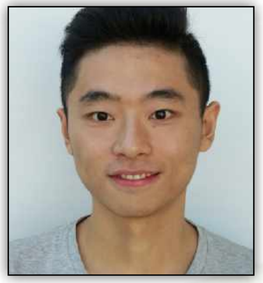
CVPR SEATTLE WASHINGTON JUNE 16-18 2020

CvxNet

Learnable Convex Decomposition


Boyang Deng


Kyle Genova


Sofien Bouaziz


Soroosh Yazdani


Geoffrey Hinton


Andrea Tagliasacchi

Google Research

Family of 3D representations

Tatarchenko et al. 2017

{voxels}

Tulsiani et al. 2017

{boxes}

Groueix et al. 2018

{patches}

Genova et al. 2019

{gaussians}

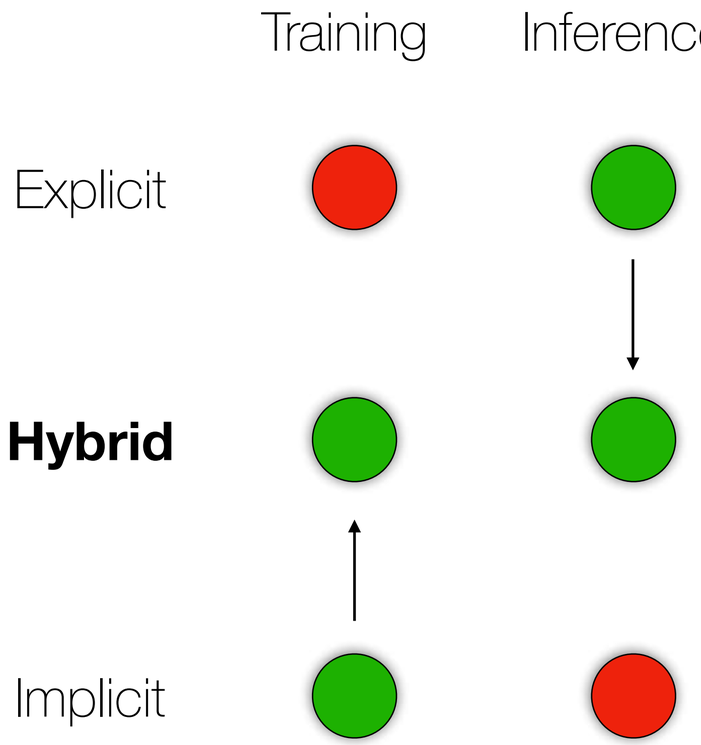
Mescheder et al. 2019

implicits

Deng et al. 2020

{convexes}

What is the «best» 3D representation?



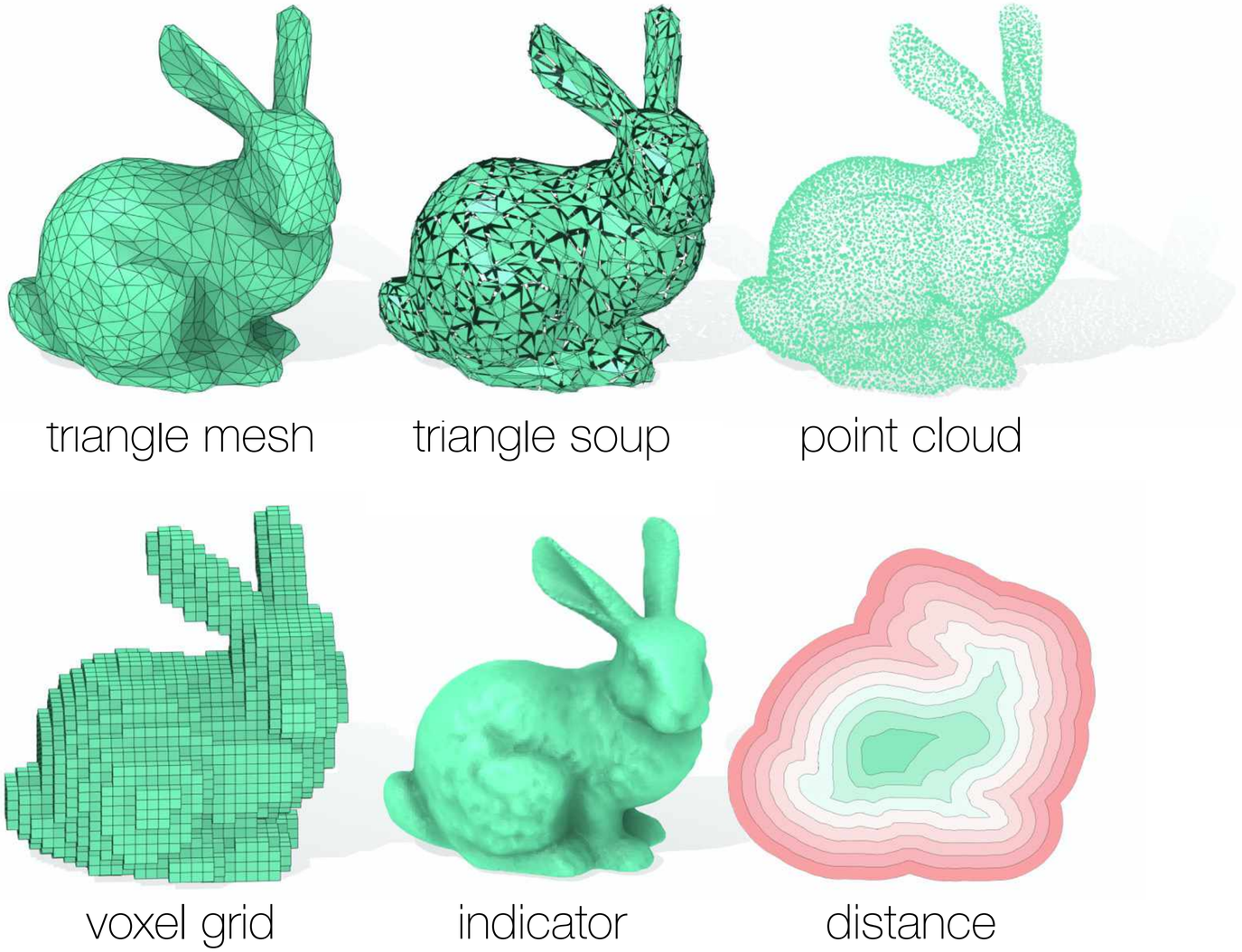
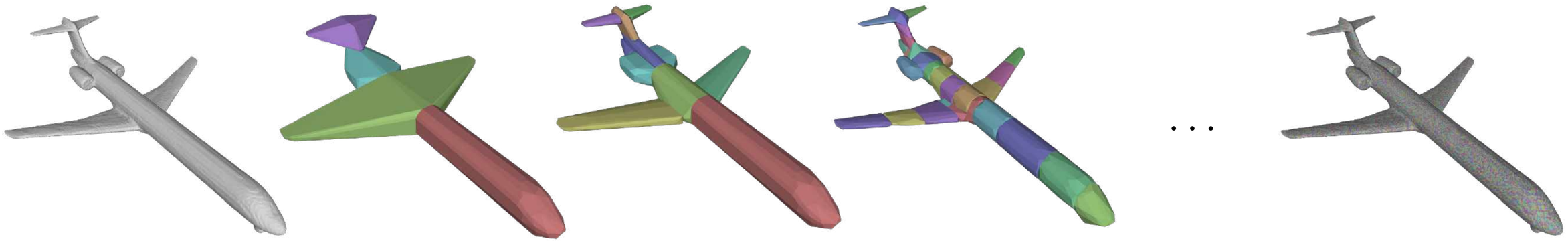


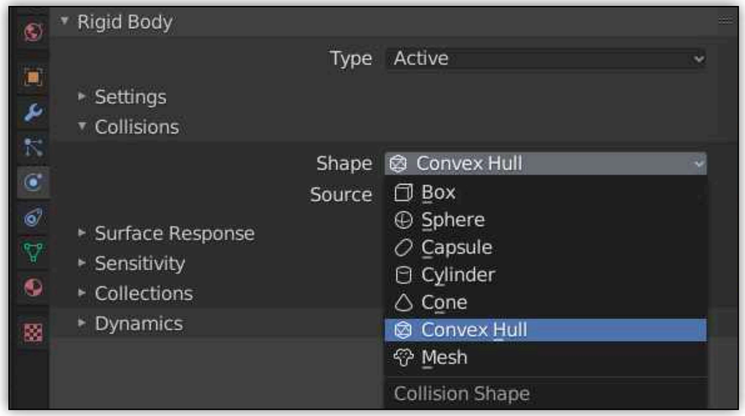
Image courtesy of Barill et al. «Fast Winding Numbers for Soups and Clouds» SIGGRAPH 2018

Convexes: why are they relevant?



input — K=4 — K=8 — K=32 — universal approximators — K=∞ →

rigid body physics →



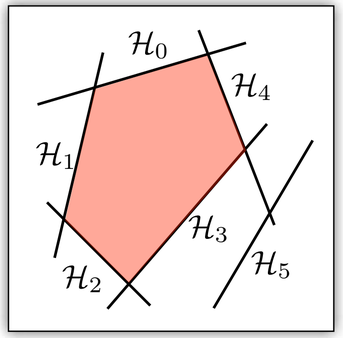
Convex Hulls (standard)

Meshes (the «wild west»)

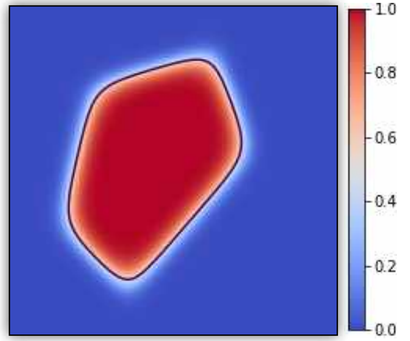
decomposition computed by «v-hacd» – https://github.com/kmammou/v-hacd

Universal approximator of convex domains

$$\cdots \quad \mathcal{E}_\omega \quad \underbrace{\begin{bmatrix} n_0^x & n_0^y & d_0 \\ n_1^x & n_1^y & d_1 \\ \cdots & \cdots & \cdots \\ n_5^x & n_5^y & d_5 \end{bmatrix}}_{\text{input: half-spaces tensor}} \equiv \begin{bmatrix} \mathcal{H}_0 \\ \mathcal{H}_1 \\ \cdots \\ \mathcal{H}_5 \end{bmatrix} \equiv$$

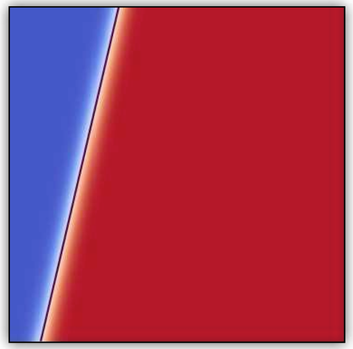


$$\mathcal{C}(\mathbf{x}|\boldsymbol{\beta}) \rightarrow$$

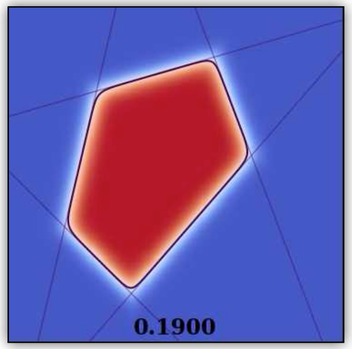


output: function

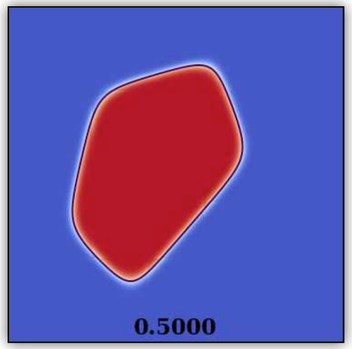
$$\underbrace{\text{Sigmoid}(-\sigma\text{LogSumExp}\{\delta(\mathbf{n_h}\cdot\mathbf{x}+d_h)\})}_{\mathcal{C}(\mathbf{x}|\boldsymbol{\beta})}$$



permutation invariance

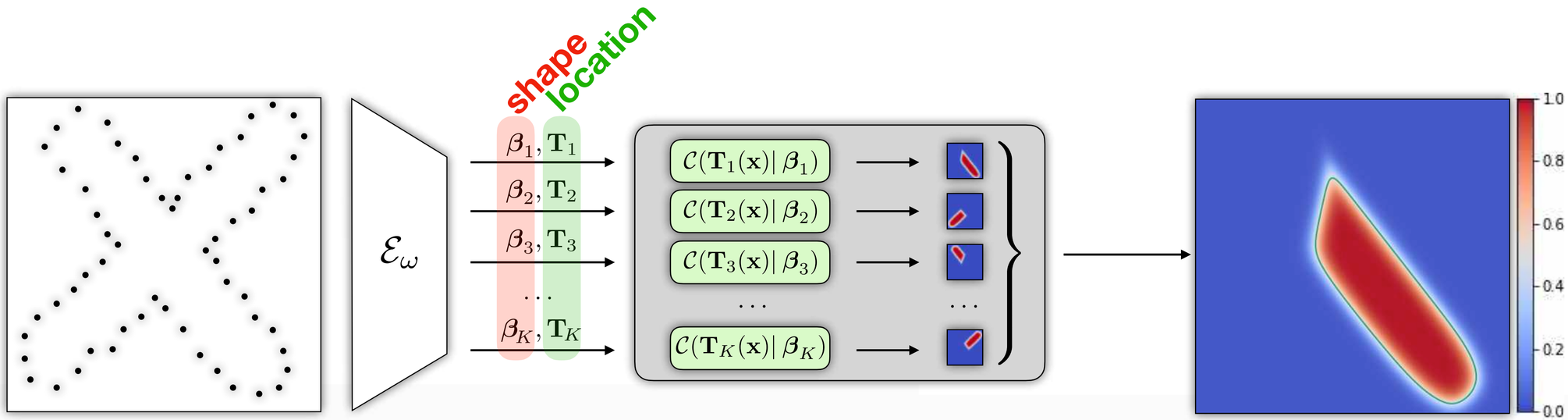


δ controls smoothness



σ controls indicator dropoff

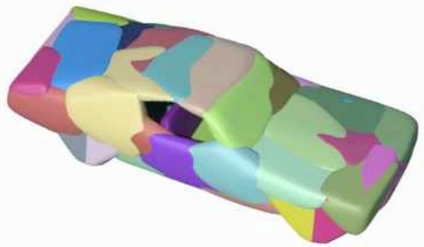
… and non-convex domains!

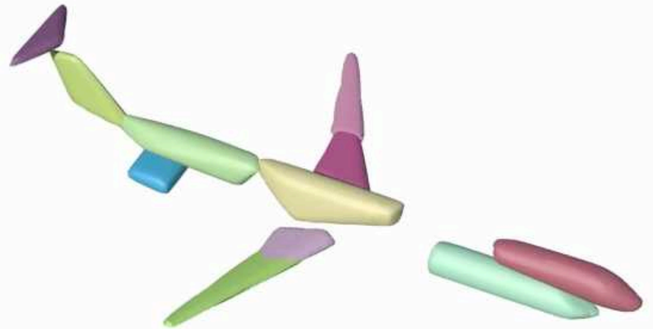


input: image, cloud, …

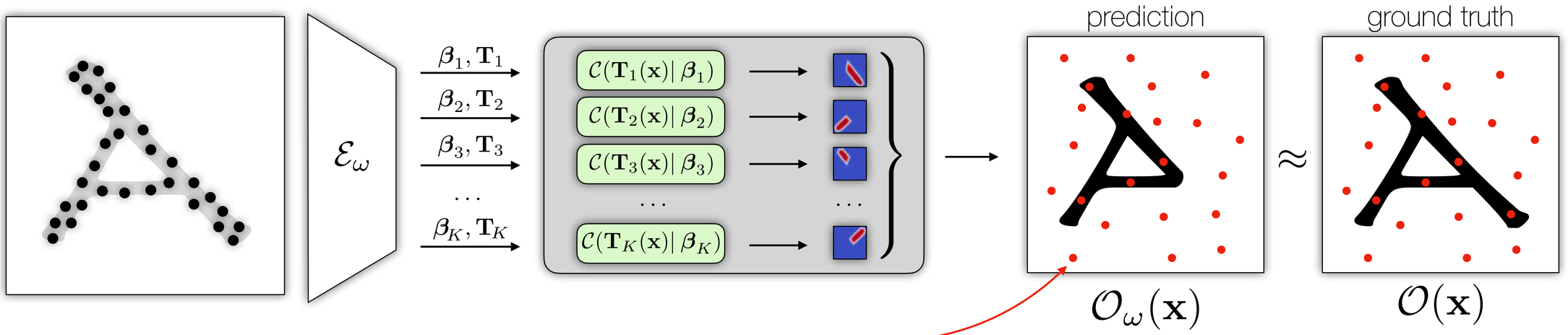
output: function

$$\mathcal{O}_\omega(\mathbf{x}) = \max_k \{\mathcal{C}(\mathbf{T}_k(\mathbf{x})|\boldsymbol{\beta}_k)\}$$





Implicit functions @ training time



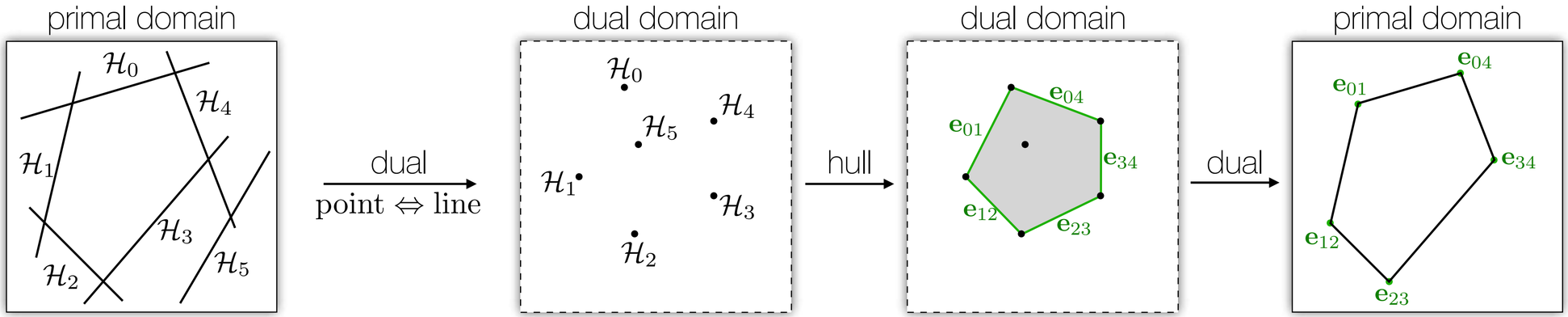
$$\mathcal{L}_{\text{approx}}(\omega) = \mathbb{E}_{\mathbf{x}\sim\mathbb{R}^2} \left\| \mathcal{O}_\omega(\mathbf{x}) - \mathcal{O}(\mathbf{x}) \right\|^2$$

$\mathcal{L}_{\text{decomp}}(\omega) =$ convexes should not overlap

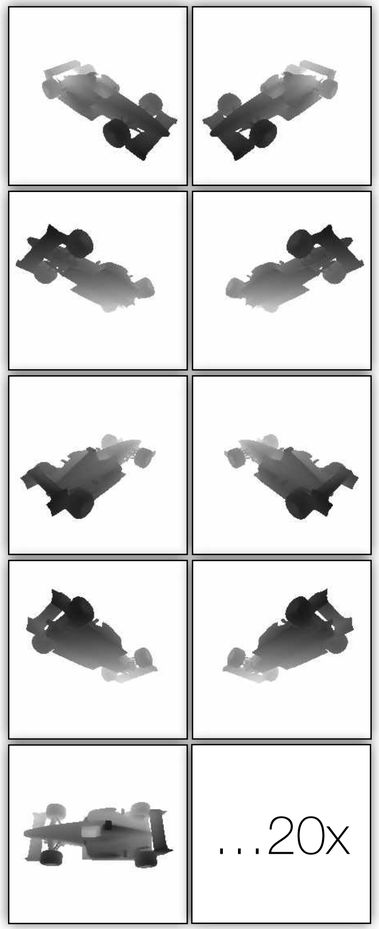
$\mathcal{L}_{\text{unique}}(\omega) =$ only one way to represent a convex

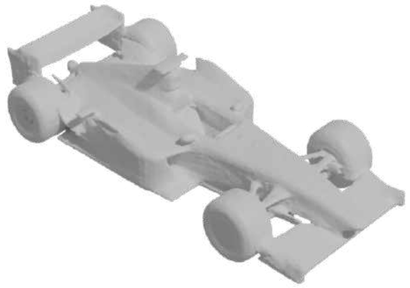
$\mathcal{L}_{\text{merged}}(\omega) =$ prevents vanishing gradients

Polygonal meshes @ inference time



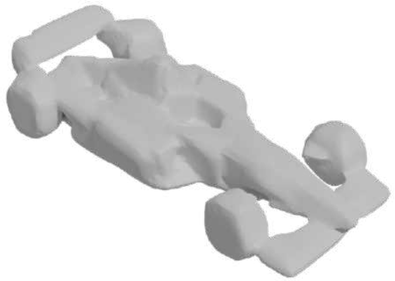
Multi view reconstruction – {Depth}→3D





Ground Truth
Mesh

SIF
[Genova et al. 2019]

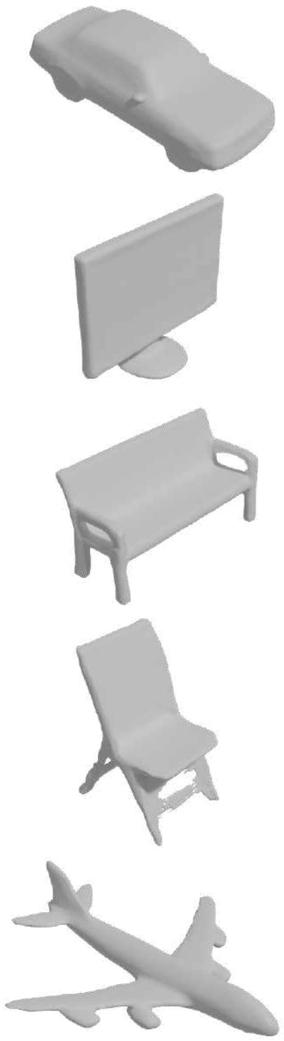


CvxNet
[Deng et al. 2020]

OccNet
[Mescheder et al. 2019]

Category	F-Score		
	OccNet	SIF	CvxNet
airplane	79.52	71.40	**84.68**
bench	71.98	58.35	**77.68**
cabinet	71.31	59.26	**76.09**
car	69.64	56.58	**77.75**
chair	63.14	42.37	**65.39**
display	63.76	56.26	**71.41**
lamp	**51.60**	35.01	51.37
speaker	58.09	47.39	**60.24**
rifle	78.52	70.01	**83.63**
sofa	69.66	55.22	**75.44**
table	68.80	55.66	**71.73**
phone	85.60	81.82	**89.28**
vessel	66.48	54.15	**70.77**
mean	69.08	59.02	**73.49**

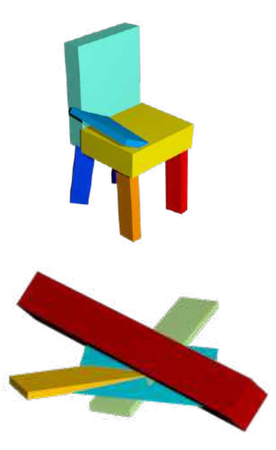
Single view reconstruction (SVR)



N/A

N/A

N/A



Input Image

Ground Truth Mesh

CvxNet [Deng et al. 2020]

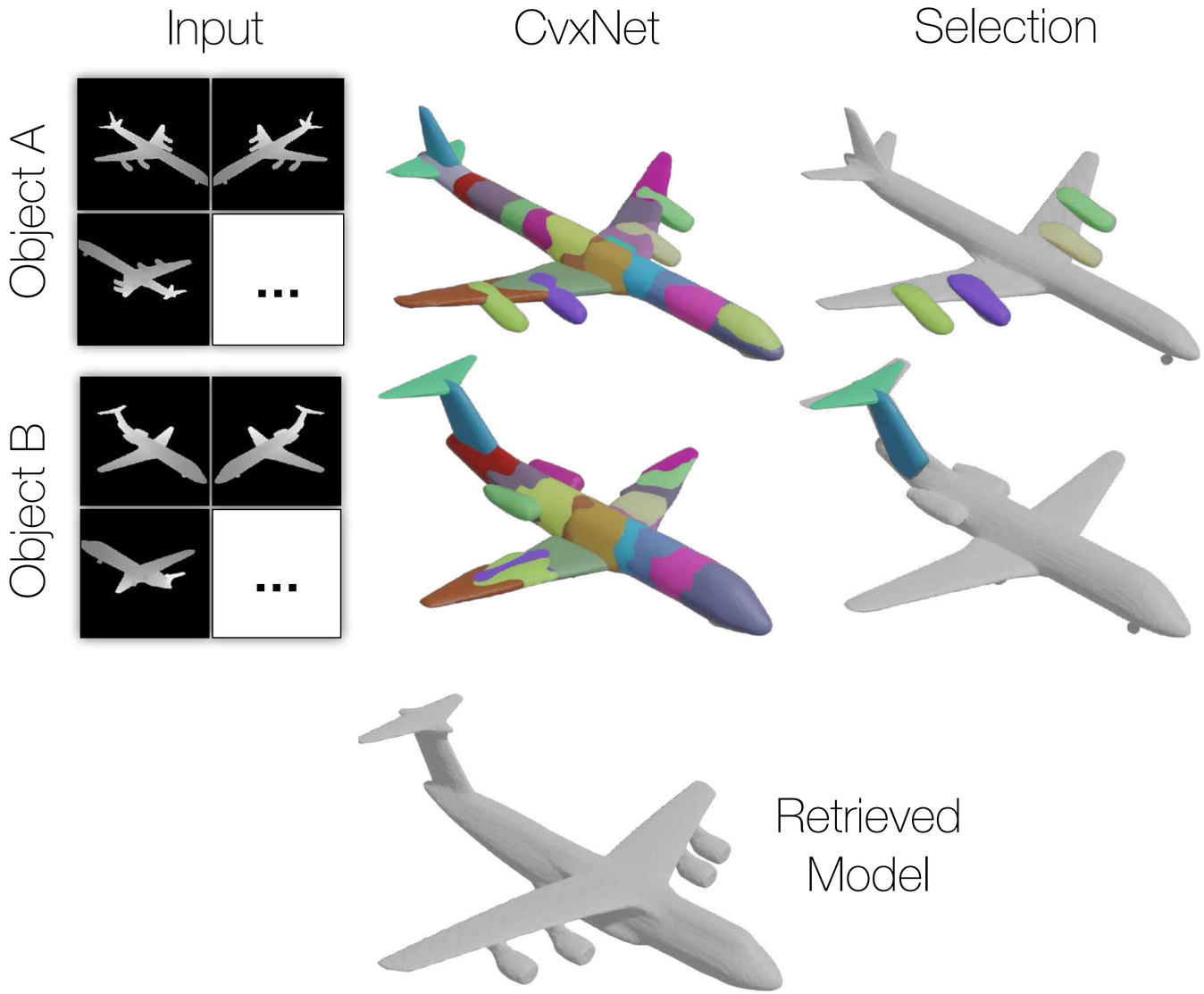
AtlasNet [Groueix et al. 2019]

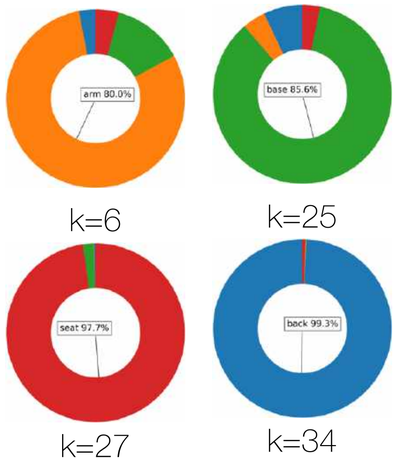
OccNet [Mescheder et al. 2019]

Structured Implicits [Genova et al. 2019]

Volumetric Primitives [Tulsiani et al. 2017]

Applications: Retrieval & Correspondence





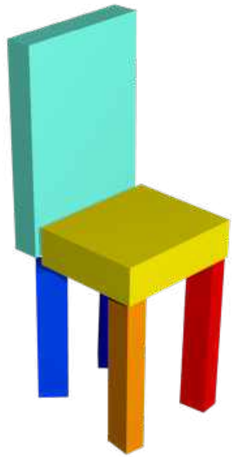
Part	Accuracy CvxNet	Accuracy BAE	BAE*
back	**91.50%**	86.36%	91.81%
arm	38.94%	**65.75%**	71.32%
base	71.95%	**88.46%**	91.75%
seat	**90.63%**	73.66%	92.91%

Ground Truth

CvxNet

Shape approximation – comparisons

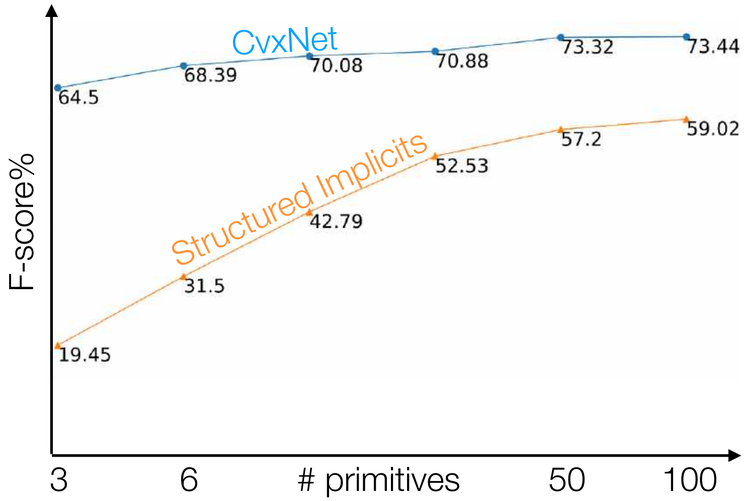
Ground Truth
Mesh

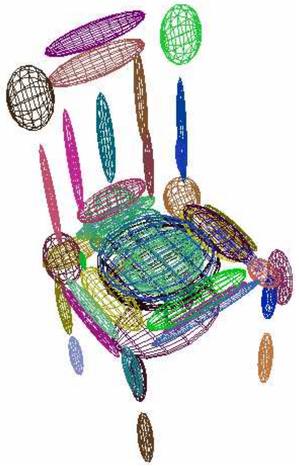


Volumetric Primitives
[Tulsiani et al. 2017]

CvxNet
[Deng et al. 2019]

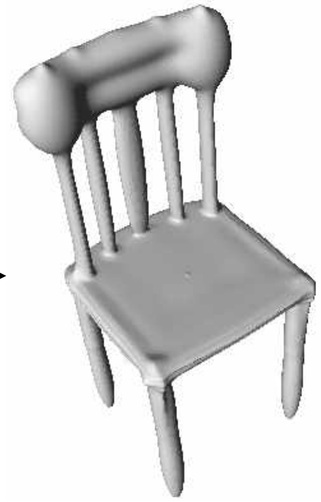
«union»





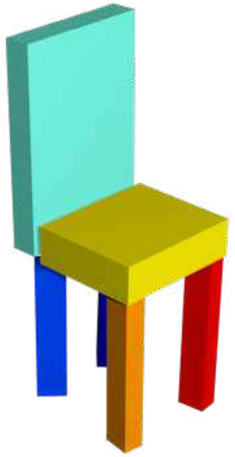
Structured Implicits
[Genova et al. 2019]

«structuring»

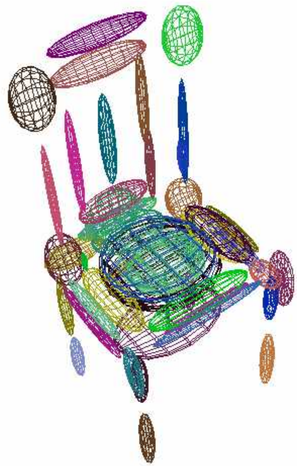


Shape approximation – quality v.s. budget

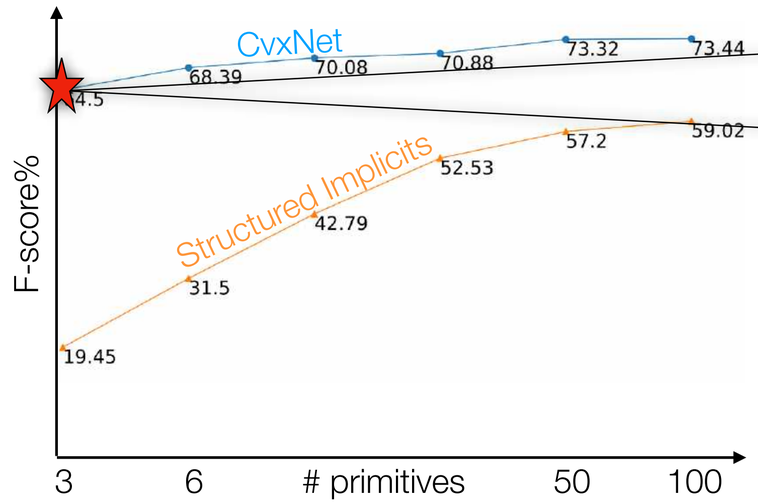
Ground Truth
Mesh



Volumetric Primitives
[Tulsiani et al. 2017]



Structured Implicits
[Genova et al. 2019]

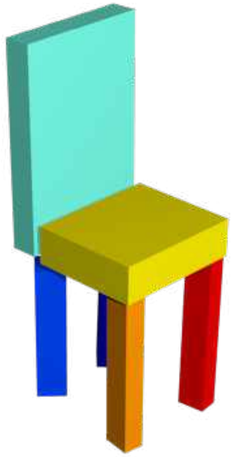


«union»

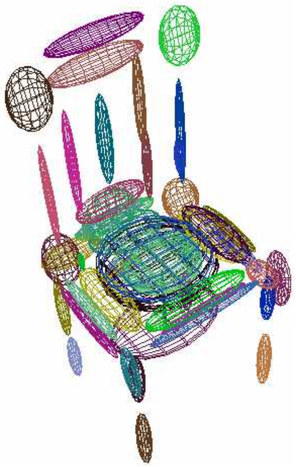
CvxNet
[Deng et al. 2019]

Shape approximation – quality v.s. budget

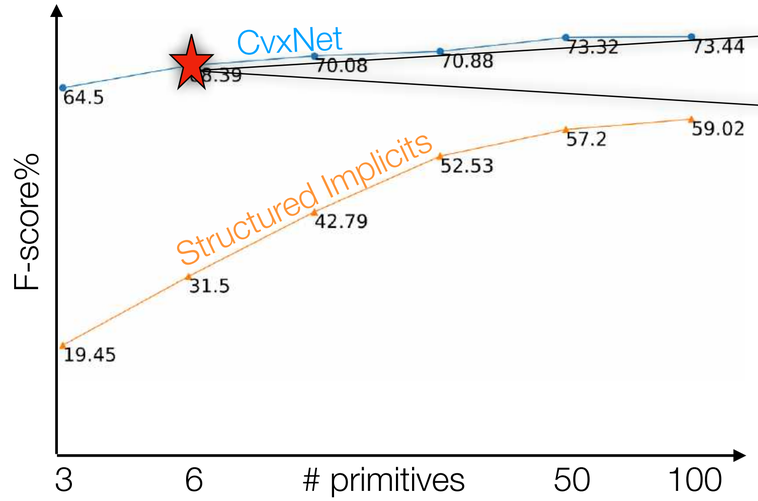
Ground Truth
Mesh



Volumetric Primitives
[Tulsiani et al. 2017]



Structured Implicits
[Genova et al. 2019]

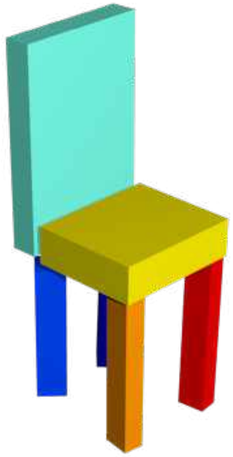


CvxNet
[Deng et al. 2019]

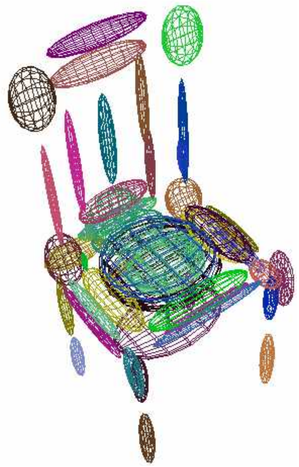
«union»

Shape approximation – quality v.s. budget

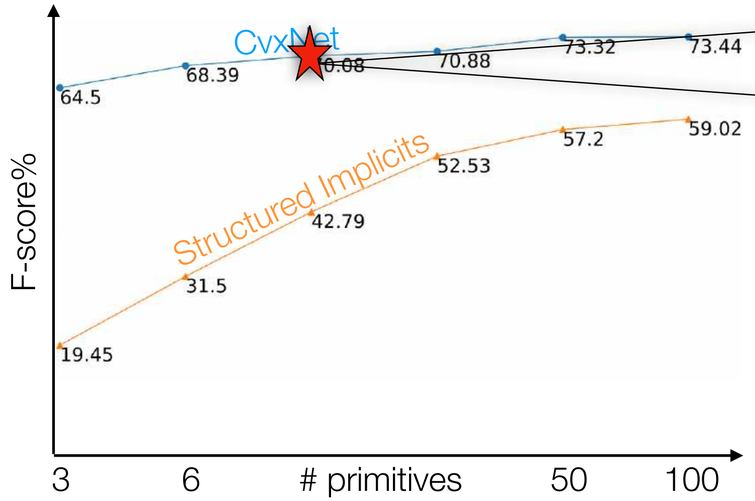
Ground Truth
Mesh



Volumetric Primitives
[Tulsiani et al. 2017]



Structured Implicits
[Genova et al. 2019]

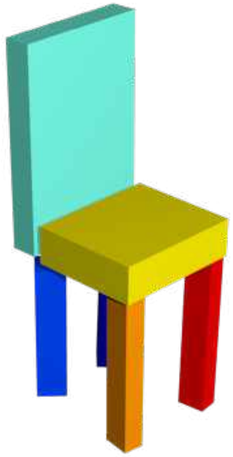


CvxNet
[Deng et al. 2019]

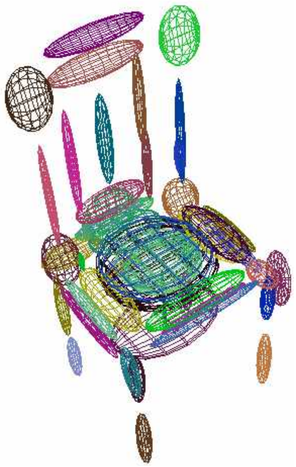
«union»

Shape approximation – quality v.s. budget

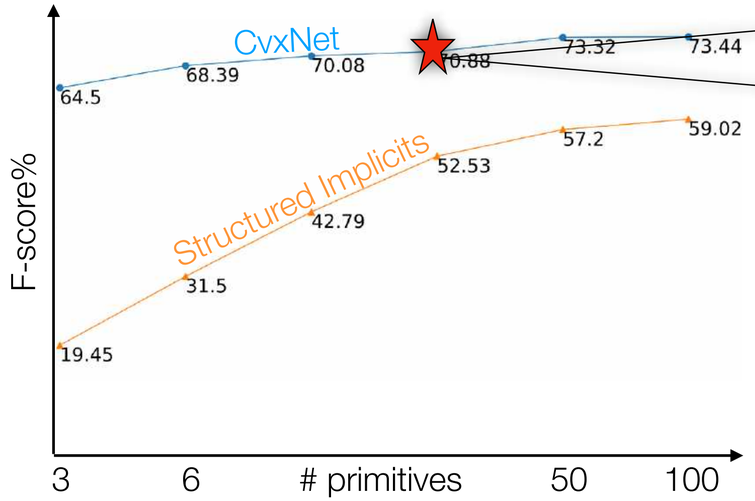
Ground Truth
Mesh



Volumetric Primitives
[Tulsiani et al. 2017]



Structured Implicits
[Genova et al. 2019]

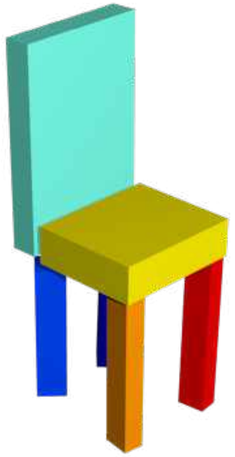


CvxNet
[Deng et al. 2019]

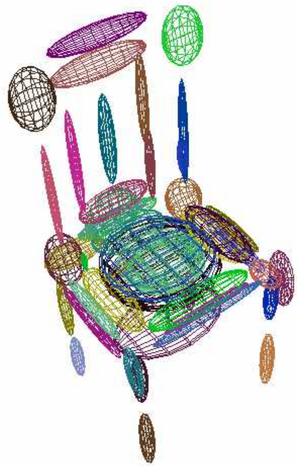
«union»

Shape approximation – quality v.s. budget

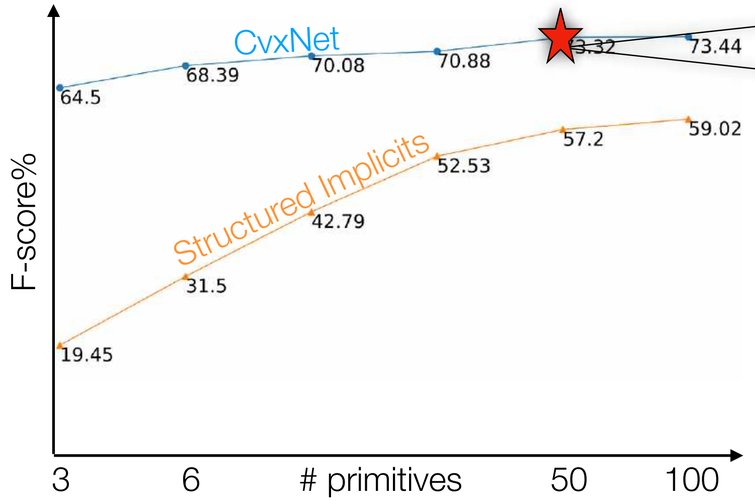
Ground Truth
Mesh



Volumetric Primitives
[Tulsiani et al. 2017]



Structured Implicits
[Genova et al. 2019]

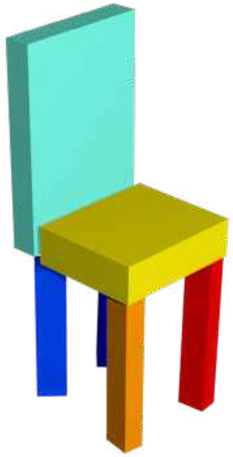


CvxNet
[Deng et al. 2019]

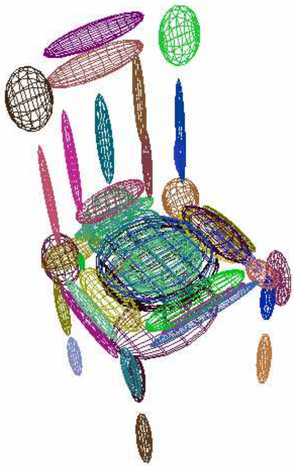
«union»

Shape approximation – quality v.s. budget

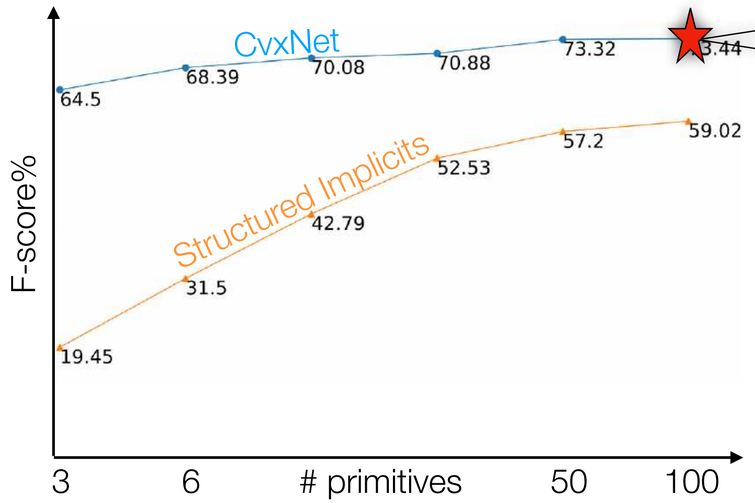
Ground Truth
Mesh



Volumetric Primitives
[Tulsiani et al. 2017]



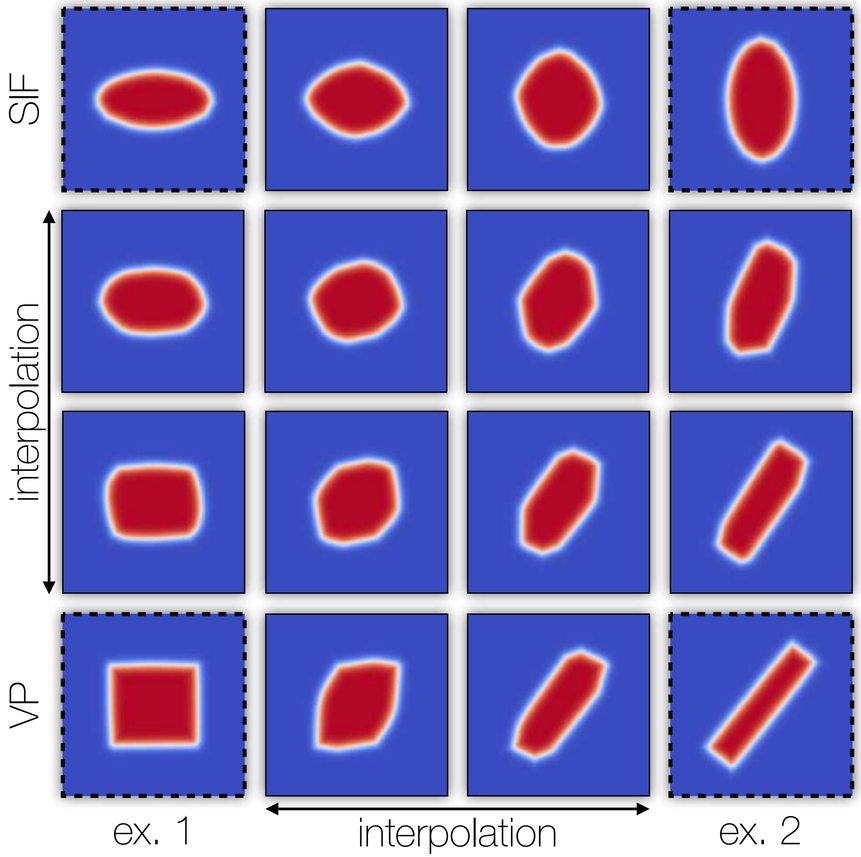
Structured Implicits
[Genova et al. 2019]



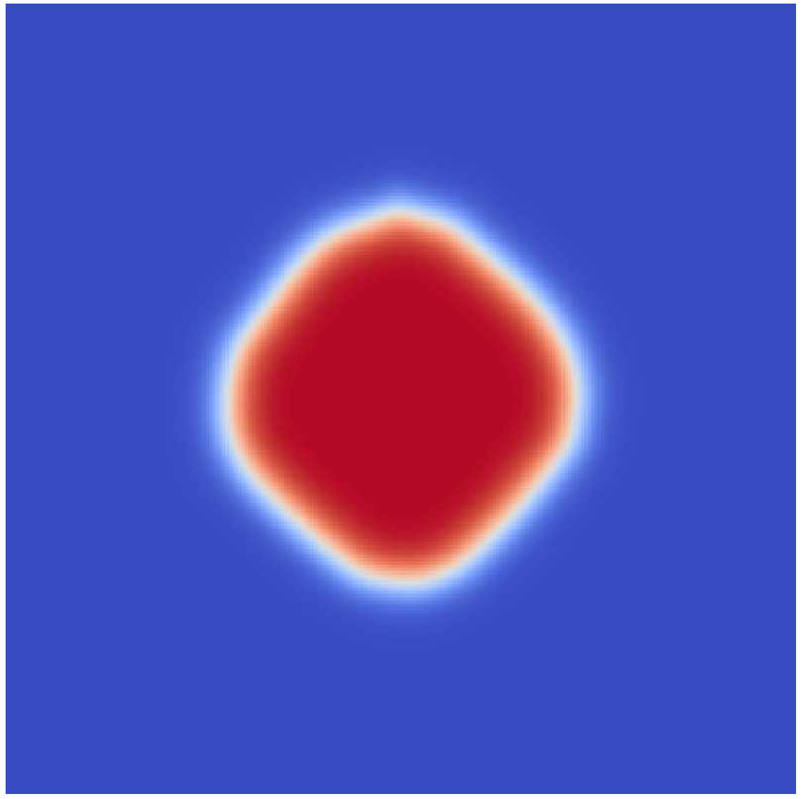
CvxNet
[Deng et al. 2019]

«union»

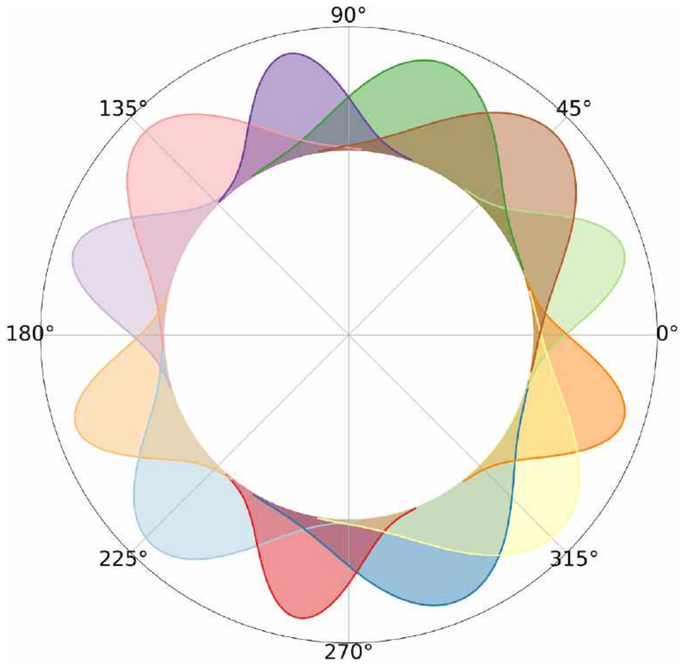
Explore the shape space of primitives



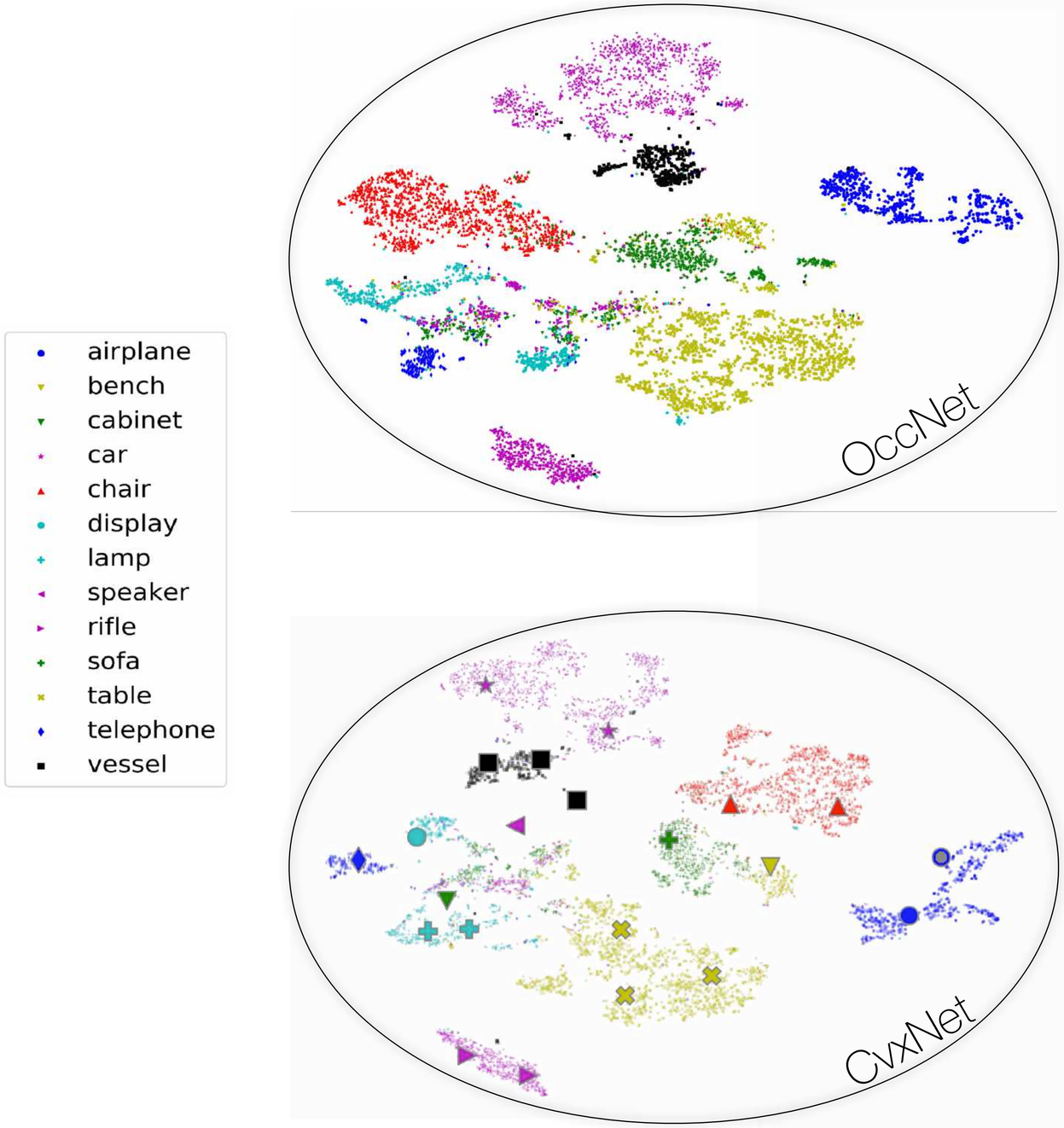
Bilinear Interpolation



Linear Interpolation



Correlation of Hyperplanes

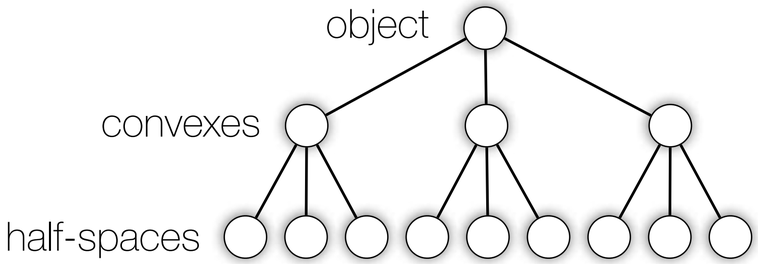

airplane
bench
cabinet
car
chair
display
lamp
speaker
rifle
sofa
table
telephone
vessel
OccNet
CvxNet

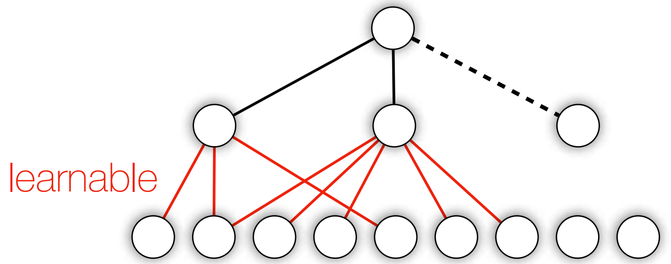
Related work @ CVPR 2020

Deng et al. «CvxNet»
smooth convexes, **fixed** graph

focus

Chen et al. «BSP-Net»
graph learning, **sharp** details





CvxNet
Learnable Convex Decomposition

Boyang Deng Kyle Genova Soroosh Yazdani
Sofien Bouaziz Geoffrey Hinton Andrea Tagliasacchi

Google Research

cvxnet.github.io