Aggregation Methods for Optical Flow Computation

Denis Fortun, Patrick Bouthemy, Charles Kervrann

SERPICO Team
Inria Rennes - Bretagne Atlantique
Campus universitaire de Beaulieu
35 042 Rennes Cedex, France
**Framework**: patch-wise approach and aggregation method for large and small displacement optical flow
Outline

1. Local and global methods for computing optical flow
   - data conservation
   - local smoothing
   - global approach and spatial regularization

2. Semi-local estimation and global aggregation
   - semi-local estimation of flow fields
   - global aggregation

3. Experimental results and application
   - comparisons of aggregation procedures
   - comparison of parametric and variational methods
   - algorithm parameters

Discussion and conclusion
1. Local and global methods for computing optical flow
Data conservation constraint

Fundamental assumption to find correspondences: conservation of brightness or image gradient over time

\[ l_2(x + w(x)) = l_1(x), \quad x \in \Omega \]

Insufficient constraint:

- "Aperture problem": 1 equation for 2 unknowns
  \[ w(x) = (u(x), v(x))^T \]
- Uniform regions: no image gradient
- Assumption violations: occlusion, brightness changes, . . .

→ Adding a spatial constraint to make the problem well-posed!
Local smoothing and spatial neighborhoods

Assume a coherent motion in a neighborhood \( V(x_0) \subset \Omega \)

**Coherent motion**: parametric motion model \( w_\theta(x) \)

- **translation**: \( w_\theta(x) = (\theta_1, \theta_4)^T \)
- **Affine motion**: \( w_\theta(x) = \begin{pmatrix} \theta_1 & \theta_2 & \theta_3 \\ \theta_4 & \theta_5 & \theta_6 \end{pmatrix} \begin{pmatrix} 1 \\ x^T \end{pmatrix} \)

**Local motion estimation of \( \theta \) over \( V(x_0) \)** (Lucas & Kanade, 1981):

\[
E_{LK}(w_\theta(x_0)) = \int_{V(x_0)} \rho_{data}(x, w_\theta) \, dx
\]

where \( \rho_{data}(x, w_\theta) = \psi(I_2(x + w_\theta(x)) - I_1(x)) \) denotes the data potential (penalizes deviations from the data constraint)
Limitations of the local smoothing

Problem: choice of the region

V must fulfill two criteria:

- Delineating a coherently moving region to ensure the validity of the parametric assumption
- Collecting enough trustful gradient information for reliable optical flow computation
1. Local and global methods for computing optical flow

**Limitations of the local smoothing**

- **square fixed-size windows**
- **regions with variable shapes**

- parametric assumption violated over motion boundaries
- "aperture problem" for small windows
- fast optimization at each pixel

- joint motion estimation and image segmentation problem
- minimization of non-convex functionals (e.g. "motion competition", Cremers & Soatto, 2005)
Global approach and spatial regularization

Spatial regularization: penalization of high gradients $|\nabla w|$

Minimization of a global energy (Brox et al., 2004):

$$E_G(w) = \int_{\Omega} \rho_{data}(x, w) + \lambda \rho_{reg}(x, w) \, dx$$

$$= \int_{\Omega} \phi(||l_2(x + w(x)) - l_1(x)||^2) + \phi(||\nabla w(x)||^2) \, dx$$

where $\phi(z^2) = \sqrt{z^2 + \epsilon^2}$ with $\epsilon = 0.001$

Optimization method: resolution of Euler-Lagrange equations $\rightarrow$ state-of-the-art methods
Limitation of the global approach

Variational optimization:
- Difficulty to handle non-convex functionals:
  - restriction to convex penalty functions
  - over-smoothing of discontinuities
- Coarse-to-fine warping to cope with large displacements:
  - over-smoothing of fast moving details

Discrete optimization:
- Discretization of the search space and multi-label assignment
- Quantization of the flow field range:
  - compromise between computational cost and accuracy
2. Semi-local estimation and global aggregation
Semi-local estimations of flow fields

Decomposition of $I_1$ into overlapping square patches with several sizes:

Set of patches containing $x$:

$$\mathcal{V}_{s_2}(x) = \{V_1, \ldots, V_4\}$$

$$\mathcal{V}(x) = \bigcup_{i=1 \ldots 4} \mathcal{V}_{s_i}(x)$$

- Computation of flow fields over overlapping patches
- Combining adaptively the patches of $\mathcal{V}(x)$ for pointwise motion estimation
Block matching to cope with large displacements

For each patch of $I_1$, find the $N$ most similar patches $\mathcal{M}_V = \{M_1, \ldots, M_N\}$ in $I_2$ (normalized cross-correlation):

- Crude estimation of translation (i.e. pixel accuracy)
- Handle large displacements of small structures (small patch sizes)
Hierarchical motion estimation

For each pair of registered patches \((V, M_i)\), compute optical flow \(w_{V,M_i}^0\) with a variational (Brox et al., 2004) or parametric (Odobez & Odobez, 1995) method (sub-pixel accuracy, optional coarse-to-fine warping)

At each pixel \(x\), we collect several hundreds of candidates:

\[ \mathcal{W}(x) = \{ w_{V,M}(x) = (x_M - x_V) + w_{V,M}^0(x) : M \in \mathcal{M}_V, V \in \mathcal{V}(x) \} \]

\[ \text{translation} \quad \text{optical flow} \]
Global aggregation

Motivations:
- Optimization of a variational energy or parametric estimation for each patch
- “Selection of the best” competing estimators at each location

Aggregation principle: minimization of a global energy to compute a unique flow field \( w : \Omega \rightarrow \mathbb{R}^2 \) from the set \( \mathcal{W} = \bigcup_{x \in \Omega} \mathcal{W}(x) \):
- Discrete optimization (\( \mathcal{W} \) considered as a discrete space)
  \[ w = \arg \min_{w \in \mathcal{W}} E_{DA}(w) \]
- Continuous optimization (\( \mathcal{W} \) used to derive a new data term)
  \[ w = \arg \min_{w \in \mathbb{R}^2|\Omega|} E_{CA}(w, \mathcal{W}) \]
Global aggregation: discrete optimization

Finite set of candidates $\mathcal{W}(x)$ at each pixel: quantization of the space of motion vectors

Adaptive selection by global energy minimization

$$E_{DA}(w) = \sum_{x \in \Omega} \psi^A_{data}(x, w) + \beta \psi^A_{reg}(x, w)$$

Data term:

$$\psi^A_{data}(x, w) = \begin{cases} 
\rho_{data}(x, w) \\
1 - \text{NCC}_\sigma(x, w) \end{cases} \text{ (normalized cross-correlation)}$$

Regularization term: ($\mathcal{N}(x):$ spatial neighborhood of pixel $x$)

$$\psi^A_{reg}(x, w) = \sum_{y \in \mathcal{N}(x)} \phi(\|w(x) - w(y)\|^2)$$
Global aggregation: discrete optimization (cont’d’)

Discrete optimization: no restriction on the form of $E_{DA}(w)$ (differentiability, convexity)

“Fusion-Move” transforms a multi-label problem into a succession of binary labeled problems (Lempitsky et al., 2008): application to the fusion of independent pre-computed motion fields (e.g. Horn & Schunk, Lucas & Kanade, ...).
Global aggregation: continuous optimization

Limitations of discrete optimization: minimizer is found in the finite set of candidates

Continuous aggregation: aggregated motion field is allowed to deviate from the local candidates to ensure global smoothness

\[ E_{CA}(w) \int_{x \in \Omega} \Psi_{data}^{CA}(x, \mathcal{W}, w) + \beta \phi(\|\nabla w(x)\|^2) \, dx \]

A new data term to compare the candidates \( w_0(x) \in \mathcal{W}(x) \):

\[ \Psi_{data}^{CA}(x, \mathcal{W}, w) = \sum_{w_0(x) \in \mathcal{W}(x)} \alpha(x, w_0) \|w(x) - w_0(x)\|^p \]

→ variational minimization
Global aggregation: continuous optimization

Remark: if $\beta = 0$, the aggregated pointwise solution is easily found if the weights $\alpha(x, w_0)$ are uniform:

- **Mean**: $\alpha(x, w_0) = |\mathcal{W}(x)|^{-1}, p = 2$
- **Median**: $\alpha(x, w_0) = |\mathcal{W}(x)|^{-1}, p = 1$

→ experimentally insufficient

Weighting according to a “confidence measure”:

$$\alpha(x, w_0) = \frac{1}{\rho_{data}(x, w_0)} = \frac{1}{\phi(|I_2(x + w_0(x)) - I_1(x)|^2)}$$

→ $\alpha(x, w_0)$ as a function of $\text{Var}(w_0)$ is possible in the parametric case (Odobez & Bouthemy, 1995)
3. Experimental results and application
Middlebury dataset and evaluation criterion

Average Angular Error (AAE) : angle between the normalized estimated flow \( \frac{1}{\sqrt{u^2+v^2+1}}(u, v, 1) \) and true flow \( \frac{1}{\sqrt{u_{true}^2+v_{true}^2+1}}(u_{true}, v_{true}, 1) \)
Comparisons of aggregation procedures ("grove")

- Discrete approach
- Continuous approach

→ Continuous approach: over-smoothing of flow fields, low robustness to outliers
Performance of aggregation ("grove")

**Global variational method**
(Brox et al., 2004)

![Image of global variational method]

\[ \text{AAE} = 5.97 \]

**Patch-wise variational method and aggregation**

![Image of patch-wise variational method and aggregation]

\[ \text{AAE} = 5.67 \]

→ **Patch-wise approach**: sharpening of the discontinuities, small detail recovering
Performance of aggregation ("rubberwhale")

**global variational method**
(Brox et al., 2004)

**patch-wise variational method and aggregation**

<table>
<thead>
<tr>
<th>Method</th>
<th>AAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global variational method</td>
<td>3.92</td>
</tr>
<tr>
<td>Patch-wise variational method and aggregation</td>
<td>3.34</td>
</tr>
</tbody>
</table>
Parametric and variational methods ("grove")

<table>
<thead>
<tr>
<th>patch-wise parametric method</th>
<th>patch-wise variational method</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAE = 5.45</td>
<td>AAE = 5.67</td>
</tr>
</tbody>
</table>

→ parametric approach: preservation of sharp discontinuities and small details
3. Experimental results and application

Parametric and variational methods ("dimetrodon")

patch-wise parametric method

\[ \text{AAE} = 2.91 \]

\[ \rightarrow \text{variational approach: smooth temporal-varying image sequences} \]

patch-wise variational method

\[ \text{AAE} = 1.79 \]
Algorithm parameters

Size of patches:

- Several sizes to capture large and small motions
- Typical values: \( S = \{9, 19, 39, 59\} \)

Overlapping rate: not critical when \( \alpha > 0.8 \)

Block matching:

- Typical value: \( N = 2 \)
- Minor improvements for \( N > 2 \)
Discussion and conclusion

Regularization parameter and window sizes: $\lambda$ and $\beta$ are constant and spatial patch-wise or pointwise adaptivity did not improve the experimental results (window size adaptation).

Complexity of the discrete aggregation ("best" algorithm):

- Number of candidates: 200-500 vectors
- Time computations: 2-3 hours (Brox et al., 2004) or several dozens of minutes (Odobez & Bouthemy, 1995) (PC Linux 3.6 Ghz)
- Acceleration: computation of competing vectors (multi-threading), "Fusion-Move" (to be investigated)

Perspectives:

- SURE and local risk minimization need to be studied more
- Occlusion detection is currently investigated