Single view 3D object reconstruction

Bastien Déchamps and Mathieu Orhan

4 December 2019

Summary

Summary

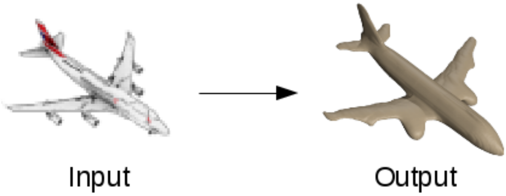
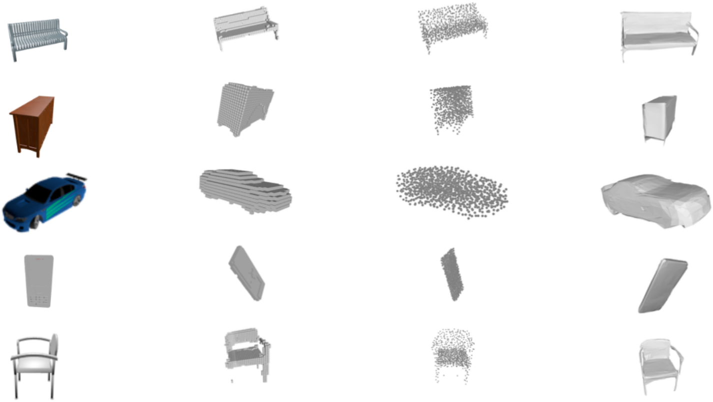


Figure – Source : bair.berkeley.edu

- Ill-posed, it requires to assume regularities, patterns and symmetries
- Still a very challenging vision task

Mesh representation

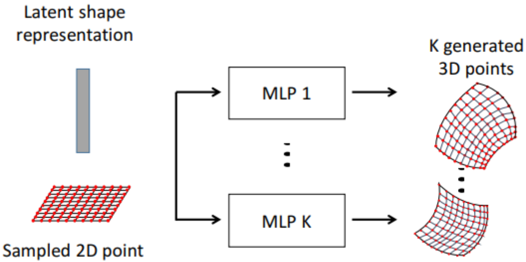


We focus on recent **deep learning** methods outputing **meshes**.

Objectives

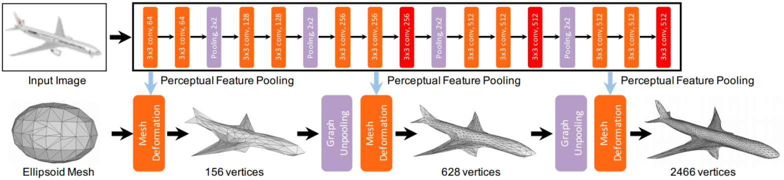
- ✓ **Compare** AtlasNet [2] and Pixel2Mesh [5] as fairly as possible
- ✓ Experiment **regularization** schemes proposed by Pixel2Mesh

Summary



Key points :

- A ResNet [3] encodes the input image
- Multiple templates are deformed **pointwisely** by MLPs to learn 2-manifolds in 3D



Key points :

- A **GNN** deforms a **template** in a coarse-to-fine fashion
- It uses **projected features** extracted by VGG-16 [4] from the image

For S_1, S_2 two point clouds :

Chamfer distance

$$d(S_1, S_2) = \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2^2 + \sum_{y \in S_2} \min_{x \in S_1} \|x - y\|_2^2$$

- Not a distance on meshes, but on point clouds $\rightarrow$ Does not directly reflect mesh quality
- Depends on data normalization, number of points, ...
- Still a very popular metric.

Mesh regularizers

Laplacian regularization

$$\mathcal{L}_{\text{lap}} = \sum_{p} \|\delta'_p - \delta_p\|_2^2$$

where $\delta_p = p - \frac{1}{|\mathcal{N}(p)|} \sum_{k \in \mathcal{N}(p)} k$ and δ'_p is the same quantity after deformation.

Expectation : enforce a global movement and avoid self intersections

Edge Length regularization

$$\mathcal{L}_{\text{edge}} = \sum_{p} \sum_{k \in \mathcal{N}(p)} \|p - k\|_2^2$$

Expectation : prevents outliers

Summary

Dataset

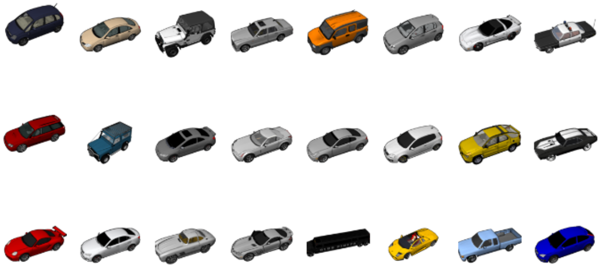


Figure – Sample of the dataset

- Compare both methods on the same data from ShapeNet [1]
- Training only on **cars**
- Same normalization before metrics computation

Quantitative comparison

	Chamfer	F1-score
Atlas (1 templ.)	5.11	0.41
Atlas (5 templ.)	4.93	0.43
Atlas (25 templ.)	4.95	0.42
P2M	**3.98**	**0.54**

Table – Performance of both methods (cars)

Qualitative comparison

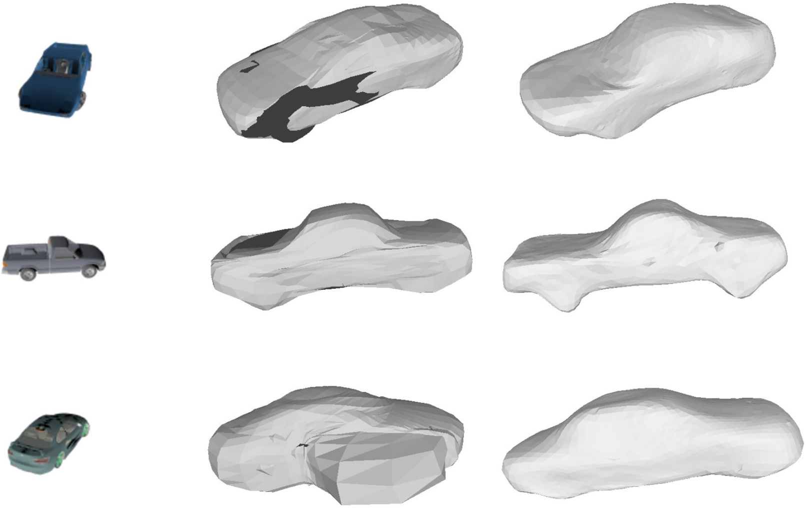
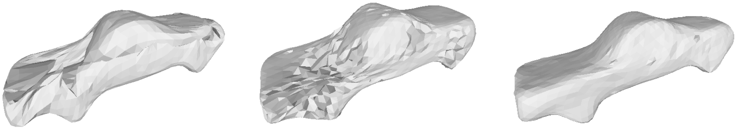


Figure – Meshes outputs by AtlasNet (middle) and Pixel2Mesh (right)

Pixel2Mesh regularization

	Chamfer	F1-score
P2M (base)	3.98	0.54
P2M (-edge)	**3.76**	**0.55**
P2M (-lapl.)	3.95	0.53

Table – Pixel2Mesh, ablation study (cars)



(a) wt. edge-length (b) wt. Laplacian (c) both reg.

AtlasNet regularization

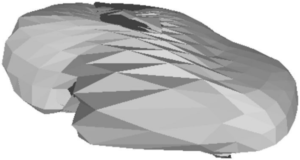
Edge-length regularization :

$$\mathcal{L} = \mathcal{L}_{\text{chamfer}} + \lambda \mathcal{L}_{\text{edge}}$$

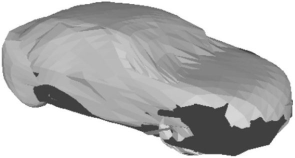
	Chamfer	F1-score
Atlas (5)	4.93	0.42
Atlas (5) + reg	**4.89**	**0.43**

Table – AtlasNet, regularization effect.

(a) input



(b) no regularization



(c) regularization

Summary

Summary

A. X. Chang, T. Funkhouser, L. Guibas, P. Hanrahan, Q. Huang, Z. Li, S. Savarese, M. Savva, S. Song, H. Su, J. Xiao, L. Yi, and F. Yu.
Shapenet : An information-rich 3d model repository, 2015.
cite arxiv :1512.03012.

T. Groueix, M. Fisher, V. G. Kim, B. Russell, and M. Aubry.
AtlasNet : A Papier-Mâché Approach to Learning 3D Surface Generation.
In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2018.

K. He, X. Zhang, S. Ren, and J. Sun.
Deep residual learning for image recognition, 2015.

K. Simonyan and A. Zisserman.
Very deep convolutional networks for large-scale image recognition, 2014.

N. Wang, Y. Zhang, Z. Li, Y. Fu, W. Liu, and Y.-G. Jiang.
Pixel2mesh : Generating 3d mesh models from single rgb images.
In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 52–67, 2018.