State-Of-The-Art on Optical Flow

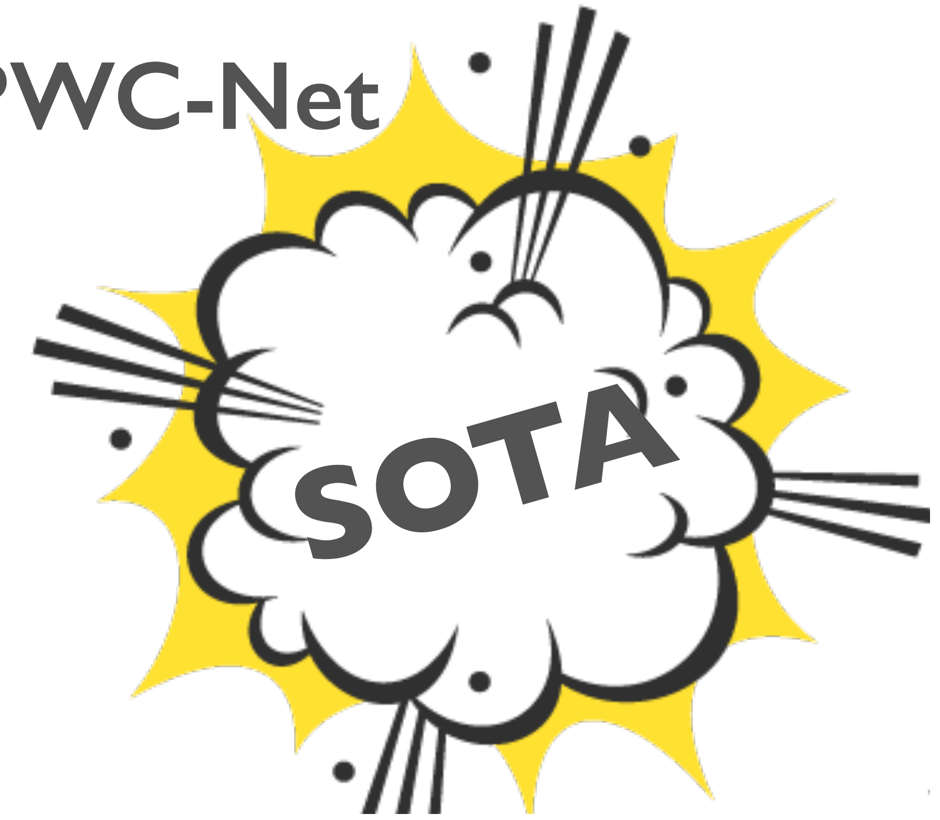
A Quick Intro

- B. Sc. (2005 - 2010)
  Bilkent University
- M. Sc. (2010 - 2012)
  Bogazici University
- Ph. D. (2013 - 2017)
  MPI for Intelligent Systems
- Postdoc (2017 - 2019)
  University of Oxford
- Asst. Prof. (2019 - current)
  Koc University

- Stereo Matching using Object Knowledge
- 3D Reconstruction by Exploiting Object Similarity
- Learning Features for Optical Flow Estimation
- Optical Flow Reference Data using High-speed Cameras
- End-to-end training of Optical Flow and Action Recognition
- Unsupervised Learning of Multi-frame Optical Flow and Occlusions
- Multi-object Tracking

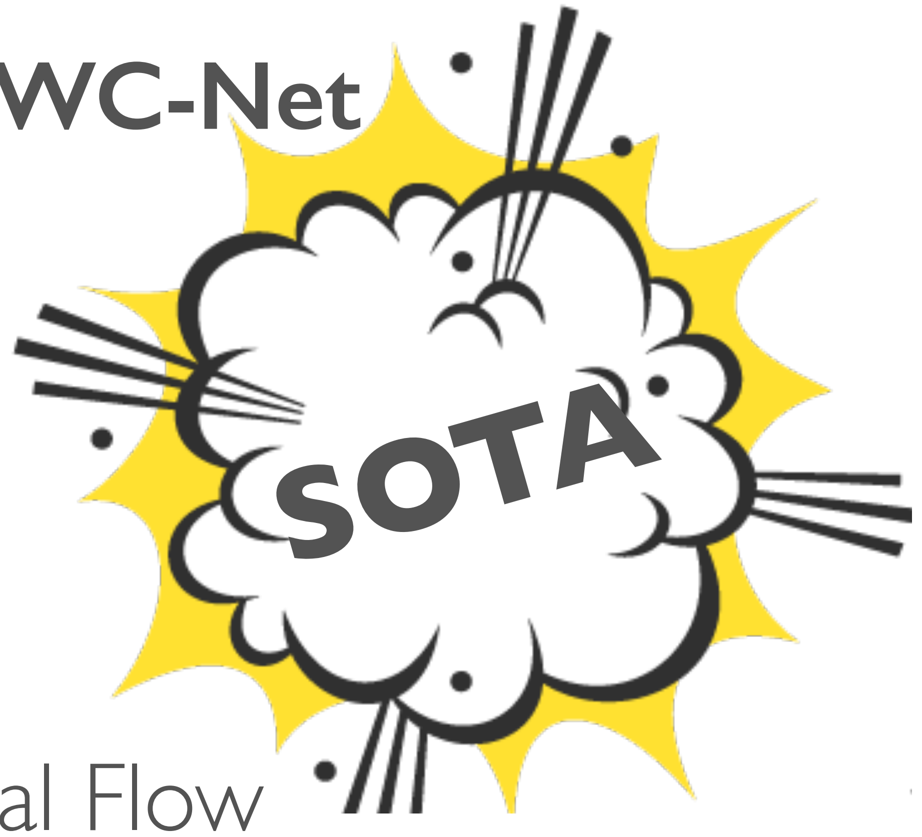
Outline

- Introduction
- End-to-End Learning of Optical Flow
    - Pyramid, Warping, and Cost Volume: **PWC-Net**
    - +Occlusion Handling: **MaskFlownet**



Outline

- Introduction
- End-to-End Learning of Optical Flow
  - Pyramid, Warping, and Cost Volume: **PWC-Net**
  - +Occlusion Handling: **MaskFlownet**

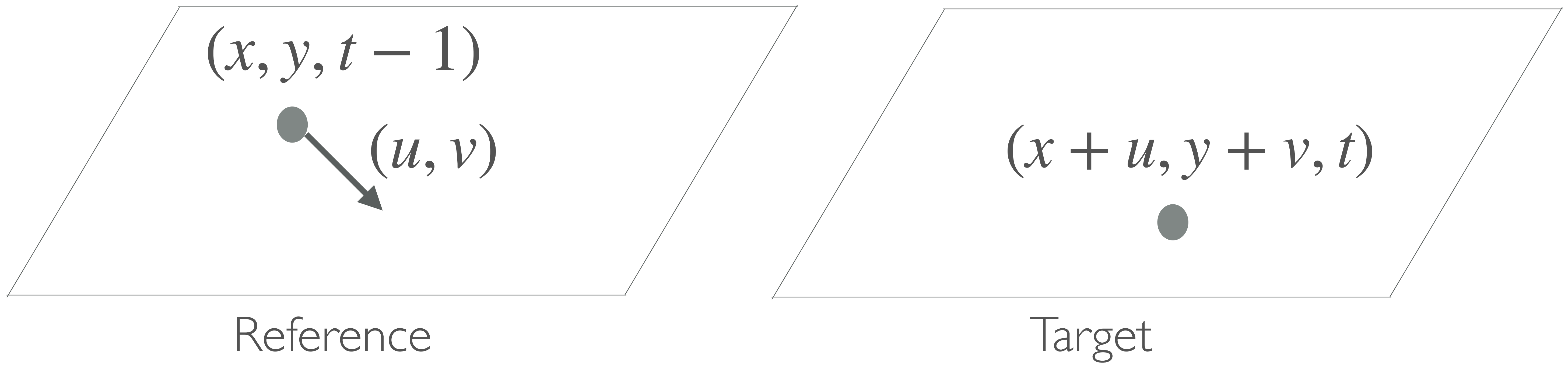


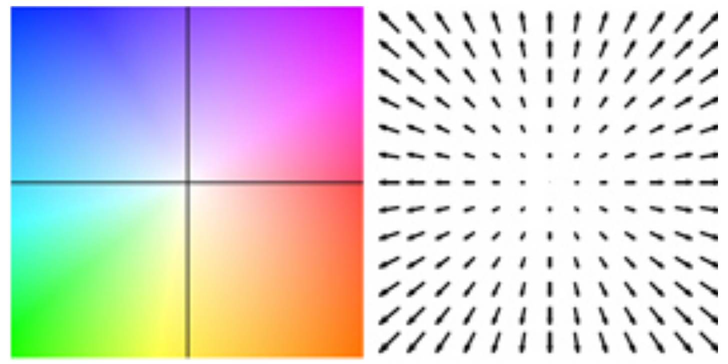
- The Problem of Data
  - A New Dataset for Real-World Optical Flow
  - Unsupervised Learning of Multi-frame Optical Flow with Occlusions
    - Data Distillation
    - Good Practices (by google)

What is Optical Flow?

The motion of pixels between consecutive frames in time.

a 2D flow vector at each pixel

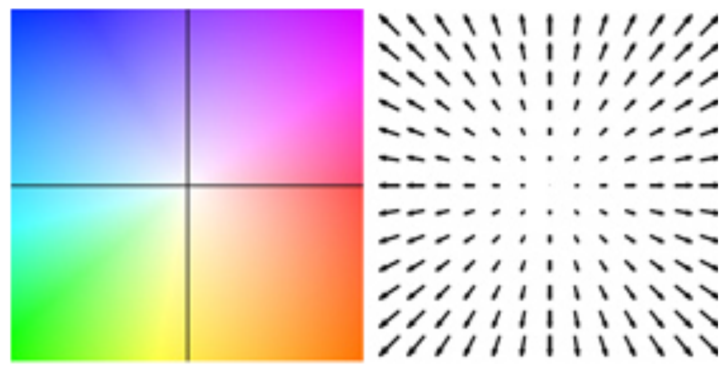




What is Optical Flow?

The motion of pixels between consecutive frames in time.

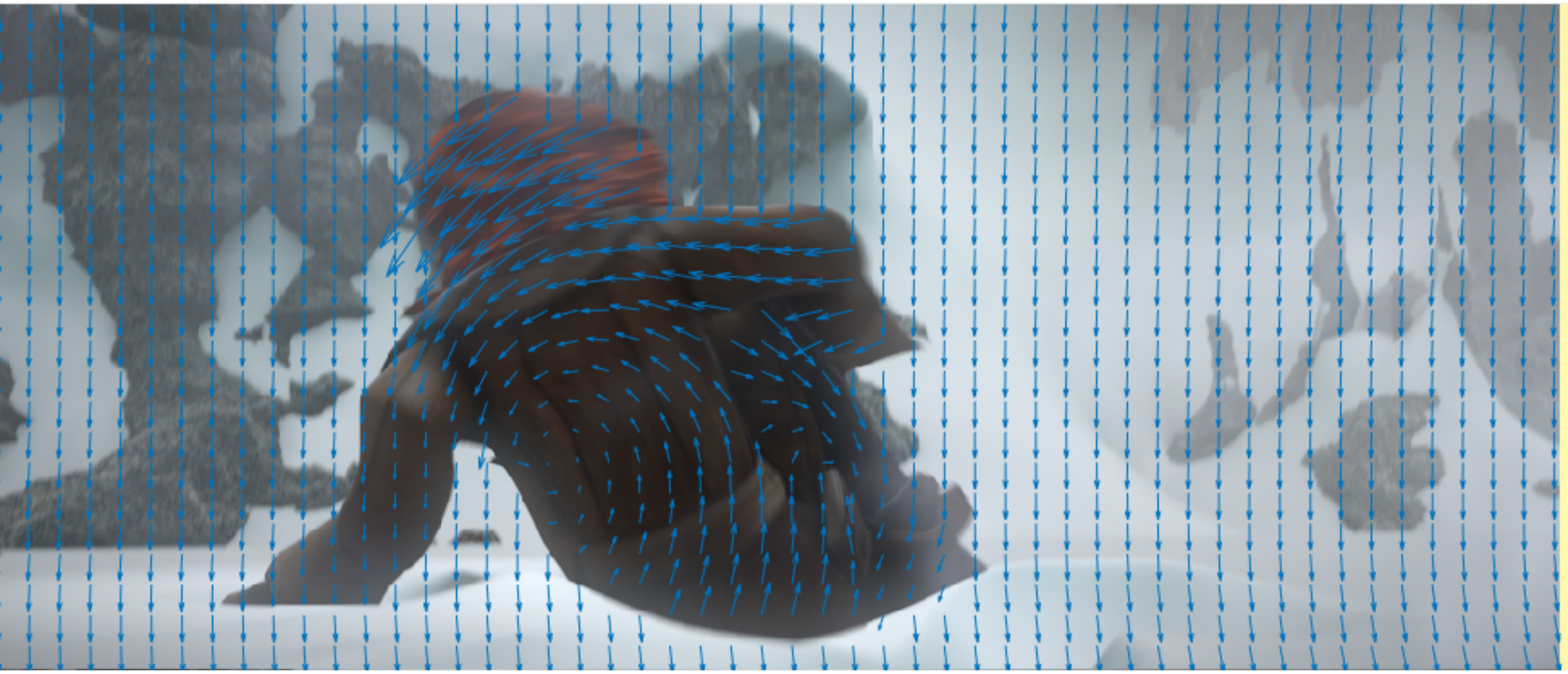
a 2D flow vector at each pixel

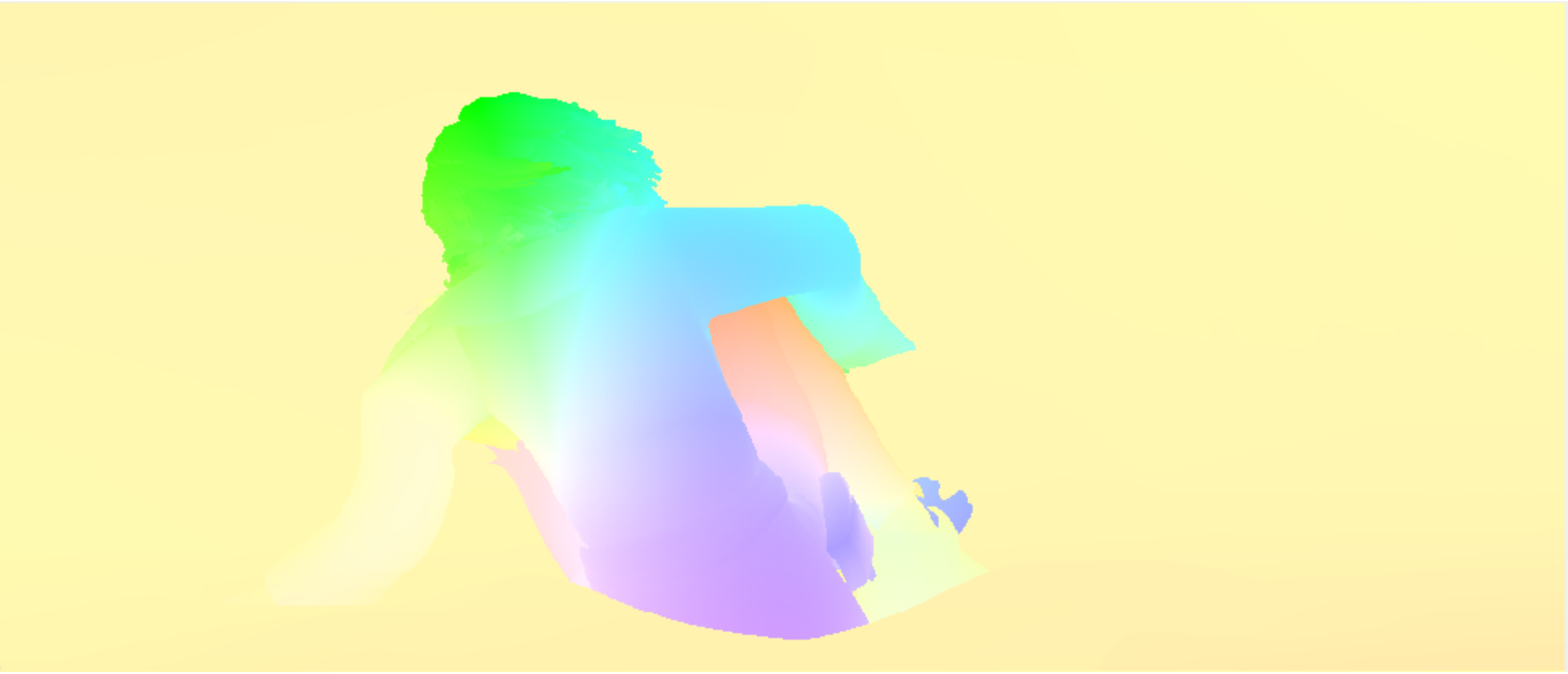


What is Optical Flow?

The motion of pixels between consecutive frames in time.

a 2D flow vector at each pixel





Why Optical Flow?

- Processing of video sequences
- Action recognition, object detection and tracking,
  video denoising, frame interpolation and video summarization
- Autonomous driving
  ego-motion estimation, structure from motion

How to Solve Optical Flow?

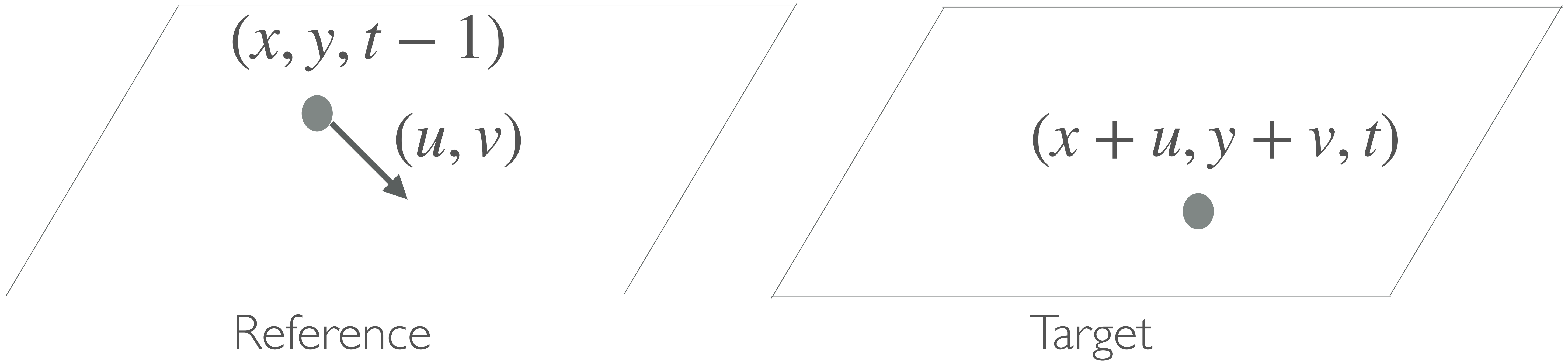
- Given a pixel in t-1, look for nearby pixels of the same color in t

How to Solve Optical Flow?

- Given a pixel in t-1, look for nearby pixels of the same color in t
- Color constancy: a point in t-1 looks the same in t.
- Small motion: points do not move very far.

How to Solve Optical Flow?

Brightness constancy assumption



$$I(x, y, t-1) = I(x+u, y+v, t)$$

How to Solve Optical Flow?

Linearizing the right side using Taylor Expansion:

$$I(x, y, t-1) = I(x+u, y+v, t)$$

$$I(x, y, t-1) = I(x, y, t) + I_x\, u + I_y\, v$$

$$I_x\, u + I_y\, v + I_t \approx 0$$

How to Solve Optical Flow?

Aperture problem: 2 unknowns, 1 equation

$$I_x\, u + I_y\, v + I_t \approx 0$$

How to Solve Optical Flow?

Aperture problem: 2 unknowns, 1 equation

$$I_x\, u + I_y\, v + I_t \approx 0$$

Solution: Additional constraints

Traditional Methods

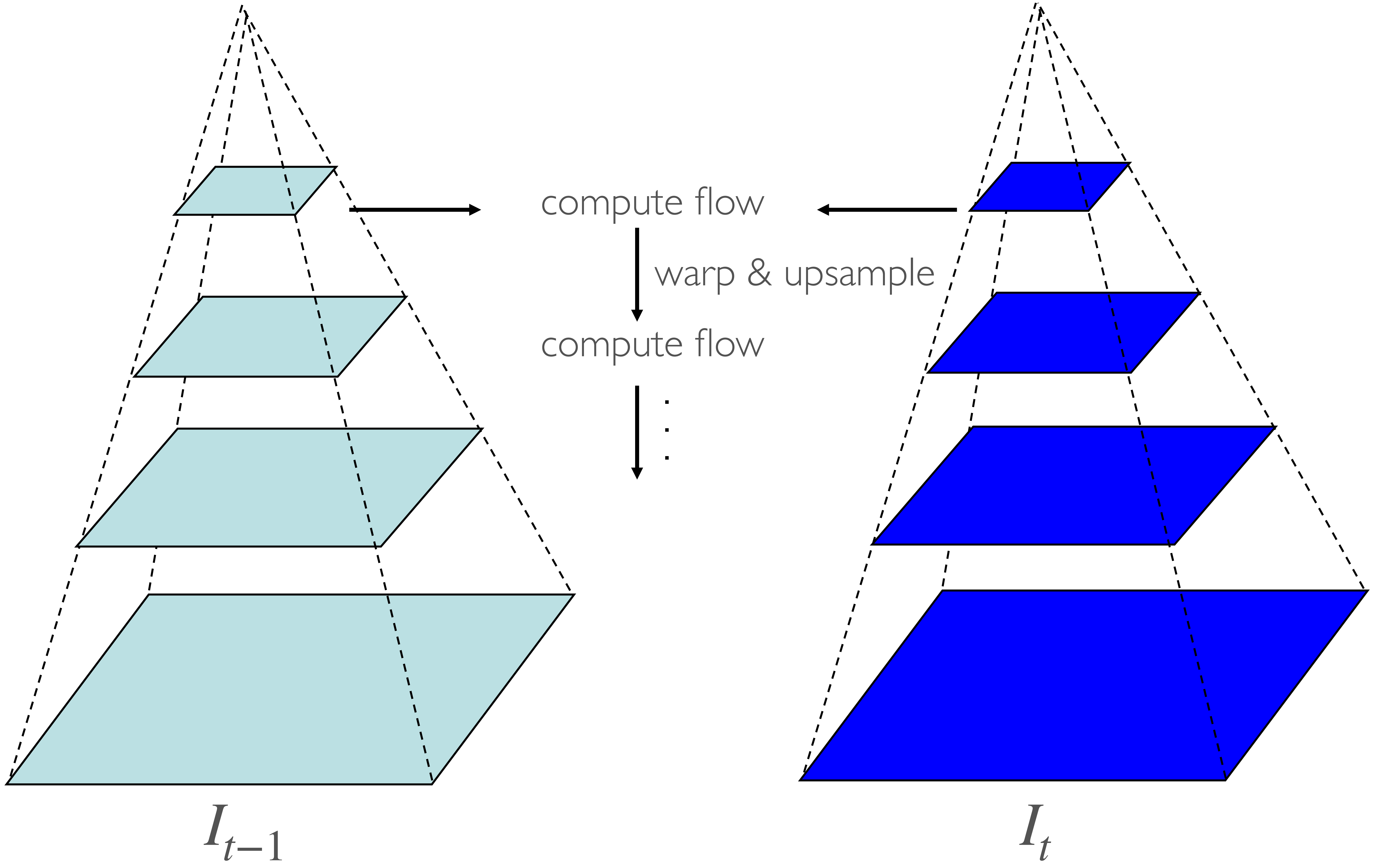
- **Lukas-Kanade**: Spatial coherence constraint
  same flow within a window

- **Horn**&**Schunck**: Smoothness constraint
  neighbouring pixels having similar flow

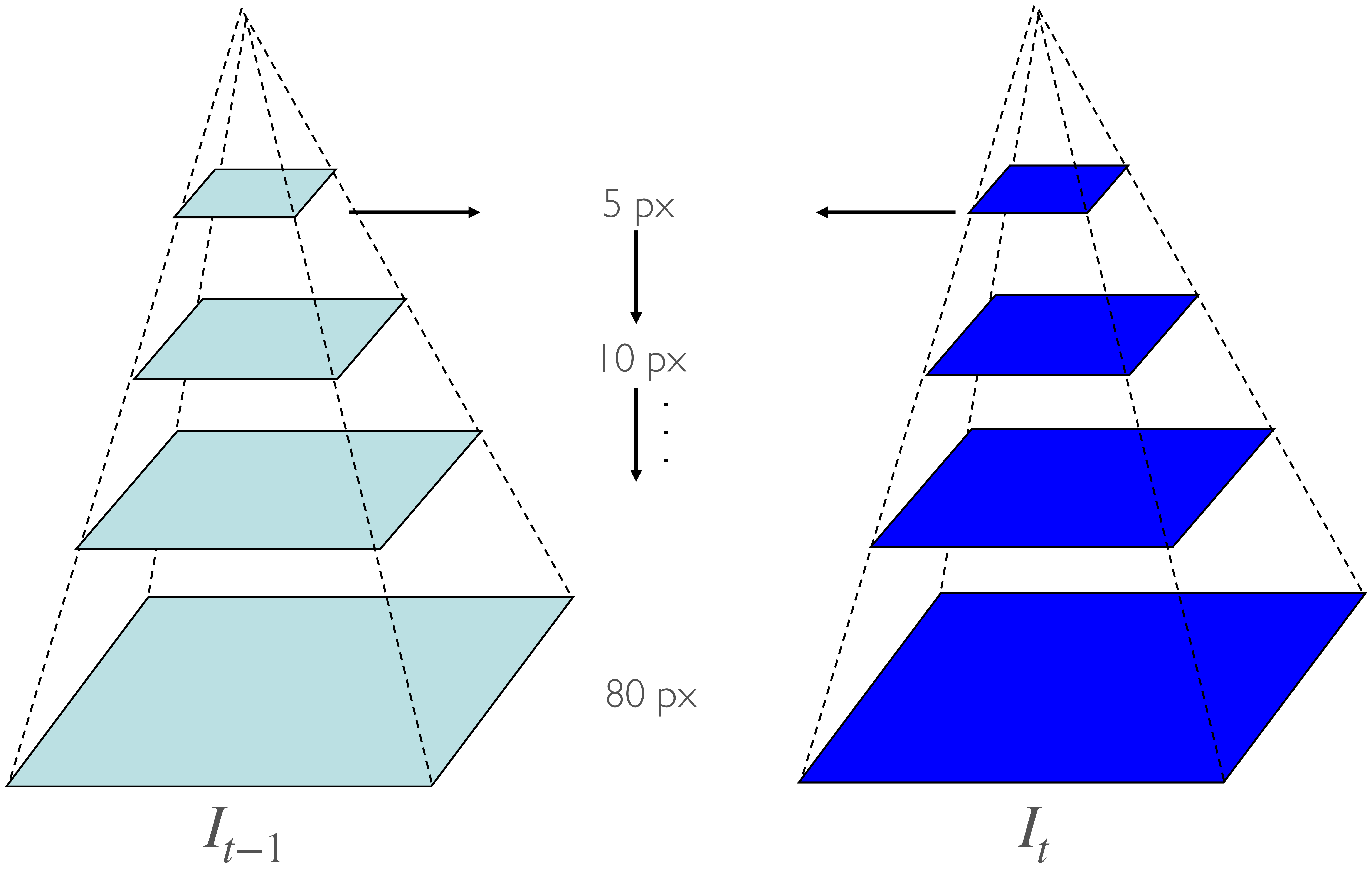
B. Lucas and T. Kanade. An Iterative Image Registration Technique with an Application to Stereo Vision. IJCAI 1981.
B. K. P. Horn and B. G. Schunck. Determining Optical Flow. AI, 1981.

The Next Challenge

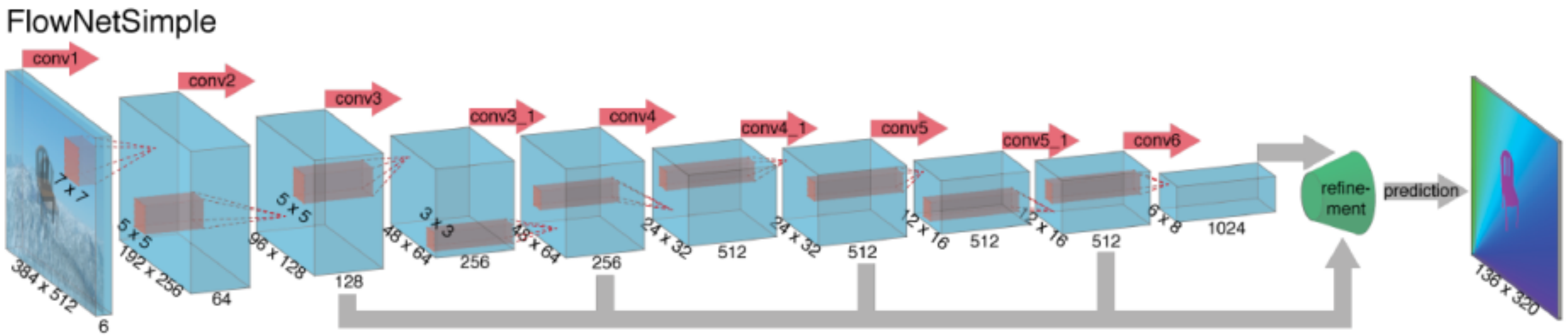
Coarse-to-Fine Strategy



Coarse-to-Fine Strategy

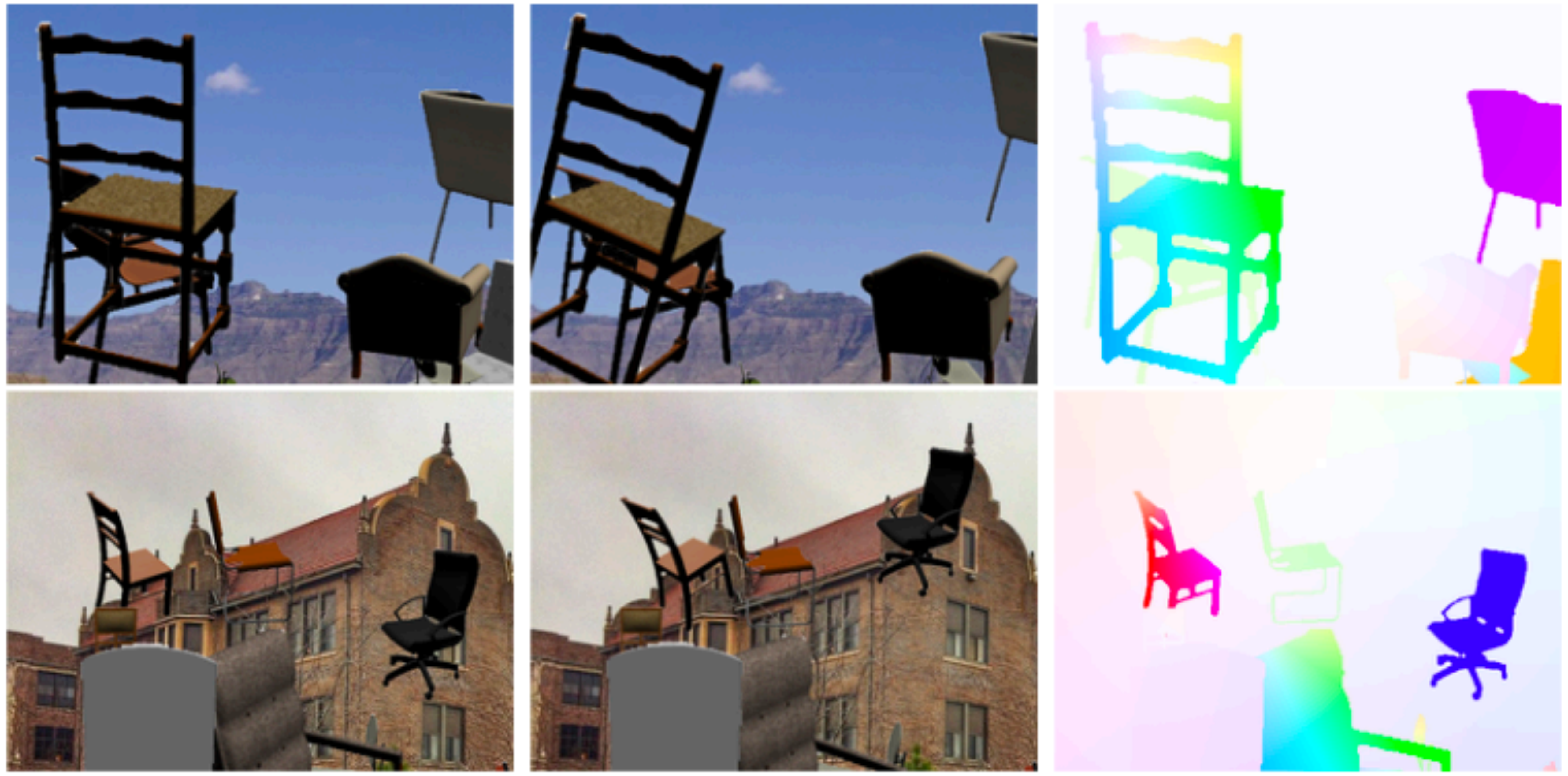


E2E Learning of Optical Flow: FlowNet



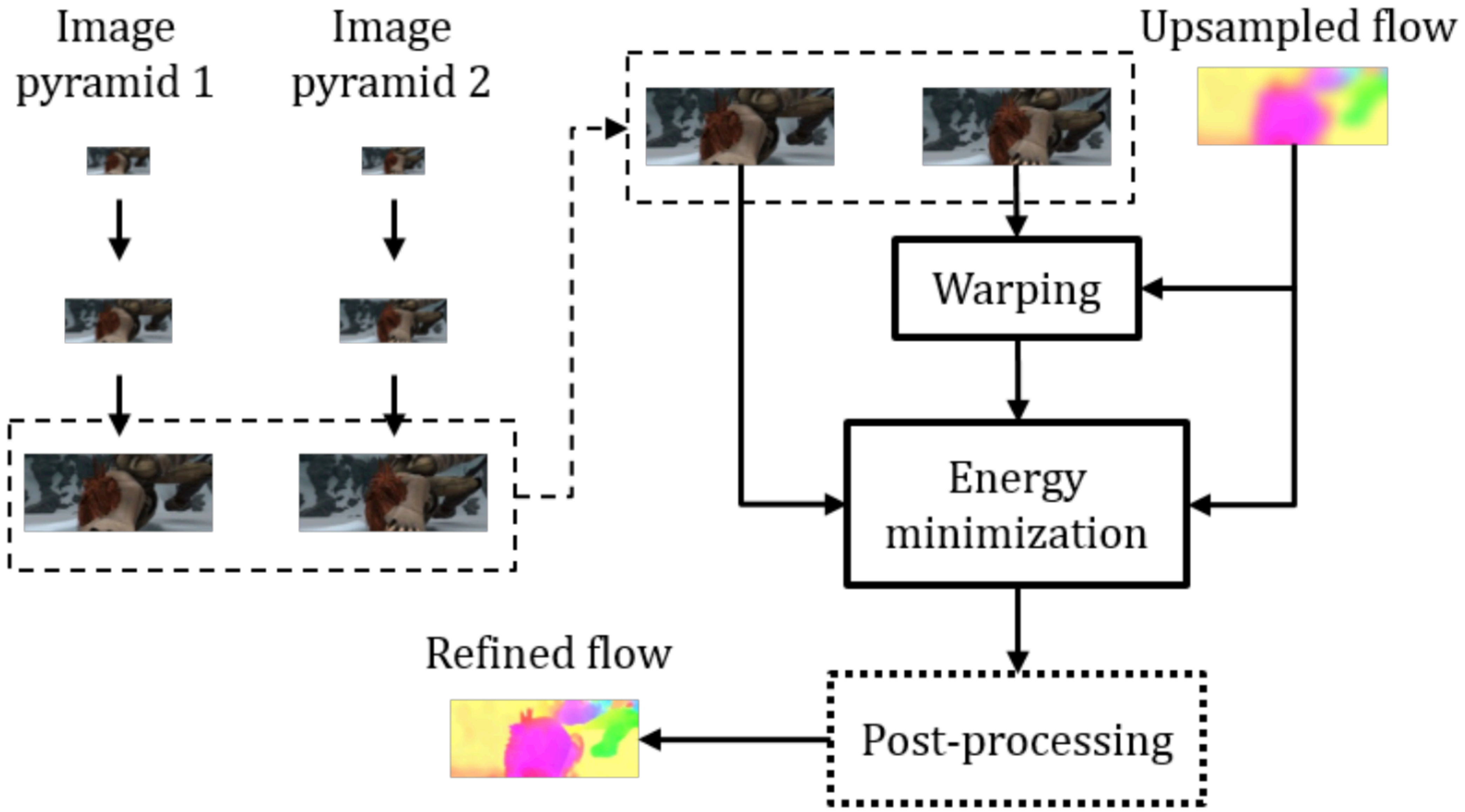
P. Fischer, A. Dosovitskiy, E. Ilg, P. Häusser, C. Hazırbaş, V. Golkov, P. van der Smagt, D. Cremers, and T. Brox. FlowNet: Learning Optical Flow with Convolutional Networks. ICCV 2015.

E2E Learning of Optical Flow: FlyingChairs



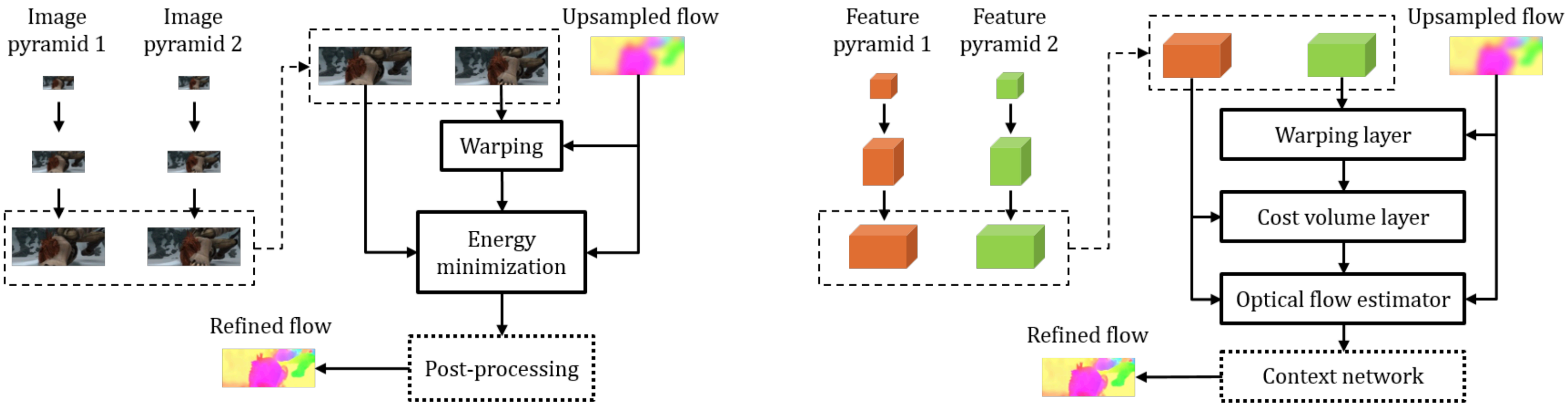
P. Fischer, A. Dosovitskiy, E. Ilg, P. Häusser, C. Hazırbaş, V. Golkov, P. van der Smagt, D. Cremers, and T. Brox. FlowNet: Learning Optical Flow with Convolutional Networks. ICCV 2015.

E2E Learning of Optical Flow: PWC-Net



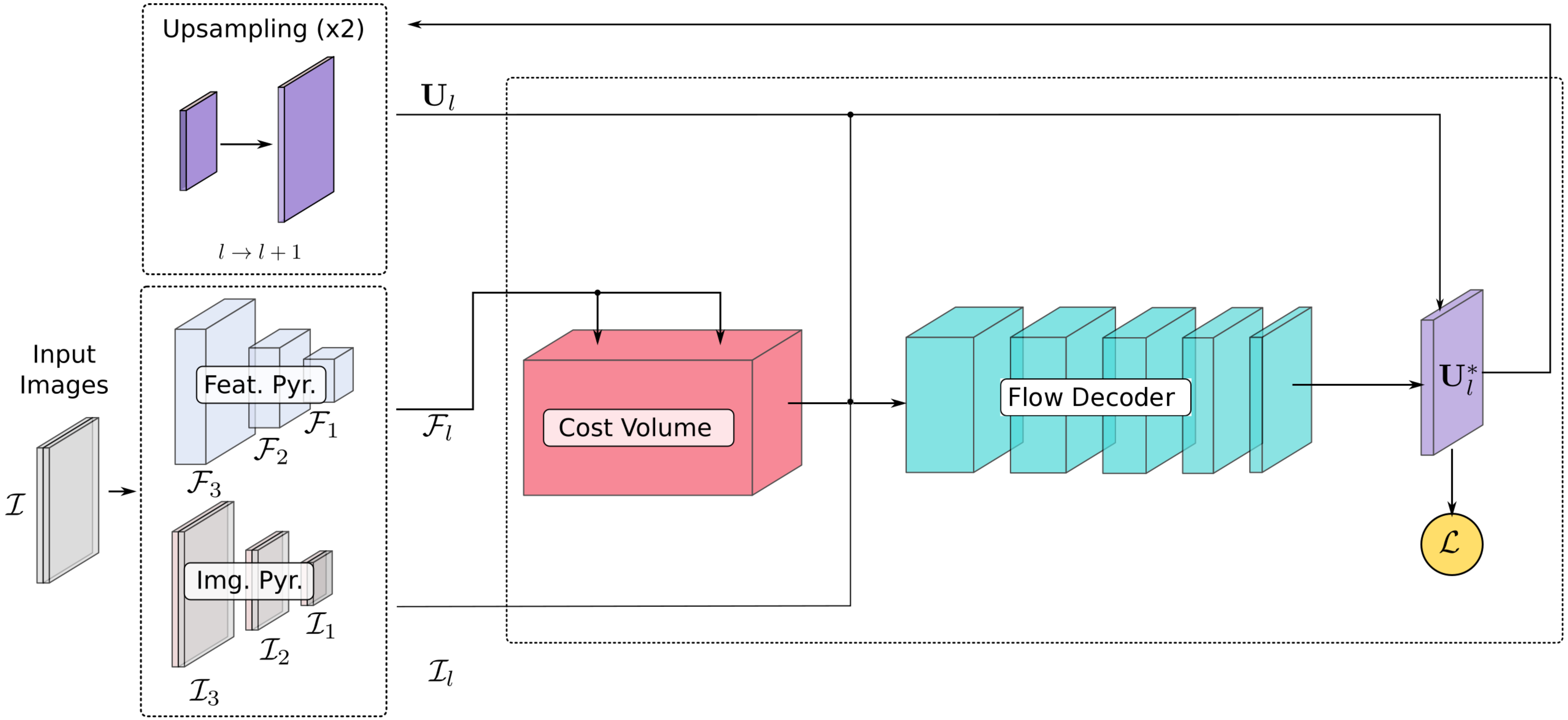
D. Sun, X. Yang, M. Liu, and J. Kautz. PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume. CVPR, 2018.

E2E Learning of Optical Flow: PWC-Net



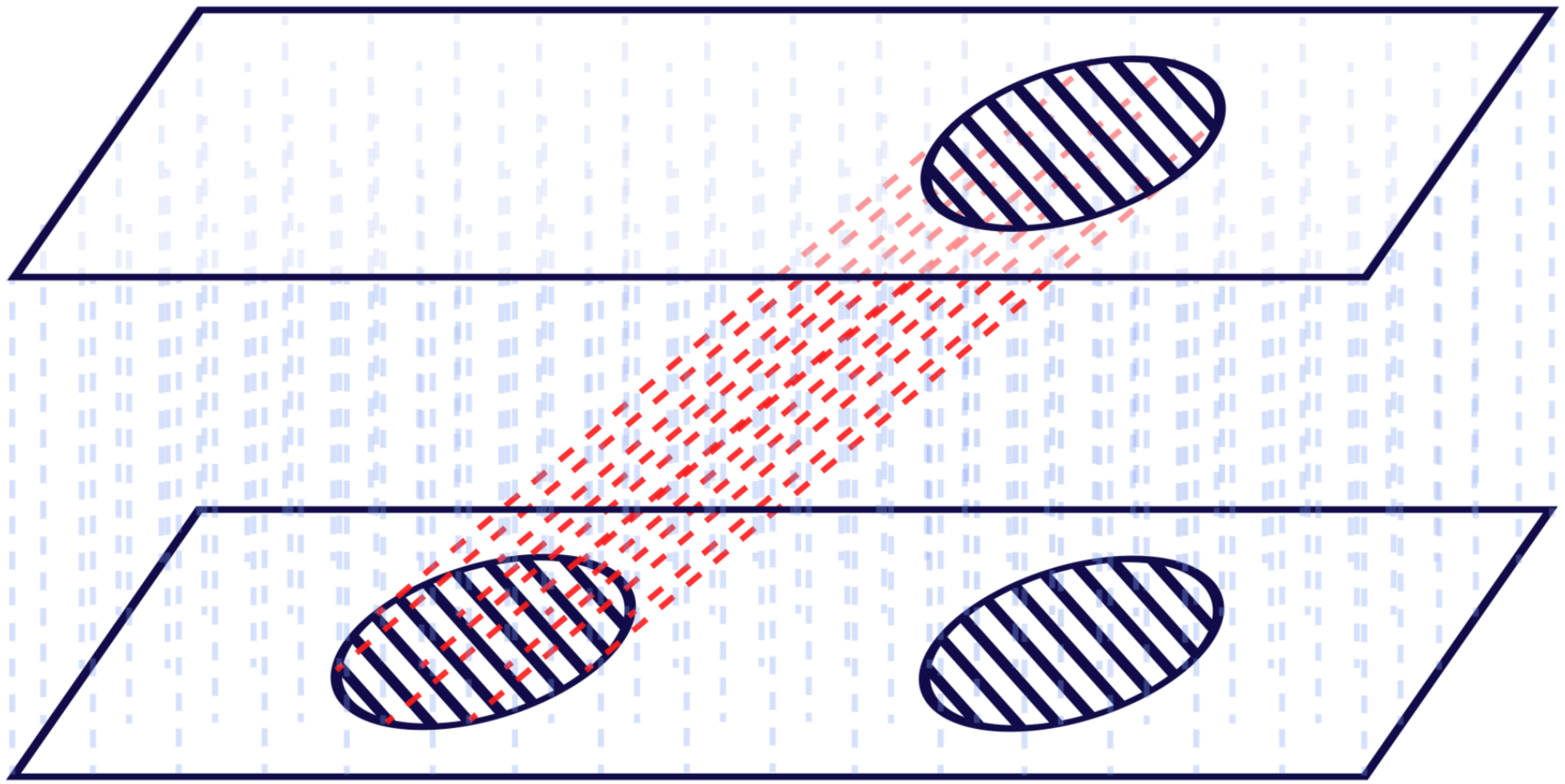
D. Sun, X. Yang, M. Liu, and J. Kautz. PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume. CVPR, 2018.

E2E Learning of Optical Flow: PWC-Net



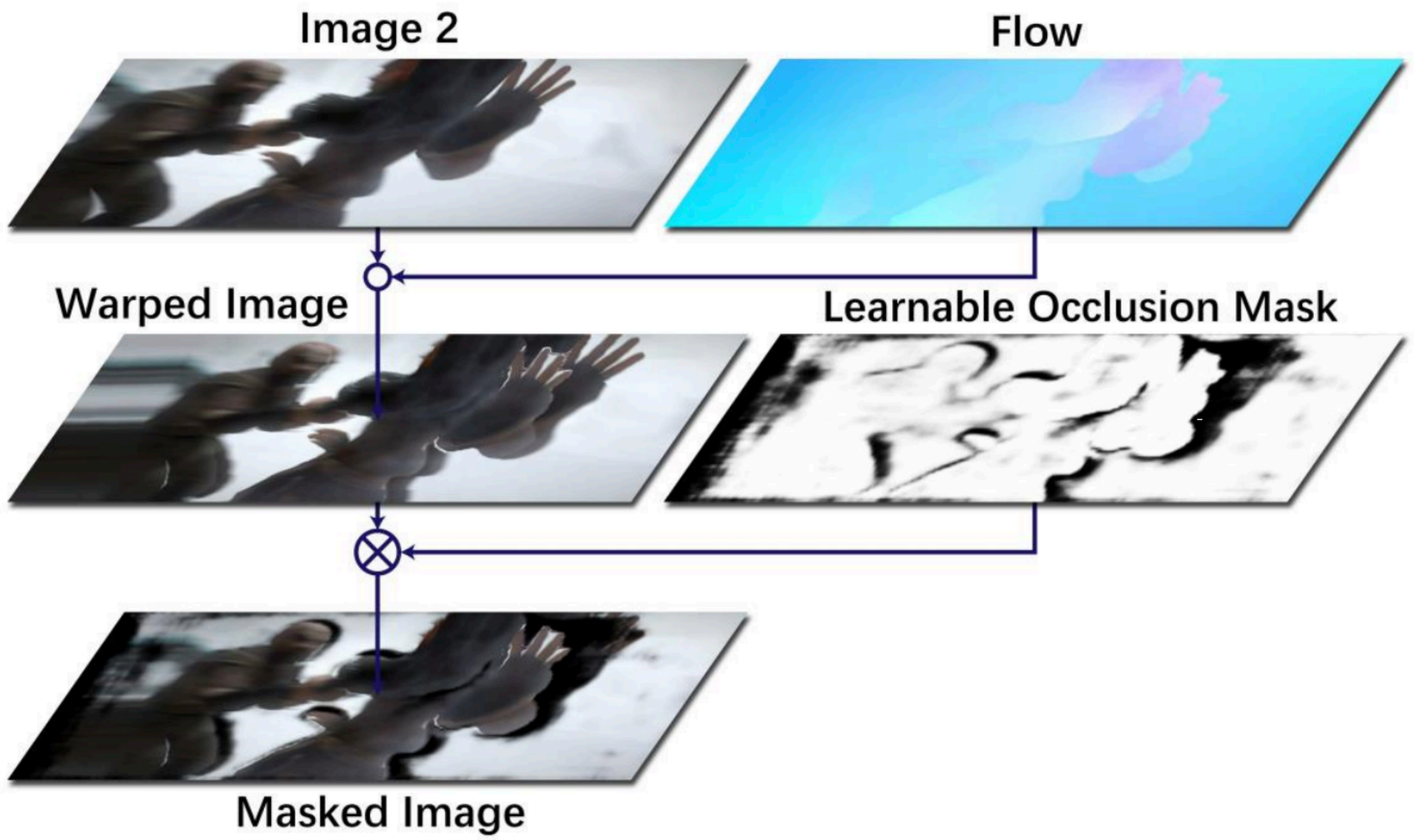
D. Sun, X. Yang, M. Liu, and J. Kautz. PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume. CVPR, 2018.

E2E Learning of Optical Flow: Occlusions



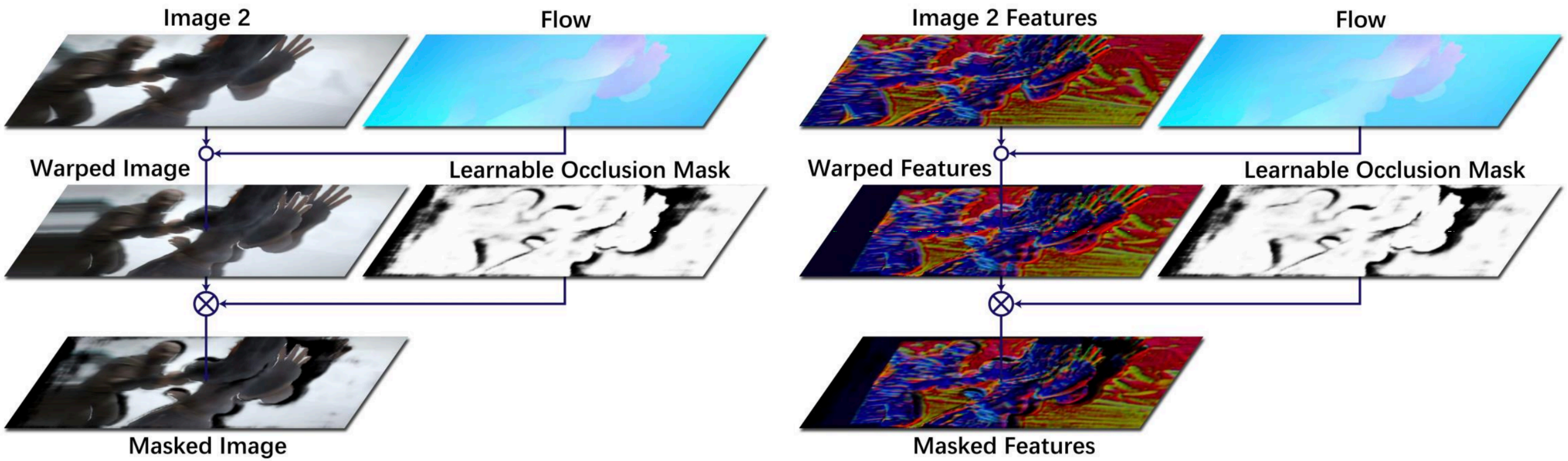
S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

E2E Learning of Optical Flow: MaskFlownet



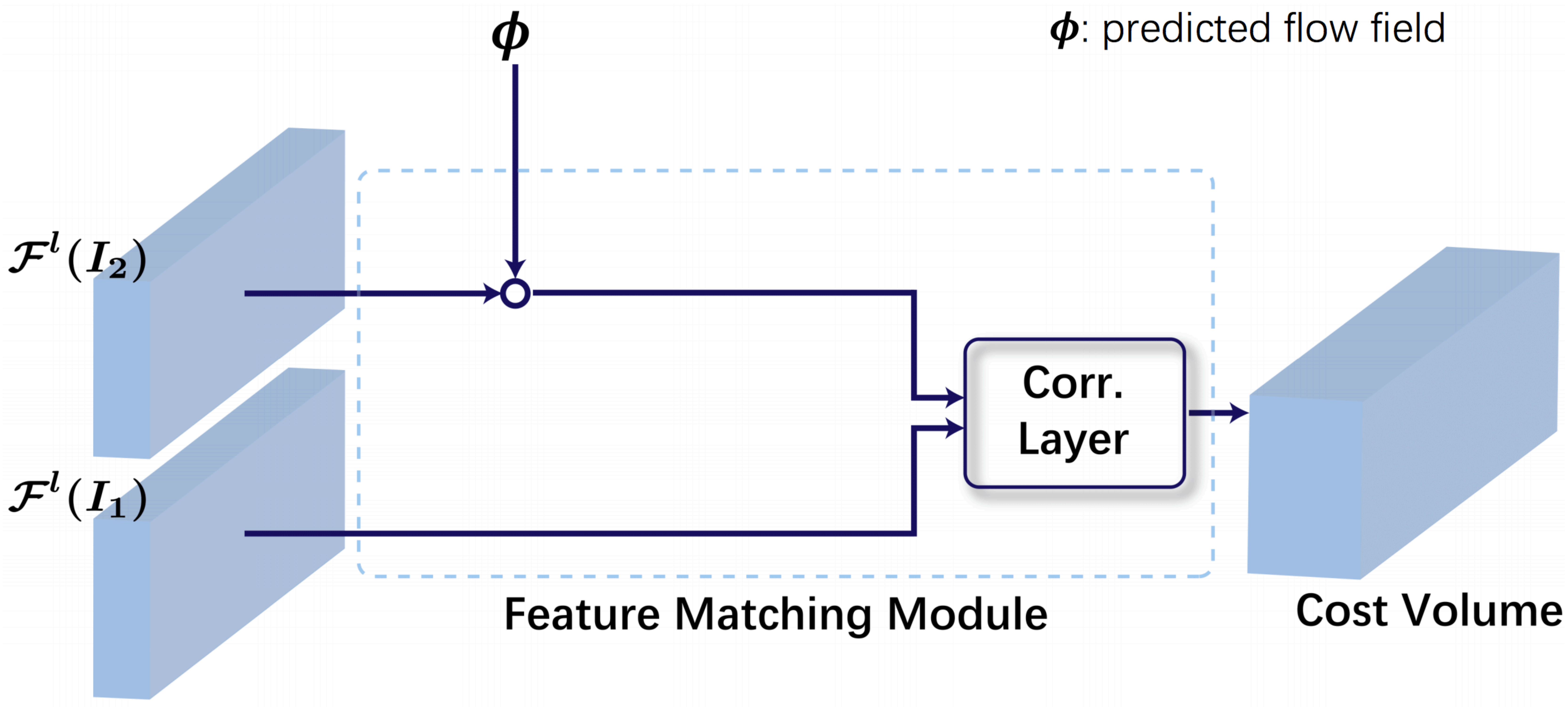
S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

E2E Learning of Optical Flow: MaskFlownet



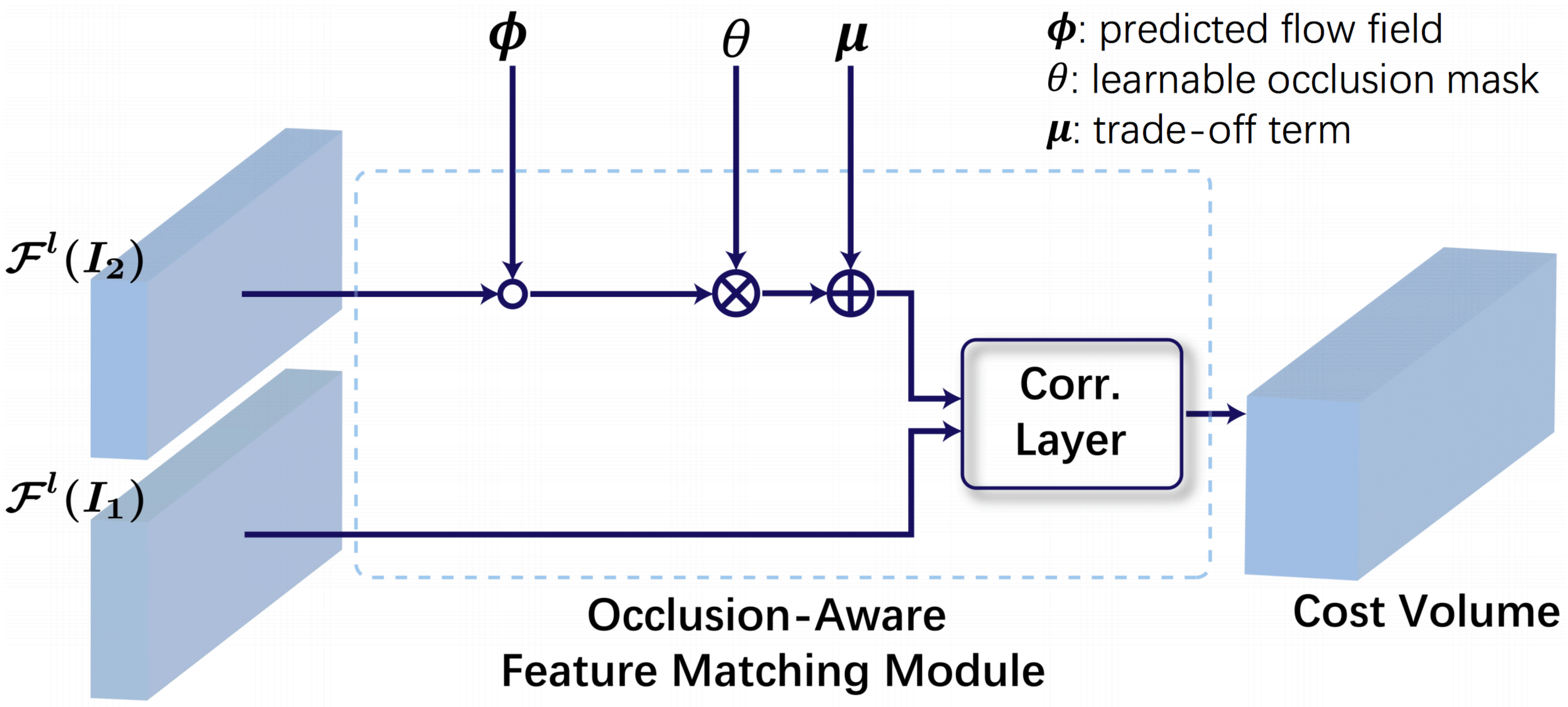
S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

MaskFlownet: Feature Matching



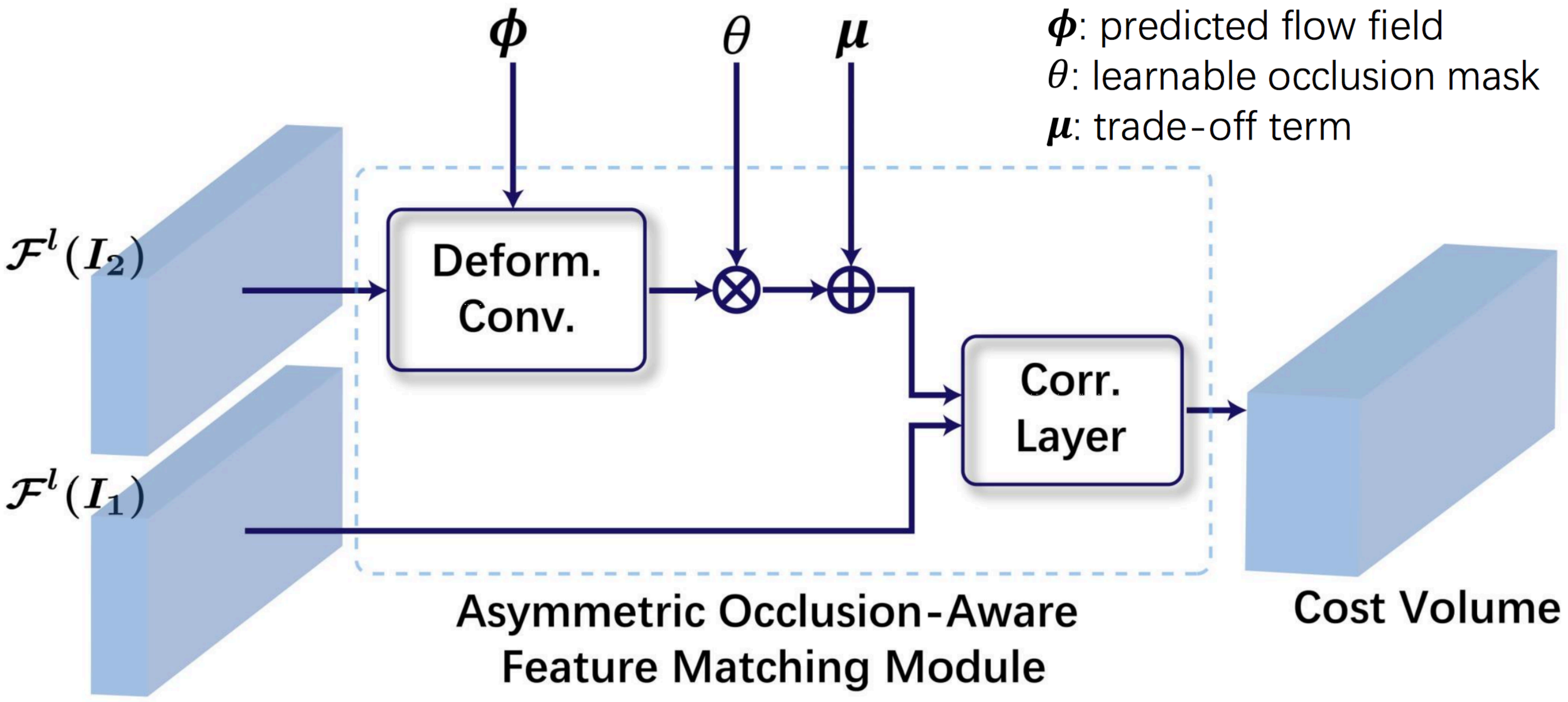
S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

MaskFlownet: Occ-Aware Feature Matching



S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

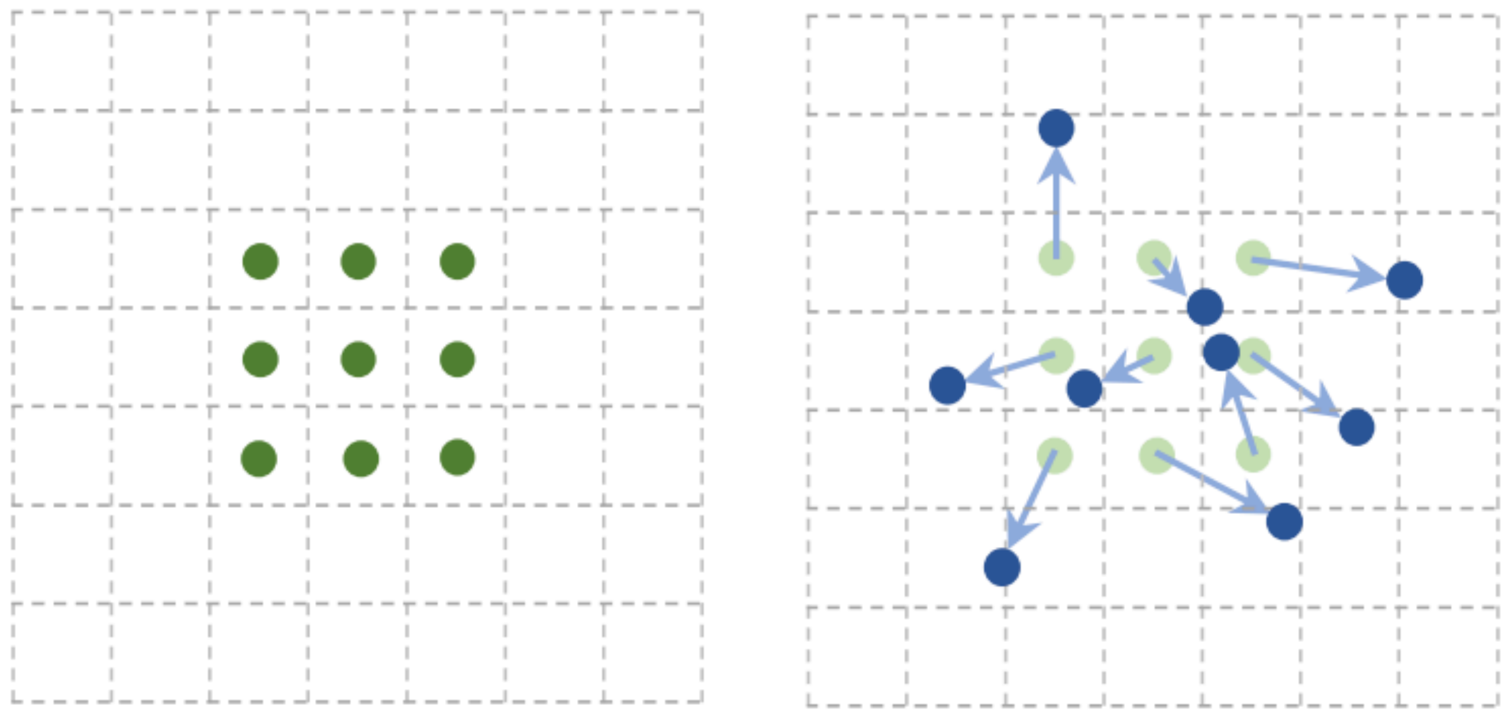
MaskFlownet: Asymmetric Occ-Aware Feature Matching



S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

Deformable Convolution

Dilated convolutions with a ***learned offset***

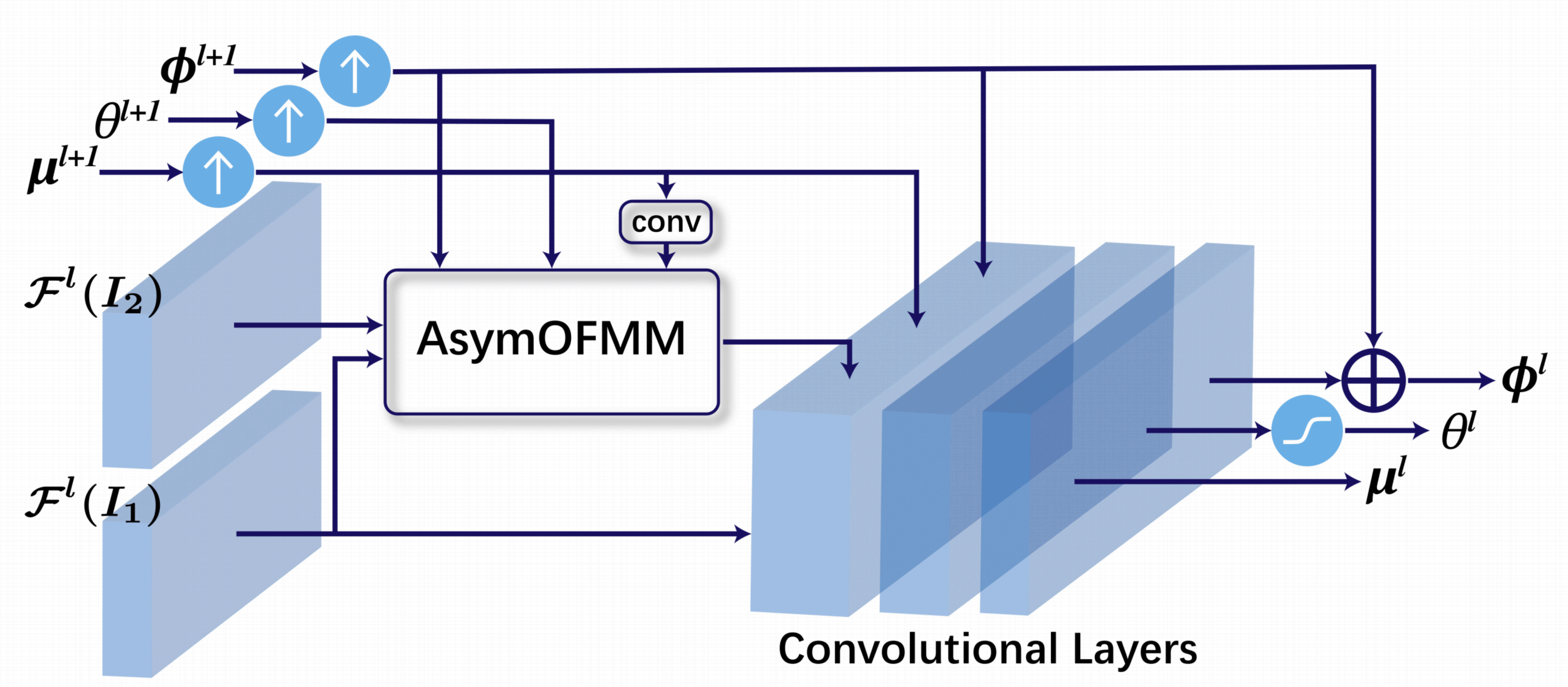


(a) Regular

(b) Deformable

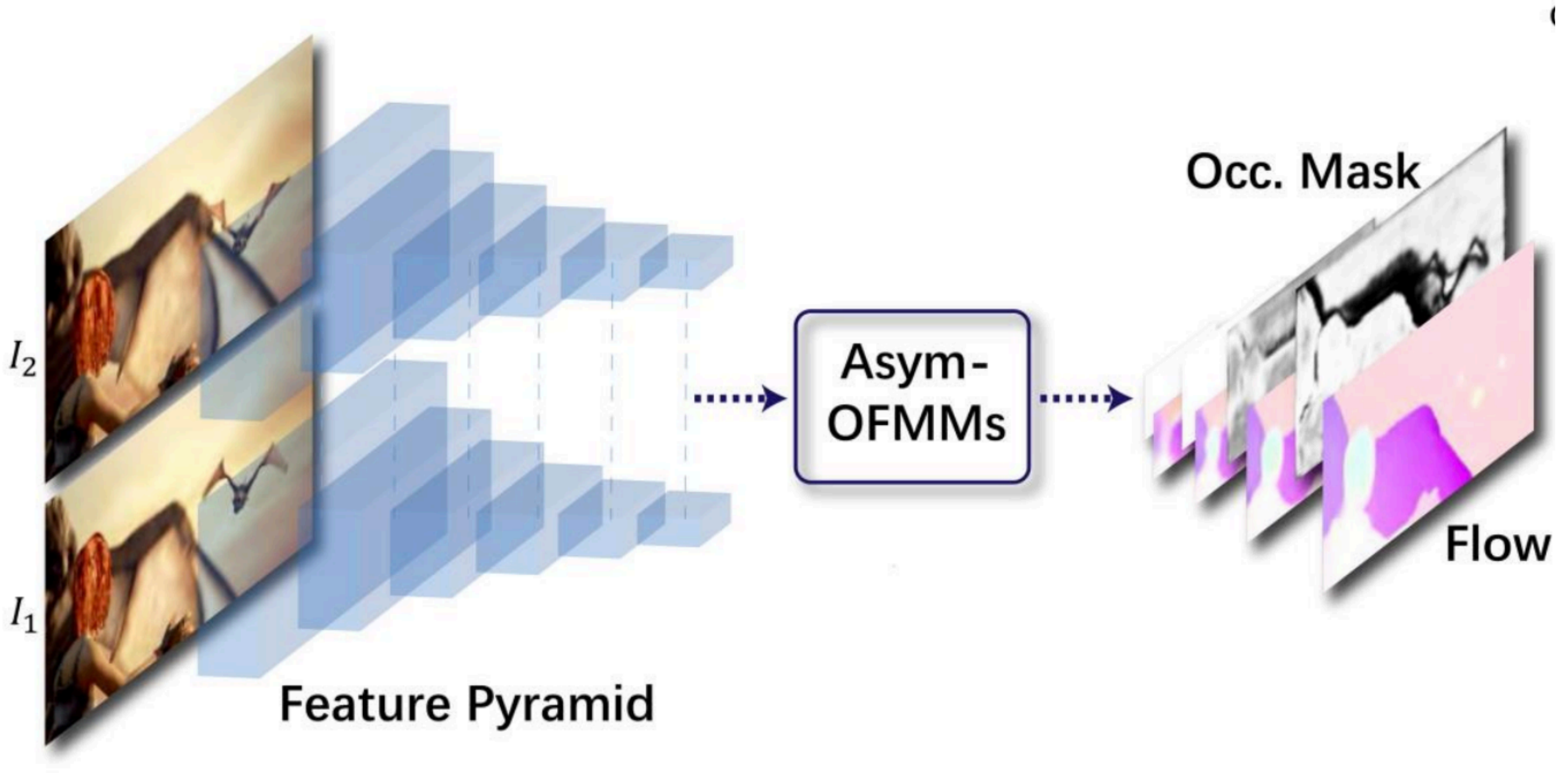
J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, and Y. Wei. Deformable Convolutional Networks. ICCV, 2017.

MaskFlownet: Network Connections



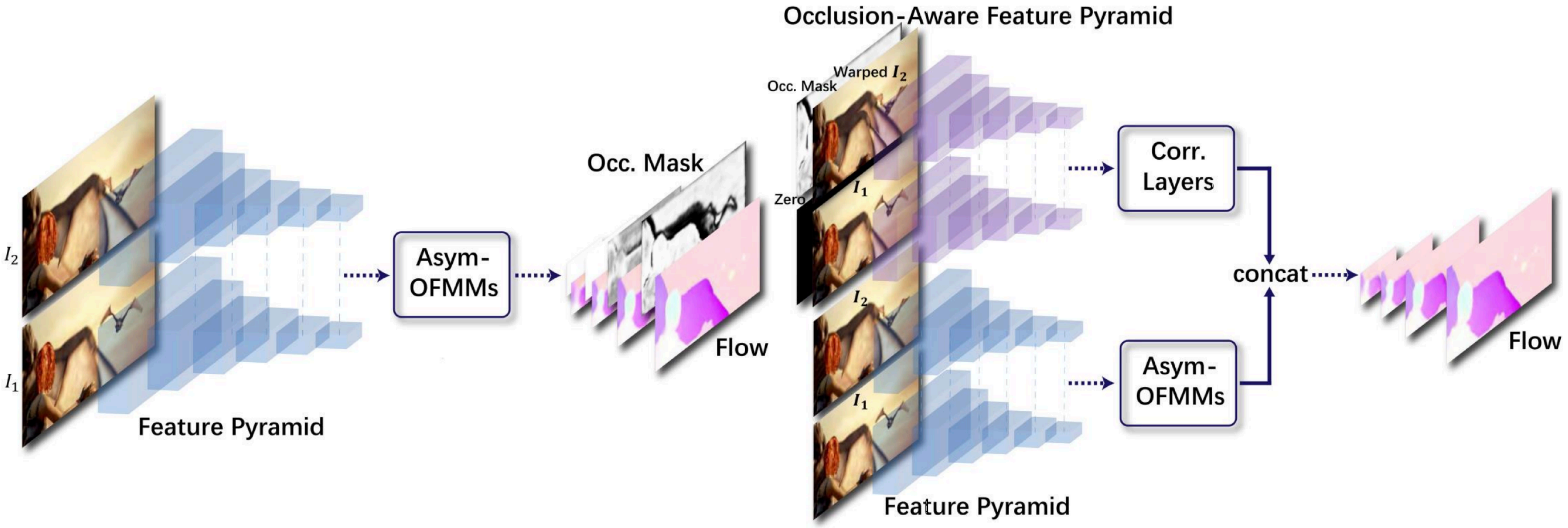
S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

E2E Learning of Optical Flow: MaskFlownet



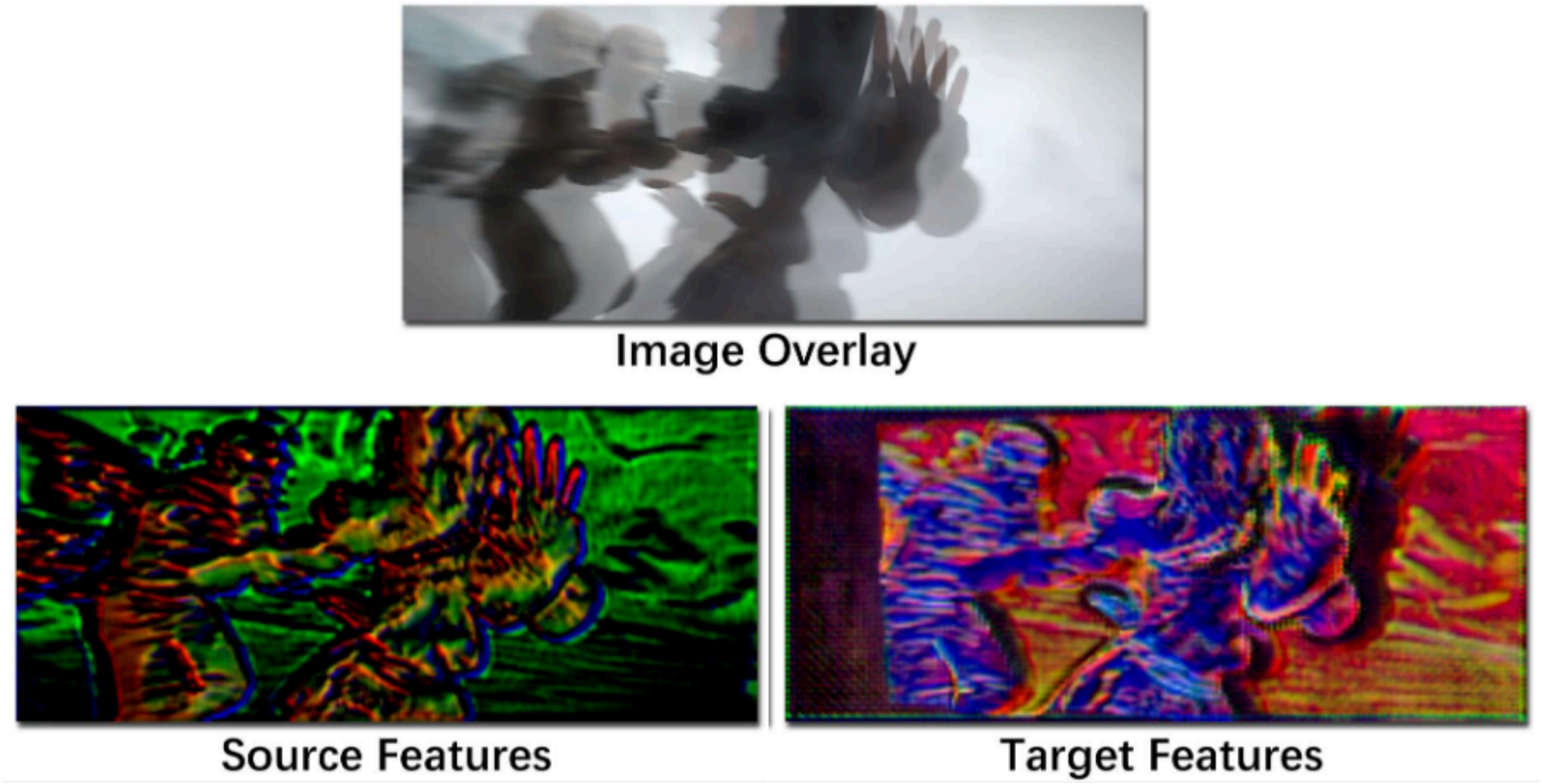
S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

E2E Learning of Optical Flow: MaskFlownet



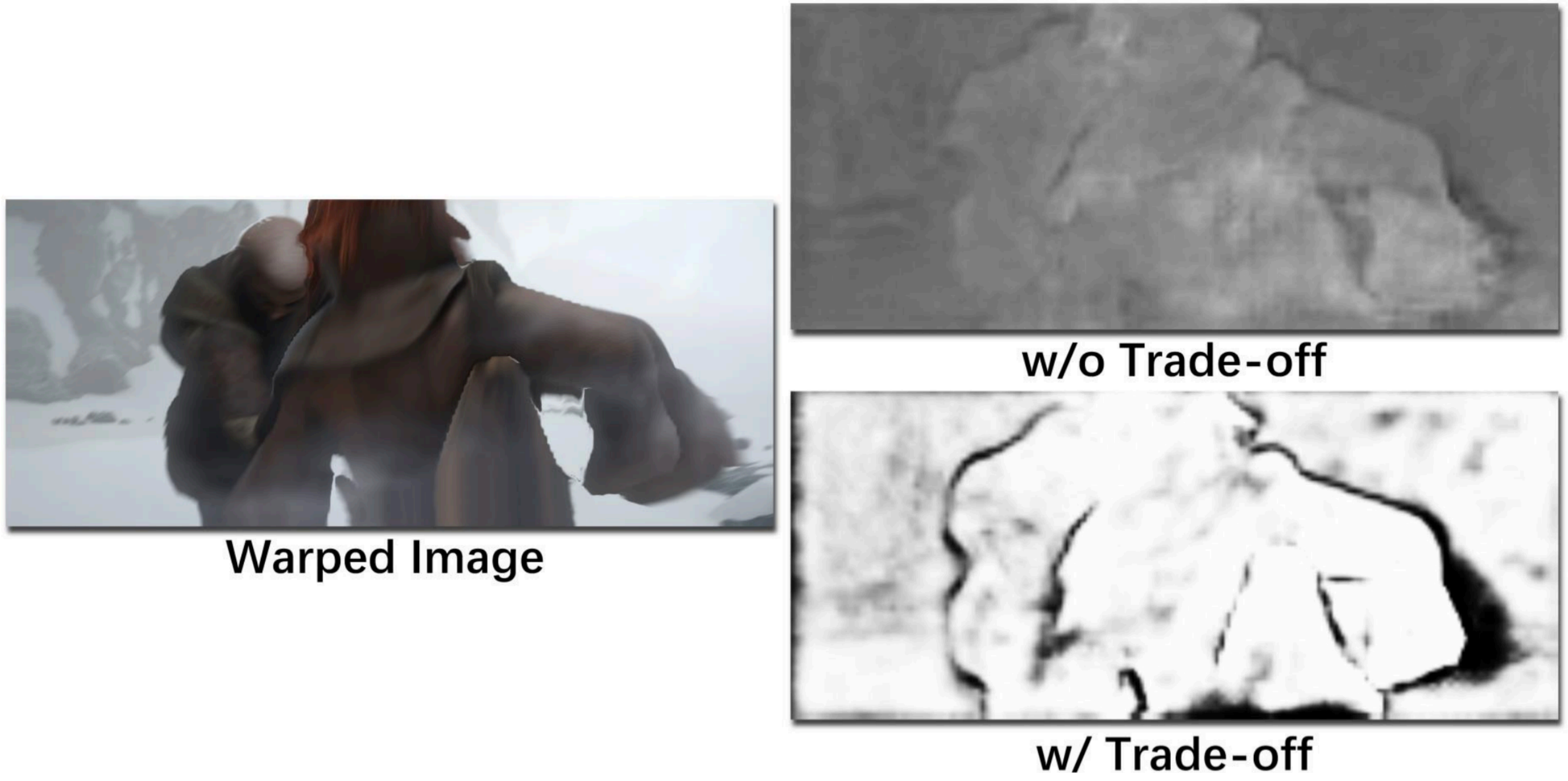
S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

E2E Learning of Optical Flow: MaskFlownet



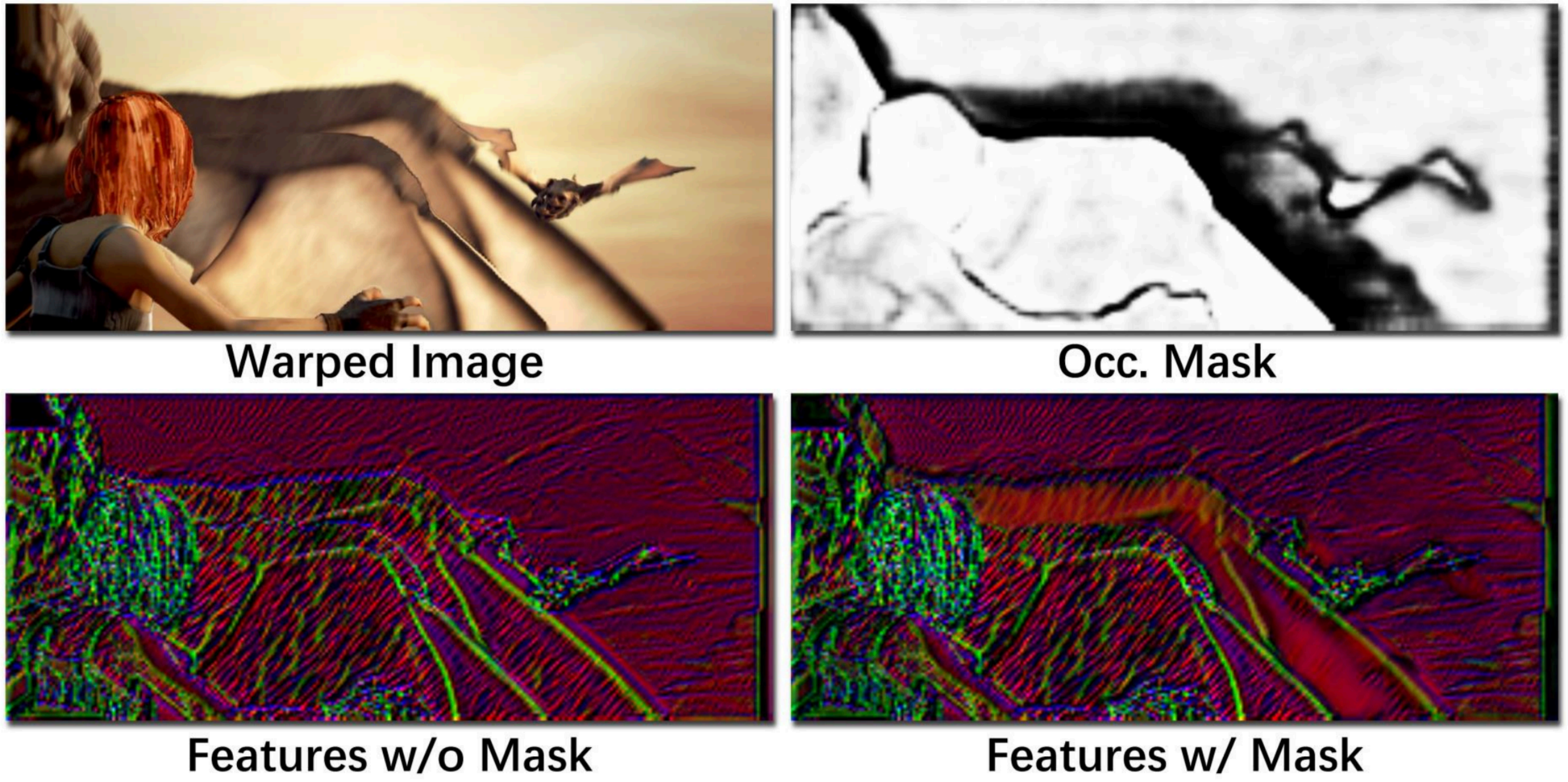
S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

E2E Learning of Optical Flow: MaskFlownet



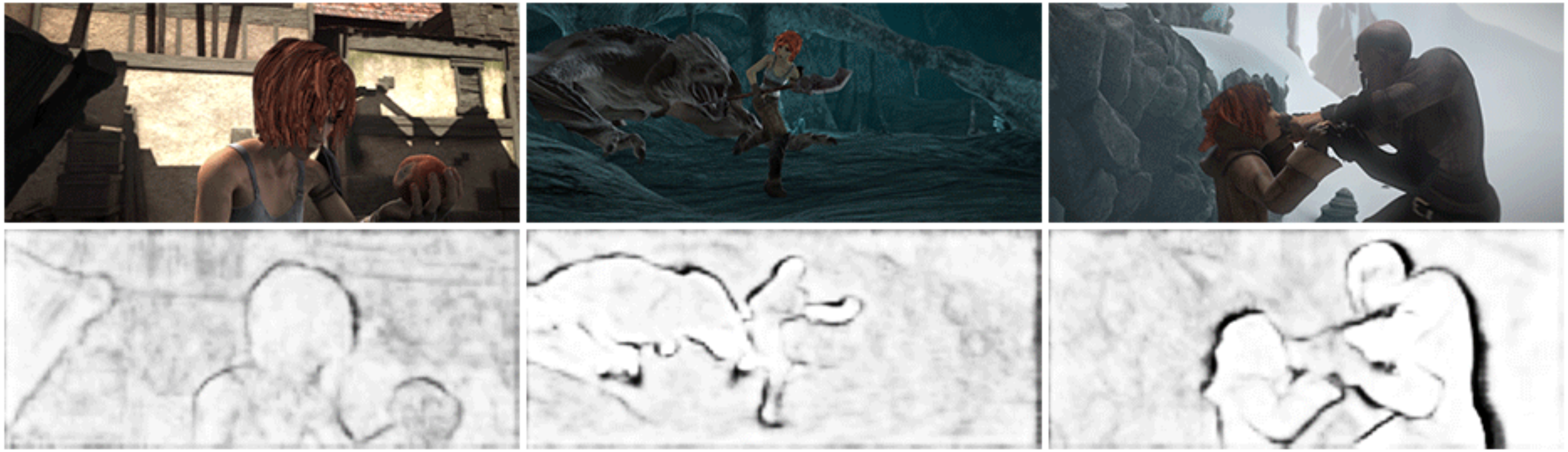
S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

E2E Learning of Optical Flow: MaskFlownet



S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

E2E Learning of Optical Flow: MaskFlownet

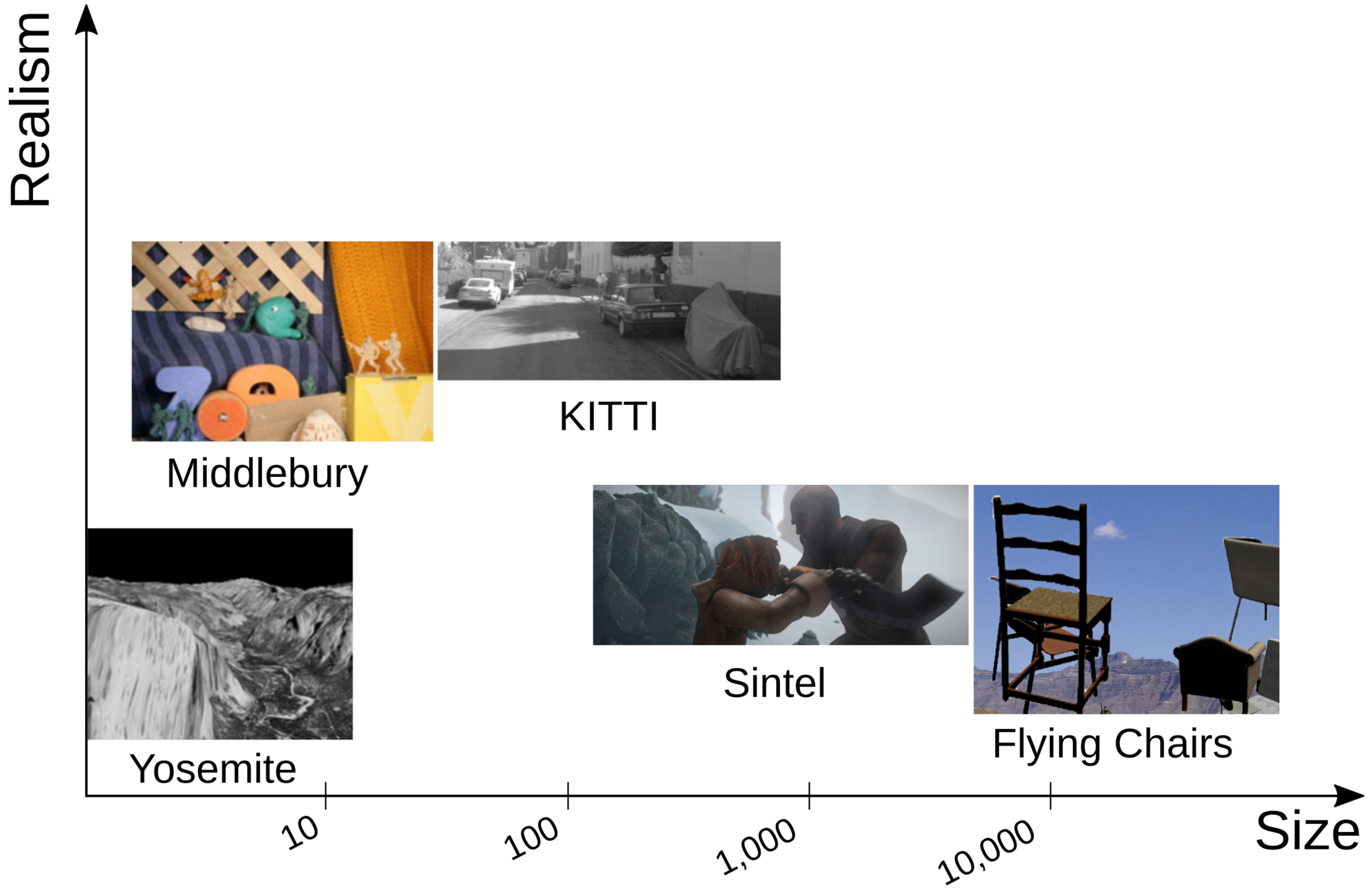


S. Zhao, Y. Sheng, Y. Dong, E. I-C. Chang, and Y. Xu. MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask. CVPR, 2020.

The Problem of Data

- Learning Optical Flow requires large datasets.
- Obtaining real-world annotated datasets is hard.

The Problem of Data



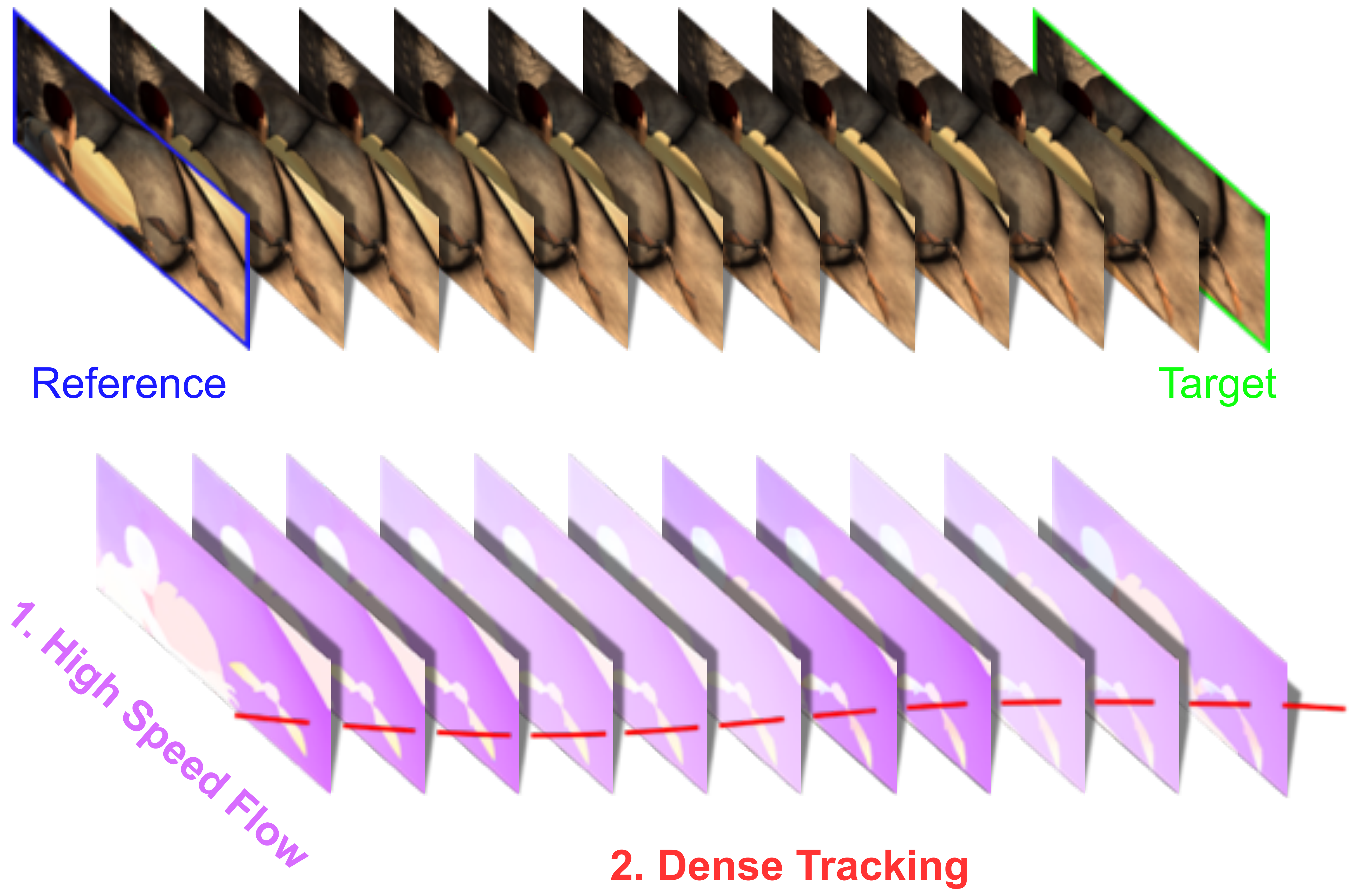
Real-World Optical Flow

J. Janai, F. Guney, J. Wulff, M. Black, and A. Geiger. Slow Flow: Exploiting High-Speed Cameras for Accurate and Diverse Optical Flow Reference Data. CVPR, 2017.

Real-World Optical Flow

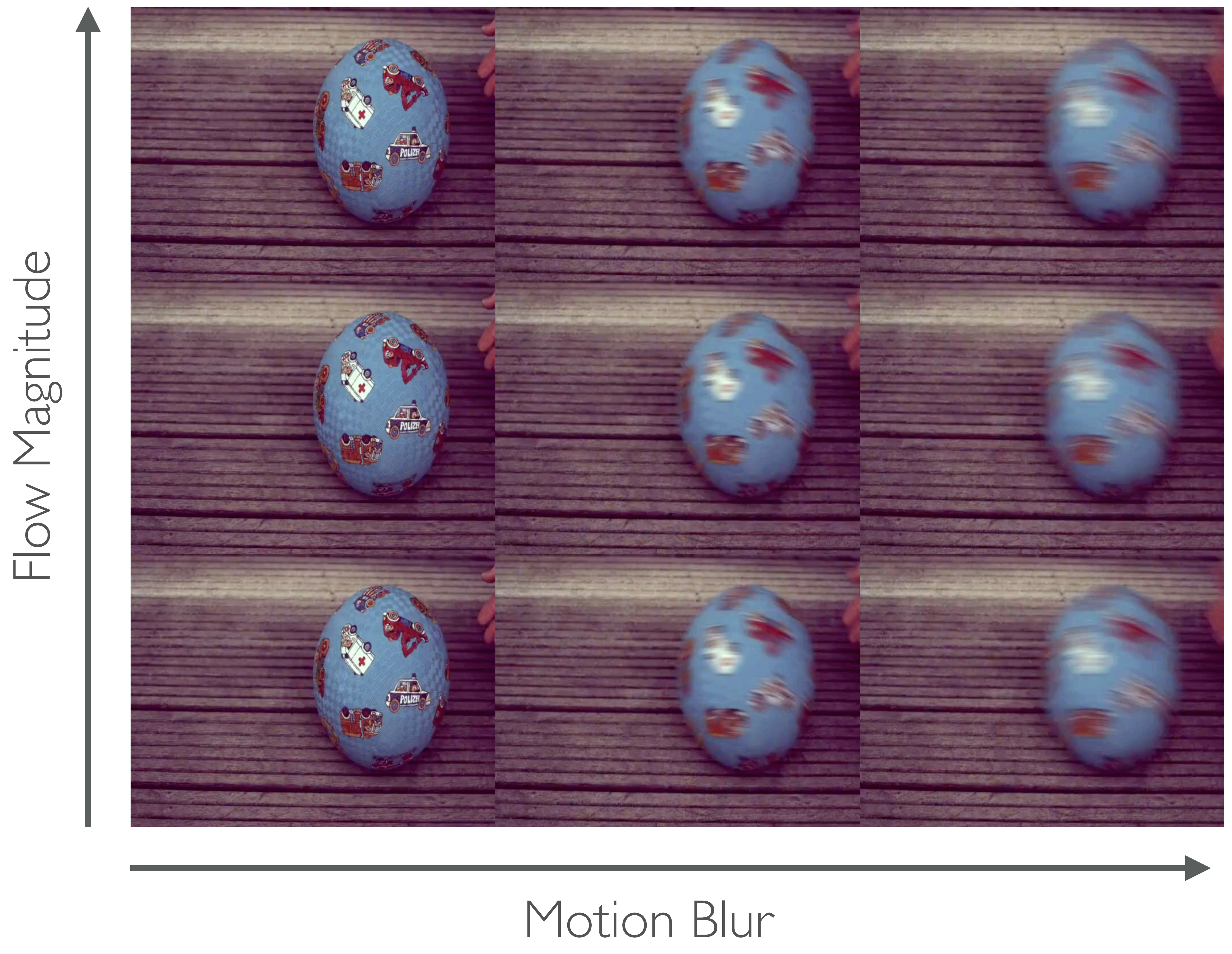
J. Janai, F. Guney, J. Wulff, M. Black, and A. Geiger. Slow Flow: Exploiting High-Speed Cameras for Accurate and Diverse Optical Flow Reference Data. CVPR, 2017.

Exploiting High-Frame Rate for Real-World Optical Flow



J. Janai, F. Guney, J. Wulff, M. Black, and A. Geiger. Slow Flow: Exploiting High-Speed Cameras for Accurate and Diverse Optical Flow Reference Data. CVPR, 2017.

A Systematic Evaluation



J. Janai, F. Guney, J. Wulff, M. Black, and A. Geiger. Slow Flow: Exploiting High-Speed Cameras for Accurate and Diverse Optical Flow Reference Data. CVPR, 2017.

The Problem of Data

- Learning Optical Flow requires large datasets.

- Obtaining real-world annotated datasets is hard.

The Problem of Data

- Learning Optical Flow requires large datasets.
- Obtaining real-world annotated datasets is hard.
- Unsupervised learning allows to use unlabelled data.

Unsupervised Learning

Photometric Loss

- Estimate flow

- Warp target frame
  according to estimated flow

Reference

Target

Unsupervised Learning

Photometric Loss

- Estimate flow
- Warp target frame according to estimated flow
- Minimize distance between warped and reference frames

Reference

Target

$$\sum_{\mathbf{p}\in\Omega}\left(\hat{\mathbf{I}}\left(\mathbf{p}+\mathbf{u}\left(\mathbf{p}\right)\right)-\mathbf{I}_R\left(\mathbf{p}\right)\right)^2$$

Domain of Reference

Warped Target

Reference

Unsupervised Learning

Photometric Loss

- Estimate flow
- Warp target frame according to estimated flow
- Minimize distance between warped and reference frames

Reference

Target

Robust function

$$\sum_{\mathbf{p}\in\Omega} \delta\left(\hat{\mathbf{I}}\left(\mathbf{p}+\mathbf{u}\left(\mathbf{p}\right)\right), \mathbf{I}_R\left(\mathbf{p}\right)\right)$$

Domain of Reference

Warped Target

Reference

A Baseline

Take the SOTA method and train unsupervised.

Ground-truth

Unsupervised
PWC-Net

D. Sun, X. Yang, M. Liu, and J. Kautz. PWC-Net: CNNs for Optical Flow Using Pyramid, Warping, and Cost Volume. CVPR, 2018.

Unsupervised Learning

Photometric Loss

- Estimate flow
- Warp target frame according to estimated flow
- Minimize distance between warped and reference frames

Reference

Target

Warped Target

Unsupervised Learning

Photometric Loss

- Estimate flow
- Warp target frame according to estimated flow
- Minimize distance between warped and reference frames

Reference

Target

Warped Target
with *Ghosting Effect*

The Problem of Data

- Learning Optical Flow requires large datasets.
- Obtaining real-world annotated datasets is hard.
- Unsupervised learning allows to use unlabelled data.
- Photometric loss is misleading due to *Ghosting Effect*.

Our Solution

Exploit multiple frames

Ground-truth

Unsupervised
PWC-Net

Unsupervised
Multi PWC-Net

J. Janai, F. Guney, A. Ranjan, M. Black, A. Geiger. Unsupervised Learning of Multi-Frame Optical Flow with Occlusions. ECCV, 2018.

Our Solution

Exploit multiple frames and model occlusions

Ground-truth

Unsupervised
PWC-Net

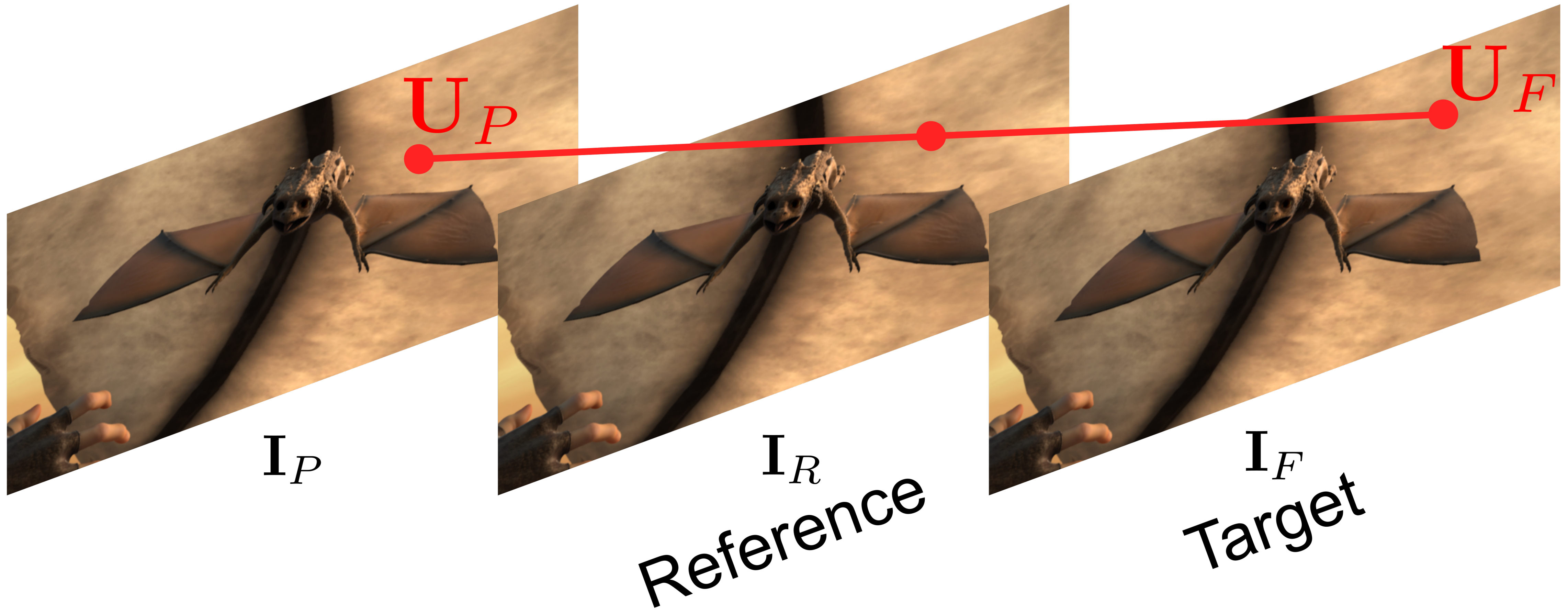
Unsupervised
Multi PWC-Net

Unsupervised
Multi PWC-Net
with Occlusions

J. Janai, F. Guney, A. Ranjan, M. Black, A. Geiger. Unsupervised Learning of Multi-Frame Optical Flow with Occlusions. ECCV, 2018.

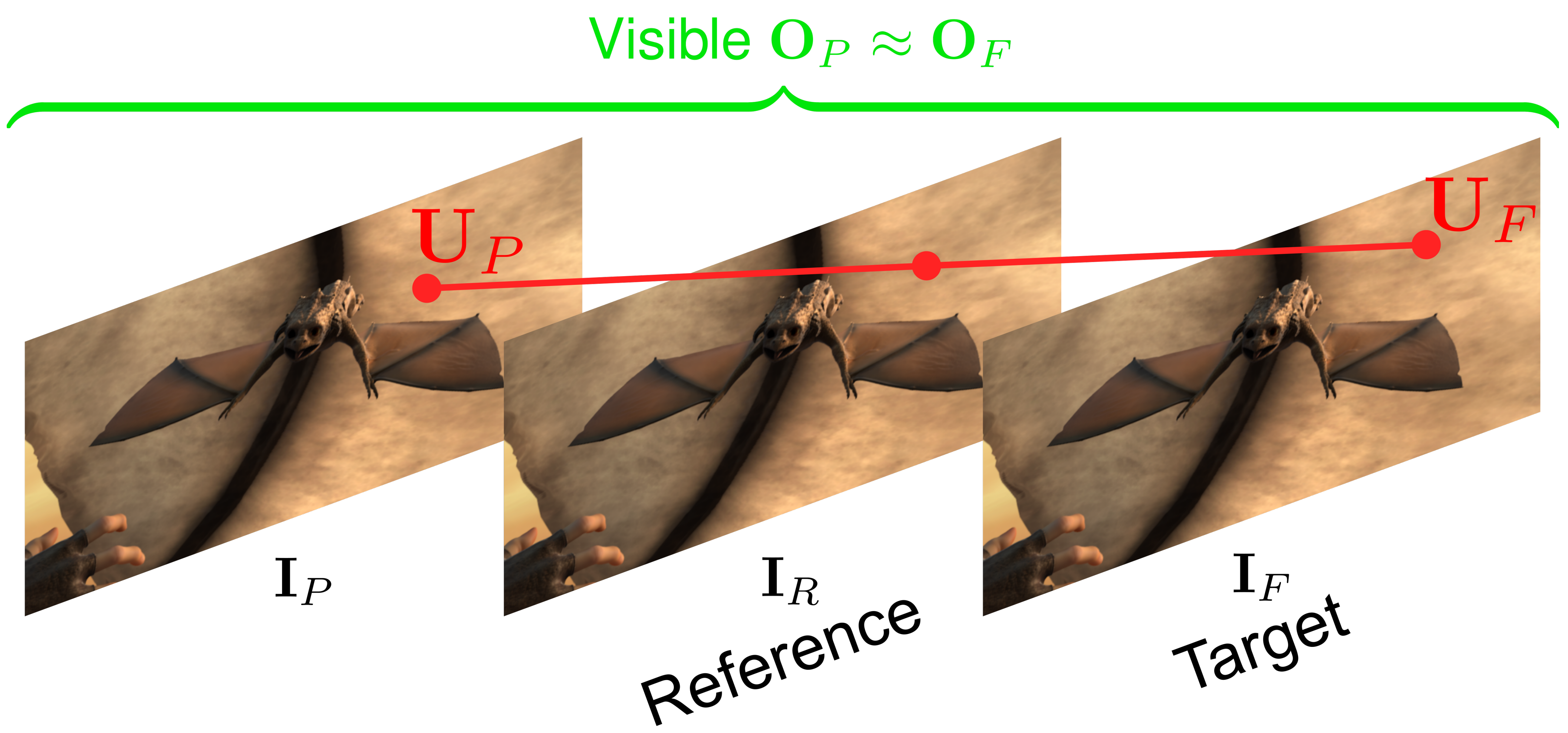
Formulation

- In a temporal window
  - Flow $\mathbf{U}_P, \mathbf{U}_F \in \mathbb{R}^{W \times H \times 2}$
  - Occlusions $\mathbf{O} \in [0,1]^{W \times H \times 2}$



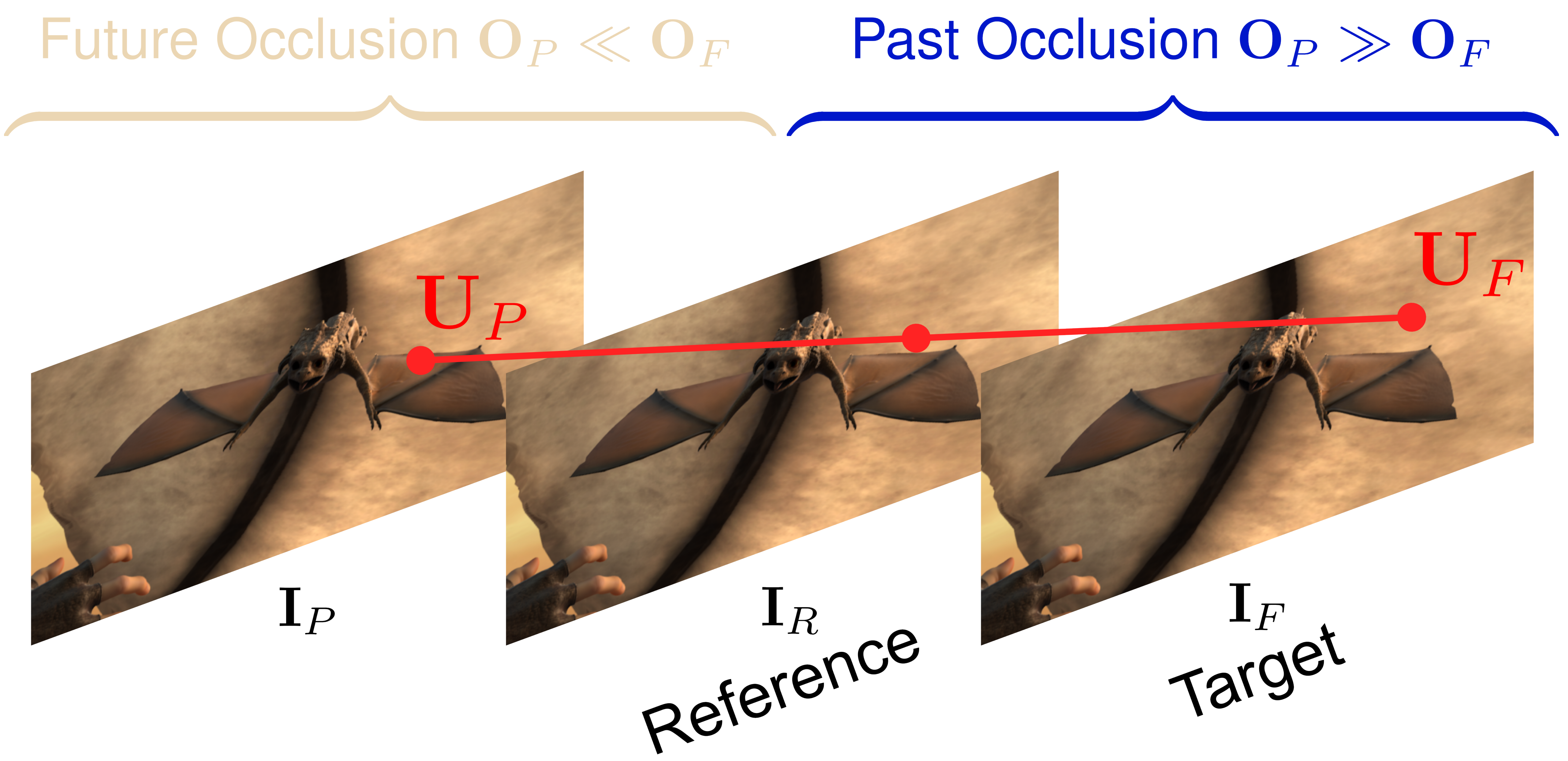
Formulation: Visible

- In a temporal window
  - Flow $\mathbf{U}_P, \mathbf{U}_F \in \mathbb{R}^{W \times H \times 2}$
  - Occlusions $\mathbf{O} \in [0,1]^{W \times H \times 2}$



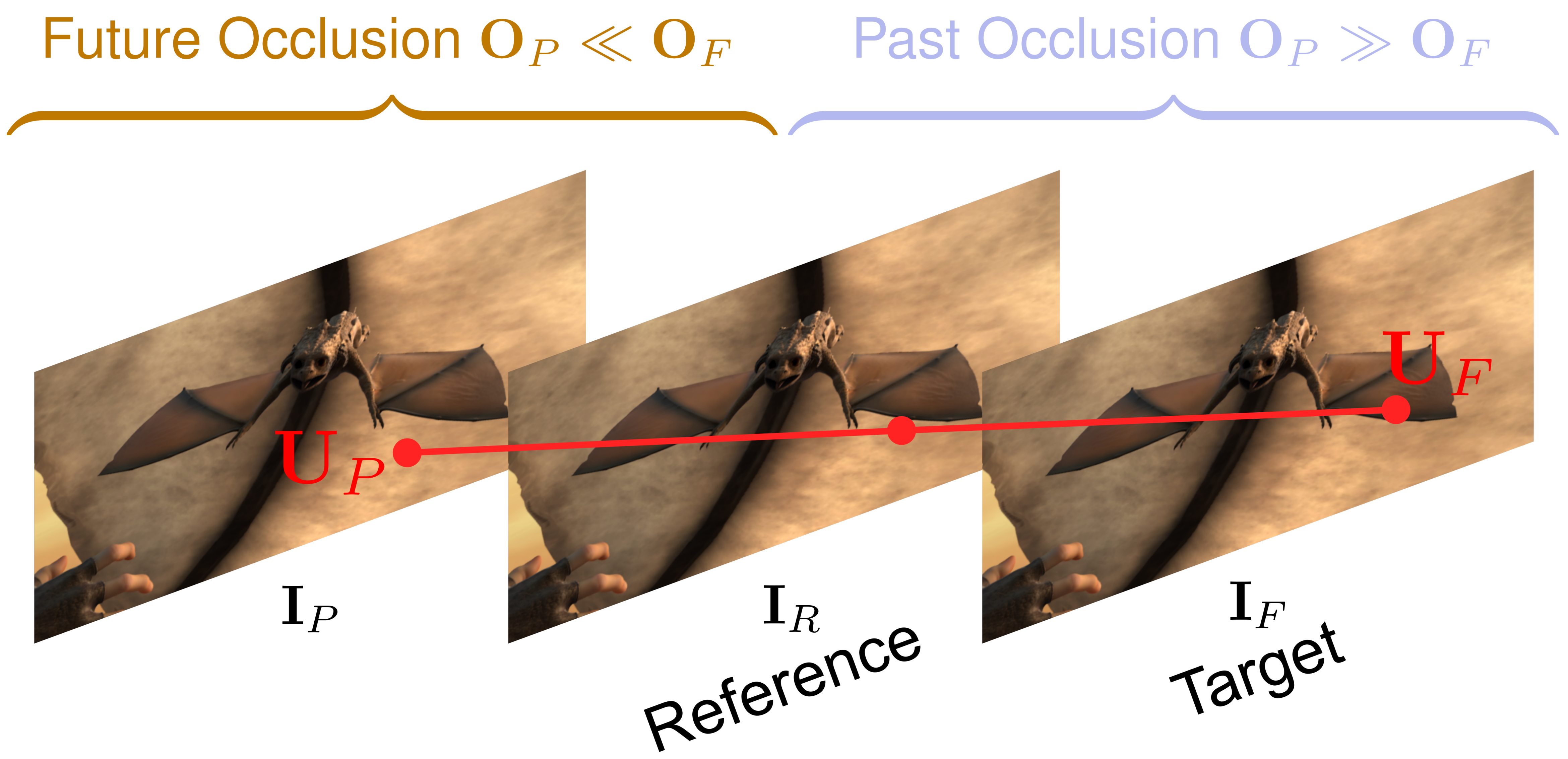
Formulation: Past Occlusion

- In a temporal window
  - Flow $\mathbf{U}_P, \mathbf{U}_F \in \mathbb{R}^{W \times H \times 2}$
  - Occlusions $\mathbf{O} \in [0,1]^{W \times H \times 2}$



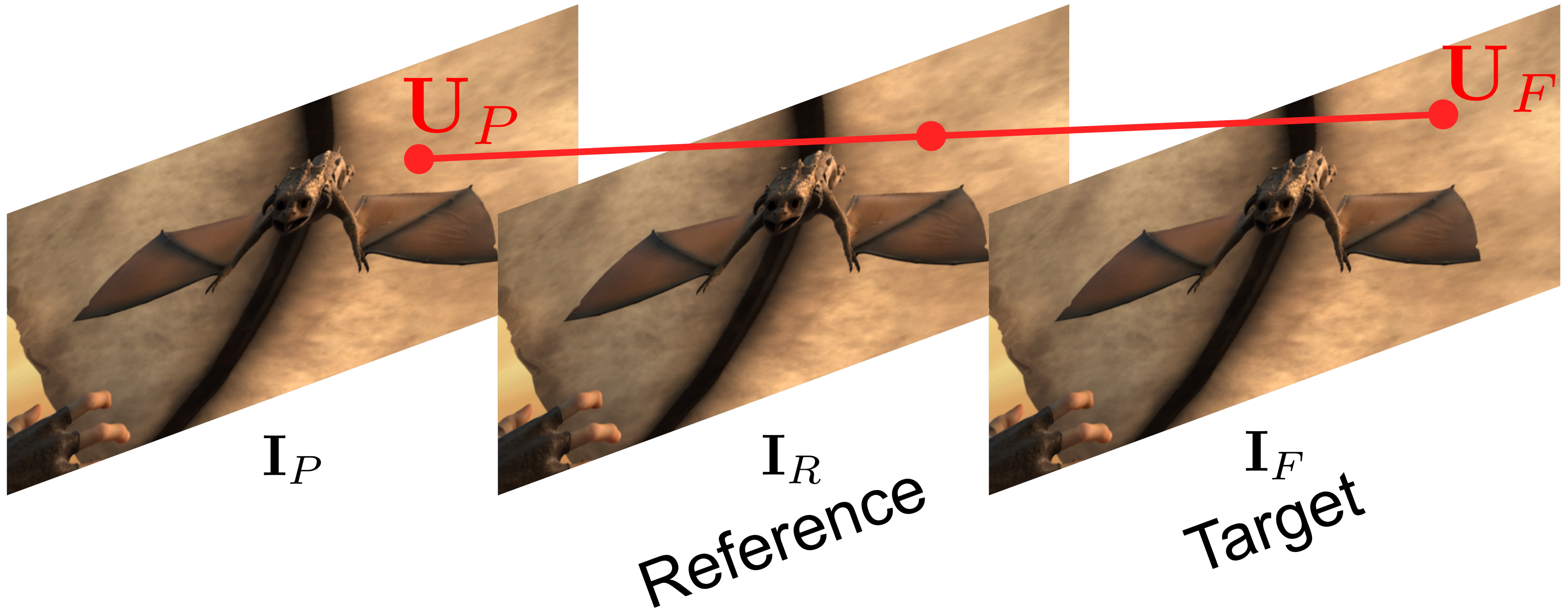
Formulation: Future Occlusion

- In a temporal window
    - Flow $\mathbf{U}_P, \mathbf{U}_F \in \mathbb{R}^{W \times H \times 2}$
    - Occlusions $\mathbf{O} \in [0,1]^{W \times H \times 2}$

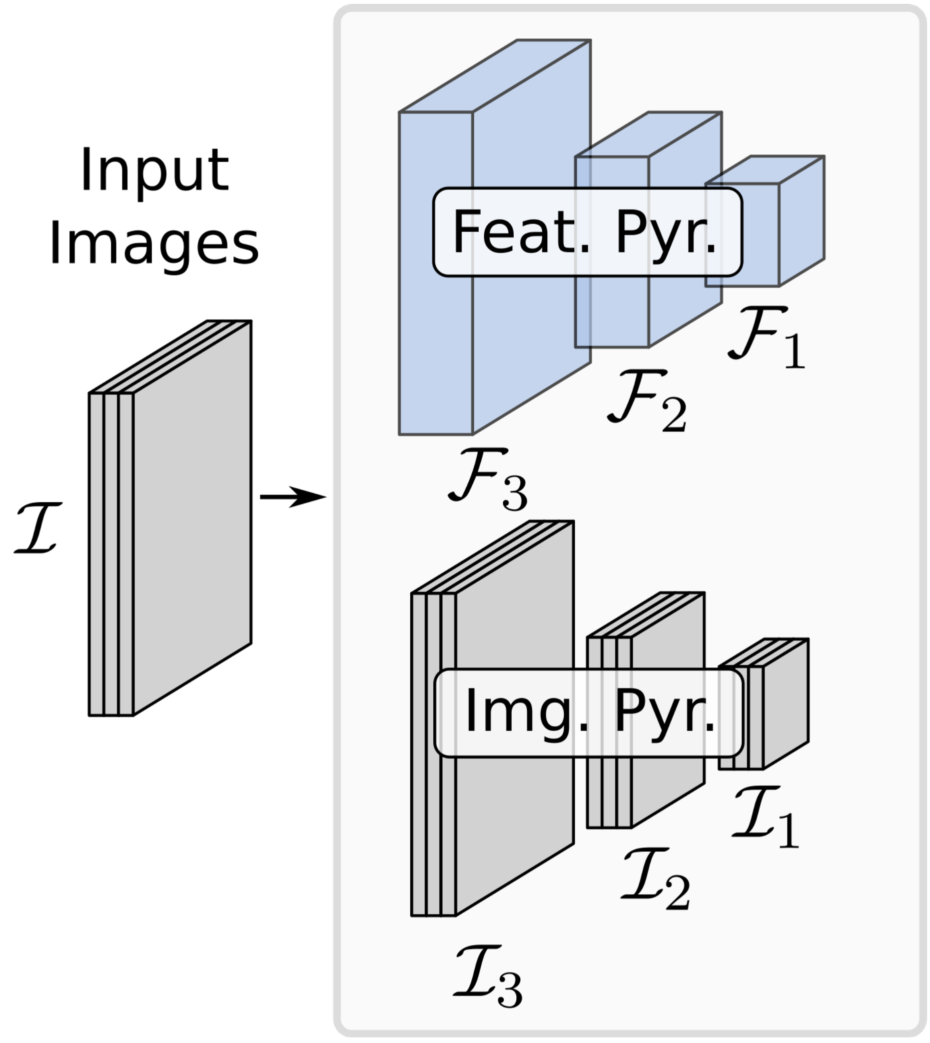


Formulation

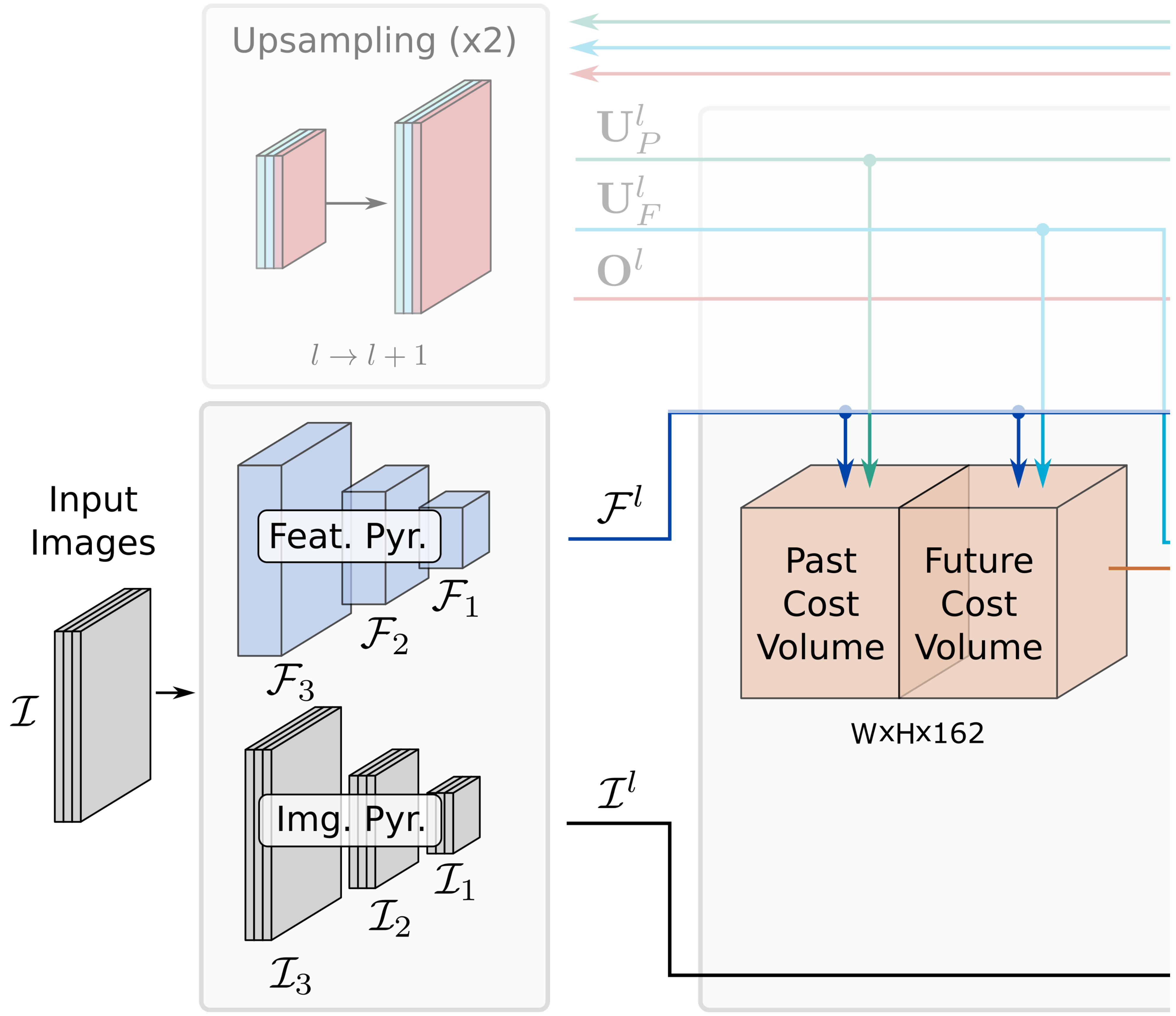
- In a temporal window
  - Flow $\mathbf{U}_P, \mathbf{U}_F \in \mathbb{R}^{W \times H \times 2}$
  - Occlusions $\mathbf{O} \in [0,1]^{W \times H \times 2}$



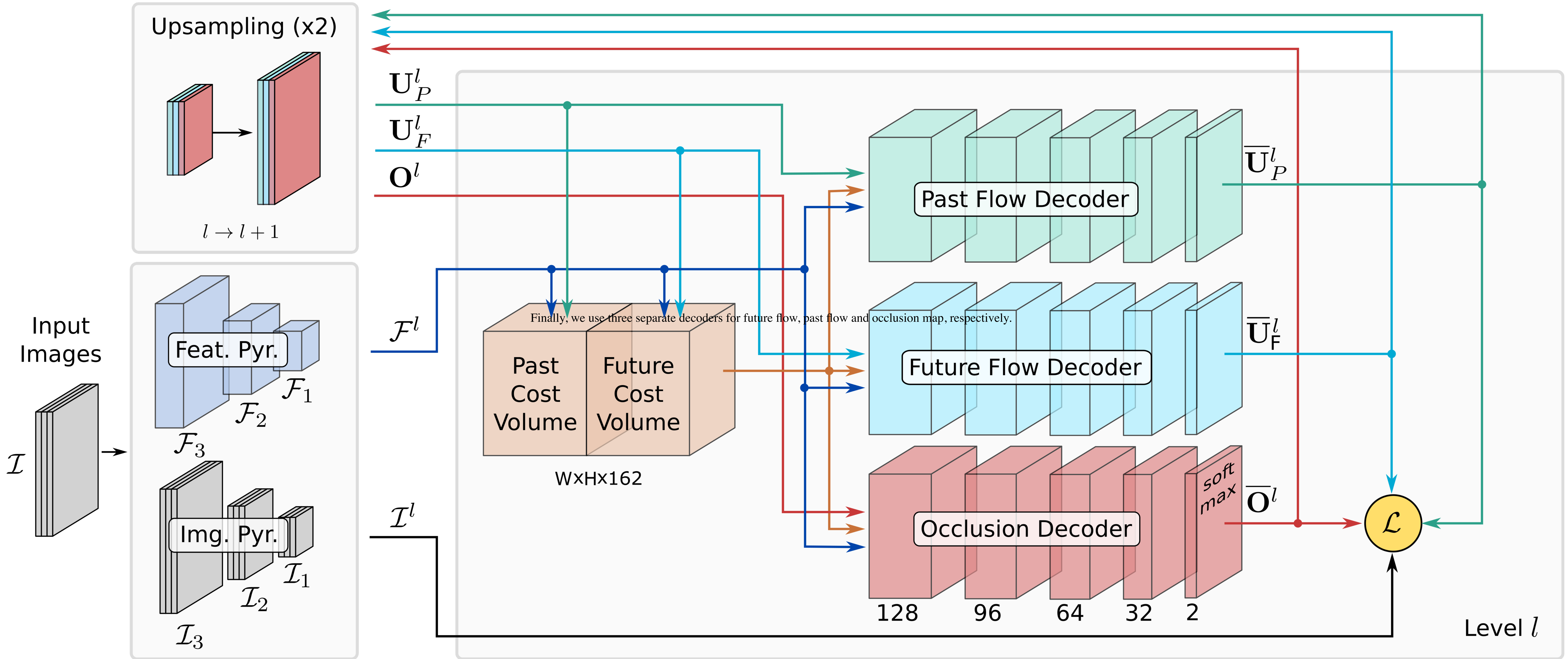
Architecture



Architecture



Architecture



Loss Function

$$\mathcal{L} = \underbrace{\mathcal{L}_P}_{\text{Photometric}} + \underbrace{\mathcal{L}_{S_F} + \mathcal{L}_{S_P}}_{\text{Flow smoothness}} + \underbrace{\mathcal{L}_S}_{\text{Occlusion smoothness}}(\boldsymbol{\theta}) + \underbrace{\mathcal{L}_O}_{\text{Occlusion prior}} + \underbrace{\mathcal{L}_{CV}(\boldsymbol{\theta})}_{\text{Constant velocity}}$$

Loss Function

$$\mathcal{L} = \underbrace{\boldsymbol{\mathcal{L}}_P}_{\textbf{Photometric}} + \underbrace{\mathcal{L}_{S_F} + \mathcal{L}_{S_P}}_{\text{Flow smoothness}} + \underbrace{\mathcal{L}_S}_{\text{Occlusion smoothness}}(\boldsymbol{\theta}) + \underbrace{\mathcal{L}_O}_{\text{Occlusion prior}} + \underbrace{\mathcal{L}_{CV}}_{\text{Constant velocity}}(\boldsymbol{\theta})$$

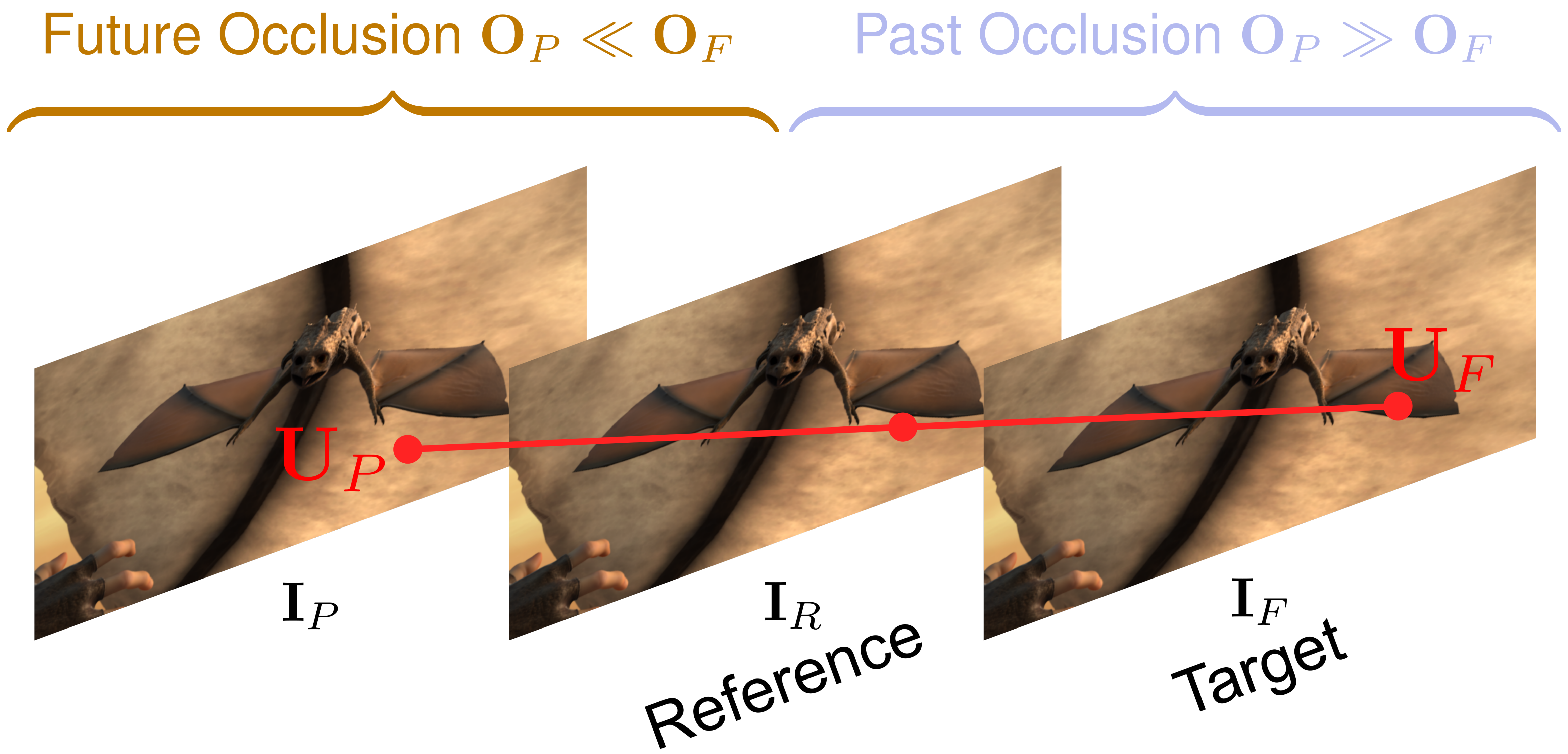
Photometric Loss

Future Occlusion — Warped Past

$$\sum_{\mathbf{p}\in\Omega} \mathbf{O}_F(\mathbf{p}) \cdot \delta\left(\hat{\mathbf{I}}_P\left(\mathbf{p}+\mathbf{u}_P(\mathbf{p})\right), \mathbf{I}_R(\mathbf{p})\right)$$

$$+\sum_{\mathbf{p}\in\Omega} \mathbf{O}_P(\mathbf{p}) \cdot \delta\left(\hat{\mathbf{I}}_F\left(\mathbf{p}+\mathbf{u}_F(\mathbf{p})\right), \mathbf{I}_R(\mathbf{p})\right)$$

Past Occlusion — Warped Future — Reference

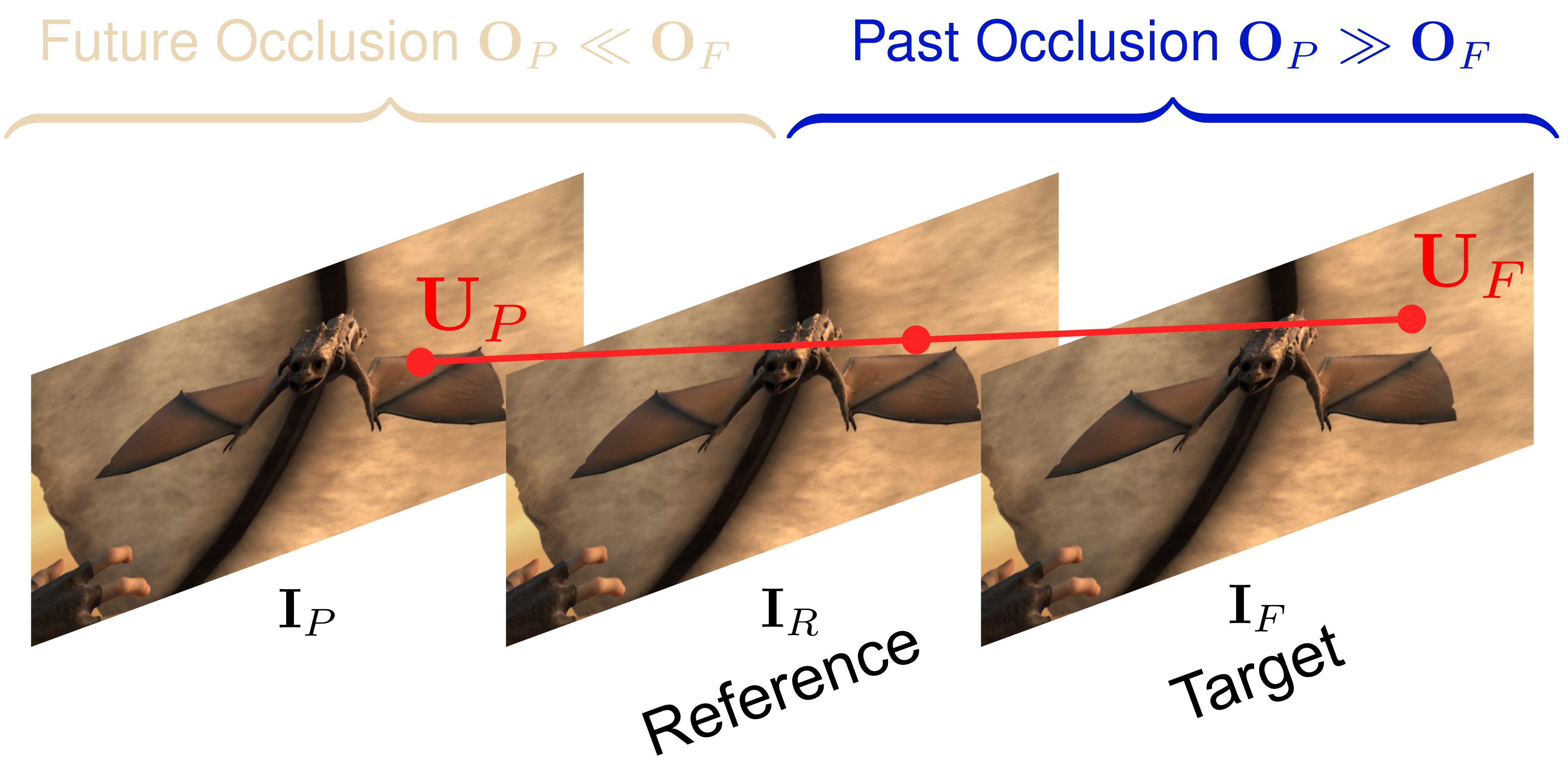


$$\mathbf{O}(p) = (\mathbf{O}_F, \mathbf{O}_P) \approx (1,0)$$

Photometric Loss

$$\sum_{\mathbf{p}\in\Omega} \overset{\text{Future Occlusion}}{\mathbf{O}_F(\mathbf{p})} \cdot \delta\left(\overset{\text{Warped Past}}{\hat{\mathbf{I}}_P\left(\mathbf{p}+\mathbf{u}_P(\mathbf{p})\right)}, \mathbf{I}_R(\mathbf{p})\right)$$

$$+\sum_{\mathbf{p}\in\Omega} \underset{\text{Past Occlusion}}{\mathbf{O}_P(\mathbf{p})} \cdot \delta\left(\underset{\text{Warped Future}}{\hat{\mathbf{I}}_F\left(\mathbf{p}+\mathbf{u}_F(\mathbf{p})\right)}, \underset{\text{Reference}}{\mathbf{I}_R(\mathbf{p})}\right)$$

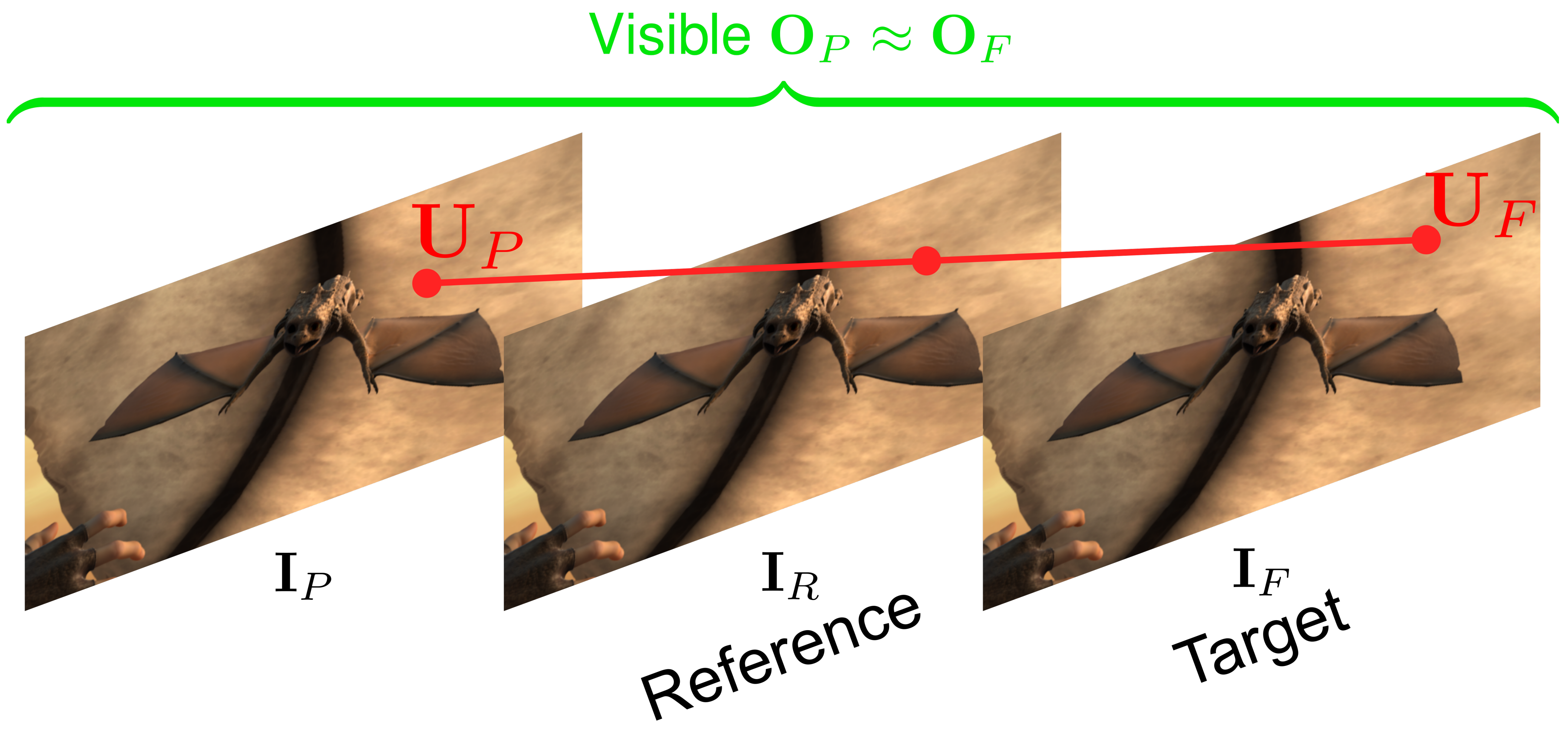


$$\mathbf{O}(p) = (\mathbf{O}_F, \mathbf{O}_P) \approx (0,1)$$

Photometric Loss

$$\sum_{\mathbf{p}\in\Omega} \overset{\text{Future Occlusion}}{\overset{\downarrow}{\mathbf{O}_F(\mathbf{p})}} \cdot \delta\left(\hat{\mathbf{I}}_P\left(\mathbf{p}+\overset{\text{Warped Past}}{\overset{\downarrow}{\mathbf{u}_P}}(\mathbf{p})\right), \mathbf{I}_R(\mathbf{p})\right)$$

$$+\sum_{\mathbf{p}\in\Omega} \underset{\text{Past Occlusion}}{\underset{\uparrow}{\mathbf{O}_P(\mathbf{p})}} \cdot \delta\left(\hat{\mathbf{I}}_F\left(\mathbf{p}+\underset{\text{Warped Future}}{\underset{\uparrow}{\mathbf{u}_F}}(\mathbf{p})\right), \underset{\text{Reference}}{\underset{\uparrow}{\mathbf{I}_R(\mathbf{p})}}\right)$$

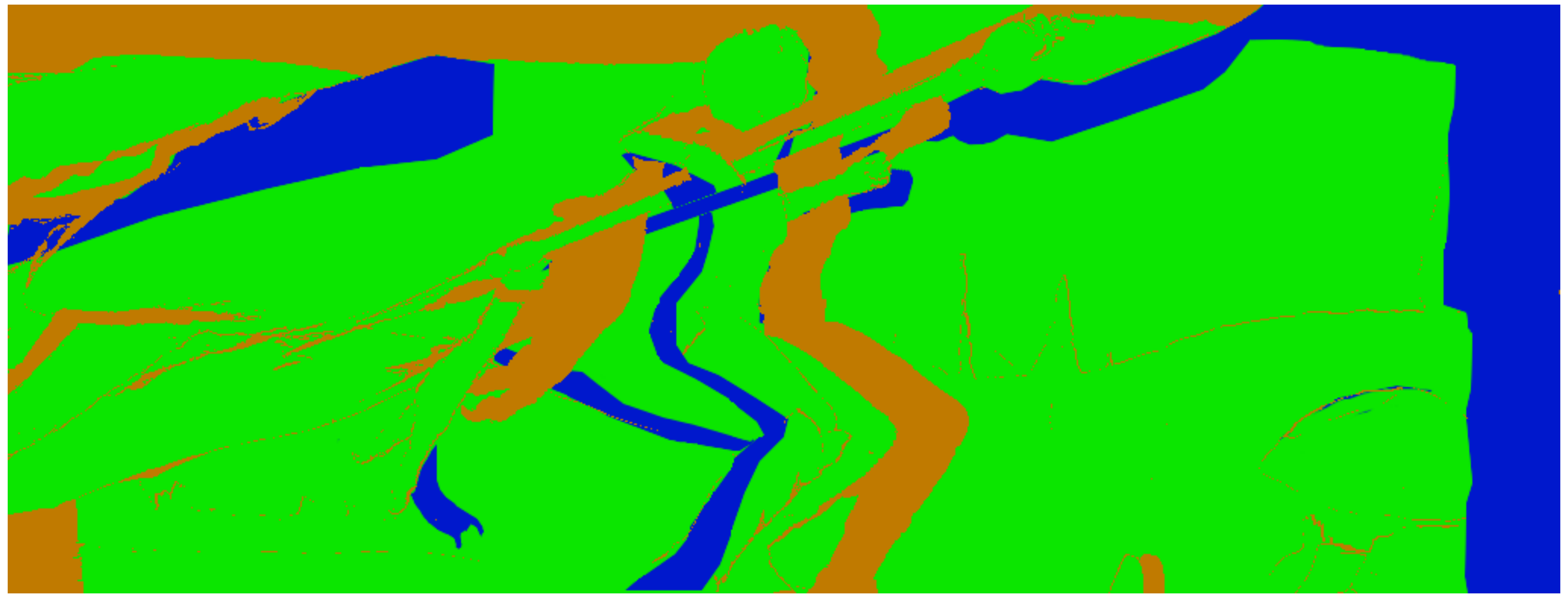


$$\mathbf{O}(p) = (\mathbf{O}_F, \mathbf{O}_P) \approx (0.5, 0.5)$$

Loss Function

$$\mathcal{L} = \underbrace{\mathcal{L}_P}_{\text{Photometric}} + \underbrace{\mathcal{L}_{S_F} + \mathcal{L}_{S_P}}_{\text{Flow smoothness}} + \underbrace{\mathcal{L}_S}_{\text{Occlusion smoothness}}(\boldsymbol{\theta}) + \underbrace{\boldsymbol{\mathcal{L}_O}}_{\textbf{Occlusion prior}} + \underbrace{\mathcal{L}_{CV}}_{\text{Constant velocity}}(\boldsymbol{\theta})$$

Occlusion Prior

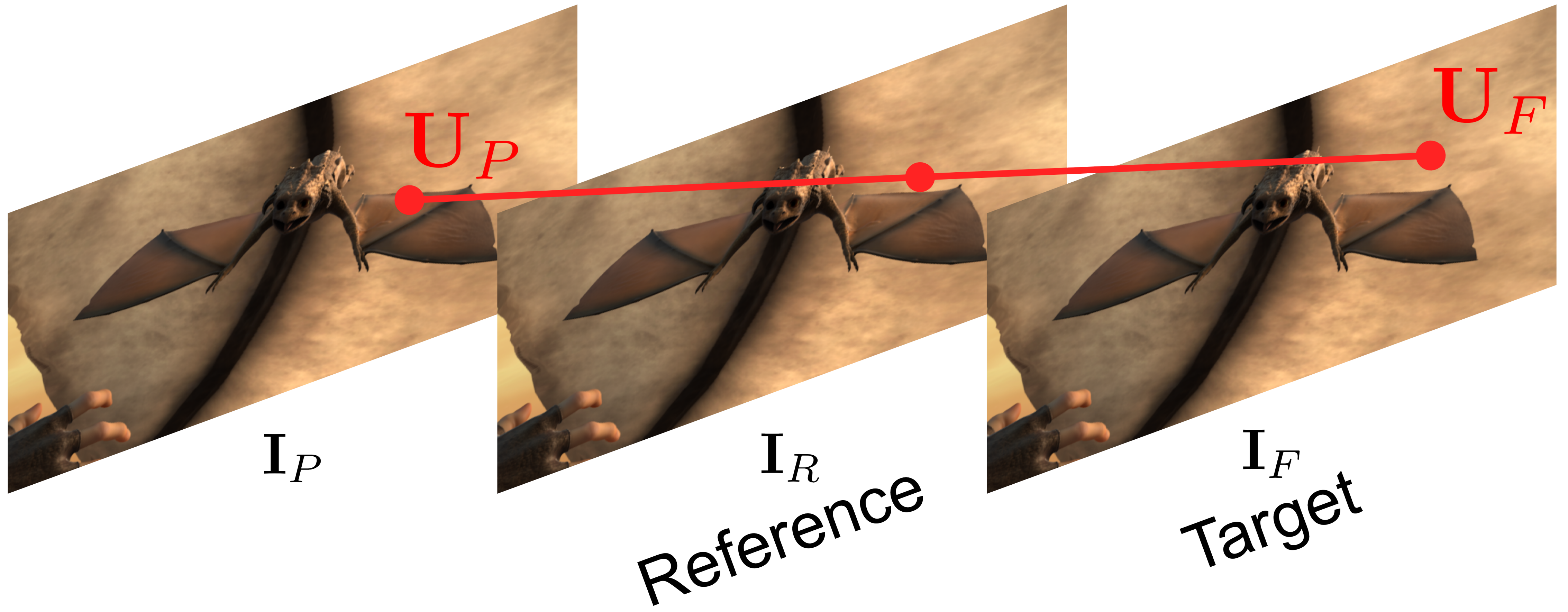


$$-\sum_{\mathbf{p}\in\Omega} \mathbf{O}_F(\mathbf{p}) \cdot \mathbf{O}_P(\mathbf{p})$$

Loss Function

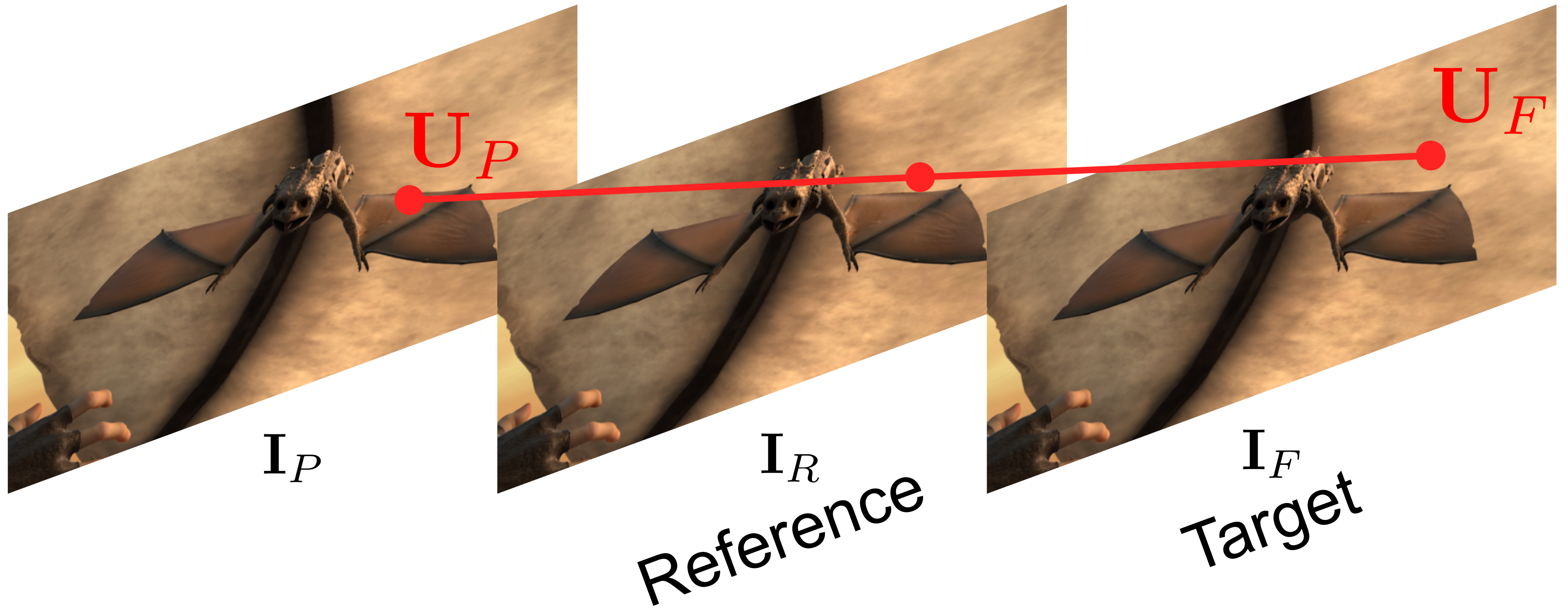
$$\mathcal{L} = \underbrace{\mathcal{L}_P}_{\text{Photometric}} + \underbrace{\mathcal{L}_{S_F} + \mathcal{L}_{S_P}}_{\text{Flow smoothness}} + \underbrace{\mathcal{L}_S(\boldsymbol{\theta})}_{\text{Occlusion smoothness}} + \underbrace{\mathcal{L}_O}_{\text{Occlusion prior}} + \underbrace{\mathcal{L}_{CV}(\boldsymbol{\theta})}_{\textbf{Constant velocity}}$$

Constant Velocity: Hard Constraint



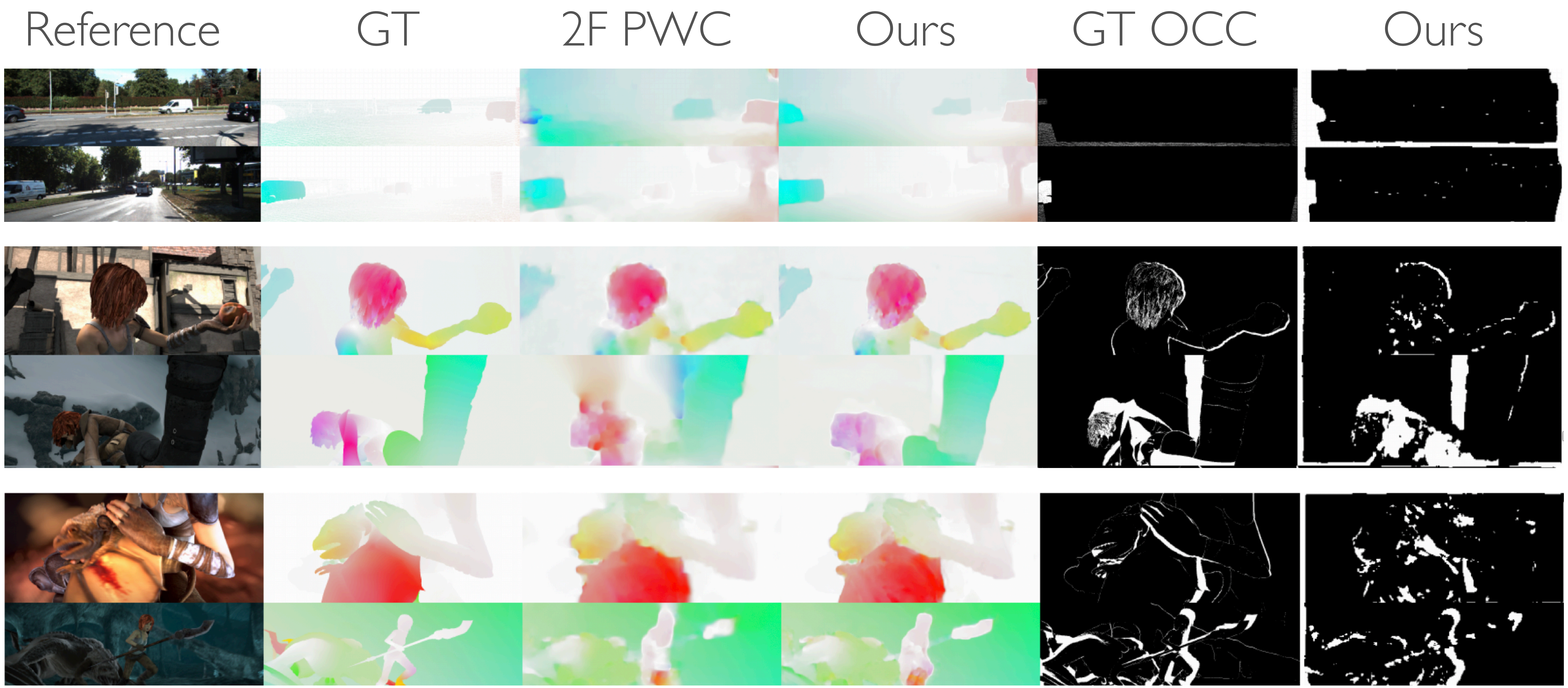
$$U(\mathbf{p}) = -\,U_P(\mathbf{p}) = \mathbf{U}_F(\mathbf{p})$$

Constant Velocity: Soft Constraint

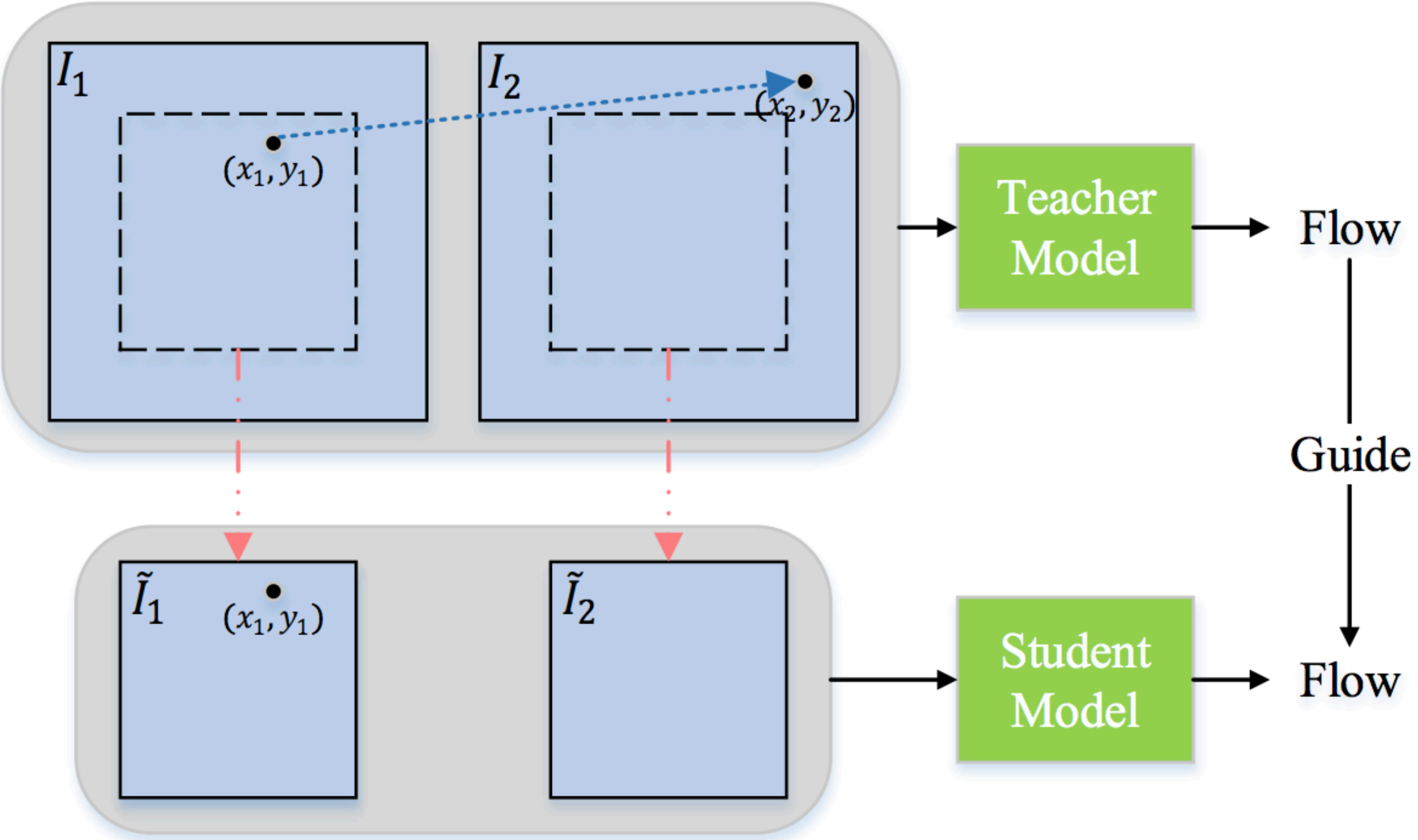


$$\sum_{\mathbf{p}\in\Omega} \rho\left(\mathbf{U}_P(\mathbf{p}) + \mathbf{U}_F(\mathbf{p})\right)$$

Visualization



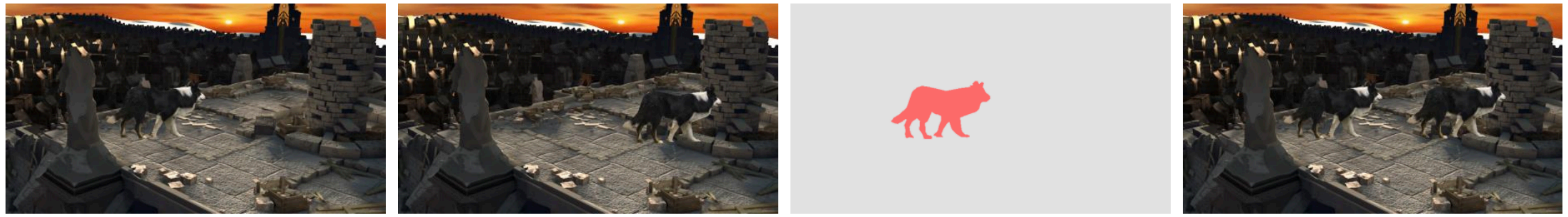
Data Distillation: DDFlow

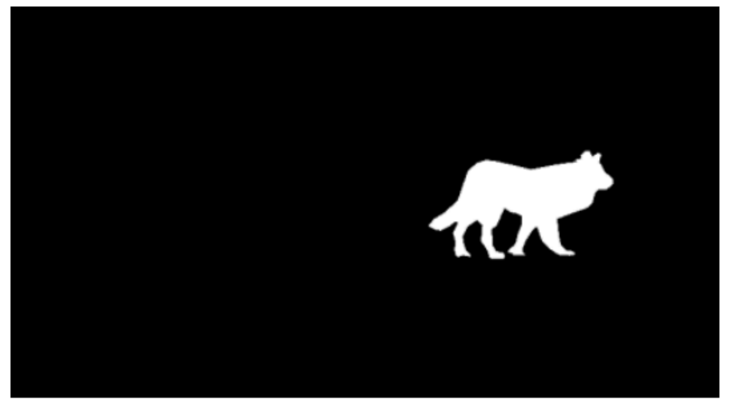


P. Liu, I. King, M. Lyu, and J Xu. DDFlow: Learning Optical Flow with Unlabelled Data Distillation. AAAI, 2019.

Data Distillation: SelFlow

(a) Reference Image I_t (b) Target Image I_{t+1} (c) Ground Truth Flow $\mathbf{w}_{t\to t+1}$ (d) Warped Target Image $I^w_{t+1\to t}$

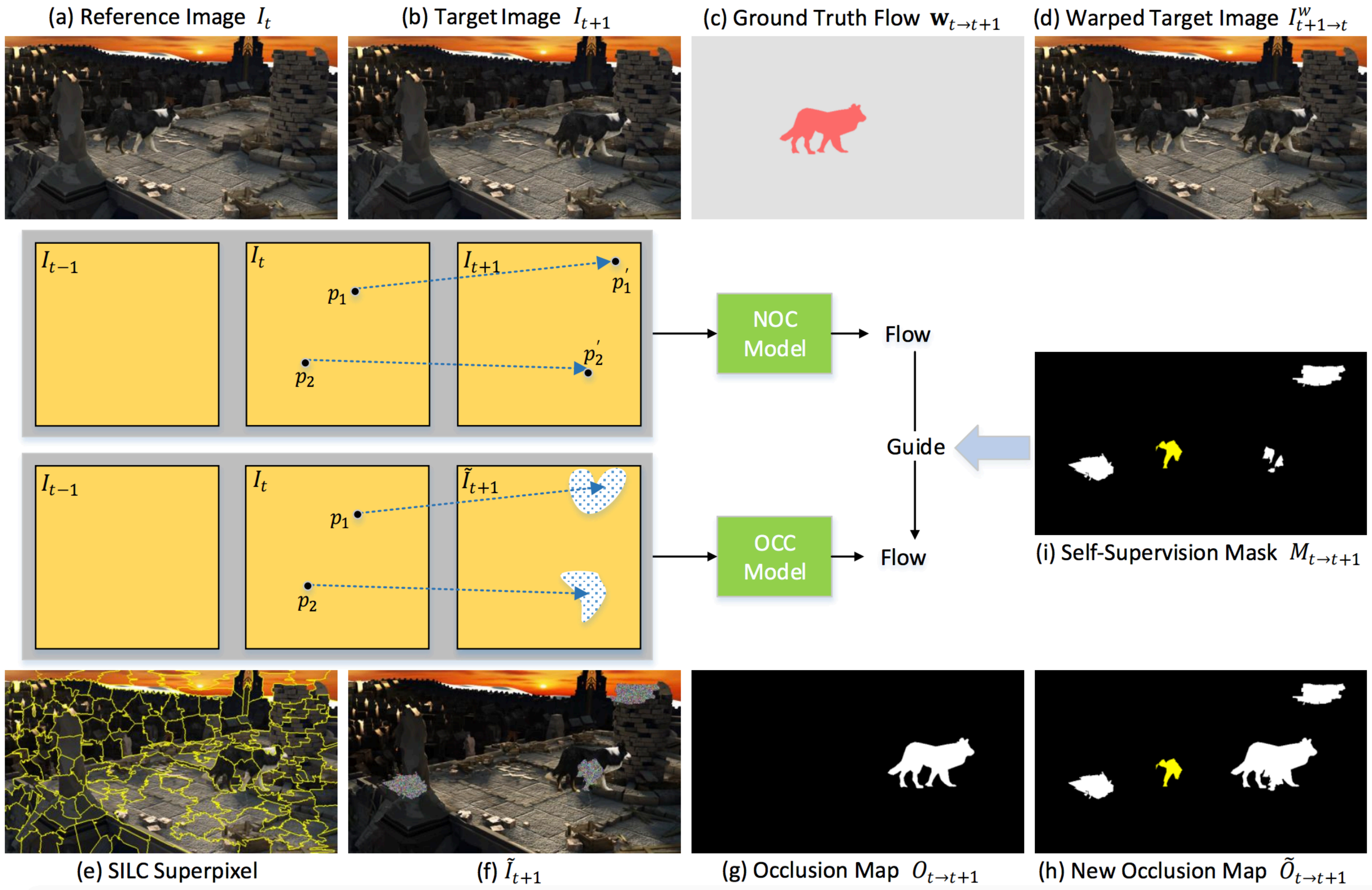




(g) Occlusion Map $O_{t\to t+1}$

P. Liu, M. Lyu, I. King, and J Xu. SelFlow: Self-Supervised Learning of Optical Flow. CVPR, 2019.

Data Distillation: SelFlow



P. Liu, M. Lyu, I. King, and J Xu. SelFlow: Self-Supervised Learning of Optical Flow. CVPR, 2019.

Good Practices

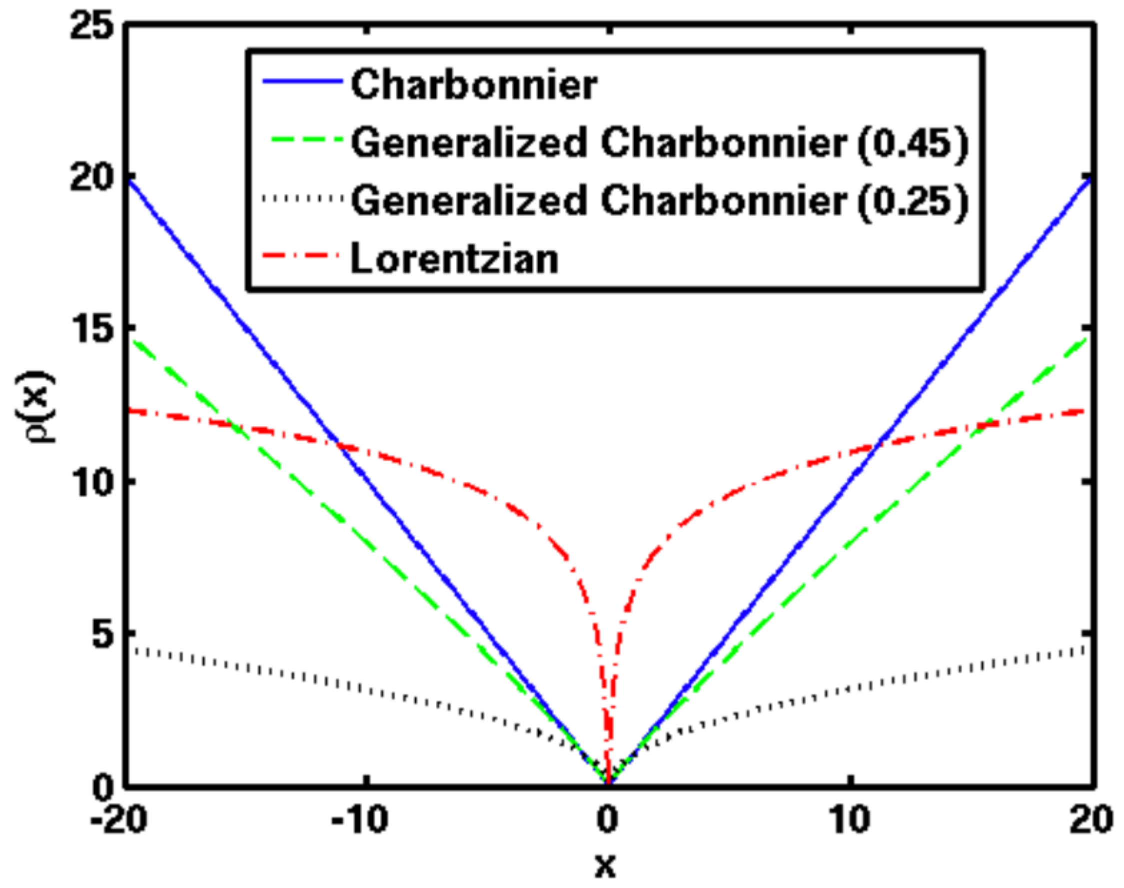
- Photometric losses
- Occlusion estimation
- Smoothness constraints
- Self-supervision (data distillation)
- Other choices
  pre-training, image resolution, data augmentation, and batch size

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Photometric Losses

- The Generalized Charbonnier

$$\rho(x) = (x^2 + \epsilon^2)^\alpha$$

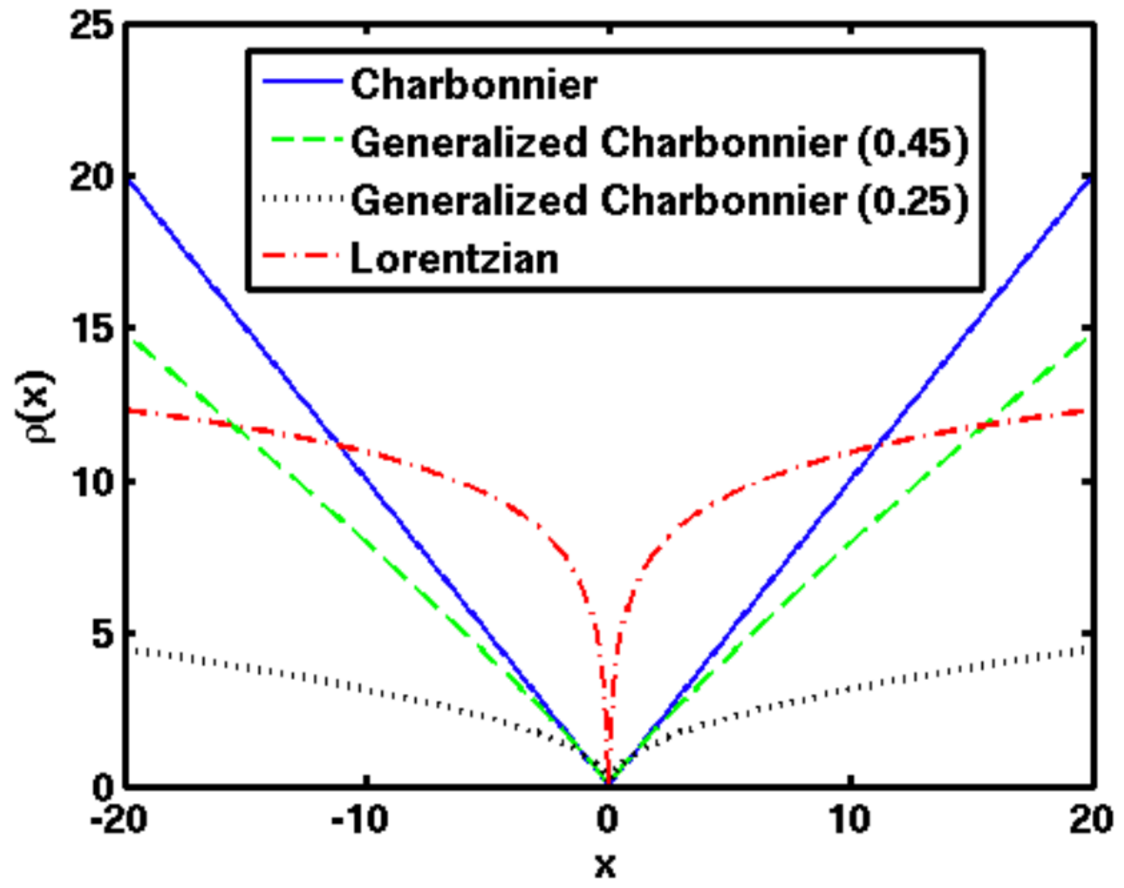


R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Photometric Losses

- The Generalized Charbonnier

$\rho(x) = (x^2 + \epsilon^2)^\alpha$



- Structural similarity index (SSIM)

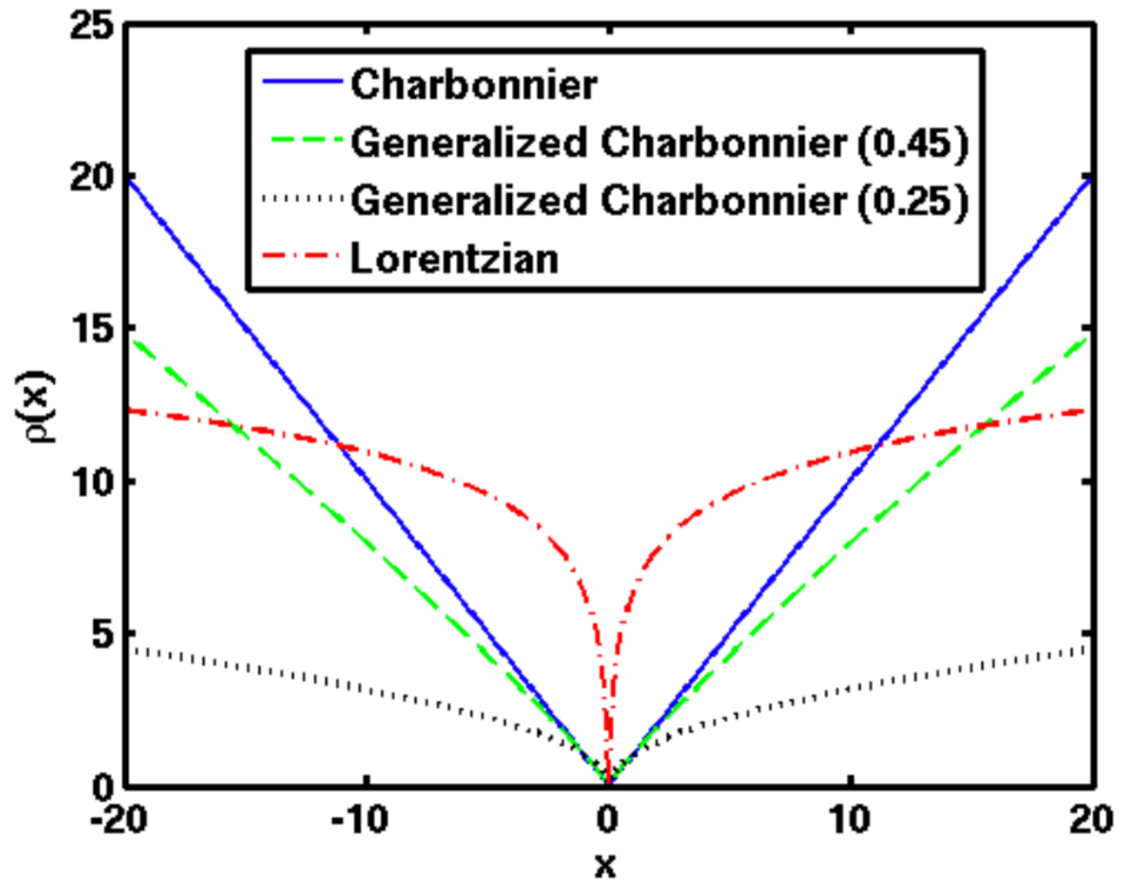
$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c1)(\sigma_x + \sigma_y + c_2)}$$

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Photometric Losses

- The Generalized Charbonnier

$$\rho(x) = (x^2 + \epsilon^2)^\alpha$$



- Structural similarity index (SSIM)

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c1)(\sigma_x + \sigma_y + c_2)}$$

- The Census loss

$$\xi(p, p') = \begin{cases} 0 & \text{if } p > p' \\ 1 & \text{if } p \leq p' \end{cases}$$

$$\begin{array}{|c|c|c|}\hline 124 & 74 & 32 \\ \hline 124 & 64 & 18 \\ \hline 157 & 116 & 84 \\ \hline \end{array} \longrightarrow \begin{array}{|c|c|c|}\hline 1 & 1 & 0 \\ \hline 1 & x & 0 \\ \hline 1 & 1 & 1 \\ \hline \end{array} \longrightarrow 11010111$$

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Occlusion Estimation

- Forward and backward flow

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Occlusion Estimation

- Forward and backward flow
- Using the range map of the backward flow

$$\mathbf{V}(x, y) = \sum_i \sum_j \mathrm{max}(0, 1 - |x - (i + \mathbf{F}_{21}^x(i, j))|) \cdot \mathrm{max}(0, 1 - |y - (j + \mathbf{F}_{21}^y(i, j))|)$$

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Occlusion Estimation

- Forward and backward flow
- Using the range map of the backward flow

$$\mathbf{V}(x, y) = \sum_i \sum_j \max(0, 1 - |x - (i + \mathbf{F}_{21}^x(i, j))|) \cdot \max(0, 1 - |y - (j + \mathbf{F}_{21}^y(i, j))|)$$

$$\mathbf{O}(x, y) = \min(1, \mathbf{V}(x, y))$$

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Occlusion Estimation

- Forward and backward flow
- Using the range map of the backward flow
  - Gradient stopping at occlusion masks

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Occlusion Estimation

- Forward and backward flow
- Using the range map of the backward flow
    - Gradient stopping at occlusion masks
- Learning a model for occlusion estimation

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Smoothness

- Edge-aware first and second order smoothness

$$\sum_{p} \exp\left(-\lambda \sum_{c} \left| \frac{\partial \mathbf{I}_c}{\partial p} \right| \right) \left| \frac{\partial^k \mathbf{F}}{\partial p^k} \right|$$

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Smoothness

- Edge-aware first and second order smoothness

$$\sum_p \exp\left(-\lambda \sum_c \left|\frac{\partial \mathbf{I}_c}{\partial p}\right|\right)\left|\frac{\partial^k \mathbf{F}}{\partial p^k}\right|$$

- Smoothness at flow resolution

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Self-Supervision

- For pixels that going out of the image boundary

1. Apply the model on the full images
2. Crop the images from each edge
3. Apply the model again
4. Use the cropped estimated flow from the full images
   as supervision for flow estimation from the cropped images

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Self-Supervision

- For pixels that going out of the image boundary
- Occlusion-weighted Charbonnier loss

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Good Practices: Self-Supervision

- For pixels that going out of the image boundary
- Occlusion-weighted Charbonnier loss
- Continual self-supervision and image resizing
    - No freezing; a single model that supervises itself
    - No gradients to teacher
    - In the original resolution

R. Jonschkowski, A. Stone, J.T. Barron, A. Gordon, K, Konolige, and A. Angelova. What Matters in Unsupervised Optical Flow. ARXIV, 2020.

Conclusion

- End-to-end learning for Optical Flow works great on the benchmarks.
- We need to start thinking about real-world scenes.
- Obtaining reference data using high-frame rate cameras might be a solution but we need to work on occlusions.
- Unsupervised Learning of multi-frame flow and occlusions is another direction. We can learn flow and occlusions from *any* data.
- The next step: explore natural scenes with unsupervised learning.

Thank you!


Joel Janai


Andreas Geiger


Michael J. Black


Anurag Ranjan


Jonas Wulff

Questions?