MirrorFlow: Exploiting Symmetries in Joint Optical Flow and Occlusion Estimation

ICCV 17
International Conference on Computer Vision 2017

Junhwa Hur   Stefan Roth
Department of Computer Science, TU Darmstadt

TECHNISCHE UNIVERSITÄT DARMSTADT


Forward optical flow

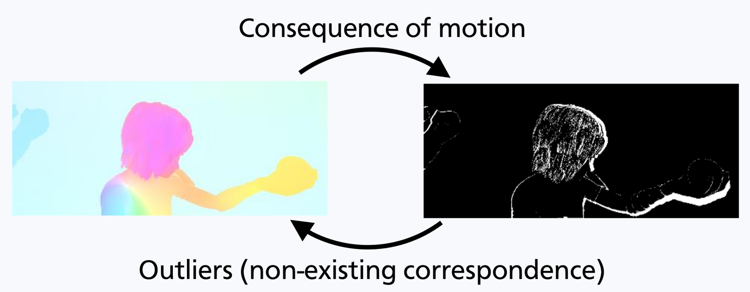

Backward optical flow


Occlusion map on the current frame


Occlusion map on the next frame

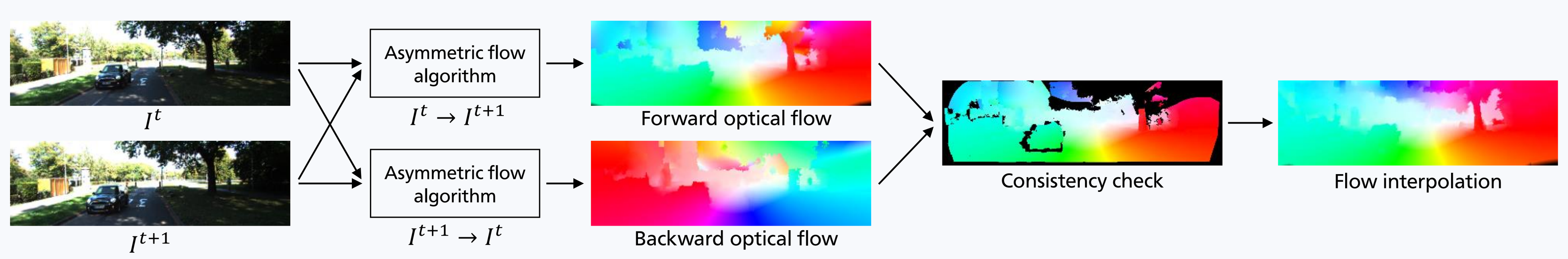
Introduction

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

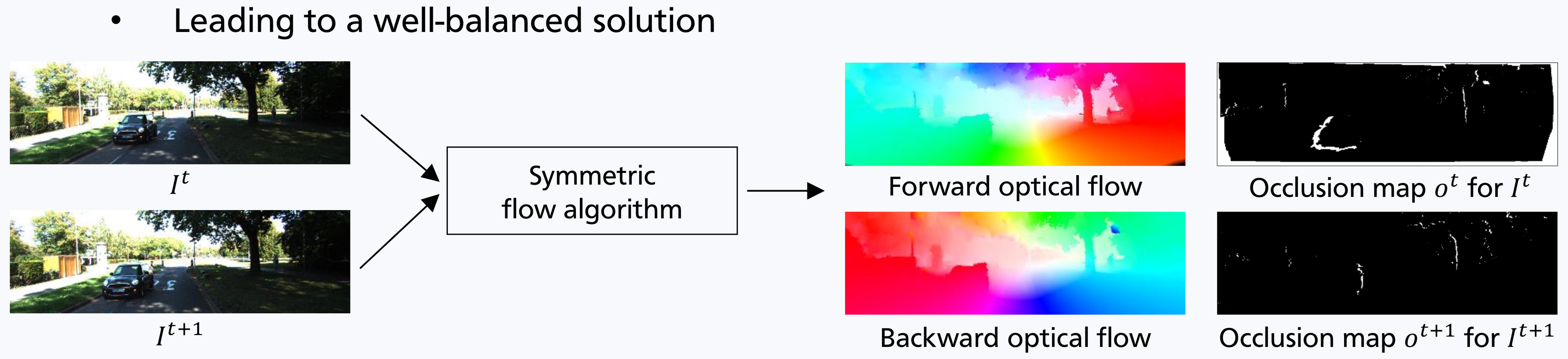


Q) How can we resolve their mutual dependency?

- **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



Acknowledgement. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC Grant Agreement No. 307942.

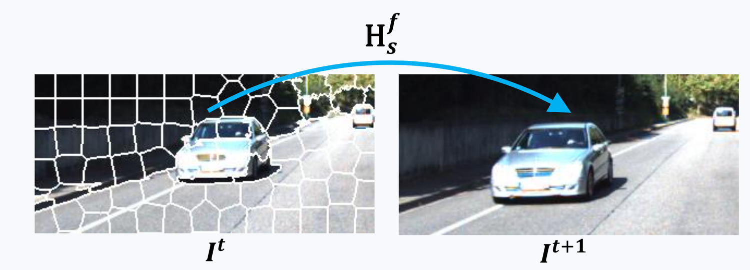
[1] E. Ilg, N. Mayer, T. Saikia, M. Keuper, A. Dosovitskiy, and T. Brox. FlowNet 2.0: Evolution of optical flow estimation with deep networks. In CVPR, 2017.
[2] M. Bai, W. Luo, K. Kundu, and R. Urtasun. Exploiting semantic information and deep matching for optical flow. In ECCV, 2016.
[3] J. Wulff, L. Sevilla-Lara, and M. J. Black. Optical flow in mostly rigid scenes. In CVPR, 2017.
[4] J. Xu, R. Ranftl, and V. Koltun. Accurate optical flow via direct cost volume processing. In CVPR, 2017.
[5] C. Bailer, K. Varanasi, and D. Stricker. CNN-based patch matching for optical flow with thresholded hinge embedding loss. In CVPR, 2017.
[6] Y. Yang and S. Soatto. S2F: Slow-to-fast interpolator flow. In CVPR, 2017.
[7] Y. Hu, Y. Li, and R. Song. Robust interpolation of correspondences for large displacement optical flow. In CVPR, 2017.

Method

- **Objective function:**

$$
\begin{aligned}
E(\mathrm{H}^f,\mathrm{H}^b,o^t,o^{t+1}) &= E_D(\mathrm{H}^f,\mathrm{H}^b,o^t,o^{t+1}) && \text{Data term}\\
&+ \lambda_P E_P(\mathrm{H}^f,\mathrm{H}^b,o^t,o^{t+1}) && \text{Pairwise term}\\
&+ \lambda_C E_C(\mathrm{H}^f,\mathrm{H}^b,o^t,o^{t+1}) && \text{Forward-backward consistency term}\\
&+ \lambda_S E_S(o^t,o^{t+1}) && \text{Occlusion-disocclusion symmetry term}
\end{aligned}
$$

H^f: Per-superpixel homography for forward motion
H^b: Per-superpixel homography for backward motion
o^t: Per-pixel occlusion map for the current frame
o^{t+1}: Per-pixel occlusion map for the next frame



$\mathrm{H}_s \in \mathcal{R}^{3\times3}$: homography motion per each superpixel

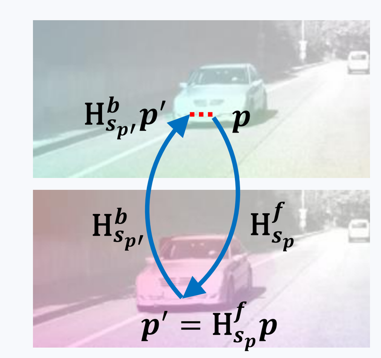
	Visible pixels ($o_p = 0$)	Occluded pixels ($o_p = 1$)
Data term	Truncated ternary census transform	Constant penalty (avoiding the trivial solution)
Pairwise term	Penalizing motion & occlusion status in 8-neighborhood	
Consistency term	Forward-backward inconsistency penalty	0
Symmetry term	Occlusion-disocclusion inconsistency penalty	

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

$$E_C(\mathrm{H}^f,\mathrm{H}^b,o^t,o^{t+1}) = \sum_{p\in I^t} C_p^f + \sum_{p\in I^{t+1}} C_p^b$$

$$C_p^f = \overline{o_p^t}\,\overline{o_{p'}^{t+1}}\,\rho_C\left(\left\|p - \mathrm{H}_{s_{p'}}^b \mathrm{H}_{s_p}^f p\right\|\right) \qquad p' = \mathrm{H}_{s_p}^f p$$

$$C_p^b = \overline{o_p^{t+1}}\,\overline{o_{p'}^t}\,\rho_C\left(\left\|p - \mathrm{H}_{s_{p'}}^f \mathrm{H}_{s_p}^b p\right\|\right) \qquad p' = \mathrm{H}_{s_p}^b p$$



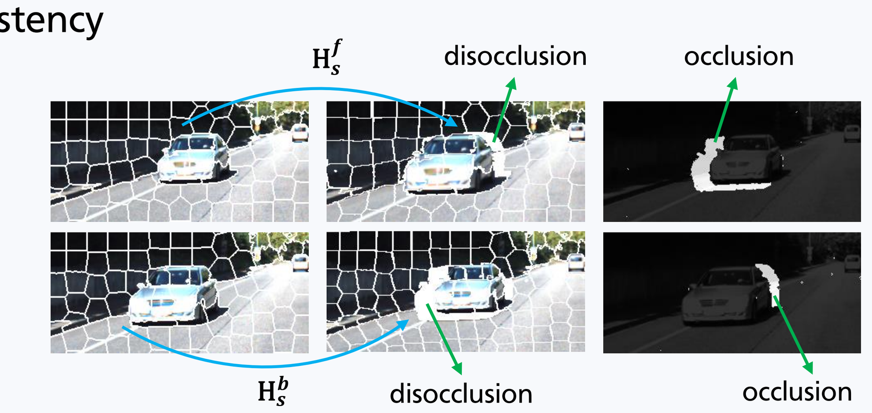
- **Occlusion-disocclusion symmetry term**
  - Penalizing occlusion and disocclusion inconsistency

$$E_S(o^t,o^{t+1}) = \sum_{p\in I^t} o_p^t \odot N_p^t + \sum_{p\in I^{t+1}} o_p^{t+1} \odot N_p^{t+1}$$

$$N_p^{t+1} = \left|\left\{p \mid p = \mathrm{H}_{s_{p'}}^f p', \quad \forall p' \in I^t\right\}\right|$$

$$N_p^t = \left|\left\{p \mid p = \mathrm{H}_{s_{p'}}^b p', \quad \forall p' \in I^{t+1}\right\}\right|$$

(Number of pixels that are mapped to p from warping)



XOR Operation:
$0 \odot 0 = 1$
$0 \odot 1 = 0$

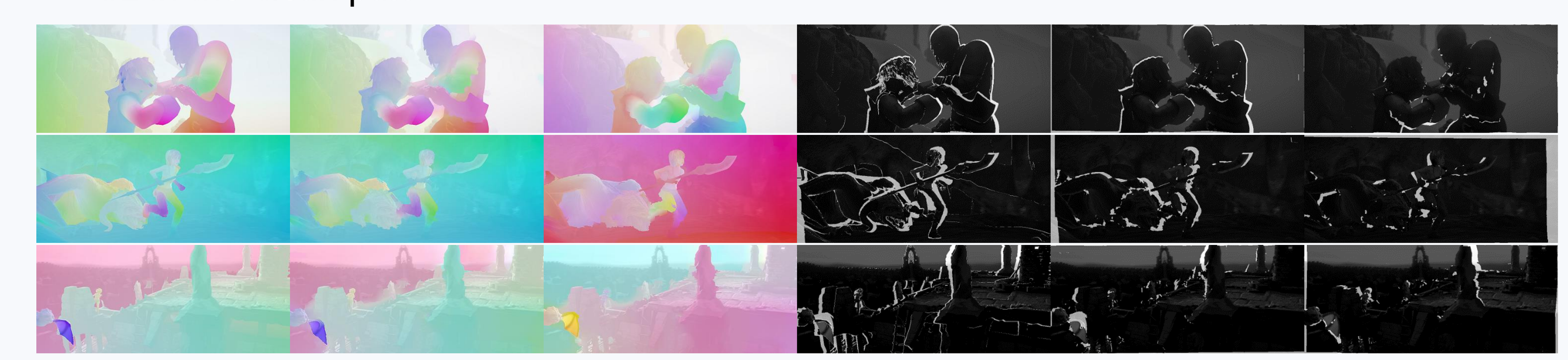
- **Optimization (block coordinate descent)**
  - Estimating H_s^f and H_s^b : discrete multi-label optimization
    - Collecting proposal homography motions
    - Fusion moves
  - Estimating o^t, o^{t+1} : binary optimization
    - Standard graph cut

Experiments

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


Ground truth optical flow · Forward optical flow · Backward optical flow · Occlusion map on the current frame · Occlusion map on the next frame

- **MPI Sintel Flow Dataset**
  - #2 in the Clean pass


Ground truth optical flow · Forward optical flow · Backward optical flow · Ground truth occlusion map · Occlusion map on the current frame · Occlusion map on the next frame

Method	Non-occluded pixels			All pixels		
	Fl-bg	Fl-fg	Fl-all	Fl-bg	Fl-fg	Fl-all
Ours (MirrorFlow)	6.24 %	12.95 %	7.46 %	8.93 %	17.07 %	**10.29 %**
FlowNet2 [1]	7.24 %	**5.60 %**	**6.94 %**	10.75 %	**8.75 %**	10.41 %
SDF [2]	**5.75 %**	18.38 %	8.04 %	**8.61 %**	23.01 %	11.01 %
MR-Flow [3]	6.86 %	17.91 %	8.86 %	10.13 %	22.51 %	12.19 %
DCFlow [4]	8.04 %	19.84 %	10.18 %	13.10 %	23.70 %	14.86 %

KITTI Optical Flow 2015   * Flow outlier rates

Method	Final pass			Clean pass		
	EPE all	EPE nocc.	EPE occ.	EPE all	EPE nocc.	EPE occ.
DCFlow [4]	**5.119**	**2.283**	28.228	3.537	1.103	23.394
FlowFieldsCNN [5]	5.363	2.303	30.313	3.778	0.996	26.469
MR-Flow [3]	5.376	2.818	**26.235**	**2.527**	**0.954**	**15.365**
S2F-IF [6]	5.417	2.549	28.795	3.500	0.988	23.986
RicFlow [7]	5.620	2.765	28.907	3.550	1.264	22.220
Ours (MirrorFlow)	6.071	3.186	29.567	3.316	1.338	19.470

MPI Sintel Flow Dataset   * EPE (End-point error)

- **Ablation study**
  - The occlusion-disocclusion symmetry term contributes the most
  - The F/B consistency term boosts the quality of the results further, but only when the symmetry term is turned on

Usage of terms		Non-occluded pixels			All pixels		
Occ/Disocc symmetry	F/B Consistency	Fl-bg	Fl-fg	Fl-all	Fl-bg	Fl-fg	Fl-all
Yes	**Yes**	**6.52 %**	**11.72 %**	**7.41 %**	**9.26 %**	**13.94 %**	**9.98 %**
Yes	No	6.73 %	11.90 %	7.62 %	9.49 %	14.04 %	10.19 %
No	**Yes**	10.41 %	18.17 %	11.74 %	13.72 %	20.40 %	14.74 %
No	No	8.39 %	14.97 %	9.51 %	11.82 %	17.41 %	12.68 %

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.