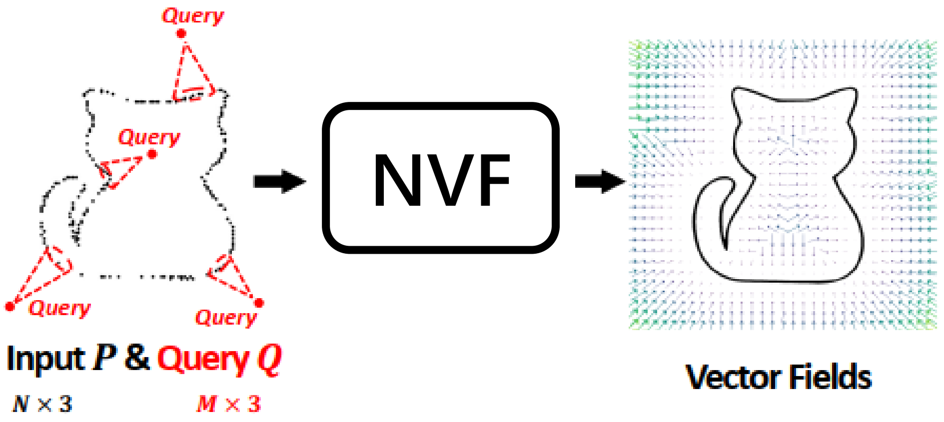
Neural Vector Fields: Implicit Representation by Explicit Learning

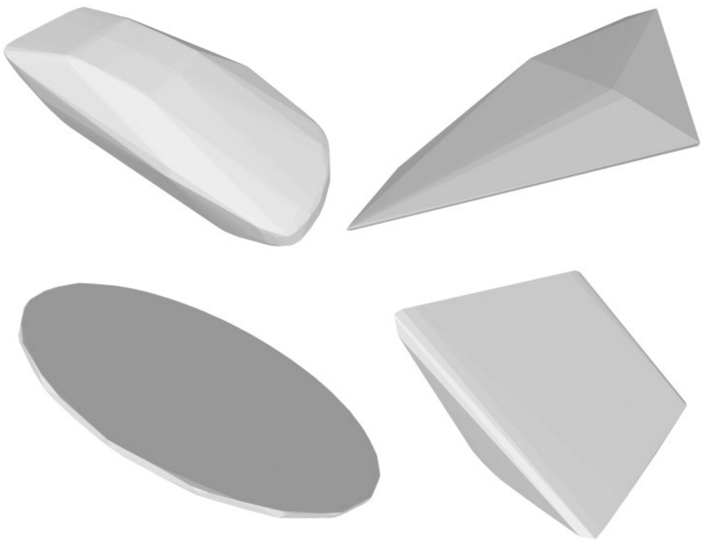
Xianghui Yang[1], Guosheng Lin[2], Zhenghao Chen[1], Luping Zhou[1]

[1]The University of Sydney,
[2]Nanyang Technological University.

Overview



Implicit representation

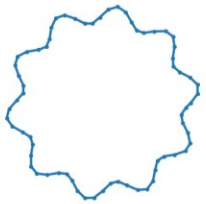


Explicit deformation

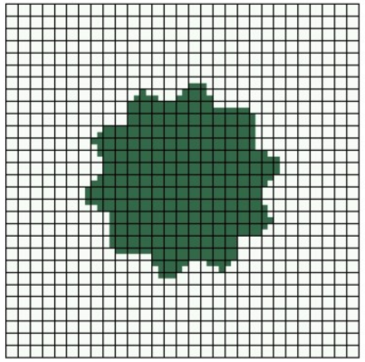
- We propose a novel differentiation-free 3D representation Neural Vector Fields.
- We propose a learned shape codebook to provide cross-object priors.
- We evaluate the effectiveness of NVF on extensive experiments, i.e., category-specific, category-agnostic, category-unseen, and cross-domain reconstructions.

Comparison

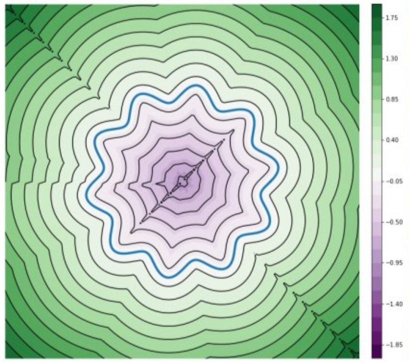
(a) Point Cloud



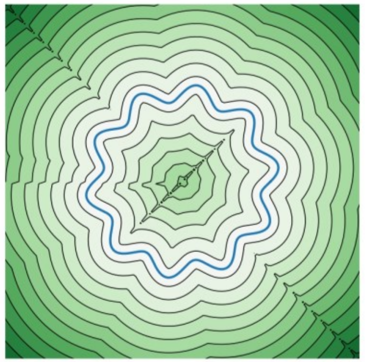
(b) Mesh



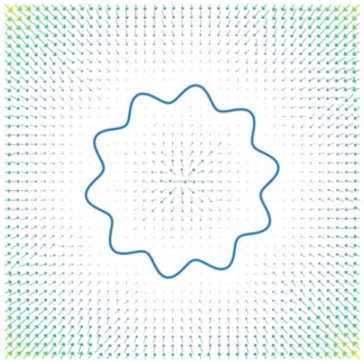
(c) Voxel/Occupancy



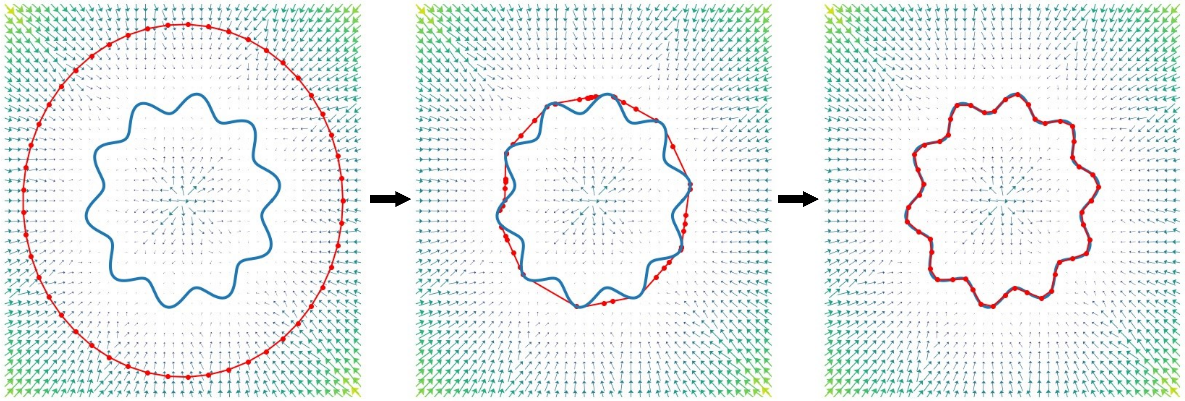
(d) Signed Distance



(e) Unsigned Distance



(f) Vector Fields



(g) Mesh Deformation via Vector Fields

$$\mathrm{NVF}(\boldsymbol{q}) = \boldsymbol{\Delta q} = \hat{\boldsymbol{q}} - \boldsymbol{q}, \forall \boldsymbol{q}$$

$\boldsymbol{q} \in \mathbb{R}^3$ is the query point

$\hat{\boldsymbol{q}} \in \mathbb{R}^3$ is the nearest point on surface

Differentiation-free

$$dist(\boldsymbol{q}) = ||\boldsymbol{\Delta q}||_2$$

$$dir(\boldsymbol{q}) = \frac{\mathbf{\Delta q}}{||\boldsymbol{\Delta q}||_2}$$

Comparison

Unsigned Distance Functions

$$\mathrm{UDF}(\boldsymbol{q}) = d, \forall \boldsymbol{q}$$

$\boldsymbol{q} \in \mathbb{R}^3$ is the query point

$\hat{\boldsymbol{q}} \in \mathbb{R}^3$ is the nearest point on surface

$$\downarrow$$

$$dist(\boldsymbol{q}) = d$$

$$dir(\boldsymbol{q}) = \boldsymbol{\nabla}\mathrm{UDF}(\boldsymbol{q})$$

Differentiation

Neural Vector Fields

$$\mathrm{NVF}(\boldsymbol{q}) = \boldsymbol{\Delta q} = \hat{\boldsymbol{q}} - \boldsymbol{q}, \forall \boldsymbol{q}$$

$\boldsymbol{q} \in \mathbb{R}^3$ is the query point

$\hat{\boldsymbol{q}} \in \mathbb{R}^3$ is the nearest point on surface

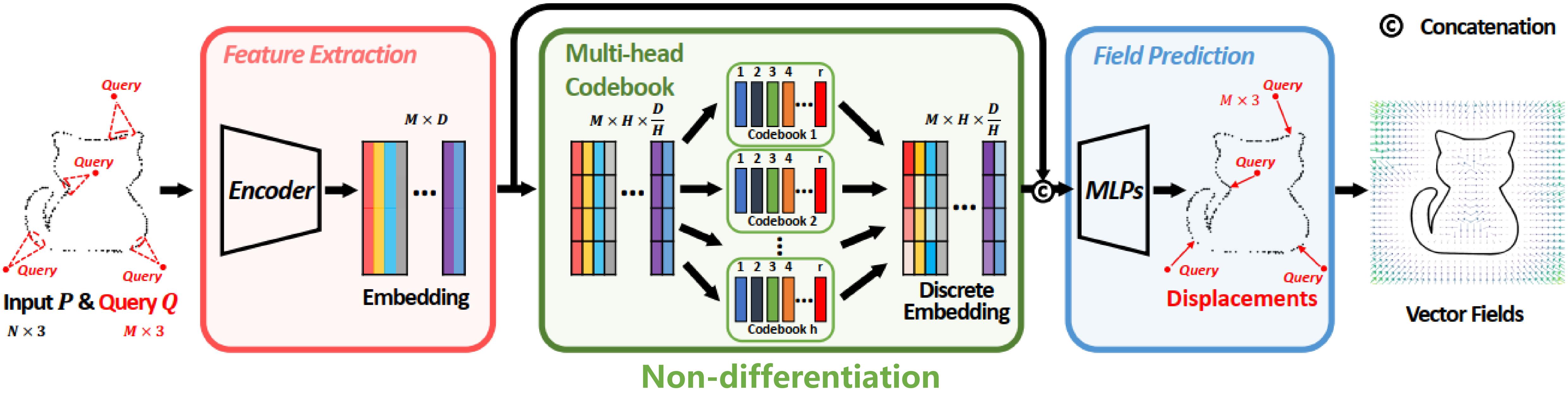
$$\downarrow$$

Differentiation-free

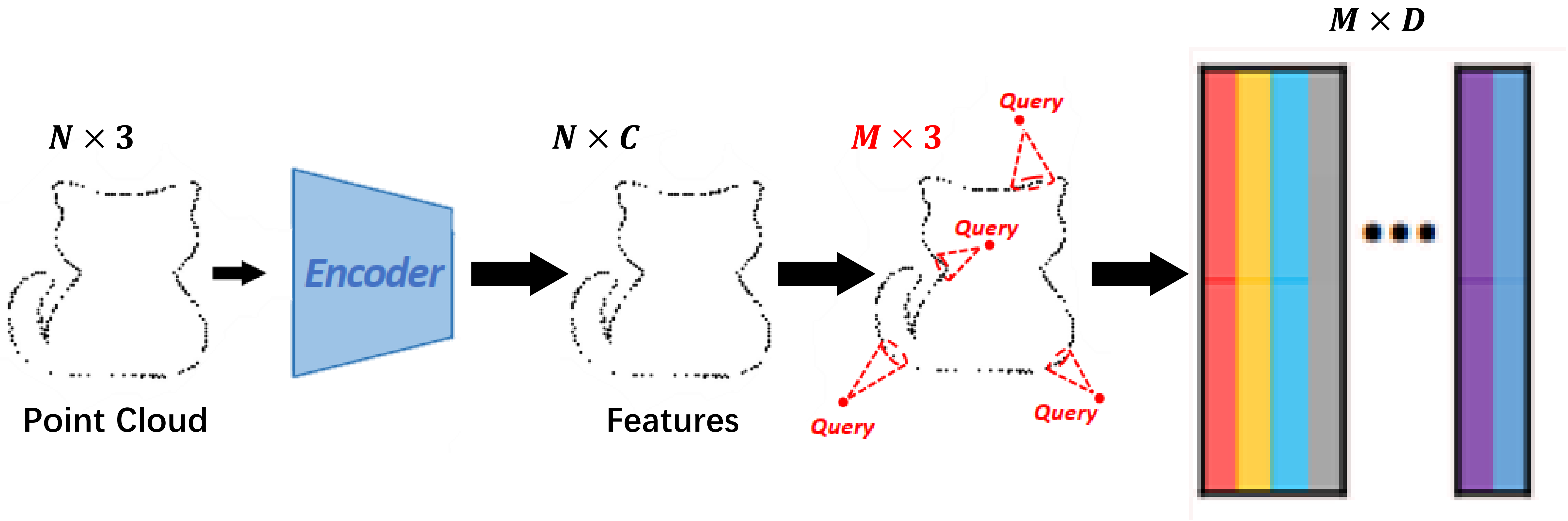
$$dist(\boldsymbol{q}) = ||\boldsymbol{\Delta q}||_{\mathbf{2}}$$

$$dir(\boldsymbol{q}) = \frac{\mathbf{\Delta q}}{||\boldsymbol{\Delta q}||_{\mathbf{2}}}$$

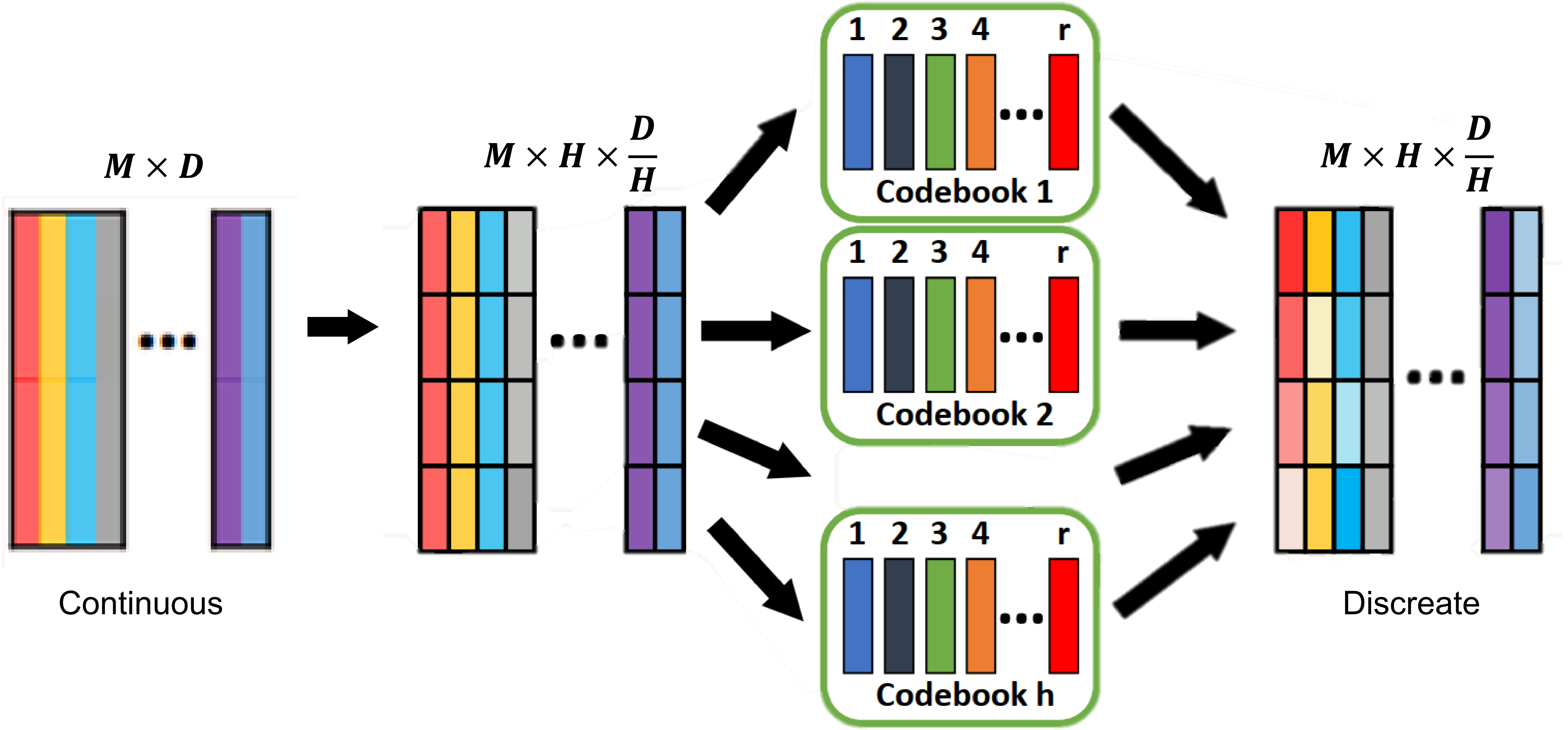
Architecture



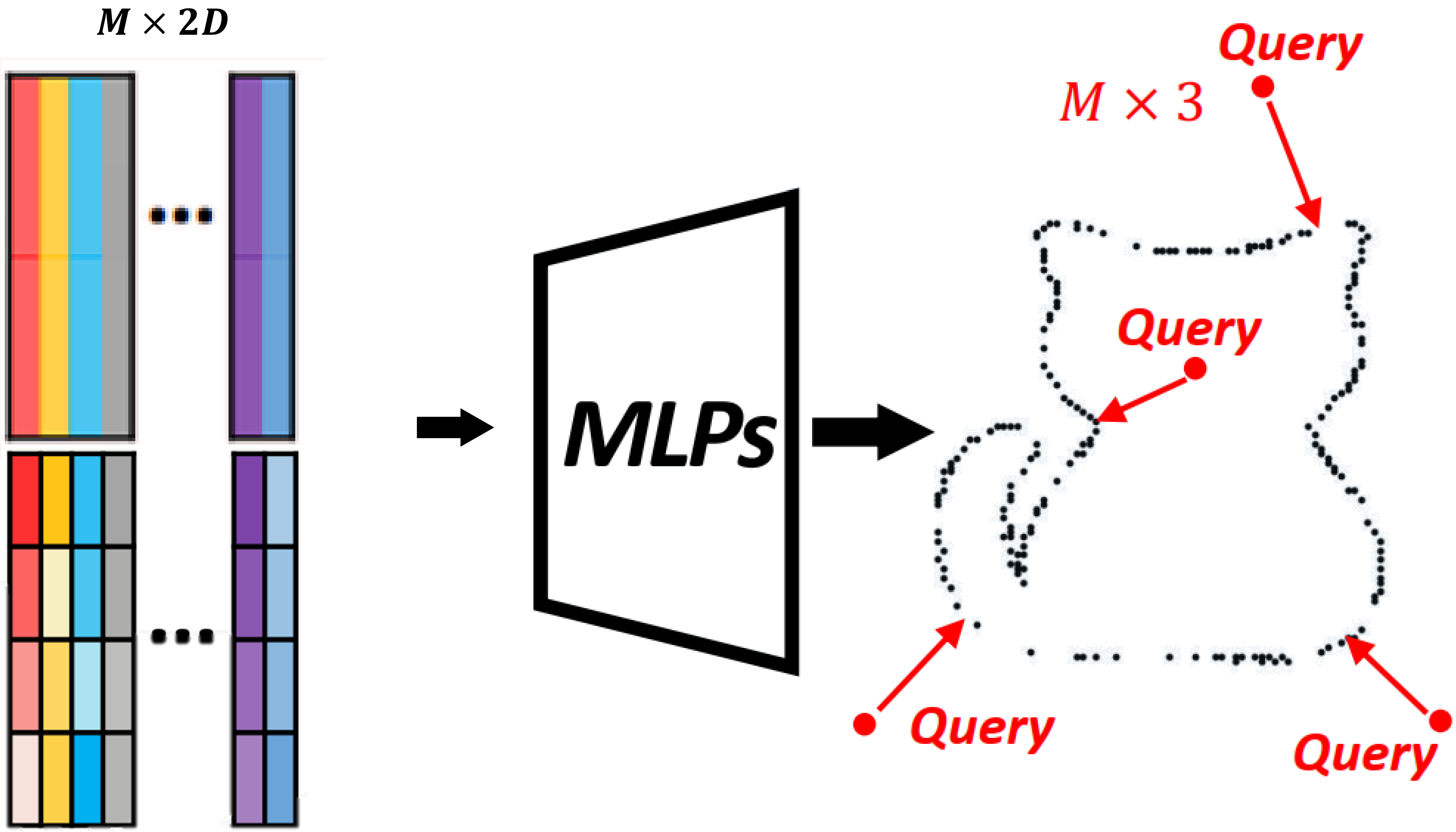
Feature Extraction



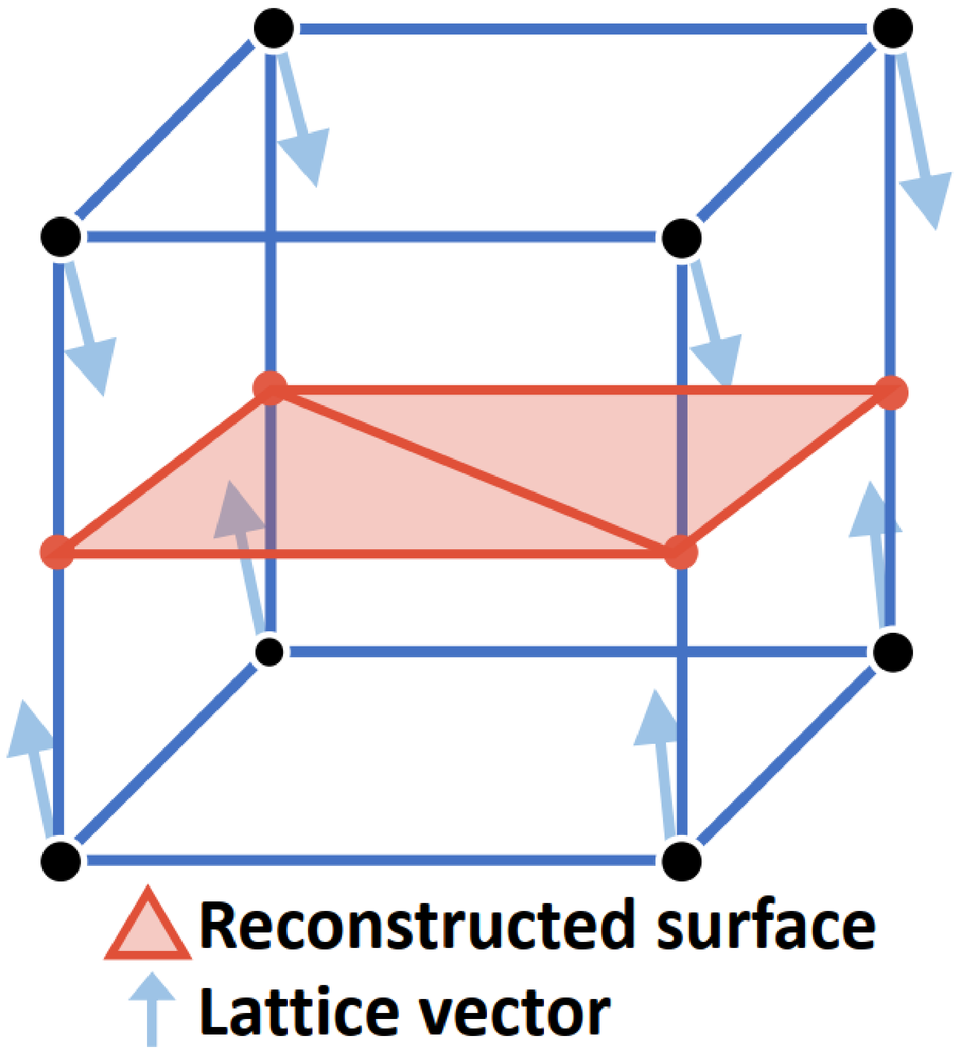
Multi-head Codebook



Field Prediction



Field Prediction



Category-specific Reconstruction

Methods	CD↓	EMD↓	$F1_{1\times10^{-5}}$	$F1_{2\times10^{-5}}$
Input	0.363	0.707	23.735	41.588
NDF [15]	0.197	1.248	64.116	84.902
GIFS [75]	0.146	0.970	54.867	79.722
Ours	**0.114**	**0.945**	**64.261**	**85.290**

Table 1. Quantitative evaluation on ShapeNet Cars. We train and evaluate our method on the raw data of the ShapeNet "Car" category. Our method achieves better performance than the state-of-the-art UDF-based methods.

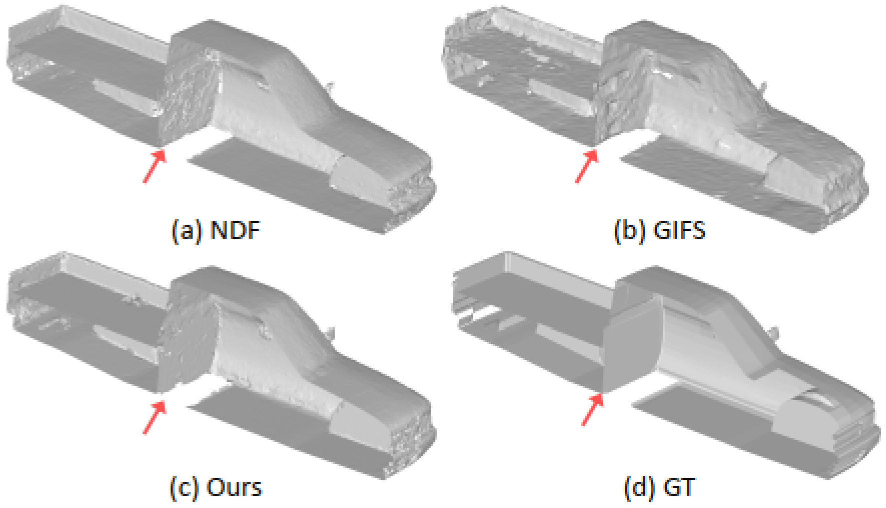
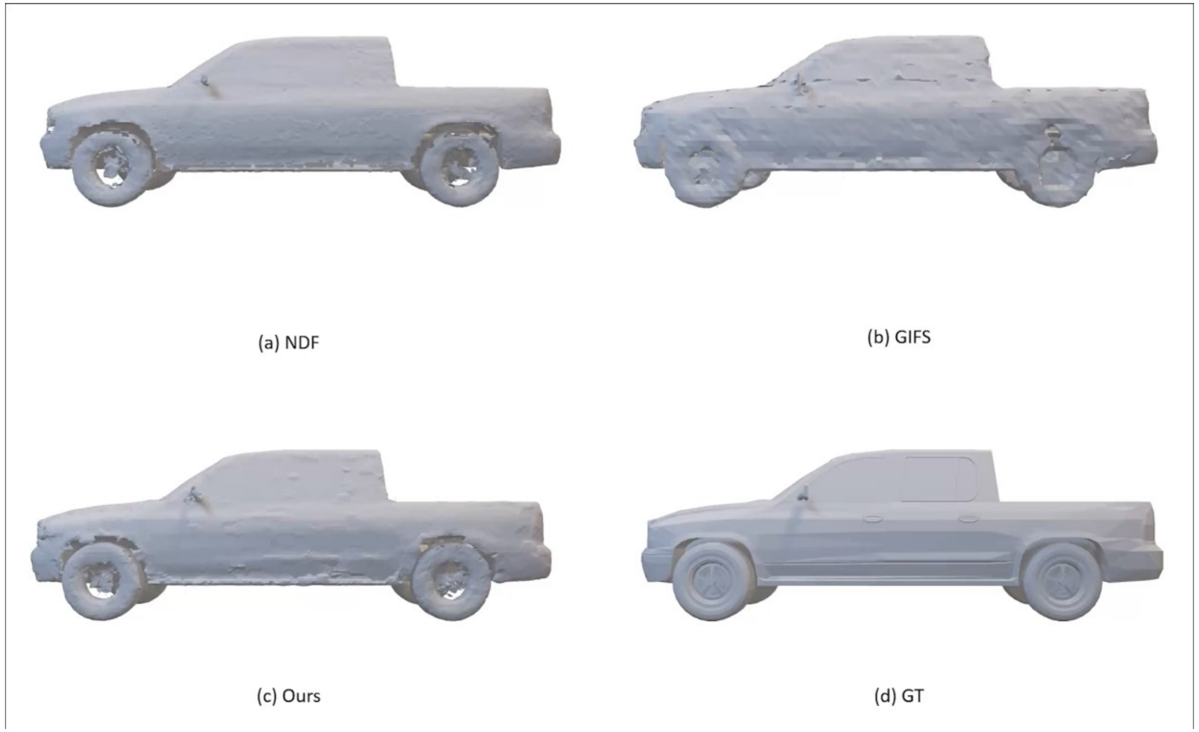


Figure 4. Qualitative visualization of Category-specific reconstruction on ShapeNet Cars. We cut parts of the shapes to visualize inner structures better.



Category-agnostic & Category-unseen Reconstruction

Methods	Base				Novel			
	CD↓	EMD↓	$F1_{2.5\times10^{-5}}$ ↑	$F1_{1\times10^{-4}}$ ↑	CD↓	EMD↓	$F1_{1\times10^{-5}}$ ↑	$F1_{2\times10^{-5}}$ ↑
Input	0.840	1.045	14.148	25.111	0.800	1.024	17.576	29.815
OccNet [44]	2.766	1.694	30.877	46.644	44.762	4.013	15.943	24.433
IF-Net [14]	0.190	1.120	65.975	85.421	0.596	1.608	61.670	81.106
NDF [15]	0.169	1.538	66.802	84.809	0.169	1.741	65.622	84.069
GIFS [75]	0.179	1.280	56.188	78.458	0.194	1.534	56.644	78.016
Ours	**0.091**	**1.079**	**78.503**	**91.408**	**0.144**	**1.145**	**80.883**	**91.836**

Table 2. Quantitative results of category-agnostic and category-unseen reconstructions on watertight shapes of ShapeNet. We train all models on the base classes, and evaluate them on the base and the novel classes, respectively.

Methods	Base				Novel			
	CD↓	EMD↓	$F1_{2.5\times10^{-5}}$ ↑	$F1_{1\times10^{-4}}$ ↑	CD↓	EMD↓	$F1_{1\times10^{-5}}$ ↑	$F1_{2\times10^{-5}}$ ↑
Input	0.317	0.867	32.875	51.105	0.289	0.843	39.902	58.092
NDF [15]	0.099	1.372	72.425	88.754	0.093	1.532	76.162	89.977
GIFS [75]	0.118	1.260	64.915	85.115	0.296	1.499	69.252	86.518
Ours	**0.085**	**1.197**	**75.372**	**90.266**	**0.078**	**1.340**	**79.723**	**91.576**

Table 3. Quantitative results of category-agnostic and category-unseen reconstructions on non-watertight shapes of ShapeNet. We train all models on the base classes and evaluate them on the base and the novel classes, respectively.

Category-agnostic & Category-unseen Reconstruction

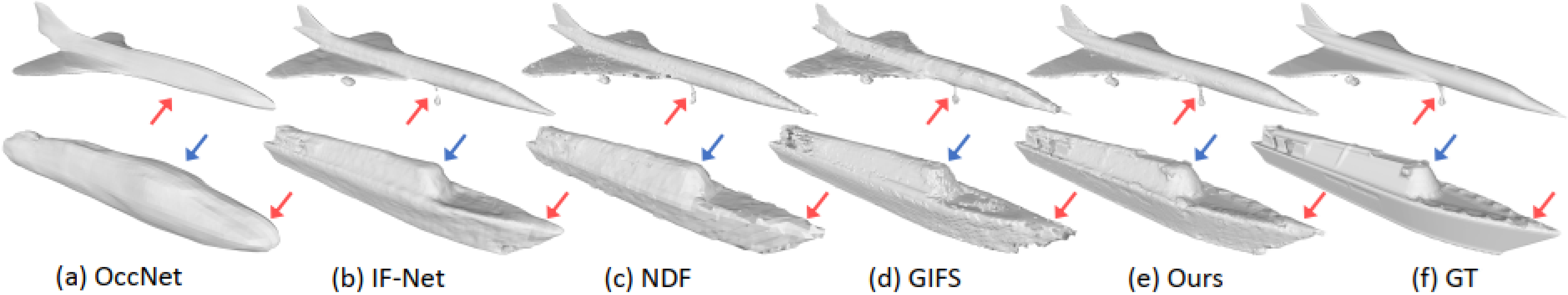


Figure 5. Visualization of category-agnostic and category-unseen reconstructions on watertight shapes from the ShapeNet dataset. The 1^{st} row, planes, is from the base classes. The 2^{rd} row, watercraft, is from the novel classes.

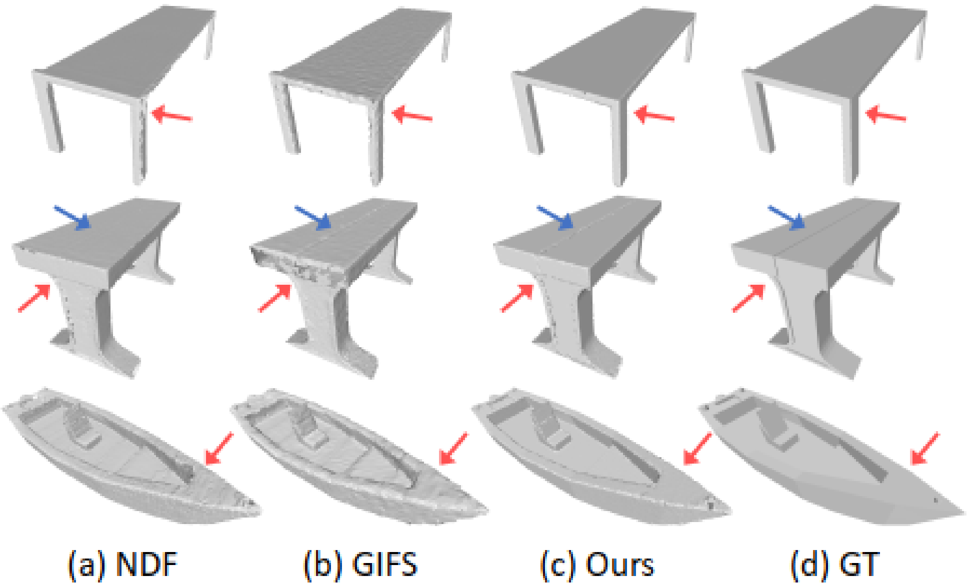
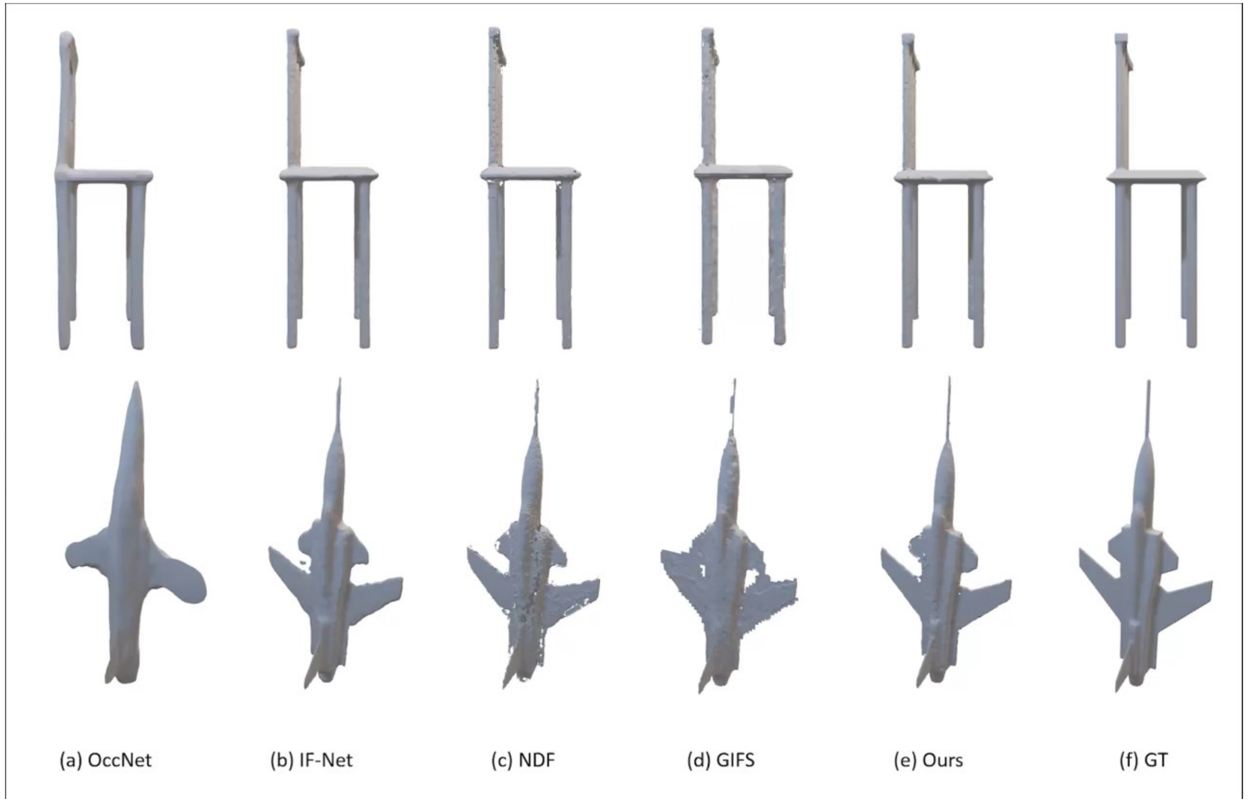
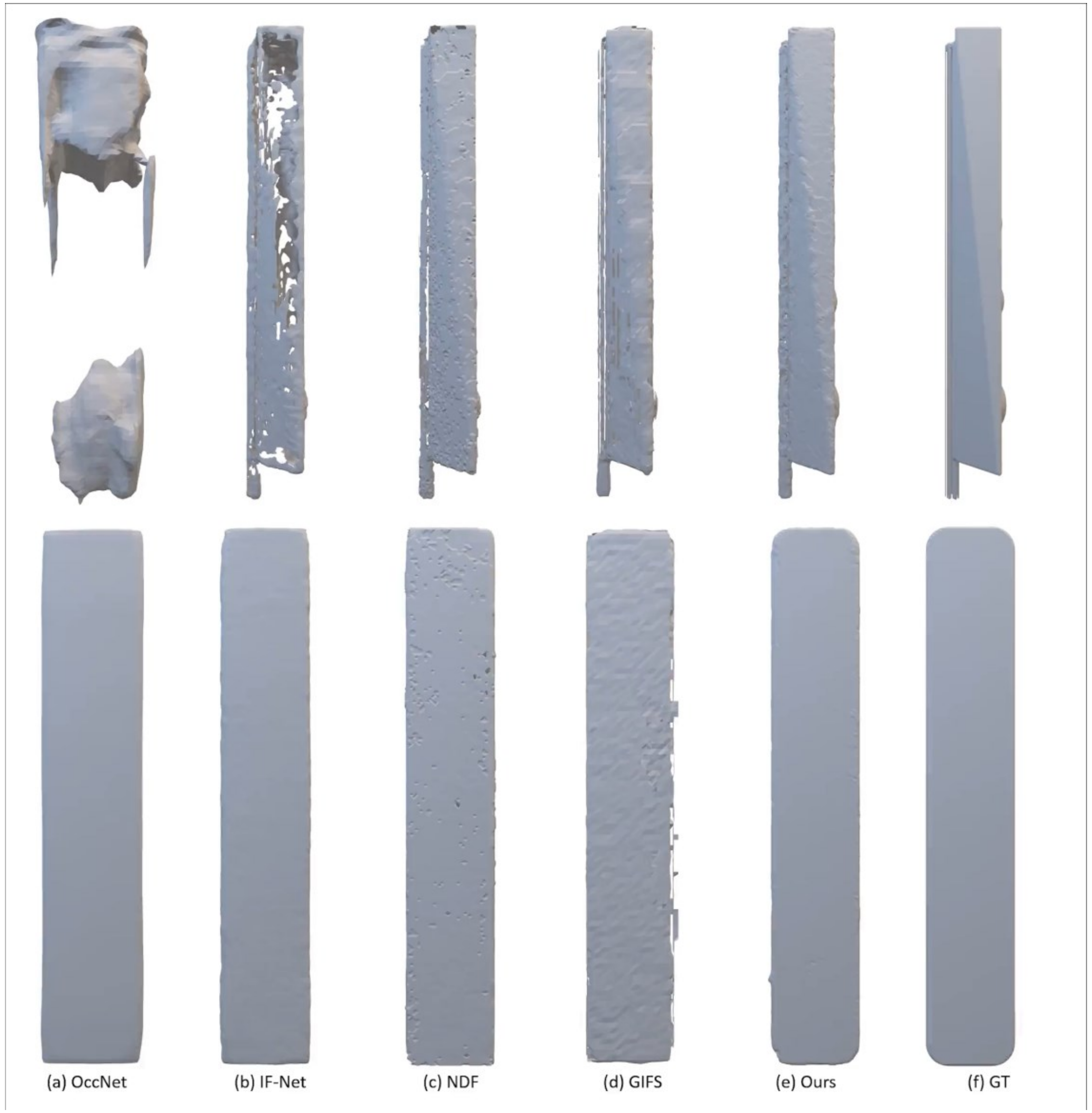


Figure 6. Visualization of category-agnostic and category-unseen reconstructions on non-watertight shapes from the ShapeNet dataset. The 1^{st} row, tables, is from the base classes. The 2^{nd} and 3^{rd} rows, benches and watercraft, are from the novel classes.

Watertight

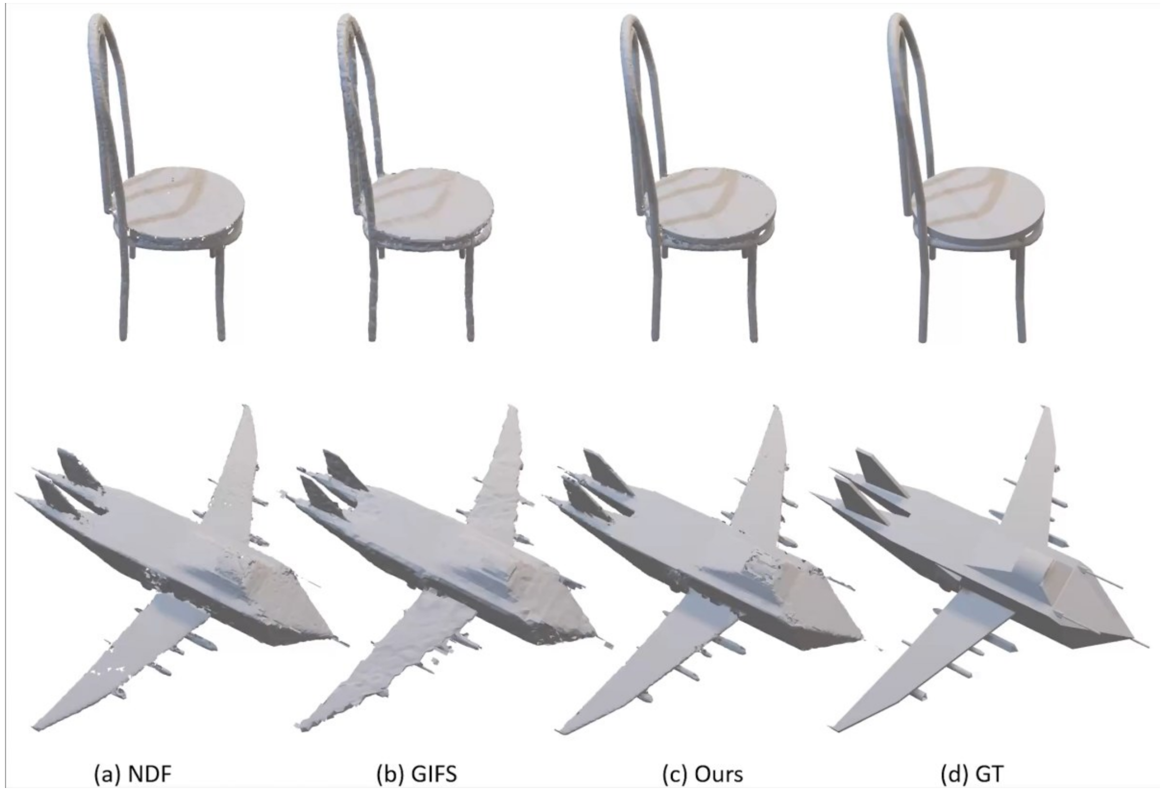


Base

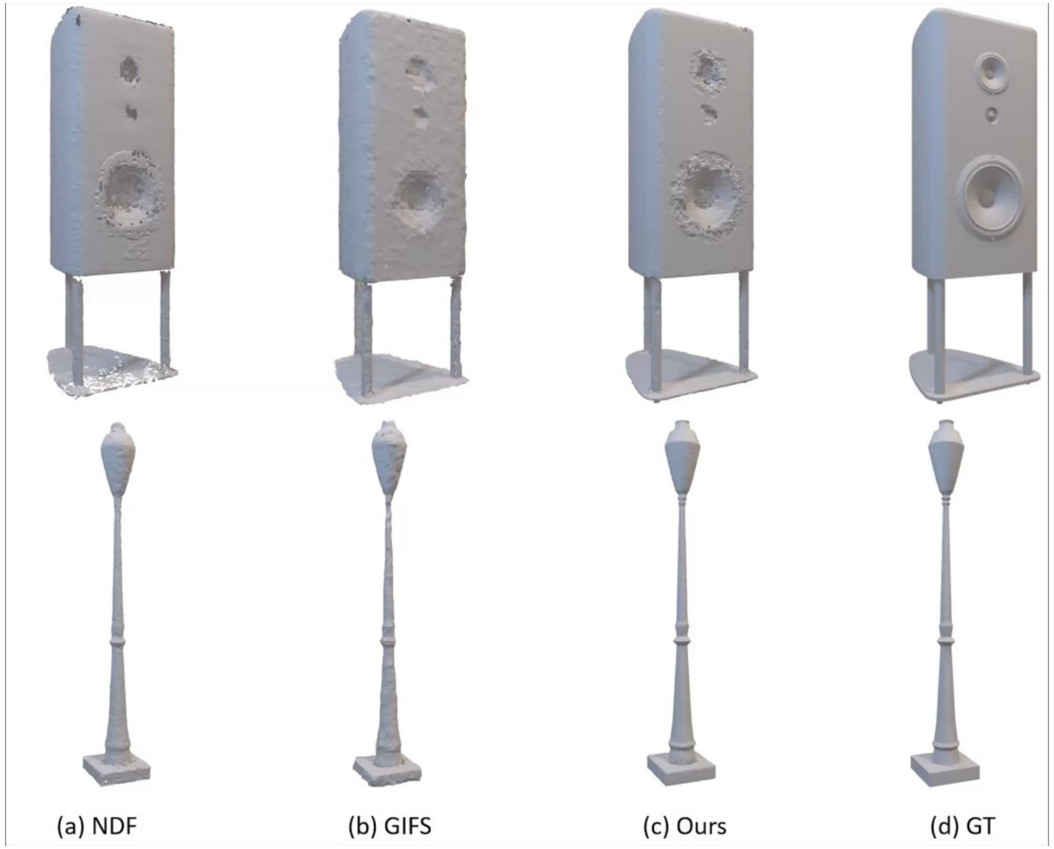


Novel

Non-watertight



Base

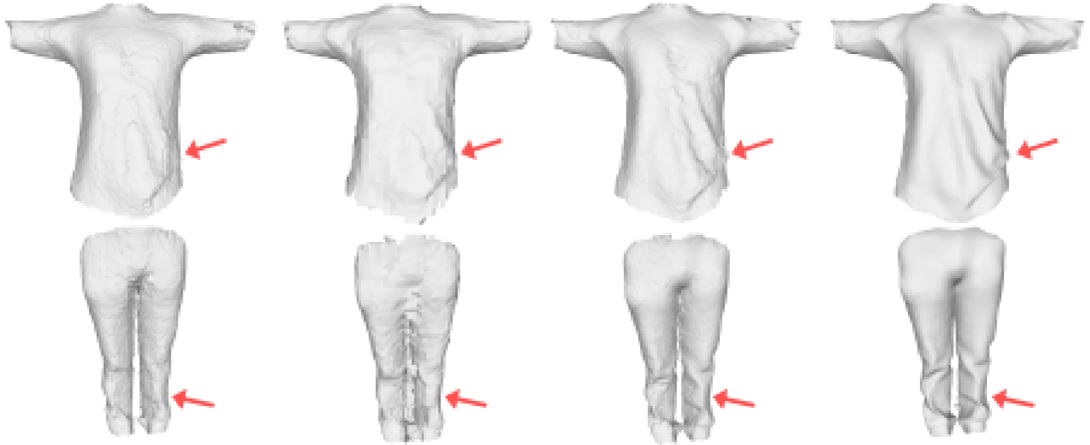


Novel

Cross-domain Reconstruction

Methods	CD↓	EMD↓	$F1_{1\times10^{-5}}$	$F1_{2\times10^{-5}}$
Input	0.124	0.157	52.189	72.969
NDF [15]	0.025	0.216	96.338	98.687
GIFS [75]	0.039	0.192	93.330	97.295
Ours	**0.014**	**0.184**	**98.499**	**99.498**

Table 4. Quantitative results of cross-domain reconstruction on MGN [6]. We train our models based on ShapeNet with the base classes and evaluate them on the raw data from MGN.



(a) NDF (b) GIFS (c) Ours (d) GT

Figure 7. Visualization of cross-domain reconstruction on the MGN dataset. All models are trained based on ShapeNet with the base classes and evaluated directly on MGN.

(a) NDF (b) GIFS (c) Ours (d) GT

Ablation Study

THE UNIVERSITY OF SYDNEY
NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

	K	Codebook	CD↓	EMD↓	$F1_{1\times10^{-5}}$	$F1_{2\times10^{-5}}$
Base	k=8	✗	0.089	1.195	74.139	89.392
	k=8	✓	0.087	**1.172**	75.104	90.057
	k=16	✗	0.121	1.212	73.642	88.962
	k=16	✓	**0.085**	1.197	**75.372**	**90.266**
Novel	k=8	✗	0.080	1.334	78.701	90.972
	k=8	✓	0.083	1.354	79.522	91.434
	k=16	✗	0.081	**1.329**	78.800	91.084
	k=16	✓	**0.078**	1.340	**79.723**	**91.576**

Table 5. Effect of feature number K and multi-head codebook. The multi-head codebook improves the performance and achieves the best for $K = 16$.

Methods	Backbone	Codebook	Runtime	Memory
NDF [15]	3D Conv	✗	0.75s	13.44G
Ours	3D Conv	✗	0.27s	9.21G
	3D Conv	✓	0.34s	9.21G
	PointTransformer	✗	0.29s	9.28G
	PointTransformer	✓	0.35s	9.28G

Table 7. Inference analysis. The runtime and memory are time cost and peak memory during the inference of 200k queries. NVF is more efficient on inference runtime and memory.

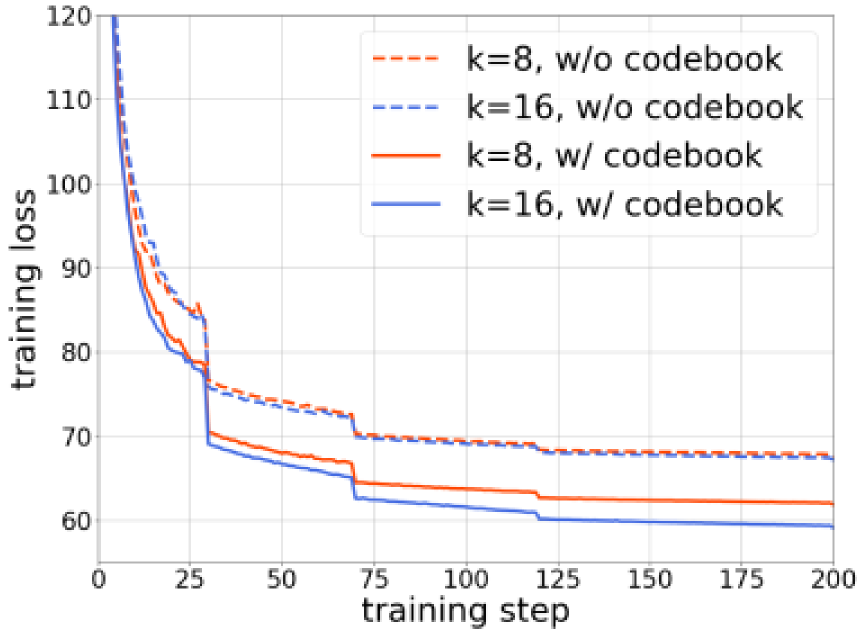


Figure 8. Training curves of models w/ and w/o codebook. The models w/ codebooks converge faster than those w/o codebooks.

Thanks

GitHub: https://github.com/Wi-sc/NVF.git

Paper: https://arxiv.org/abs/2303.04341

Email: xianghui.yang@sydney.edu.au