THE UNIVERSITY OF TEXAS AT DALLAS

3D Reconstruction

CS 4391 Introduction to Computer Vision
Professor Yapeng Tian
Department of Computer Science

A lot of slides borrowed from Prof. Yu Xiang and Prof. Andreas Geiger

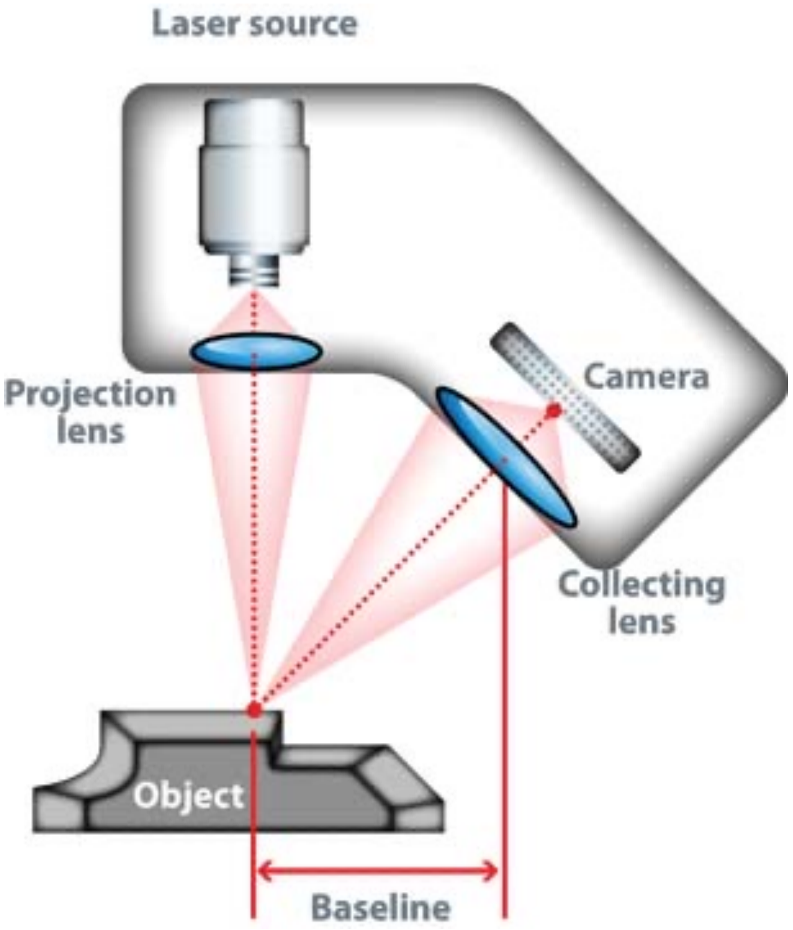
3D Reconstruction

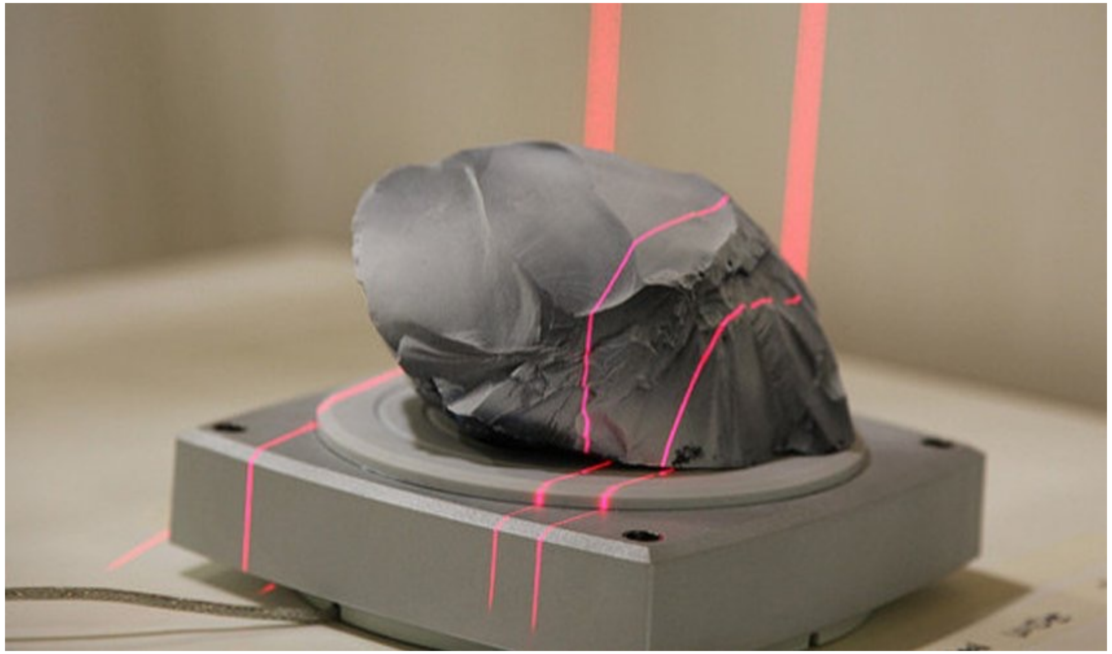
How to obtain 3D models of objects or scenes?

- Stereo matching
- SfM and SLAM
- 3D scanning
- Multi-view stereo
- 3D from a single 2D

3D Reconstruction

Triangulation-based 3D Scanner

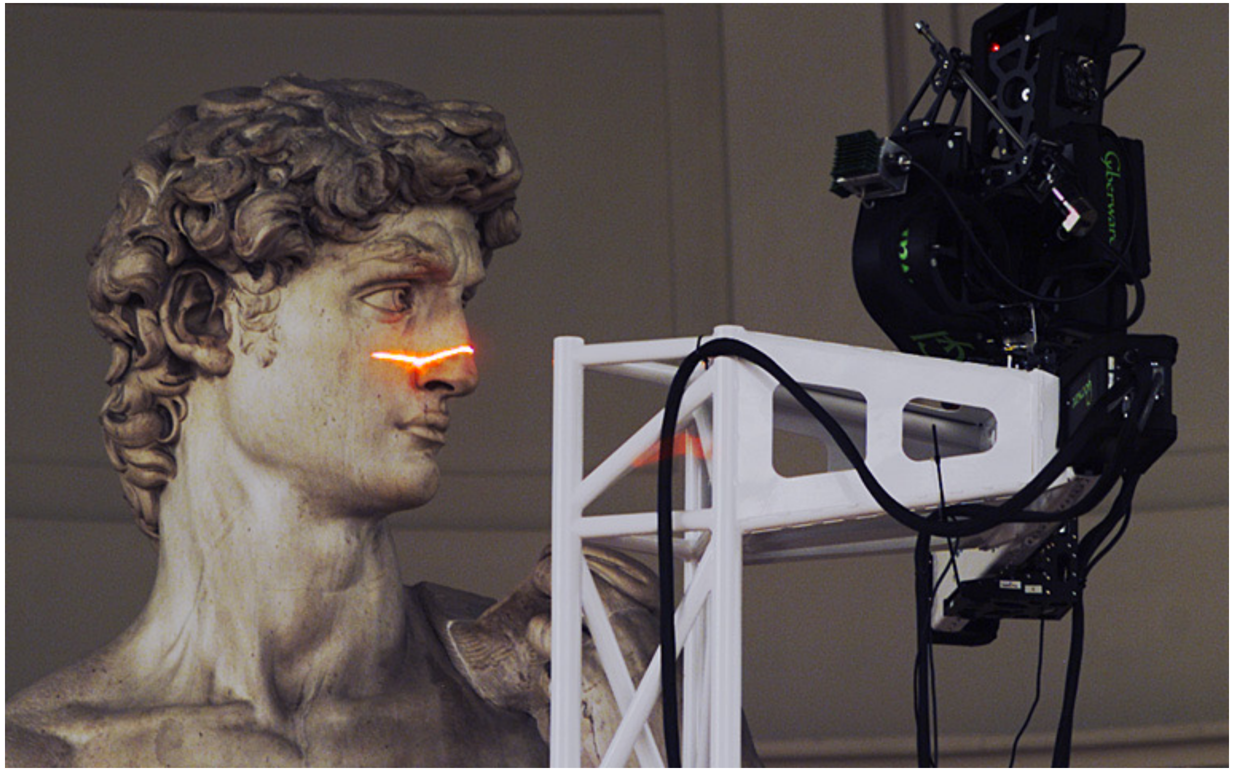




https://3dscanningservices.net/blog/need-know-3d-scanning/

Triangulation-based 3D Scanner

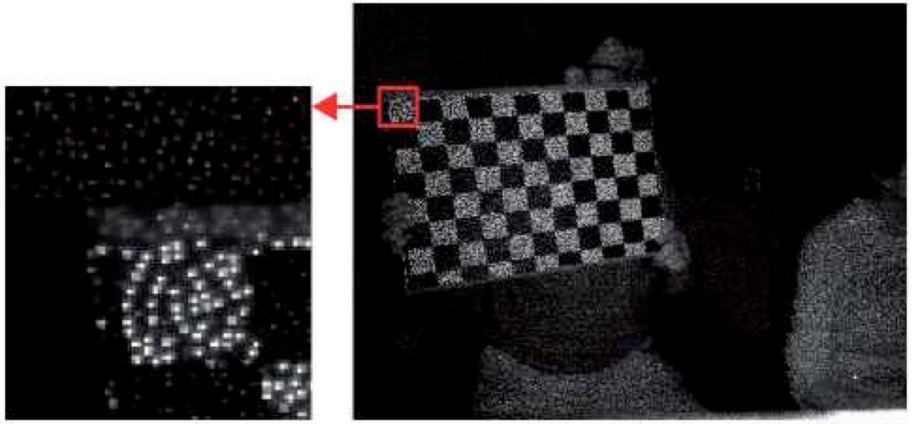
Digital Michelangelo Project (1990)



https://accademia.stanford.edu/mich/

Microsoft Kinect 1

Structured light infrared (IR)



IR stereo

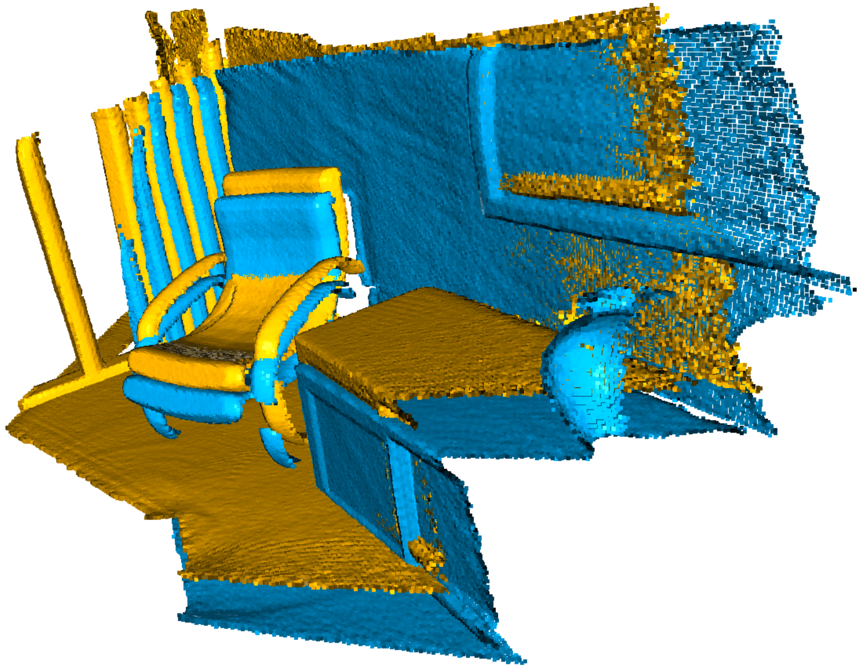
infrared (IR) speckle pattern

Range Data Merging

Each scan/capture generates a depth image or a point cloud

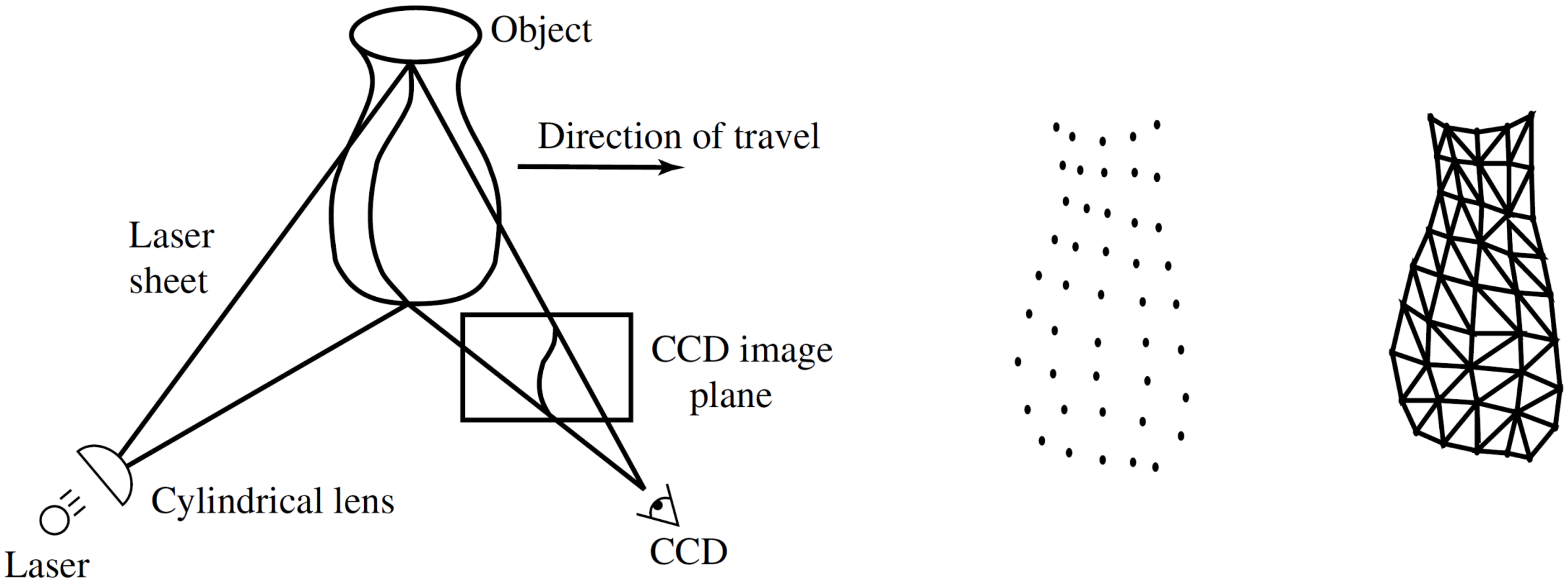
How can we combine these data into a 3D model?

- Alignment/registration
  - E.g., iterative closest point (ICP) algorithm
- Merging



http://www.open3d.org/docs/latest/tutorial/Basic/icp_registration.html

Volumetric Integration



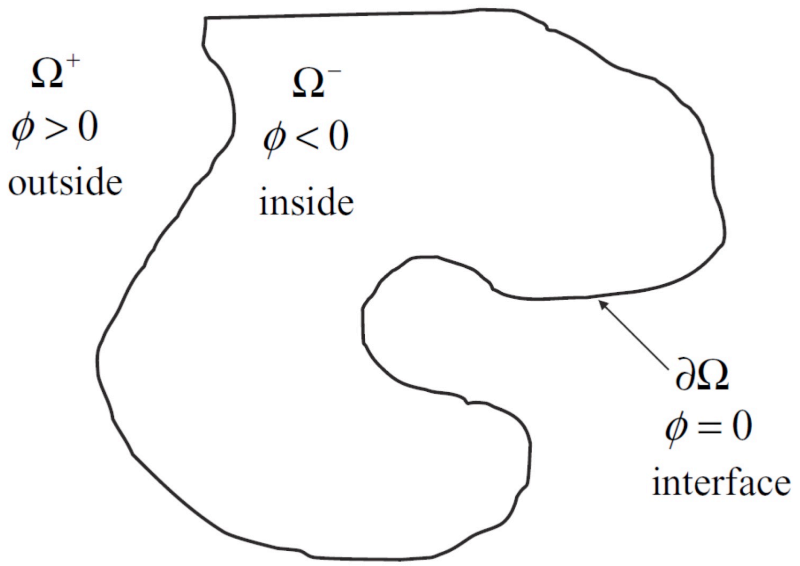
A Volumetric Method for Building Complex Models from Range Images. Curless & Levoy. SIGGRAPH'96.

Volumetric Integration

Signed Distance Function (SDF)

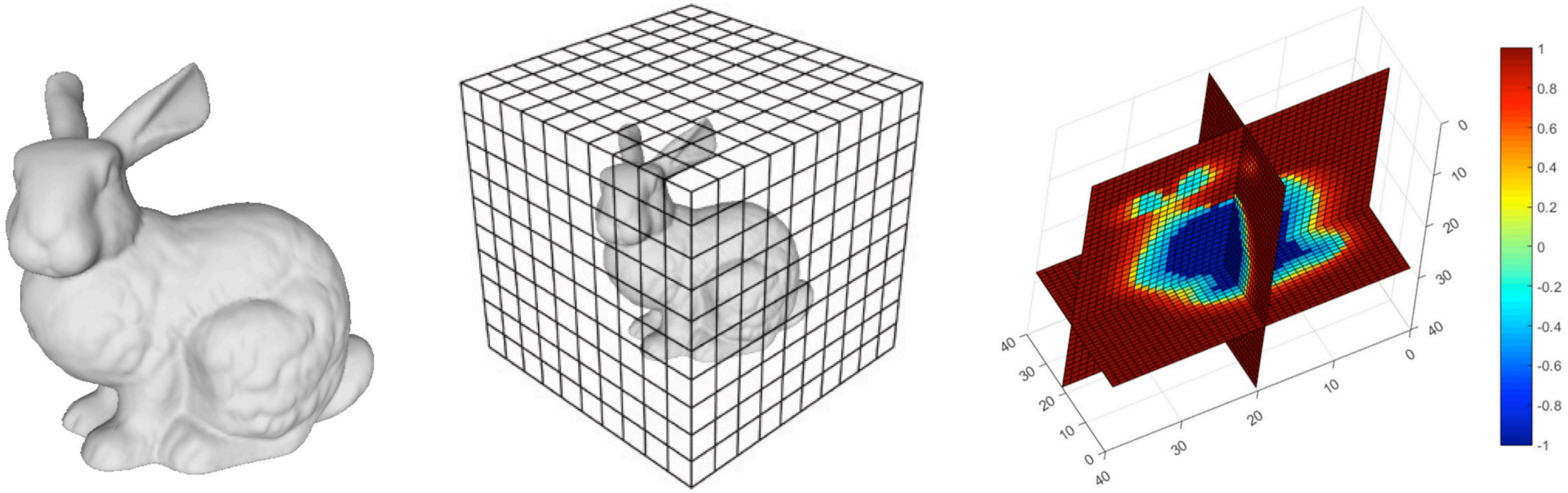
$$\phi : \Omega \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}$$

Signed distance to the closest object boundary



Volumetric Integration

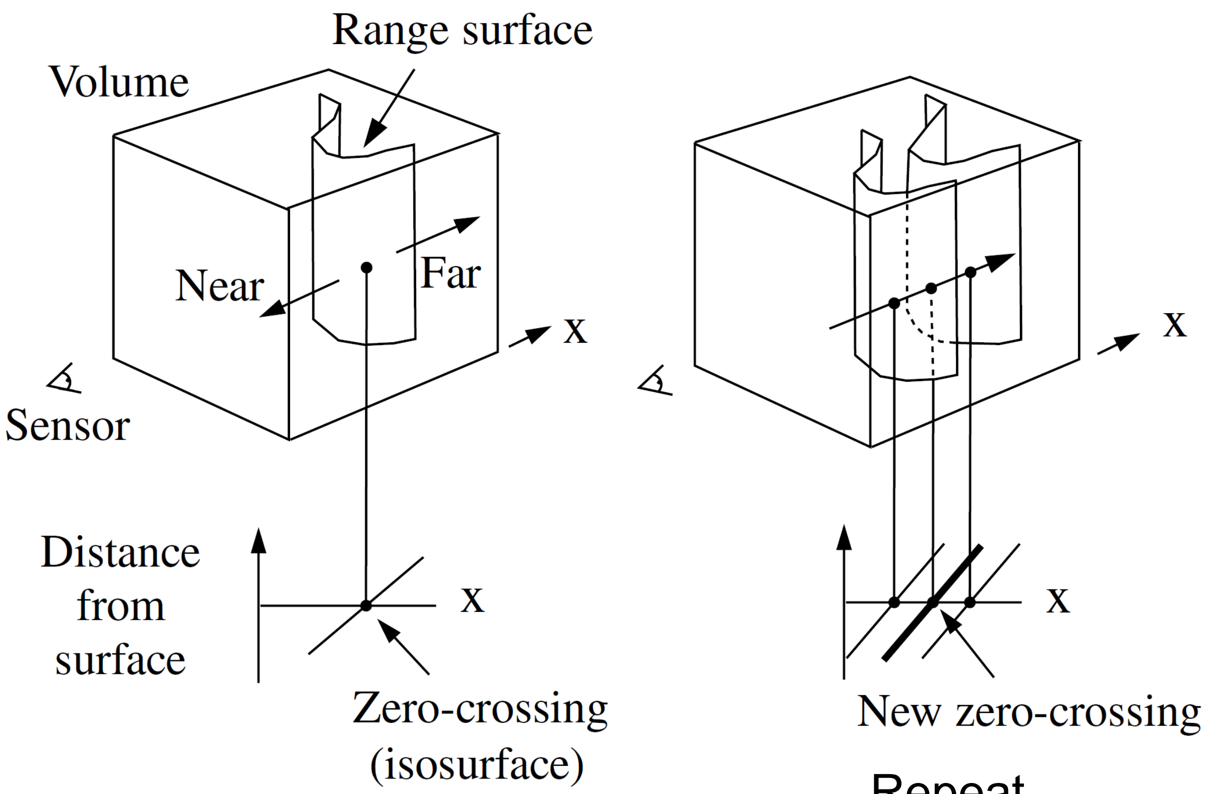
Signed Distance Function (SDF)



(a) Surface view. (b) Bounding volume. (c) Generated SDF.

Signed Distance Fields for Rigid and Deformable 3D Reconstruction. Miroslava Slavcheva.

Volumetric Integration



Repeat

SDF for the range image

$$D_{i+1}(\mathbf{x}) = \frac{W_i(\mathbf{x})D_i(\mathbf{x}) + w_{i+1}(\mathbf{x})d_{i+1}(\mathbf{x})}{W_i(\mathbf{x}) + w_{i+1}(\mathbf{x})}$$

$$W_{i+1}(\mathbf{x}) = W_i(\mathbf{x}) + w_{i+1}(\mathbf{x})$$

Weight function

A Volumetric Method for Building Complex Models from Range Images. Curless & Levoy. SIGGRAPH'96.

Volumetric Integration

Image Single scan Merged scan

A Volumetric Method for Building Complex Models from Range Images. Curless & Levoy. SIGGRAPH’96.

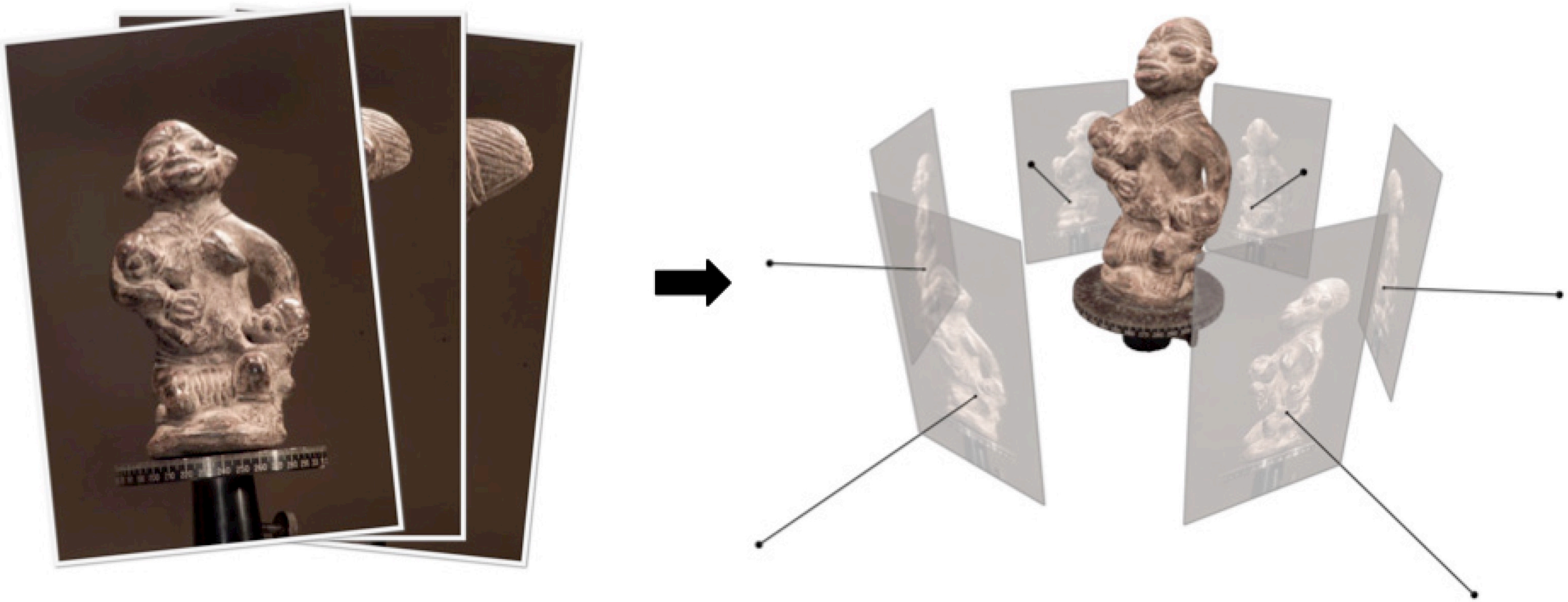
KinectFusion

Single scan

Rendered normal map

Rendered 3D model

Image-based 3D Reconstruction

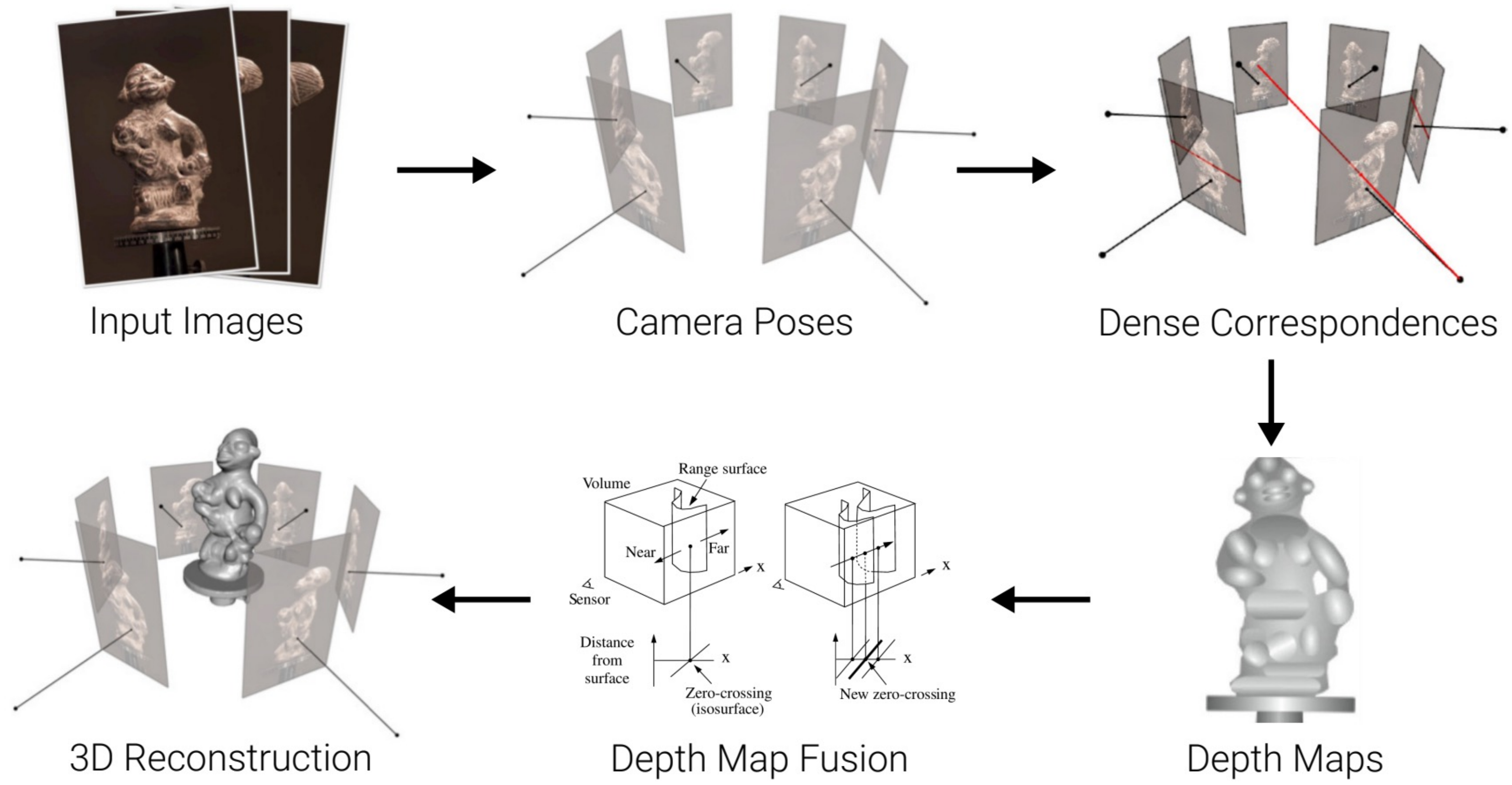


A set of images

3D model

Multi-View Stereo: A Tutorial. Yasutaka Furukawa and Carlos Hernández

Image-based 3D Reconstruction Pipeline

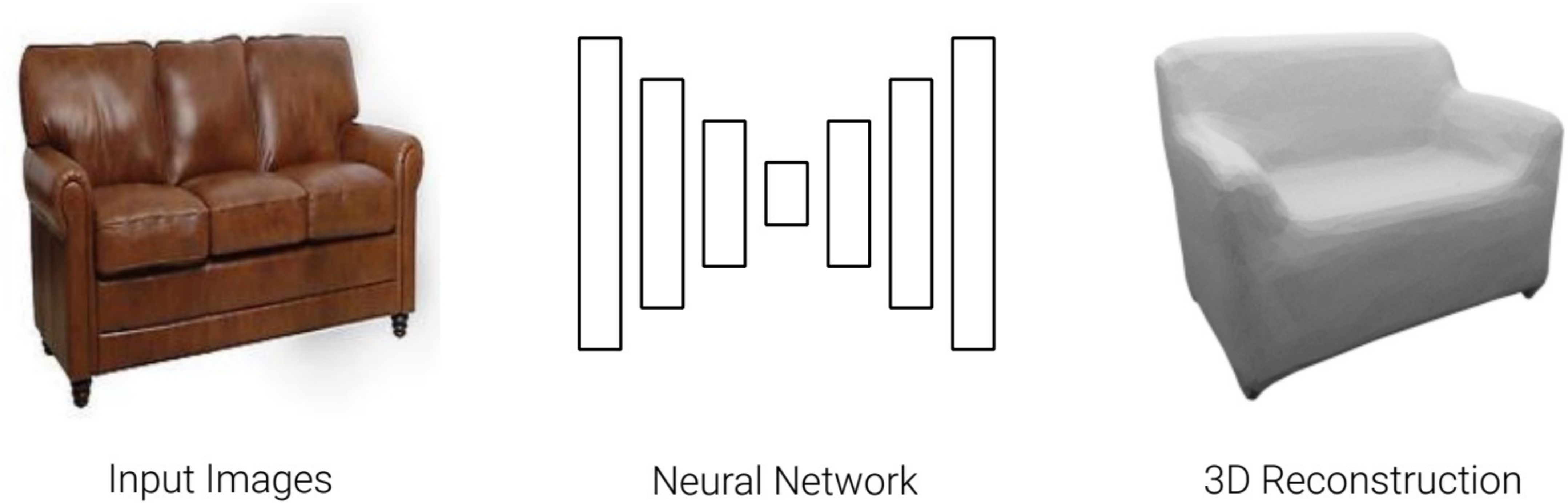


Humans recognize 3D from a **single** 2D image


BELIEVE

Can we learn to infer 3D **from a 2D image**?

3D Reconstruction from a 2D Image

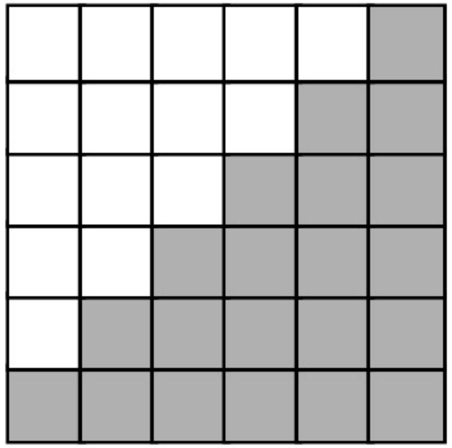


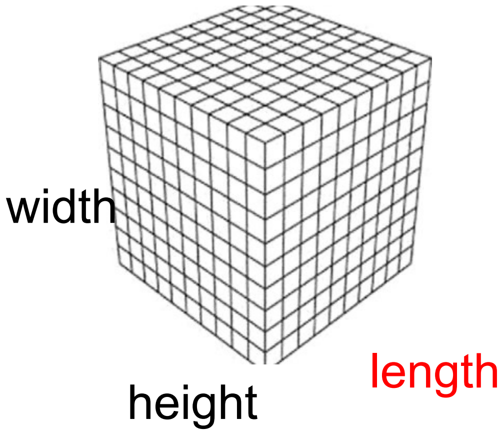
What is a good **output** 3D representation?

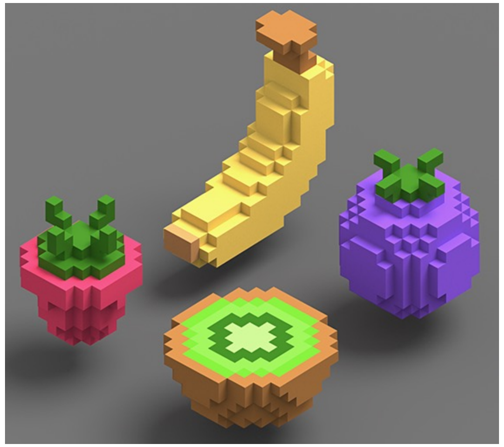
3D Representations

Voxels:

- Discretization of 3D space into grid
- Easy to process with neural networks
- Cubic memory $O(n^3)$ ⇒ limited resolution





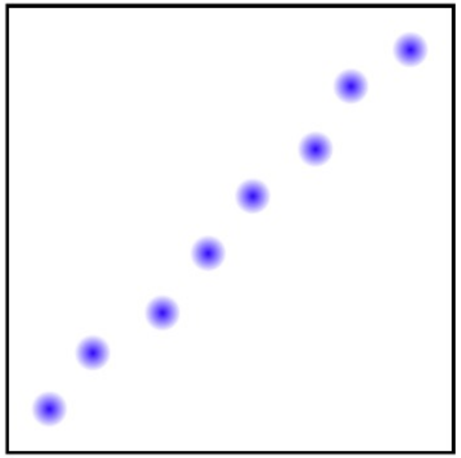


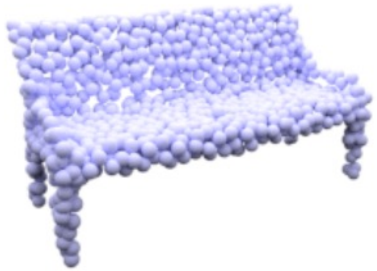
[Maturana et al., IROS 2015]

3D Representations

Points

- Discretization of surface into 3D points
- Does not model connectivity / topology
- Limited number of points
- Global shape description



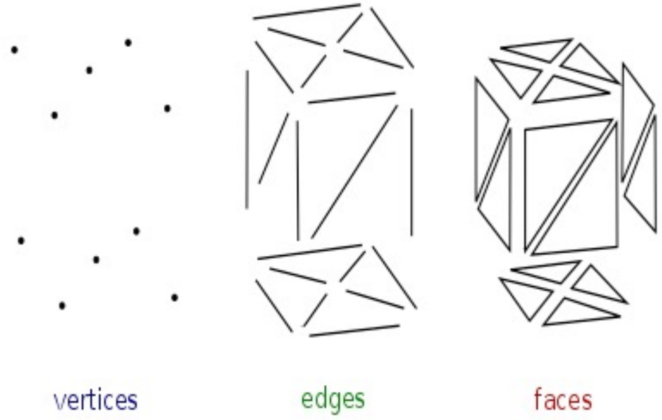


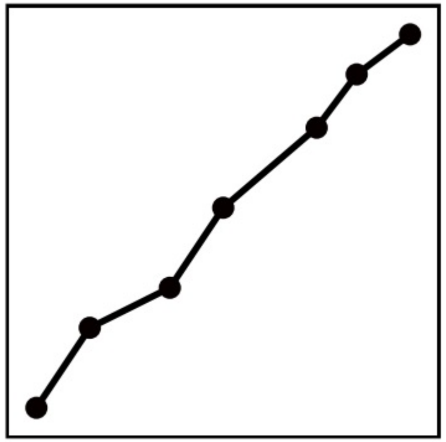
[Fan et al., CVPR 2017]

3D Representations

Meshes

- Discretization into vertices and faces
- Limited number of vertices / granularity
- Requires class-specific template – or –
- Leads to self-intersections





[Groueix et al., CVPR 2018]

3D Representations

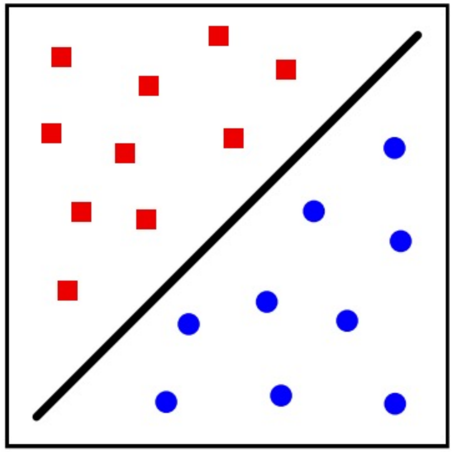
Implicit 3D representation

- Implicit representation $\Rightarrow$ No discretization
- Arbitrary topology & resolution
- Low memory footprint
- Not restricted to specific class

Occupancy Network for 3D Reconstruction

Key idea

- Do not represent 3D shape explicitly
- Instead, consider surface **implicitly** as **decision boundary** of a non-linear classifier:

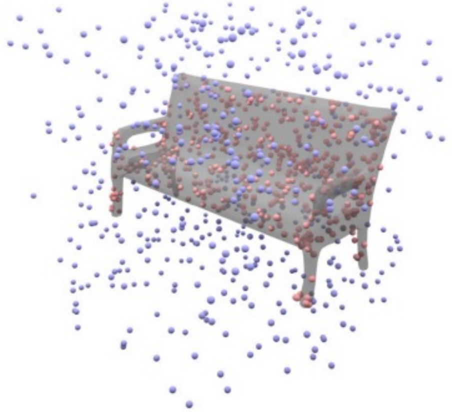


$$f_\theta : \mathbb{R}^3 \times \mathcal{X} \to [0, 1]$$

3D location

Input for 3D reconstruction: image/point cloud

Occupancy probability



Occupancy Networks: Learning 3D Reconstruction in Function Space. Mescheder et al., CVPR’19

Occupancy Network for 3D Reconstruction

Training

Binary cross-entropy loss

$$\mathcal{L}_{\mathcal{B}}(\theta) = \frac{1}{|\mathcal{B}|} \sum_{i=1}^{|\mathcal{B}|} \sum_{j=1}^{K} \mathcal{L}(f_\theta(p_{ij}, x_i), o_{ij})$$

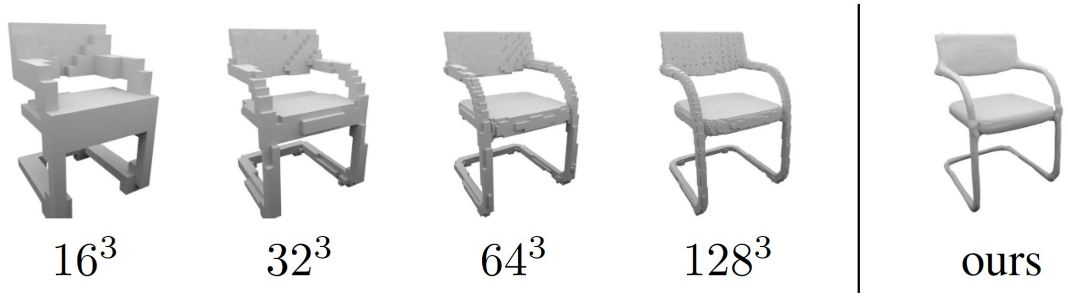
3D point j for image i

image i

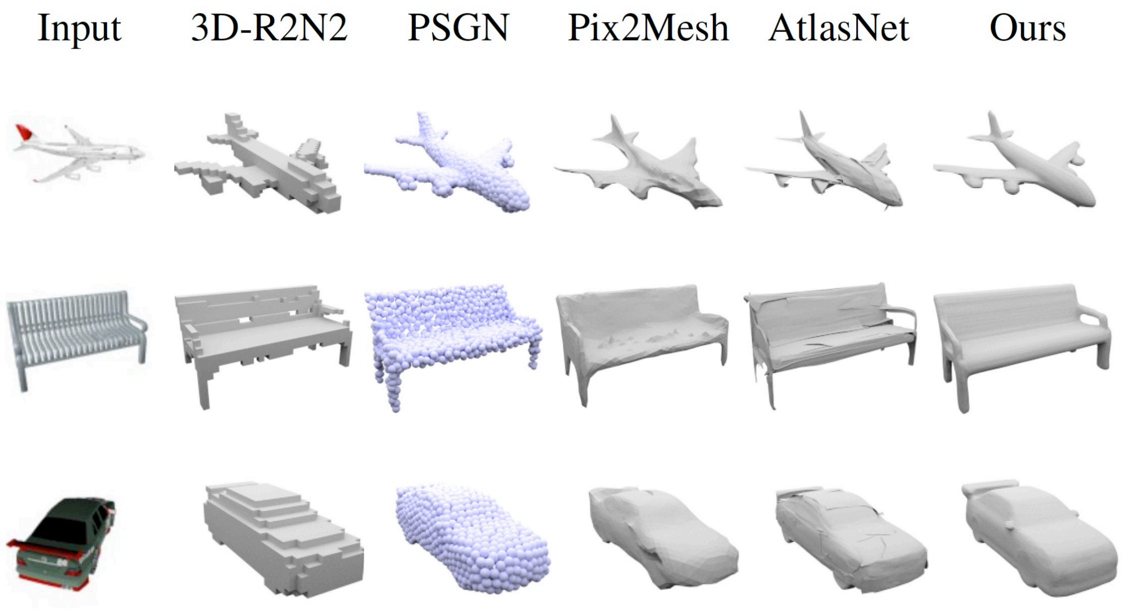
Ground truth occupancy

Occupancy Networks: Learning 3D Reconstruction in Function Space. Mescheder et al., CVPR'19

Occupancy Network for 3D Reconstruction



Continuous shape representation



Single image 3D reconstruction

Occupancy Networks: Learning 3D Reconstruction in Function Space. Mescheder et al., CVPR'19

Occupancy Networks: Learning 3D Reconstruction in Function Space. Mescheder et al., CVPR'19 [video link]

Summary

- 3D scanning and Multiview stereo pipeline

- Explicit 3D representations
  - Voxels, points, meshes

- Implicit 3D representations
  - Learn a function to represent the 3D shape (occupancy, SDFs, radiance fields)

Further Reading

- Chapter 13, Computer Vision, Richard Szeliski
- A Volumetric Method for Building Complex Models from Range Images. Curless & Levoy. SIGGRAPH'96.
- Multi-View Stereo: A Tutorial. Yasutaka Furukawa and Carlos Hernández, 2015
- Occupancy Network https://arxiv.org/abs/1812.03828
- DeepSDF https://arxiv.org/abs/1901.05103

Project Presentation

Presentation (Slides)

- Introduction: Project title, group members, problem overview (1 min)
- Method: your approach (2 mins)
- Results: your data and experimental results to showcase your method (2 mins)
- QA (1 min)

Each group has 6 minutes for the presentation and questions

- Please use slides to present your work
- Show a demo of the project if you have one
- All group members should show up

Evaluation criteria

- The grading will be based on the overall quality of the presentation in terms of content, clarity, and question answering

Presentation Order and Submission

- **The presentation order was randomly generated**
  - Set 1 (Wednesday 11/29): 8, 1, 23, 25, 13, 9, 16, 24, 15
  - Set 2 (Monday 12/04): 18, 19, 27, 11, 7, 14, 12, 17, 10
  - Set 3 (Wednesday 12/06): 3, 6, 20, 2, 21, 22, 26, 5, 4

- Please submit the following items to eLearning. You can zip all the files. I will download your submission a day prior to your presentation. To save time and prevent potential technical issues, you will use my computer for the presentation
  - (Required) Presentation slides in pdf/pptx format
  - (Optional) A demo video in mp4 format if you have one