A Benchmark and a Baseline for Robust Multi-view Depth Estimation

Philipp Schröppel, Jan Bechtold, Artemij Amiranashvili, Thomas Brox

Introduction

Robust depth estimation is an important component of autonomous driving and erroneous estimates can lead to fatal failures. Therefore, it is crucial that depth estimation works robustly in all possible scenarios.
The so-called open-world problem describes that it is not possible to cover all possible scenarios in the training data. To circumvent this, we present a benchmark that evaluates robust multi-view depth on unseen data. Further, we present a baseline model for this task that outperforms previous work.

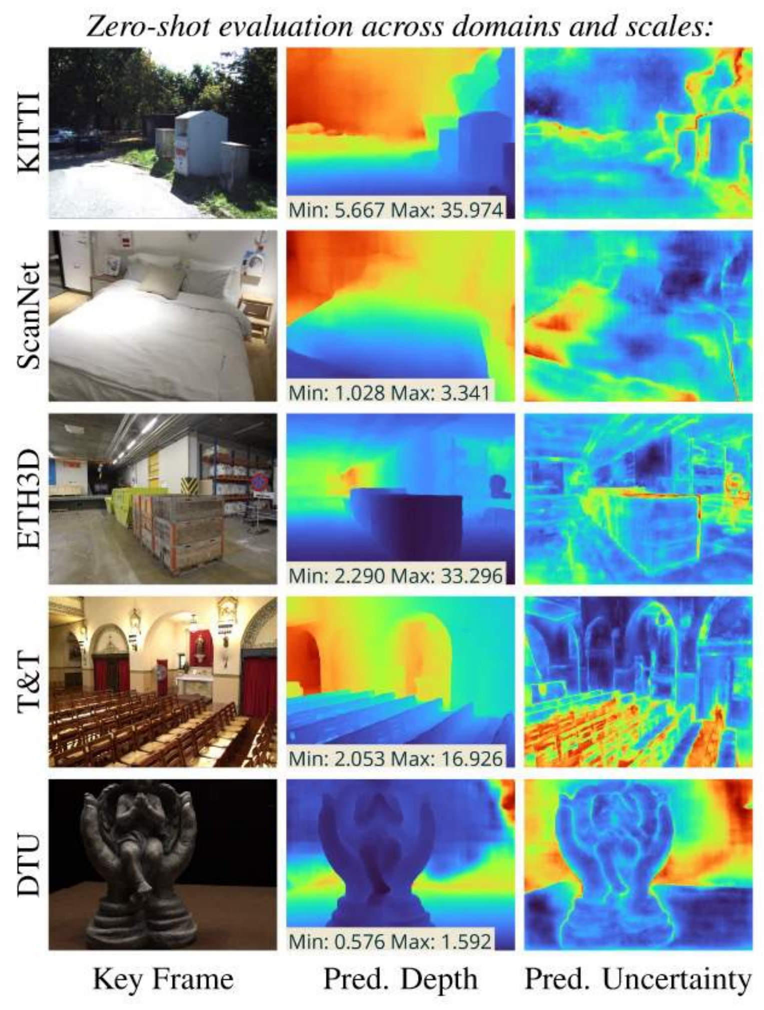

Figure 1: The Robust Multi-View Depth Benchmark evaluates robust multi-view depth estimation on arbitrary real-world data.
© University of Freiburg

The Robust Multi-view Depth Benchmark

Data: The benchmark defines test sets based on multiple diverse existing datasets: KITTI, ScanNet, ETH3D, DTU, Tanks and Temples.
Evaluation settings: The benchmark allows evaluation with different input modalities (images, intrinsics, GT poses, GT depth ranges) and an optional alignment between predicted and ground truth depths.
Metrics: Absolute Relative Error (rel), Inlier Ratio with a Threshold of 1.03 (τ), Sparsification Error Curves, AUSE, performance for different numbers of source views.

The Robust MVD Baseline Model

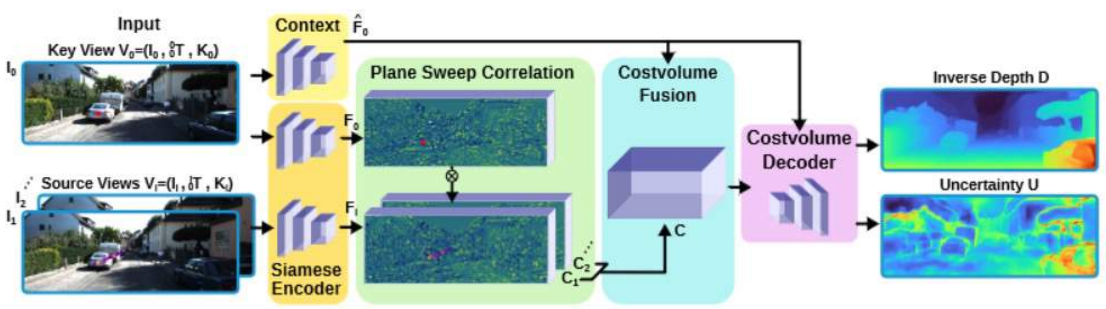

Figure 2: Robust MVD Baseline Model Architecture.
© University of Freiburg

Architecture: The architecture is based on DispNet but modified to multi-view inputs.
Training data: StaticThings3D + BlendedMVS.
Data augmentation: The model is trained with a novel augmentation strategy that we call scale augmentation.

Results

We evaluate state-of-the art models and the Robust MVD Baseline Model on the RMVD Benchmark. Results are given in Table 1.
We find that:

- All models perform significantly better on the training domain (marked in ■).
- Models perform significantly worse in the absolute scale setting (marked in ■).
- The Robust MVD Baseline Model shows more consistent performance across test sets and works in the absolute scale setting (marked in ■).

We provide code at:
https://github.com/lmb-freiburg/robustmvd

- dataloaders
- evaluation code
- code to run all models.

Link to code

Approach	GT Poses	GT Range	Align	KITTI rel ↓	KITTI τ ↑	ScanNet rel ↓	ScanNet τ ↑	ETH3D rel ↓	ETH3D τ ↑	DTU rel ↓	DTU τ ↑	T&T rel ↓	T&T τ ↑	Average rel ↓	Average τ ↑	time [s] ↓
a)																
DeMoN	✗	✗	‖t‖	15.5	15.2	**12.0**	**21.0**	17.4	15.4	21.8	16.6	13.0	23.2	16.0	18.3	**0.08**
DeepV2D KITTI	✗	✗	med	**(3.1)**	**(74.9)**	23.7	11.1	27.1	10.1	24.8	8.1	34.1	9.1	22.6	22.7	2.07
DeepV2D ScanNet	✗	✗	med	**10.0**	**36.2**	(4.4)	(54.8)	**11.8**	**29.3**	**7.7**	**33.0**	**8.9**	**46.4**	**8.6**	**39.9**	3.57
b)																
MVSNet	✓	✓	✗	22.7	36.1	24.6	20.4	35.4	31.4	**(1.8)**	(86.0)	8.3	73.0	18.6	49.4	0.07
MVSNet Inv. Depth	✓	✓	✗	18.6	30.7	22.7	20.9	21.6	35.6	**(1.8)**	(86.7)	6.5	74.6	14.2	49.7	0.32
CVP-MVSNet	✓	✓	✗	156.7	2.2	137.1	15.9	156.4	13.6	(4.0)	(68.4)	24.7	52.9	95.8	30.6	0.49
Vis-MVSNet	✓	✓	✗	**9.5**	**55.4**	8.9	33.5	**10.8**	**43.3**	**(1.8)**	**(87.4)**	**4.1**	**87.2**	**7.0**	**61.4**	0.70
PatchmatchNet	✓	✓	✗	10.8	45.8	**8.5**	**35.3**	19.1	34.8	(2.1)	(82.8)	4.8	82.9	9.1	56.3	0.28
Fast-MVSNet	✓	✓	✗	14.4	37.1	17.0	24.6	25.2	32.0	(2.5)	(81.8)	8.3	68.6	13.5	48.8	0.30
MVS2D ScanNet	✓	✓	✗	21.2	8.7	(27.2)	(5.3)	27.4	4.8	17.2	9.8	29.2	4.4	24.4	6.6	**0.04**
MVS2D DTU	✓	✓	✗	226.6	0.7	32.3	11.1	99.0	11.6	(3.6)	(64.2)	25.8	28.0	77.5	23.1	0.05
c)																
DeMoN	✓	✗	✗	16.7	13.4	75.0	0.0	19.0	16.2	23.7	11.5	17.6	18.3	30.4	11.9	0.08
DeepTAM	✓	✗	✗	68.7	0.4	(6.7)	(39.7)	20.4	19.8	58.0	9.1	40.0	12.9	38.8	16.4	0.85
DeepV2D KITTI	✓	✗	✗	(20.4)	(16.3)	25.8	8.1	30.1	9.4	24.6	8.2	38.5	9.6	27.9	10.3	1.43
DeepV2D ScanNet	✓	✗	✗	61.9	5.2	**(3.8)**	**(60.2)**	18.7	28.7	9.2	27.4	33.5	38.0	25.4	31.9	2.15
MVSNet	✓	✗	✗	14.0	35.8	1568.0	5.7	507.7	8.3	(4429.1)	(0.1)	118.2	50.7	1327.4	20.1	0.15
MVSNet Inv. Depth	✓	✗	✗	29.6	8.1	65.2	28.5	60.3	5.8	(28.7)	(48.9)	51.4	14.6	47.0	21.2	0.28
CVP-MVSNet	✓	✗	✗	158.2	1.2	2289.0	0.1	1735.3	1.2	(8314.0)	(0.0)	415.9	9.5	2582.5	2.4	0.50
Vis-MVSNet	✓	✗	✗	10.3	**54.4**	84.9	15.6	51.5	17.4	(374.2)	(1.7)	21.1	65.6	108.4	31.0	0.82
PatchmatchNet	✓	✗	✗	29.0	16.3	70.1	16.7	99.4	3.5	(82.6)	(5.6)	39.4	19.3	64.1	12.3	0.18
Fast-MVSNet	✓	✗	✗	12.1	37.4	287.1	9.4	131.2	9.6	(540.4)	(1.9)	33.9	47.2	200.9	21.1	0.35
MVS2D ScanNet	✓	✗	✗	73.4	0.0	(4.5)	(54.1)	30.7	14.4	5.0	57.9	56.4	11.1	34.0	27.5	**0.05**
MVS2D DTU	✓	✗	✗	93.3	0.0	51.5	1.6	78.0	0.0	**(1.6)**	**(92.3)**	87.5	0.0	62.4	18.8	0.06
Robust MVD Baseline	✓	✗	✗	**7.1**	41.9	**7.4**	**38.4**	**9.0**	**42.6**	**2.7**	**82.0**	**5.0**	**75.1**	**6.3**	**56.0**	0.06

Table 1: Quantitative results for the evaluated multi-view depth models on the Robust MVD Benchmark with different evaluation settings: a) no poses, no depth range, with alignment, b) with poses, with depth range, no alignment, c) with poses, no depth range, no alignment.

Partners

External partners

For more information contact:

schroepp@cs.uni-freiburg.de
brox@cs.uni-freiburg.de

KI Delta Learning is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.

www.ki-deltalearning.de @KI_Familie KI Familie

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag