Dense Lagrangian Motion Estimation With Occlusions
Susanna Ricco and Carlo Tomasi, Department of Computer Science, Duke University
{sricco, tomasi}@cs.duke.edu

Goal
Our goal is to extract long-range, dense correspondences from general video sequences.

We return a trajectory for each world point, listing its position in every frame in which it is visible.

Optimization
The best set of trajectories minimizes an objective function penalizing changes in brightness along visible trajectories, regularized by robust spatial and temporal smoothness terms.

\[ c^*, \nu^* = \arg \min_{c, \nu} \left\{ \lambda_1 E_D(c, \nu) + \lambda_2 E_S(c) + \lambda_3 E_V(\nu) \right\} \]

coefficients defining trajectories
visiblity of each trajectory
Data term anchors estimate to evidence from frames
Visibility term encourages smooth causal occlusions
Smoothness term encourages nearby trajectories to be similar

Solving for motion (occlusions fixed)
Objective function reduces to:

\[ c^* = \arg \min_{c} \int \left[ \left( I(p, t) - f(p, t \mid I) \right) \right] dt + \sum_k \left( \phi(p) \right) \left( \phi(p) \right) \]

This multiframe extension of traditional optical flow with explicit occlusions is solved using the Split Bregman method.

Solving for occlusions (motion fixed)
Occlusions occur when two different world points project to the same pixel in the camera, i.e., when the trajectories intersect.

To compute local occlusion estimates:
1. Find the controlling trajectory for each video pixel from candidates that pass through it.
2. Label a trajectory \( p \) as occluded at time \( t \) if the controlling trajectory at \( \mathbf{x}(p,t) \) is significantly different.

Local occlusion estimates are denoted with robust, weighted total variation temporal and spatial smoothness regularizers.

Results
Results on synthetic sequence show importance of explicit occlusion reasoning.

<table>
<thead>
<tr>
<th>Method</th>
<th>Error</th>
<th>Recall</th>
<th>Precision</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>synthetic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ours</td>
<td>0.975</td>
<td>1.000</td>
<td>0.991</td>
<td>0.991</td>
</tr>
<tr>
<td>LDOF [1]</td>
<td>0.826</td>
<td>0.816</td>
<td>0.816</td>
<td>0.816</td>
</tr>
<tr>
<td>rigid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ours</td>
<td>0.783</td>
<td>0.783</td>
<td>0.783</td>
<td>0.783</td>
</tr>
<tr>
<td>LDOF [1]</td>
<td>0.629</td>
<td>0.629</td>
<td>0.629</td>
<td>0.629</td>
</tr>
<tr>
<td>non-rigid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ours</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>LDOF [1]</td>
<td>0.973</td>
<td>0.973</td>
<td>0.973</td>
<td>0.973</td>
</tr>
</tbody>
</table>