Upgrading Optical Flow to 3D Scene Flow through Optical Expansion

Gengshan Yang[1], Deva Ramanan[1,2]

[1]Robotics Institute, Carnegie Mellon University

[2]Argo AI

Input from a drone Optical flow Optical expansion

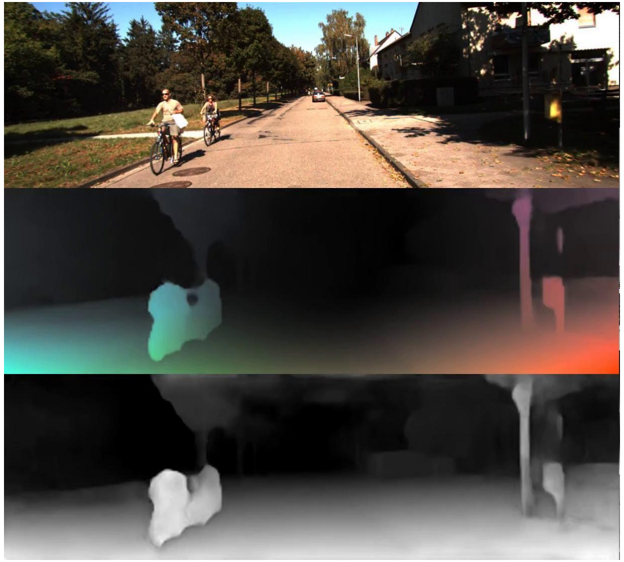
"Red-Tailed Hawk vs. Drone. Hawk Wins." Video from The New York Times.

CVPR 2020

Monocular 3D Scene Motion Estimation

Problem: Estimate the 3D motion of dynamic scene elements using a monocular camera.

Input (drone) | Optical flow | Optical expansion



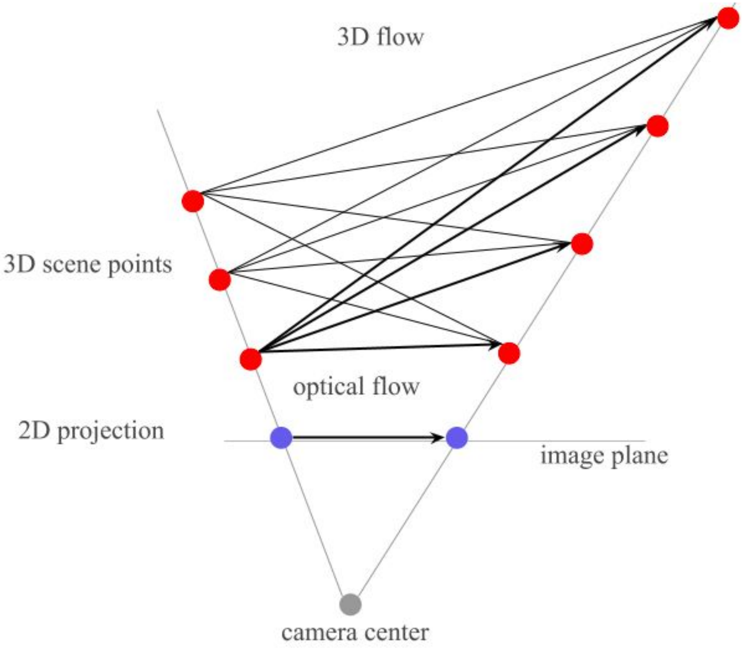
Input (car)

Optical flow

Optical expansion

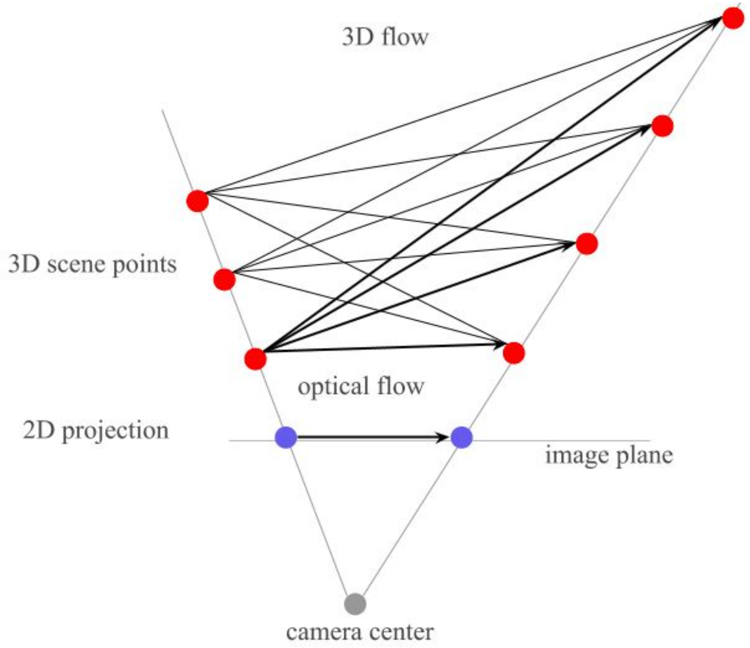
Monocular 3D Scene Motion Estimation

Challenge: Infinite pairs of 3D points correspond to the 2D flow observation.



Monocular 3D Scene Motion Estimation

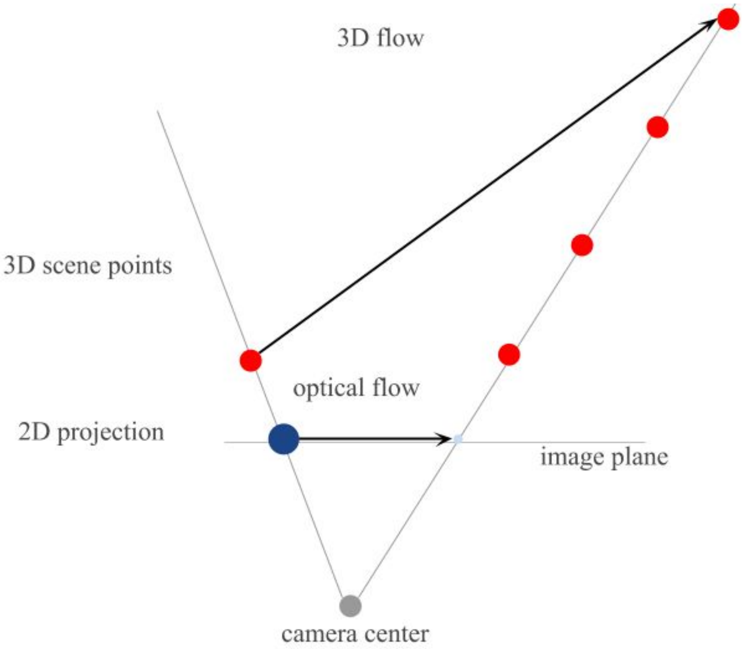
Challenge: Infinite pairs of 3D points correspond to the 2D flow observation.



Prior work

- Motion prior [Kumar *et al.*, ICCV 17, ...]
- Data-driven depth prior [Brickwedde *et al.*, ICCV 19, ...]

Optical Expansion and Motion-in-depth

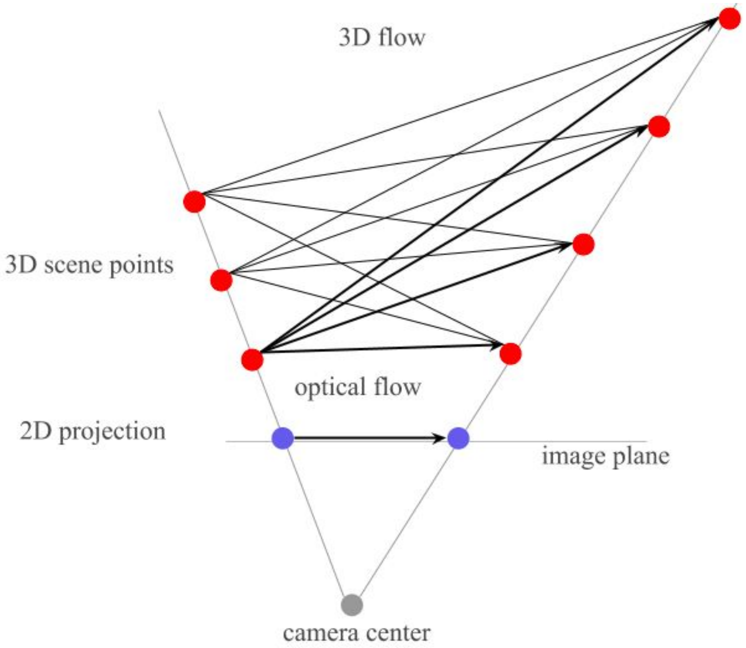


Change of perceptual size corresponds to change of physical depth.

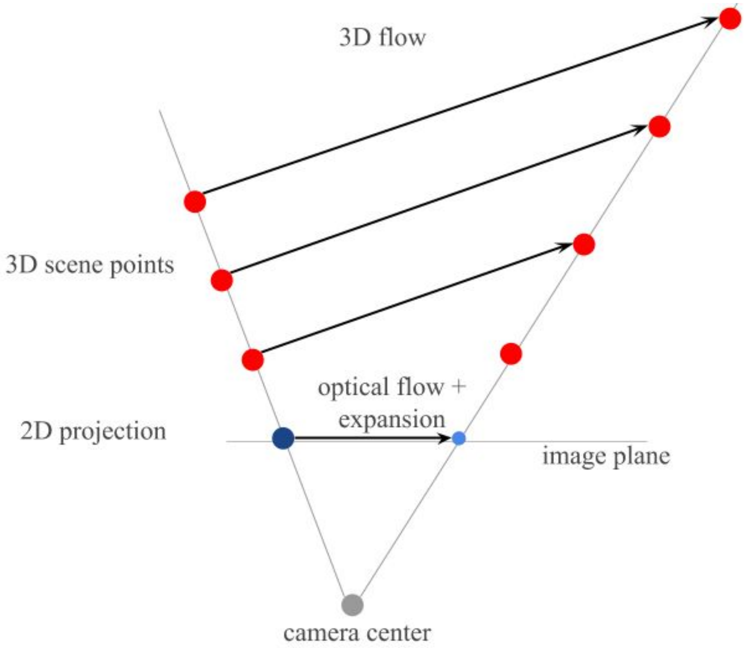
Swanston, Michael T., and Walter C. Gogel. "Perceived size and motion in depth from optical expansion." *Perception & psychophysics* 39.5 (1986): 309-326.

Upgrading Optical Flow to 3D Scene Flow through Optical Expansion
Gengshan Yang, Deva Ramanan. CVPR 2020.

Upgrading to 3D Scene Flow

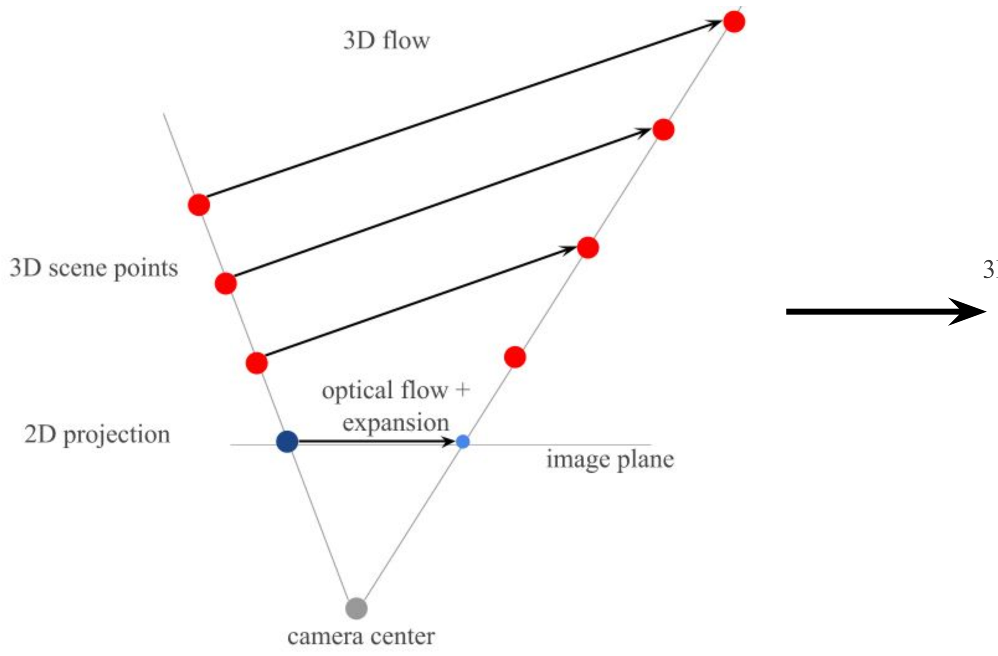


(a) with optical flow

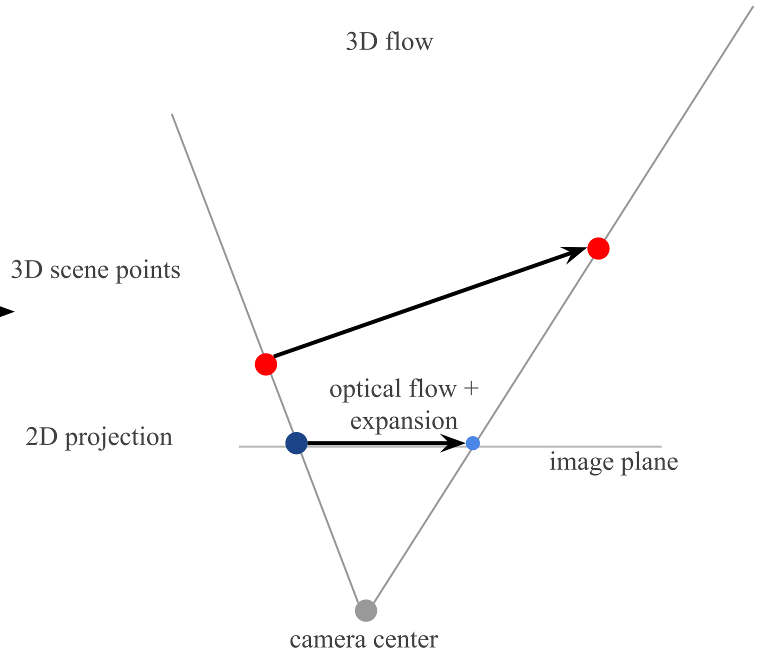


(b) optical flow + expansion

Upgrading to 3D Scene Flow



(b) optical flow + expansion



(c) optical flow + expansion + 1st or 2nd depth

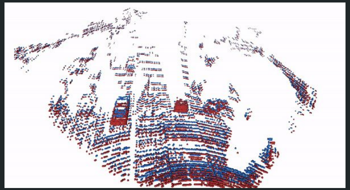
Pipeline Overview

Input frame pair

Optical Flow Estimation

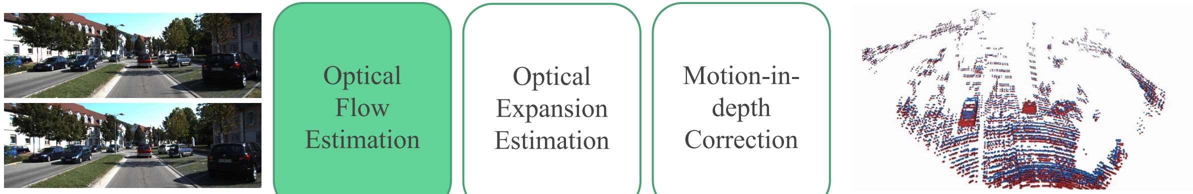
Optical Expansion Estimation

Motion-in-depth Correction



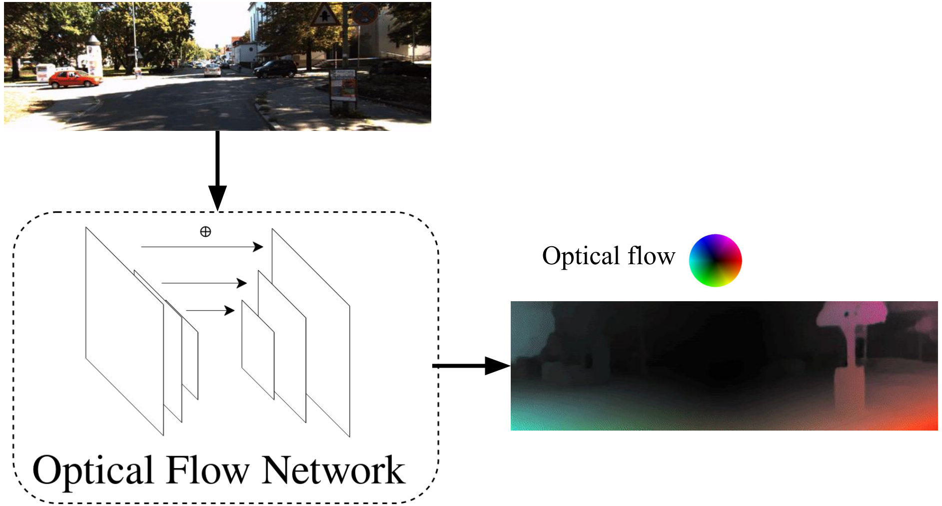
Output 3D scene flow

Upgrading Optical Flow to 3D Scene Flow through Optical Expansion
Gengshan Yang, Deva Ramanan. CVPR 2020.

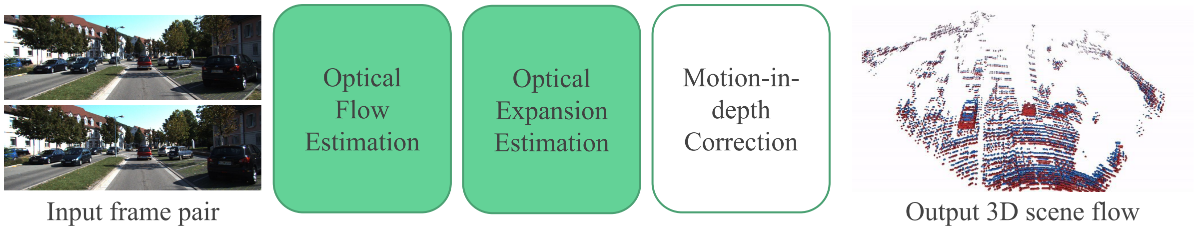


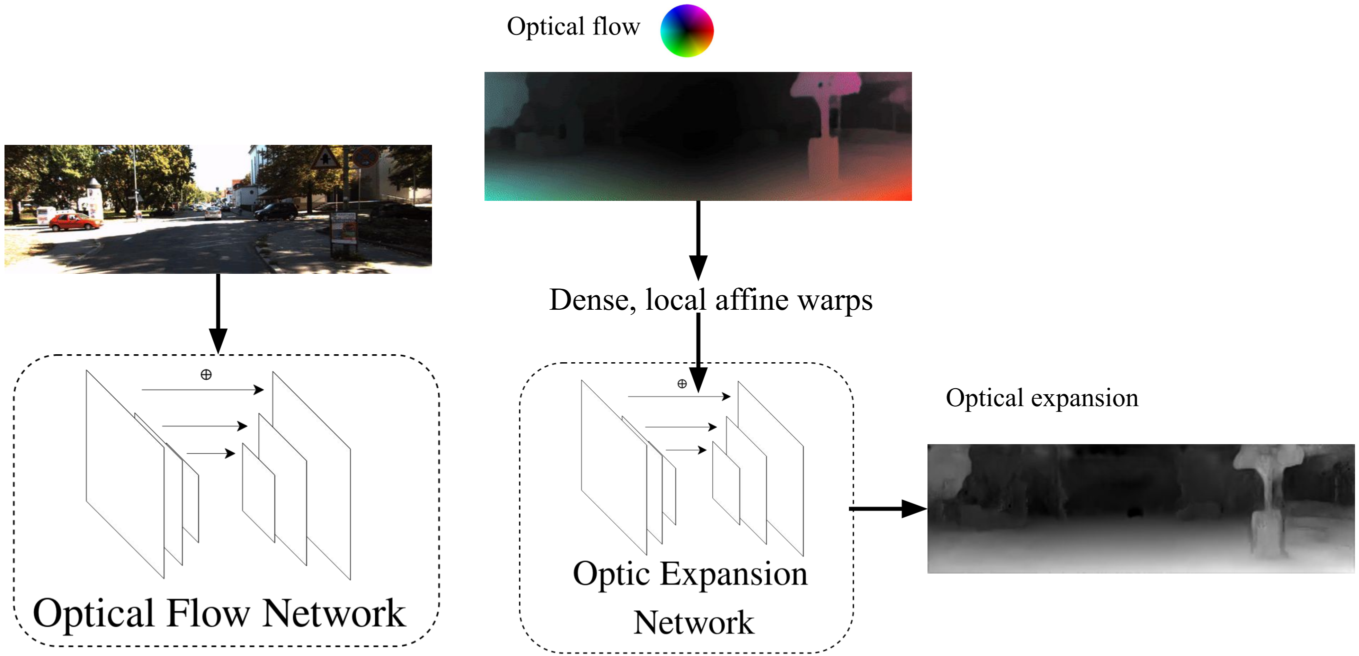
Input frame pair

Output 3D scene flow

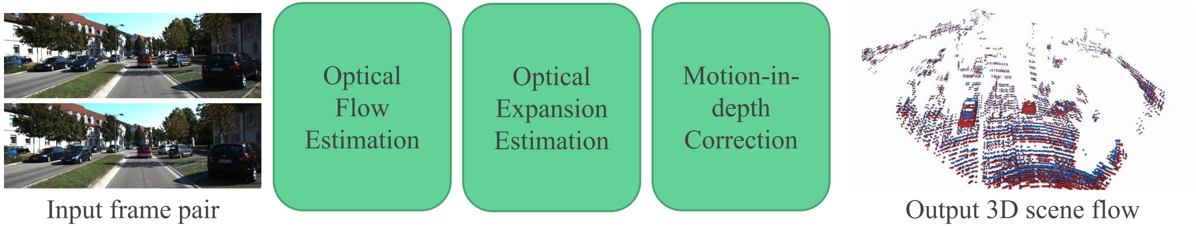


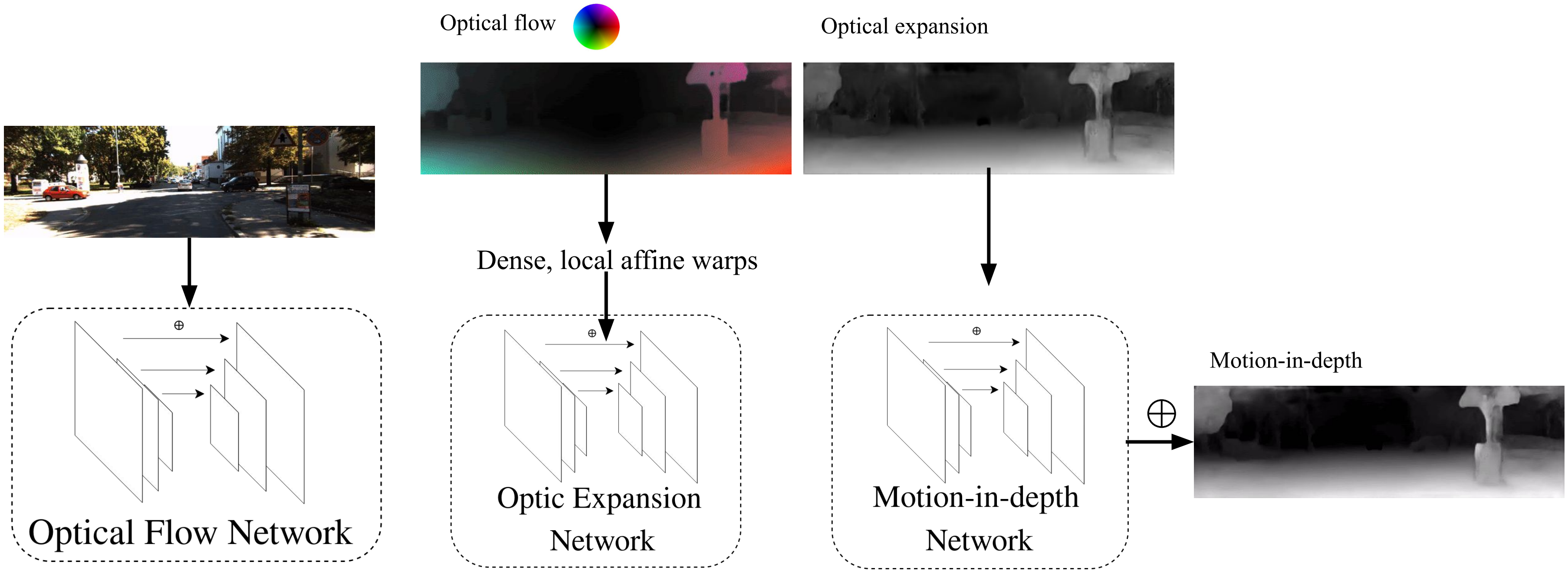
Upgrading Optical Flow to 3D Scene Flow through Optical Expansion
Gengshan Yang, Deva Ramanan. CVPR 2020.

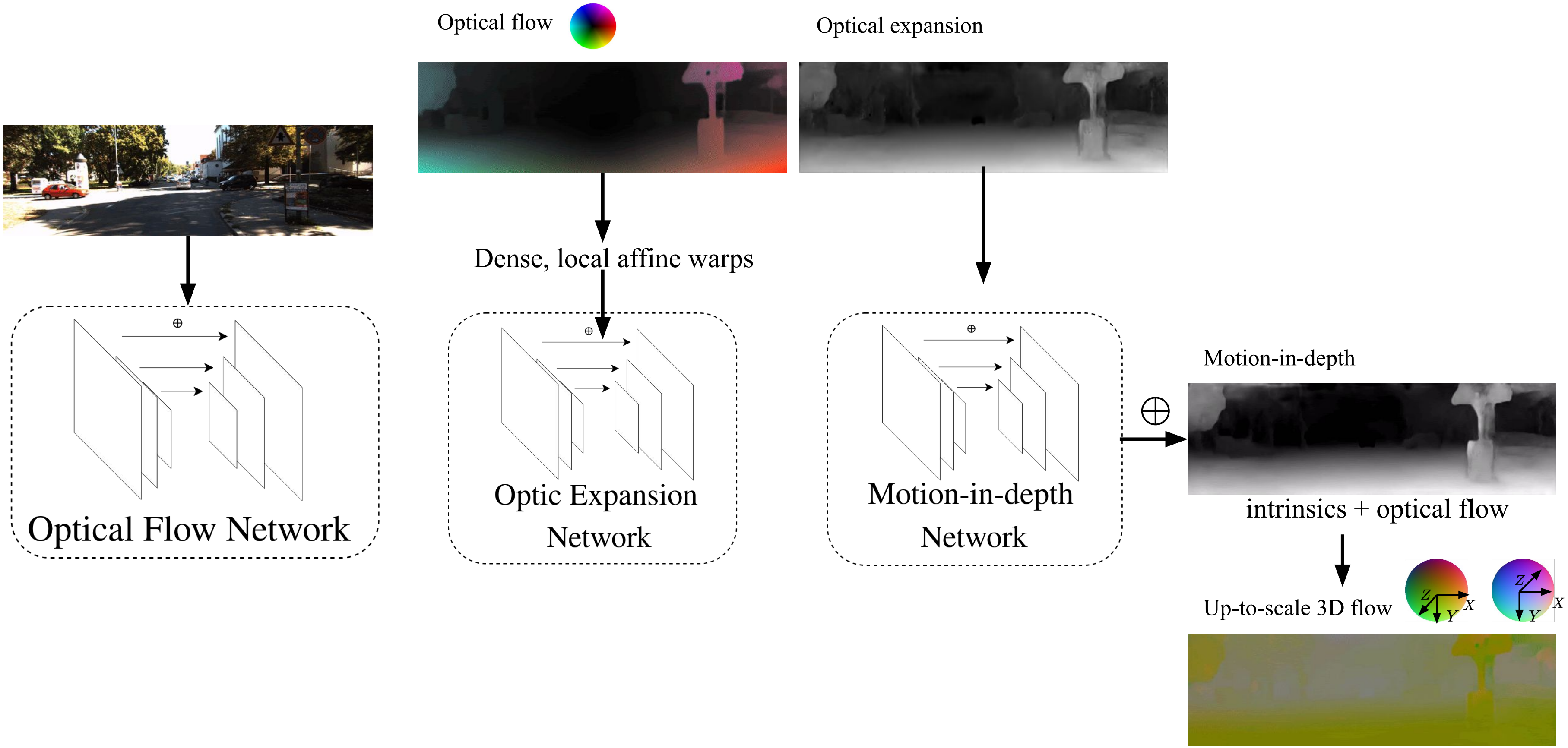


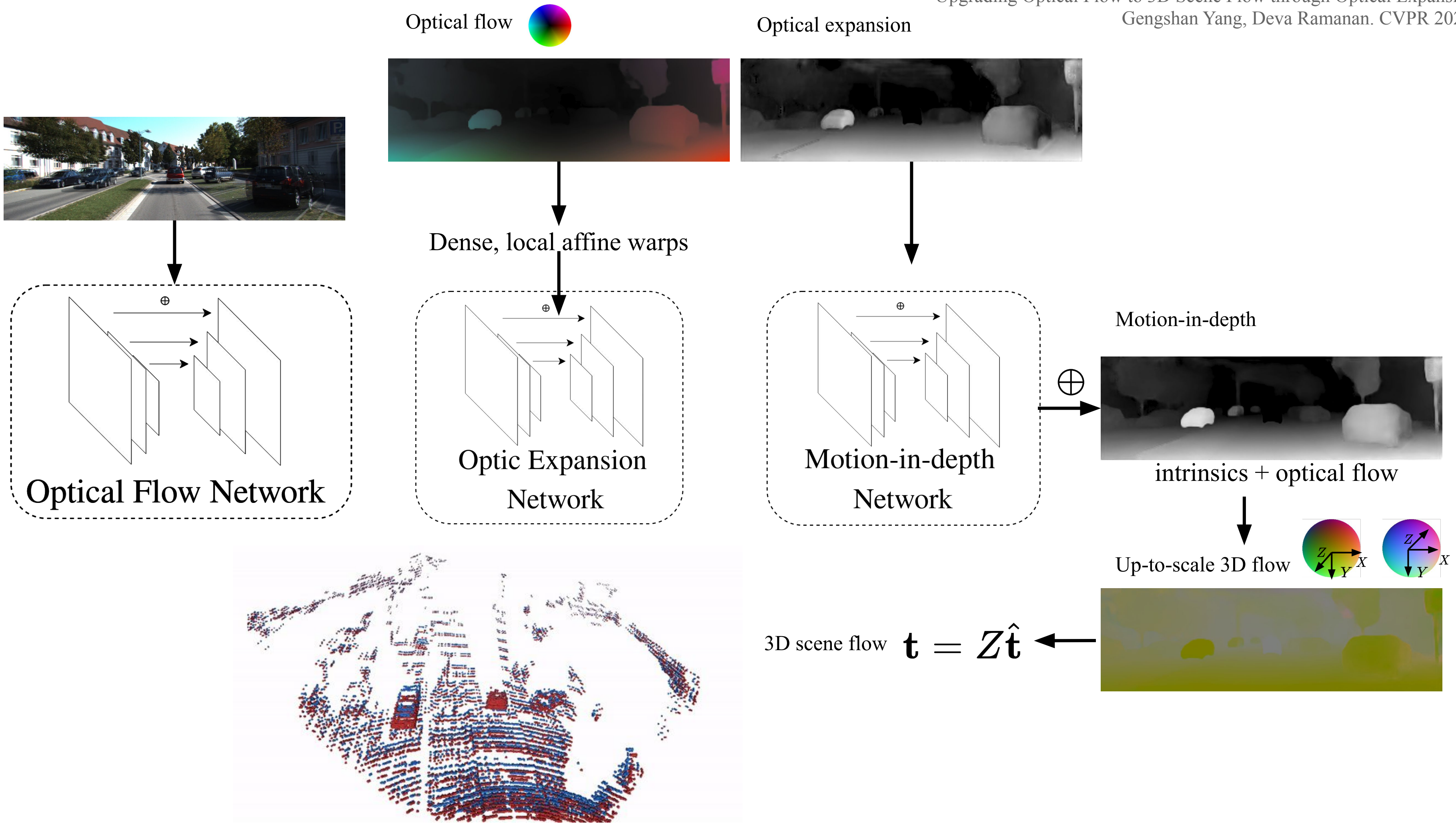


Upgrading Optical Flow to 3D Scene Flow through Optical Expansion
Gengshan Yang, Deva Ramanan. CVPR 2020.

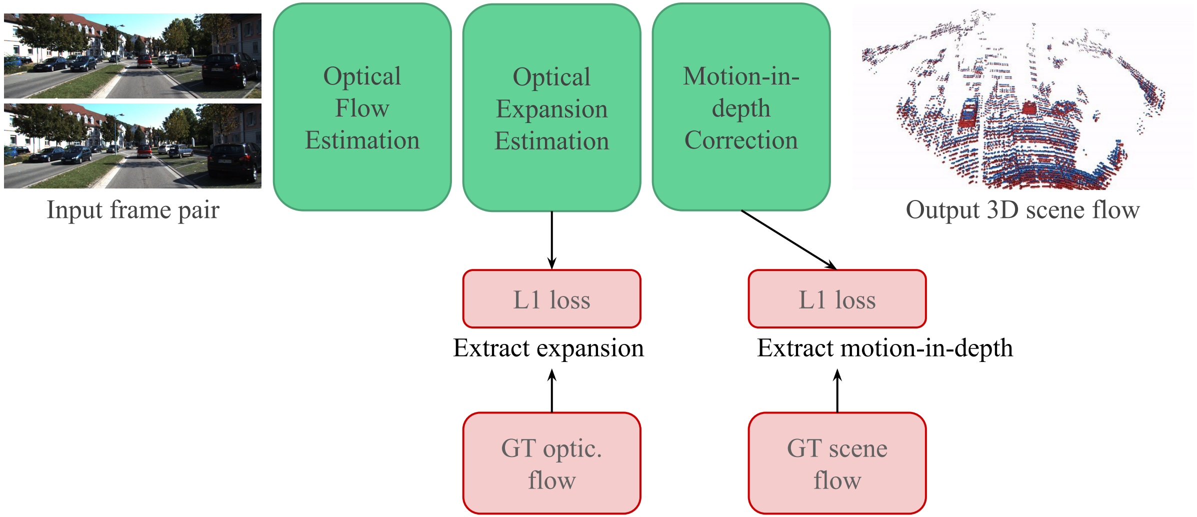





Optical flow
Optical expansion
Dense, local affine warps
Optical Flow Network
Optic Expansion Network
Motion-in-depth Network
Motion-in-depth
intrinsics + optical flow
Up-to-scale 3D flow
Z
X
Y
Z
X
Y

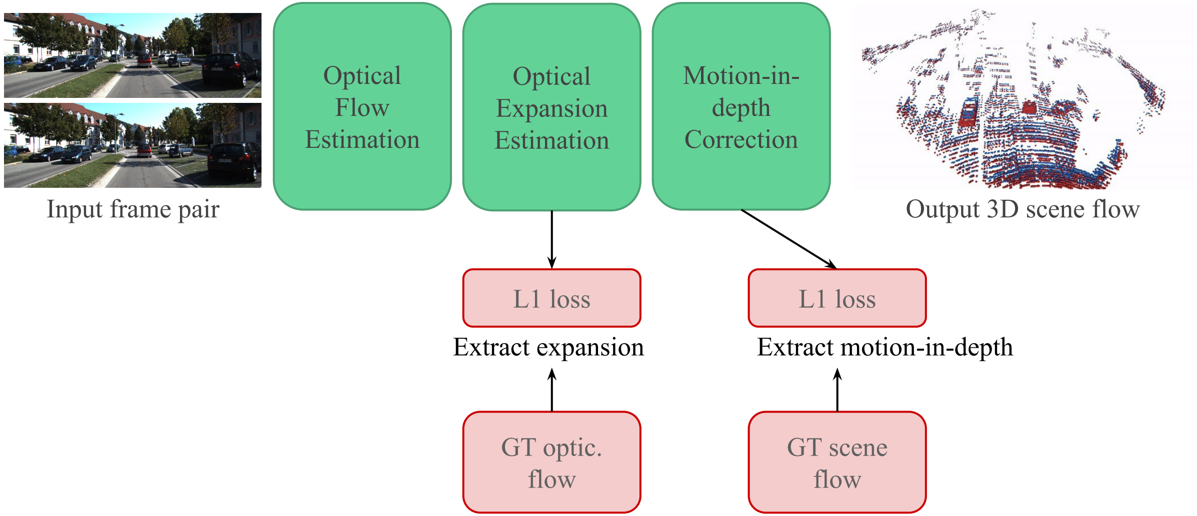

Optical flow
Optical expansion
Dense, local affine warps
Optical Flow Network
Optic Expansion Network
Motion-in-depth Network
Motion-in-depth
intrinsics + optical flow
Up-to-scale 3D flow
3D scene flow $\mathbf{t} = Z\hat{\mathbf{t}}$

Learning for 3D Scene Flow Upgrade



Multi-task losses for optical expansion and motion-in-depth estimation.

Learning for 3D Scene Flow Upgrade

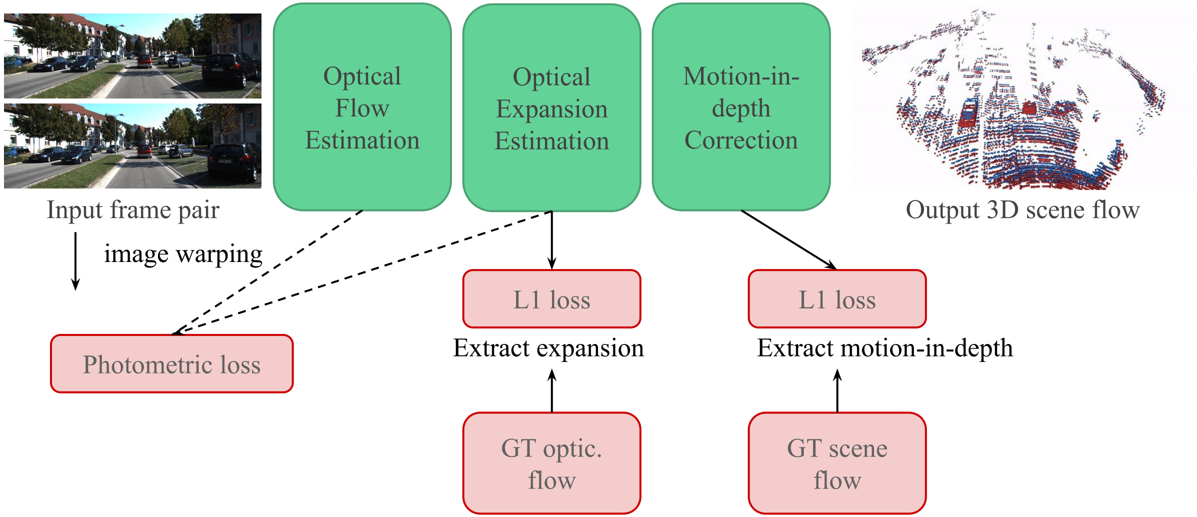


Training procedure

1. Pre-train with synthetic Scene Flow Datasets [CVPR 2016]
2. Fine-tune on target domain data, KITTI [JPRS 2018].

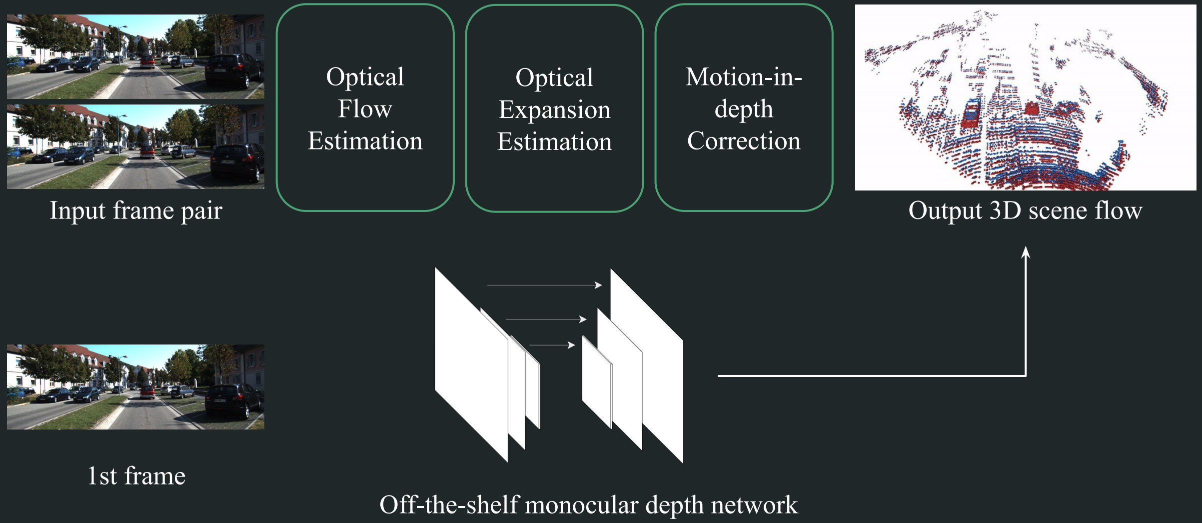
Upgrading Optical Flow to 3D Scene Flow through Optical Expansion
Gengshan Yang, Deva Ramanan. CVPR 2020.

Learning for 3D Scene Flow Upgrade

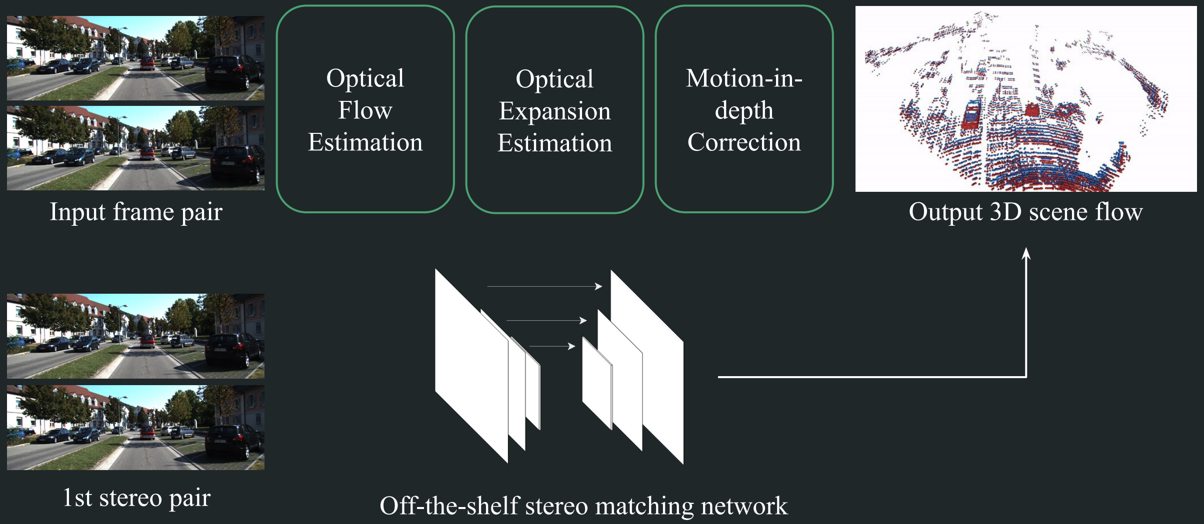


Self-supervised training for optical expansion and optical flow estimation.

Application: Monocular Scene Flow



Application: Stereo Scene Flow



Monocular / Stereo Scene Flow

Input

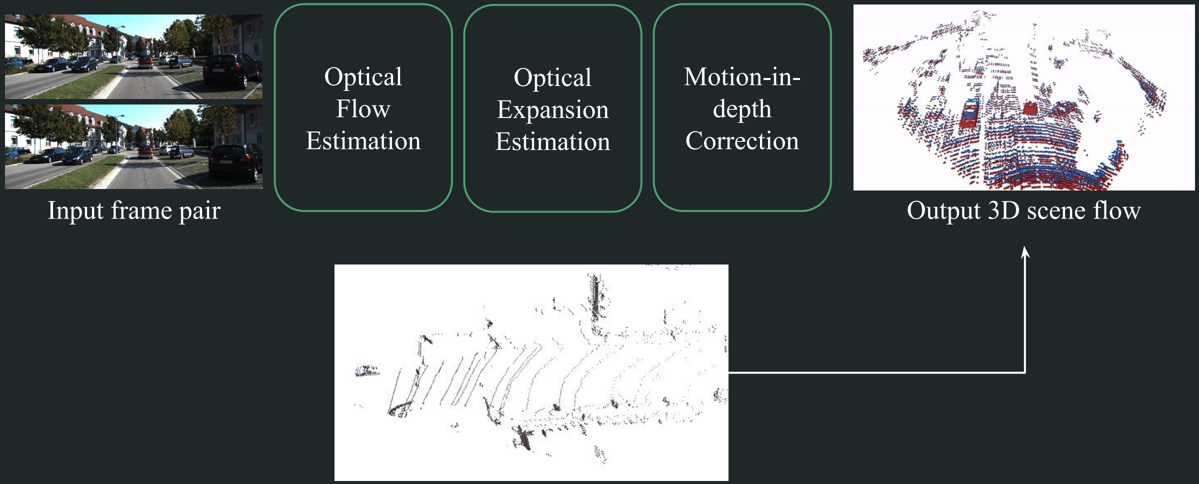
Relative depth change
(motion-in-depth)

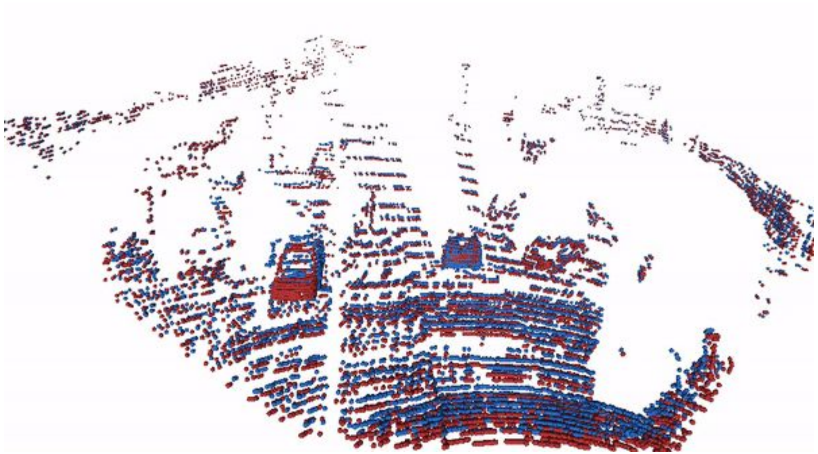
"Flow warping" [1] FlowNet-3 [2] Ours

SOTA monocular and stereo scene flow performance on foreground objects of KITTI leaderboard.

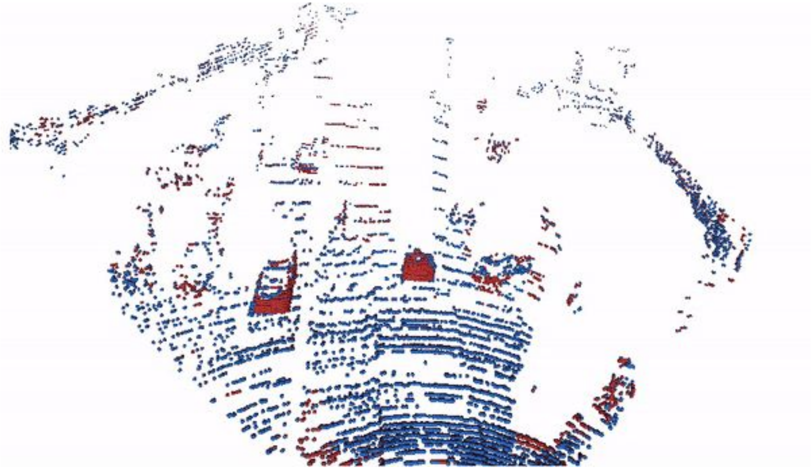
[1] Schuster, René, et al. "Combining stereo disparity and optical flow for basic scene flow." Commercial Vehicle Technology 2018. Springer Vieweg, Wiesbaden, 2018. 90-101.
[2] Ilg, Eddy, et al. "Occlusions, motion and depth boundaries with a generic network for disparity, optical flow or scene flow estimation." ECCV. 2018.

Application: LiDAR Scene Flow





Result of HPLFlowNet



Our result

- High-accuracy than state-of-the-art lidar-only methods
- Can be computed before the next LiDAR sweep is captured

Gu, Xiuye, et al. "HPLFlowNet: Hierarchical Permutohedral Lattice FlowNet for Scene Flow Estimation on Large-scale Point Clouds." CVPR. 2019.

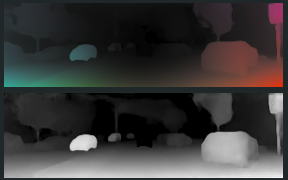
Application: Two-frame SFM

Input frame pair

Optical Flow Estimation

Optical Expansion Estimation

Motion-in-depth Correction



up-to-scale 3D flow

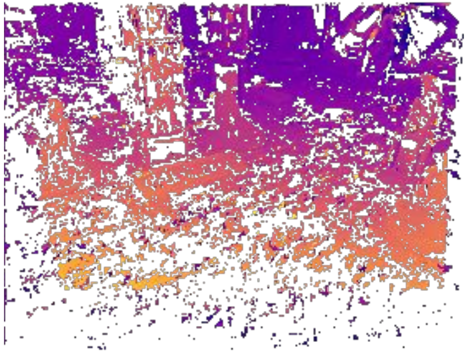
Rigid structure

Inverse depth

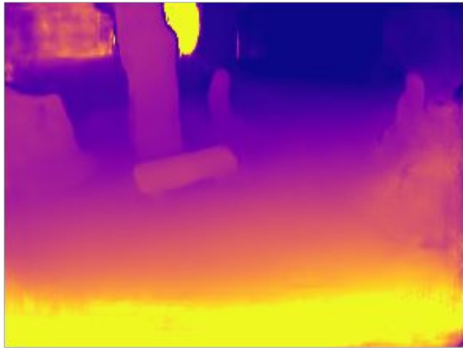
Residual error

Iteration 0 | Iteration 1 | • • • | Iteration 5

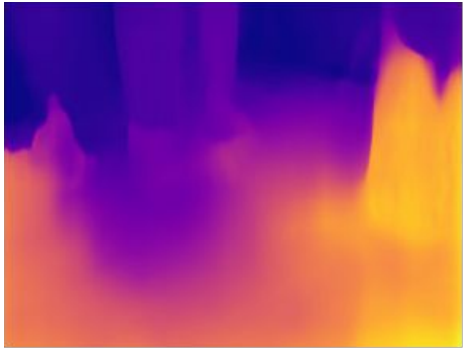
Overlaid two frames [1]



Result from COLMAP [2] (two view)



Ours



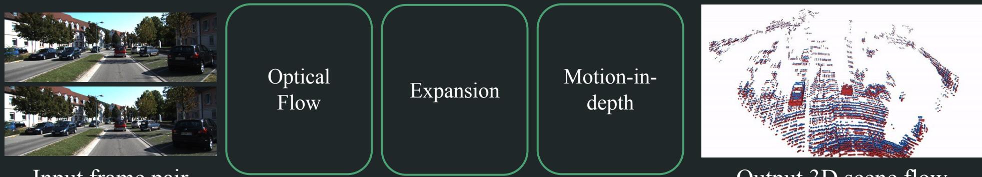
Result from MonoDepth2 [3]

Residual error

[1] Antonini, Amado, et al. "The blackbird dataset: A large-scale dataset for UAV perception in aggressive flight." arXiv preprint arXiv:1810.01987 (2018).
[2] Schonberger, Johannes L., and Jan-Michael Frahm. "Structure-from-motion revisited." CVPR. 2016.
[3] Godard, Clément, et al. "Digging into self-supervised monocular depth estimation." ICCV. 2019.

Thanks! More in our paper ...

- Formalism for upgrading 2D optical flow to 3D scene flow



Input frame pair

Output 3D scene flow

- Optical expansion is the crucial ingredient enabling the above

- If you are using optical-flow-for-X, consider using optical-expansion as well!