MirrorFlow: Exploiting Symmetries in Joint Optical Flow and Occlusion Estimation

Junhwa Hur
Department of Computer Science, TU Darmstadt

Stefan Roth

Introduction

• Continued challenges for optical flow estimation: Large motion, illumination changes, and occlusion
• Optical Flow and occlusion estimation: well-known chicken and egg problem

Q) How can we resolve their mutual dependency?
• Conventional optical flow approaches:
  • Treat occlusions as outliers (violation of the basic optical flow assumptions)
  • Asymmetric algorithm + post-processing
  • Problem: occasional failures during post-processing are inevitable and irreversible

Our approach:
• Occlusion estimated in a joint objective
• An integrative, symmetric approach (no post-processing required)
  • Jointly estimate forward/backward flow, occlusion maps in both views
  • Exploiting two key symmetry properties: forward/backward & occlusion/disocclusion
  • Coupling two different problem domains
  • Better utilizing the available image evidence
  • Leading to a well-balanced solution

Method

• Objective function:
  \[ E(\mathbf{u}, \mathbf{v}, \mathbf{o}, \mathbf{s}) = E_D(\mathbf{u}, \mathbf{v}, \mathbf{o}, \mathbf{s}) + \lambda E_F(\mathbf{u}, \mathbf{v}, \mathbf{s}) + \lambda E_C(\mathbf{u}, \mathbf{v}, \mathbf{o}, \mathbf{s}) + \lambda E_O(\mathbf{u}, \mathbf{v}, \mathbf{o}) \]
  \[ \mathbf{o} = \{0, 1\} \text{ per pixel} \]

\[ E_D(\mathbf{u}, \mathbf{v}, \mathbf{o}, \mathbf{s}) \text{ Data term} \]

\[ E_F(\mathbf{u}, \mathbf{v}, \mathbf{s}) \text{ Forward term} \]

\[ E_C(\mathbf{u}, \mathbf{v}, \mathbf{o}, \mathbf{s}) \text{ Consistency term} \]

\[ E_O(\mathbf{u}, \mathbf{v}, \mathbf{o}) \text{ Occlusion term} \]

\[ \mathbf{H} \text{: Per-superpixel homography for forward motion} \]

\[ \mathbf{O} \text{: Per-superpixel homography for backward motion} \]

\[ \mathbf{o} = \{0, 1\} \text{ per pixel} \]

• Forward-backward consistency term
  • Penalizing motion inconsistency only for visible pixels in both views

\[ E_F(\mathbf{u}, \mathbf{v}, \mathbf{s}) = \sum_{b \in B} \sum_{p \in b} c^F_p \left( \mathbf{I}_{1,b} - \mathbf{H}_{1,b} \mathbf{I}_{2,b} \right) \]

\[ E_C(\mathbf{u}, \mathbf{v}, \mathbf{s}) = \sum_{b \in B} \sum_{p \in b} c^C_p \left( \mathbf{I}_{1,b} - \mathbf{H}_{1,b} \mathbf{I}_{2,b} \right) \]

• Occlusion-disocclusion symmetry term
  • Penalizing occlusion and disocclusion inconsistency

\[ E_O(\mathbf{u}, \mathbf{v}, \mathbf{o}) = \sum_{b \in B} \sum_{p \in b} c^O_p \left( \mathbf{I}_{1,b} - \mathbf{H}_{1,b} \mathbf{I}_{2,b} \right) \]

• Optimization (block coordinate descent)
  • Estimating \( \mathbf{H} \) and \( \mathbf{O} \) : discrete multi-label optimization
  • Collecting proposal homography motions
  • Fusion moves
  • Estimating \( \mathbf{o} \): binary optimization
  • Standard graph cut

Experiments

• KITTI Optical Flow 2015
  • #1 two-frame based optical flow method
  • Lowest flow error rates in occluded regions

• MPI Sintel Flow Dataset
  • #2 in the Clean pass

Conclusions & Future Work

• Our joint, symmetric design yields significant improvements in flow estimation accuracy, especially in occluded areas.
• Leading results on public benchmarks even without employing any semantic knowledge or learning of appearance descriptors.
• Enabling to handle non-rigid motion, exploiting semantic knowledge, and using learned descriptors are natural next steps.