$, $\boldsymbol{\chi}^{(k)} = (\chi_1^{(k)}, \chi_2^{(k)}, \cdots, \chi_n^{(k)})$ and $\boldsymbol{c}^{(k)}$, then find $\boldsymbol{\chi}^{(k+1)} \in \mathcal{S}$ and $\boldsymbol{c}^{(k+1)} \in \mathbb{R}^n$ sequentially such that

$$\begin{aligned} \boldsymbol{\chi}^{(k+1)} &= \underset{\boldsymbol{\chi}\in\mathcal{S}}{\arg\min}\, \mathcal{E}_\tau(\boldsymbol{\chi}, \boldsymbol{c}^{(k)}), \\ \boldsymbol{c}^{(k+1)} &= \underset{\boldsymbol{c}\in\mathbb{R}^n}{\arg\min}\, \mathcal{E}_\tau(\boldsymbol{\chi}^{(k+1)}, \boldsymbol{c}). \end{aligned}$$

The c-subproblem

Note that the energy functional is given by

$$\mathcal{E}_\tau(\chi, \boldsymbol{c}) = \sum_{i=1}^{n} \Big(\lambda \sum_{j=1, j\neq i}^{n} \sqrt{\frac{\pi}{\tau}} \int_\Omega \chi_i(\boldsymbol{x}) G_\tau(\boldsymbol{x}) * \chi_j(\boldsymbol{x})\, d\boldsymbol{x} + \int_\Omega \chi_i(\boldsymbol{x}) g_i(\boldsymbol{x})\, d\boldsymbol{x} \Big).$$

Then

$$\min_{\boldsymbol{c}\in\mathbb{R}^n} \mathcal{E}_\tau(\boldsymbol{\chi}^{(k+1)}, \boldsymbol{c}) = \min_{\boldsymbol{c}\in\mathbb{R}^n} \int_\Omega \chi_i^{(k+1)}(\boldsymbol{x})(f(\boldsymbol{x}) - c_i)^2\, d\boldsymbol{x}$$

Letting

$$\frac{\partial}{\partial c_i} \int_\Omega \chi_i^{(k+1)}(\boldsymbol{x})(f(\boldsymbol{x}) - c_i)^2\, d\boldsymbol{x} = 0,$$

we have

$$-2 \int_\Omega \chi_i^{(k+1)}(\boldsymbol{x})(f(\boldsymbol{x}) - c_i)\, d\boldsymbol{x} = 0 \Longrightarrow c_i = \frac{\int_\Omega \chi_i^{(k+1)}(\boldsymbol{x}) f(\boldsymbol{x})\, d\boldsymbol{x}}{\int_\Omega \chi_i^{(k+1)}(\boldsymbol{x})\, d\boldsymbol{x}}.$$

The χ-subproblem

Consider the χ-subproblem:

$$\chi^{(k+1)} = \arg\min_{\chi \in \mathcal{S}} \mathcal{E}_\tau(\chi, \boldsymbol{c}^{(k)}).$$

Note that the minimization problem is a non-convex problem since the characteristic function set $\mathcal{S}$ is not a convex set. In order to circumvent this drawback, we define the convex hull $\mathcal{K}$ of $\mathcal{S}$ by

$$\mathcal{K} = \left\{ \chi \in (BV(\Omega))^n : 0 \le \chi_i(\boldsymbol{x}) \le 1, \sum_{i=1}^{n} \chi_i(\boldsymbol{x}) = 1 \ \forall \boldsymbol{x} \in \Omega \setminus \mathcal{C} \right\}.$$

Then we consider the convex relaxed minimization problem instead:

$$\min_{\chi \in \mathcal{K}} \mathcal{E}_\tau(\chi, \boldsymbol{c}^{(k)}).$$

The χ-subproblem (cont'd)

In [WLWW-JCP2017], the authors proved that:

Assume that $\boldsymbol{\chi}^* \in \mathcal{K}$ *is a minimizer of* $\mathcal{E}_\tau(\boldsymbol{\chi}, \boldsymbol{c}^{(k)})$ *on* $\mathcal{K}$*, i.e.,*

$$\mathcal{E}_\tau(\boldsymbol{\chi}^*, \boldsymbol{c}^{(k)}) = \min_{\chi \in \mathcal{K}} \mathcal{E}_\tau(\boldsymbol{\chi}, \boldsymbol{c}^{(k)}).$$

Then $\boldsymbol{\chi}^* \in \mathcal{S}$ *and hence it is also a minimizer of* $\mathcal{E}_\tau(\boldsymbol{\chi}, \boldsymbol{c}^{(k)})$ *on* $\mathcal{S}$*, i.e.,*

$$\mathcal{E}_\tau(\boldsymbol{\chi}^*, \boldsymbol{c}^{(k)}) = \min_{\chi \in \mathcal{S}} \mathcal{E}_\tau(\boldsymbol{\chi}, \boldsymbol{c}^{(k)}).$$

Another approach is to show that $\mathcal{E}_\tau(\boldsymbol{\chi}, \boldsymbol{c}^{(k)})$ is a concave functional on the convex set $\mathcal{K}$. Then minimizers can only be attained at the boundary points of the convex set $\mathcal{K}$, i.e., the subset $\mathcal{S}$.

How to solve the χ-subproblem

Linearizing $\mathcal{E}_\tau(\chi, c^{(k)})$ at $\chi^{(k)}$, we obtain

$$\begin{aligned}\mathcal{E}_\tau(\chi, c^{(k)}) &\approx \mathcal{E}_\tau(\chi^{(k)}, c^{(k)}) + \sum_{i=1}^{n} \int_\Omega \frac{\delta \mathcal{E}_\tau}{\delta \chi_i}\Big|_{\chi=\chi^{(k)}} \left(\chi_i(x) - \chi_i^{(k)}(x)\right) dx \\ &:= \mathcal{E}_\tau(\chi^{(k)}, c^{(k)}) + \sum_{i=1}^{n} \int_\Omega \varphi_i^{(k)}(x) \left(\chi_i(x) - \chi_i^{(k)}(x)\right) dx,\end{aligned}$$

where function $\varphi_i^{(k)}$ is given by

$$0 \le \varphi_i^{(k)}(x) := 2\lambda \sqrt{\frac{\pi}{\tau}} \sum_{j=1, i\neq j}^{n} G_\tau(x) * \chi_j^{(k)}(x) + g_i^{(k)}(x).$$

How to solve the χ-subproblem (cont'd)

Dropping the constant terms in $\mathcal{E}_\tau(\boldsymbol{\chi}, \boldsymbol{c}^{(k)})$, then the $\boldsymbol{\chi}$-subproblem becomes

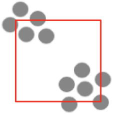
$$\boldsymbol{\chi}^{(k+1)} = \arg\min_{\boldsymbol{\chi} \in \mathcal{K}} \sum_{i=1}^{n} \int_\Omega \varphi_i^{(k)}(\boldsymbol{x}) \chi_i(\boldsymbol{x})\, d\boldsymbol{x}.$$

Because $\varphi_i^{(k)}(\boldsymbol{x}) \geq 0$ and $\chi_i(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \Omega$, the minimizer $\boldsymbol{\chi}^{(k+1)}$ of the above problem can be easily attained at

$$\chi_i^{(k+1)}(\boldsymbol{x}) \quad = \quad \begin{cases} 1, & \text{if } \varphi_i^{(k)}(\boldsymbol{x}) = \min\limits_{1 \leq \ell \leq n} \varphi_\ell^{(k)}(\boldsymbol{x}), \\ 0, & \text{otherwise,} \end{cases}$$

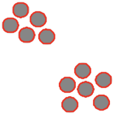
$$\text{for } i = 1, 2, \cdots, n \text{ and } \boldsymbol{x} \in \Omega \setminus \mathcal{C}.$$

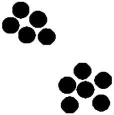
Numerical experiment #1



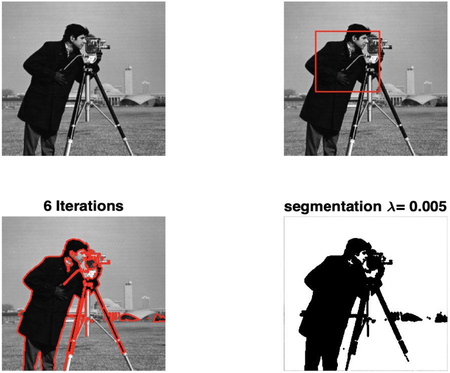
5 Iterations

segmentation λ= 0.002





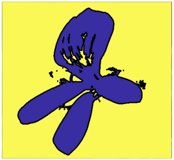
Numerical experiment #2



Numerical experiment #3

17 Iterations

segmentation λ= 0.005



Numerical experiment #4