

SfM-TTR: Using Structure from Motion for Test-Time Refinement of Single-View Depth Networks (Supplementary Material)

Sergio Izquierdo Javier Civera
I3A, University of Zaragoza, Spain
{izquierdo, jcivera}@unizar.es

1. Results on Sequences without SfM

Tables 1, 2 and 3 show results of the KITTI evaluation considering only sequences with SfM convergence. Specifically, on all sequences from the Eigen test split *except* of:

- 2011_09_26_drive_0020_sync,
- 2011_09_26_drive_0048_sync,
- 2011_09_26_drive_0052_sync,
- 2011_09_26_drive_0056_sync, and
- 2011_10_03_drive_0047_sync.

References

- [1] Shariq Farooq Bhat, Ibraheem Alhashim, and Peter Wonka. AdaBins: Depth Estimation using Adaptive Bins. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4009–4018, 2021. 2
- [2] Jamie Watson, Oisin Mac Aodha, Victor Prisacariu, Gabriel Brostow, and Michael Firman. The temporal opportunist: Self-supervised multi-frame monocular depth. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1164–1174, 2021. 2
- [3] Jiaying Yan, Hong Zhao, Penghui Bu, and YuSheng Jin. Channel-wise attention-based network for self-supervised monocular depth estimation. In *2021 International Conference on 3D Vision (3DV)*, pages 464–473. IEEE, 2021. 2
- [4] Hang Zhou, David Greenwood, and Sarah Taylor. Self-supervised monocular depth estimation with internal feature fusion. In *British Machine Vision Conference (BMVC)*, 2021. 2

TTR	Method	Abs Rel ↓	Sq Rel ↓	RMSE ↓	RMSE log ↓	$\delta < 1.25 \uparrow$	$\delta < 1.25^2 \uparrow$	$\delta < 1.25^3 \uparrow$
✗	AdaBins [1]	0.075	0.345	3.238	0.114	0.937	0.989	0.997
✓	AdaBins [1] + SfM-TTR	0.061	0.204	2.208	0.091	0.970	0.993	0.998
✗	ManyDepth [2]	0.060	0.326	3.096	0.099	0.955	0.992	0.997
✓	ManyDepth [2] + Ph-TTR	0.053	0.263	2.852	0.091	0.964	0.992	0.998
✓	ManyDepth [2] + SfM-TTR	0.054	0.270	2.571	0.090	0.971	0.993	0.997
✗	CADepth [3]	0.079	0.413	3.489	0.120	0.931	0.988	0.997
✓	CADepth [3] + Ph-TTR	0.101	0.616	4.193	0.149	0.889	0.975	0.994
✓	CADepth [3] + SfM-TTR	0.067	0.318	2.728	0.106	0.957	0.989	0.996
✗	DIFFNet [4]	0.071	0.367	3.278	0.110	0.944	0.990	0.997
✓	DIFFNet [4] + Ph-TTR	0.054	0.292	2.927	0.093	0.962	0.991	0.997
✓	DIFFNet [4] + SfM-TTR	0.055	0.261	2.514	0.090	0.971	0.993	0.997

Table 1. **Quantitative results with new KITTI ground truth, Eigen split on sequences with SfM and no cropping.** Best results per model in **bold**, best results across all self-supervised models underlined.

TTR	Method	Abs Rel ↓	Sq Rel ↓	RMSE ↓	RMSE log ↓	$\delta < 1.25 \uparrow$	$\delta < 1.25^2 \uparrow$	$\delta < 1.25^3 \uparrow$
✗	AdaBins [1]	0.060	0.204	2.429	0.091	0.961	0.995	0.999
✓	AdaBins [1] + SfM-TTR	0.055	0.149	1.871	0.081	0.976	0.994	0.998
✗	ManyDepth [2]	0.057	0.289	2.967	0.093	0.959	0.993	0.998
✓	ManyDepth [2] + Ph-TTR	0.051	0.237	2.742	0.087	0.968	0.994	0.998
✓	ManyDepth [2] + SfM-TTR	0.050	0.235	2.446	0.083	0.974	0.994	0.998
✗	CADepth [3]	0.073	0.368	3.333	0.113	0.938	0.990	0.997
✓	CADepth [3] + Ph-TTR	0.093	0.552	3.998	0.140	0.901	0.979	0.995
✓	CADepth [3] + SfM-TTR	0.058	0.269	2.585	0.095	0.965	0.992	0.997
✗	DIFFNet [4]	0.066	0.325	3.122	0.103	0.950	0.992	0.998
✓	DIFFNet [4] + Ph-TTR	0.052	0.264	2.812	0.089	0.965	0.992	0.998
✓	DIFFNet [4] + SfM-TTR	0.050	0.218	2.390	0.083	0.974	0.994	0.998

Table 2. **Quantitative results with new KITTI ground truth, Eigen split on sequences with SfM and Eigen cropping.** Best results per model in **bold**, best results across all self-supervised models underlined.

TTR	Method	Abs Rel ↓	Sq Rel ↓	RMSE ↓	RMSE log ↓	$\delta < 1.25 \uparrow$	$\delta < 1.25^2 \uparrow$	$\delta < 1.25^3 \uparrow$
✗	AdaBins [1]	0.086	0.476	3.609	0.164	0.916	0.970	0.985
✓	AdaBins [1] + SfM-TTR	0.089	0.454	3.308	0.163	0.925	0.970	0.984
✗	ManyDepth [2]	0.089	0.690	4.184	0.165	0.916	0.968	0.984
✓	ManyDepth [2] + Ph-TTR	0.085	0.646	4.031	0.164	0.919	0.968	0.984
✓	ManyDepth [2] + SfM-TTR	0.086	0.708	3.975	0.162	0.924	0.969	0.984
✗	CADepth [3]	0.102	0.729	4.415	0.176	0.896	0.966	0.984
✓	CADepth [3] + Ph-TTR	0.120	0.916	4.944	0.195	0.857	0.955	0.982
✓	CADepth [3] + SfM-TTR	0.092	0.693	3.985	0.168	0.915	0.966	0.983
✗	DIFFNet [4]	0.096	0.712	4.318	0.171	0.907	0.967	0.984
✓	DIFFNet [4] + Ph-TTR	0.084	0.665	4.128	0.164	0.920	0.968	0.984
✓	DIFFNet [4] + SfM-TTR	0.083	0.649	3.872	0.160	0.928	0.970	0.984

Table 3. **Quantitative results with Eigen (old) KITTI ground truth, Eigen split on sequences with SfM and Eigen cropping.** Best results per model in **bold**, best results across all self-supervised models underlined.