Shape Synthesis from Sketches via Procedural Models and Convolutional Networks

Haibin Huang, Evangelos Kalogerakis, Ersin Yumer, Radomir Mech

Presentation by Julien

Procedural Modeling

Procedural Modeling

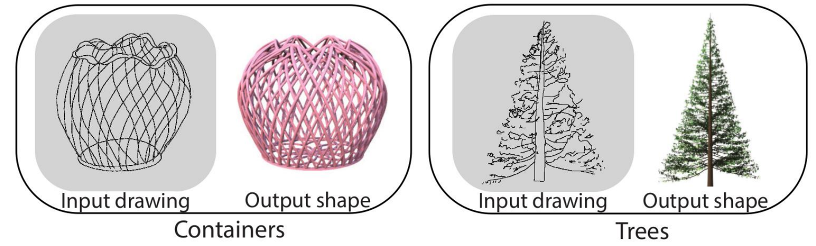
Synthesize 2D or 3D models with a set of parameters.

Figure 1: Procedurally Modeled Chairs (from Sven Havemann, Wikipedia)

Idea: Learn to translate human sketches to Procedural Modeling parameters.

Human Sketches

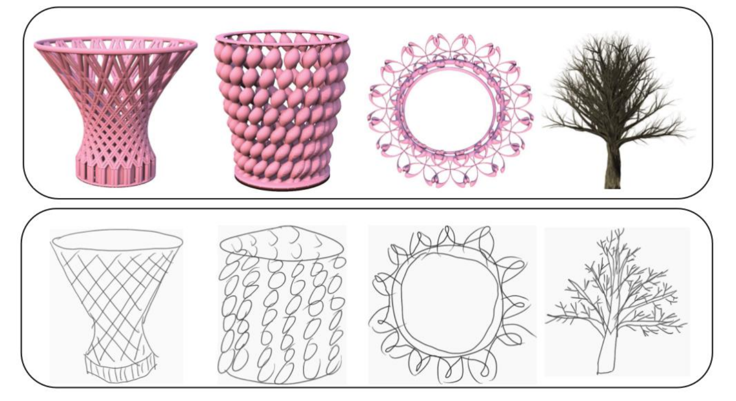
A challenge of this approach is dealing with dramatic abstractions, simplifications, and exaggerations humans make when drawing objects.



A Convolutional Neural Network Architecture

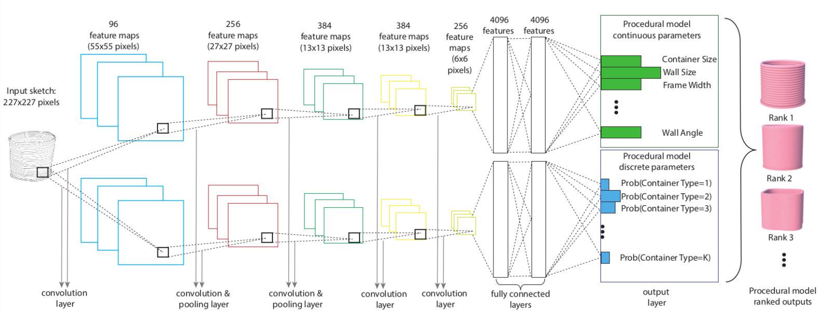
The Data

With a user study and careful curation of generated sketches, a synthetic dataset can be created for a model.



CNN Architecture

Composed of two distinct AlexNet CNN architectures.



One with additional layer for regression of continuous parameters, the other with additional layer for classification of discrete parameters.

Pre-training

AlexNet pre-trained on ImageNet gives significant performance increase for free.

Image classification is a similar task, similar visual features will be useful for sketch $\rightarrow$ shape parameter task.

The network can then be fine tuned by training end to end.

Regression and Classification Layers

The regression layer estimates a normalized parameter c within the range $[0,1]$ using a sigmoid.

$$O_c = \frac{1}{1 + \exp(-\mathbf{w}_c \cdot \mathbf{h}_L - b_c)}$$

The classification layer estimates probabilities of each value of a discrete parameter D_r using a softmax.

$$P(D_r = d) = \frac{\exp(\mathbf{w}_{d,r} \cdot \mathbf{h}_L + b_{d,r})}{\exp(\sum_{d'} \mathbf{w}_{d',r} \cdot \mathbf{h}_L + b_{d',r})}$$

Loss Functions

Loss function for C continuous parameters over S sketches in training data. Uses L^2 loss with L^2 regularization.

$$E_{\text{reg}}(\theta_1) = \sum_{s=1}^{S} \sum_{c=1}^{C} \mathbf{1}_{\delta_{c,s}} ||O_{c,s}(\theta_1) - \hat{O}_{c,s}||^2 + \lambda_1 ||\theta_1||^2$$

Loss function for R discrete parameters over S sketches in training data. Uses logistic loss with L^2 regularization.

$$E_{\text{class}}(\theta_2) = \sum_{s=1}^{S} \sum_{r=1}^{R} \log(P(D_{s,r} = \hat{d}_{s,r} \mid \theta_2)) + \lambda_2 ||\theta_2||^2$$

Results

At Runtime

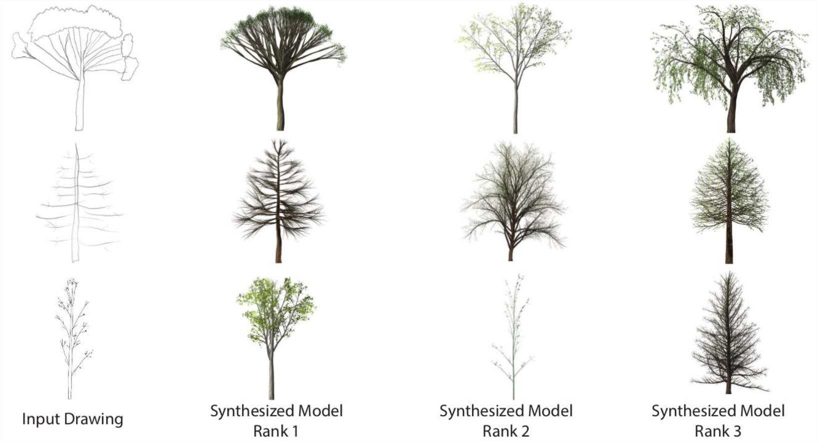
GPU execution can predict PM parameters given a sketch in 1 to 2 seconds.

This is fast enough for near-interactive use.

Method	Containers	Trees	Jewelry	Average
Nearest neighbors (Fisher)	33.3%	26.7%	46.7%	35.6%
Nearest neighbors (CNN)	26.7%	33.3%	40.0%	33.3%
SVM (Fisher)	33.3%	46.7%	60.0%	46.7%
SVM (CNN)	40.0%	33.3%	53.3%	42.2%
Single CNN	66.7%	60.0%	80.0%	68.9%
No pretraining	20.0%	13.3%	20.0%	17.8%
Our method	**80.0%**	**73.3%**	**86.7%**	**80.0%**

Figure 2: Top-3 classification accuracy for discrete parameters

Results


Input Drawing
Synthesized Model
Rank 1
Synthesized Model
Rank 2
Synthesized Model
Rank 3

Results