

Structure-Aware Shape Synthesis

Elena Balashova¹, Vivek Singh², Jiangping Wang², Brian Teixeira², Terrence Chen², Thomas Funkhouser¹

1 - Princeton University, NJ, USA, 2 - Siemens Healthineers, Medical Imaging Technologies, Princeton, NJ, USA

Goal

Synthesize realistic shapes that respect structure

Shape Only Synthesis

Structure-Aware Synthesis

Approach

Structure Decoder
Given a shape, generate its structure

Shape Structure Consistency

Compute consistency between a shape and a structure

max. of shape voxel grid within a Gaussian sphere around the landmark coordinate

Consistency Definition

$$L_{consist} = \sum_{s \in S} f_\gamma(s, f_\beta(s))$$

predicted structure

shape s sampled from manifold

Consistency-Aware Training

Train a structure-aware shape generator

Solution:
Sample shapes from manifold and use shape-structure consistency as loss

Evaluate shape-structure consistency loss $L_{consist}$

Training Pipeline

Sampled Random Chair

(b) Shape-Structure Consistency

(c) Shape Encoder-Generator Loss

(d) Structure Detector Model

(e) Structure Detector Loss

Shape Model

3D Shape Encoder

3D Shape Generator

Structure Detector

Structure Detector

Shape Structure Consistency

Shape Loss

Landmark Correctness Loss

Landmark Consistency Loss

Applications

Shape Interpolation

Shape Completion

Landmark Prediction

Bibliography

[1] J. Wu, T. Xue, J. J. Lim, Y. Tian, J. B. Tenenbaum, A. Torralba, and W. T. Freeman. Single image 3D interpreter network. In European Conference on Computer Vision, pages 365–382. Springer, 2016.

[2] R. Girdhar, D. F. Fouhey, M. Rodriguez, and A. Gupta. Learning a predictable and generative vector representation for objects. In European Conference on Computer Vision pages 484–499. Springer, 2016.

[3] L. Ge, H. Liang, J. Yuan, and D. Thalmann. 3d convolutional neural networks for efficient and robust hand pose estimation from single depth images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, volume 1, page 5, 2017.

Evaluation

Metric	Shape Only	Structure - Aware
Dense input	0.22	0.25
Sparse input	0.05	0.14

Table 1. Quantitative evaluation on the scan completion task. The reported numbers are average IoU computed with respect to the ground truth shape corresponding to the scan.

Figure 1. Comparison between 3D VAE-GAN [1] (an adversarial approach), 3D VAE, and our Structure-Aware method. The GAN model:

- generates slightly less noisy results
- does not always generate structurally meaningful outputs.

Region	3D VAE	Structure-Aware
All	~0.15	~0.55
Leg Chair Region	~0.10	~0.40
Back	~0.15	~0.60
Seat	~0.20	~0.55

Figure 2. Evaluation of average shape structure consistency scores by chair region. The shape structure consistency score is in range [0,1], where higher numbers indicate better performance. Note that for both methods, the most difficult region is the leg area, and the easiest region is the seat area.