Evaluation of CNN-based Single-Image Depth Estimation Methods(CVPR 18)

Tobias Koch, Lukas Liebel, Friedrich Fraundorfer, Marco Körner

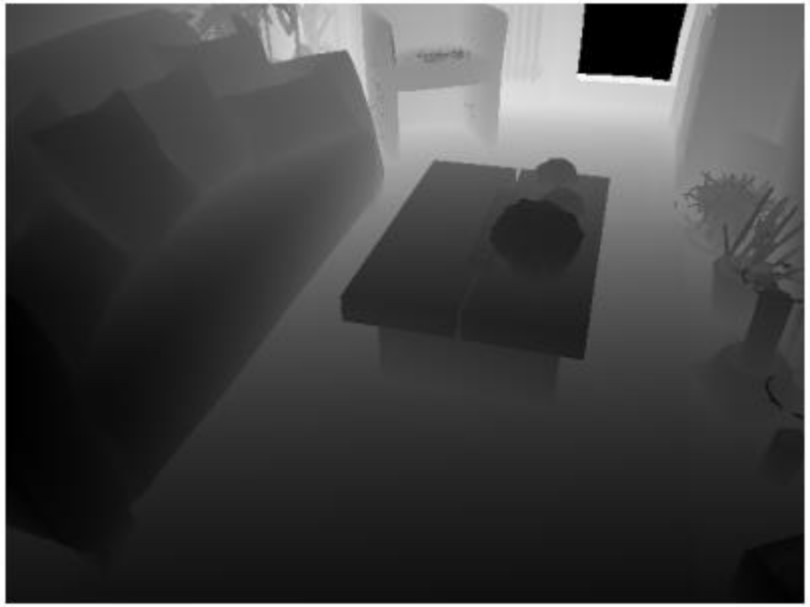
20184269 Taehee Kim

18. 11. 20. CS688 Student Presentation

Main Topic

- Single image -> Depthmap estimation
- Application: Shape, depth aware image retrieval

RGB Image



Depthmap

Introduction

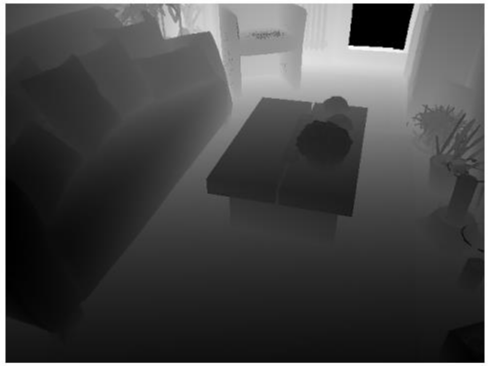
Problem

Goals

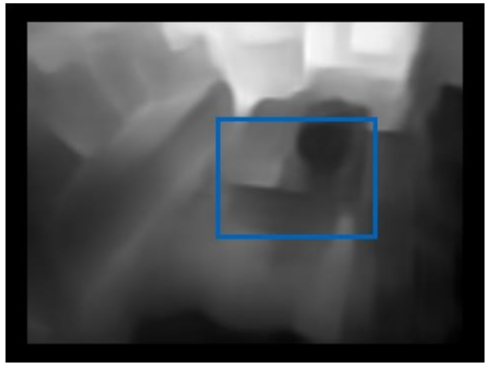
Problem

- Error metrics does not reflect detailed structures
- No sufficient dataset for training

(a) RGB image



(b) Depth map



(c) Prediction

(d) Prediction detail

Positively evaluated but poor details

Goals

1. Introduce a set of new **error metrics**
2. Present a new **dataset** from laser scan
3. **Evaluate** state-of-art methods

1. Error Metrics

Commonly Used Error Metrics
Hard Examples
Requirements for Good Metric
Planarity, Orientation Metric
Depth Boundary Metric

Commonly Used Error Metrics

Threshold: $\%$ of y such that $\max(\frac{y_i}{y_i^*}, \frac{y_i^*}{y_i}) = \sigma < thr$

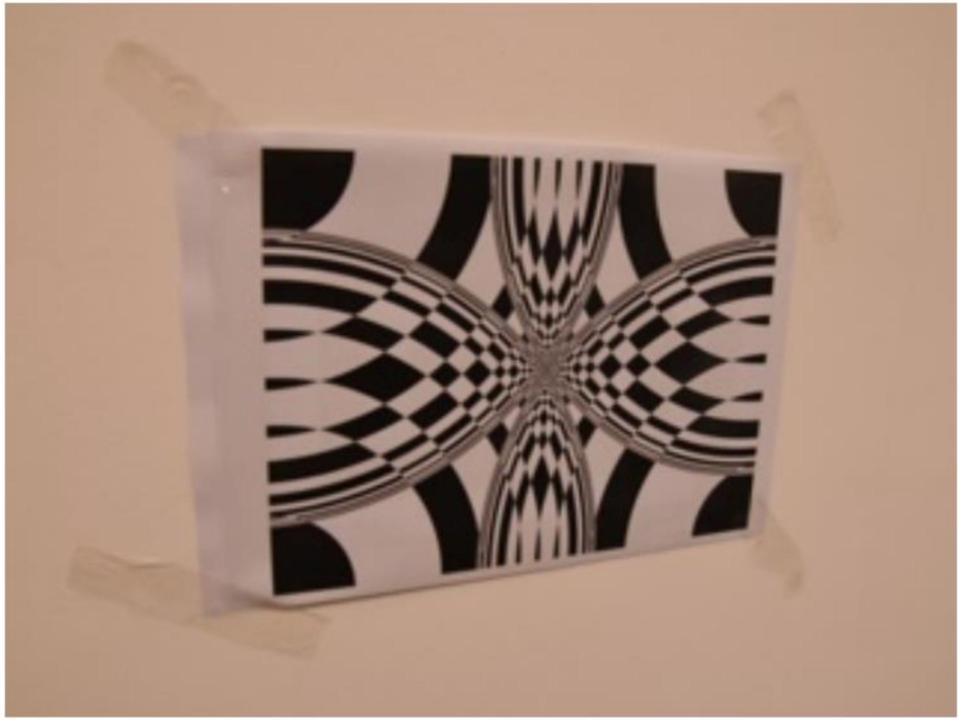
Absolute rel. diff.: $\mathrm{rel} = \frac{1}{T} \sum_{i,j} \left| y_{i,j} - y_{i,j}^* \right| / y_{i,j}^*$

Squared rel. diff.: $\mathrm{srel} = \frac{1}{T} \sum_{i,j} \left| y_{i,j} - y_{i,j}^* \right|^2 / y_{i,j}^*$

RMS (linear): $\mathrm{RMS} = \sqrt{\frac{1}{T} \sum_{i,j} \left| y_{i,j} - y_{i,j}^* \right|^2}$

RMS (log): $\log_{10} = \sqrt{\frac{1}{T} \sum_{i,j} \left| \log y_{i,j} - \log y_{i,j}^* \right|^2}$

Hard Examples



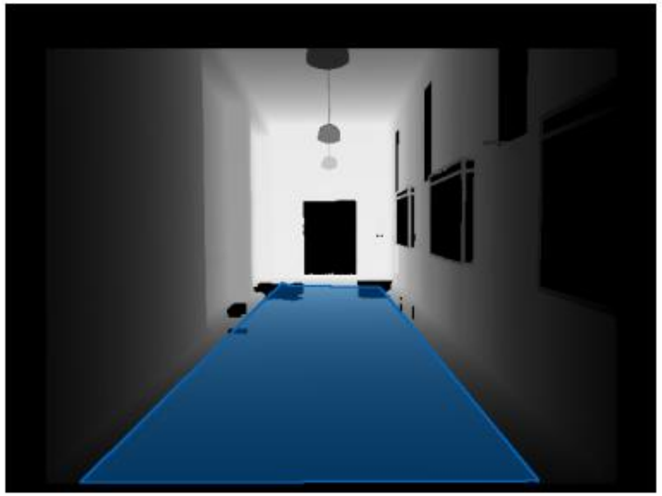
Paint? Bumps?

Reflection? Shallow Region??

Requirements for Good Metric

(Overall accuracy)+

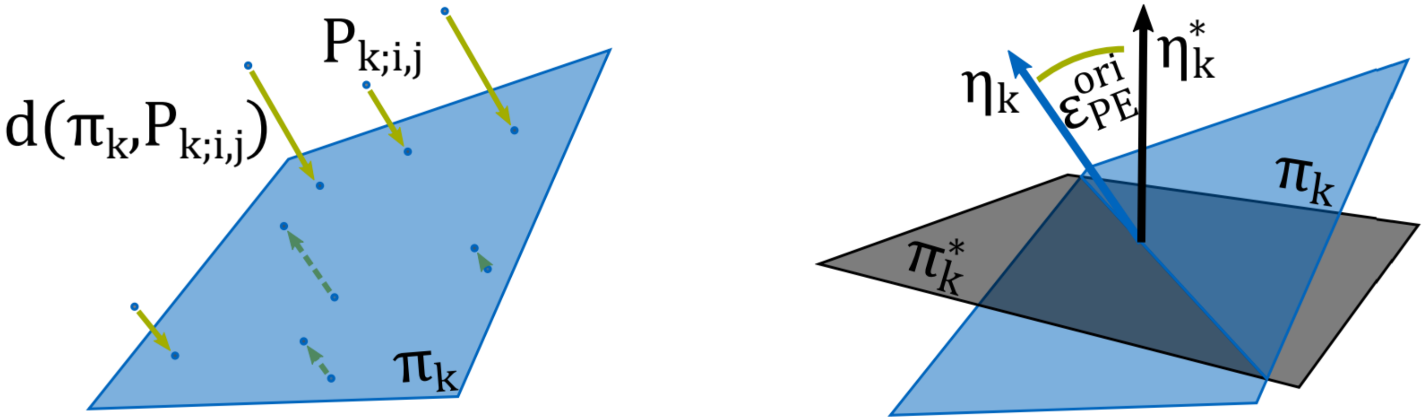
- Capture planarity
- Orientation of surface
- Depth Discontinuity(edge) location



Planarity, Orientation Metric

- Annotated Plane: $\pi_k^* = (\eta_k^*, d_k^*)$ (normal vector, origin)
- Project depthmap $\boldsymbol{Y_k}$ to 3D points $\boldsymbol{P_{k;i,j}}$

$$\varepsilon_{\text{PE}}^{\text{plan}}(\boldsymbol{Y}_k) = \mathbb{V}\left[\sum_{\boldsymbol{P}_{k;i,j} \in \mathcal{P}_k} d\left(\boldsymbol{\pi}_k, \boldsymbol{P}_{k;i,j}\right)\right] \qquad \varepsilon_{\text{PE}}^{\text{orie}}(\boldsymbol{Y}_k) = \operatorname{acos}\left(\boldsymbol{\eta}_k^{\top} \cdot \boldsymbol{\eta}_k^*\right)$$

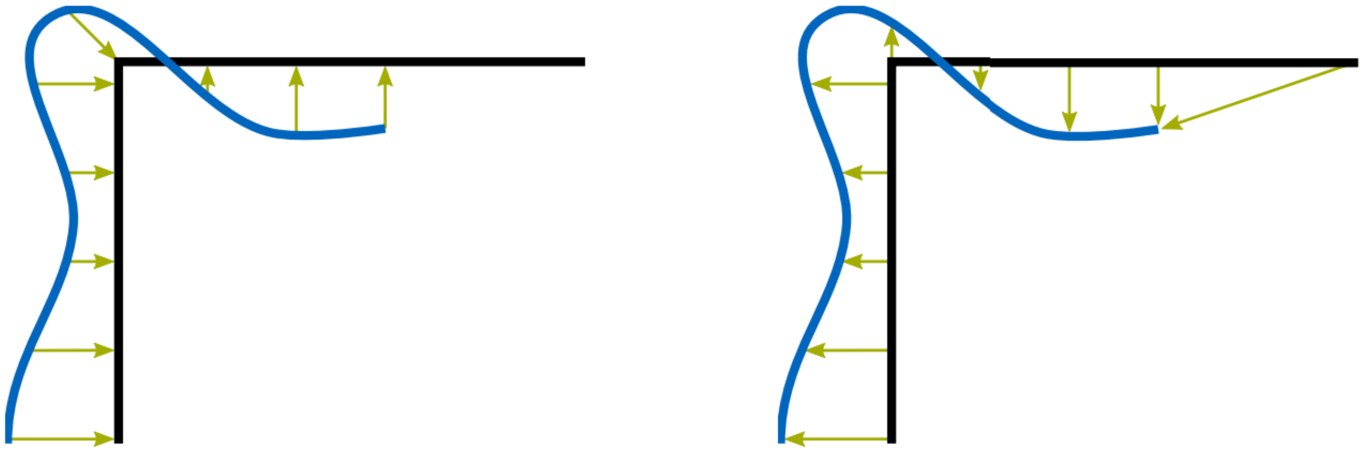


(a) $\varepsilon_{\text{PE}}^{\text{plan}}$ **(b)** $\varepsilon_{\text{PE}}^{\text{orie}}$

Depth Boundary Metric

- Edge prediction using “Structured Edge ”
- Euclidian distance between Structured Edge and Ground T.

$$\varepsilon_{\mathrm{DBE}}^{\mathrm{acc}}(\boldsymbol{Y}) = \frac{1}{\sum_i \sum_j y_{\mathrm{bin};i,j}} \sum_i \sum_j e_{i,j}^{*} \cdot y_{\mathrm{bin};i,j}$$



(c) $\varepsilon_{\mathrm{DBE}}^{\mathrm{acc}}$ **(d)** $\varepsilon_{\mathrm{DBE}}^{\mathrm{comp}}$

2. Dataset

Existing Datasets

Data Acquisition

Proposed Dataset: IBims-1

Existing Datasets

- Multiple laser scan (ETH3D, Tanks&Temples, …)
  - ➢ Occlusion
- Custom Built-in 3D scanner (Kitti)
  - ➢ Low Resolution
- Active RGB-D sensors (NYU depth v2, Matterport3D)
  - ➢ Short range, erroneous specular surface

Data Acquisition

- DSLR + Single laser scanner
- Custom tripod to align optical center

(a) Laser scanner

(b) Camera

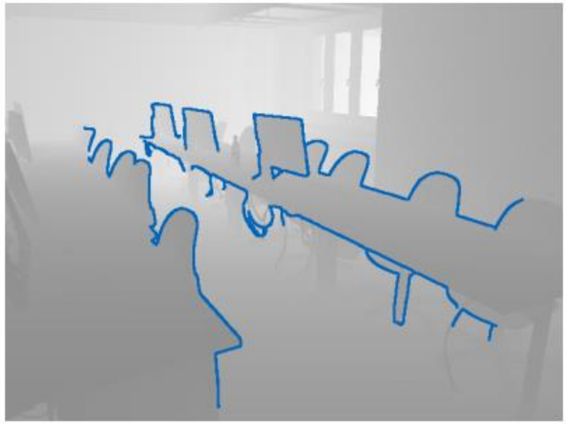
Proposed Dataset: IBims-1

- High-resolution RGB-D with annotations
- Object masks and edges

(a) Camera image

(b) Ground truth

(c) Masks



(d) Distinct edges

3. Evaluation

Previous Works
CNN Based Depth Estimation(Eigen et el)
Quantitative Evaluations
Qualitative Evaluations

Previous Works

- Eigen et el. First CNN based approach.
- Liu et el. CNN + conditional random fields(CRF).
- Laina et el. Fully convolutional network
- Li et el. Two-streamed CNN for depth and depth gradients
- Xu et el. Integrate multiple CNN using CRF

CNN Based Depth Estimation(Eigen et el)

Coarse-scale network
(Global level)

Fine-scale network
(Local level)

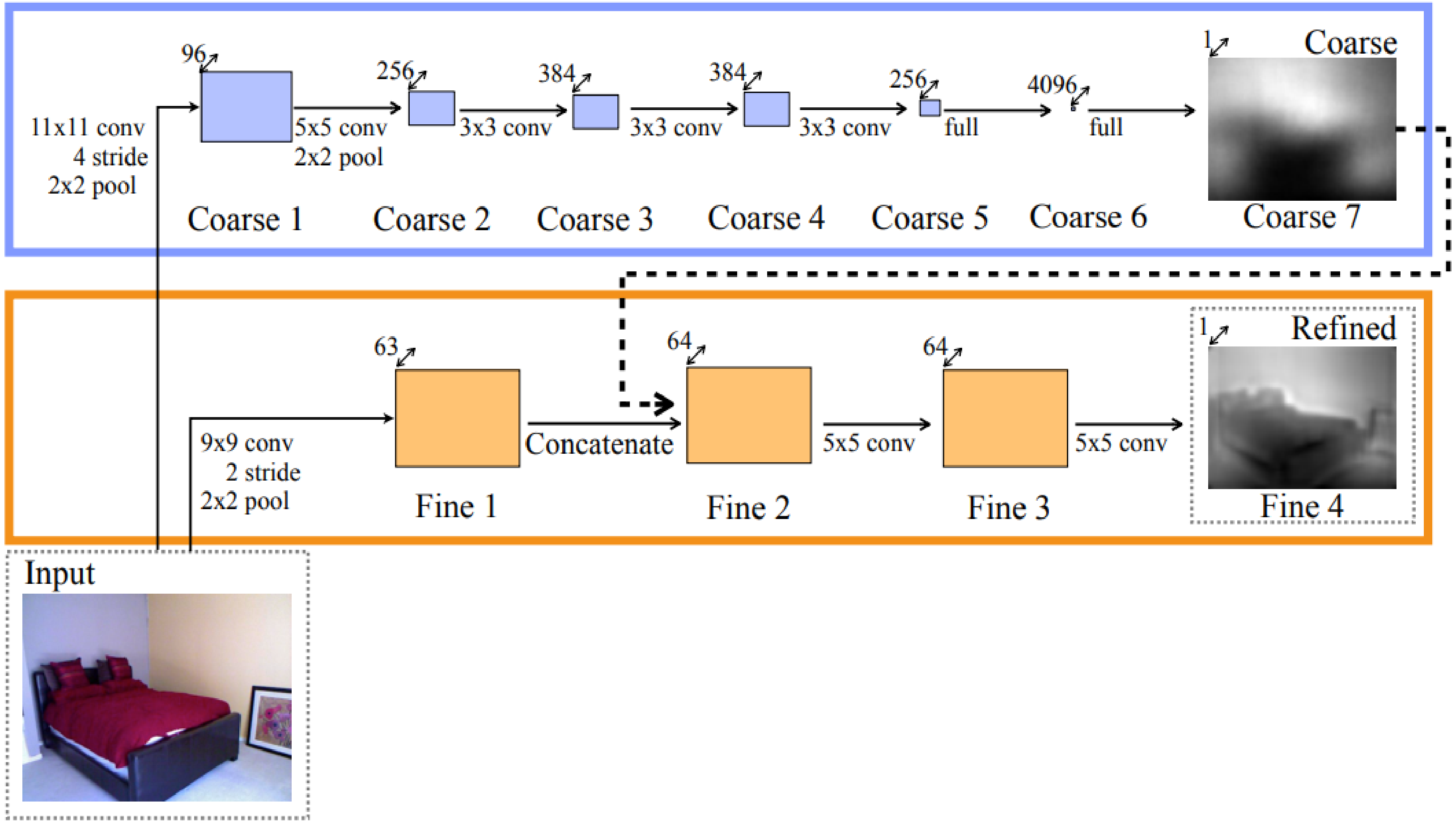


Figure from Eigen et el. “Depth Map Prediction from a Single Image using a Multi-Scale Deep Network”

Quantitative Evaluations

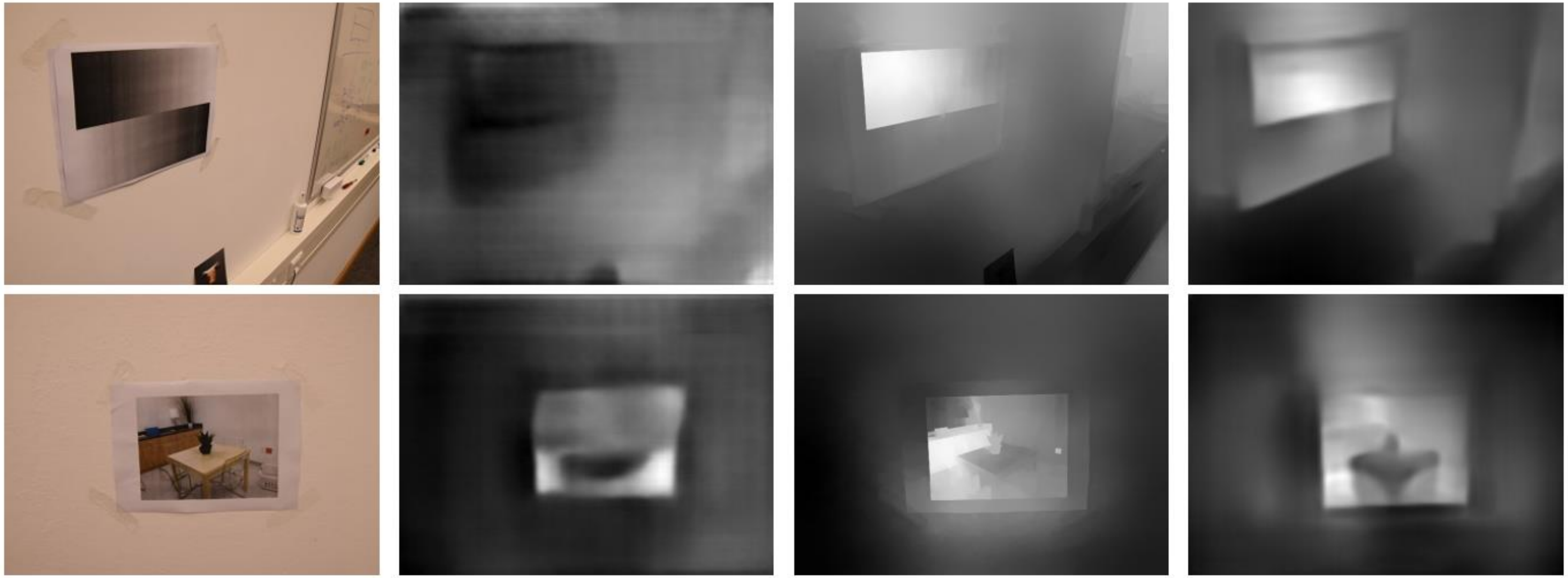
- Li et el is best with standard metrics, but not with proposed metrics

Table 3: Quantitative results for standard metrics and proposed PE, DBE, and DDE metrics on `IBims-1` applying different SIDE methods

Method	Standard Metrics ($\sigma_i = 1.25^i$)						PE (in m/°)		DBE (in px)		DDE (in %)		
	rel	$\log_{10}$	RMS	σ_1	σ_2	σ_3	ε_{PE}^{plan}	ε_{PE}^{orie}	ε_{DBE}^{acc}	ε_{DBE}^{comp}	ε_{DDE}^{0}	ε_{DDE}^{-}	ε_{DDE}^{+}
Eigen [7]	0.36	0.22	2.92	0.35	0.63	0.79	0.18	33.27	3.60	48.08	64.53	32.31	3.15
Eigen (AlexNet) [6]	0.32	0.18	2.63	0.42	0.72	0.82	0.21	26.64	3.01	32.00	74.65	21.51	3.84
Eigen (VGG) [6]	0.29	0.17	2.59	0.47	0.73	0.85	**0.17**	**21.64**	3.16	27.47	75.10	23.44	**1.46**
Laina [16]	0.27	0.16	2.42	0.56	0.76	0.84	0.22	32.02	4.58	38.41	77.12	20.89	1.99
Liu [20]	0.33	0.17	2.51	0.46	0.73	0.84	0.22	31.90	**2.32**	**16.85**	77.27	**16.38**	6.35
Li [19]	**0.25**	**0.14**	**2.32**	**0.58**	**0.79**	**0.86**	0.20	26.67	2.36	21.02	**80.99**	16.44	2.57

Qualitative Evaluations

• Laina et el seems poor, Liu et el seems good (Proposed metrics well represent these points)



(a) RGB **(b)** Laina et al. [16] **(c)** Liu et al. [20] **(d)** Eigen [6]