DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation

Date: 2021/1/26

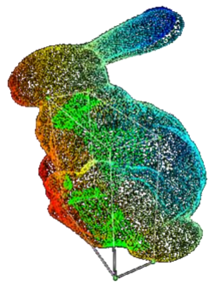
Presenter: Tianchang Shen

Instructor: Animesh Garg

UNIVERSITY OF TORONTO

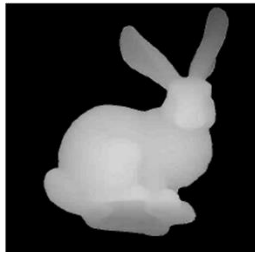
Motivation and Main Problem

2D or Partial 3D



Point Cloud

Image



Depth

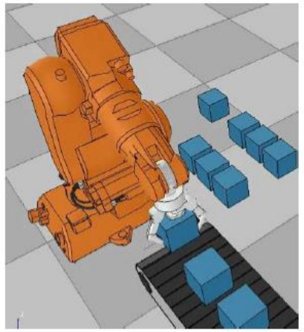
Reconstruction

Graphics

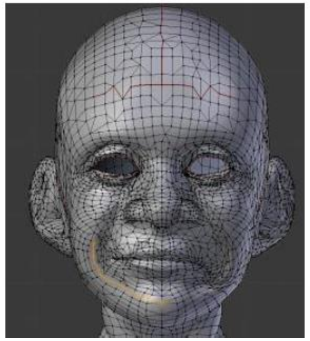
3D Model

Downstream Applications

Robotics



Simulation

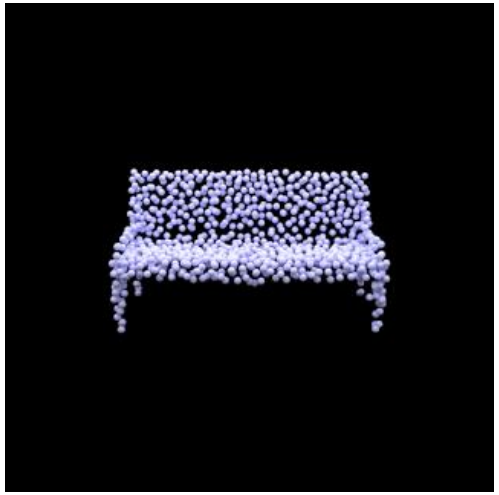


Content Creation

Motivation and Main Problem

3D Representations prior to DeepSDF

Voxel (Choy et al. 2016)



Point Cloud (Fan et al. 2017)

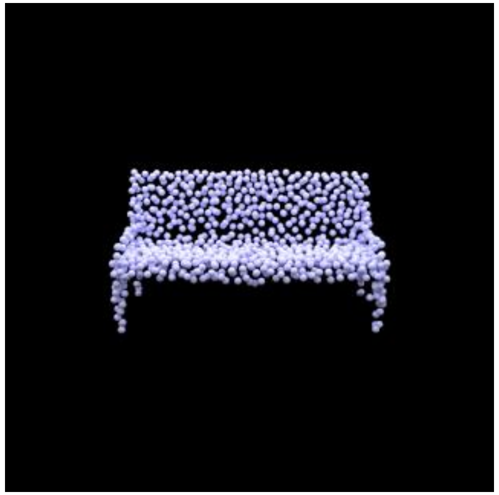
Mesh (Groueix et al. 2017)

Motivation and Main Problem

3D Representations prior to DeepSDF

Voxel (Choy et al. 2016)

Cubically growing compute and memory requirements



Point Cloud (Fan et al. 2017)

Do not describe surface

Mesh (Groueix et al. 2017)

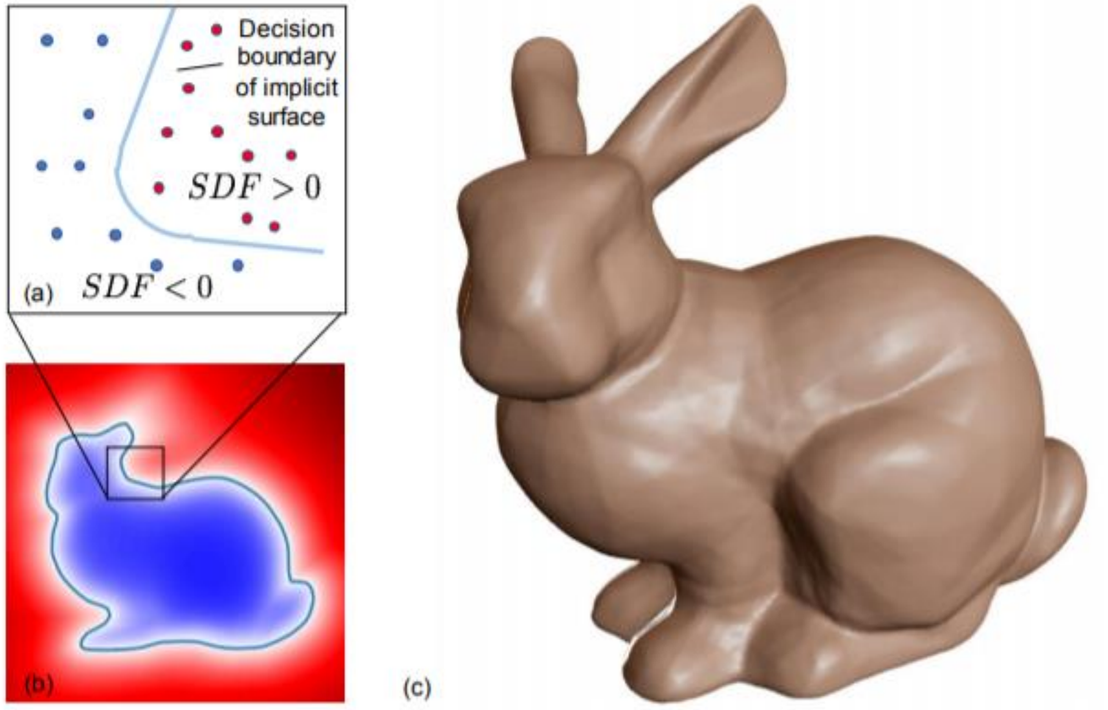
Limited to the typology of the template

Motivation and Main Problem

Key insight

- Directly regress SDF
- The surface is implicitly represented by the zero-level set

$$f_\theta(x, y, z) \approx SDF(x, y, z)$$



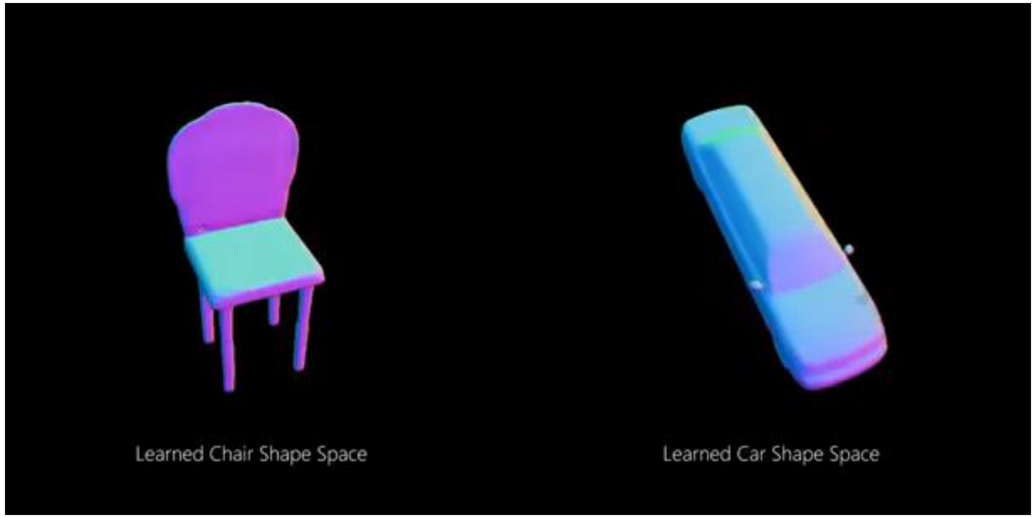
Contribution

Prior work of 3D representations lacks the ability of representing fine surface details.

In this paper, the authors proposed

- a new 3D representation that is efficient, expressive and continuous.
- a learning method for 3D shapes based on a probabilistic auto-decoder

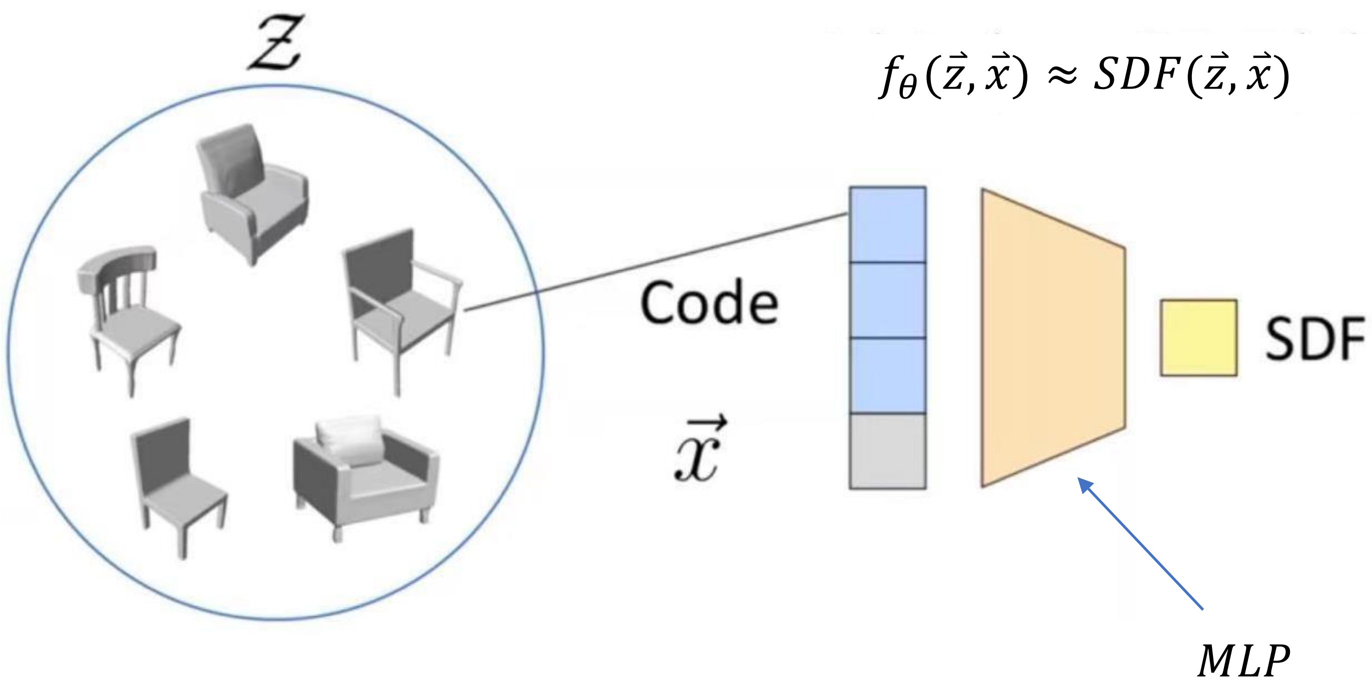
Further, they demonstrate the application of their formulation by obtaining SOTA results on shape reconstruction and completion



Problem Setting

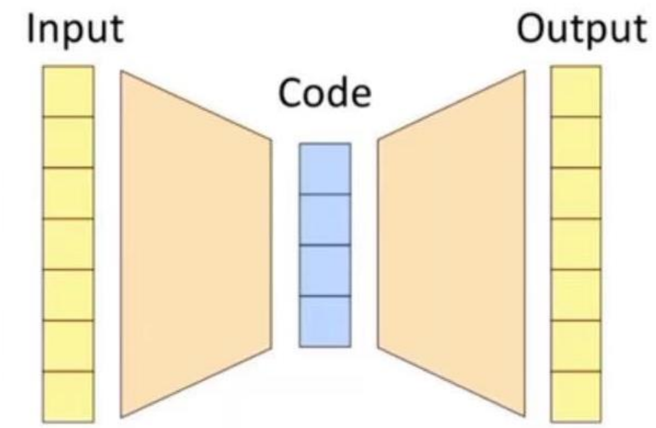
Learning Shape Conditioned Reconstruction with a Continuous Implicit Surface

$$f_\theta(\vec{z}, \vec{x}) \approx SDF(\vec{z}, \vec{x})$$



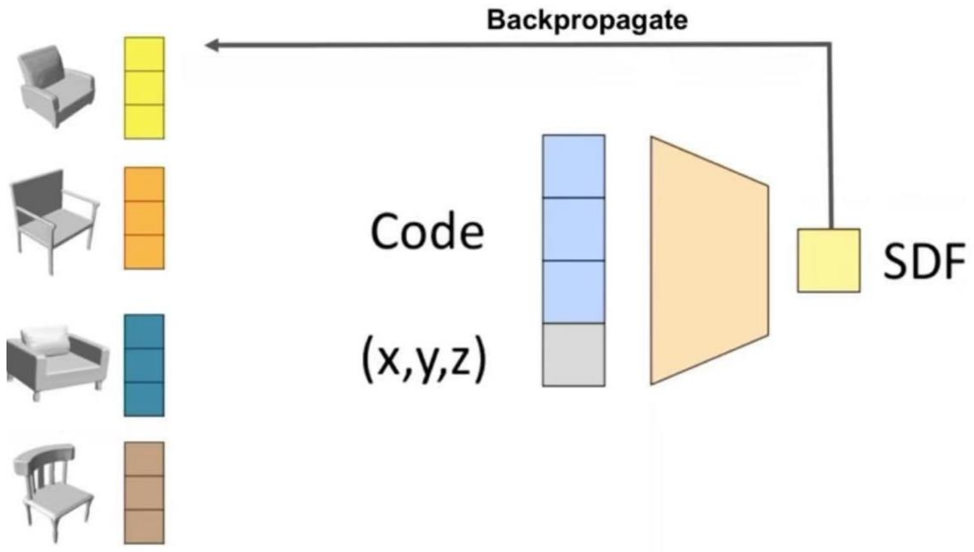
Learning

Coding Multiple Shapes with Auto Decoder



Auto Encoder

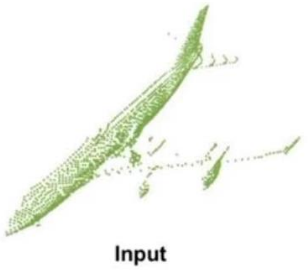
It is not straightforward to design and encoder on SDF

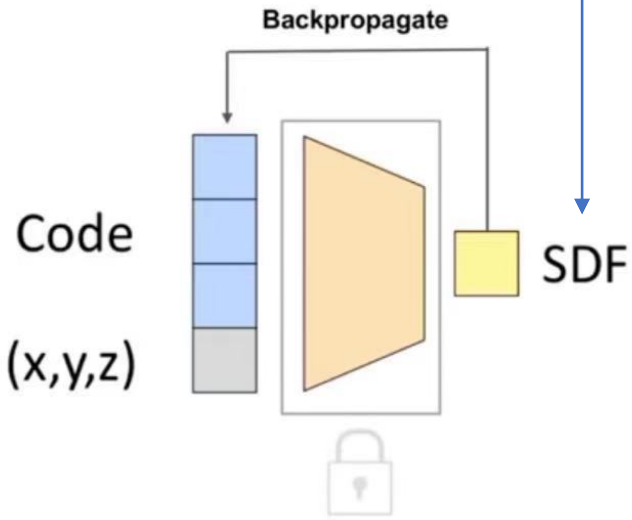


Auto Decoder

Directly optimize latent code

Inference



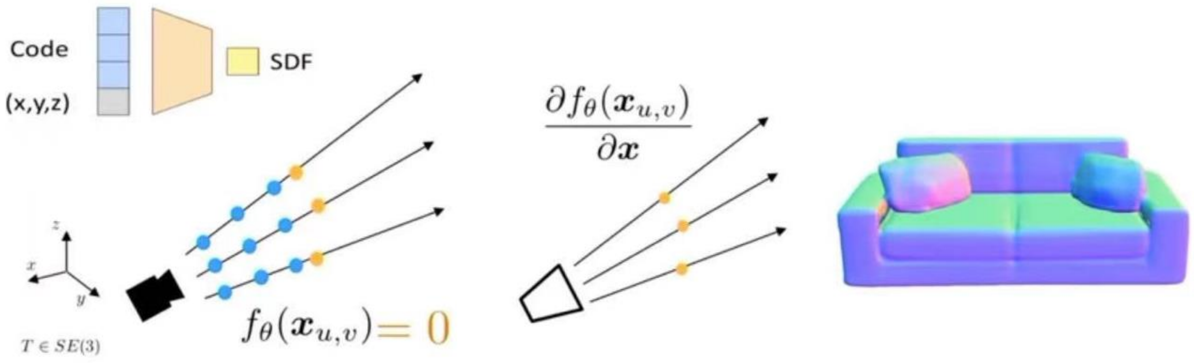


Compute SDF on visible parts
Optimize latent code while fixing decoder

Inference with optimized latent code to complete shape

Surface Extraction

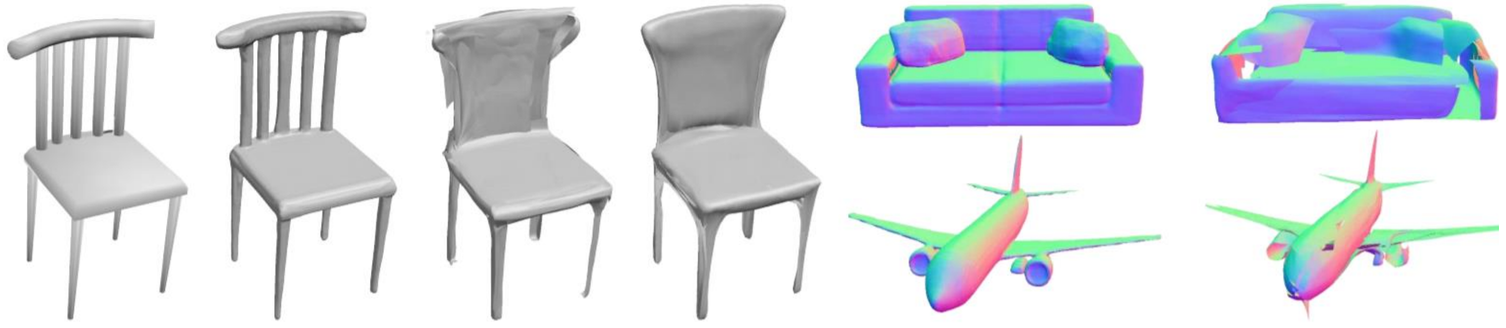
1. Ray casting



2. Marching Cubes – faster, however imposes quantization error due to fixed grid size

Results and Discussion

- Representing unknown shapes

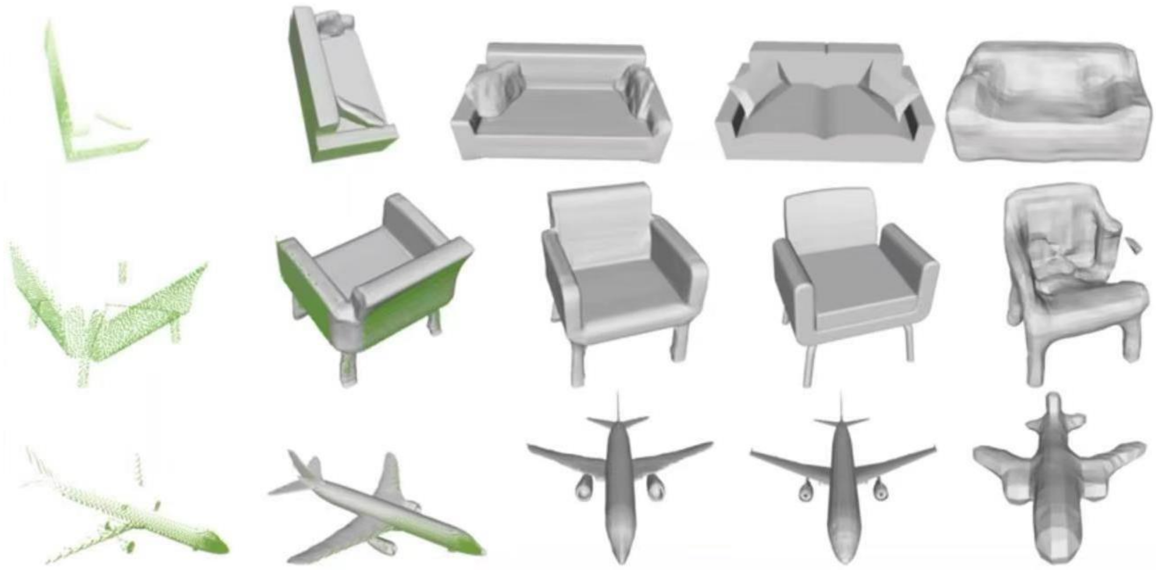


(a) Ground-truth (b) Our Result (c) [19]-25 patch (d) [19]-sphere (e) Our Result (f) [19]-25 patch

CD, mean	chair	plane	table	lamp	sofa
AtlasNet-Sph.	0.752	0.188	0.725	2.381	0.445
AtlasNet-25	0.368	0.216	**0.328**	1.182	0.411
DeepSDF	**0.204**	**0.143**	0.553	**0.832**	**0.132**
CD, median					
AtlasNet-Sph.	0.511	0.079	0.389	2.180	0.330
AtlasNet-25	0.276	0.065	0.195	0.993	0.311
DeepSDF	**0.072**	**0.036**	**0.068**	**0.219**	**0.088**
EMD, mean					
AtlasNet-Sph.	0.071	0.038	0.060	0.085	0.050
AtlasNet-25	0.064	0.041	0.073	0.062	0.063
DeepSDF	**0.049**	**0.033**	**0.050**	**0.059**	**0.047**
Mesh acc., mean					
AtlasNet-Sph.	0.033	0.013	0.032	0.054	0.017
AtlasNet-25	0.018	0.013	0.014	0.042	0.017
DeepSDF	**0.009**	**0.004**	**0.012**	**0.013**	**0.004**

Results and Discussion

- Shape completion from partial range scans
  - Notice that, the same trained model can be applied to different reconstruction tasks, unlike Octnet.



(a) Input Depth (b) Completion (ours) (c) Second View (ours) (d) Ground truth (e) 3D-EPN

	lower is better				*higher is better*	
Method \Metric	CD, med.	CD, mean	EMD	Mesh acc.	Mesh comp.	Cos sim.
chair						
3D-EPN	2.25	2.83	0.084	0.059	0.209	0.752
DeepSDF	**1.28**	**2.11**	**0.071**	**0.049**	**0.500**	**0.766**
plane						
3D-EPN	1.63	2.19	0.063	0.040	0.165	0.710
DeepSDF	**0.37**	**1.16**	**0.049**	**0.032**	**0.722**	**0.823**
sofa						
3D-EPN	2.03	2.18	0.071	0.049	0.254	0.742
DeepSDF	**0.82**	**1.59**	**0.059**	**0.041**	**0.541**	**0.810**

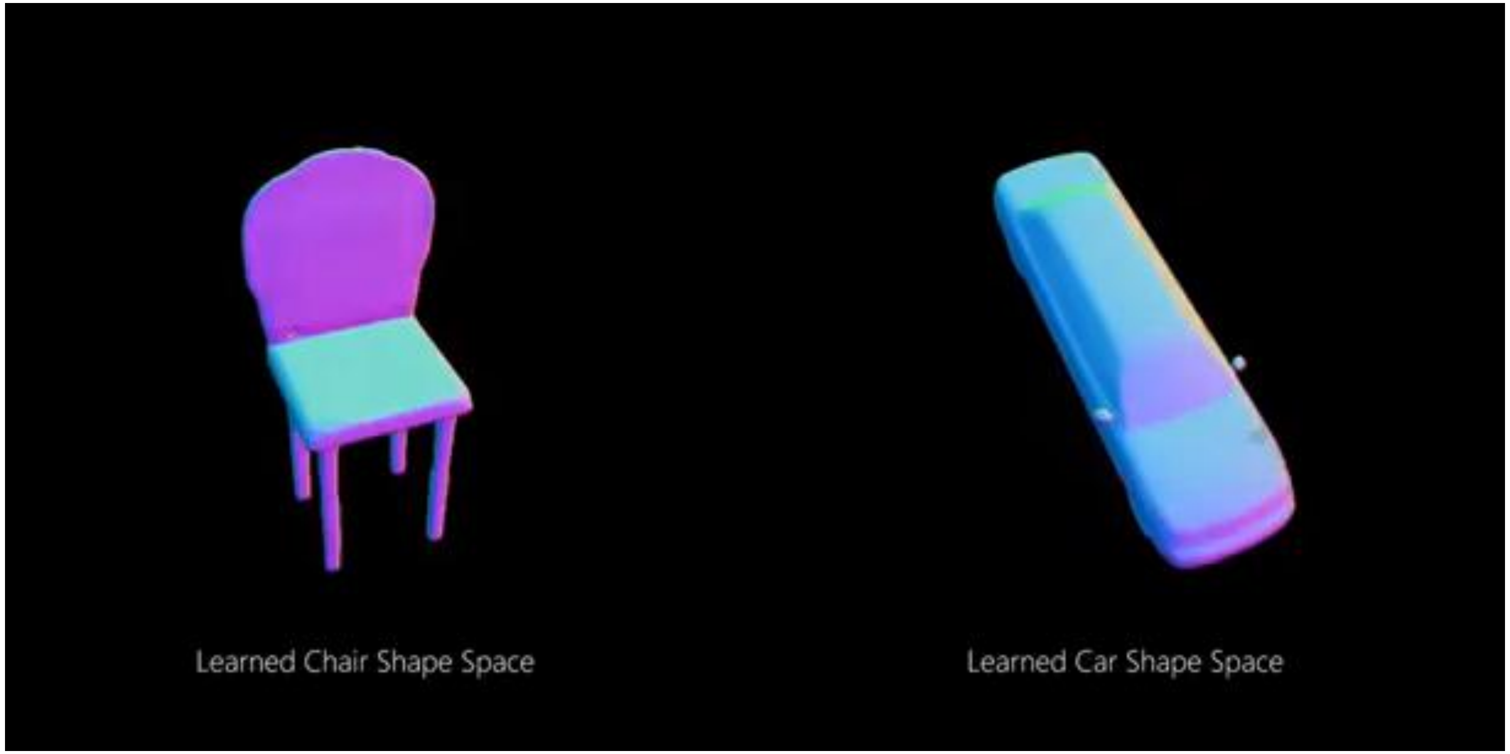
Results and Discussion

- Overview of the benchmarked methods

Method	Type	Discretization	Complex topologies	Closed surfaces	Surface normals	Model size (GB)	Inf. time (s)	Eval. tasks
3D-EPN [15]	Voxel SDF	32^3 voxels	✓	✓	✓	0.42	-	C
OGN [49]	Octree	256^3 voxels	✓	✓		0.54	0.32	K
AtlasNet -Sphere [19]	Parametric mesh	1 patch		✓		0.015	**0.01**	K, U
AtlasNet -25 [19]	Parametric mesh	25 patches	✓			0.172	0.32	K, U
DeepSDF (ours)	Continuous SDF	**none**	✓	✓	✓	**0.0074**	9.72	K, U, C

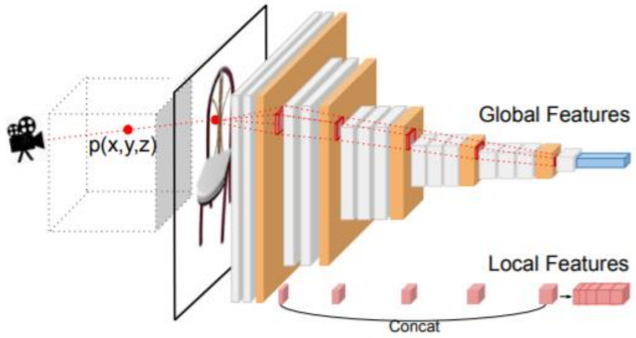
Results and Discussion

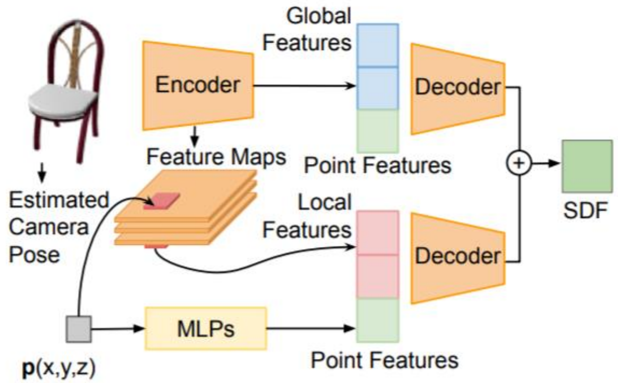
- Feature Space Interpolation



Limitation

- Inference need optimizing latent code with SDF -> not applicable to 2D observation.
  - DISN (Wang et al. 2019) addressed this problem with a novel image encoder.





  - The inference time is slow even assuming models are in their canonical pose.

Contribution

Prior work of 3D representations lacks the ability of representing fine surface details.

In this paper, the authors proposed

- a new 3D representation that is efficient, expressive and continuous.
- a learning method for 3D shapes based on a probabilistic auto-decoder

Further, they demonstrate the application of their formulation by obtaining SOTA results on shape reconstruction and completion