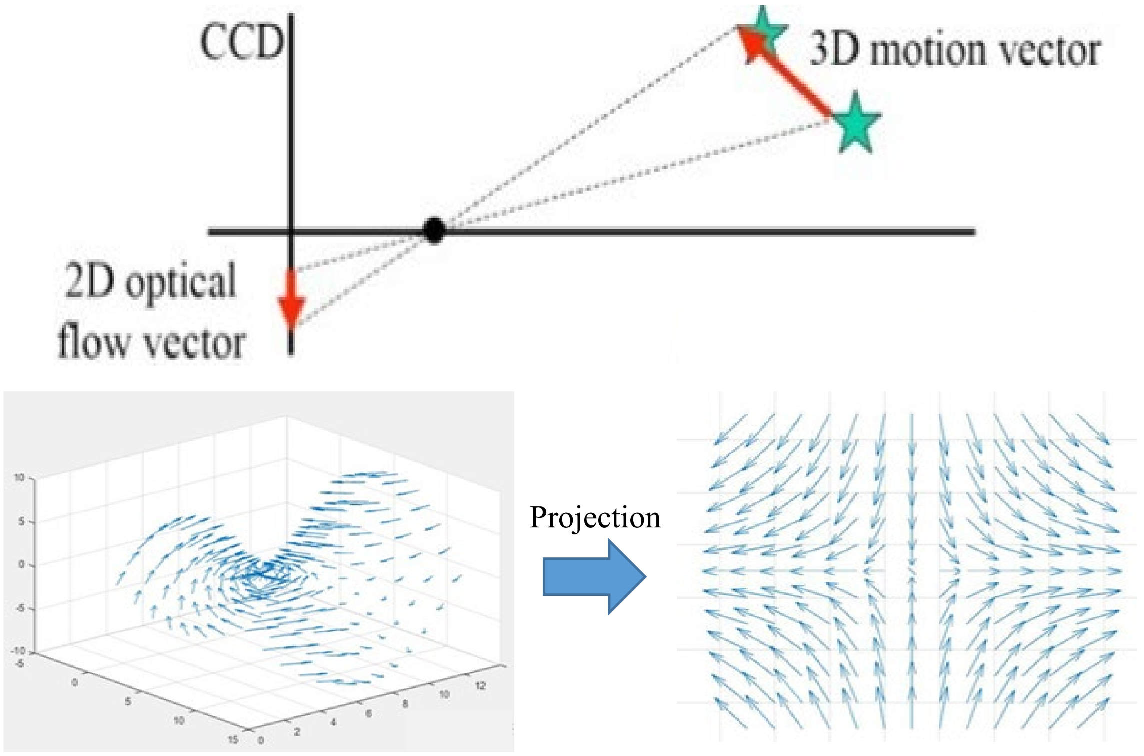
Unsupervised Cumulative Domain Adaptation for Foggy Scene Optical Flow

Hanyu Zhou[1], Yi Chang[1*], Wending Yan[2], Luxin Yan[1]

[1] Huazhong University of Science and Technology
[2] Huawei International Co. Ltd.

CVPR 2023

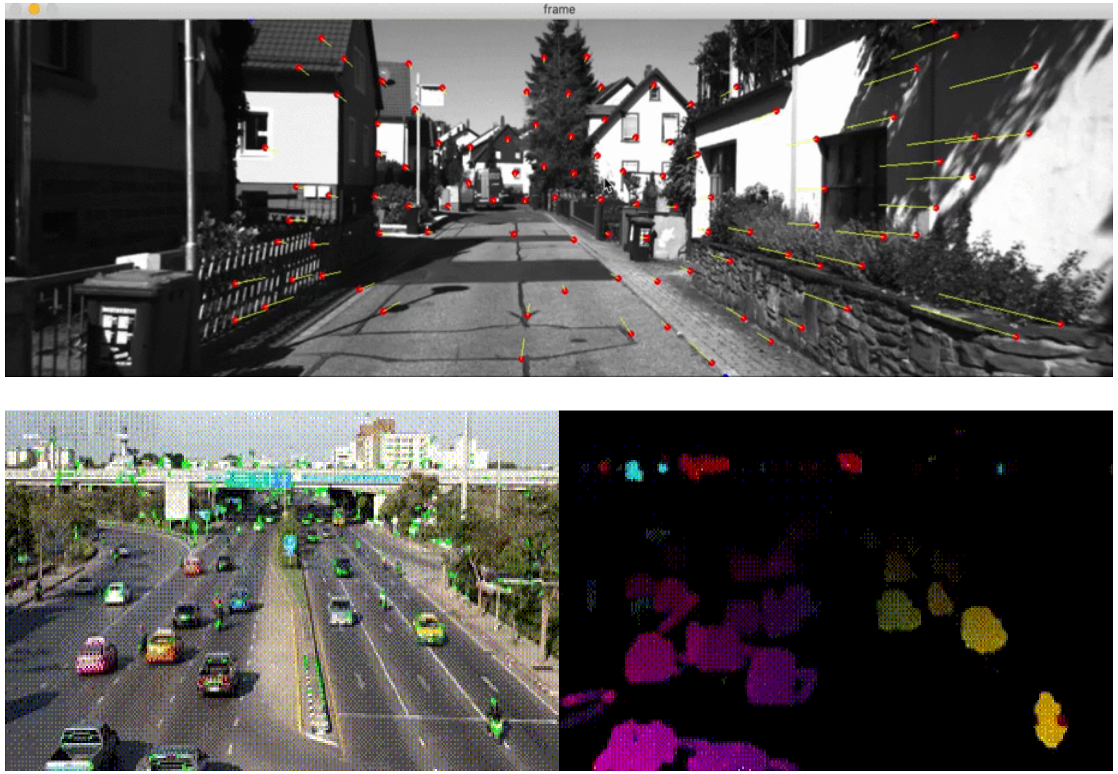
Optical Flow: Motion Estimation Tool



3D motion field

2D motion field

Optical Flow Schematic



Optical Flow Visualization

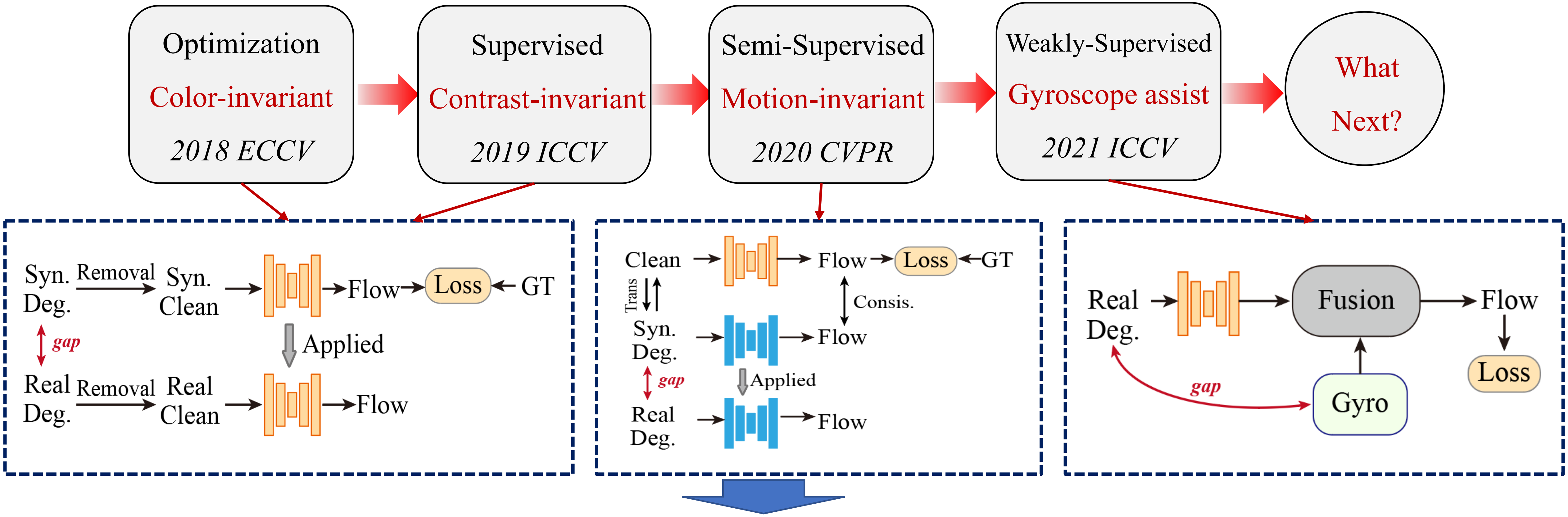
Problem

Clean Scene

Foggy Scene

Optical flow suffers degradation under foggy scenes

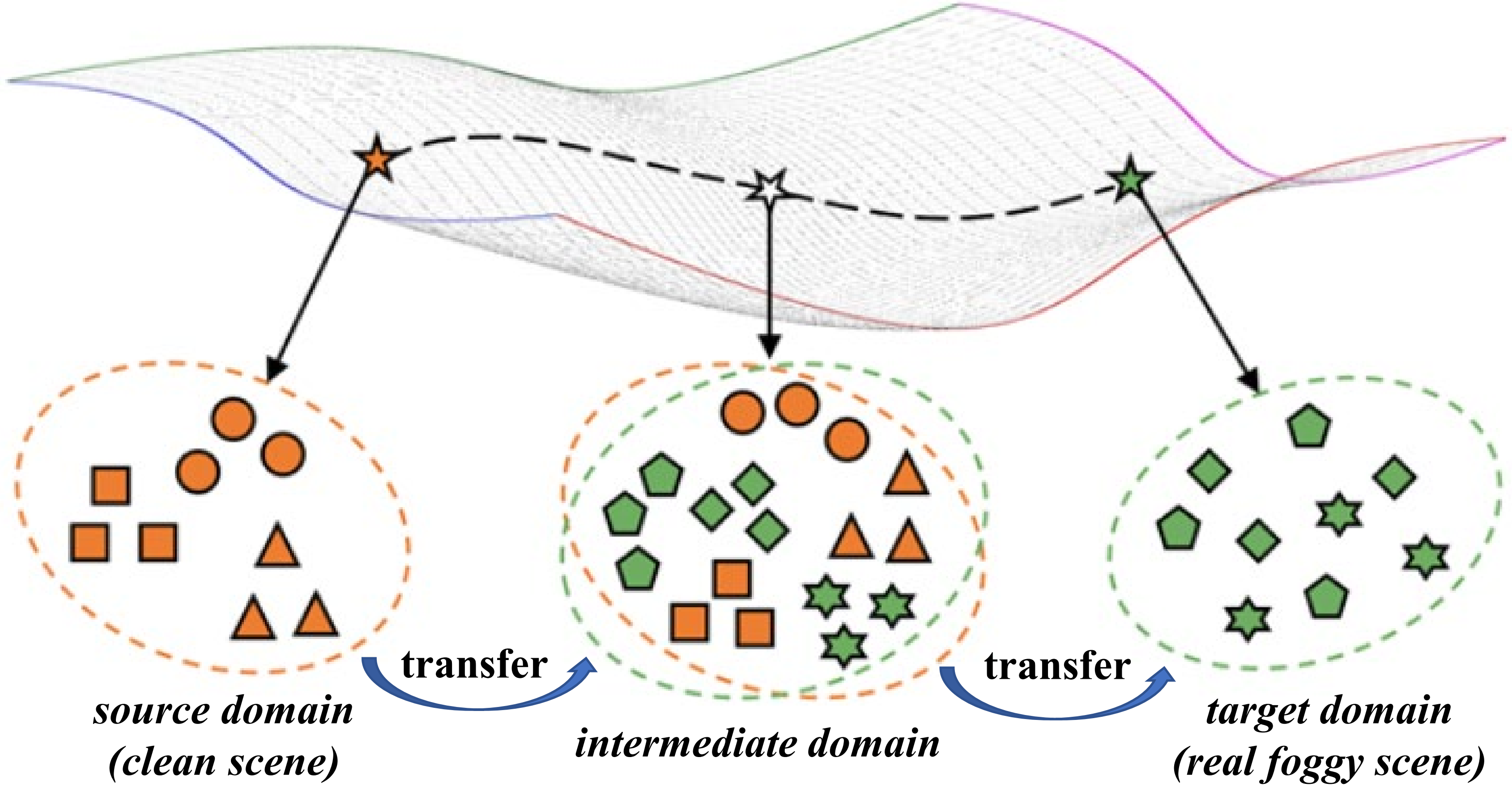
Related Work



One-stage solution: they mainly seeks degradation-invariant features to transfer knowledge from clean to degraded domain.

- ☐ There exists **synthetic-to-real domain gap.**
- ☐ Obtaining **optical flow GT** is difficult under foggy scenes.

Transfer Mechanism: Two-Stage Strategy



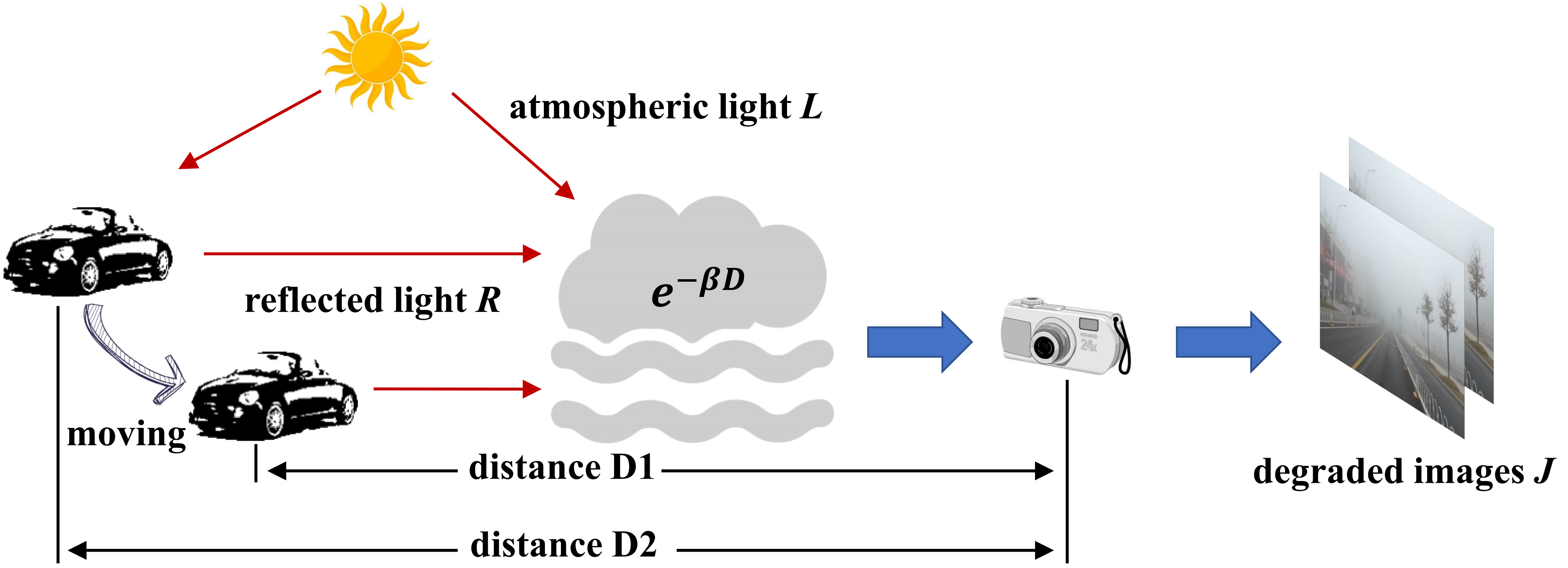
- Two-stage strategy: progressively transfer the motion knowledge from source domain to intermediate domain, and to target domain in an unsupervised manner.

closing dual domain gaps

without GT

How to do ?

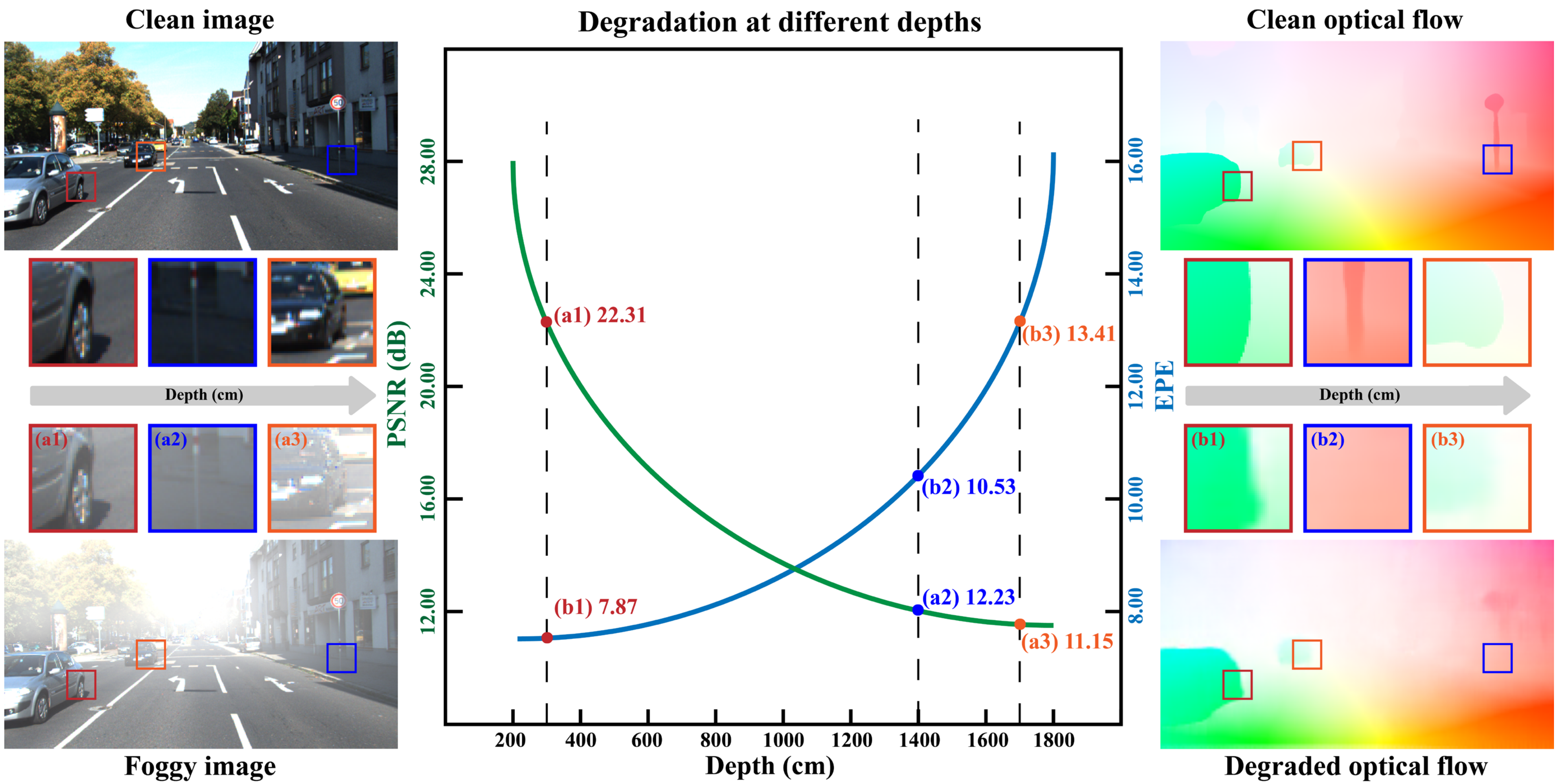
Motion under Foggy Scene



atmospheric scattering model: $L(x) = e^{-\beta D(x)} R(x) + L(1 - e^{-\beta D(x)})$

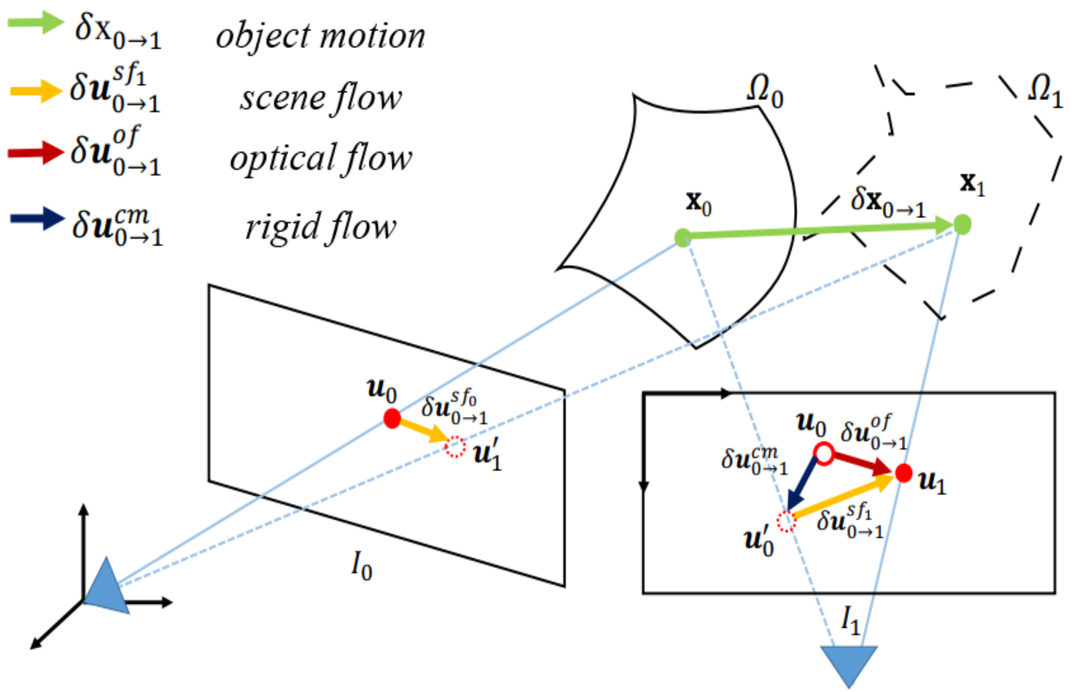
Fog is highly related scene depth

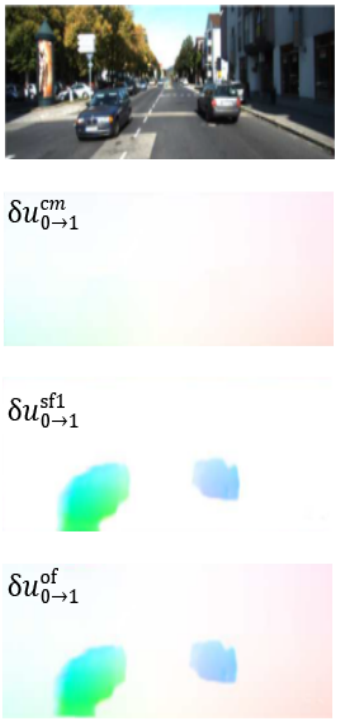
Motivation Ⅰ: What Affects Optical Flow?



The deeper depth, the inferior optical flow

Depth: The Key to Optical Flow





$$\delta u^{of}_{0\to1} = \delta u^{sf_1}_{0\to1} + \delta u^{cm}_{0\to1}$$

Only ego motion

$$\delta u^{of}_{0\to1} = \delta u^{cm}_{0\to1}$$

$$x_t = D_t(p_t)K^{-1}p_t$$

Pose transform

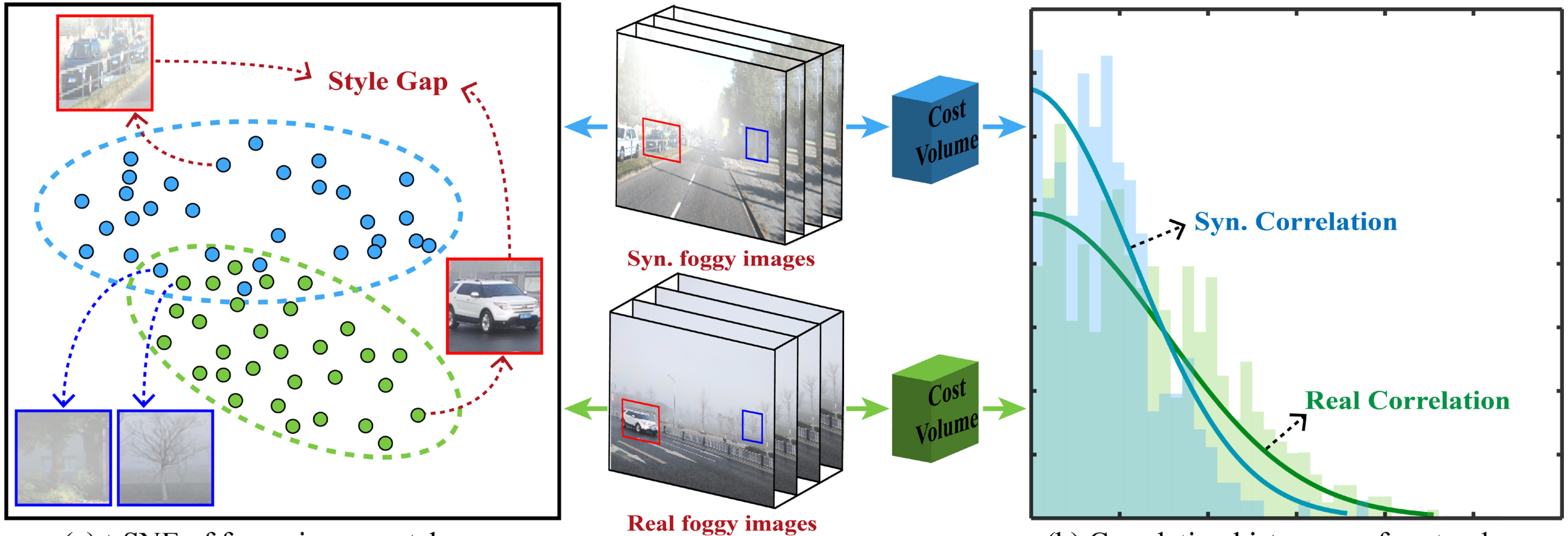
$$p_{t+1} = KPD_t(p_t)K^{-1}p_t$$

$$p_{t+1} - p_t$$

$$\delta u^{of}_{0\to1} = (KPD_t(p_t)K^{-1} - 1)p_t$$

There exists a 2D-3D geometry projection relationship between depth and optical flow

Motivation Ⅱ: What is Invariant?



(a) t-SNE of foggy image patches

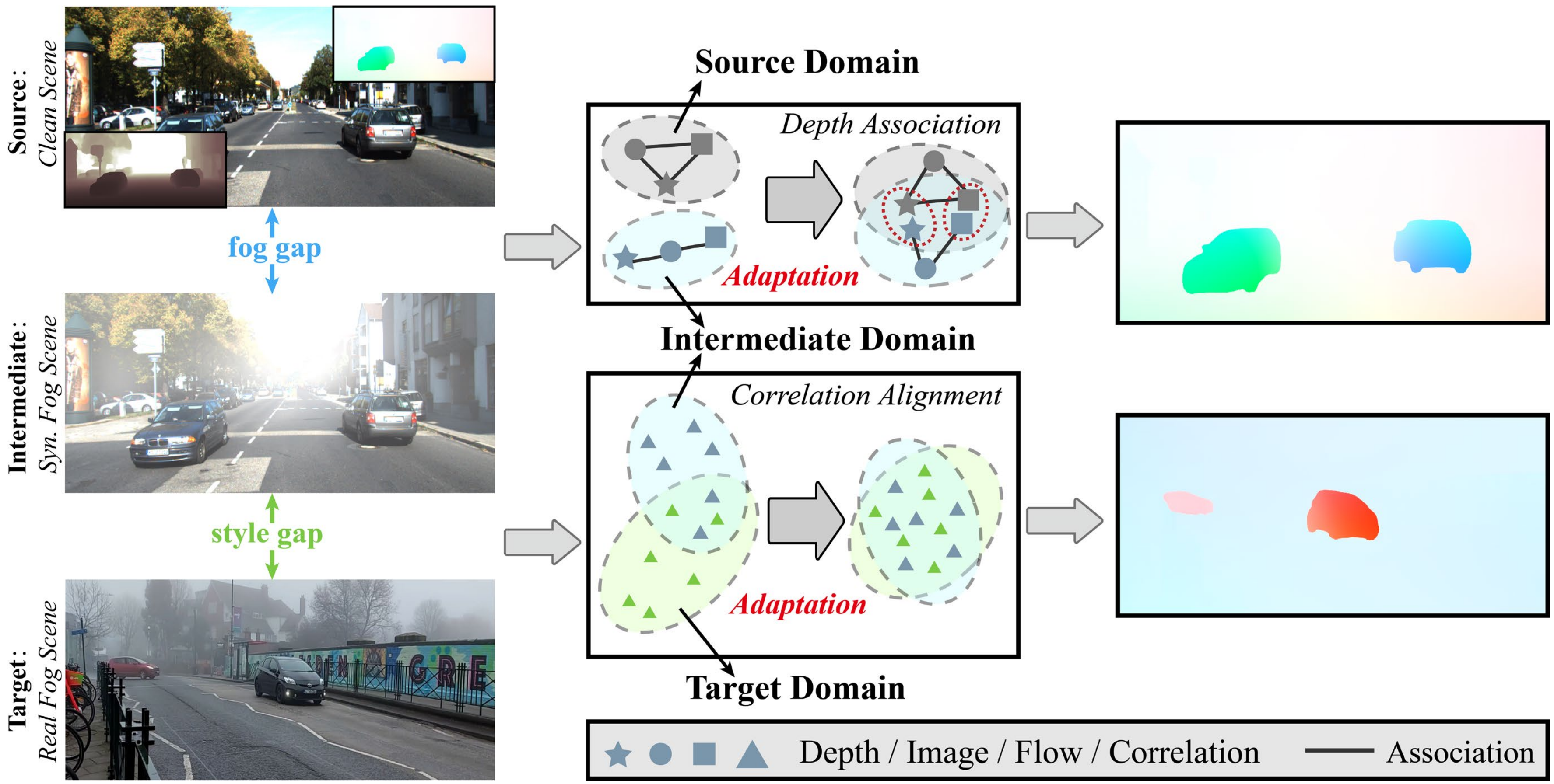
(b) Correlation histogram of cost volume

- ☐ There exists an **obvious style gap** between synthetic and real foggy images.
- ☐ Both the domains share a similar **correlation distribution**.

Cost volume benefits to bridging the synthetic-to-real domain gap

Main Idea: Cumulative Domain Adaptation

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

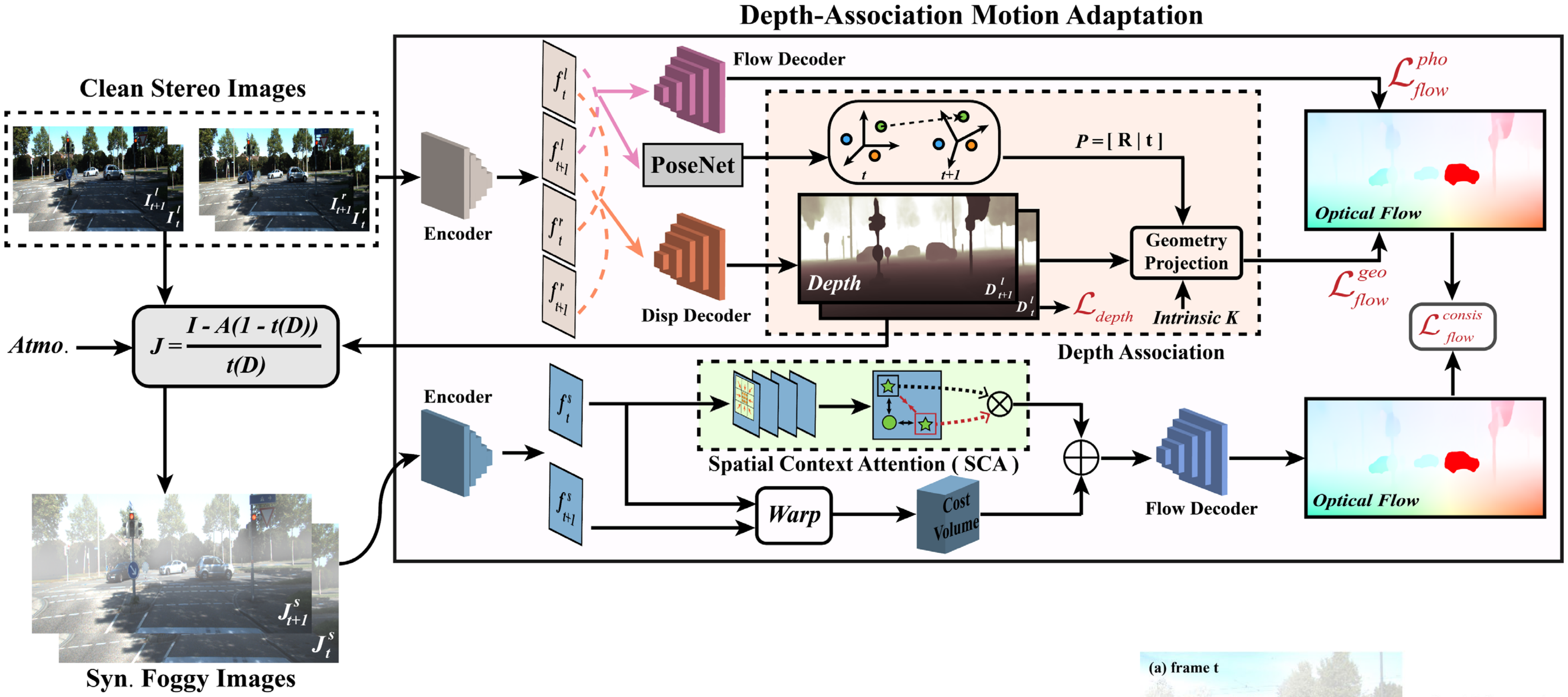


Depth Association Motion Adaptation → Clean-to-Foggy Gap

Correlation Alignment Motion Adaptation → Synthetic-to-Real Gap

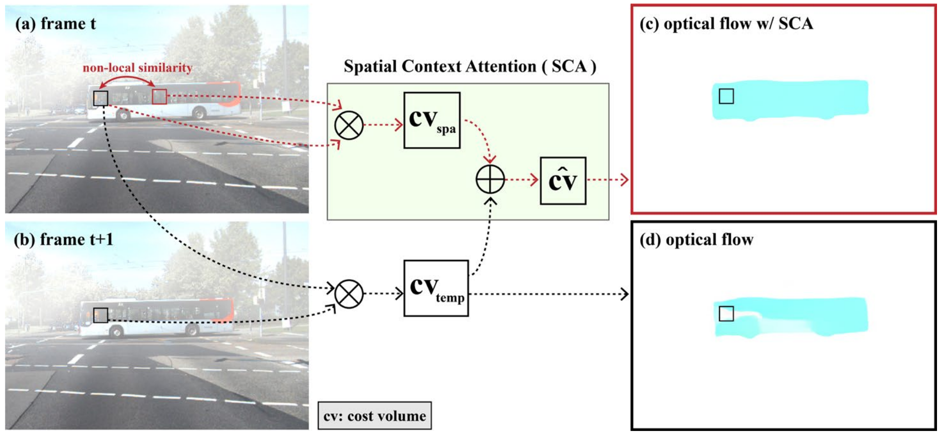
UCDA-Flow: Cumulative Domain Adaptation

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA



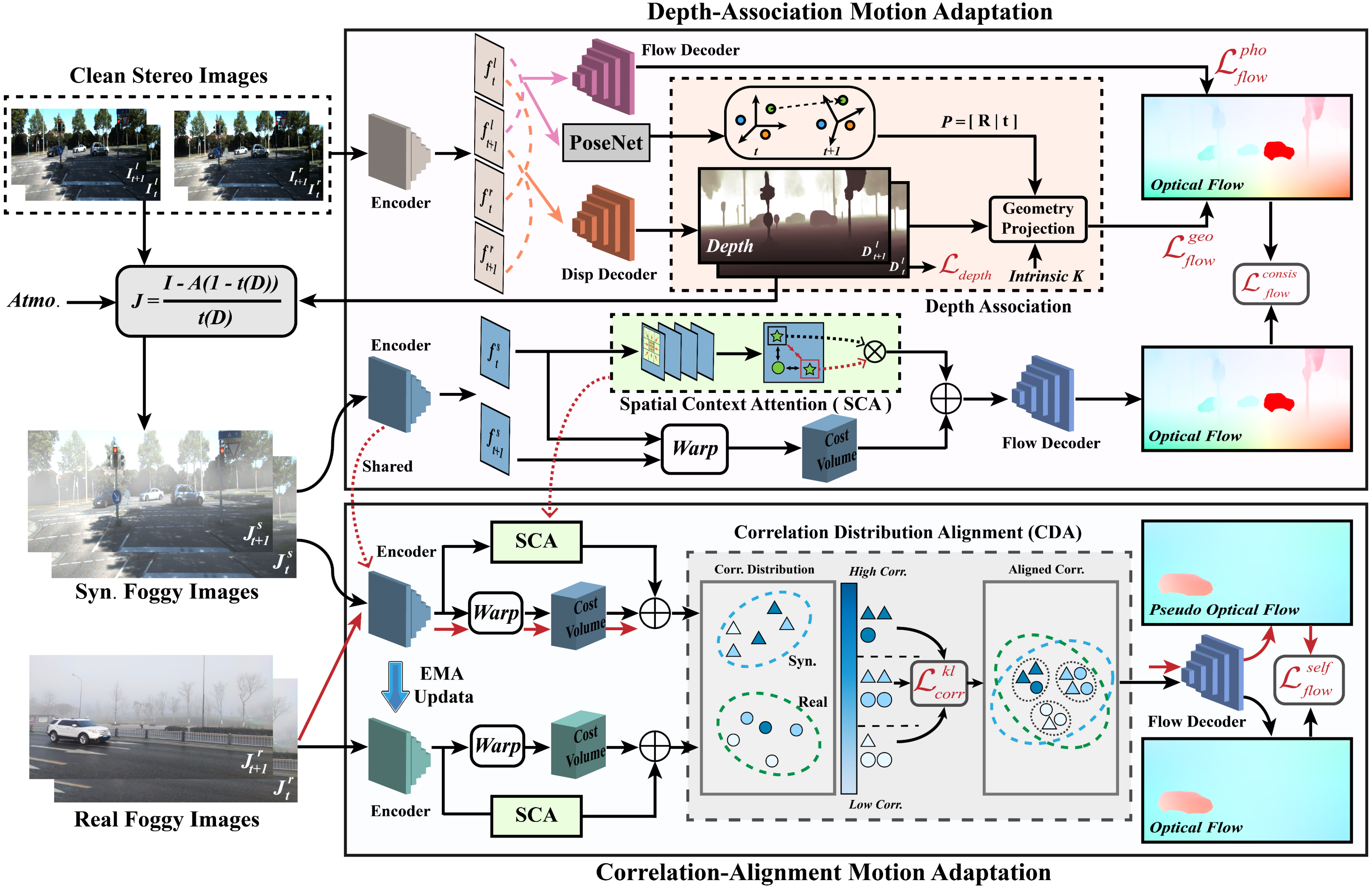
$$\mathcal{L}^{geo}_{flow} = \sum ||\boldsymbol{F} - \boldsymbol{F}_{rigid}||_1 \odot (1 - V) / \sum (1 - V),$$

$$\mathcal{L}^{consis}_{flow} = \sum ||\boldsymbol{F}_{syn} - \boldsymbol{F}||_1.$$



UCDA-Flow: Cumulative Domain Adaptation

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA



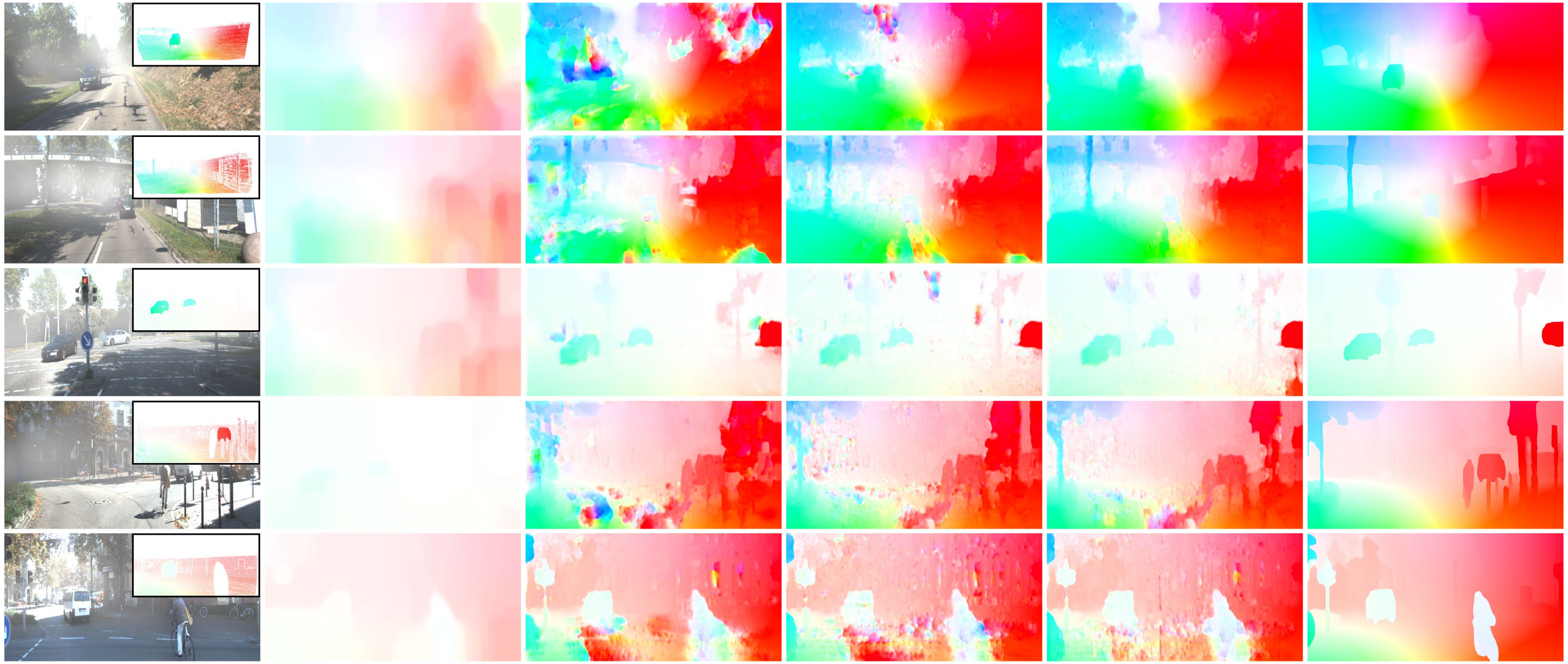
$$\mathcal{L}^{kl}_{corr} = \sum_{i=1}^{k} p_{r,i} log \frac{p_{r,i}}{p_{s,i}}$$

Total Loss

$$\mathcal{L} = \underbrace{\lambda_1 \mathcal{L}_{depth} + \lambda_2 \mathcal{L}_{flow}^{pho}}_{\text{Learn the knowledge from clean domain}} + \underbrace{\lambda_3 \mathcal{L}_{flow}^{geo} + \lambda_4 \mathcal{L}_{flow}^{consis}}_{\text{Transfer motion knowledge from clean-to-foggy domain}} + \underbrace{\lambda_5 \mathcal{L}_{flow}^{self} + \lambda_6 \mathcal{L}_{corr}^{kl}}_{\text{Transfer motion knowledge from synthetic-to-real domain}}$$

Comparison on Synthetic Dataset

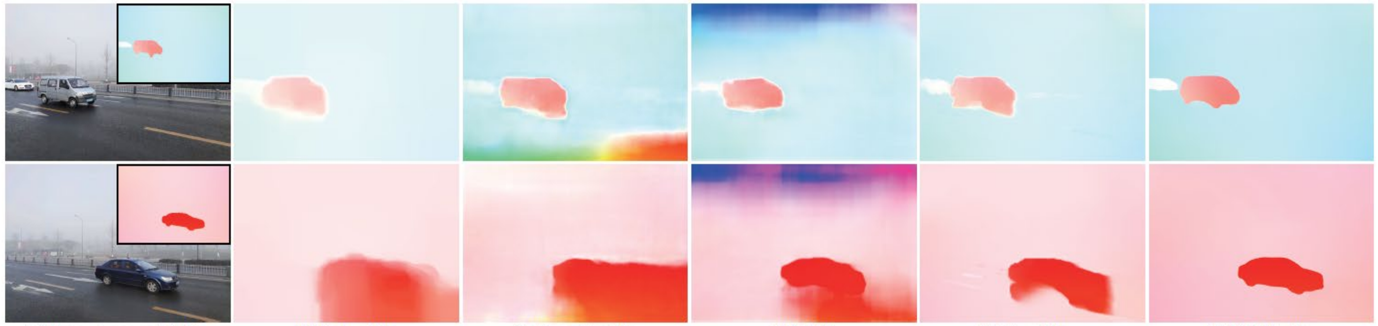
Method		RobustFlow	DenseFogFlow	UFlow			Selflow			SMURF	UCDA-Flow
				-	FFA-Net +	AECR-Net +	-	FFA-Net +	AECR-Net +		
LF-KITTI	EPE	23.48	6.82	14.33	14.21	11.66	13.42	13.15	10.06	10.48	**5.94**
	F1-all	81.54%	39.18%	56.96%	56.38%	50.92%	55.37%	54.83%	48.74%	47.60%	**34.11%**
DF-KITTI	EPE	25.32	8.03	16.55	15.97	12.16	15.84	14.93	11.21	11.56	**6.29**
	F1-all	85.77%	41.73%	62.84%	61.69%	53.17%	58.81%	57.06%	50.25%	51.39%	**36.25%**



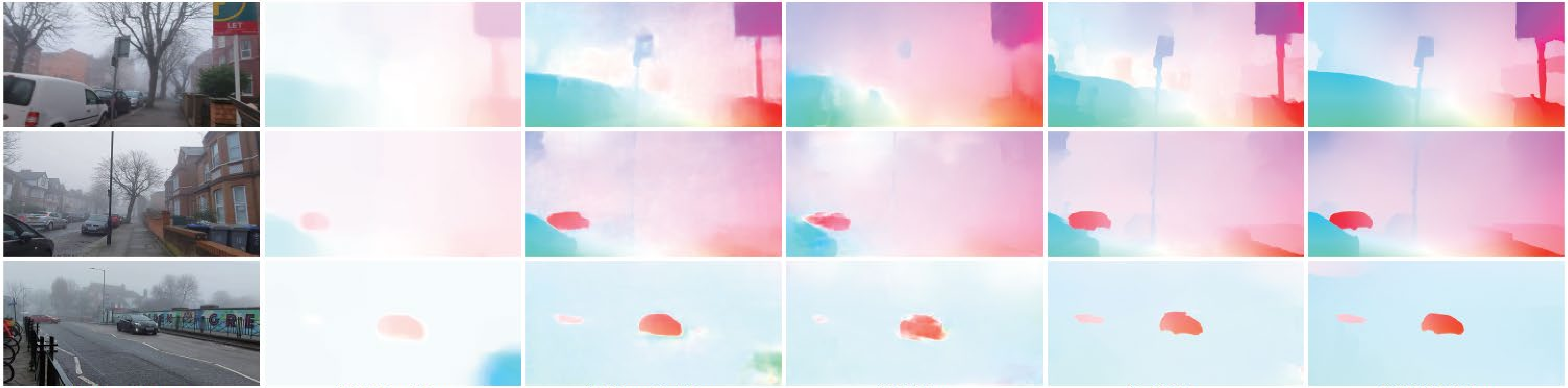
(a) Foggy images / GT (b) RobustFlow (c) UFlow (d) Defog + UFlow (e) Defog + Selflow (f) UCDA-Flow

Comparison on Real Foggy Images

Method		RobustFlow	UFlow	GMA		DenseFogFlow	GyroFlow	UCDA-Flow
				-	ssl*			
Fog-GOF	EPE	12.25	2.97	1.63	1.69	1.78	0.95	0.81
	F1-all	80.93%	30.82%	14.25%	15.11%	16.41%	9.13%	7.18
Dense-Fog	EPE	13.48	6.21	3.68	3.81	4.32	-	2.94
	F1-all	79.31%	62.45%	33.18%	35.20%	41.26%	-%	28.67%



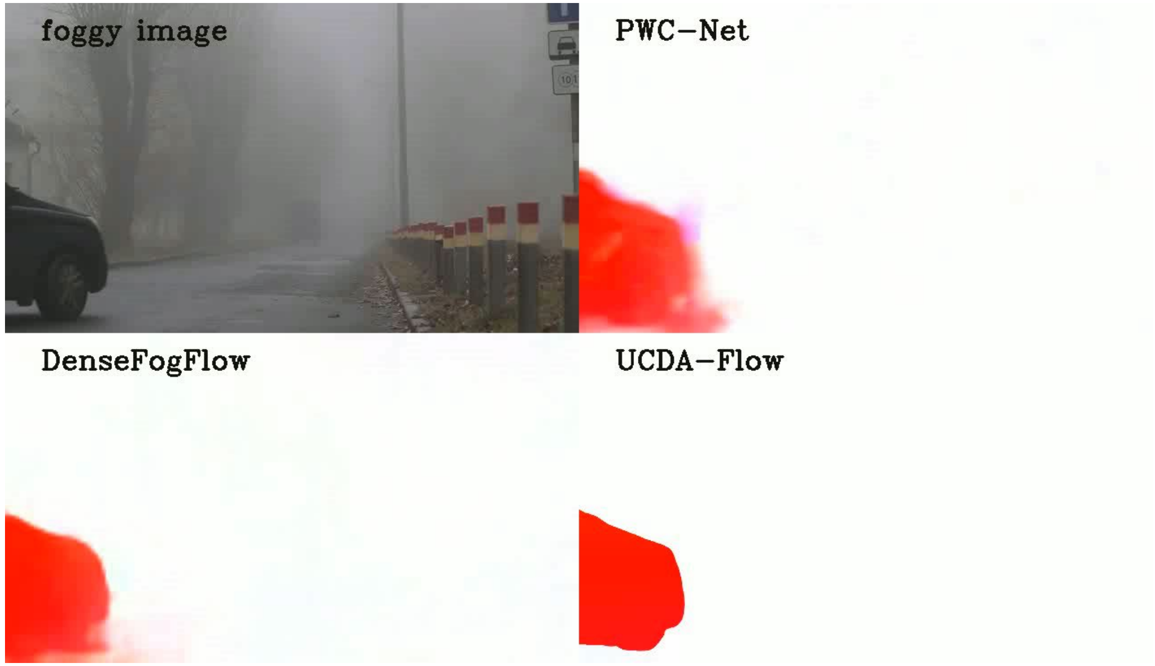
(a) Foggy images / GT (b) RobustFlow (c) DenseFogFlow (d) UFlow (e) GyroFlow (f) UCDA-Flow



(a) Foggy images (b) RobustFlow (c) DenseFogFlow (d) UFlow (e) GMA (f) UCDA-Flow

*: *'ssl' denotes self-supervised training strategy*

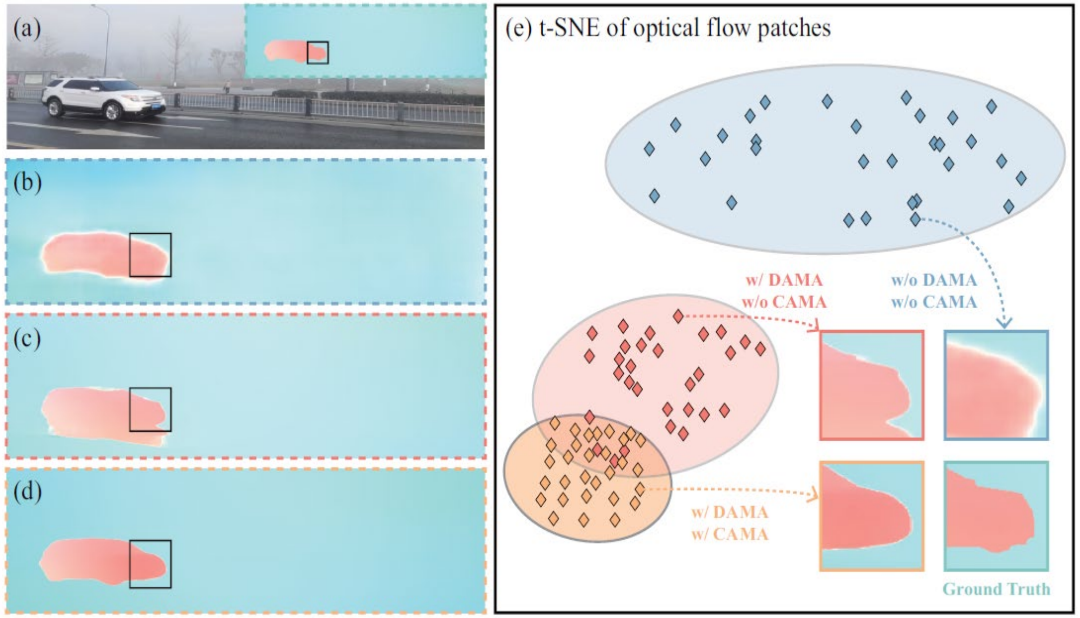
UCDA-Flow under Real Foggy Scenes



Static Scene

Dynamic Scene

Ablation Study

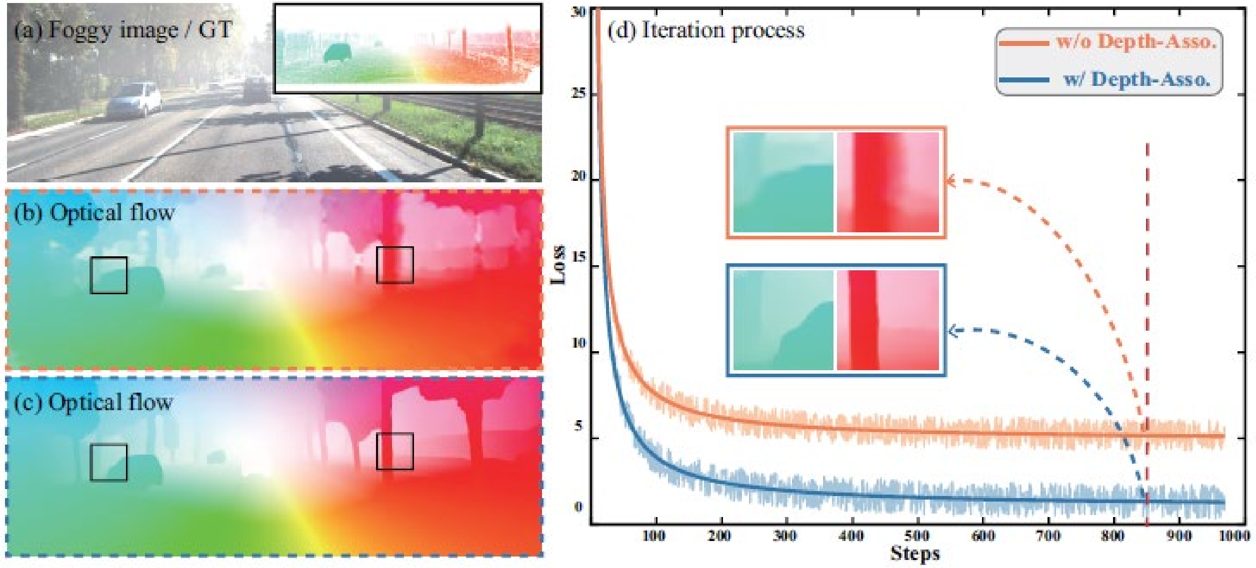


- Effectiveness of Cumulative Adaptation Architecture.

L_{flow}^{consis}	L_{flow}^{geo}	L_{flow}^{self}	L_{corr}^{kl}	EPE	F1-all
×	×	×	×	2.92	30.94%
×	√	×	×	2.88	30.20%
√	×	×	×	1.59	14.03%
√	√	×	×	1.35	11.27%
√	√	√	×	1.27	10.76%
√	√	×	√	0.92	8.81%
√	√	√	√	**0.81**	**7.18%**

- Effectiveness of Cumulative Adaptation Losses.

Discussion



- How dose the Depth Improve Optical Flow ?

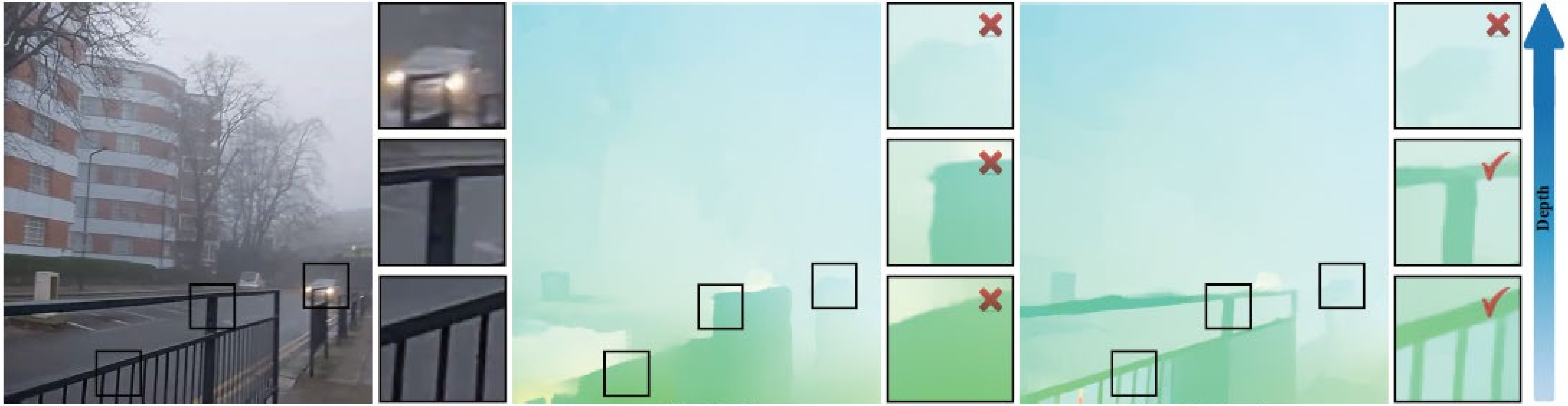
EMA	SCA	CDA	EPE	F1-all
×	×	×	1.38	12.06%
√	×	×	1.36	11.43%
√	√	×	1.27	10.76%
√	√	√	**0.81**	**7.18%**

- Effect of Modules in CAMA Stage on Optical Flow.

Method		EPE	F1-all
GAN-Based		1.43	13.10%
Depth-Based	Monocular	0.92	8.83%
	Pseudo-GT	0.83	7.45%
	Stereo (Ours)	**0.81**	**7.18%**

- Why Associate Depth with Fog ?

Limitation



(a) Foggy image (b) GMA (c) UCDA-Flow

- Compared with the state-of-the-art optical flow method GMA, UCDA-Flow obtains the clearer motion boundary in the nearby regions, but fails for the too-distant moving objects under foggy scenes.

Thanks