Chapter 9: Variational Methods: A Short Intro

Konrad Koniarski

2016/07/26

1 Variational Methods

1.1 Variational Methods

Variational methods are aa class of optimization methods. They are popular because they allow to solve many problems in a mathematically transparent manner. Instead of implementing a heuristic sequency of processing step (as was commonly done in the 1980's), one clarifies beforehand what properties an 'optimal' solution should have.

Variational methods are particularly popular for **infinite-dimensional** problems and **spatially continuous** representations.

Particular applications are:

- Image denoising and image restoration
- Image segmentation
- motion estimation and optical flow
- Spatially dense multiple view reconstruction
- Tracking

1.2 Advantages of Variational Methods

Variational methods have **many advantages** over heuristic multi-step approaches (such as the Canny edge detector):

- A mathematical analysis of the considered cost function allows to make statements on the **existence** and **uniqueness** of solutions.
- Approaches with multiple processing steps are difficult to modify. All steps rely one the input from a previous step. Exchanging one module by another typically requires to re-engineer the entire processing pipeline.
- Variational methods make **all modeling assumptions transparent**, there are not hiddne assumptions.
- Variational methods typically have **fewer tuning parameters**. In addition, the effect of respective parameters is clear.
- Variational methods are **easily fused** - one simply adds respective energies / cost functions.

2 Variational Image Smoothing

2.1 Example: Variational Image Smoothing

Let f: $\Omega\mathbb{R}$ be a **grayvalue input image** on the domain $\Omega \subset \mathbb{R}^2$. We assume that the observed image arises by some 'true' image corrupted by additive noise. We are interested in a denoised version u of the input image f.

The approximation u should fulfill two properties:

- It should be as **similar** as possible to f.
- It should be **spatially smooth** (i.e. 'Noise-free').

Both of these criteria can be entered in a **cost function** of the form

$$E(u) = E_{data}(u, f) + E_{smoothness}(u) \tag{1}$$

The first term measures the similarity of f and u. The second one measures the smoothness of the (hypothetical) function u. Most variational approaches have the above form. They merely differ in the specific form of the data (similarity) term and the regularity (or smoothness) term.

2.2 example: Variational Image Smoothing

For denoising a grayvalue image $f : \Omega \subset \mathbb{R}^2 \to R$, specific examples of data and smoothness term are:

$$E_{data}(u, f) = \int_\Omega (u(x) - f(x))^2 dx \tag{2}$$

and

$$E_{smoothness}(u) = \int_\Omega |\nabla u(x)|^2 dx \tag{3}$$

where $\nabla = (\partial/\partial x, \partial/\partial y)^T$ denotes the **spatial gradient**.

Minimizing the weighted sum of data and smoothness term

$$E(u) = \int (u(x) - f(x))^2 dx + \lambda \int |\nabla(x)|^2 dx, \lambda > 0 \tag{4}$$

leads to a **smooth approximation** $u : \Omega \to \mathbb{R}$ of the input image.

Such energies which assign a real value to a function are called a functionals. How does one minimize **functionals** where the argument is a function $u(x)$ (rather then a finite number of parameters) ?

3 Euler-Lagrange Equation

3.1 Functional Minimization and Euler-Lagrange Equation

- As a **necessary condition** for minimizers of a functional the associate **Euler-Lagrange equation** must hold. For a fuctional of the form

$$E(u) = \int \mathcal{L}(u, u')dx \tag{5}$$

it is given by

$$\frac{dE}{du} = \frac{\partial \mathcal{L}}{\partial u} - \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial u'} = 0 \tag{6}$$

- The contral idea of vatiational methods is therefore to determine **solutions of the Euler-lagrange equation** of a given function. **For general non-convex functionals this is a difficult problem**.
- Another solution is to start with an (appropriate) function $u_0(x)$ and to modify it step by step such that in each iteration the value of the functional is decreased. Such methods are called **descent methods**.

4 Gradient Descent

4.1 Gradient Descent

One specific descent methods is called **gradient descent** or **steepest descent**. The key idea is to start from an initialization $u(x, t = 0)$ and iteratively march in direction of the nagative energy gradient. For the class of functionals considered above, the gradient descent is given by the following **partiall differential equations**:

$$\begin{cases} u(x,0) = u_0(x) \\ \frac{\partial u(x,t)}{\partial t} = \frac{dE}{du} = -\frac{d\mathcal{L}}{du} + \frac{\partial}{\partial x}\frac{d\mathcal{L}}{du'} \end{cases} \tag{7}$$

Specifically for $\mathcal{L}(u, u') = \frac{1}{2}(u(x) - f(x))^2 + \frac{\lambda}{2}|u'(s)|^2$ this means:

$$\frac{\partial u}{\partial t} = (f - u) + \lambda u'' \tag{8}$$

If the gradient descent evolution converges: $\partial u/\partial t = -\frac{dE}{du} = 0$, then we have found a solution for the Euler-Lagrange equation.

5 Adaptive Smoothing

5.1 Image Smoothing by Gradient Descent

$$E(u) = \int (f - u)^2 dx + \lambda \int |\nabla u|^2 dx \to min. \tag{9}$$

$$E(u) = \int |\nabla u|^2 dx \to min. \tag{10}$$

5.2 Discontinuity-preserving Smoothing

$$E(u) = \int |\nabla u|^2 dx \to min. \tag{11}$$

$$E(u) = \int |\nabla u| dx \to min. \tag{12}$$