On Joint Estimation of Pose, Geometry and svBRDF from a Handheld Scanner

Carolin Schmitt[1,2,*] Simon Donné[1,2,*] Gernot Riegler[3] Vladlen Koltun[3] Andreas Geiger[1,2]

[1]MPI for Intelligent Systems, Tübingen [2]University of Tübingen [3]Intel Intelligent Systems Lab

MAX PLANCK INSTITUTE FOR INTELLIGENT SYSTEMS

intel

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

imprs-is

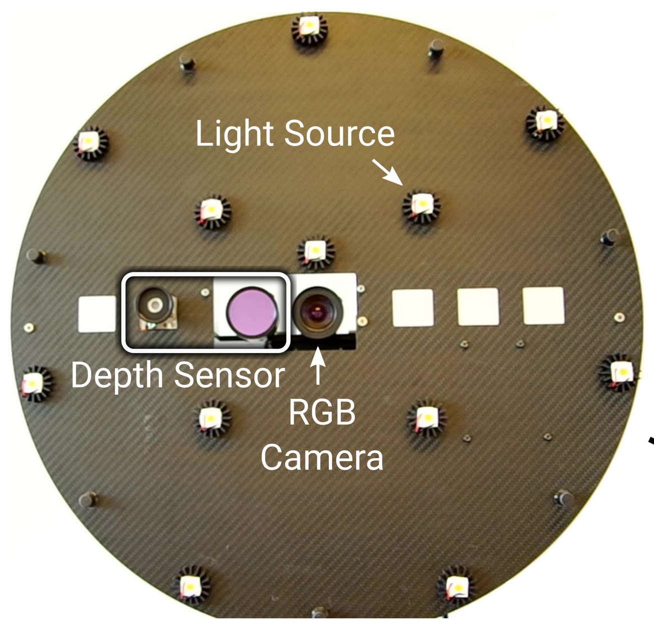
Motivation

Structured light sensing (e.g. Kinect) enabled large-scale 3D reconstructions. Yet, the **level of realism is limited** since physical light transport is not modeled. Illumination effects such as specular reflections or shadows are baked into the texture.

Reflections and geometry are linked strongly. To use that mutual information, object geometry and material properties are best inferred jointly. We propose a novel formulation for **joint pose, geometry and svBRDF estimation**.

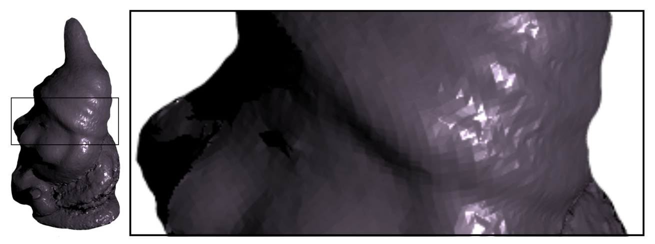
While existing approaches use alternating optimization procedures, we propose a **single objective** that can be optimized using off-the-shelf gradient-based solvers.

Handheld Sensor



Custom built sensor rig.

RGB input images captured under known point light illumination.



Complete but coarse initial depth.

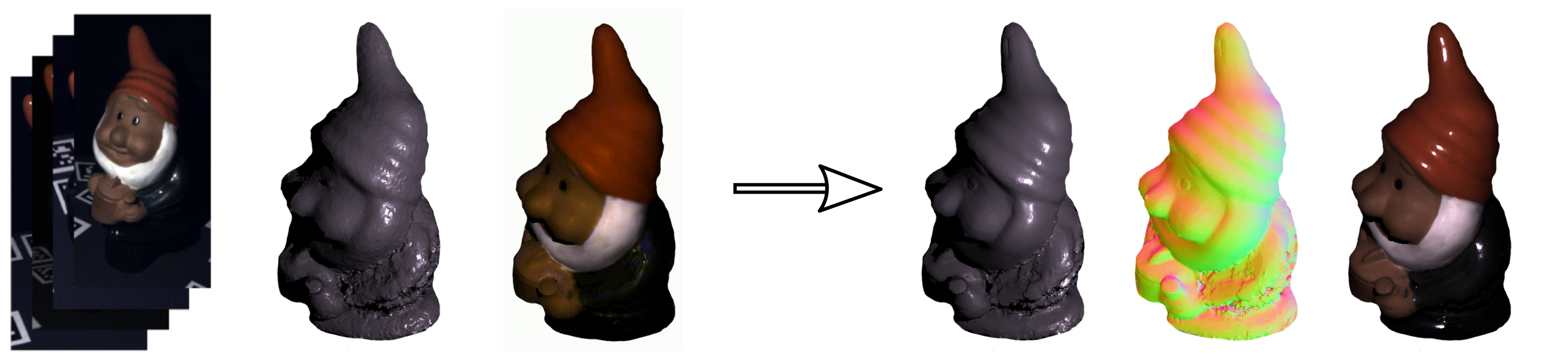
+ Fully portable, no laboratory setup for material estimation
+ For arbitrary geometry, no prior assumptions like planarity

Materials ⟷ Geometry

→ Traditionally accurate geometry reconstruction requires known appearance properties.

← Likewise, accurate appearance estimation requires very well known geometry.

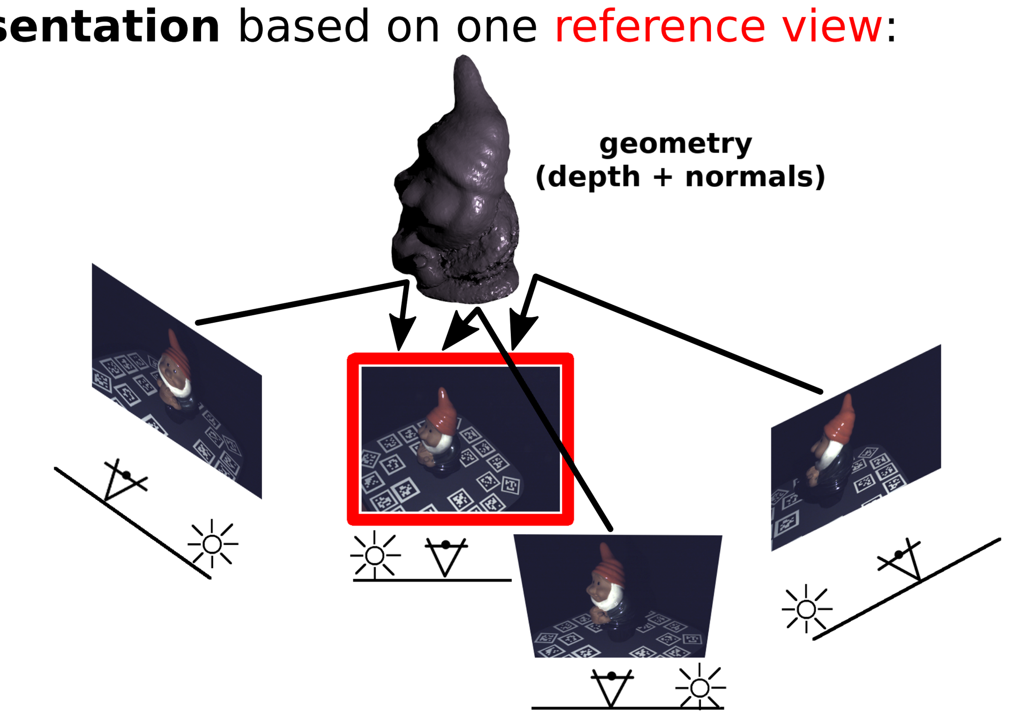
↔ Joint estimation of geometry and materials requires only a rough initialization for both.



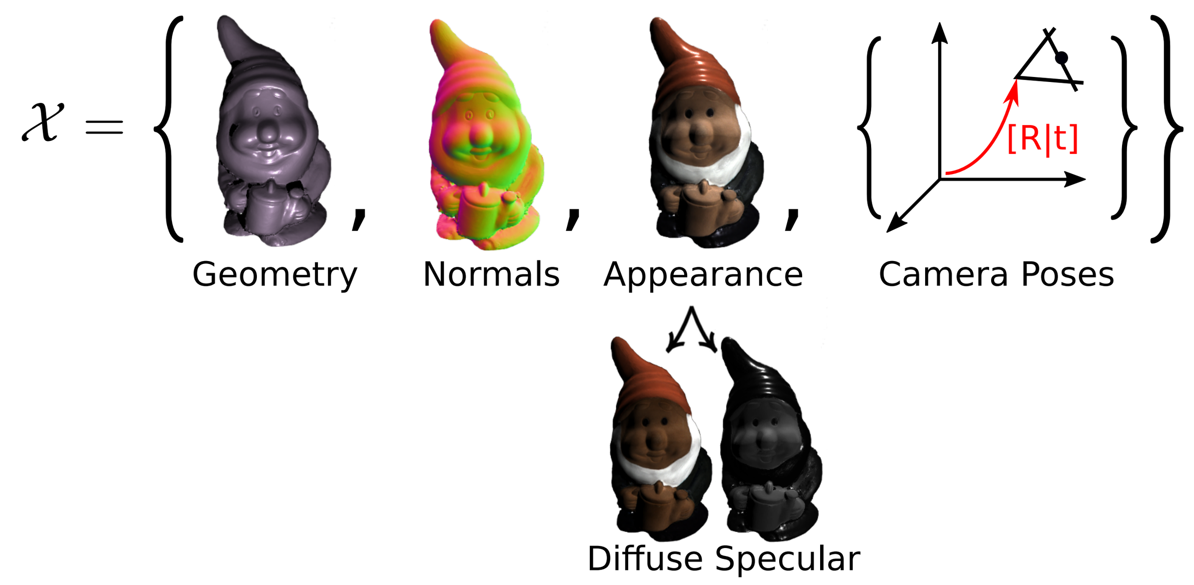
Input RGB Initial Depth and Appearance Refined Geometry, Normals and Appearance

Joint Formulation

2.5D scene representation based on one reference view:



Optimized variables:

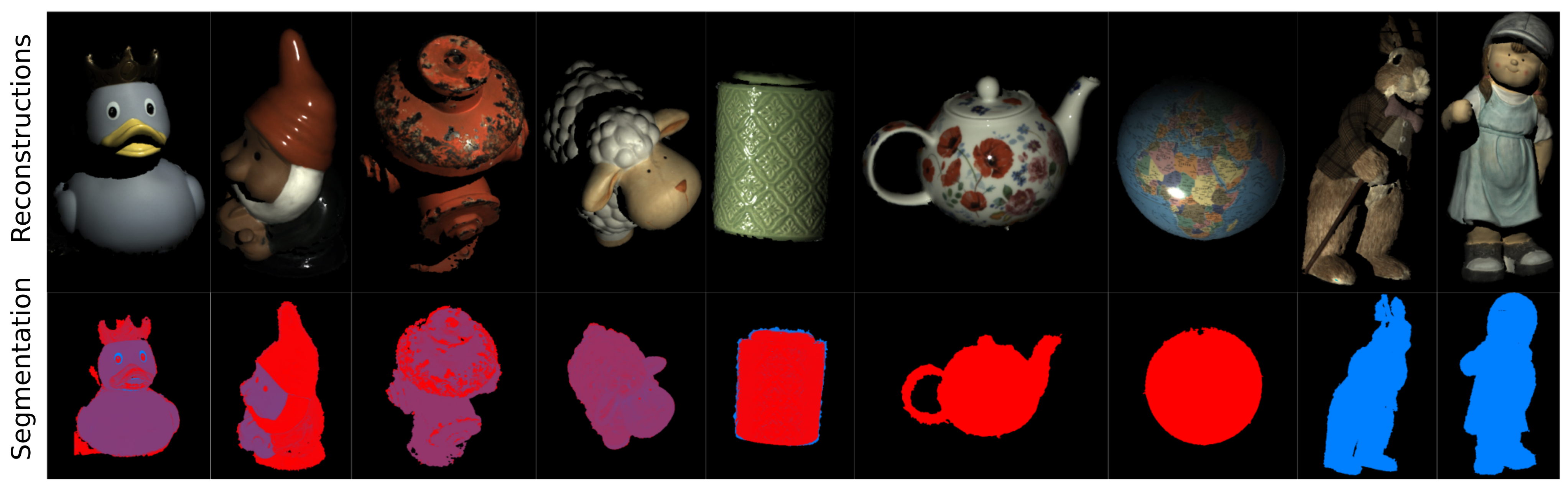


A single objective function minimized by off-the-shelf gradient-based solvers.

$$\mathcal{X}^* = \text{argmin}_{\mathcal{X}} \underbrace{\psi_{\mathcal{P}}}_{\text{photoconsistency}} + \underbrace{\psi_{\mathcal{G}} + \psi_{\mathcal{D}} + \psi_{\mathcal{N}}}_{\text{geometry regularizer}} + \underbrace{\psi_{\mathcal{M}}}_{\text{material smoothness}}$$

Material Assignments

$\psi_{\mathcal{M}}$ integrates material assignments into the optimization process as a differentiable regularizer. This naturally leads to a **semantically meaningful material segmentation**.



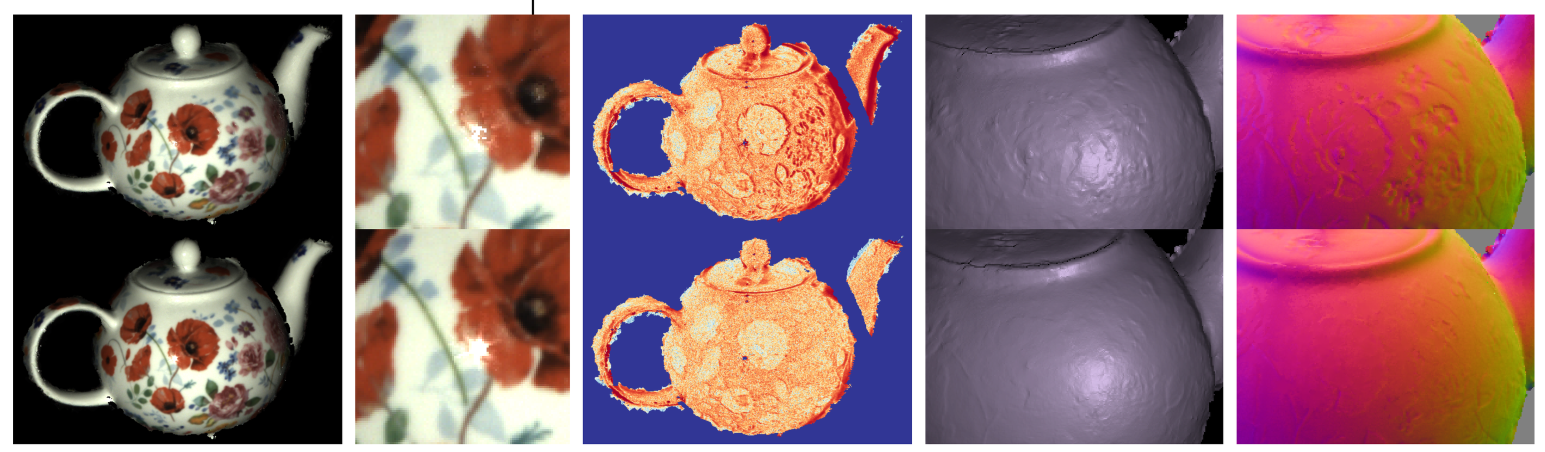
Duck ($T=2$) Gnome ($T=2$) Hydrant ($T=2$) Sheep ($T=2$) Vase ($T=2$) Teapot ($T=1$) Globe ($T=1$) Rabbit ($T=1$) Girl ($T=1$)

$T =$ number of specular base materials; independently determined by our model.

Pose Ablation Study

Refining the initial poses is crucial for recovering fine texture details. Top: using input poses, bottom: optimizing the poses.

Photometric Test Error	Overall	Specularity Regions	Other Regions
Fixed Poses	1.210	3.349	1.151
Full Model	1.138	3.243	1.081

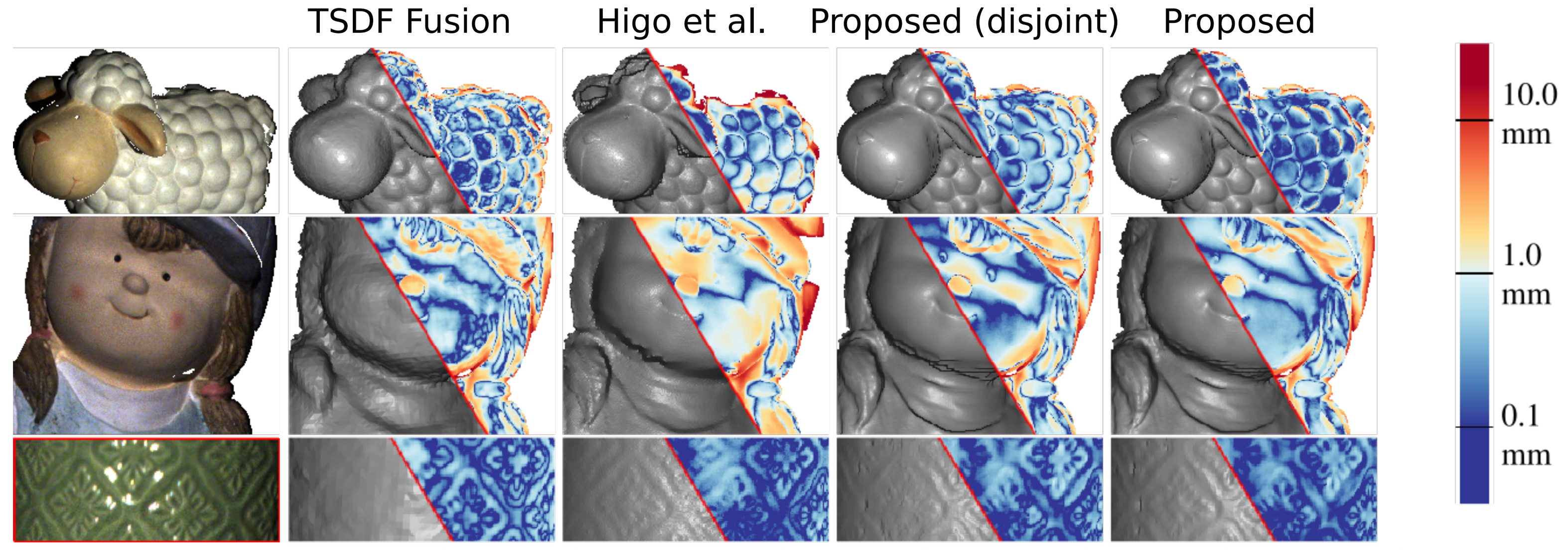


Estimation Crop Error Geometry Normals

Geometric Evaluation

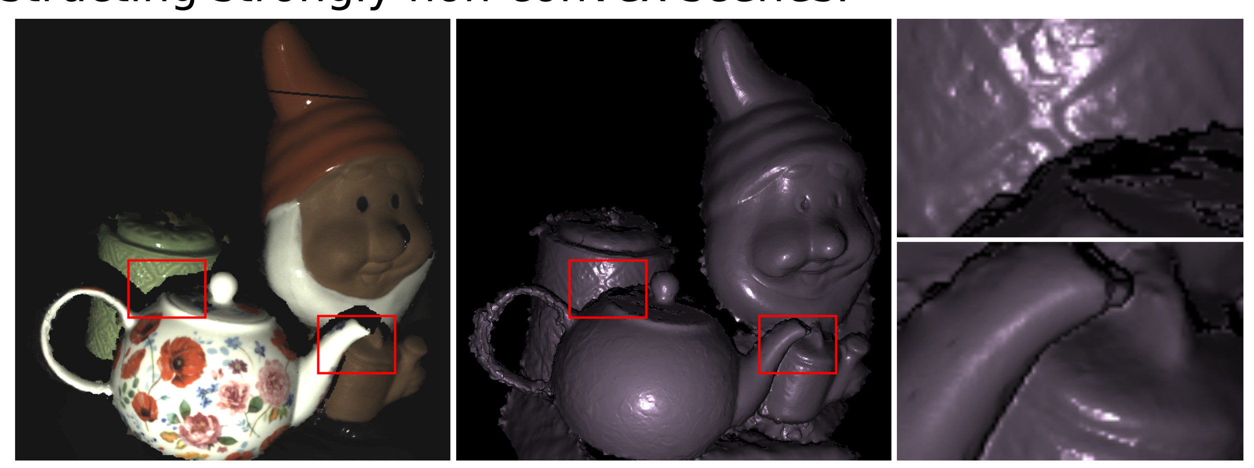
Qualitatively and quantitatively, the proposed method recovers **more details** with **less noise** and a **lower error** than the baselines.

- Proposed (disjoint): alternating instead of joint optimization of geometry and materials.
- Artec Spider scans serve as ground truth.



		duck	pineapple	girl	gnome	sheep	hydrant	rabbit
AEA (mm)	TSDF Fusion	0.81	1.24	1.11	0.73	0.79	1.35	2.16
	Higo et al.	2.65	1.05	1.59	1.60	2.09	1.81	2.85
	Proposed (disjoint)	0.81	1.00	1.06	0.65	0.64	1.18	2.25
	Proposed	0.80	1.00	1.00	0.64	0.57	1.16	2.16
AAE (°)	TSDF Fusion	6.75	12.09	11.40	7.64	8.38	11.67	24.42
	Higo et al.	7.77	10.62	11.30	9.24	8.65	15.27	27.67
	Proposed (disjoint)	6.01	9.13	9.81	6.41	6.66	9.59	23.01
	Proposed	5.17	8.98	8.73	5.74	5.60	8.50	23.50

The **explicit shadow and occlusion modeling** enables reconstructing strongly non-convex scenes.



Reconstruction Geometry Crops of Object Boundaries

https://github.com/autonomousvision/handheld_svbrdf_geometry

{carolin.schmitt, simon.donne}@tue.mpg.de