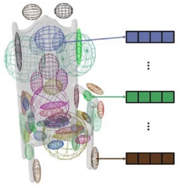
Presentation's Topics

Presentation's Topics

- Local Deep Implicit Function - LDIF

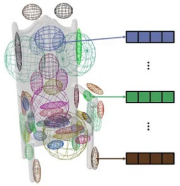
Dec 2019 and CVPR 2020



Kyle Genova, Forrester Cole, **Avneesh Sud**, Aaron Sarna, **Thomas Funkhouser**

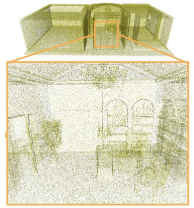
Presentation's Topics

- Local Deep Implicit Function - LDIF *Dec 2019 and CVPR 2020*



Kyle Genova, Forrester Cole, **Avneesh Sud**, Aaron Sarna, **Thomas Funkhouser**

- Local Implicit Grid - LIG *Mar 2020 and CVPR 2020*



Chiyu Max Jiang, **Avneesh Sud**, Ameesh Makadia, Jingwei Huang, Matthias Nießner, **Thomas Funkhouser**

Local Deep Implicit Functions for 3D Shape (LDIF)

CVPR 2020

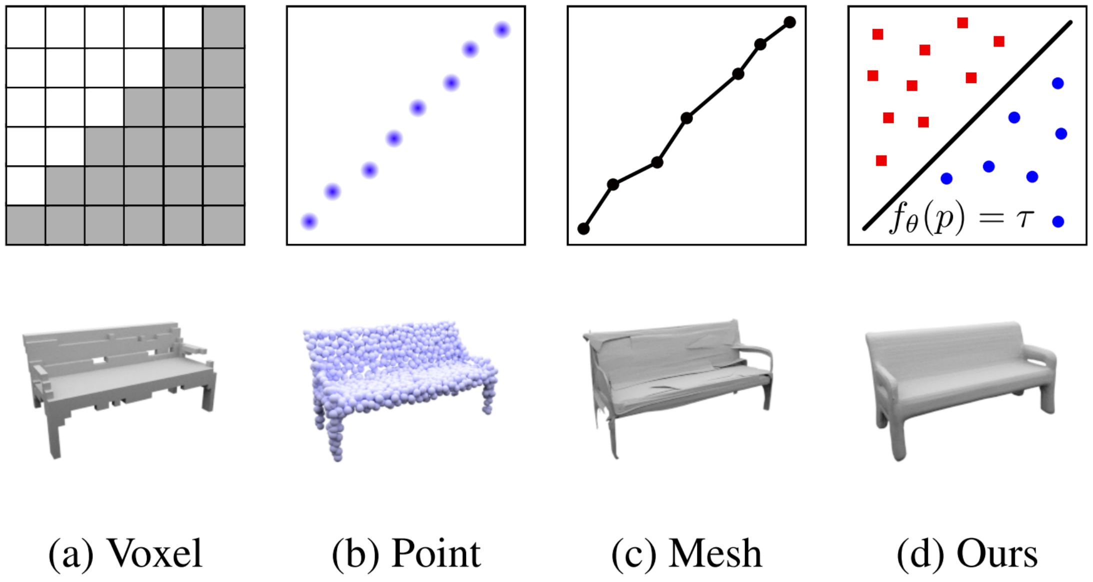
Credit to CS348I and specifically to Sharon Zhang, Chaitanya Patel, David Braun, Eric Zelikman

Motivation

Deep Implicit Functions **DIF**

(OccNet, IM-NET, DeepSDF, DISN)

- Do NOT generalize well outside training classes
- Do NOT capture fine details
- Hard to scale



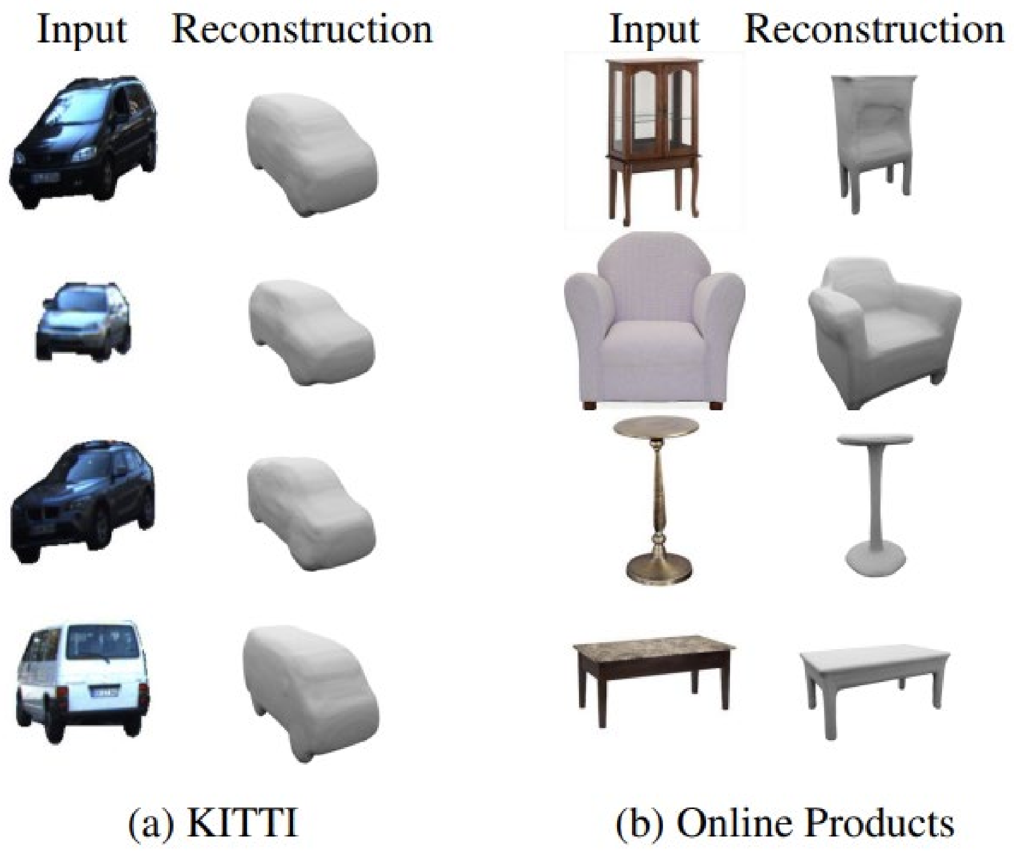
Occupancy Networks: Learning 3D Reconstruction in Function Space
Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger

Motivation

Deep Implicit Functions **DIF**

(OccNet, IM-NET, DeepSDF, DISN)

- **Do NOT generalize well outside training classes**
- Do NOT capture fine details
- Hard to scale



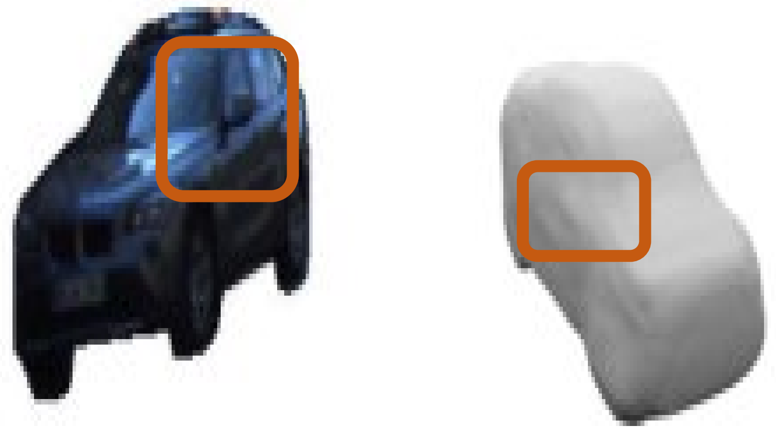
Occupancy Networks: Learning 3D Reconstruction in Function Space
Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger

Motivation

Deep Implicit Functions **DIF**

(OccNet, IM-NET, DeepSDF, DISN)

- Do NOT generalize well outside training classes
- **Do NOT capture fine details**
- Hard to scale



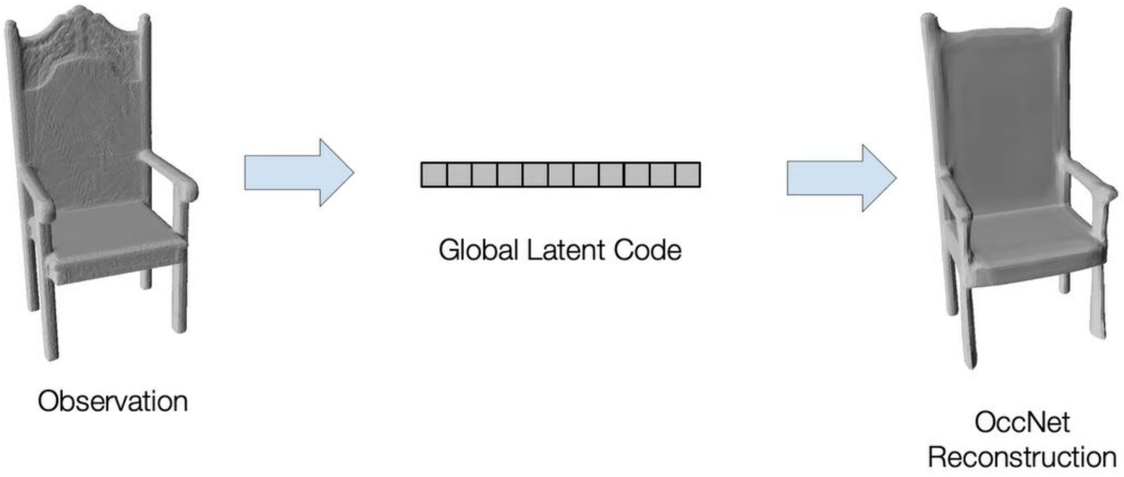
Occupancy Networks: Learning 3D Reconstruction in Function Space
Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger

Motivation

Deep Implicit Functions **DIF**

(OccNet, IM-NET, DeepSDF, DISN)

- Do NOT generalize well outside training classes
- Do NOT capture fine details
- **Hard to scale**



Occupancy Networks: Learning 3D Reconstruction in Function Space
Lars Mescheder, Michael Oechsle, Michael Niemeyer, Sebastian Nowozin, Andreas Geiger

SIF: Structured Implicit Function

3D position

$$F(\mathbf{x}, \boldsymbol{\Theta}) = \overbrace{\sum_{i \in [N]} f_i(\mathbf{x}, \theta_i)}^{\textit{Sum of the local influence}}$$

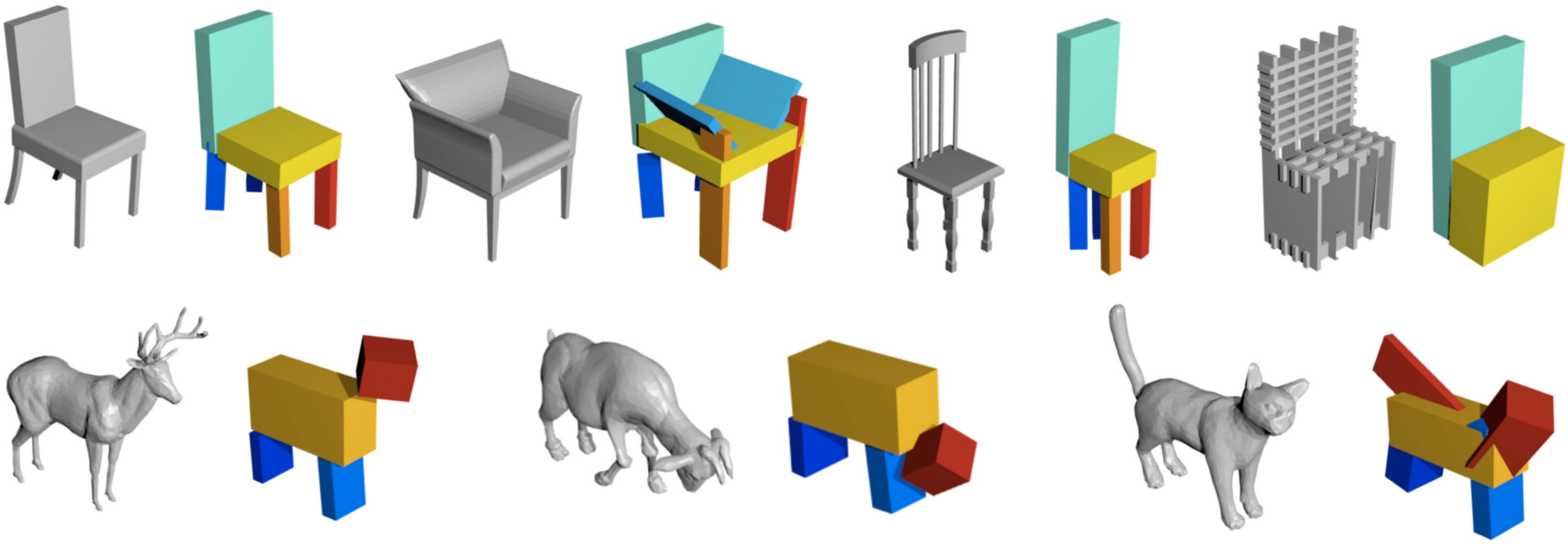
Vector of Template Parameters

$$f_i(\mathbf{x}, \theta_i) = c_i \exp \overbrace{\left(\sum_{d \in \{x,y,z\}} \frac{-(\mathbf{p}_{i,d} - \mathbf{x}_d)^2}{2\mathbf{r}_{i,d}^2} \right)}^{\textit{Set of squished and stretched 3D blobs}}$$

Learning Shape Templates with Structured Implicit Functions

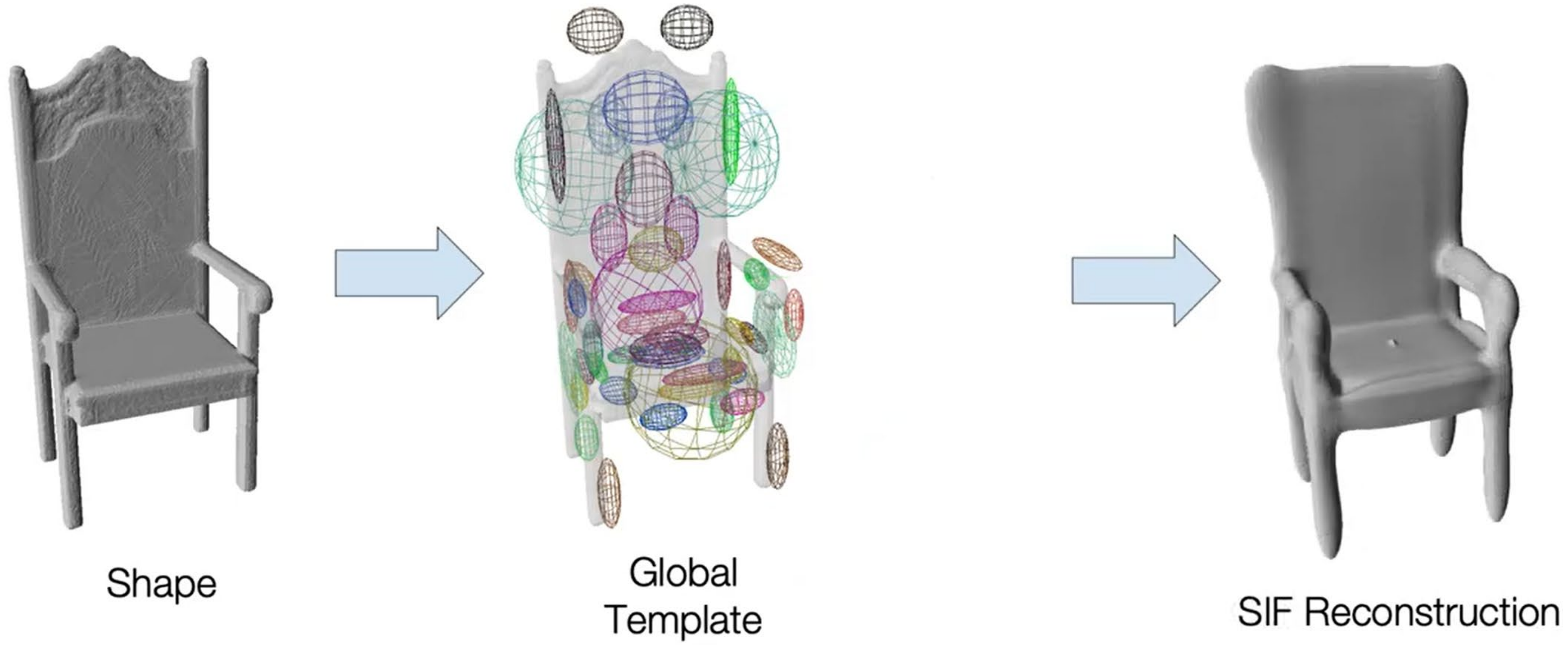
[**Kyle Genova**](), [**Forrester Cole**](), [Daniel Vlasic](), [**Aaron Sarna**](), [William T. Freeman](), [**Thomas Funkhouser**]()

VP: Volumetric Primitives

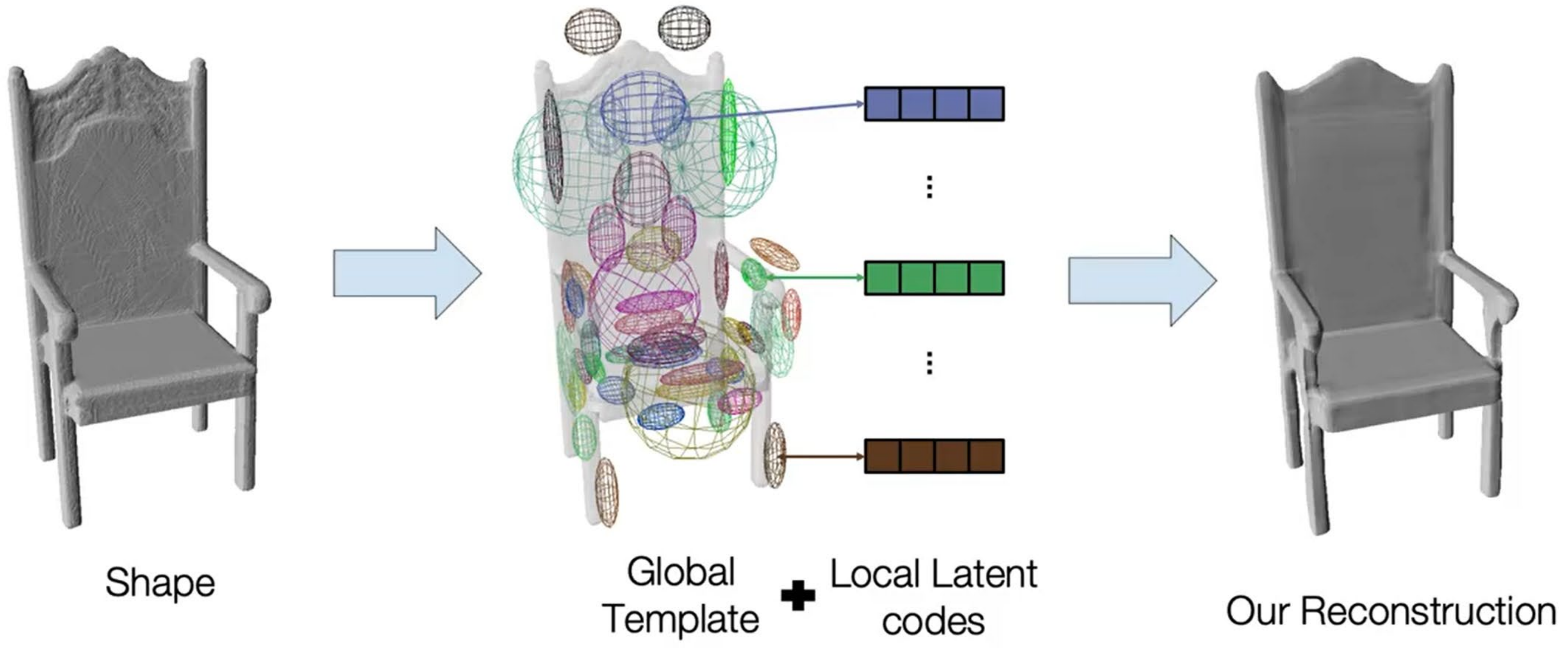


Learning Shape Abstractions by Assembling Volumetric Primitives
Shubham Tulsiani, Hao Su, Leonidas J. Guibas, Alexei A. Efros, Jitendra Malik

LDIF: is similar to SIF but…



LDIF: Local Deep Implicit Functions



LDIF: Local Deep Implicit Functions

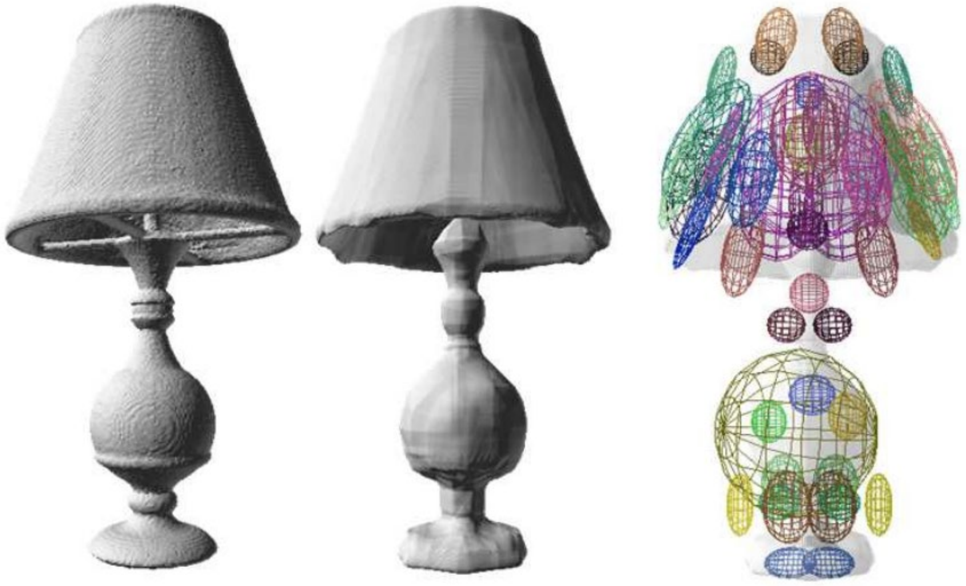
$$\mathrm{LDIF}(\mathbf{x}, \mathbf{\Theta}, \mathbf{Z}) = \overbrace{\sum_{i \in [N]} g(\mathbf{x}, \theta_i)(1 + f(\mathbf{x}, \mathbf{z}_i))}^{\textit{Set of Local DIF blended to a SIF template (Overlapping)}}$$

Analytic shape Variable

DIF (fine details)

Shape elements

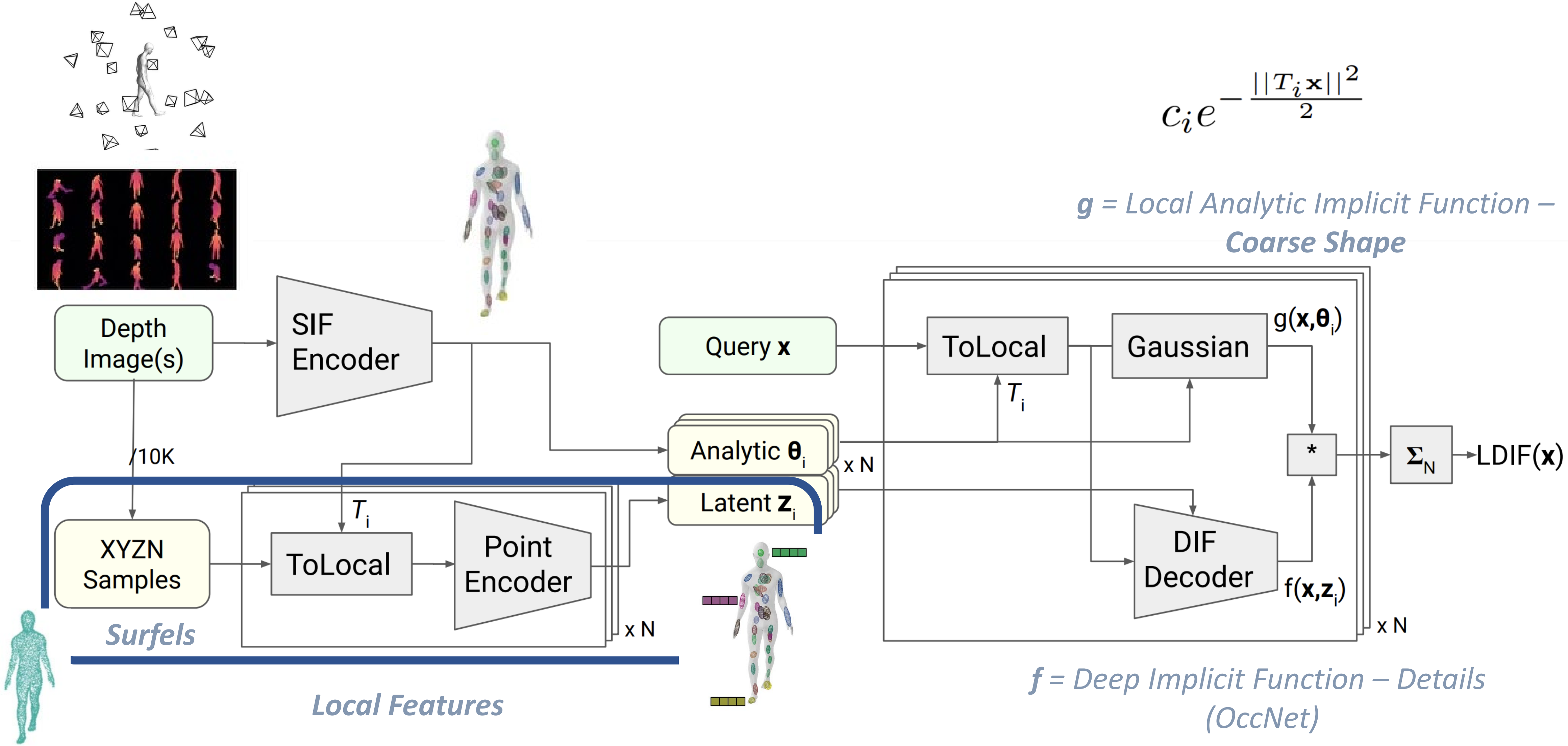
$$\overbrace{g(\mathbf{x}, \theta_i) = c_i e^{-\frac{||T_i \mathbf{x}||^2}{2}}}^{\textit{Coarse SIF Template Oriented Anisotropic Gaussian}}$$



- *1 Variable for a scale constant c*
- *9 Variables for **T** Affine Transformation with*
    - *3 for a center point **p***
    - *3 radii **r***
    - *3 Euler angles **e***

Network

$$c_i e^{-\frac{||T_i \mathbf{x}||^2}{2}}$$



Losses

$$L(\mathbf{\Theta}, \mathbf{Z}) = w_P L_P(\mathbf{\Theta}, \mathbf{Z}) + w_C L_C(\mathbf{\Theta})$$

Point Sample Loss

$$L_P(\mathbf{\Theta}, \mathbf{Z}) = \frac{1}{|\mathcal{C}|} \sum_{\mathbf{x_i} \in \mathcal{C}} w_i \|\mathtt{sig}(\alpha \mathrm{LDIF}(\mathbf{x_i}, \mathbf{\Theta}, \mathbf{Z})) - I(\mathbf{x_i})\|$$

Shape Element Center Loss

$$L_C(\mathbf{\Theta}) = \begin{cases} \sum_{\theta_\mathbf{i} \in \mathbf{\Theta}} G(\mathbf{p}_i)^2 & G(\mathbf{p}_i) > \beta \\ 0 & G(\mathbf{p}_i) \leq \beta \end{cases}$$

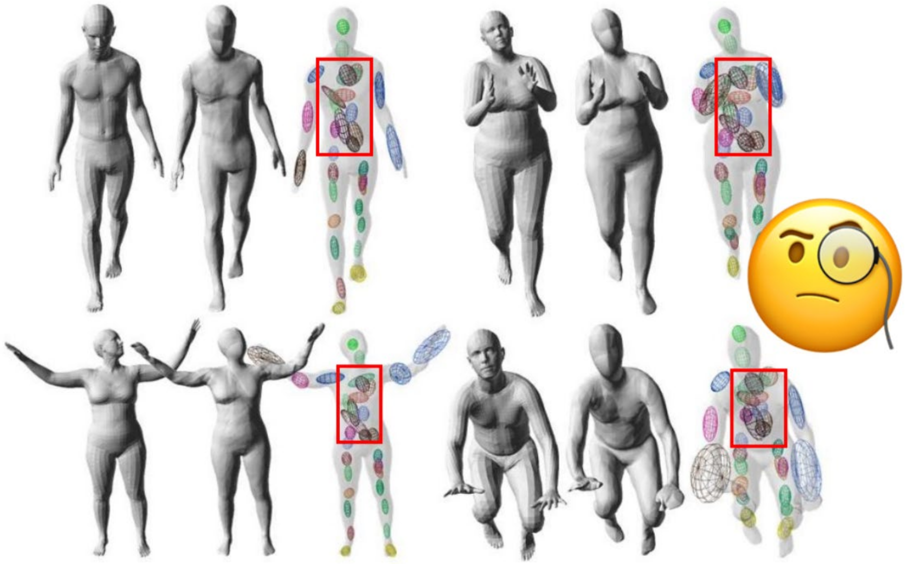
Experiments

PROS

- Compact Storage (1% of OccNet parameters)
- Efficient Computation
- Consistent for similar shapes
- Generalize well for unseen categories (F-score +16)

CONS

- Flat Set of Local Region, no multiresolution hierarchy
- Uses known camera poses for reconstruction from depth images (and calibration)
- Constant number of local regions



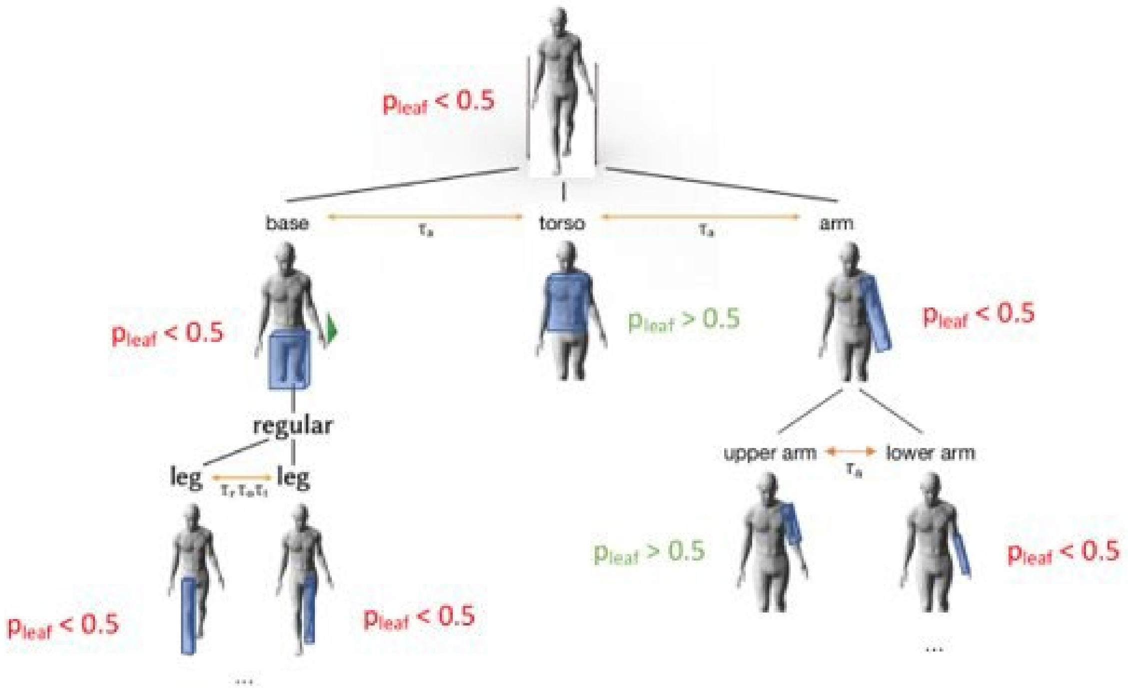
Credit to Sharon Zhang – PhD Student

PROS

- Compact Storage (1% of OccNet parameters)
- Efficient Computation
- Consistent for similar shapes
- Generalize well for unseen categories (F-score +16)

CONS

- Flat Set of Local Region, no multiresolution hierarchy
- Uses known camera poses for reconstruction from depth images (and calibration)
- Constant number of local regions



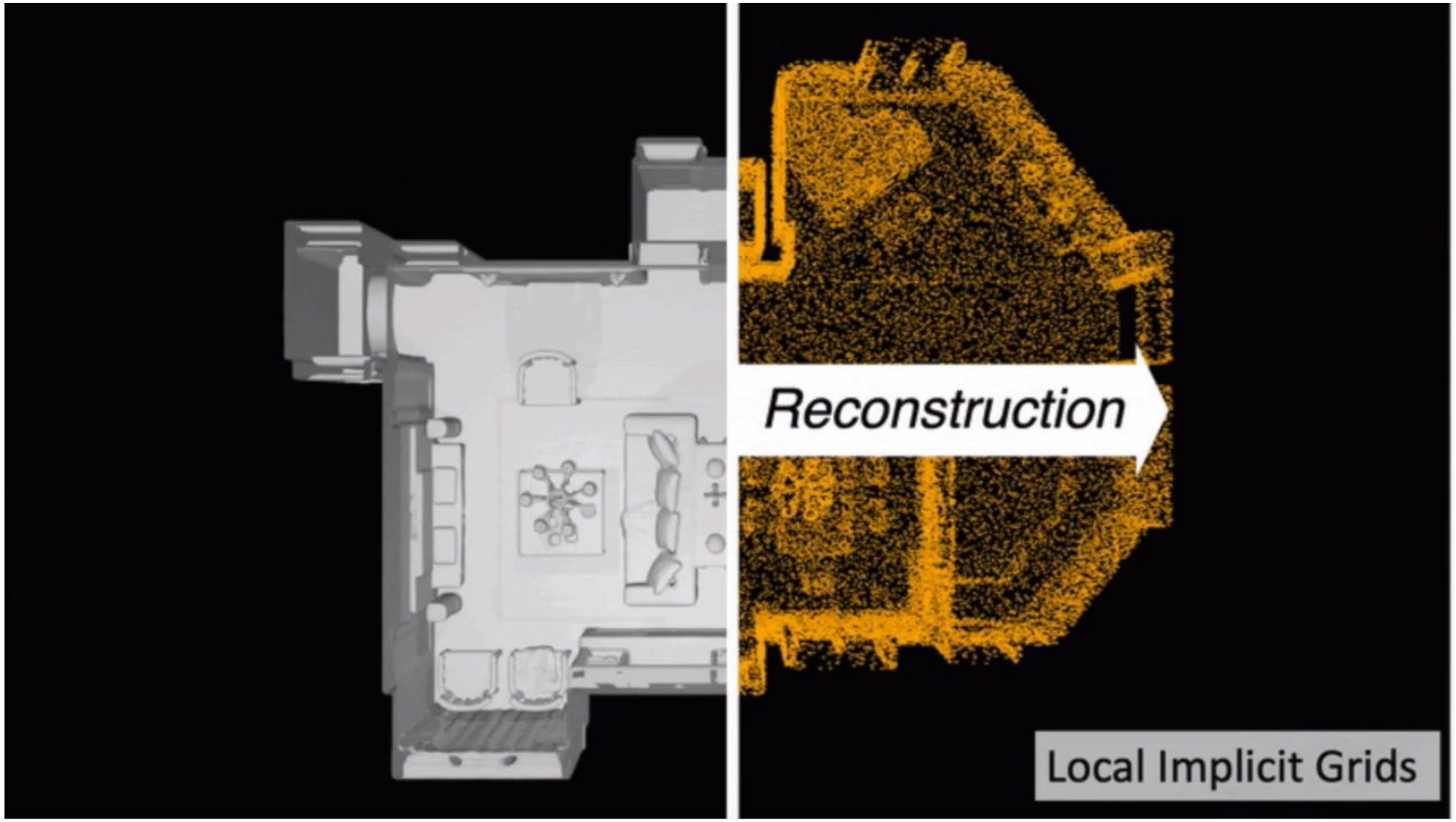
Credit to Sharon Zhang – PhD Student

Local Implicit Grid Representations for 3D Scenes (LIG)

CVPR 2020

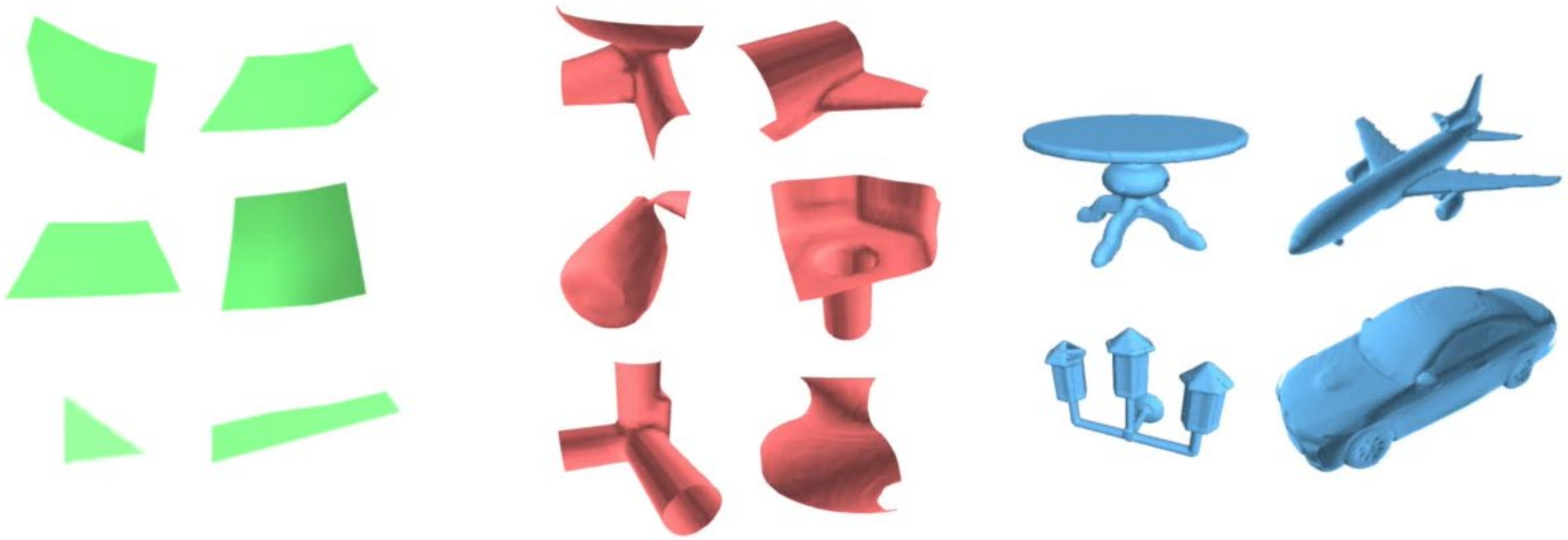
Motivation

- **Scalable**
- Efficient



Motivation

- Scalable
- **Efficient**



(a) Micro Scale

Trivial abstraction

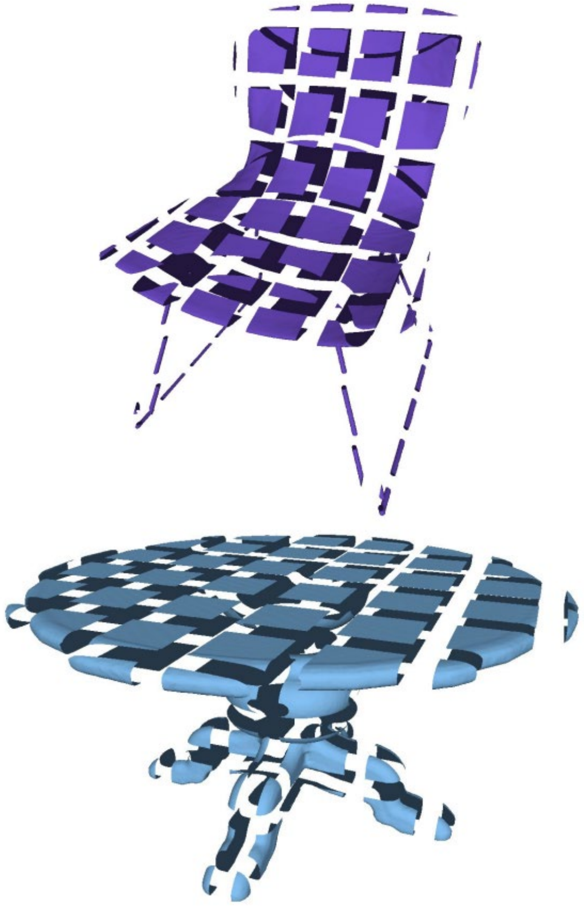
(b) Part Scale

Generalizable abstraction

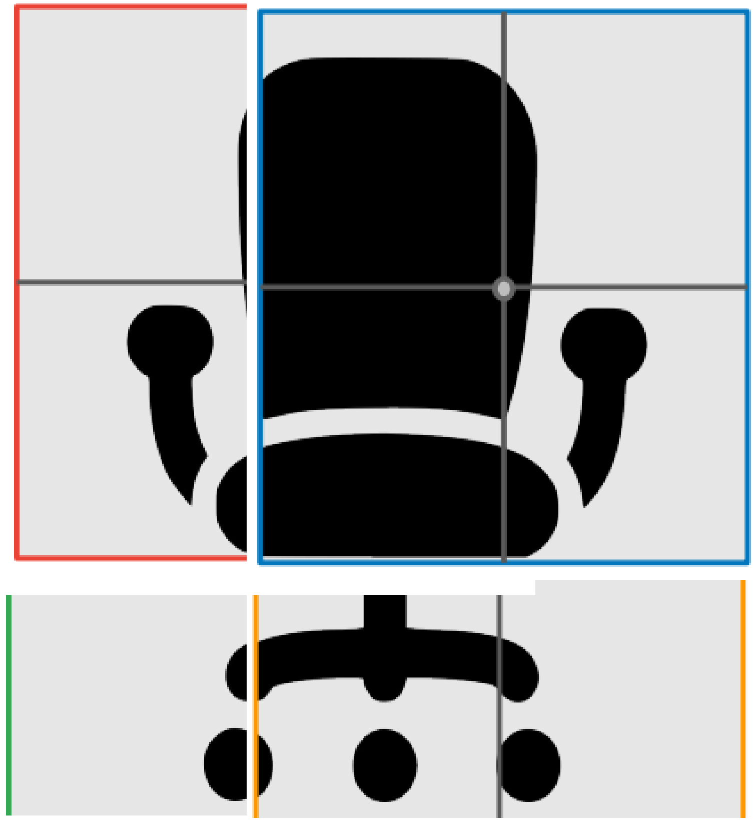
(c) Object Scale

Does not generalize

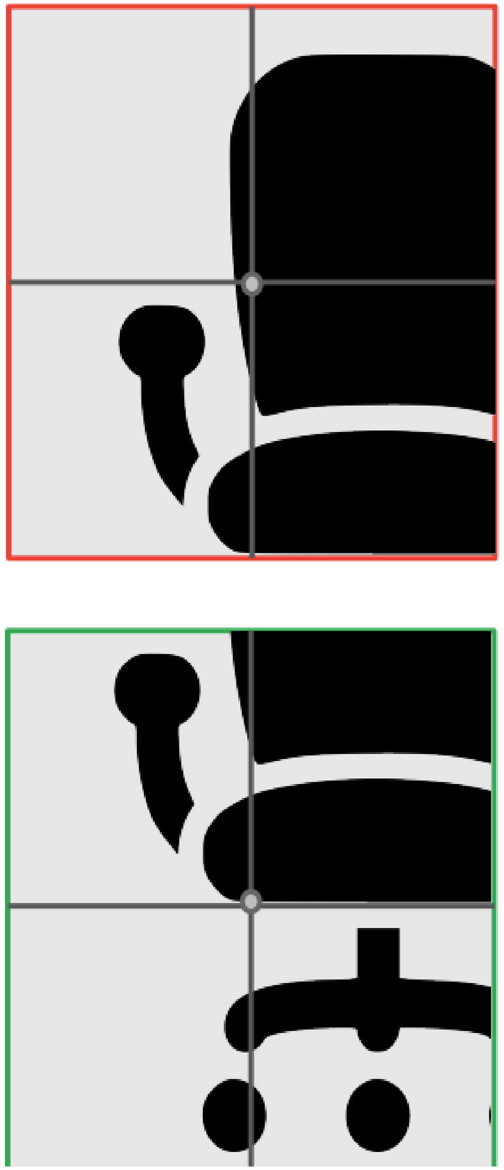
Method

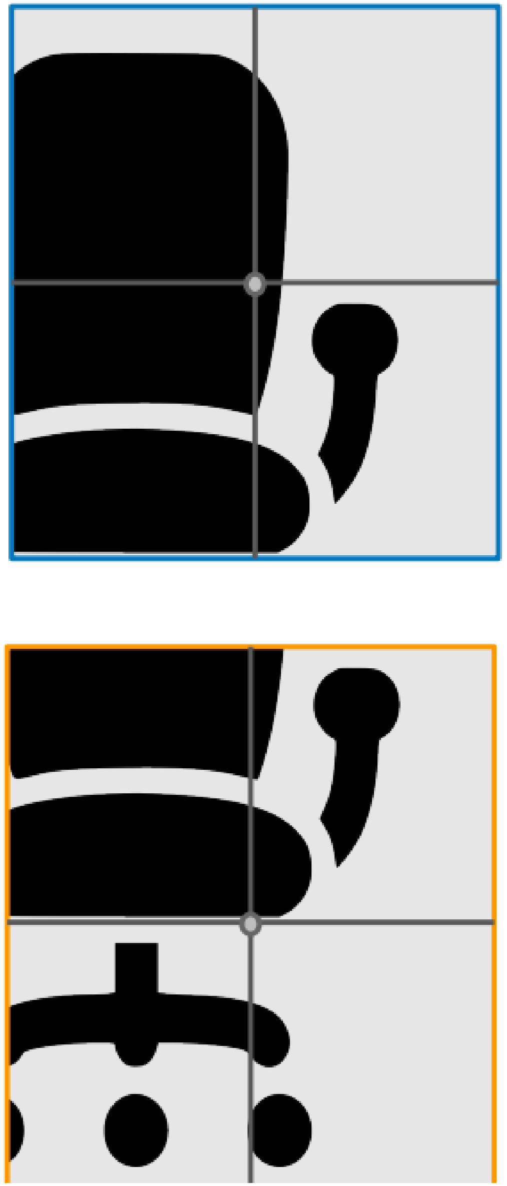


LDIF	LIG

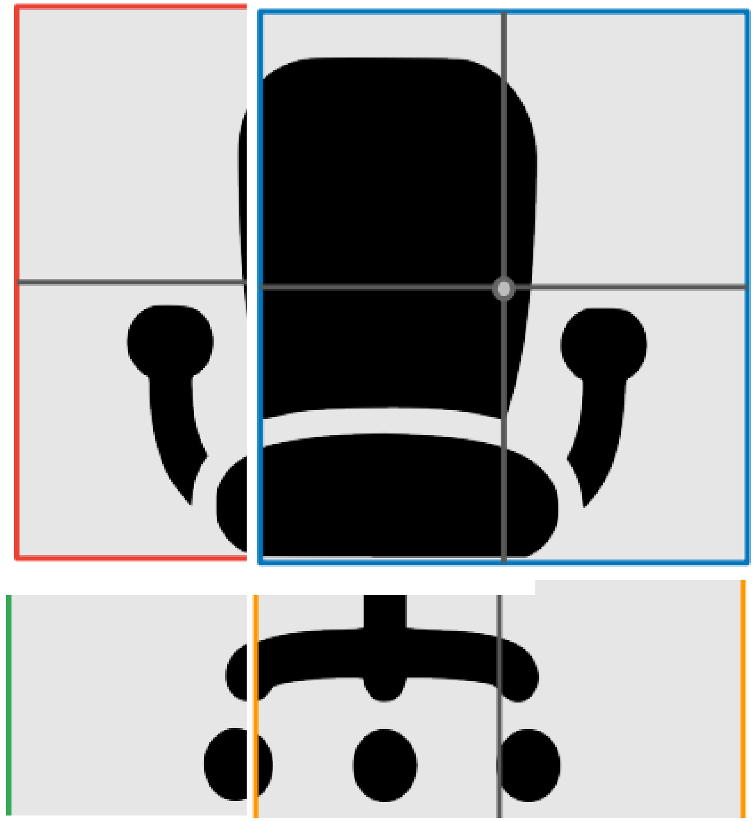


LDIF	LIG

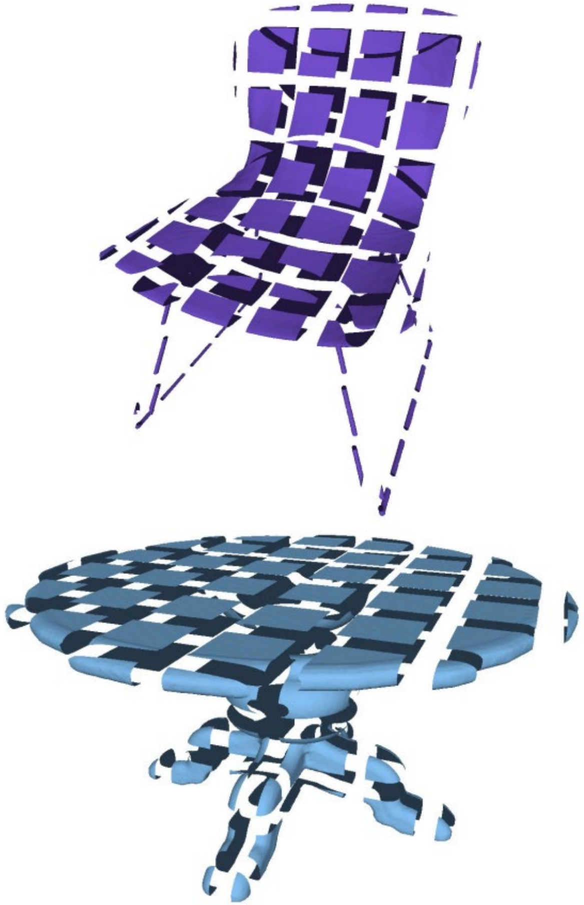


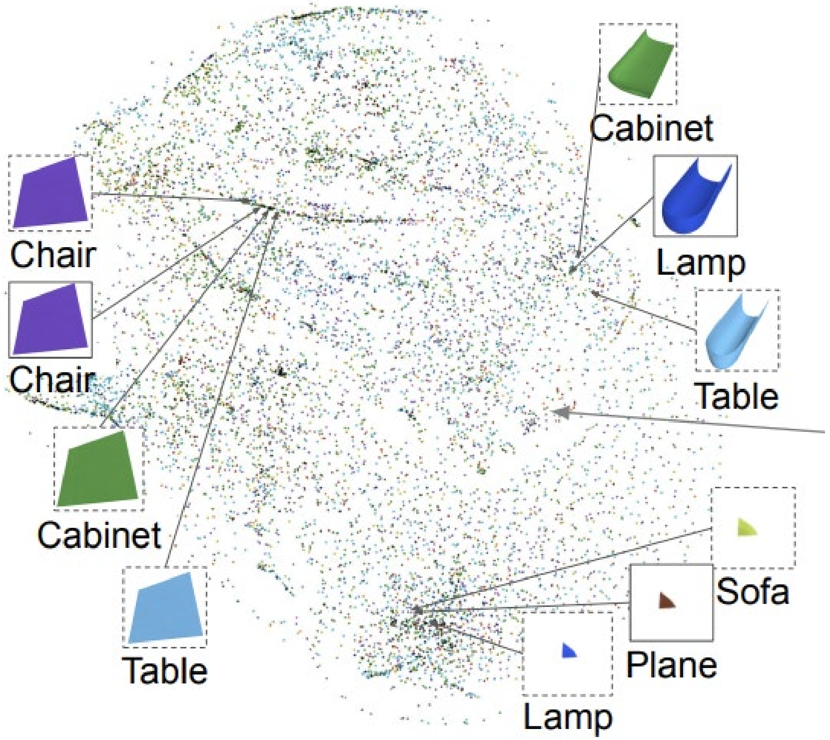


LDIF	LIG

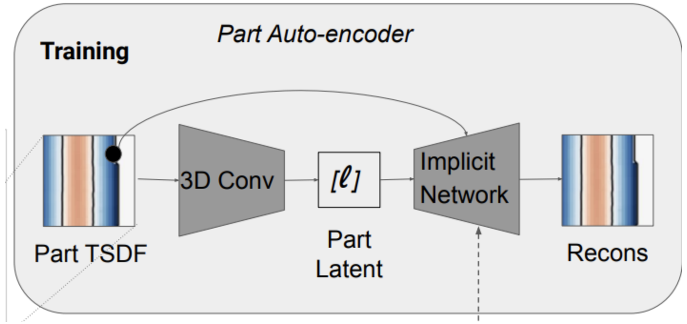


Method

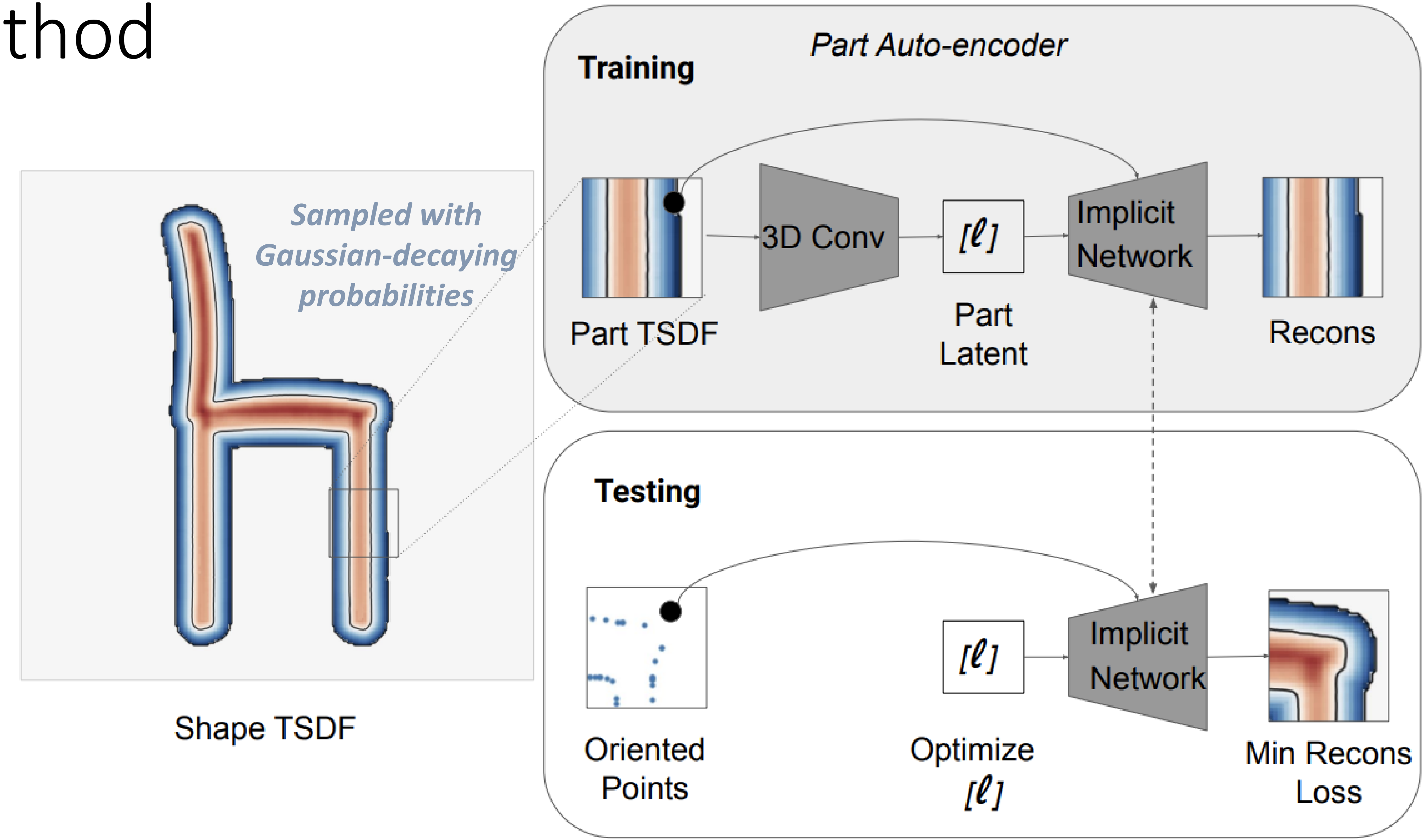




Embedding



Method


Sampled with
Gaussian-decaying
probabilities
Shape TSDF
Training
Part Auto-encoder
Part TSDF
3D Conv
[ℓ]
Part
Latent
Implicit
Network
Recons
Testing
Oriented
Points
[ℓ]
Optimize
[ℓ]
Implicit
Network
Min Recons
Loss

Losses for Part AutoEncoder

$$\mathcal{L}(\theta_e, \theta_d) = \overbrace{\frac{1}{|\mathcal{P}||\mathcal{B}|} \sum_{i \in \mathcal{P}} \sum_{j \in \mathcal{B}} \mathcal{L}_c(D_{\theta_d}(\boldsymbol{x}_{i,j}, E_{\theta_e}(g_i)), \mathrm{sign}(\boldsymbol{x}_{i,j}))}^{\text{Binary Cross Entropy Loss (with logits)}} \underbrace{+ \lambda ||E_{\theta_e}(g_i)||_2}_{\text{Regularization Loss}}$$

Set of training parts

Mini-Batch

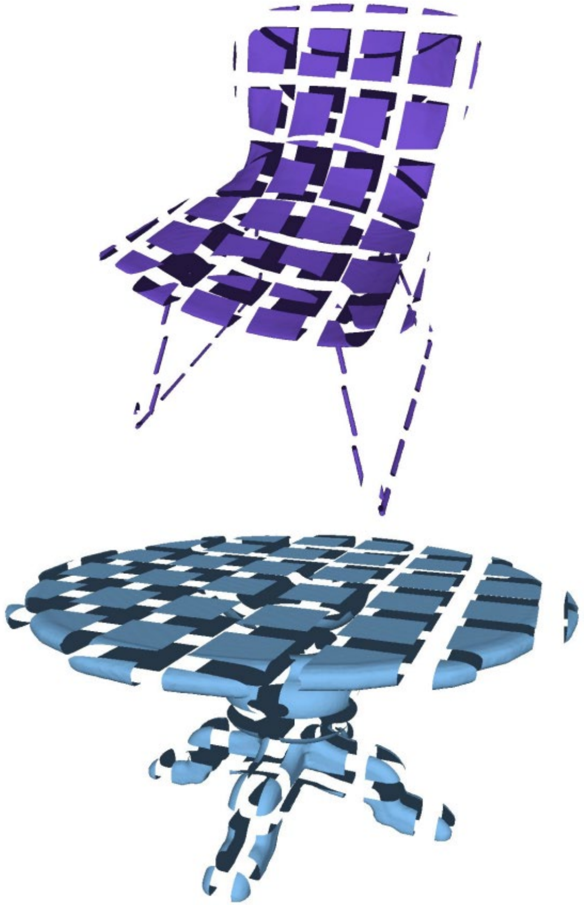
Implicit Decoder

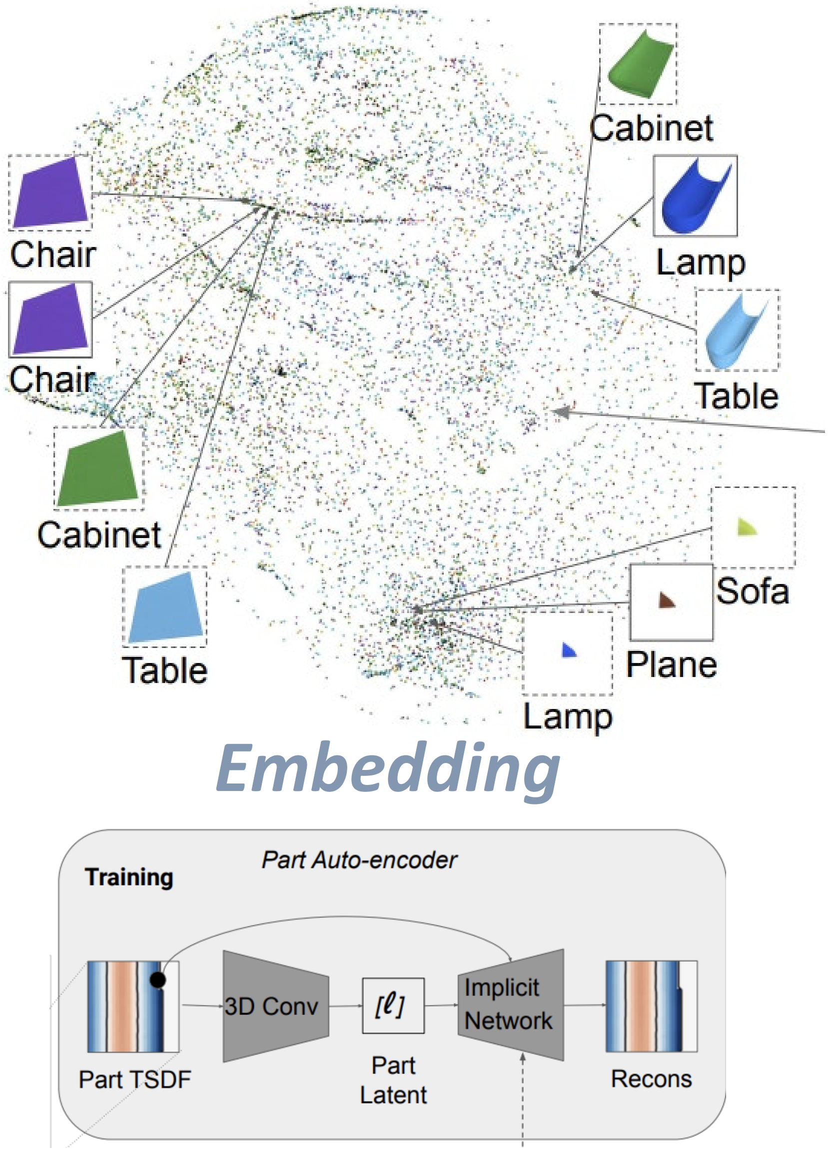
Conv Encoder

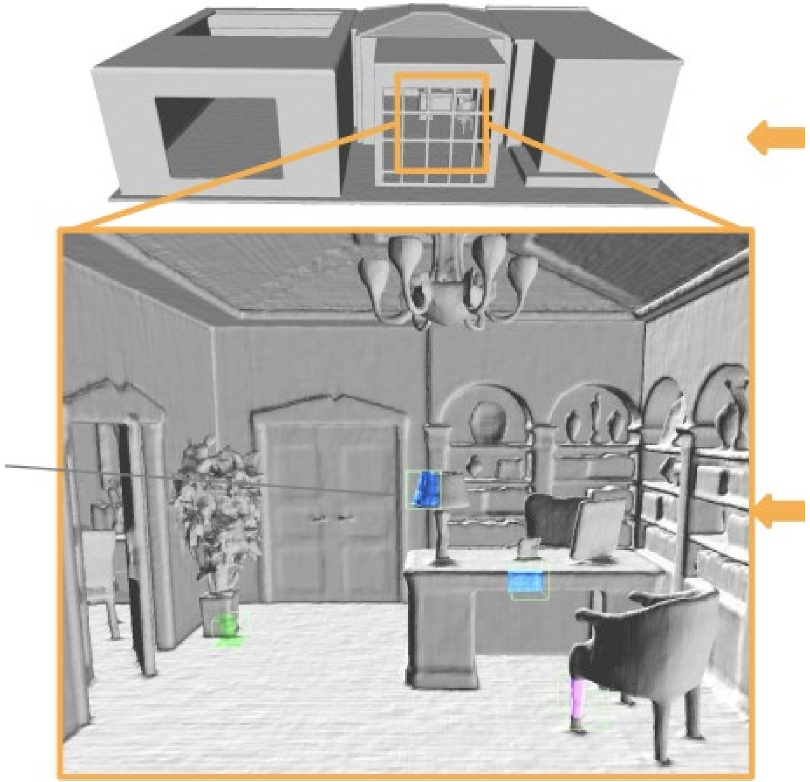
TSDF grid

LDIF	LIG

Method







LIG: Local Implicit Grid

Latent Code of the part

$$f(\boldsymbol{x}, \boldsymbol{c}_i) = D_{\theta_d}(\boldsymbol{c}_i, \frac{2}{s}(\boldsymbol{x} - \boldsymbol{x}_i))$$

Query location

Single Voxel Grid Cell Center

Part Scale

LIG: Overlapping

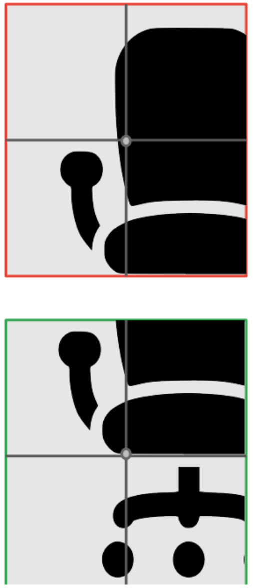
Latent Code of the part

$$f(\boldsymbol{x}, \boldsymbol{c}_i) = D_{\theta_d}(\boldsymbol{c}_i, \frac{2}{s}(\boldsymbol{x} - \boldsymbol{x}_i))$$

Query location

Part Scale

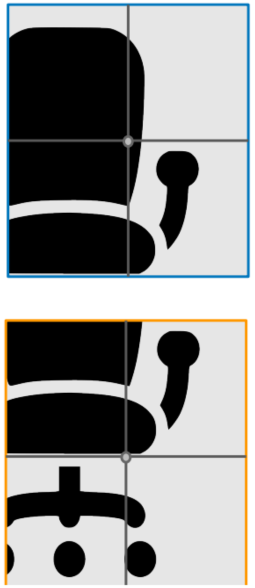
Single Voxel Grid Cell Center



Overlapping Latent Grid

$$f(\boldsymbol{x}, \{\boldsymbol{c}_j | j \in \mathcal{N}\}) = \sum_{j \in \mathcal{N}} w_j D_{\theta_d}(\boldsymbol{c}_j, \frac{2}{s}(\boldsymbol{x} - \boldsymbol{x}_j))$$

Trilinear Interpolation weight



LIG: Inference Optimization

Classification Loss

$$\arg\min_{\boldsymbol{c}\in\mathcal{G}} \sum_{i\in\mathcal{B}} \sum_{j\in\mathcal{N}_i} \mathcal{L}_c(f(\boldsymbol{x_i}, \{\boldsymbol{c}_j | j \in \mathcal{N}\}), \mathrm{sign}(\boldsymbol{x}_i)) + \lambda||\boldsymbol{c}_j||_2$$

Latent code in each grid cell

Latent Grid Cell

Sample Points

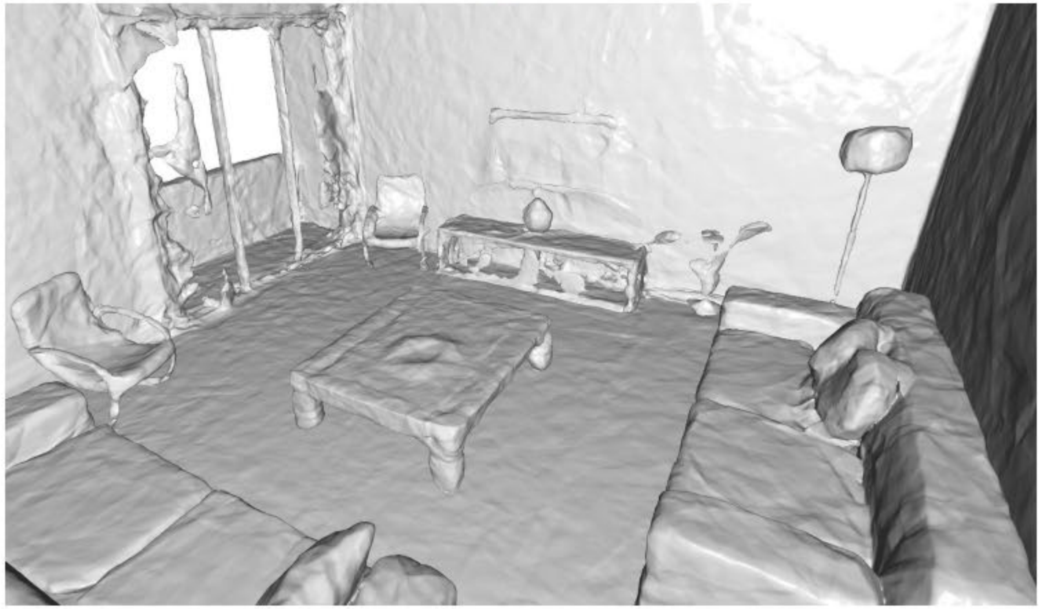
LDIF	LIG

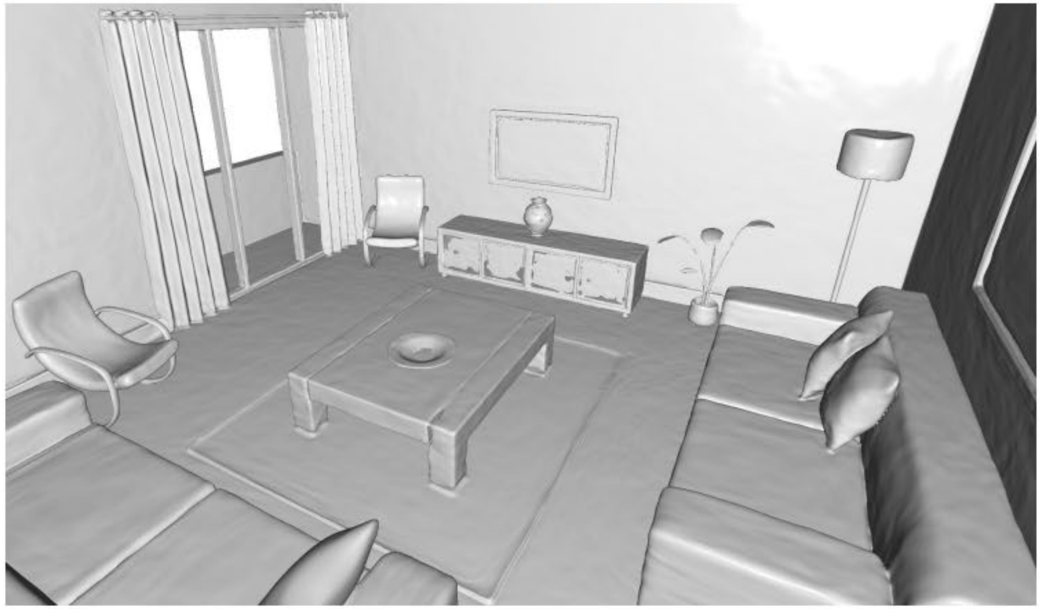
PROS

- Good Reconstruction for unseen classes
- Scalable Reconstruction for scene
- Good for sharp edges and thin structures

CONS

- Lose Part semantic
- Noise in the reconstruction





Implicit Neural Representations with Periodic Activation Functions

Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, Gordon Wetzstein