Computer
Vision
Group

Probabilistic Labeling Cost for High-Accuracy Multi-View Reconstruction

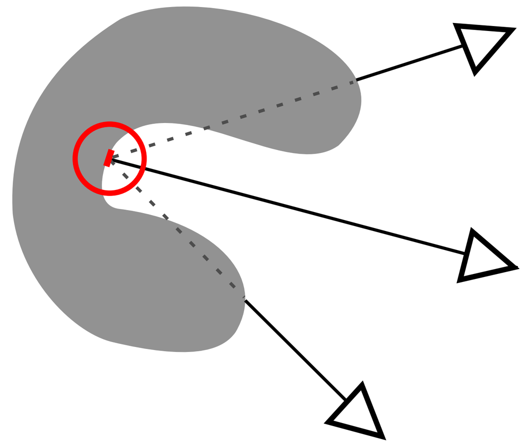
Ilya Kostrikov, Esther Horbert, Bastian Leibe

RWTH AACHEN UNIVERSITY

Abstract

We propose a novel labeling cost for multi-view reconstruction. Existing approaches use data terms that are vulnerable to common challenges, such as low-textured regions or specularities. Our new probabilistic method implicitly discards outliers and can be shown to become more exact the closer to the true object surface. Our approach is simple to implement, can be readily integrated into many existing multi-view stereo approaches, and achieves top results among all published methods on the Middlebury Dino Sparse dataset.

Motivation

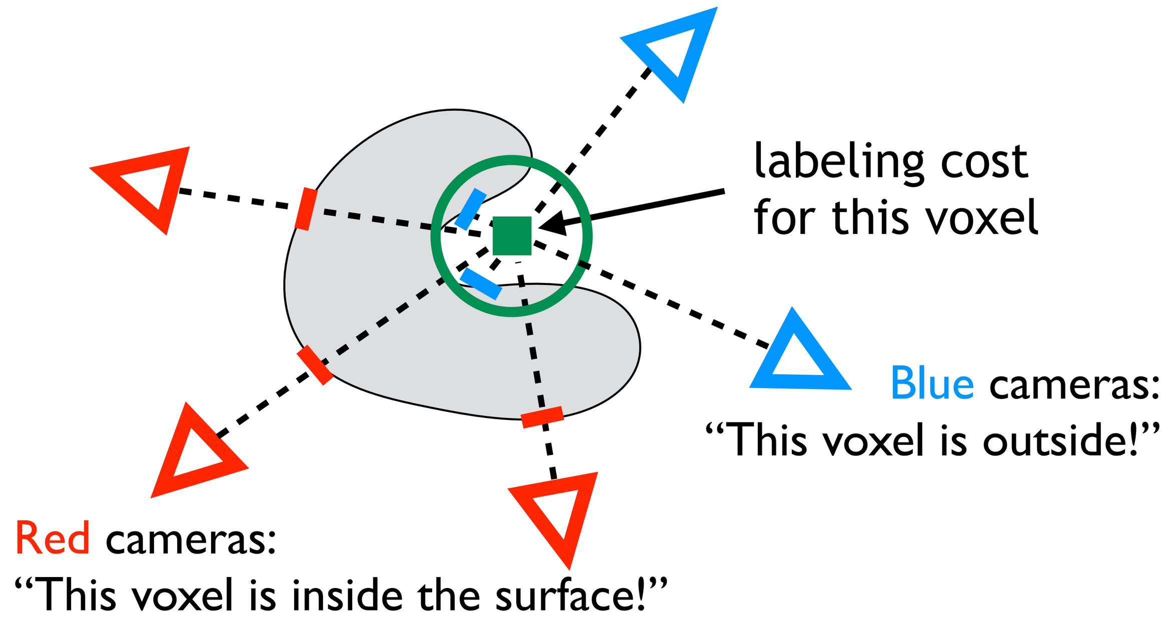


Many papers make assumptions about labeling considering only a single camera. In the context of multi-view reconstruction such assumptions are not valid.

Main Idea

Not all cameras can provide information about all voxels

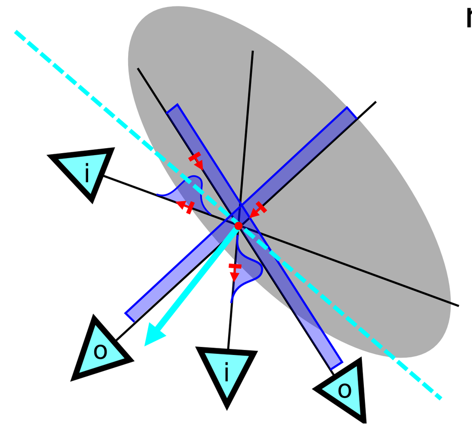
- Many voxels are not visible for a camera
- Especially in low textured areas some cameras provide unreliable information – outliers
- Observations closer to the voxel in question are more reliable



We select the cameras with the ***spatially closest*** observations for 3D reconstruction

Comparison with Related Work

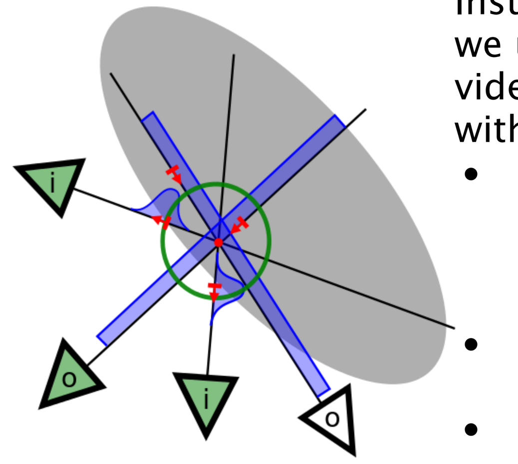
Kolev et al IJCV'09



Select all cameras within 60° of surface normal. Problems:

- Need surface to estimate normal to estimate surface
- Need to apply iterative scheme (slow and propagates errors)
- Not robust to outliers since angle does not provide any insight if camera is an outlier

Our method



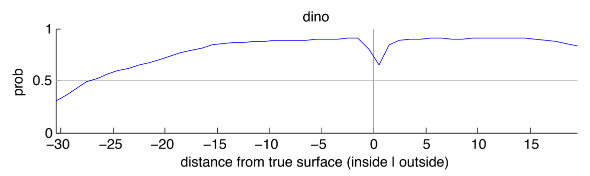
Instead of using all cameras we use a subset that provides the correct labeling with a high probability:

- For each voxel we select the k (e.g. 3) cameras with the **spatially closest** observations
- Subset does not depend on the appearance of surface
- Good reconstructions even in low textured regions

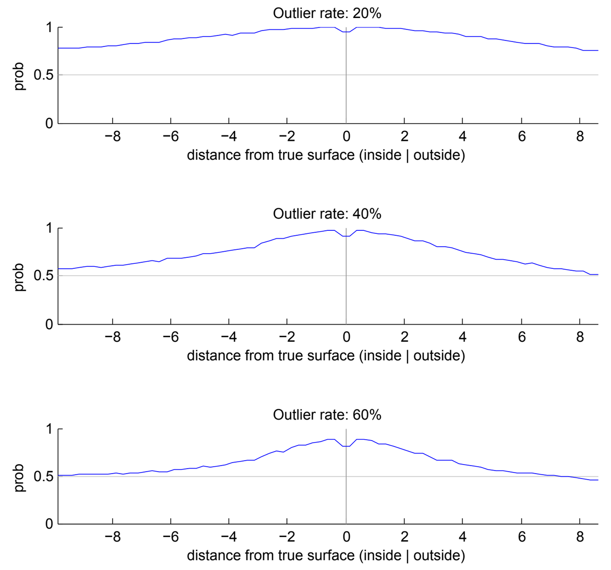
Accuracy of Camera Selection

Influence of outliers on the labeling:
Measure percentage of correct decision as a function of the distance from the surface.

dino dataset



Synthetic experiment



Useful property:
Close to the object surface:

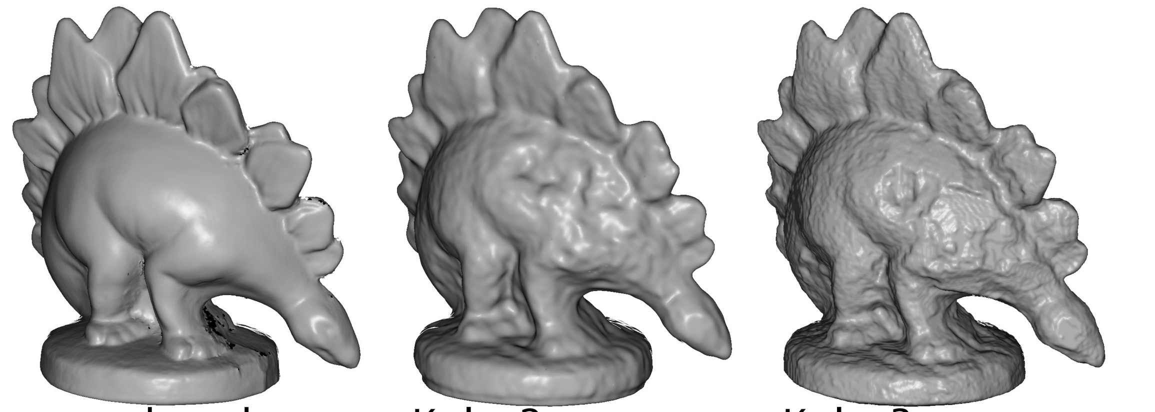
- Method selects cameras that have correct observations with a high probability
- No knowledge about the location of the surface necessary!

Further away from the surface:

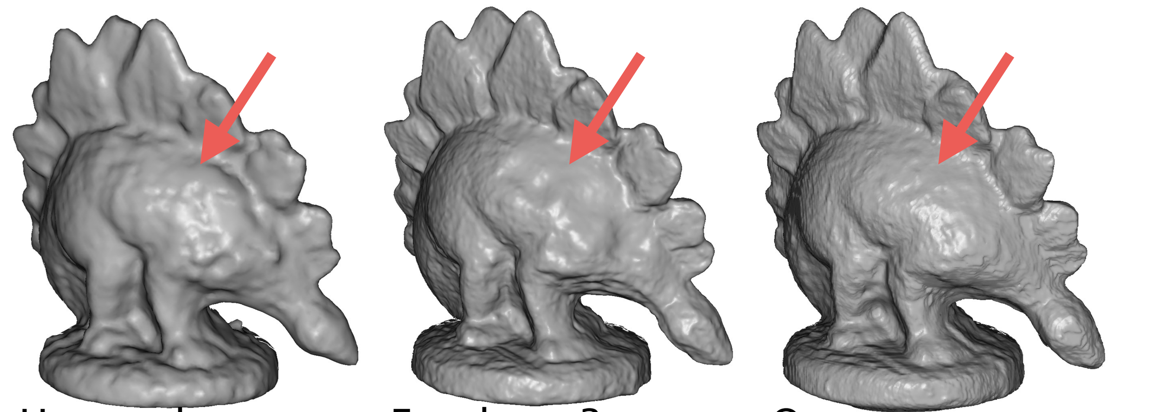
- No need for precise labeling here
- Outliers will not harm the overall result
- Noisy labels that are far from the surface will be smoothed by the framework

Middlebury Benchmark Results

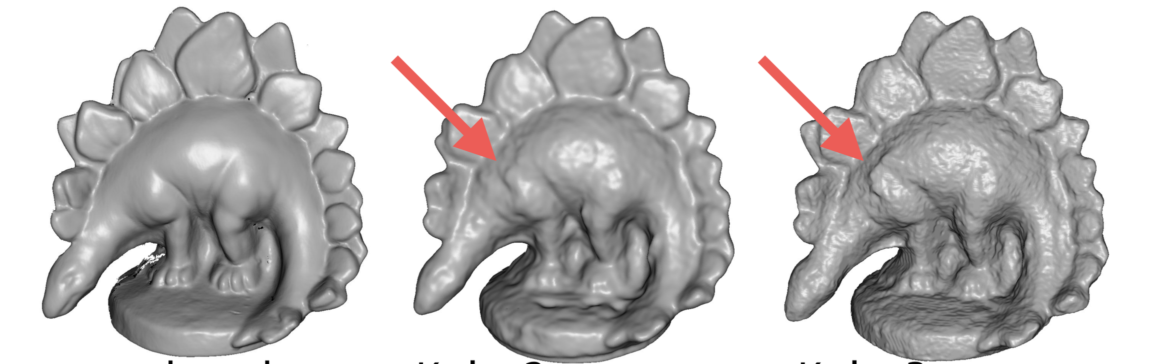
Sort By	Temple Full 312 views Acc [mm]	Temple Full Comp [%]	Temple Ring 47 views Acc [mm]	Temple Ring Comp [%]	Temple Sparse 16 views Acc [mm]	Temple Sparse Comp [%]	Dino Full 363 views Acc [mm]	Dino Full Comp [%]	Dino Ring 48 views Acc [mm]	Dino Ring Comp [%]	Dino Sparse 16 views Acc [mm]	Dino Sparse Comp [%]
Kostrikov			0.57	99.1	0.79	95.8			0.35	99.6	0.37	99.3
Furukawa 2	0.54	99.3	0.55	99.1	0.62	99.2	0.32	99.9	0.33	99.6	0.42	99.2
Zaharescu			0.55	99.2	0.78	95.8			0.42	98.6	0.45	99.2
Furukawa 3	0.49	99.6	0.47	99.6	0.63	99.3	0.33	99.8	0.28	99.8	0.37	99.2
Tsinghua_BBNC											0.3	99.1
Liu2					0.65	96.9					0.51	98.7
Kolev3			0.7	98.3	0.97	92.7			0.42	99.5	0.48	98.6
Schroers	0.57	99.1	0.64	96.4	2.12	62.9	0.33	99.7	0.33	99.7	0.54	98.6
Hernandez	0.36	99.7	0.52	99.5	0.75	95.3	0.49	99.6	0.45	97.9	0.6	98.5
Hongxing	0.83	95.7	0.79	96.3	0.97	93.9	0.62	96.3	0.5	99.1	0.52	98.4
Liu					0.96	89.6					0.59	98.3
Kolev2			0.72	97.8	1.04	91.8			0.43	99.4	0.53	98.3



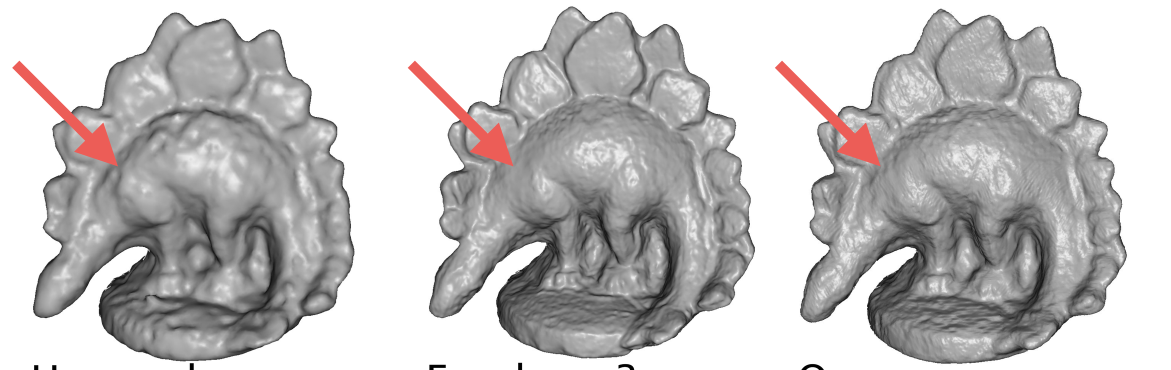
ground truth Kolev2 Kolev3



Hernandez Furukawa3 Ours



ground truth Kolev2 Kolev3



Hernandez Furukawa3 Ours

Volumetric 3D Reconstruction

3D segmentation into foreground and background

$$u : V \to \{0,1\}$$

Labeling costs $\rho_{obj}(\mathbf{x})$ and $\rho_{bck}(\mathbf{x})$

$$E(u) = \underbrace{\int_V \rho(\mathbf{x})|\nabla u(\mathbf{x})|dx}_{\text{smoothing term}} + \lambda \underbrace{\int_V (\rho_{obj}(\mathbf{x}) - \rho_{bck}(\mathbf{x}))u(\mathbf{x})dx}_{\text{labeling cost}}$$

Perform relaxation

$$u : V \to [0,1]$$

Extract object by thresholding

$$u_{min}(\mathbf{x}) = \arg\min_{u(\mathbf{x})} E(u)$$

$$u_\nu(\mathbf{x}) = \mathbf{1}\{u_{min}(\mathbf{x}) \geq \nu\}$$

Standard discontinuity cost $\rho(\mathbf{x})$

We will concentrate on the labeling costs $\rho_{obj}(\mathbf{x})$ and $\rho_{bck}(\mathbf{x})$

Find minimum with Primal-Dual method:

$$\xi_{i,j,k}^{(t+1)} = \Pi_K(\xi_{i,j,k}^{(t)} + \eta\nabla\bar{u}_{i,j,k}^{(t)})$$

$$u_{i,j,k}^{(t+1)} = \Pi_{[0,1]}(u_{i,j,k}^{(t)} + \theta(div(\xi_{i,j,k}^{(t+1)}) - b_{i,j,k}))$$

$$\bar{u}_{i,j,k}^{(t+1)} = 2u_{i,j,k}^{(t+1)} - u_{i,j,k}^{(t)}$$

Labeling Costs

For each voxel $\mathbf{x}$ and for each camera j with location $\mathbf{c}_j$:
Estimate the intersection of camera the view ray r with the surface

$$r_{j,\mathbf{x}}(t) = \mathbf{c}_j + \frac{\mathbf{x} - \mathbf{c}_j}{||\mathbf{x} - \mathbf{c}_j||}t$$

Best consistency score between camera j and all others:

$$S_{j,\mathbf{x},max} = \max_t S_j(r_{j,\mathbf{x}}(t))$$

Corresponding position along the ray $r \to$ depth:

$$t_{j,\mathbf{x},max} = \arg\max_t S_j(r_{j,\mathbf{x}}(t))$$

We can use these values to compute the costs:

$$\rho_{bck}^{i_j}(\mathbf{x}) = -\log(\mu^{\mathbf{1}\{t_{j,\mathbf{x},max} > t_{j,\mathbf{x}}\}}(1-\mu)^{(1-\mathbf{1}\{t_{j,\mathbf{x},max} > t_{j,\mathbf{x}}\})})$$

$$\rho_{obj}^{i_j}(\mathbf{x}) = -\log(\mu^{\mathbf{1}\{t_{j,\mathbf{x},max} < t_{j,\mathbf{x}}\}}(1-\mu)^{(1-\mathbf{1}\{t_{j,\mathbf{x},max} < t_{j,\mathbf{x}}\})})$$

$$\rho_{bck}(\mathbf{x}) - \rho_{obj}(\mathbf{x}) = -\log\prod_{j=1}^{k}\frac{P(B_{i_1}\cdots B_{i_k}|C_{i_j})}{P(O_{i_1}\cdots O_{i_k}|C_{i_j})}$$

$$= \sum_{j=1}^{k}\rho_{bck}^{i_j}(\mathbf{x}) - \sum_{j=1}^{k}\rho_{obj}^{i_j}(\mathbf{x})$$

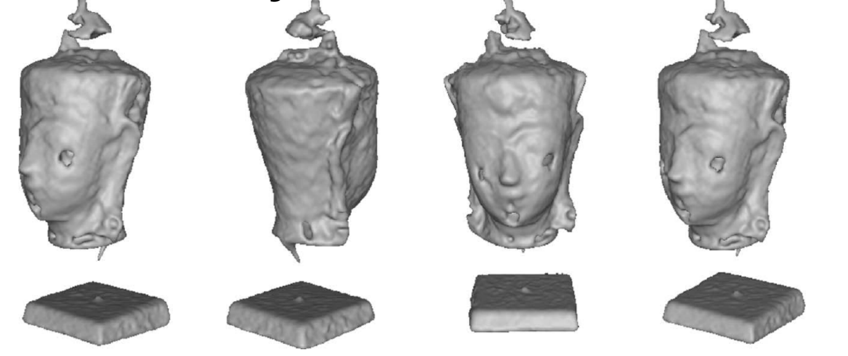
Note that we select this particular subset of cameras:

$$N_d(\mathbf{x}) = \{j \in \{1,\ldots,N\} \mid |t_{j,\mathbf{x},max} - t_{j,\mathbf{x}}| \leq d\}$$

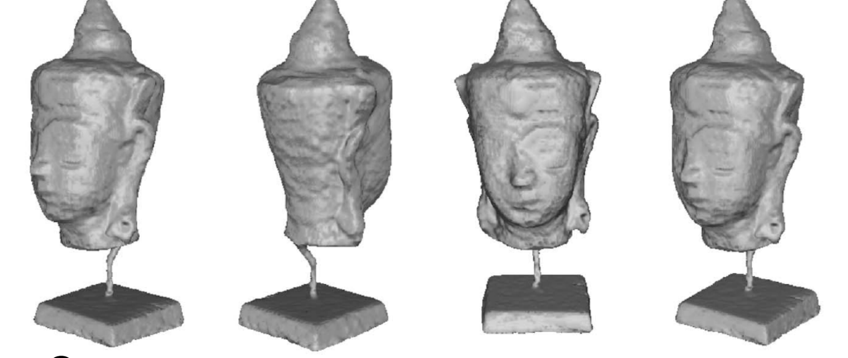
$$d_{min}(\mathbf{x}) = \min_d \text{s. t. } |N_d(\mathbf{x})| \geq k$$

Results Comparison

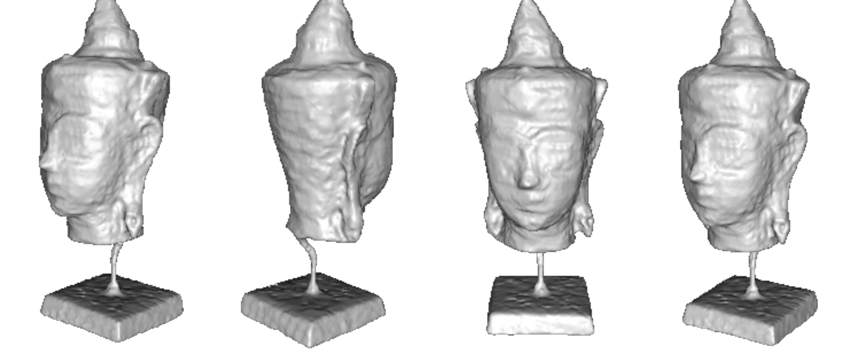
Kolev et al IJCV'09



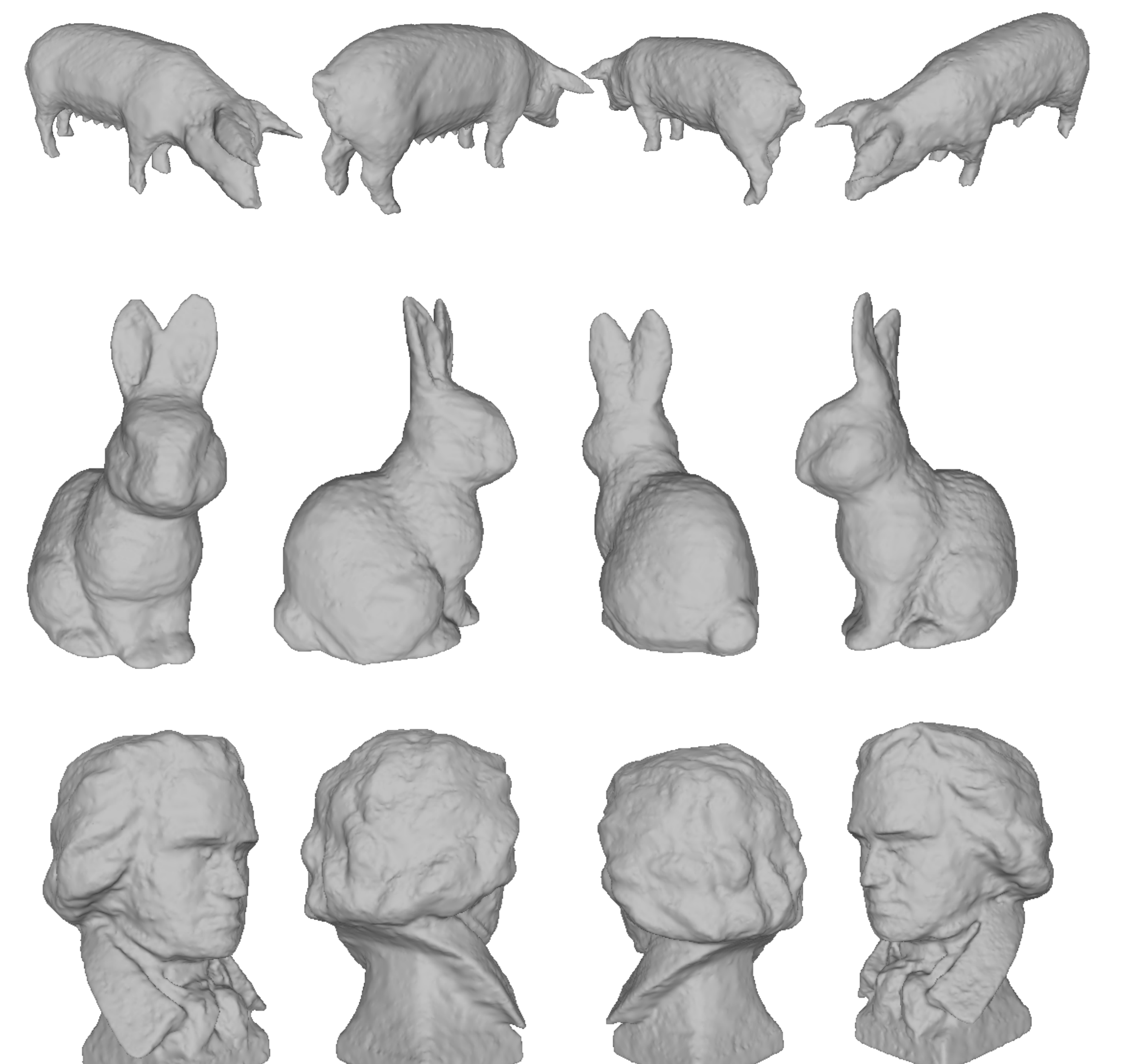
Cremers et al PAMI'11



Ours



Results



ilya.kostrikov@rwth-aachen.de {horbert,leibe}@vision.rwth-aachen.de

www.vision.rwth-aachen.de