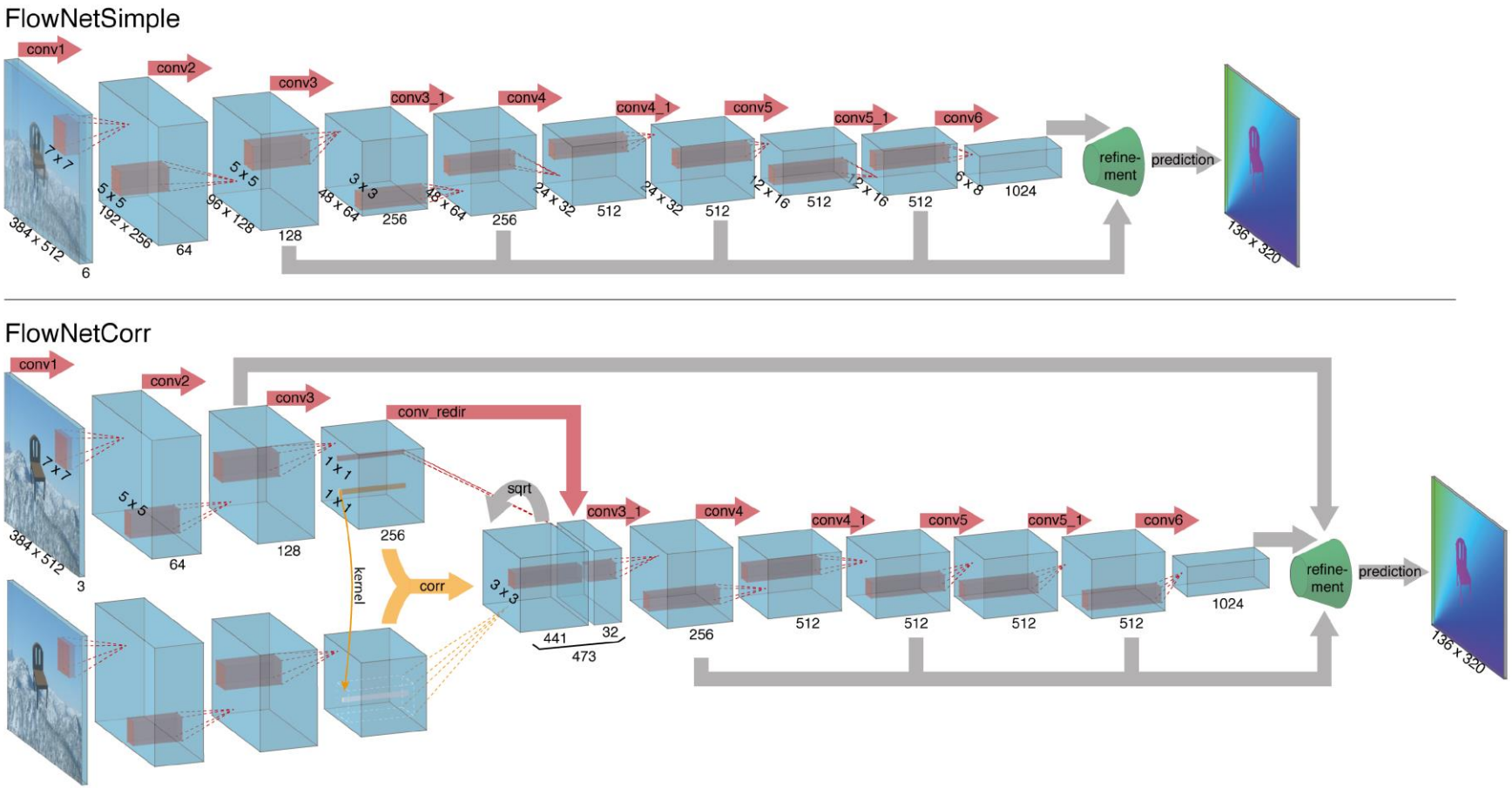
MaskFlownet: Asymmetric Feature Matching with Learnable Occlusion Mask

2020 CVPR Oral

Shengyu Zhao*, Yilun Sheng*, Yue Dong, Eric I-Chao Chang, Yan Xu

Optical Flow Estimation

FlowNet


FlowNetSimple
conv1
conv2
conv3
conv3_1
conv4
conv4_1
conv5
conv5_1
conv6
7 x 7
5 x 5
5 x 5
3 x 3
384 x 512
6
192 x 256
64
96 x 128
128
48 x 64
256
48 x 64
256
24 x 32
512
24 x 32
512
12 x 16
512
12 x 16
512
6 x 8
1024
refine-ment
prediction
136 x 320
FlowNetCorr
conv1
conv2
conv3
conv_redir
7 x 7
5 x 5
1 x 1
1 x 1
384 x 512
3
64
128
256
sqrt
kernel
corr
conv3_1
conv4
conv4_1
conv5
conv5_1
conv6
3 x 3
441
32
473
256
512
512
512
512
1024
refine-ment
prediction
136 x 320

FlyingChairs

FlowNet2

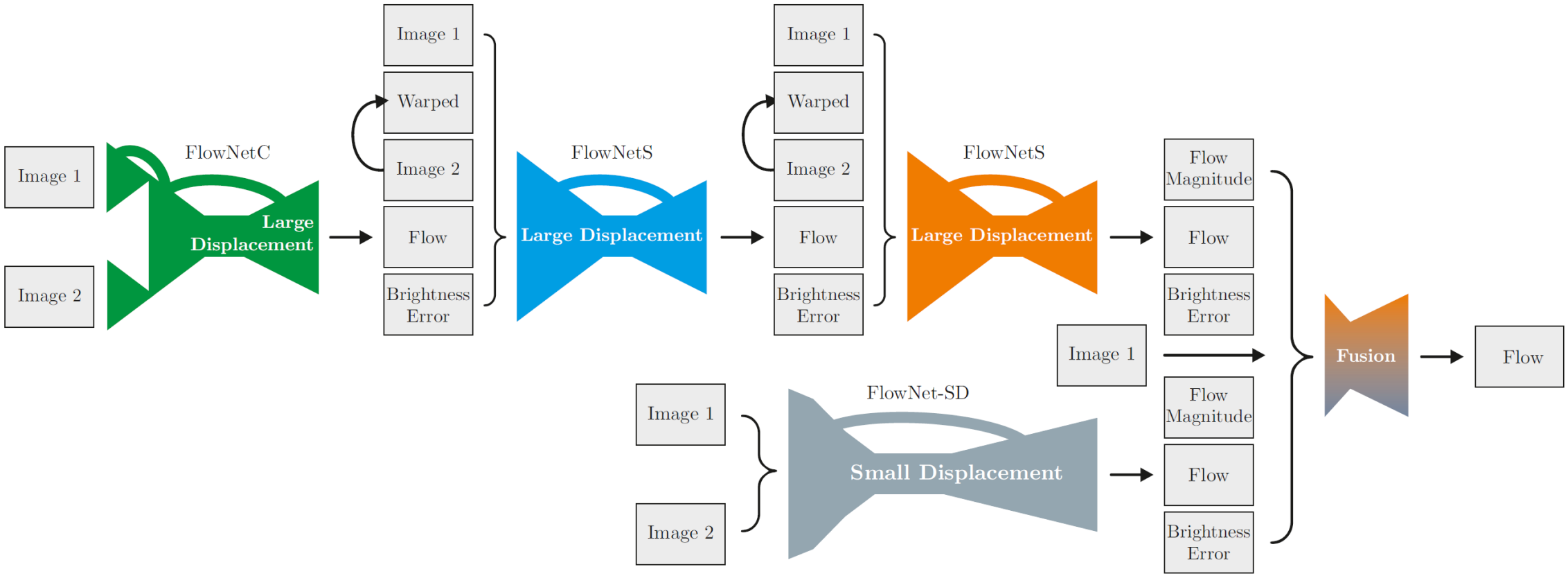
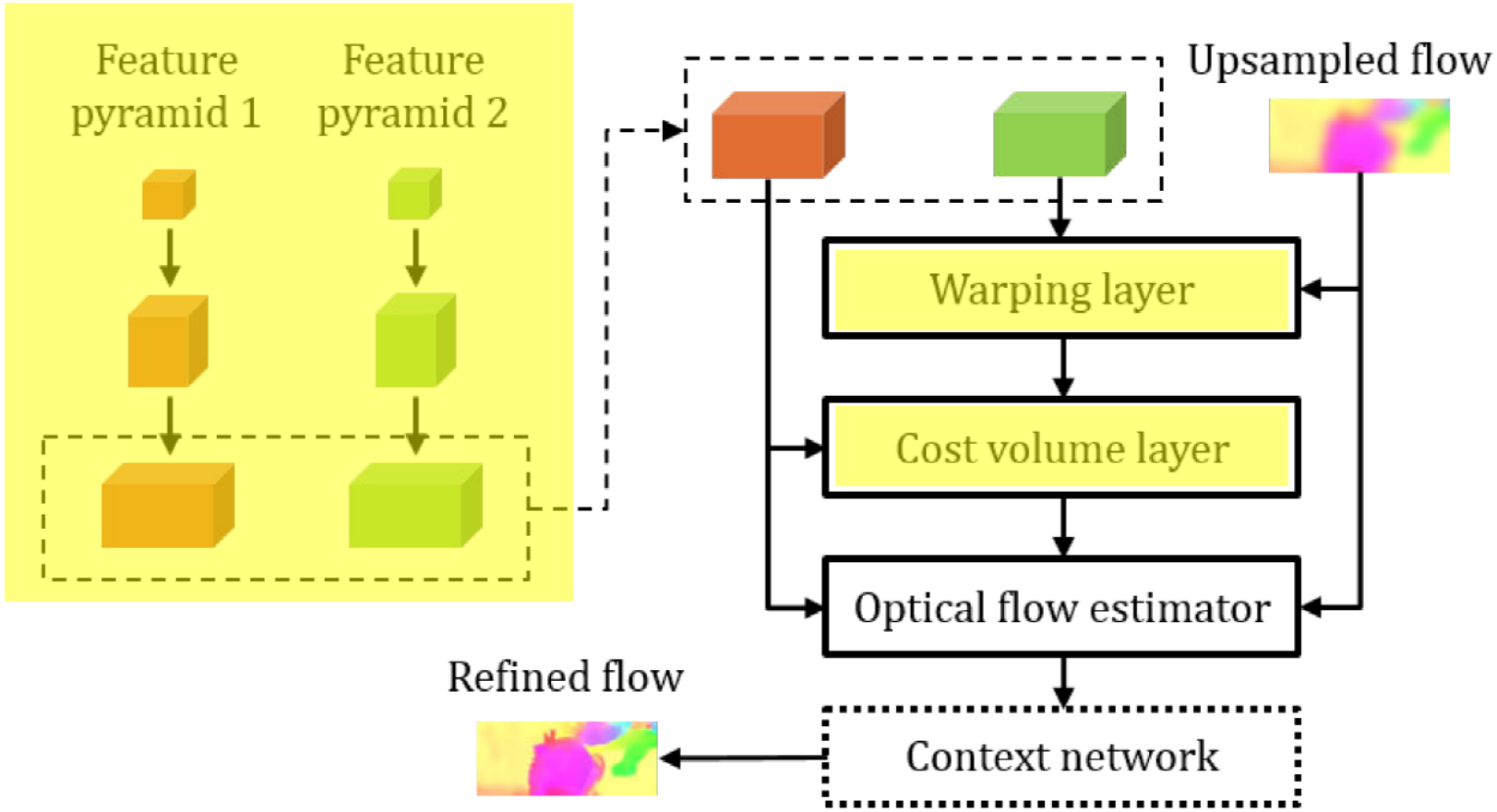


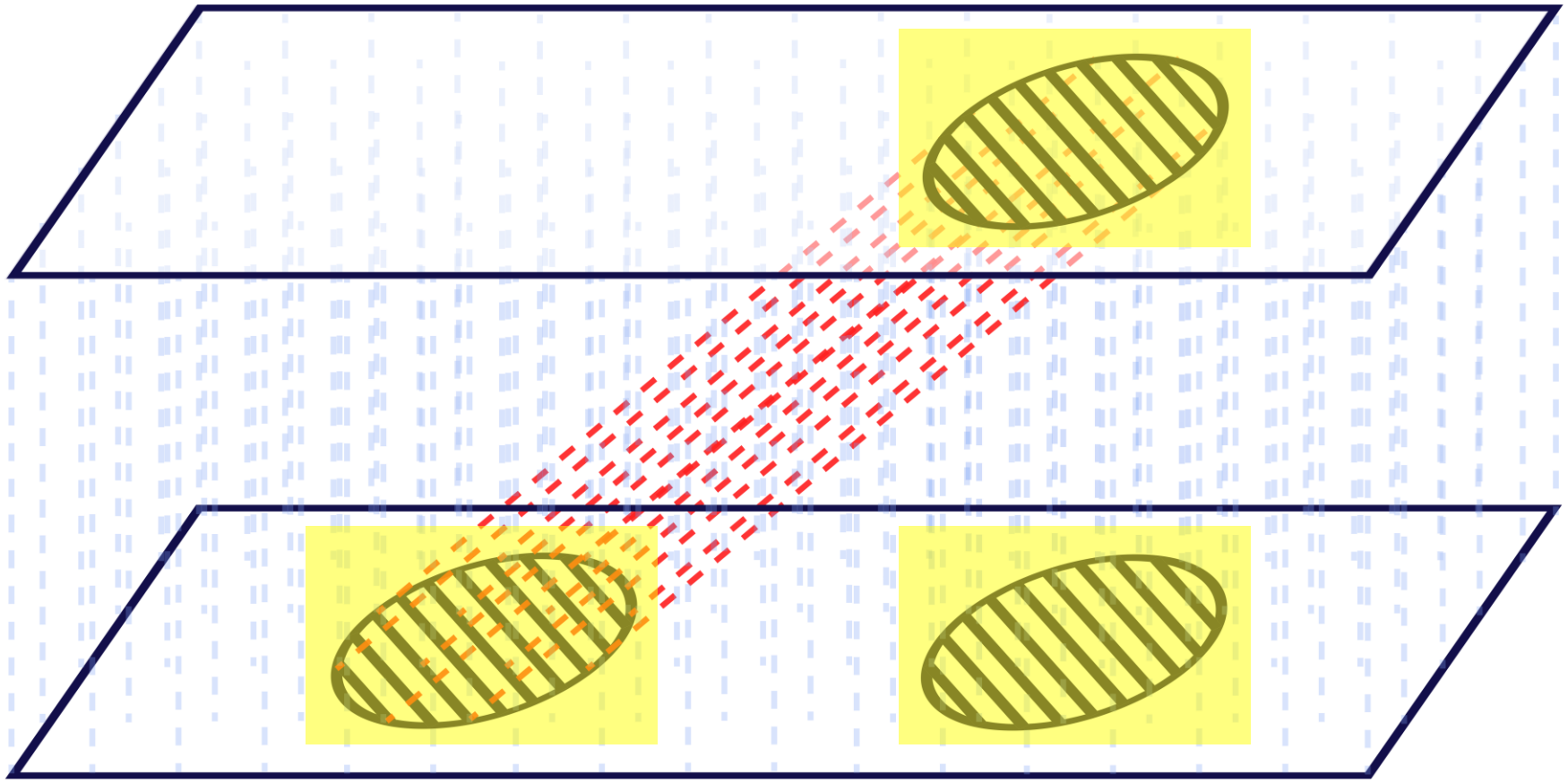
Image 1
Image 2
FlowNetC
Large Displacement
Image 1
Warped
Image 2
Flow
Brightness Error
FlowNetS
Large Displacement
Image 1
Warped
Image 2
Flow
Brightness Error
FlowNetS
Large Displacement
Flow Magnitude
Flow
Brightness Error
Image 1
Image 1
Image 2
FlowNet-SD
Small Displacement
Flow Magnitude
Flow
Brightness Error
Fusion
Flow

PWC-Net



Sun, Deqing, et al. "Pwc-net: Cnns for optical flow using pyramid, warping, and cost volume." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.

Motivation



Motivation

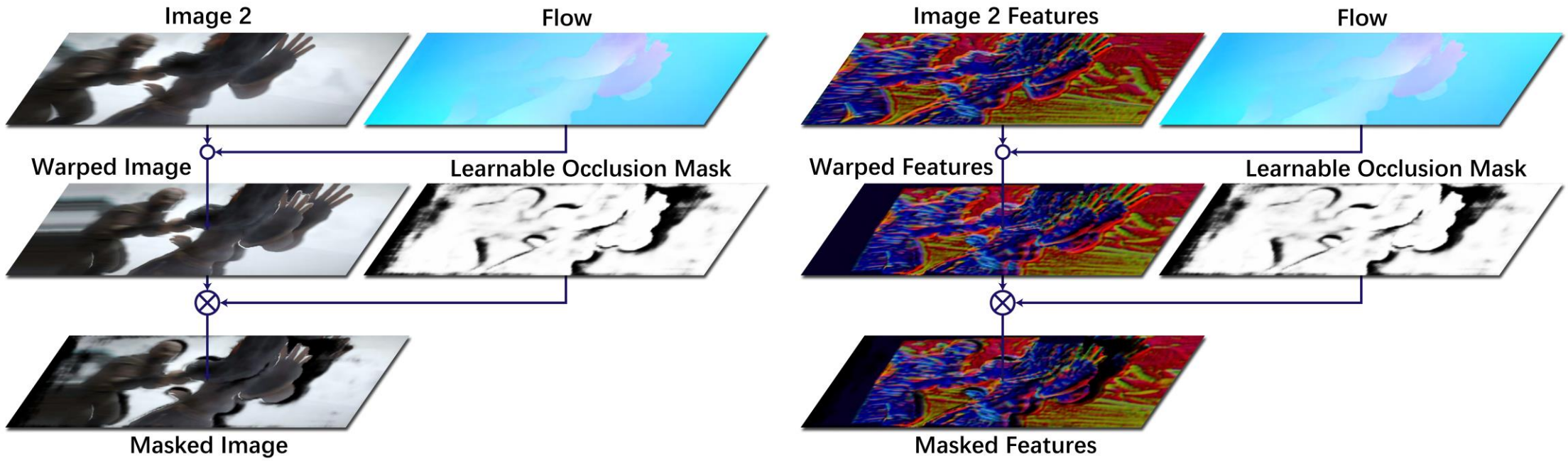
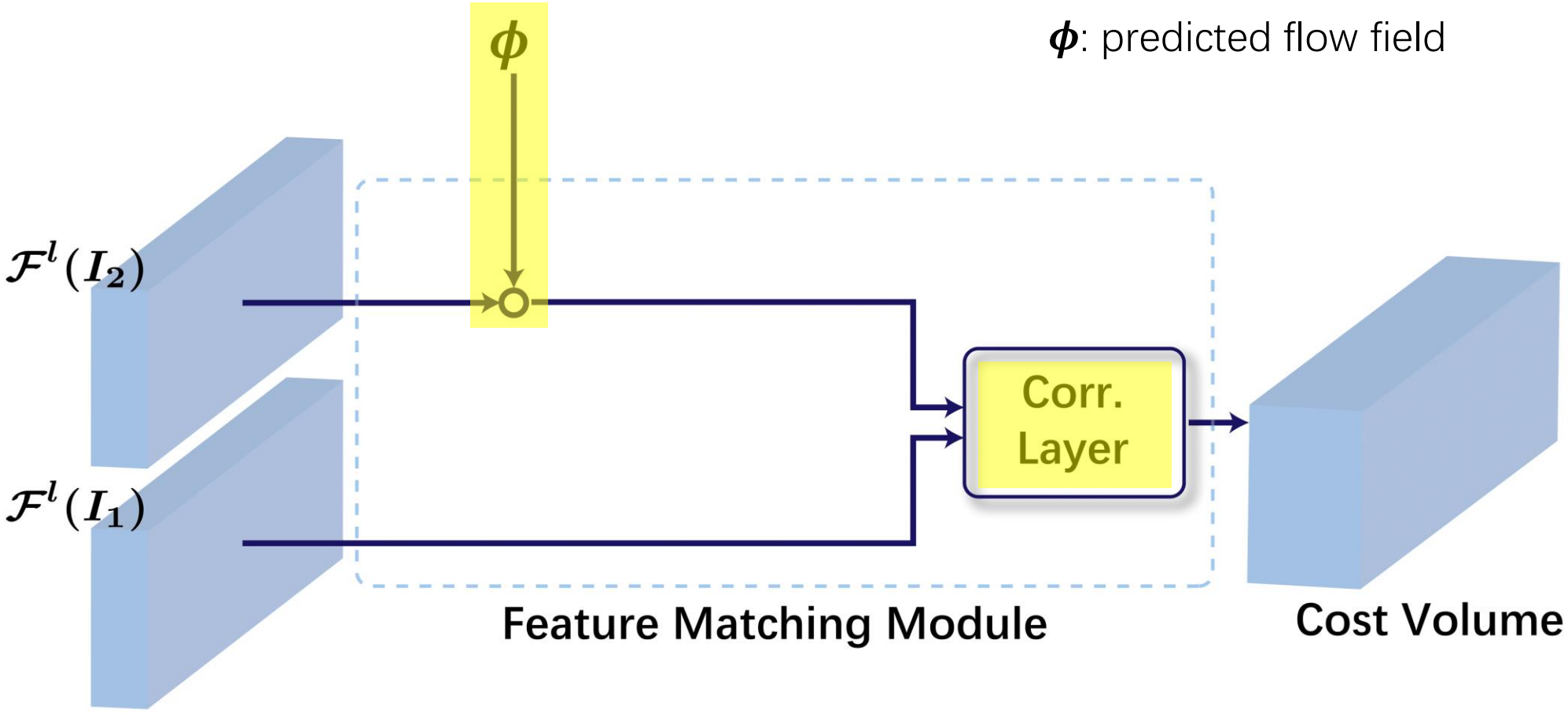


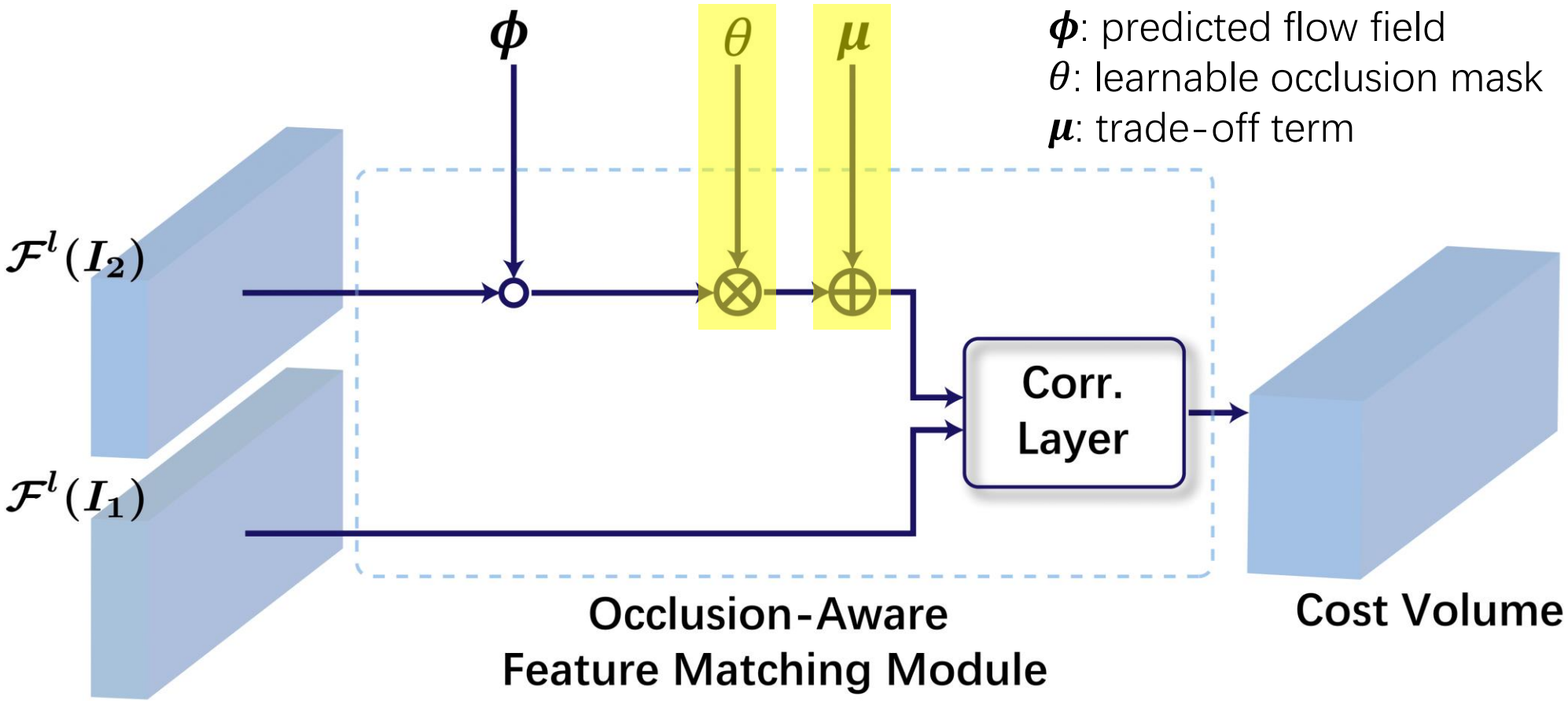
Image 2
Flow
Warped Image
Learnable Occlusion Mask
Masked Image
Image 2 Features
Flow
Warped Features
Learnable Occlusion Mask
Masked Features

Occlusion-Aware Feature Matching

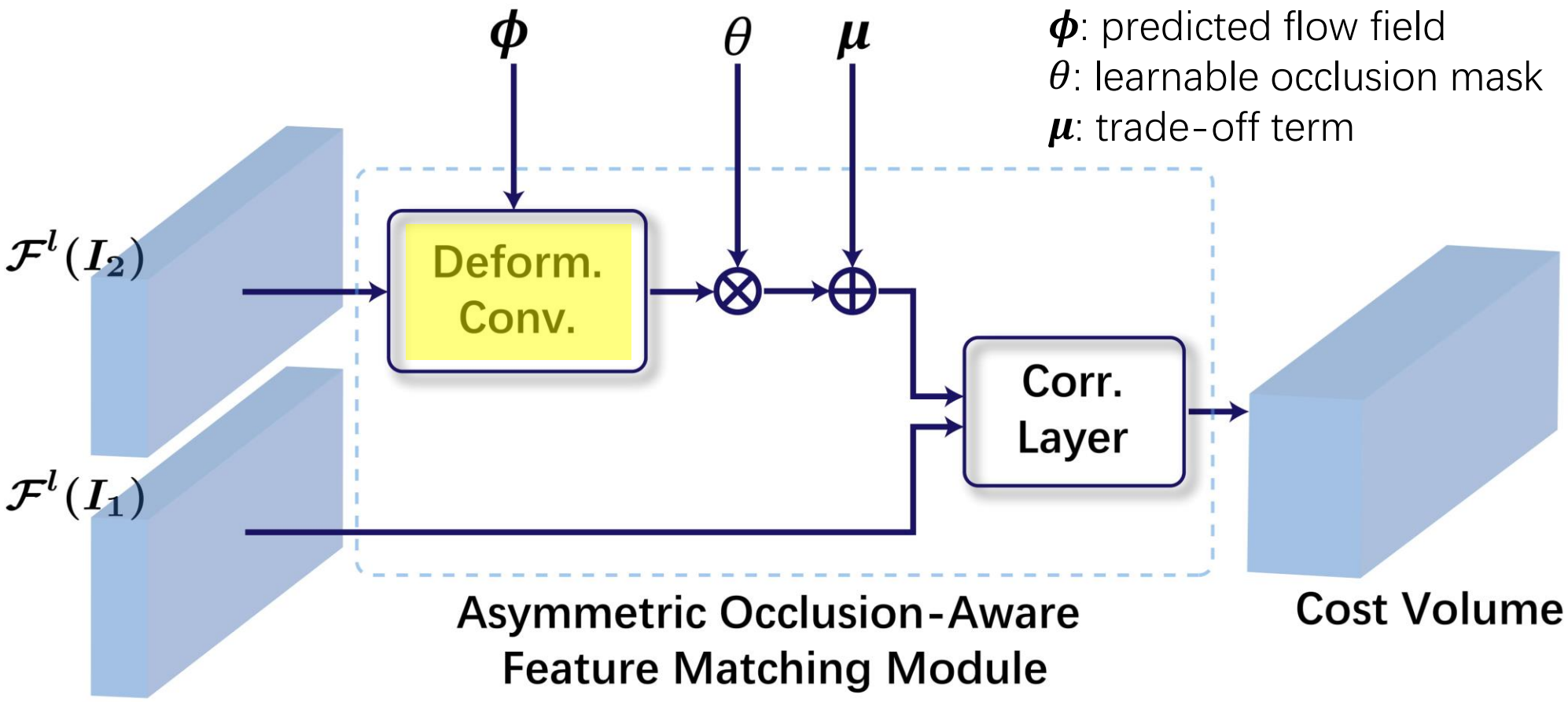
Feature Matching Module (FMM)



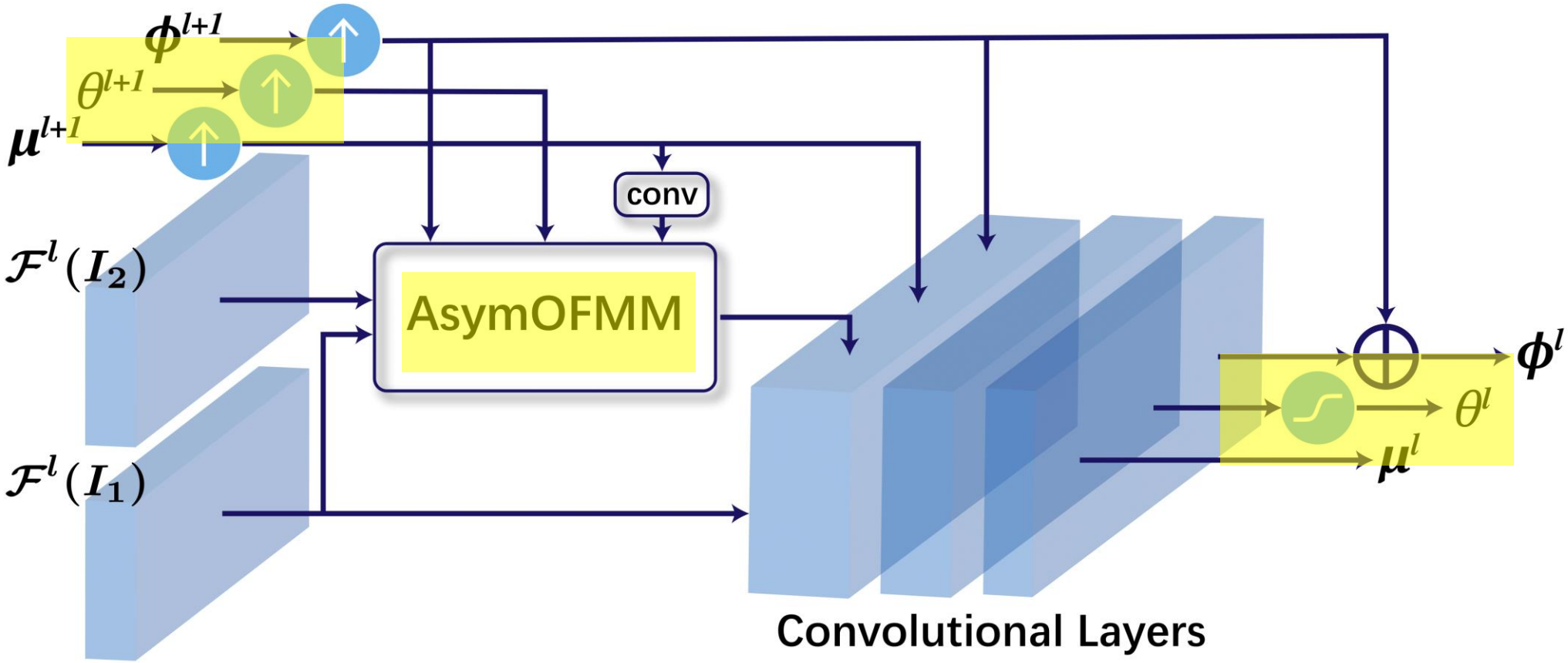
Occlusion-aware Feature Matching Module (OFMM)



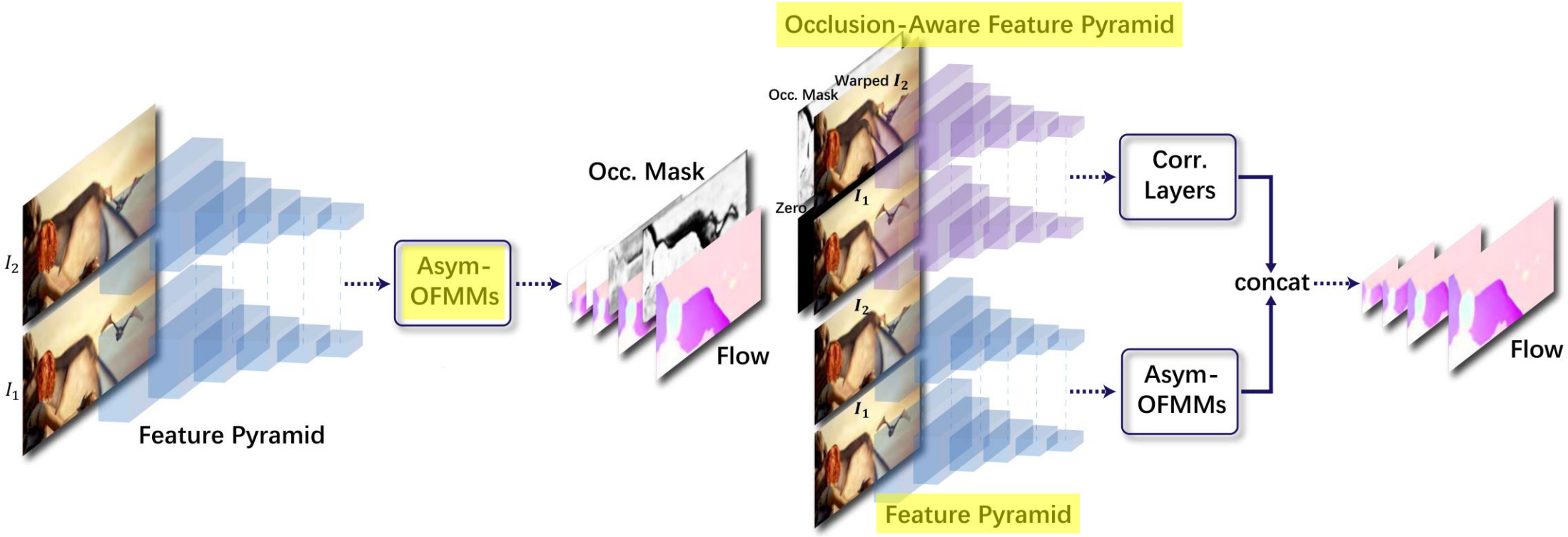
Asymmetric Occlusion-Aware Feature Matching Module (AsymOFMM)



Network Connections



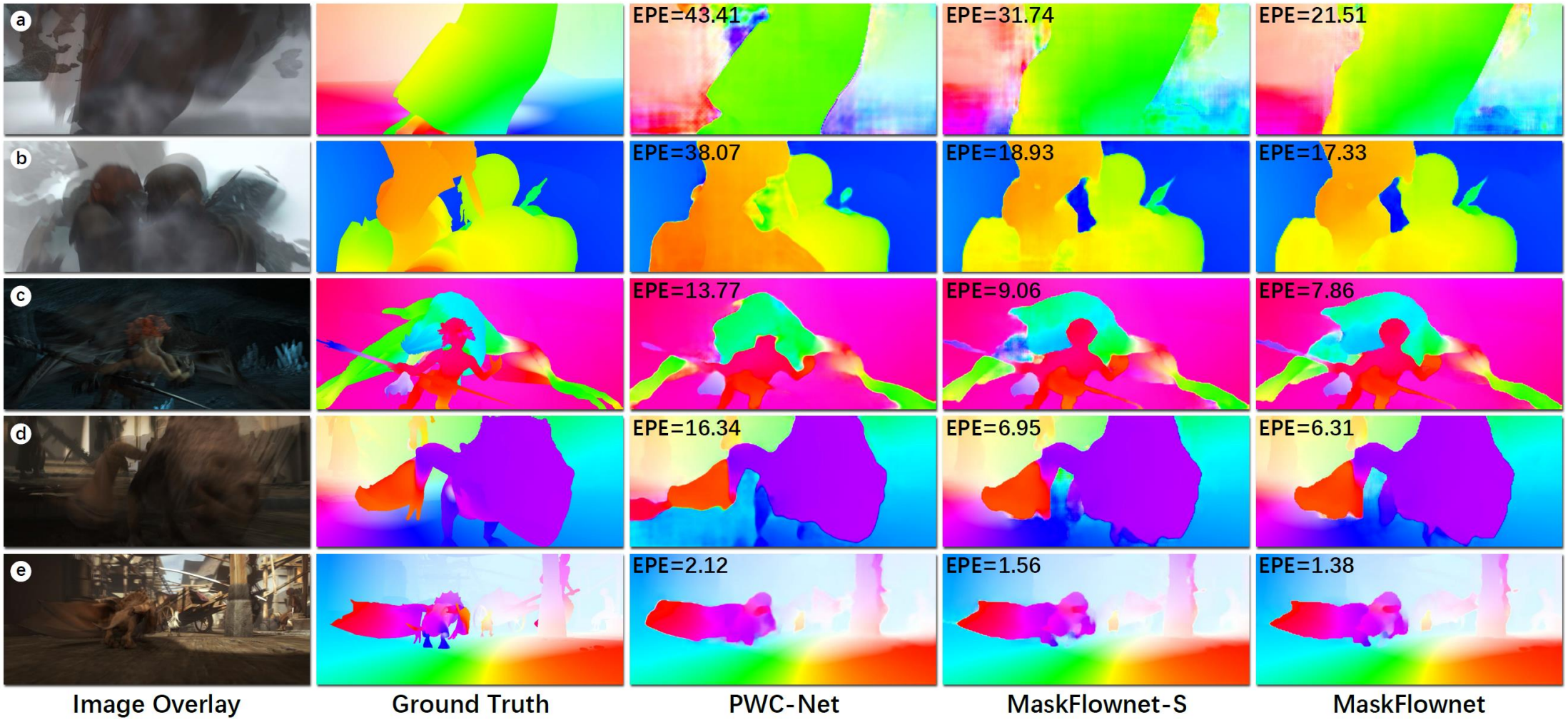
Overall Architecture



Main Results

Method	Time	Sintel *clean*		Sintel *final*		KITTI 2012		KITTI 2015	
	(s)	AEPE *train*	AEPE *test*	AEPE *train*	AEPE *test*	AEPE *train*	AEPE *test*	Fl-all *train*	Fl-all *test*
FlowNetS [8]	0.01	3.66	6.16	4.76	7.22	6.07	7.6	-	-
FlowNetC [8]	0.05	3.57	6.08	5.25	7.88	7.31	-	-	-
FlowNet2 [14]	0.12	**2.02**	3.96	**3.14**	6.02	4.09	1.8	28.20%	11.48%
SpyNet [26]	0.16	4.12	6.64	5.57	8.36	9.12	4.1	-	35.07%
MR-Flow [37]	480	-	2.53	-	5.38	-	-	-	12.19%
LiteFlowNet [11]	0.09	2.48	4.54	4.04	5.38	4.00	1.6	28.50%	9.38%
LiteFlowNet2 [12]	0.04	2.24	3.45	3.78	4.90	3.42	1.4	25.88%	7.74%
PWC-Net [31]	0.03	2.55	3.86	3.93	5.13	4.14	1.7	33.67%	9.60%
PWC-Net+ [30]	0.03	-	3.45	-	4.60	-	1.5	-	7.90%
SelFlow [18]	0.09	-	3.74	-	4.26	-	1.5	-	8.42%
VCN [39]	0.03	2.21	2.81	3.62	4.40	-	-	25.1%	6.30%
MaskFlownet-S	0.03	2.33	2.77	3.72	4.38	3.21	**1.1**	23.58%	6.81%
MaskFlownet	0.06	2.25	**2.52**	3.61	**4.17**	**2.94**	**1.1**	**23.14%**	**6.11%**

Qualitative Comparison



Main Result

Ablation Study

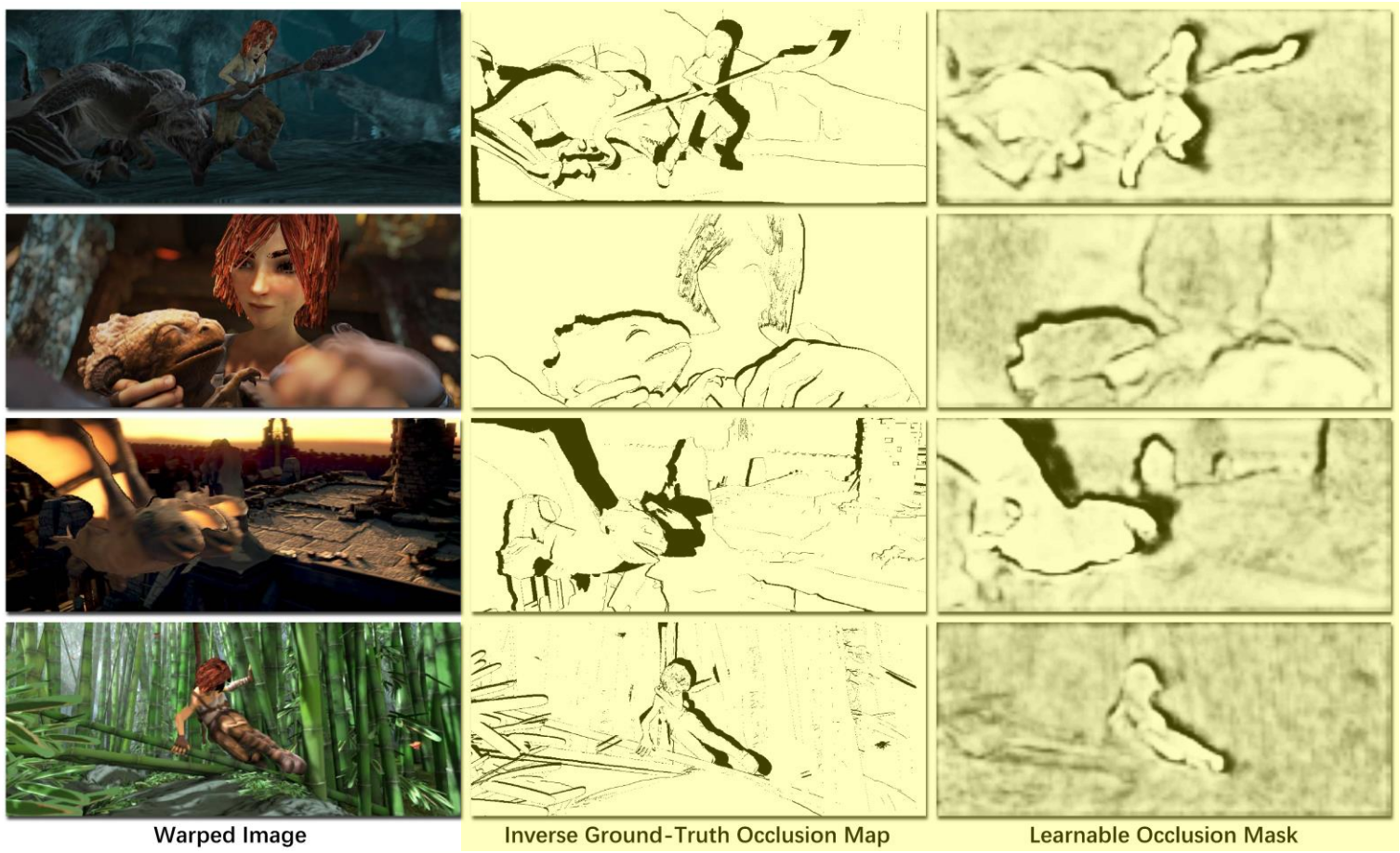
Module	Trained on Chairs			Things3D		Sintel	
	Chairs	Sintel (*train*)		Sintel (*train*)		Sintel (*val*)	
	test	*clean*	*final*	*clean*	*final*	*clean*	*final*
FMM	1.61	3.25	4.59	2.55	4.05	3.02	4.70
OFMM	1.62	3.20	4.50	2.52	4.01	3.06	4.52
AsymOFMM	**1.56**	**2.88**	**4.25**	**2.33**	**3.72**	**2.70**	**4.07**

Network	Tuned on Sintel	
	Sintel (*val*)	
	clean	*final*
MaskFlownet-S	2.70	4.07
+ single pyramid w/o mask	2.53	3.90
+ single pyramid w/ mask	2.55	3.88
+ dual pyramids w/o mask	**2.52**	3.85
+ dual pyramids w/ mask	**2.52**	**3.83**

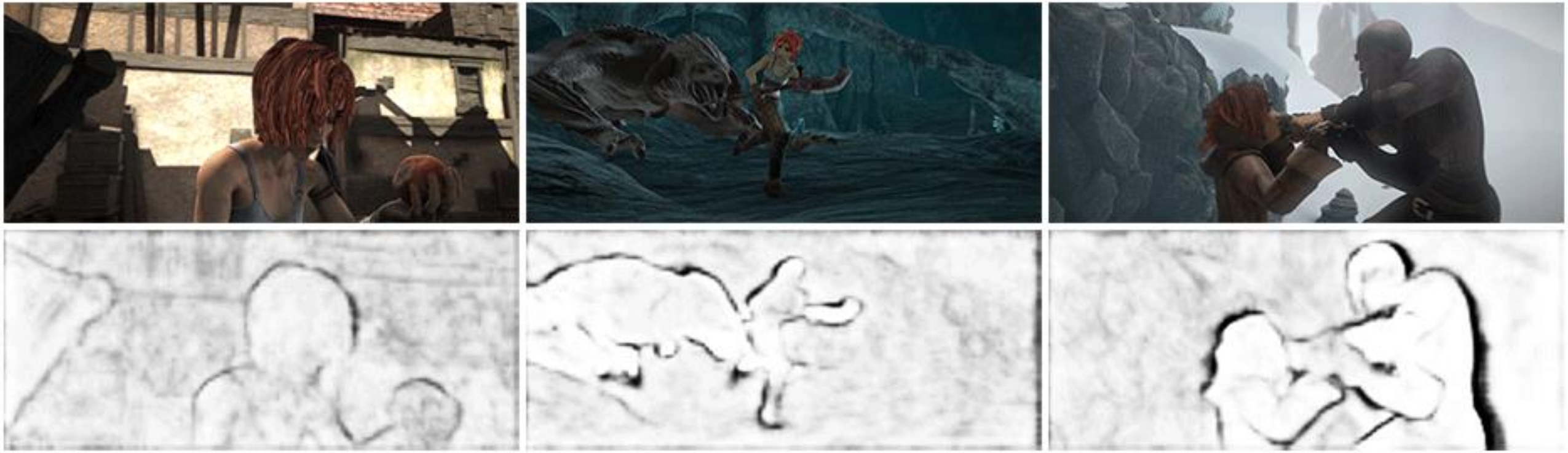
Module	Trained on Chairs			Things3D	
	Chairs	Sintel (*train*)		Sintel (*train*)	
	test	*clean*	*final*	*clean*	*final*
OFMM	1.62	3.20	4.50	2.52	4.01
+ sym-conv	1.61	3.33	4.64	2.54	3.84
+ asym-conv	**1.52**	2.96	4.29	2.41	3.85
+ deform-conv	1.56	**2.88**	**4.25**	**2.33**	**3.72**

Module (AsymOFMM)	Trained on Chairs		
	Chairs	Sintel (*train*)	
	test	*clean*	*final*
w/o mask w/o trade-off	1.58	3.08	4.29
w/ mask w/o trade-off	1.60	3.06	4.32
w/o mask w/ trade-off	1.58	2.97	4.30
(w/ mask w/ trade-off)	**1.56**	**2.88**	**4.25**

Learnable Occlusion Mask



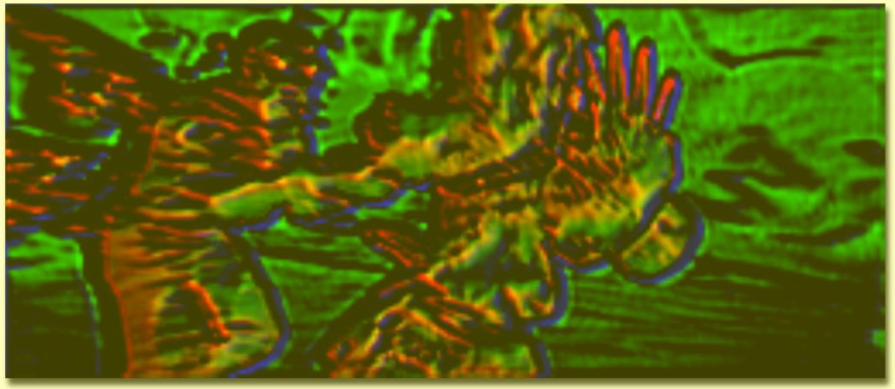
Learnable Occlusion Mask



Asymmetricity

Image Overlay

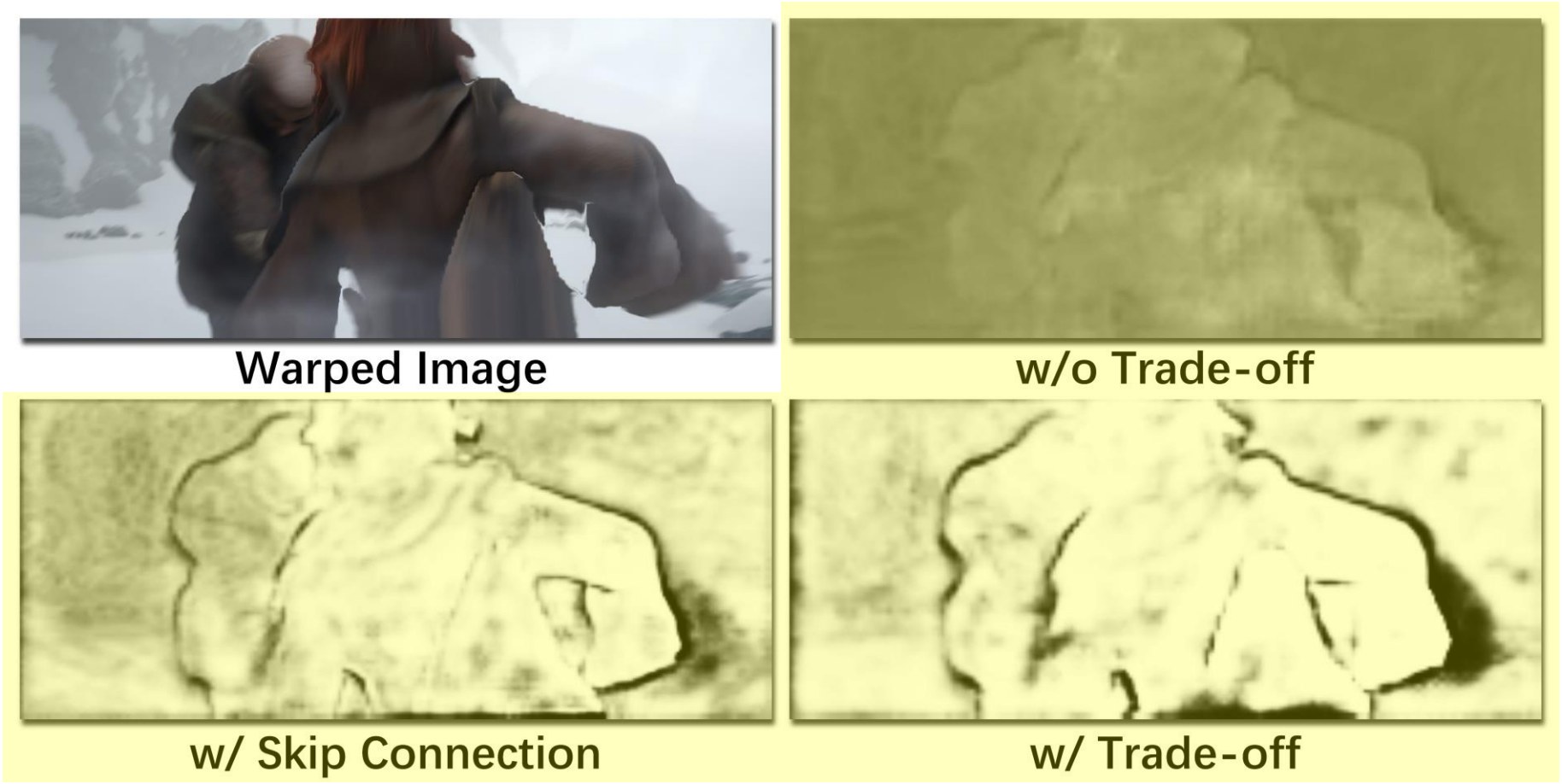
Masked Features



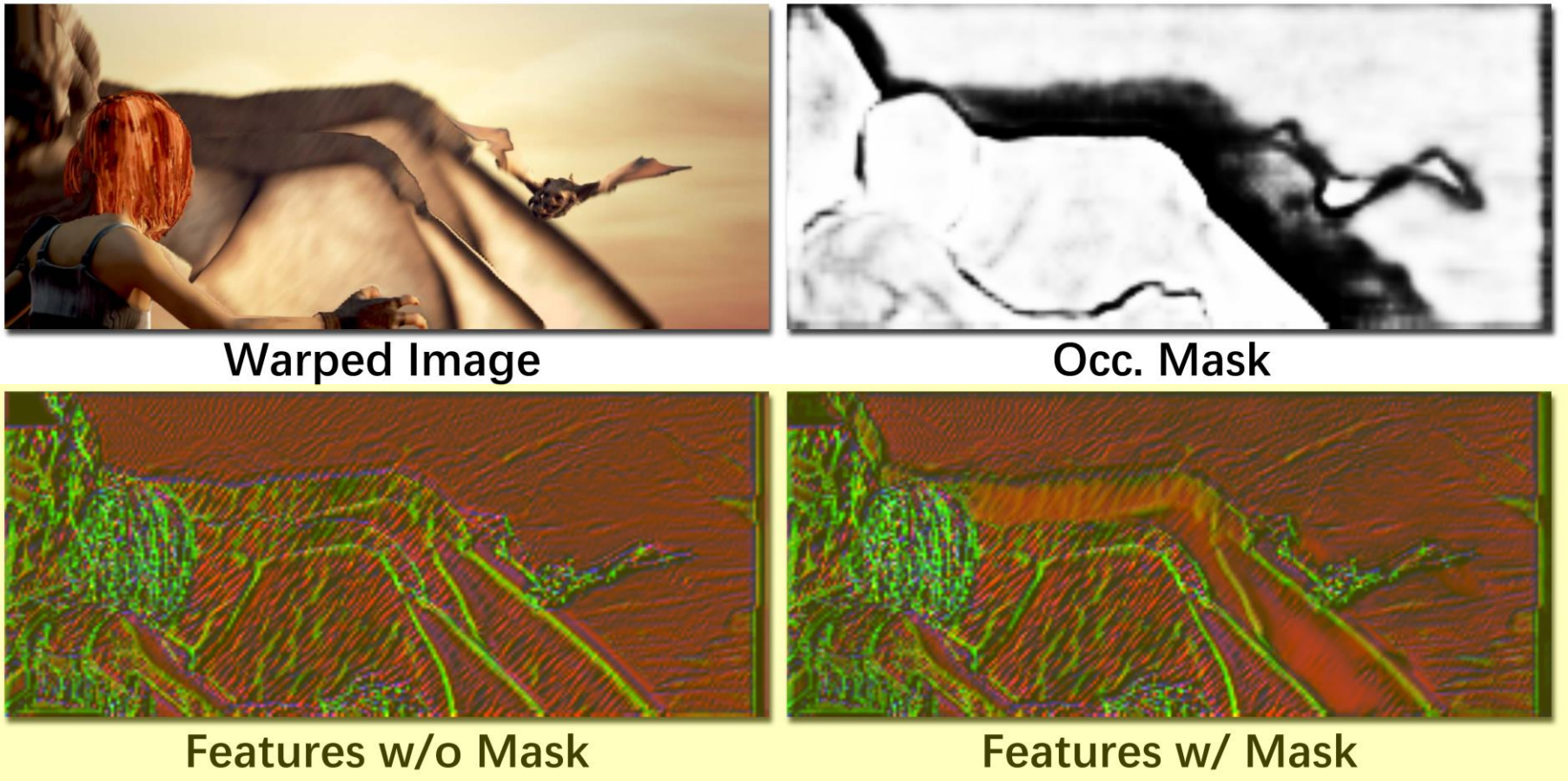
Source Features

Target Features

The Trade-off Term

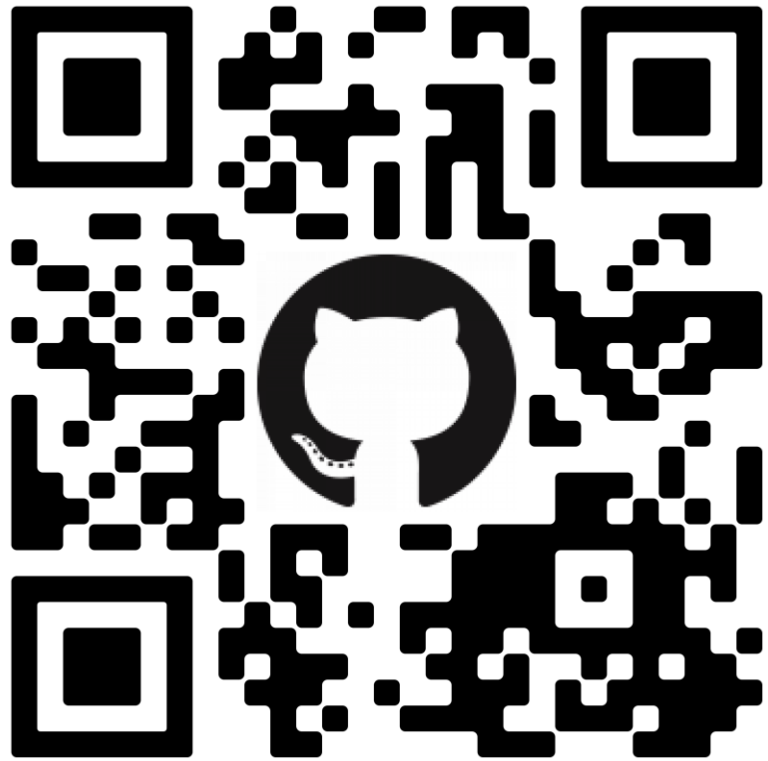


Occlusion-aware Feature Pyramid



Conclusion

- Learning occlusion mask without any explicit supervision.
- Asymmetric design of feature matching module.
- Easily integrable and nearly costless.
- New perspective on how to deal with occlusions for both supervised and unsupervised optical flow estimation.



https://github.com/microsoft/MaskFlownet

https://www.msra.cn/zh-cn/news/features/cvpr-2020-maskflownet