Occupancy Networks: Learning 3D Reconstruction in Function Space

Andreas Geiger

Autonomous Vision Group
University of Tübingen / MPI for Intelligent Systems

University of Tübingen
MPI for Intelligent Systems

Autonomous Vision Group

Occupancy Networks:
Learning 3D Reconstruction in Function Space

[Mescheder, Oechsle, Niemeyer, Nowozing & Geiger, CVPR 2019]





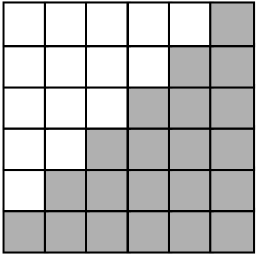
What is a good **output** representation?

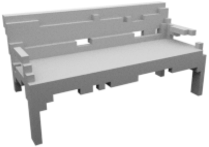
3D Representations

Voxels:

- **Discretization** of 3D space into grid
- Easy to process with neural networks
- Cubic memory $O(n^3) \Rightarrow$ limited resolution
- Manhattan world bias

[Maturana et al., IROS 2015]



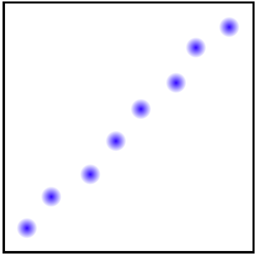


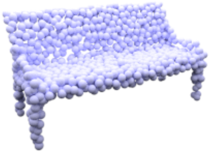
3D Representations

Points:

- **Discretization** of surface into 3D points
- Does not model connectivity / topology
- Limited number of points
- Global shape description

[Fan et al., CVPR 2017]



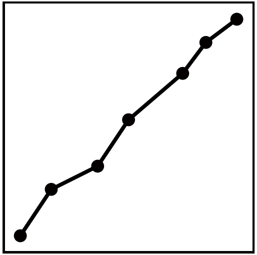


3D Representations

Meshes:

- **Discretization** into vertices and faces
- Limited number of vertices / granularity
- Requires class-specific template – or –
- Leads to self-intersections

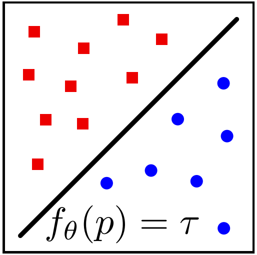
[Groueix et al., CVPR 2018]



3D Representations

This work:

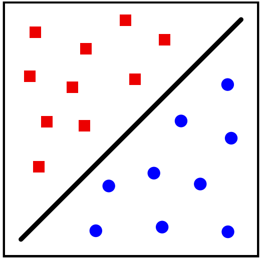
- Implicit representation $\Rightarrow$ **No discretization**
- Arbitrary topology & resolution
- Low memory footprint
- Not restricted to specific class



Occupancy Networks

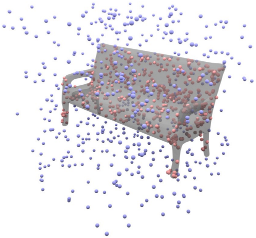
Key Idea:

- Do not represent 3D shape explicitly
- Instead, consider surface **implicitly** as **decision boundary** of a non-linear classifier:



$$f_\theta : \mathbb{R}^3 \times \mathcal{X} \rightarrow [0,1]$$

3D Location — Condition (eg, Image) — Occupancy Probability



Concurrent work:

- DeepSDF [Park et al., CVPR 2019]
- IM-NET [Chen et al., CVPR 2019]

Mescheder, Oechsle, Niemeyer, Nowozin and Geiger: Occupancy Networks: Learning 3D Reconstruction in Function Space. CVPR, 2019.

Training Objective

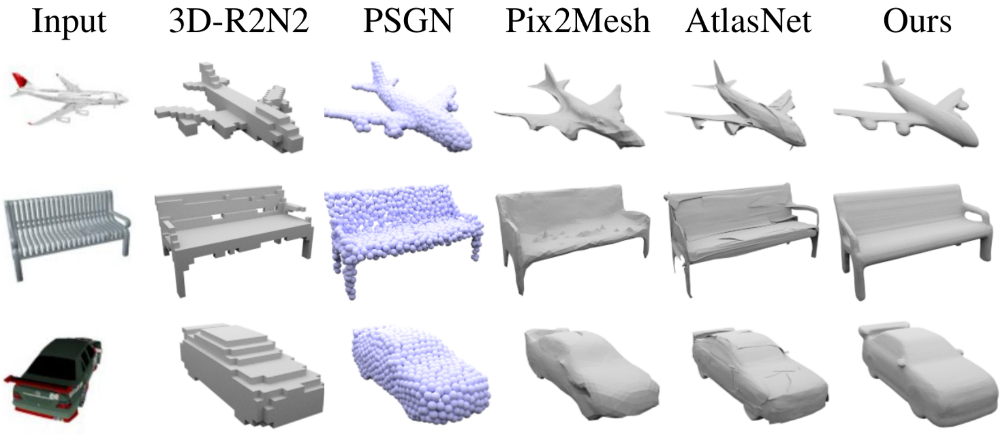
Occupancy Network:Variational Occupancy Encoder:

$$\mathcal{L}(\theta, \psi) = \sum_{j=1}^{K} \text{BCE}(f_\theta(p_{ij}, z_i), o_{ij}) + KL\left[q_\psi(z|(p_{ij}, o_{ij})_{j=1:K}) \,\|\, p_0(z)\right]$$

- K: Randomly sampled 3D points ($K = 2048$)
- BCE: Cross-entropy loss
- q_ψ: Encoder

Mescheder, Oechsle, Niemeyer, Nowozin and Geiger: Occupancy Networks: Learning 3D Reconstruction in Function Space. CVPR, 2019.

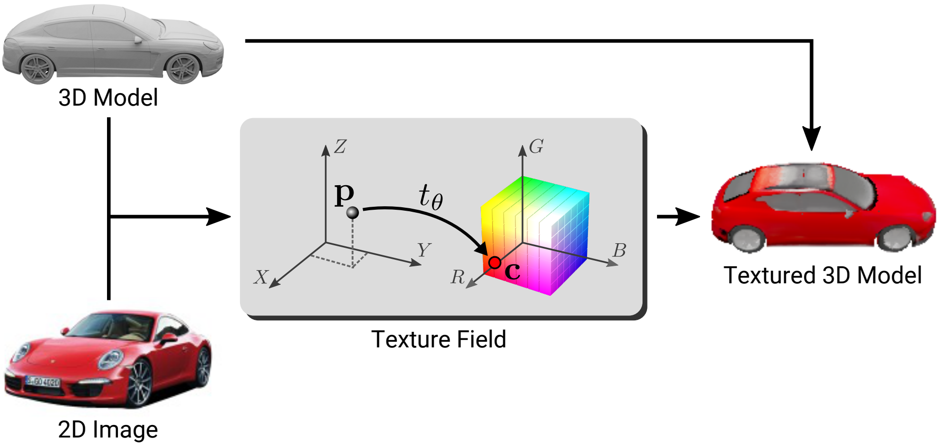
Results



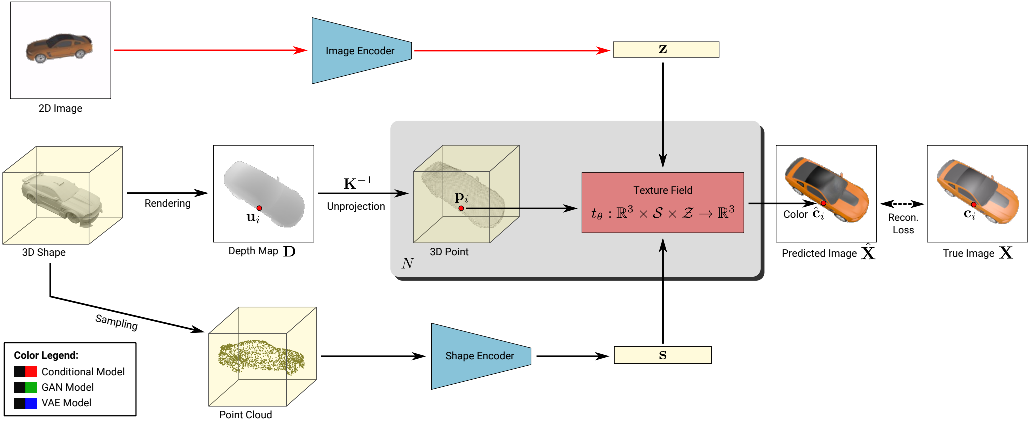
Mescheder, Oechsle, Niemeyer, Nowozin and Geiger: Occupancy Networks: Learning 3D Reconstruction in Function Space. CVPR, 2019.

Can we also learn about object **appearance**?

Texture Fields

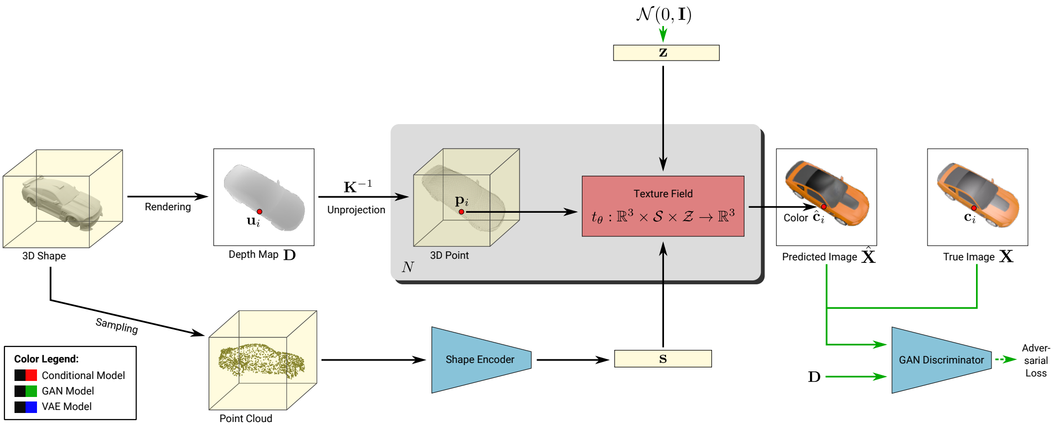


Texture Fields



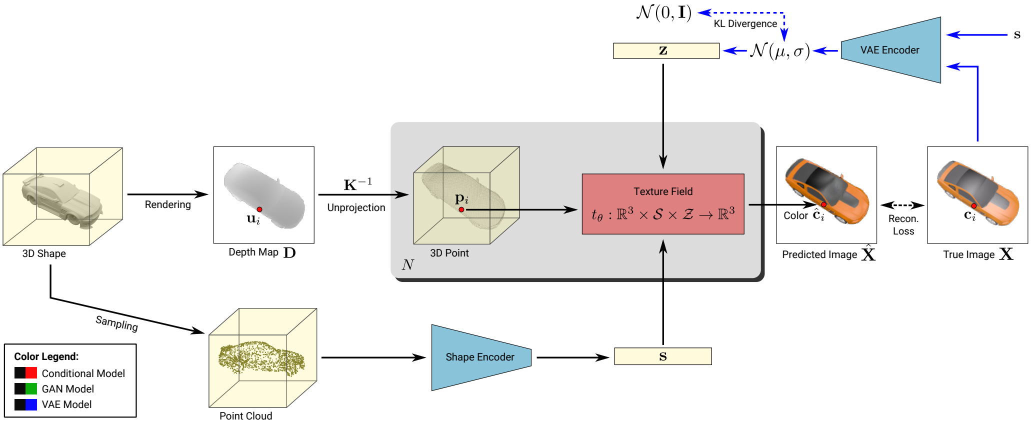
Oechsle, Mescheder, Niemeyer, Strauss and Geiger: Texture Fields: Learning Texture Representations in Function Space. ICCV, 2019.

Texture Fields



Oechsle, Mescheder, Niemeyer, Strauss and Geiger: Texture Fields: Learning Texture Representations in Function Space. ICCV, 2019.

Texture Fields



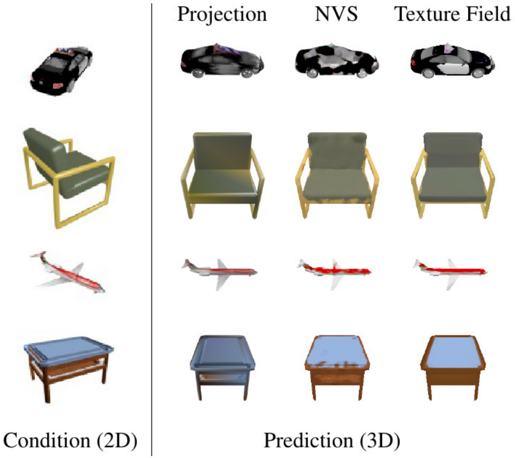
Oechsle, Mescheder, Niemeyer, Strauss and Geiger: Texture Fields: Learning Texture Representations in Function Space. ICCV, 2019.

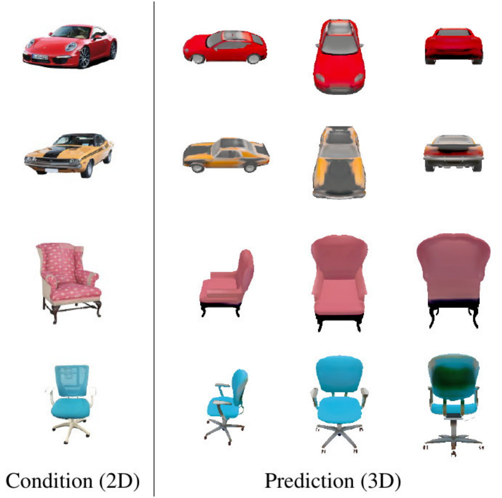
Representation Power

- Ground truth vs. Texture Field vs. Voxelization

Oechsle, Mescheder, Niemeyer, Strauss and Geiger: Texture Fields: Learning Texture Representations in Function Space. ICCV, 2019.

Results

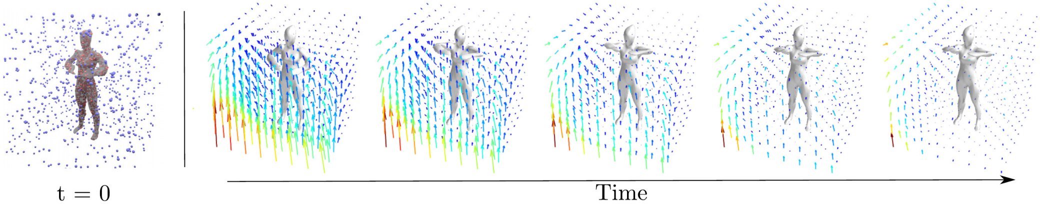




Oechsle, Mescheder, Niemeyer, Strauss and Geiger: Texture Fields: Learning Texture Representations in Function Space. ICCV, 2019.

What about object **motion**?

Occupancy Flow

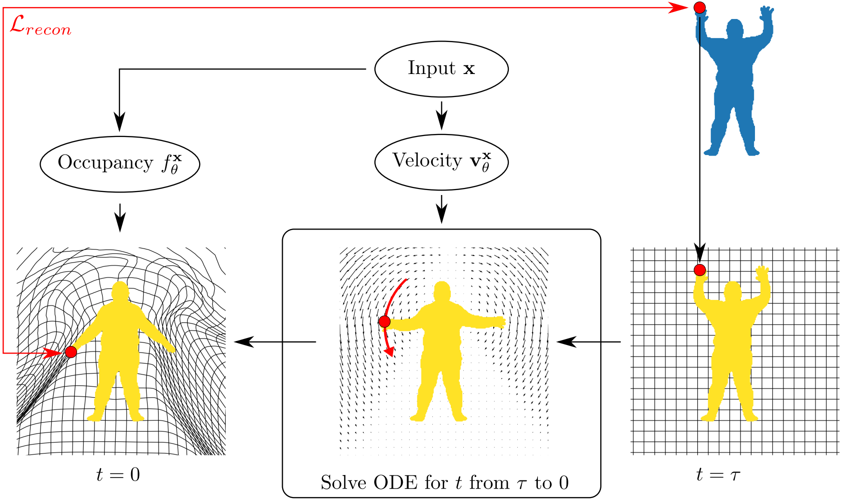


- Extending Occupancy Networks to 4D is hard (curse of dimensionality)
- Represent shape at $t = 0$ using a 3D Occupancy Network
- Represent motion by temporally and spatially continuous vector field
- Relationship between 3D trajectory $\mathbf{s}$ and velocity $\mathbf{v}$ given by (differentiable) ODE:

$$\frac{\partial \mathbf{s}(t)}{\partial t} = \mathbf{v}(\mathbf{s}(t), t)$$

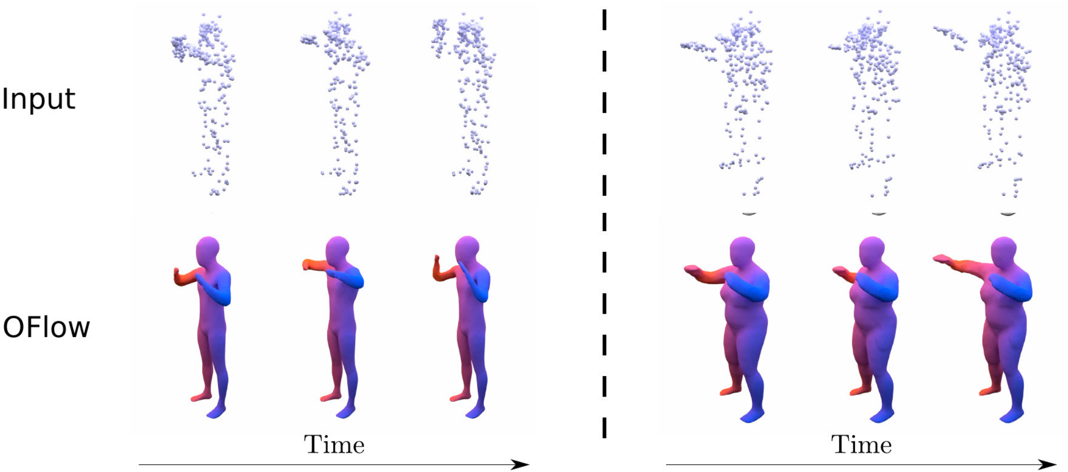
Niemeyer, Mescheder, Oechsle and Geiger: Occupancy Flow: 4D Reconstruction by Learning Particle Dynamics. ICCV, 2019.

Occupancy Flow



Niemeyer, Mescheder, Oechsle and Geiger: Occupancy Flow: 4D Reconstruction by Learning Particle Dynamics. ICCV, 2019.

Results



- No correspondences needed ⇒ implicitly established by our model!

Niemeyer, Mescheder, Oechsle and Geiger: Occupancy Flow: 4D Reconstruction by Learning Particle Dynamics. ICCV, 2019.

Summary

Thank you!

http://autonomousvision.github.io

European Research Council
Established by the European Commission

Federal Ministry of Education and Research

Microsoft Research