Exploring Generative 3D Shapes Using Autoencoder Networks

Nobuyuki Umetani
Motivation

• Machine learning is booming
Motivation

• Machine learning for 3D shape
  – Understanding a class of shapes

In my understanding, shapes of cars look like..
Motivation

• Machine learning for 3D shapes
  – Finding low dimensional manifold in space
Synthesizing New Shapes

- Mapping low dimensional manifold to the shapes
Preview of the Result: Slider bar

video speed x3
Preview of the Result: Random

FPS: 59.76
Interactive Exploration

video speed x3
Parameterization problem

- Shape need to be represented by fixed dimensional vector/tensor
Parameterization problem

- Triangle mesh / NURBS are **not suitable** for ML
  - Topology / #points are not constant
Related Work: Voxel model

• Difficult to synthesize **detailed 3D shape**
  – Expensive memory cost, noise...

[FPNN, Li et al., 2017]

[O-CNN, Wang et al., 2017]

[Girdhar et al., 2016]

[3D GAN, 2016]

[OctNet, Riegler et al., 2017]

[liu et al., 2017]
Related Work: Multi-view model

- Difficult to synthesize 3D data

[Su et al., 2017]  
[Tulsiani et al., 2017]  
[Zhou et al., 2016]

[Soltani et al. 2017]  
[Park et al., 2017]  
[Lun et al. 2017]
Related Work: Point-based model

- Difficult to synthesize detailed 3D shape
  - Hard to define surface

[PointNet, Qi et al, 2017]
[PointNet++, Qi et al, 2017]
[Fan et al, 2017]

[DeformNet, Kurenkov et.al, 2017]
[Gadelha et al., 2017]
Our Approach: Mesh Representation

- Quad mesh with constant topology
- Deforming a template mesh into input shape
Our Approach: 3D Shape as Height Field

- Storing XYZ coordinates is redundant
- Height field from a cube in its normal
Naïve Projection Leads to Distortion

- Distortion means biased sampling density
- Not robust to the occlusion

Never to be sampled

occlusion

Never to be sampled
Hierarchical Projection

- We repeat subdivision and projection
- Key observation: concave shape is locally convex
Hierarchical Projection in 3D

Level 0
Level 1
Level 2
Level 3
Level 4
Level 5

XYZ positions

Parameter vector

height
+0.5m

-0.2m
Challenge

• Projecting points into polygon soup
  • Inverted triangle
  • Intersections
  • Gaps and holes
  • Internal structures
Depth Field Techniques

- Depth field can be efficiently computed with GPU
- More orthogonal projection is more trusted
Parameterization
Over 1,200 Car Shapes from ShapeNet

Data is available on my web page:  http://nobuyuki-umetani.com/

[chang et al. 2015]
Autoencoder Network

• Input and output of network is as same as possible

Minimize difference
Autoencoder Network

- We could train the network preserving shapes
Autoencoder Network

• Nonlinear dimensional reduction
Autoencoder Network

- We use Sigmoid function for activation
- Decoder takes value between [0,1]
Autoencoder Network

$$\{0.3 \ldots 0.1, \ldots 0.9\}^T$$

10 dim

6165 dim
Interactive Exploration

• Pulling operation to guide the synthesis

Minimize length

Optimize parameters
Interactive Exploration

• Our parameterization is close form
  – Easily differentiable for grad-based optimization

update
\{0.3 \ldots 0.1, \ldots 0.9\}^T
Live Demo!
Limitation

• Only handles one-class of shape
  – Difficult to handle different topologies

• Only handles nearly convex shapes
  – Starting from a template mesh other than cube
  – combinations of several meshes
Future work

• Advanced generation framework
  – GAN, VAE


• Filter operations on the subdivision
  – Convolution operation
  – Gaussian Pyramid / Laplacian Pyramid

  http://library.wolfram.com/infoncenter/Demos/4532/
Acknowledgement

• Anonymous reviewers

• Ryan Schmidt and Alec Jacobson for many advice
Exploring Generative 3D Shapes Using Autoencoder Networks

Nobuyuki Umetani