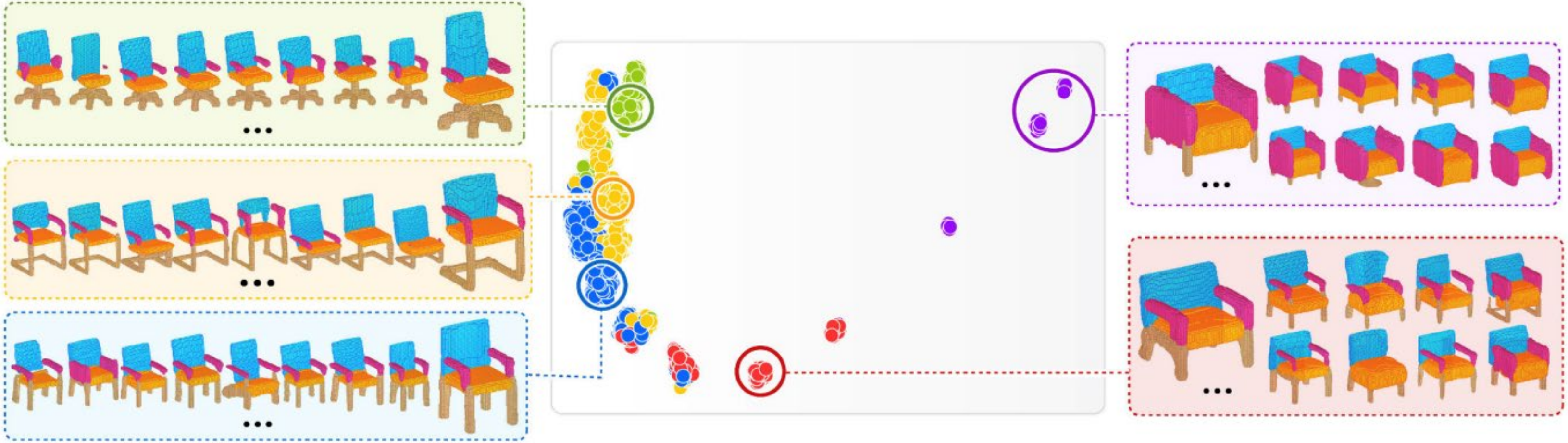
Hierarchical Generation of Structure and Geometry

Presenter: Boxiao Pan

Global-to-Local Generative Model for 3D Shapes

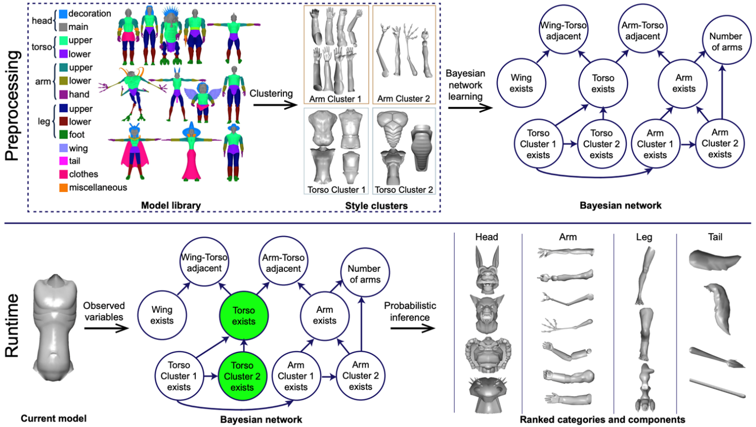
Hao Wang*, Nadav Schor*, Ruizhen Hu, Haibin Huang, Daniel Cohen-Or, Hui Huang. SIGGRAPH Asia 2018.

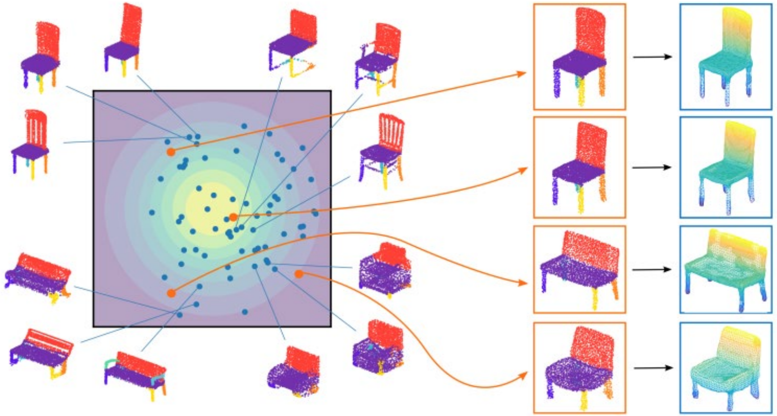
Problem statement



Given a collection of 3D semantically **segmented** shapes, generate new shapes from the same distribution.

Previous works





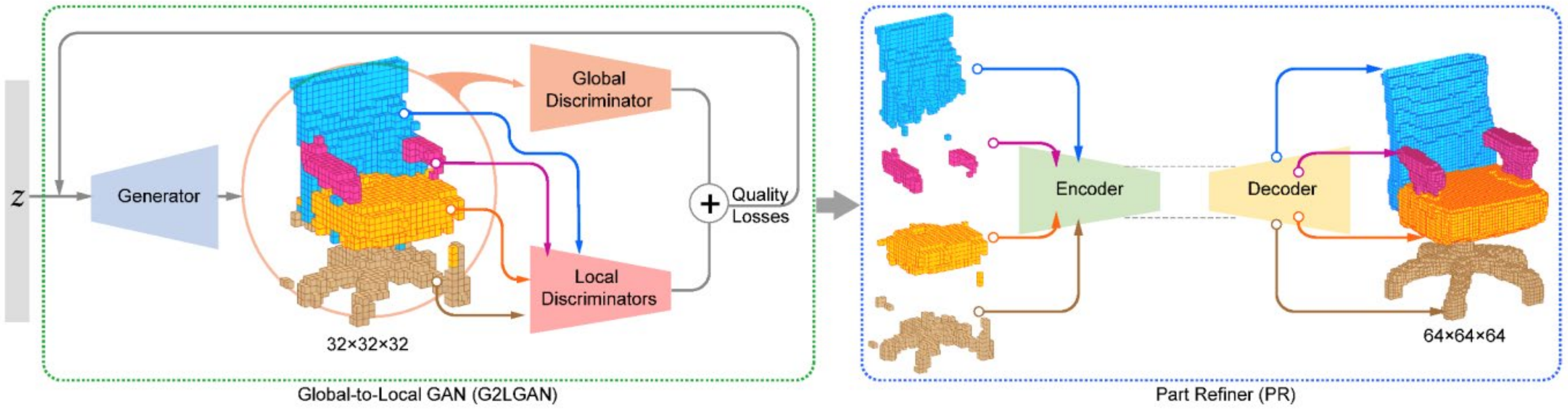
Classical approach [Chaudhuri*, Kalogerakis* et al. 2011]

- Compose by probabilistically combine existing parts
- Not able to create new parts

Deep learning-based approach [Nash and Williams, 2017]

- Assume dense point correspondences for objects
- Only attend to global shapes

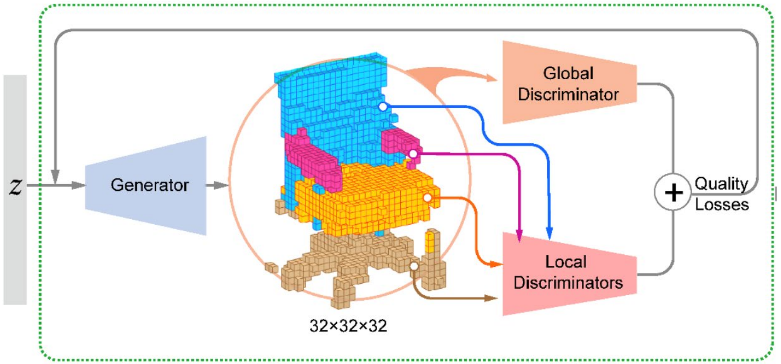
Method



- Global-to-Local GAN: synthesizes a 32^3 model
- Part Refiner: upsamples the model to 64^3, and refines the individual parts

Global-to-Local GAN

- Takes a random noise vector to produce a 32^3 segmented shape.
- Problems:
  - Generated shape is rough and often contains floating disconnected voxels.
  - Part voxels often penetrate and intermix with adjacent parts.
- Solutions:
  - Multiple part-level discriminators
  - Smoothness loss that improves shape smoothness and alleviate voxel disconnection issue
  - Purity loss that prevents part intermingling



Global-to-Local GAN (G2LGAN)

Global-to-Local GAN

- Local discriminator: one for each semantic part. Trained jointly with the global discriminator.
- Smoothness loss - floating voxels:

$$\mathcal{L}_s = \sum_x (S_c(x))^2,$$

  where $S_c(x)$ is the sum of L1 distance on the label vector between voxel x and its adjacent voxels (background / not background).
- Purity loss - part intermingling: Similar to smoothness loss, but compares the part labels.

0	1	0	2	1	1	1
1	**4**	**3**	**3**	0	1	1
0	**3**	**3**	**4**	1	1	1
1	2	0	1	0	1	1
1	1	1	1	1	1	1
1	2	2	1	2	2	1
1	2	1	0	1	2	1

(a) smoothness

-	-	-	-	1	0	-
-	0	-	1	**4**	1	-
-	-	1	-	2	1	-
-	0	1	0	2	1	-
-	1	0	1	**3**	2	-
-	1	-	-	-	1	-
-	0	-	-	-	0	-

(b) purity

0	1	0	2	2	1	1
1	**4**	**3**	**4**	**4**	2	1
0	**3**	**4**	**4**	3	2	1
1	2	1	1	2	2	1
1	2	1	2	**4**	3	1
1	3	2	1	2	3	1
1	2	1	0	1	2	1

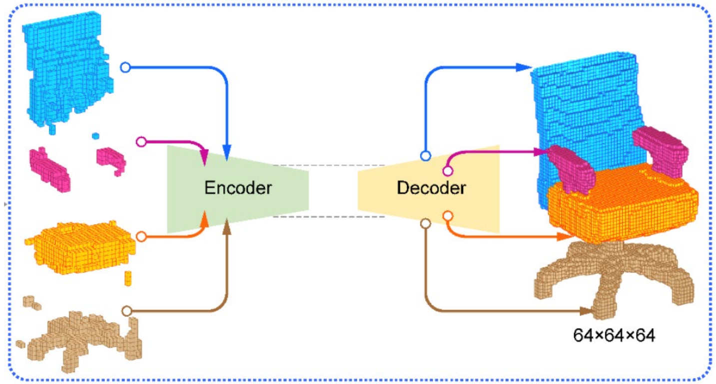
(c) combination

Floating voxels

Part intermingling

Part Refiner

- Goal: upsample and refine parts.
- Encode and then decode to reconstruct input parts at a higher resolution.
- Upsample: pairs of 32^3 and 64^3 models from the original training set.
- Refinement: pairs of 32^3 generated model and its nearest-neighbor 64^3 ground-truth model.
- Refinement: additional pairs of noise (very few voxels) and empty space.



Part Refiner (PR)

Results - shape generation

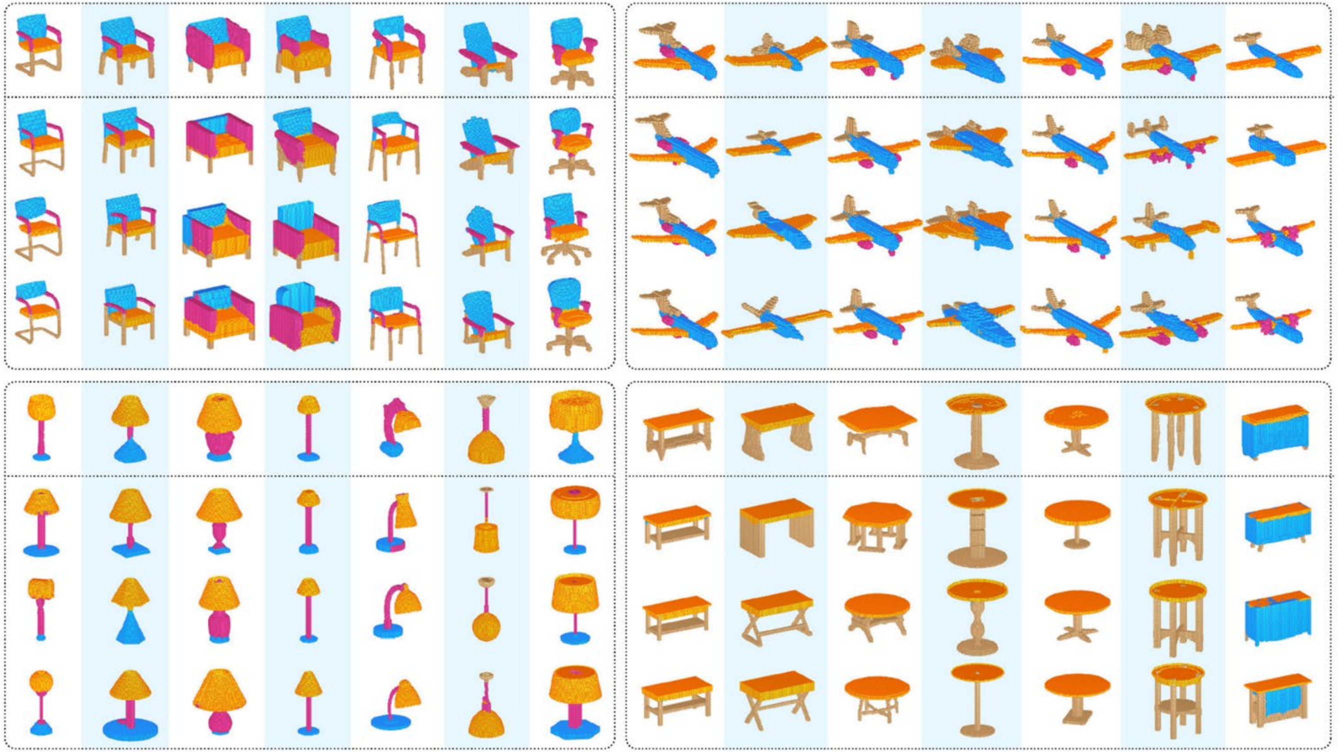


Fig. 4. A gallery of our generated Chairs, Airplanes, Lamps, and Tables shown above, with their 3-nearest-neighbors retrieved from the training set.

Results - part refinement

Fig. 5. PR Improvement. For each category, we present four examples of the improvement achieved by the PR. Shapes generated by G2LGAN are shown on the top row, and their PR enhanced versions are provided underneath for a clear comparison.

Results - shape interpolation

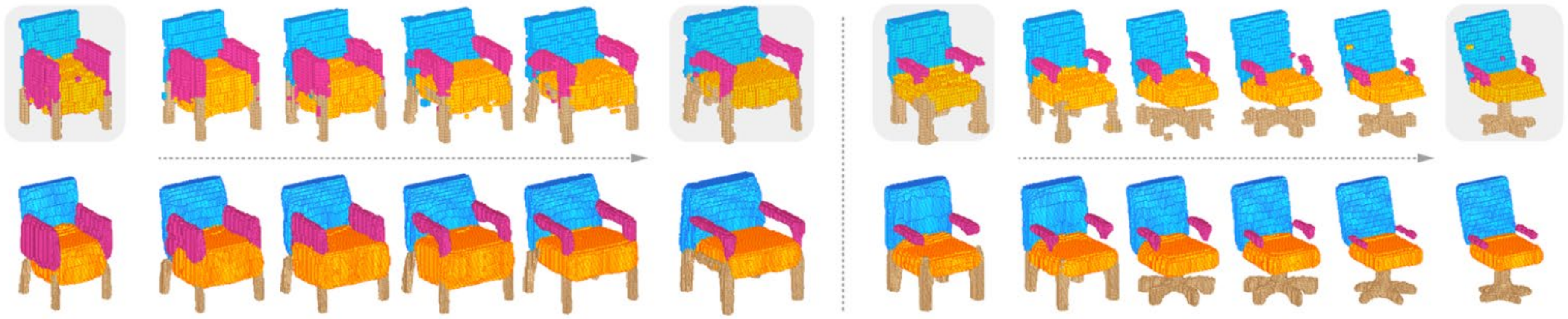


Fig. 9. Interpolating between two pairs of chairs at the left and right, respectively. Shapes generated by G2LGAN are shown on the top row, and their PR enhanced versions are provided underneath. The PR results stay sharp and clear throughout the different stages, where usual latent space interpolation results, such as G2LGAN, have some artifacts and noise. This results in a clear-cut transition between the different stages in the PR interpolation, which can be seen from the arm-rests of the left example and the legs in the right example.

Results - ablation study

Table 1. Inception and Symmetry comparisons of generated chairs. Both baselines receive lower scores in all categories, in comparison to our models, G2LGAN and G2LGAN+PR. Local Discriminators (LD) better improve the Baseline in the symmetry score for Leg and Arm while the Quality Losses (QL) better improve the other categories.

Models	Inception	Symmetry				
		Shape	Back	Seat	Leg	Arm
3DGAN	5.84	0.70	0.71	0.76	0.35	0.16
Baseline (BL)	5.55	0.78	0.76	0.80	0.55	0.51
BL & LD	5.62	0.80	0.78	0.80	0.64	0.63
BL & QL	5.99	0.82	0.82	0.85	0.60	0.53
G2LGAN	6.00	0.84	0.83	0.85	0.63	**0.66**
G2LGAN+PR	**6.17**	**0.91**	**0.93**	**0.94**	**0.71**	0.64
GT	8.16	0.96	0.93	0.94	0.84	0.84

SDM-NET: Deep Generative Network for Structured Deformable Mesh

Lin Gao, Jie Yang, Tong Wu, Yu-Jie Yuan,
Hongbo Fu, Yu-Kun Lai, Hao Zhang. SIGGRAPH Asia 2019.

Idea

- Compared to voxels and points popularly employed by previous works for shape generation, triangle meshes provide better visual quality and controllability.
- Previous works on mesh generation tend to produce low-quality outputs with artifacts, and often fail to model complicated shaped objects.

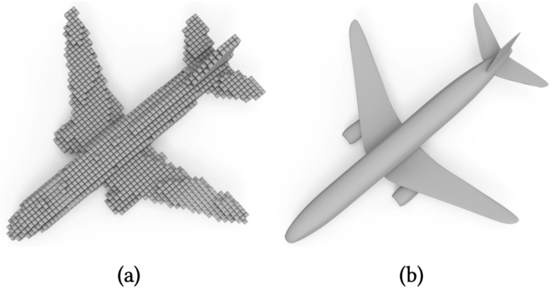


Fig. 12. Visual comparison between the global-to-local method [Wang et al. 2018a] (a) and our technique (b) for shape generation.

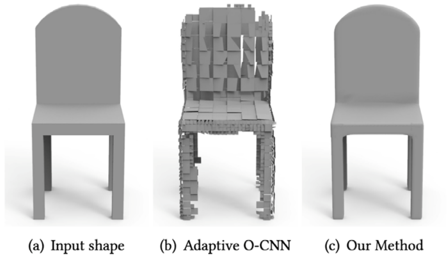
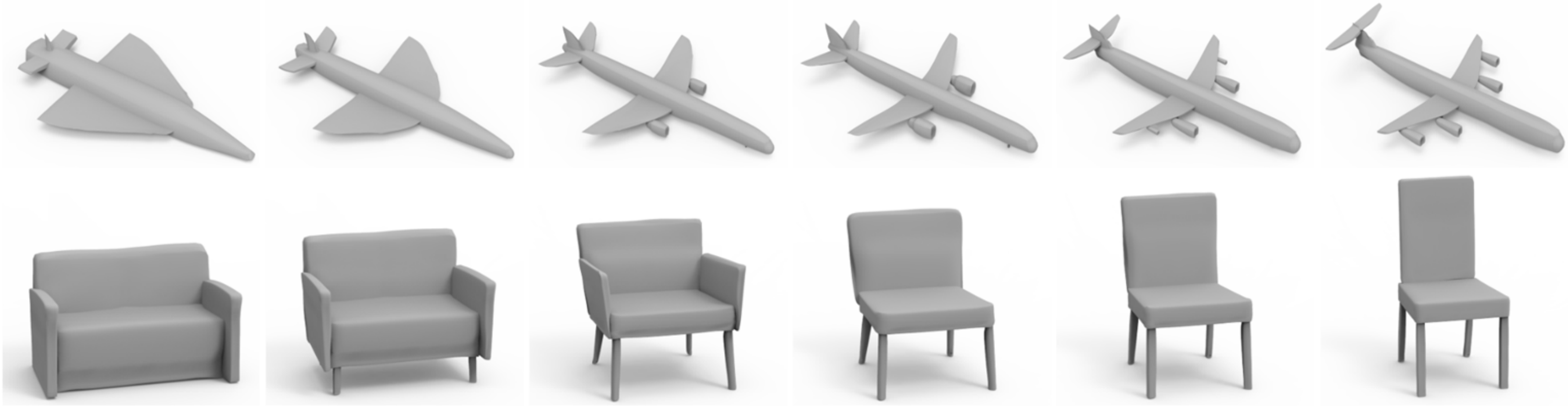


Fig. 10. Visual comparison of the decoded shapes with Adaptive O-CNN [Wang et al. 2018b] and our method. Compared with Adaptive O-CNN, while the planar regions of the chair can be decoded by both methods, the curved regions such as the top of the chair back can be recovered only by our method.

Idea



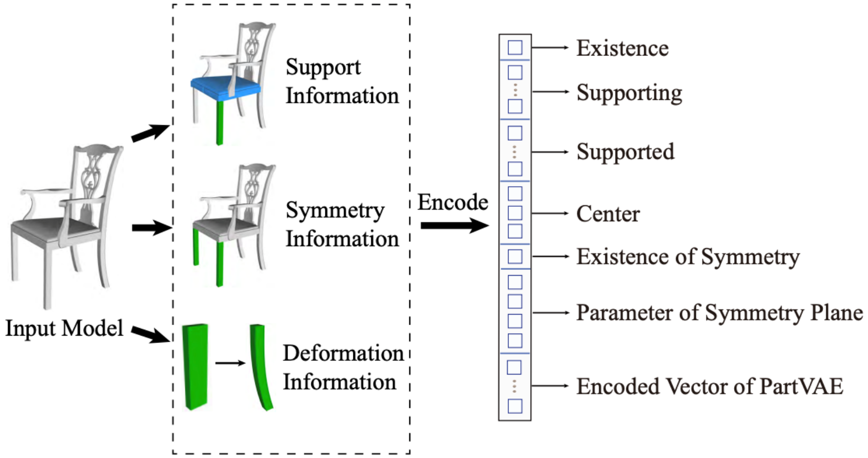
The goal is to generate structured meshes composed of deformable parts.

Method

- A local VAE that encodes geometry at a part level (PartVAE).
- A global VAE that encodes the structure and geometry of the overall shape (SP-VAE).
- A support-based part connection optimization procedure which further enables that the generated shapes are plausible and physically valid.

Part encoding

- Structure encoding:
  - Support
  - Symmetry
- Geometry encoding:
  - A latent vector that encodes the deformation process from a bounding box to the part.



Geometry part encoding

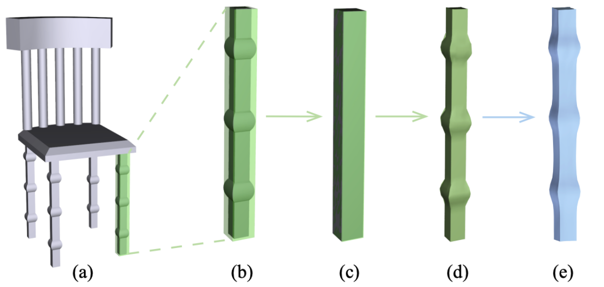
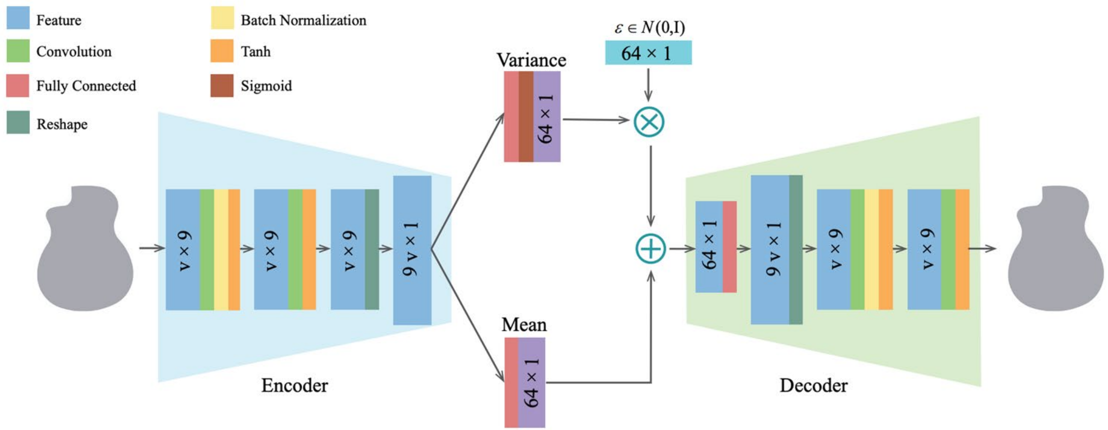
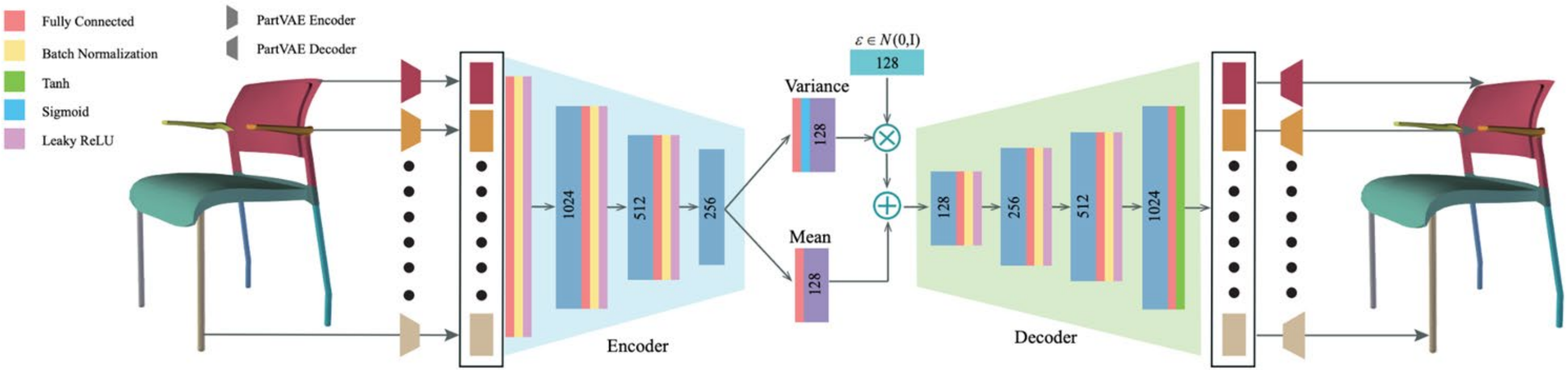


Fig. 3. An example of representing a chair leg with detailed geometry by a deformable bounding box. (a) a chair with one of its leg parts highlighted, (b) the targeted part in (a) with the bounding box overlaid, (c) the bounding box used as the template, (d) deformed bounding box after non-rigid registration, (e) recovered shape using PartVAE.



- Part deformations are done by applying non-rigid registration from the bounding box to the part, which are represented by a 9-dimensional vector that characterizes the local deformation of each vertex.
- This deformation vector is then used to train the PartVAE, which encodes local part deformations.

Global shape encoding



- The support, symmetry and geometry encoding for parts are used to train the global SP-VAE, which encodes the structure of the overall shape.
- New shapes can then be generated by sampling in the learned latent space.

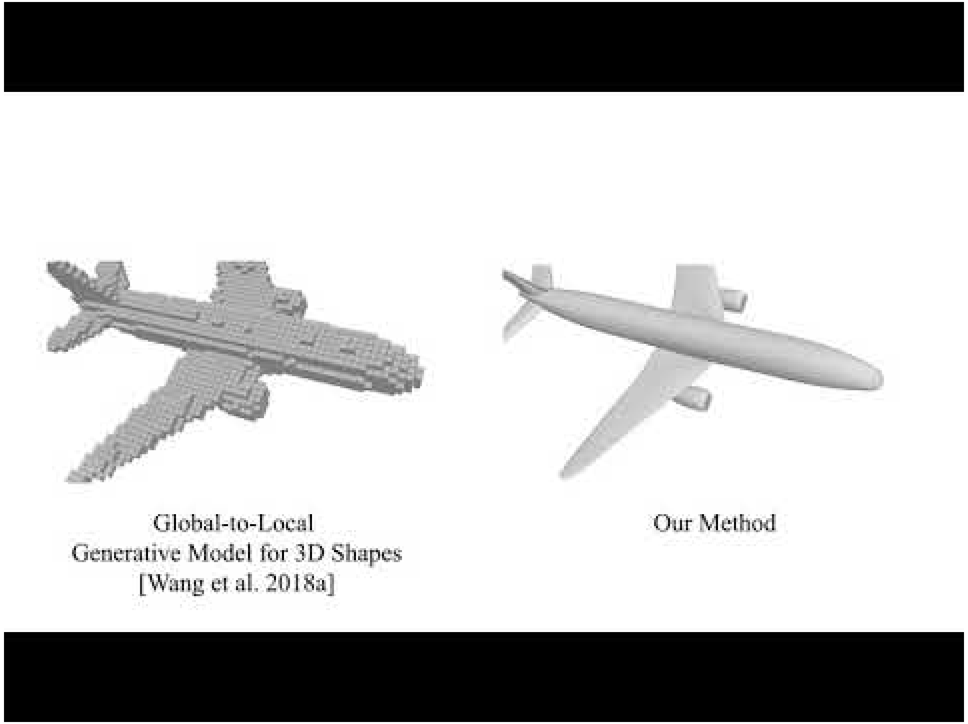
Global optimization for refinement

- Minimize the changes between the optimized parts' position / scale and the original position / scale:

$$\sum_i \|\mathbf{p}_i' - \mathbf{p}_i\|^2 + \alpha\|\mathbf{q}_i' - \mathbf{q}_i\|^2$$

  while satisfying a number of constraints:
  - Symmetry constraint
  - Equal length constraint
  - Support relationship constraint
  - Stable support constraint

Results


Global-to-Local
Generative Model for 3D Shapes
[Wang et al. 2018a]
Our Method

Thank you!