Variational Methods

Computer Vision, Lecture 15
Michael Felsberg
Computer Vision Laboratory
Department of Electrical Engineering

Optimization: Overview

<table>
<thead>
<tr>
<th>Function</th>
<th>Output (codomain / target set)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Set</td>
</tr>
<tr>
<td></td>
<td>Continuous</td>
</tr>
<tr>
<td></td>
<td>Discrete</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input (domain of definition)</th>
<th>Lecture 15</th>
<th>Lecture 13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discrete</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ex: diffusion
ex: level-set segmentation

Repetition: Vector Analysis

- Nabla operator $\nabla = \left[ \frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right]$
- On a scalar function $\nabla f = \text{grad} f = \left[ \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]$
- On a vector field $\langle \nabla | f \rangle = \nabla^T f = \text{div} f = \partial_x f_1 + \partial_y f_2$
- Laplace $\Delta = \nabla^2 = \langle \nabla | \nabla \rangle = \text{div} \text{grad} = \partial_x^2 + \partial_y^2$
- Green’s first identity $x = \left[ \begin{array}{c} x \\ y \end{array} \right]$
- $\int_\Omega (f \Delta g + \langle \nabla f | \nabla g \rangle) \, dx = \int_{\partial \Omega} f \langle \nabla g | n \rangle \, dS = 0$

Revisit: Diffusion

- Lecture on image enhancement:
  $f_x = \frac{\partial}{\partial x} f = \text{div} (\mathbf{D}(\nabla f) \nabla f) = \langle \nabla | \mathbf{D}(\nabla f) \nabla f \rangle$
- Consider scalar diffusivities $\mathbf{D}(\nabla f) \mapsto d(\nabla f)$
- Can diffusion be related to the iterations in an optimization process?
Evolution Equation

- diffusion is an evolution process starting from the original image:
  - initial value problem (IVP)
- discrete steps: gradient descent steps (forward Newton scheme) on a
  - boundary value problem (BVP)
- BVP is obtained by variational calculus from a continuous objective function

Variational Methods

- Minimize the local integral of a Lagrange function
  \[ L(f, f_x, f_y, x, y) \]
  \[ \varepsilon(f) = \int_{\Omega} L(f, \nabla f, x) \, dx \]
- gives the Euler-Lagrange equation on \( \Omega \)
  \[ L_f - \text{div} \, L_{\nabla f} = L_f - \partial_x L_{f_x} - \partial_y L_{f_y} = 0 \quad \forall x, y \]
- if we require \( \langle \nabla f | n \rangle = 0 \) on \( \partial \Omega \)

Insight: EL Equation

- for all test functions \( g \), the Gâteaux derivative
  \[ \langle \delta \varepsilon(f), g \rangle = \left. \frac{d \varepsilon(f + \theta g)}{d \theta} \right|_{\theta=0} = \frac{\delta \varepsilon(f + \theta g) - \delta \varepsilon(f)}{\theta} \]
  must vanish (scalar product in function space)
- Inserting the Lagrangian gives
  \[ \langle \delta \varepsilon(f), g \rangle = \int_{\Omega} \left. \frac{\delta L(f + \theta g, \nabla (f + \theta g), x)}{\theta} \right|_{\theta=0} \, dx \]
  \[ = \langle L_f(f, \nabla f, x), g \rangle + \langle L_{\nabla f}(f, \nabla f, x), \nabla g \rangle \]
- Note
  \[ h'(y) = h(a) + (y - a)^T \nabla h(a) + O((y - a)^2) \]

Insight: EL Equation

- use homogeneity of Green's first identity
  \[ \langle L_{\nabla f}, \nabla g \rangle + \langle \text{div} L_{\nabla f}, g \rangle = \int_{\Omega} \left( \frac{\partial}{\partial n} \right) (L_{\nabla f})_n \, dS = 0 \]
  to rewrite
  \[ \langle L_{\nabla f}, \nabla g \rangle = -\langle \text{div} L_{\nabla f}, g \rangle \]
- thus
  \[ \langle \delta \varepsilon(f), g \rangle = \langle L_f - \text{div} L_{\nabla f}, g \rangle \]
- and we obtain the necessary condition (for all \( x \))
  \[ L_f - \text{div} L_{\nabla f} = 0 \]
### Linear Regularization

- Minimizing $\varepsilon(f) = \frac{1}{2} \int_{\Omega} f^2 + f_x^2 \, dx \, dy$
  
  i.e. no data term $L(f, f_x, f_y, x, y) = L(f, f_x, x, y)$

- Gives the Euler-Lagrange equation
  
  $$(\beta_x f_x + \beta_y f_y) = \Delta f = 0$$

- Such that gradient descent gives $f^{(s+1)} = f^{(s)} + \alpha \Delta f^{(s)}$

- Converges towards trivial solution

### Non-Linear Regularization

- Minimizing $\varepsilon(f) = \int_{\Omega} \psi(|\nabla f|) \, dx \, dy$

  special case: $\psi() = \text{id}() \Rightarrow \psi() = 1$

- Gives the Euler-Lagrange equation

  $$\beta_x \frac{\partial}{\partial x} \left( \frac{\psi'(|\nabla f|)}{|\nabla f|} f_x \right) + \beta_y \frac{\partial}{\partial y} \left( \frac{\psi'(|\nabla f|)}{|\nabla f|} f_y \right) = \text{div} \left( \frac{\psi'(|\nabla f|)}{|\nabla f|} \nabla f \right) = 0$$

- Such that gradient descent gives $f^{(s+1)} = f^{(s)} + \alpha \text{div} \left( \frac{\psi'(|\nabla f^{(s)}|)}{|\nabla f^{(s)}|} \nabla f^{(s)} \right)$

### Exemple: Perona-Malik Flow

- Special cases: $\psi(|\nabla f|) = -K^2/2 \cdot \exp(-|\nabla f|^2/K^2)$

  $$\Rightarrow \psi'(|\nabla f|) = \frac{|\nabla f|}{|\nabla f|} \exp(-|\nabla f|^2/K^2)$$

  $$\psi(|\nabla f|) = K^2/2 \cdot \log(K^2 + |\nabla f|^2)$$

  $$\Rightarrow \psi'(|\nabla f|) = |\nabla f| (1 + |\nabla f|^2/K^2)^{-1}$$

- Such that gradient descent gives Perona-Malik Flow

  $$f^{(s+1)} = f^{(s)} + \alpha \text{div} \left( \frac{\psi'(|\nabla f^{(s)}|)}{|\nabla f^{(s)}|} \nabla f^{(s)} \right)$$

### Interpretation

- Diffusion is an evolution over "time" s

  $$f^{(s+1)} = f^{(s)} + \alpha \text{div} \left( \frac{\psi'(|\nabla f^{(s)}|)}{|\nabla f^{(s)}|} \nabla f^{(s)} \right)$$

- Starts at the measured image (IVP)

- Converges towards DC signal

- Critical parameter 1: "stopping time"

- Critical parameter 2: $\alpha$

- Several examples in the enhancement lecture
Beyond Diffusion

- In what follows: add data term to minimization problem
- Converges towards non-trivial solution
- IVP with standard forward Euler scheme

\[ e(f) = \frac{1}{2} \int \left( f - f_0 \right)^2 + \lambda \| \nabla f \|_1 \, dx \, dy \]

\[ f - f_0 - \lambda \Delta f = 0 \]

\[ f^{(n+1)} = f^{(n)} - \alpha \left( f^{(n)} - f_0 - \lambda \text{div} \left( \frac{\| \nabla f^{(n)} \|_1}{\| \nabla f^{(n)} \|_1} \nabla f^{(n)} \right) \right) \]

\[ = (1 - \alpha) f^{(n)} + \alpha (f_0 + \lambda \text{div} \left( f^{(n)} \right)) \]

Linear Restoration

- Minimizing

\[ e(f) = \frac{1}{2} \int \left( f - f_0 \right)^2 + \lambda \| \nabla f \|_2^2 \, dx \, dy \]

\[ L(f, f_0, x, y) \]

- Gives the Euler-Lagrange equation

\[ f - f_0 - \lambda \Delta f = 0 \]

\[ L_f \text{ div}(L_{f_0}, L_{f_n}) \]

- Such that gradient descent gives

\[ f^{(n+1)} = f^{(n)} - \alpha \left( f^{(n)} - f_0 - \lambda \Delta f^{(n)} \right) \]

\[ = (1 - \alpha) f^{(n)} + \alpha f_0 + \lambda \Delta f^{(n)} \]

Non-Linear Restoration

- Minimizing

\[ e(f) = \int \frac{1}{2} (f - f_0)^2 + \lambda \| \nabla f \|_2^2 \, dx \, dy \]

- Gives the Euler-Lagrange equation

\[ f - f_0 - \lambda \text{div} \left( \frac{\| \nabla f \|_2}{\| \nabla f \|_2} \nabla f \right) = 0 \]

- Such that gradient descent gives

\[ f^{(n+1)} = f^{(n)} - \alpha \left( f^{(n)} - f_0 - \lambda \text{div} \left( \frac{1}{\| \nabla f \|_2} \nabla f^{(n)} \right) \right) \]

\[ = (1 - \alpha) f^{(n)} + \alpha (f_0 + \lambda \text{div} \left( f^{(n)} \right)) \]

Special Case: TV/ROF

- Minimizing

\[ e(f) = \int \left( f - f_0 \right)^2 + \lambda \| \nabla f \|_1 \, dx \, dy \]

- Gives the Euler-Lagrange equation

\[ f - f_0 - \lambda \text{div} \left( \frac{1}{\| \nabla f \|_2} \nabla f \right) = 0 \]

- Such that gradient descent gives

\[ f^{(n+1)} = f^{(n)} - \alpha \left( f^{(n)} - f_0 - \lambda \text{div} \left( \frac{1}{\| \nabla f \|_2} \nabla f^{(n)} \right) \right) \]

\[ = (1 - \alpha) f^{(n)} + \alpha (f_0 + \lambda \text{div} \left( f^{(n)} \right)) \]
Example (lecture 13)

- Parameters: $\alpha=0.0005, \lambda=0.5, \text{noise}(0,0.001)$

Explicit vs Implicit

- All gradients so far are based on the previous estimate; the time discretization leads to an explicit scheme (least calculations, easiest)
- If the gradients are based on the new estimate, we obtain an implicit scheme (always stable, large time steps)
- If the gradients are based on both, we obtain the Crank-Nicolson scheme (always stable, small time steps)

Interpretation

- Restauration is an IVP
- Uses the measured image as input in each iteration
- Converges towards non-trivial solution
- Critical parameter 1: "meta" parameter $\lambda$
- Critical parameter 2: $\alpha$

The Data Term

- Data term can be used to describe the measurement model
- Leads to non-trivial iterations
Deblurring

• Minimizing
  \[ \varepsilon(f) = \frac{1}{2} \int_{\Omega} (g * f - f_0)^2 + \lambda (f_x^2 + f_y^2) \, d\Omega \]

• Gives the Euler-Lagrange equation
  \[ g(-x) * (g * f - f_0) - \lambda \Delta f = 0 \]

• Such that gradient descent gives
  \[ f^{(x+1)} = f^{(x)} - \alpha (g(-x) * (g * f^{(x)} - f_0) - \lambda \Delta f^{(x)}) \]

Comments

• \( g \): point spread function (PSF)
• \( g(-x) \): correlation operator / adjoint operator
• even symmetry PSF: self adjoint
• definition of adjoint operator \( \langle x | A| y \rangle = \langle A^* x | y \rangle \)

• Example from lecture 13

Indirect Measurements

• Similar to target tracking, where observations might be different from states
• We observe image information but apply the variational framework to estimate other fields
• Two examples here: optical flow and segmentation (binary partition)
Optical Flow $ f = [f_1, f_2]^T $

- Minimizing $ \mathcal{E}(f) = \frac{1}{2} \int_{\Omega} ((f_1 \nabla g)^2 + (f_2 \nabla g)^2) + \lambda (|f_1|^2 + |f_2|^2) \, dx \, dy $ gives the Euler-Lagrange equation (HS!)
- Laplacian is approximately $ \Delta f = 0 $

\[
\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}
\]

Optical Flow

- Plugging into the EL-equation gives

\[
(\lambda + \nabla g \nabla g^T) f = \lambda \ddot{f} - \dot{g} \nabla g
\]

- Explicitly solving for $ f $ results in

\[
(\lambda + \nabla g \nabla g^T) f = (\lambda + \nabla g \nabla g^T) \ddot{f} - (\nabla g \nabla g^T \dot{f} + \nabla g \dot{g})
\]

\[
= (\lambda + \nabla g \nabla g^T) \ddot{f} - \frac{\lambda + \nabla g \nabla g^T}{\lambda + \nabla g \nabla g} \nabla g (\nabla g \nabla g^T + \dot{g})
\]

\[
\begin{aligned}
\ddot{f} &= \frac{\ddot{f}}{} - \frac{1}{\lambda + \nabla g \nabla g} \nabla g (\nabla g \nabla g^T + \dot{g}) \\
\dot{f} &= \ddot{f} - \frac{1}{\lambda + \nabla g \nabla g} \nabla g (\nabla g \nabla g^T + \dot{g})
\end{aligned}
\]

Optical Flow

- Iterating the solution

\[
\ddot{f} = \ddot{f} - \frac{1}{\lambda + \nabla g \nabla g} \nabla g (\nabla g \nabla g^T + \dot{g})
\]

- Results in the Horn & Schunck iteration

\[
\dot{f}^{t+1} = \ddot{f}^{t} - \frac{1}{\lambda + |\nabla g|^2} (\nabla g \nabla g^T + \dot{g}) \nabla g
\]

- Significant improvement: use median instead of $ \ddot{f} $ !

Demonstration
Segmentation / Contours

• Segmentation function (level-set function) to be optimized
• Negative / positive in background / object region
• Contour is the zero-level

\[ E(\phi) = \int_{\Omega} (f_2 - f_1)\phi + \lambda |\nabla \phi| \, dx \]

- \( \phi \) is the (regularized) Heaviside function
- \( f \) are weights computed from the image (e.g. squared deviation from certain greyscale)
- EL equation
  \[ \delta(\phi) \left( f_2 - f_1 + \lambda \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) \right) = 0 \]

- Problem: (regularized) delta function \( \delta \)

Segmentation / Contours

• Omitting delta-function
• Original solution remains solution
• Corresponds to minimizing
\[ E(\phi) = \int_{\Omega} (f_2 - f_1)\phi + \lambda |\nabla \phi| \, dx \]
• Non-existence of minimizer (!)

Segmentation / Contours

• Chan-Vese energy minimized of level-set function \( \phi \)
  \[ E(\phi) = \int_{\Omega} (H(\phi) - 1) f_2 - H(\phi) f_1 + \lambda |\nabla H(\phi)| \, dx \]
• \( H \) is the (regularized) Heaviside function
• \( f \) are weights computed from the image (e.g. squared deviation from certain greyscale)

Segmentation / Contours

• Binary function instead of level-set function
• becomes Ising model (lecture 13)
\[ E(\phi) = -\int_{\Omega_2} f_2 \, dx - \int_{\Omega_1} f_1 \, dx + \lambda |C| \]
• Hard to solve – use relaxation
  – Binary function replaced by smooth approximation
  – After optimization apply threshold
• Discrete optimization (lecture 13)
Examples

Over-Segmentation / Superpixels
- So far: attempt for semantic segmentation
- Alternative: over-segmentation based on stationarity of image process
  - MSER (lecture 8)
  - Superpixel algorithms – clustering in 5D \((x, y, R, G, B)\)
  - Left: contour-relaxed superpixels
  - Right: SLIC

Alternative Contour Methods
- Popular application:
  - Geodesic active contours
  - Snakes
- Contour parametrized as
  \[ \mathbf{v}(s) = [x(s), y(s)] \quad s \in [0, 1] \]
- Usually approximated as spline
- Option: Fourier descriptors
Geodesic Active Contours

- Consider a curve moving in time
  \[ \mathbf{v}(s, t) = \{x(s, t), y(s, t)\} \]
- Let the curve develop according to the inward normal \( \mathbf{n} \) and the curvature \( c \)
  \[ \frac{\partial \mathbf{v}}{\partial t} = V(c)\mathbf{n} \]

Geodesic Active Contours

- Assume level set function \( \phi(x, y, t) \) such that \( \phi(\mathbf{v}(s, t), t) = 0 \)
- Negative inside and positive outside gives
  \[ \mathbf{n} = -\frac{\nabla \phi}{|\nabla \phi|} \]
- Plug in normal into evolution equation gives
  \[ \frac{\partial \mathbf{v}}{\partial t} = -\frac{V(c)\nabla \phi}{|\nabla \phi|} \]

Geodesic Active Contours

- What remains is to re-write l.h.s. of
  \[ \frac{\partial \mathbf{v}}{\partial t} = -\frac{V(c)\nabla \phi}{|\nabla \phi|} \]
- Time derivative of \( \phi(\mathbf{v}(s, t), \xi) \) gives
  \[ \frac{\partial \phi}{\partial t} + \nabla \phi \cdot \frac{\partial \mathbf{v}}{\partial t} = 0 \]
- Such that
  \[ \frac{\partial \phi}{\partial t} = V(c)|\nabla \phi| \]
- Level-set equation

Snake Function

- Energy function consists of typically 3 terms:
  - internal energy
  - image forces
  - external constraint forces
  \[ e(\mathbf{v}(s)) = \int_0^1 E_{int}(\mathbf{v}(s)) + E_{image}(\mathbf{v}(s)) + E_{con}(\mathbf{v}(s)) \, ds \]
Limitations

- Initialization close to solution
- Problems at concave regions

GVF Snakes

- Gradient vector flow snakes
- GVF used as external force
- GVF field computation related to optical flow approach

GVF Field

- Minimizing (GVF: f)
  \[ \phi(f') = \frac{1}{2} \int_{\Omega} |f' - \nabla g|^2 |\nabla g|^2 + \lambda (|\nabla f_1|^2 + |\nabla f_2|^2) \, dx \, dy \]
- Gives the Euler-Lagrange equations
  \[ (f' - \nabla g)|\nabla g|^2 - \lambda \Delta f = 0 \]
- Such that gradient descent gives
  \[ f^{(n+1)} = f^{(n)} - \alpha \left( (f^{(n)} - \nabla g)|\nabla g|^2 - \lambda \Delta f^{(n)} \right) \]

Examples

- No concavity problem
- No initialization problem