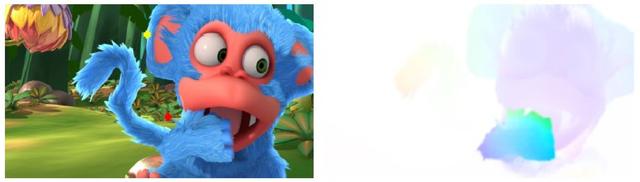
OAS-Net: Occlusion Aware Sampling Network for Accurate Optical Flow

IEEE ICASSP 2021
Paper Number: 1070

Lingtong Kong, Xiaohang Yang, Jie Yang
Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University

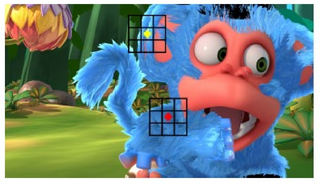
Motivation:

Existing deep optical flow networks have achieved satisfactory results by mostly employing a pyramidal coarse-to-fine paradigm, where a key process is to adopt warped target feature based on previous flow prediction to correlate with source feature for building 3D matching cost volume. However, the warping operation can lead to troublesome ghosting problem that results in ambiguity. Moreover, occluded areas are treated equally with non occluded regions in most existing works, which may cause performance degradation.

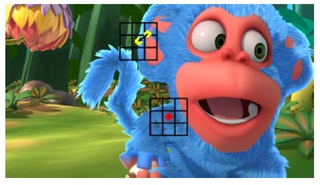


(a) Source Image

(b) Previous Optical Flow Prediction



(c) Warping Based Cost Volume



(d) Sampling Based Cost Volume

Innovations:

We propose a lightweight yet efficient optical flow network, named **OAS-Net** (occlusion aware sampling network) for accurate optical flow with following contributions:

1. A new **sampling based correlation layer** is employed without noisy warping operation.
2. A novel **occlusion aware module** is presented to make raw cost volume conscious of occluded regions.
3. A **shared flow and occlusion awareness decoder** is adopted for structure compactness.

Sampling Based Correlation:

Previous warping based cost volume suffers from duplicate artifacts, which is also known as ghosting. To solve the problem, we build cost volume with a novel sampling based approach that avoids warping and ghosting issues. Original flow based warping operation is replaced by directly sampling flow guided searching grids in the target pyramid feature. These two types of correlation can be formulated as:

$$c^k(\mathbf{x},\mathbf{d}) = f_1^k(\mathbf{x}) \cdot f_{warp}^k(\mathbf{x}+\mathbf{d})$$

$$c^k(\mathbf{x},\mathbf{d}) = f_1^k(\mathbf{x}) \cdot f_2^k(\mathbf{x}+\mathbf{f}+\mathbf{d})$$

Note that the yellow cross position in the right figure becomes distinct when using proposed sampling based correlation operation.

Occlusion Aware Module:

Although sampling based cost volume can bypass annoying ghosting feature, occluded regions are unrecoverable no matter to take either correlation methods. As shown in the left figure, the red cross pointing bush in the source image is covered by the moving arm in the target image. To endow flow estimation with occlusion awareness, we present a novel occlusion aware module to better handle two types of matching regions in raw cost volume separately, which can be formulated as:

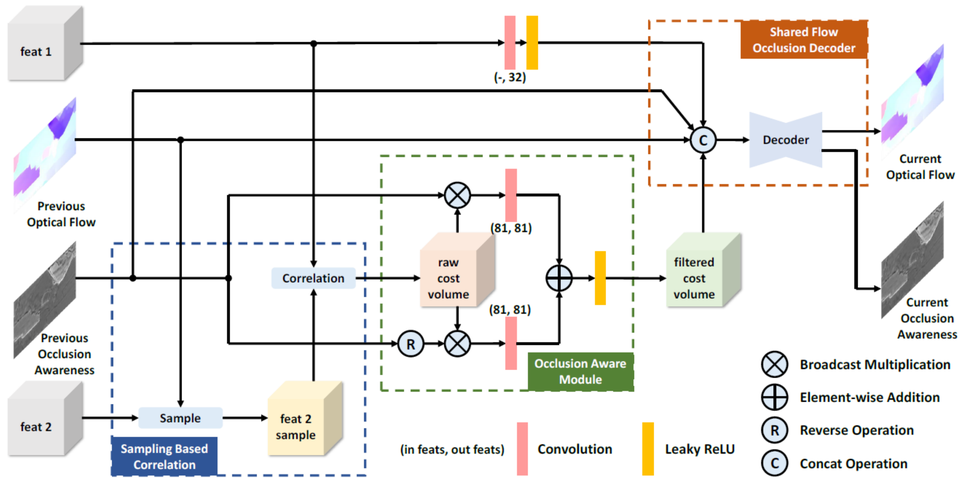
$$c_{oa}^k(\mathbf{x},\mathbf{d}) = \mathrm{lrelu}(\mathrm{conv}_1(O(\mathbf{x}) \otimes c^k(\mathbf{x},\mathbf{d})) \oplus \mathrm{conv}_2((\mathbf{1}-O(\mathbf{x})) \otimes c^k(\mathbf{x},\mathbf{d})))$$

Proposed occlusion awareness map can also be interpreted as occlusion probability map, which can be viewed as one type of self-attention mechanism.

Ablation Study:

Correlation Method	Occlusion Awareness	Sintel Final	KITTI 2012
Warping		4.05	4.62
Warping	✓	3.98	4.37
Sampling		3.86	4.44
Sampling	✓	**3.79**	**4.11**

Overall Structure:



OAS-Net estimates optical flow in a coarse-to-fine manner and adopt a shared flow and occlusion awareness decoder for structure compactness.

Benchmark Results:

Method	Sintel Clean train	Sintel Clean test	Sintel Final train	Sintel Final test	KITTI 2012 train	KITTI 2012 test	Parameters (M)	Time (s)
FlowNetC [4]	4.31	6.85	5.87	8.51	9.35	-	39.18	0.050
SPyNet [7]	4.12	6.64	5.57	8.36	9.12	4.7	**1.20**	0.055
FlowNet2 [5]	**2.02**	4.16	**3.14**	5.74	4.09	1.8	162.49	0.120
LiteFlowNet [8]	2.48	4.54	4.04	5.38	**4.00**	1.6	5.37	0.055
PWC-Net [6]	2.55	4.39	3.93	5.04	4.14	1.7	8.75	0.035
IRR-PWC [9]	-	3.84	-	**4.58**	-	1.6	6.36	0.150
HD3-Flow [12]	3.84	4.79	8.77	4.67	4.65	**1.4**	39.56	0.080
Devon [11]	-	4.34	-	6.35	-	2.6	-	0.050
OAS-Net (Ours)	2.55	**3.65**	3.79	5.01	4.11	**1.4**	6.16	**0.030**

OAS-Net performs better than all the other approaches on Sintel Clean test dataset and surpasses PWC-Net on Sinel Final test set slightly. It achieves the same best result as HD3-Flow on KITTI 2012 test benchmark and excels LiteFlowNet and IRR-PWC.

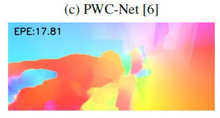
Visual Samples:

(a) First Image

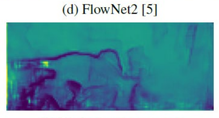
(b) Ground Truth

(c) PWC-Net [6]

(d) FlowNet2 [5]



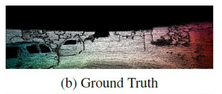
(e) OAS-Net



(f) Occlusion Awareness

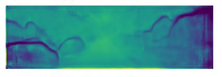
It can be seen that our occlusion awareness maps have correctly emphasized the probable occluded regions, and help to improve optical flow accuracy.

(a) First Image



(b) Ground Truth

(c) Optical Flow



(d) Occlusion Awareness