Variational Optical Flow Estimation

ICCV 2009 Tutorial, Kyoto, Japan

Andrés Bruhn
Mathematical Image Analysis Group
Saarland University
Saarbrücken, Germany
bruhn@mia.uni-saarland.de

Thomas Brox
Computer Vision Group
U.C. Berkeley
Berkeley, US
brox@eecs.berkeley.edu

What is the Optical Flow Problem?

- Given
  - two or more frames of an image sequence

- Wanted
  - displacement field between two consecutive frames $\rightarrow$ optical flow

What is Optical Flow Good for?

- Extraction of Motion Information
  - robot navigation/driver assistance
  - surveillance/tracking
  - action recognition

- Processing of Image Sequences
  - video compression
  - ego motion compensation

- Related Correspondence Problems
  - stereo reconstruction
  - structure-from-motion
  - medical image registration

Why Variational Methods?

- Advantages w.r.t. Modeling
  - transparent modeling
  - formulation as optimization problem

- Advantages w.r.t. Computation
  - unique minimizer and well-posedness
  - real-time capable numerical schemes

- Advantages w.r.t. Quality
  - dense flow fields with sub-pixel precision
  - most accurate results in the literature

These are the reasons why variational methods are so successful!
Outline of this Tutorial

- **Part I: Variational Basics (Andrés Bruhn)**
  - Continuous modeling
  - Method of Horn and Schunck

- **Part II: Modeling Aspects (Thomas Brox)**
  - Motion discontinuities
  - Robust data terms
  - Large displacements

- **Part III: Efficient Numerics (Andrés Bruhn)**
  - Improved non-hierarchical solvers
  - Linear and nonlinear multigrid
  - Implementations on parallel hardware

### Continuous Modeling

#### Given
- continuous image sequence \( I_0(x, y, t) \) location \( (x, y) \in \Omega \) time \( t \in [0, T] \)

#### Wanted
- interframe displacement field \( w(x, y, t) = \begin{pmatrix} u(x, y, t) \\ v(x, y, t) \end{pmatrix} \rightarrow \text{optical flow} \)

### Standard Preprocessing

- **Idea:** In order to reduce the influence of noise and outliers, we convolve \( I_0 \) with a Gaussian \( K_\sigma \) of mean \( \mu = 0 \) and standard deviation \( \sigma \)

\[
I(x, y, t) = K_\sigma * I_0(x, y, t)
\]

- image sequence becomes infinitely many times differentiable, i.e. \( I \in C^\infty \)
- allows to estimate larger displacements due to the blurring of objects

Important for methods that rely on the computation of image derivatives!
The Linearized Gray Value Constancy Assumption

- **Idea:** In order to retrieve corresponding pixels in subsequent frames, we assume that their gray value does not change over time:

\[
I(x + u, y + v, t + 1) - I(x, y, t) = 0.
\]

The Gray Value Constancy Assumption

- **Idea:** If \( u \) and \( v \) are small and \( I \) is sufficiently smooth, one may linearize this constancy assumption via a first-order Taylor expansion around the point \((x, y, t)\):

\[
I(x + u, y + v, t + 1) \approx I(x, y, t) + I_x(x, y, t)u + I_y(x, y, t)v + I_t(x, y, t)\Delta t
\]

\[
\rightarrow I_xu + I_yv + I_t = 0.
\]

This constraint is the **brightness constancy constraint equation (BCCE)**. In general such constraints on the flow are called optical flow constraints (OFCs).

Continuous Modeling (3)

- The BCCE provides only one equation for determining two unknowns
- Ill-posed problem with infinitely many solutions
- Only the flow component in direction of the image gradient can be computed,
  the so-called **normal flow**:

\[
(u, v)^T = -\nabla I = -\frac{I_x}{\sqrt{I_x^2 + I_y^2}}.
\]

- This problem is referred to as the **aperture problem**. It can be illustrated as

![Color Plot](image1.png)

\[
\text{Case I: } |\nabla I| \neq 0 \Rightarrow \text{Aperture problem} \quad \text{Case II: } |\nabla I| = 0 \Rightarrow \text{No estimation possible}
\]

Continuous Modeling (6)

- **Given:** estimated flow field \( w^e \) and ground truth flow field \( w^i \)

- **Spatiotemporal Average Angular Error (AAE):**
  - Consider angle and magnitude by using the **spatiotemporal angle**
  
  \[
  AAE = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \arccos \left( \frac{w_{i,j}^e \cdot w_{i,j}^i}{\|w_{i,j}^e\| \|w_{i,j}^i\|} \right).
  \]

- **Average Endpoint Error (AEE):**
  - Consider the **Euclidean distance** between the vectors

\[
AEE = \frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} \|w_{i,j}^e - w_{i,j}^i\|.
\]
Variational Optical Flow Computation

What is a Functional?

- **Known**: A function maps an input value to an output value, e.g.
  \[ f(x, y) = x^2 + y^2 \, . \]

- **New**: A functional maps an input function to an output value, e.g.
  \[ E(f(x, y)) = \frac{1}{|\Omega|} \int f(x, y) \, dx \, dy \, . \]

**Remarks**: Functionals

- can be used to rate the quality of a function w.r.t. certain assumptions
- form the basis of variational optical flow methods

Principle of Variational Optical Flow Methods

- **Idea**: Compute displacement field as minimizer of a suitable energy functional:
  \[ E(u, v) = \int_{\Omega} D(u, v) + \alpha \, S(u, v) \, \, dx \, dy \, . \]
  - **data term** \( D(u, v) \) penalizes deviations from constancy assumptions
  - **smoothness term** \( S(u, v) \) penalizes dev. from smoothness of the solution
  - regularization parameter \( \alpha > 0 \) determines the degree of smoothness

**Remarks**: The minimizing functions \( u \) and \( v \)

- fit best to all model assumptions (smallest value for the energy functional)
- can be seen as a compromise between all (partly contradictive) assumptions

The Method of Horn and Schunck

- **Idea**: Assume overall smoothness of the resulting flow field

- The method of Horn and Schunck computes the optical flow as minimizer of (Horn/Schunck AI 1981)
  \[ E(w) = \int_{\Omega} \left( I_x u + I_y v + I_t \right)^2 + \alpha \, (|\nabla u|^2 + |\nabla v|^2) \, \, dx \, dy \, . \]
  - **data term** penalizes deviations from the linearized brightness constancy assumption (BCCE)
  - **smoothness term** penalizes deviations from smoothness of the flow field, i.e. from variations of the functions \( u \) and \( v \) given by their first derivatives

Why variational methods can compute a solution everywhere?
The Filling-In-Effect

- **Observation**: If no information is available, i.e. $|\nabla f| \approx 0$, the flow functions $u$ and $v$ have hardly any influence on the contribution of the data term

$$ \left( f_u u + f_v v + f_t \right)^2 \approx f_t^2 . $$

- **Consequence**: The flow functions $u$ and $v$ adapt to the local solution(s) of the neighborhood to fulfill at least the smoothness term $\rightarrow \text{filling-in-effect}$. 

edge information

filling-in

The Method of Horn and Schunck (2)

Minimization and Discretization (1)

**Minimization of Continuous Energy Functionals**

- **Idea**: Similar strategy as for ordinary functions $\rightarrow$ derive necessary conditions
- These necessary conditions are called **Euler-Lagrange equations**. They state that the first variation of the energy functional must vanish ($\approx$ first derivative).
  - E.g. (Elsgolc 1961, Gelfand/Fomin 2000)
- For a typical optical flow energy functional of type

$$ E(u, v) = \int_{\Omega} F(x, y, u, v, u_x, u_y, v_x, v_y) \, dx \, dy $$

the Euler-Lagrange equations are given by the following system of PDEs

$$ 0 = \frac{\partial}{\partial x} F_{ux} - \frac{\partial}{\partial y} F_{uy} , $$

$$ 0 = \frac{\partial}{\partial x} F_{vx} - \frac{\partial}{\partial y} F_{vy} , $$

with the associated **boundary conditions** $\mathbf{n}^T \begin{pmatrix} F_{ux} \\ F_{vy} \end{pmatrix} = 0$ and $\mathbf{n}^T \begin{pmatrix} F_{vx} \\ F_{vy} \end{pmatrix} = 0$.

The Method of Horn and Schunck (3)

**The Motion Tensor Notation**

- **Idea**: Rewrite a linearized quadratic data term in a more compact way
- **Example**: Linearized gray value constancy assumption (BCCE)

$$ (I_u + I_v + I_t)^2 = (w^T \nabla_3 I)^2 = w^T \nabla_3 I \nabla_3 I^T w = w^T J w $$

yields a single quadratic form with the $3 \times 3$ **motion tensor**

$$ J = \begin{pmatrix} J_{11} & J_{12} & J_{13} \\ J_{12} & J_{22} & J_{23} \\ J_{13} & J_{23} & J_{33} \end{pmatrix} = \begin{pmatrix} I_{x}^2 & I_{x} I_{y} & I_{x} I_{r} \\ I_{y} I_{x} & I_{y}^2 & I_{y} I_{r} \\ I_{r} I_{x} & I_{r} I_{y} & I_{r}^2 \end{pmatrix} = \nabla_3 I \nabla_3 I^T . $$

- **Application**: In motion tensor notation the Horn and Schunck method reads

$$ E(w) = \int_{\Omega} \text{data term} + \alpha \left( |\nabla u|^2 + |\nabla v|^2 \right) \, dx \, dy . $$

Minimization and Discretization (2)

**How Do These Equations Look Like for the Method of Horn and Schunck?**

- For the Method of Horn and Schunck $F(x, y, u, v, u_x, u_y, v_x, v_y)$ is given by

$$ F = w^T J w + \alpha \left( |\nabla u|^2 + |\nabla v|^2 \right) $$

$$ = J_{11} u^2 + J_{22} v^2 + J_{33} + 2 J_{12} u v + 2 J_{13} u + 2 J_{23} v + \alpha \left( u_x^2 + u_y^2 + v_x^2 + v_y^2 \right) . $$

- The required partial derivatives can then be computed as

$$ F_u = 2 J_{11} u + 2 J_{12} v + 2 J_{13} , \quad F_u = \alpha 2 u_x , \quad F_u = \alpha 2 u_y , $$

$$ F_v = 2 J_{13} u + 2 J_{22} v + 2 J_{23} , \quad F_v = \alpha 2 v_x , \quad F_v = \alpha 2 v_y . $$

- As necessary condition for a minimizer this yields the Euler–Lagrange equations

$$ 0 = F_u - \frac{\partial}{\partial x} F_{ux} - \frac{\partial}{\partial y} F_{uy} = \partial \left( J_{11} u + J_{12} v + J_{13} - \alpha \left( u_x^2 + u_y^2 \right) \right) $$

$$ 0 = F_v - \frac{\partial}{\partial x} F_{vx} - \frac{\partial}{\partial y} F_{vy} = \partial \left( J_{12} u + J_{22} v + J_{23} - \alpha \left( v_x^2 + v_y^2 \right) \right) \Delta v $$

with (reflecting) **Neumann boundary conditions** $n^T \nabla u = 0$ and $n^T \nabla v = 0$. 

The Method of Horn and Schunck (4)
Minimization and Discretization (5)

Existence and Uniqueness of the Minimizer

- Strictly convex energy functionals
  - fulfill for all \( \alpha \in [0, \ldots, 1] \) the inequality:
    \[
    E(\alpha u_1 + (1-\alpha)u_2) < \alpha E(u_1) + (1-\alpha)E(u_2) \,.
    \]
  - have at most one solution which is **unique**
    if it exists (global minimizer)

- Further properties of strictly convex variational optical flow methods
  (Schnörr JMV 1994, Weickert/Schnörr IJCV 2001)
  - existence of a solution
  - solution depends continuously on the input data

Well-posedness (in the sense of Hadamard)

Minimization and Discretization (6)

Structure of the Linear System

- This linear system of equations \( Ax = b \) has the following block structure

\[
\begin{pmatrix}
J_{11} & J_{12} & J_{13} & \cdots & J_{1N} \\
J_{21} & J_{22} & J_{23} & \cdots & J_{2N} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
J_{N1} & J_{N2} & J_{N3} & \cdots & J_{NN}
\end{pmatrix}
- \begin{pmatrix}
\alpha & 0 & 0 & \cdots & 0 \\
0 & \alpha & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \alpha
\end{pmatrix}
- \begin{pmatrix}
\frac{u_{i+1,j} - u_{i,j}}{h_x^2} + \frac{u_{i,j+1} - u_{i,j}}{h_y^2} & -\frac{1}{h_x^2} & -\frac{1}{h_x^2} & \cdots & -\frac{1}{h_x^2} \\
-\frac{1}{h_x^2} & \frac{u_{i+1,j} - u_{i,j}}{h_x^2} + \frac{u_{i,j+1} - u_{i,j}}{h_y^2} & -\frac{1}{h_x^2} & \cdots & -\frac{1}{h_x^2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
-\frac{1}{h_x^2} & -\frac{1}{h_x^2} & -\frac{1}{h_x^2} & \cdots & \frac{u_{i+1,j} - u_{i,j}}{h_x^2} + \frac{u_{i,j+1} - u_{i,j}}{h_y^2}
\end{pmatrix}
- \begin{pmatrix}
x_1 \\
x_2 \\
\vdots \\
x_N
\end{pmatrix}
- \begin{pmatrix}
b_1 \\
b_2 \\
\vdots \\
b_N
\end{pmatrix}
\]

- **smoothness term** only contributes to block main diagonals
- **data term** also contributes to block off-diagonals

- For non-constant input images the matrix \( A \) is positive definite

- For an image with 1M pixels, the matrix \( A \) has \( 4 \cdot 10^{12} \) entries. Assuming 32-bit float precision this requires 16 Terabyte memory \( \rightarrow \) store only non-zero entries. Direct Gauss-Elimination with complexity \( O(n^3) \) is not practicable.
Solving Linear Systems of Equations

How Can We Solve The Linear System of Equation $Ax = b$?

- **Idea:** Find a cheap but accurate approximation of $A^{-1}$ via the decomposition (e.g. Young 1971, Saad 1996)
  
  $$A = A_1 + A_2$$

- Introduce fixed point iteration of type
  
  $$A_1 x^{k+1} = b - A_2 x^k$$

  $$\Leftrightarrow x^{k+1} = A_1^{-1}(b - A_2 x^k)$$

- In each iteration a linear system of equations with matrix $A_1$ has to be solved
  - $A_1^{-1}$ should be a reasonable approximation of $A^{-1}$
  - $A_1^{-1}$ should be cheap to compute, i.e. the system should be simple to solve

Frequent Approach

- Use matrix decomposition of type
  
  $$A = D - L - U$$

  - $D$ is the diagonal part of $A$
  - $L$ is the strictly lower triangular part of $A$
  - $U$ is the strictly upper triangular part of $A$

- For the method of Horn and Schunck this yields
  
  $$A = \begin{pmatrix}
  \frac{1}{h_x^2} & \frac{-1}{h_x} & \frac{1}{h_x} & 0 & \cdots & 0 & 0 \\
  \frac{-1}{h_x} & 1 & \frac{-1}{h_x} & \frac{1}{h_x^2} & \cdots & 0 & 0 \\
  \frac{1}{h_x} & \frac{-1}{h_x} & 1 & \frac{-1}{h_x} & \cdots & 0 & 0 \\
  0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
  \end{pmatrix} - \alpha \begin{pmatrix}
  \frac{1}{h_y^2} & \frac{-1}{h_y} & \frac{1}{h_y} & 0 & \cdots & 0 & 0 \\
  \frac{-1}{h_y} & 1 & \frac{-1}{h_y} & \frac{1}{h_y^2} & \cdots & 0 & 0 \\
  \frac{1}{h_y} & \frac{-1}{h_y} & 1 & \frac{-1}{h_y} & \cdots & 0 & 0 \\
  0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
  \end{pmatrix}$$

Iterative Solvers (3)

- In terms of the discretized Euler-Lagrange equations we obtain
  
  $$0 = \left| J_{11} \right|_{i,j} u_{i,j} + \left| J_{12} \right|_{i,j} v_{i,j} + \left| J_{21} \right|_{i,j}$$

  $$+ \alpha \sum_{l \in \mathbb{R}^X \cap \mathbb{N}^Y} \sum_{(i,j) \in \mathbb{N}^X \cap \mathbb{N}^Y} \frac{1}{h_x^2} u_{i,j} - \alpha \sum_{l \in \mathbb{R}^X \cap \mathbb{N}^Y} \sum_{(i,j) \in \mathbb{N}^X \cap \mathbb{N}^Y} \frac{1}{h_y^2} w_{i,j}$$

  $$0 = \left| J_{22} \right|_{i,j} u_{i,j} + \left| J_{22} \right|_{i,j} v_{i,j} + \left| J_{22} \right|_{i,j}$$

  $$+ \alpha \sum_{l \in \mathbb{R}^X \cap \mathbb{N}^Y} \sum_{(i,j) \in \mathbb{N}^X \cap \mathbb{N}^Y} \frac{1}{h_x^2} v_{i,j} - \alpha \sum_{l \in \mathbb{R}^X \cap \mathbb{N}^Y} \sum_{(i,j) \in \mathbb{N}^X \cap \mathbb{N}^Y} \frac{1}{h_y^2} w_{i,j}$$

  for $i = 1, \ldots, N$ and $j = 1, \ldots, M$.

- Notation for the neighborhood
  - $\mathcal{N}^-_l(i,j)$ denotes the set of neighbors of pixel $i,j$ in direction of axis $l$ that have a **smaller index** (will be updated before the central pixel)
  - $\mathcal{N}^+_l(i,j)$ denotes the set of neighbors of pixel $i,j$ in direction of axis $l$ that have a **larger index** (will be updated after the central pixel)

The Jacobi Method

- Set $A_1 = D$, since diagonal matrices are simple to invert. $A_2 = -L - U$.

- Yields the fixed point iteration
  
  $$x^{k+1} = D^{-1}( b + (L + U) x^k ) \Leftrightarrow x^{k+1} = \frac{1}{\alpha} \left( b - \sum_{j < i} a_{ij} x^k - \sum_{j > i} a_{ij} x^k \right)$$

- For the Horn and Schunck method the Jacobi iteration for the pixel $i,j$ reads
  
  $$u_{i,j}^{k+1} = \left( -\left| J_{11} \right|_{i,j} - \left( \left| J_{12} \right|_{i,j} v_{i,j} - \alpha \sum_{l \in \mathbb{R}^X \cap \mathbb{N}^Y} \sum_{(i,j) \in \mathbb{N}^X \cap \mathbb{N}^Y} \frac{1}{h_x} u_{i,j} - \alpha \sum_{l \in \mathbb{R}^X \cap \mathbb{N}^Y} \sum_{(i,j) \in \mathbb{N}^X \cap \mathbb{N}^Y} \frac{1}{h_y} w_{i,j} \right) \right)$$

  $$v_{i,j}^{k+1} = \left( -\left| J_{22} \right|_{i,j} - \left( \left| J_{22} \right|_{i,j} u_{i,j} - \alpha \sum_{l \in \mathbb{R}^X \cap \mathbb{N}^Y} \sum_{(i,j) \in \mathbb{N}^X \cap \mathbb{N}^Y} \frac{1}{h_x} v_{i,j} - \alpha \sum_{l \in \mathbb{R}^X \cap \mathbb{N}^Y} \sum_{(i,j) \in \mathbb{N}^X \cap \mathbb{N}^Y} \frac{1}{h_y} w_{i,j} \right) \right)$$
The Gauß-Seidel Method

- Set $A_1 = D - L$, since this triangular matrix is a better approximation to $A$ than the diagonal $D$ alone. Triangular matrices are still simple to invert. $A_2 = -U$.
- Yields the fixed point iteration

$$x^{k+1} = (D-L)^{-1}(b + U^T x^k)$$

\(\iff\)

$$x_i^{k+1} = \frac{1}{a_{ii}} \left( b_i - \sum_{j<i} a_{ij} x_j^k - \sum_{j>i} a_{ij} x_j^k \right).$$

- The corresponding Gauß-Seidel iteration for the pixel $i, j$ reads

$$u_{i,j}^{k+1} = \frac{-[d_{1,i,j} - \left( [d_{1,j}] - \alpha \sum_{x,y} \sum_{i,j} v_{i,j}^{k} \right) \left( [d_{1,j}] + \alpha \sum_{x,y} \sum_{i,j} v_{i,j}^{k} \right)]}{[d_{2,i,j} + \alpha \sum_{x,y} \sum_{i,j} v_{i,j}^{k}}.$$
Summary

- Variational methods compute optical flow as minimizer of an energy functional
- They make use of global smoothness assumptions on the solution to overcome the aperture problem (filling-in-effect by the smoothness term → dense results)
- They are minimized by solving their (discretized) Euler-Lagrange equations
- They offer many advantages such as
  - transparent modeling
  - dense flow fields
  - well-posedness
  - sub-pixel precision
- The method of Horn and Schunck is the simplest variational approach
- There are many adaptations/modifications of this basic method possible that improve the quality and the performance even further (see PART II-III)