Selective-Stereo: Adaptive Frequency Information Selection for Stereo Matching

Supplementary Material

Method	KITTI 2012 (Reflective Regions)			
	2-noc	2-all	3-noc	3-all
AcfNet [42]	11.17	13.13	6.93	8.52
LEAStereo [9]	9.66	11.40	5.35	6.50
ACVNet [35]	11.42	13.53	7.03	8.67
RAFT-Stereo [17]	8.41	9.87	5.40	6.48
PCWNet [25]	8.94	10.71	4.99	6.20
LaC + GANet [18]	10.40	12.21	6.02	7.34
CREStereo [16]	9.71	11.26	6.27	7.27
IGEV-Stereo [34]	7.29	8.48	4.11	4.76
Selective-RAFT (Ours)	7.19	7.96	4.35	4.68
Selective-IGEV (Ours)	**6.73**	**7.84**	**3.79**	**4.38**

Table 9. Quantitative evaluation on KITTI 2012 reflective regions (ill-posed regions).

6. Results on Ill-posed Regions

To demonstrate the ability of our method to handle ill-posed regions, we evaluate it on KITTI 2012 reflective regions. As shown in Tab. 9, our Selective-RAFT outperforms RAFT-Stereo [17] by almost 20% and our Selective-IGEV ranks 1^{st} among all publish methods. It shows that our method can fuse information at different frequencies to overcome ill-posed regions. Specifically, large kernels capture global low-frequency information to tackle local ambiguities, while small kernels capture local high-frequency information to maintain detailed structures. When encountering reflective regions, weights of large kernels increase to fuse more global information, while the network still fuses local information captured by small kernels to maintain details.