3D ShapeNets: A Deep Representation for Volumetric Shape Modeling

by Wu, Song, Khosla, Yu, Zhang, Tang, Xiao

presented by Abhishek Sinha

3D Shape Prior

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D Shape Prior

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D Shape Prior

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Outline

Outline

- Problem

Outline

- Problem
- Motivation

Outline

- Problem
- Motivation
- Desirable Properties for Representation

Outline

- Problem
- Motivation
- Desirable Properties for Representation
- Architecture

Outline

- Problem
- Motivation
- Desirable Properties for Representation
- Architecture
- Dataset

Outline

- Problem
- Motivation
- Desirable Properties for Representation
- Architecture
- Dataset
- Applications

Outline

- Problem
- Motivation
- Desirable Properties for Representation
- Architecture
- Dataset
- Applications
- Extensions

Outline

- Problem
- Motivation
- Desirable Properties for Representation
- Architecture
- Dataset
- Applications
- Extensions
- Discussion Points

Problem

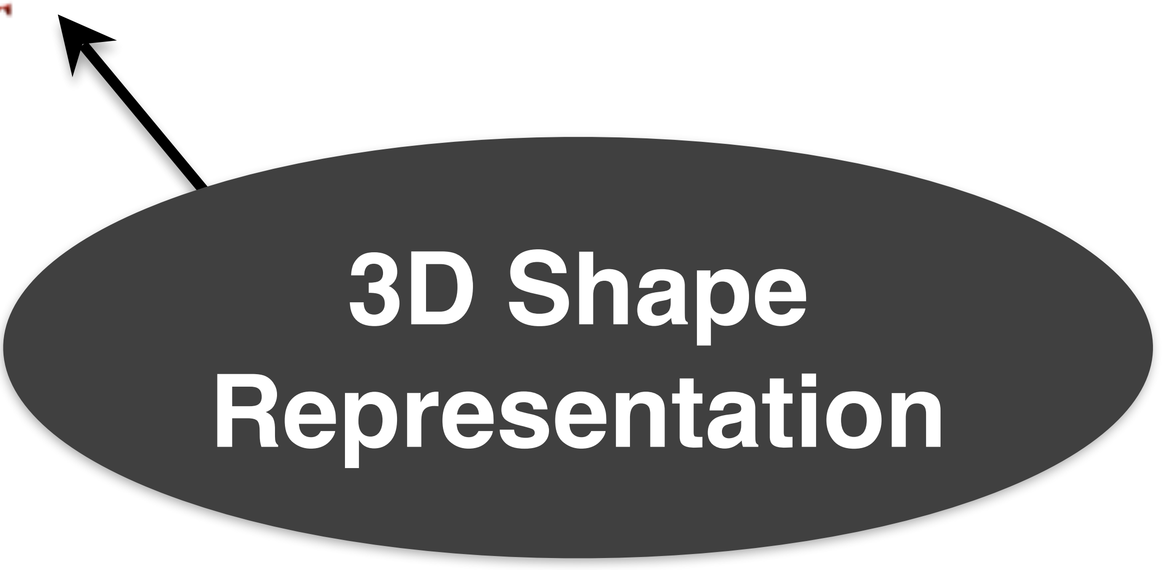
Learn ‘Useful’ 3D shape representations from images

Motivation

3D Shape Representation

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Shape Synthesis

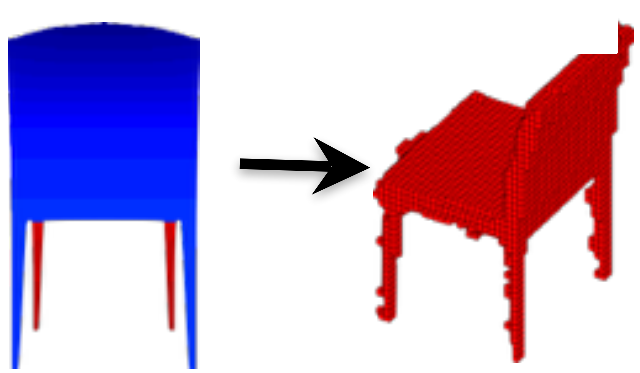


Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

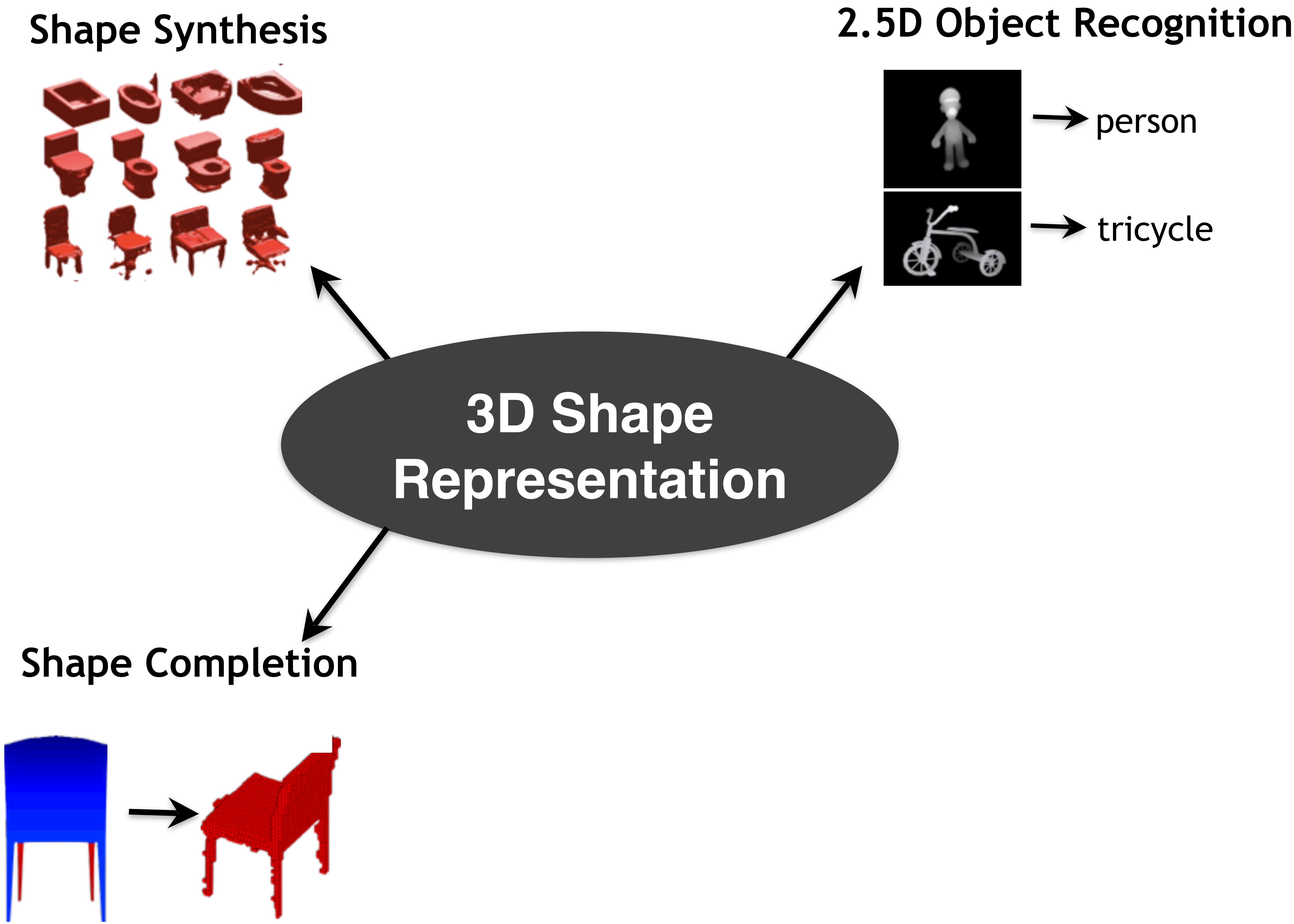
Shape Synthesis

3D Shape Representation

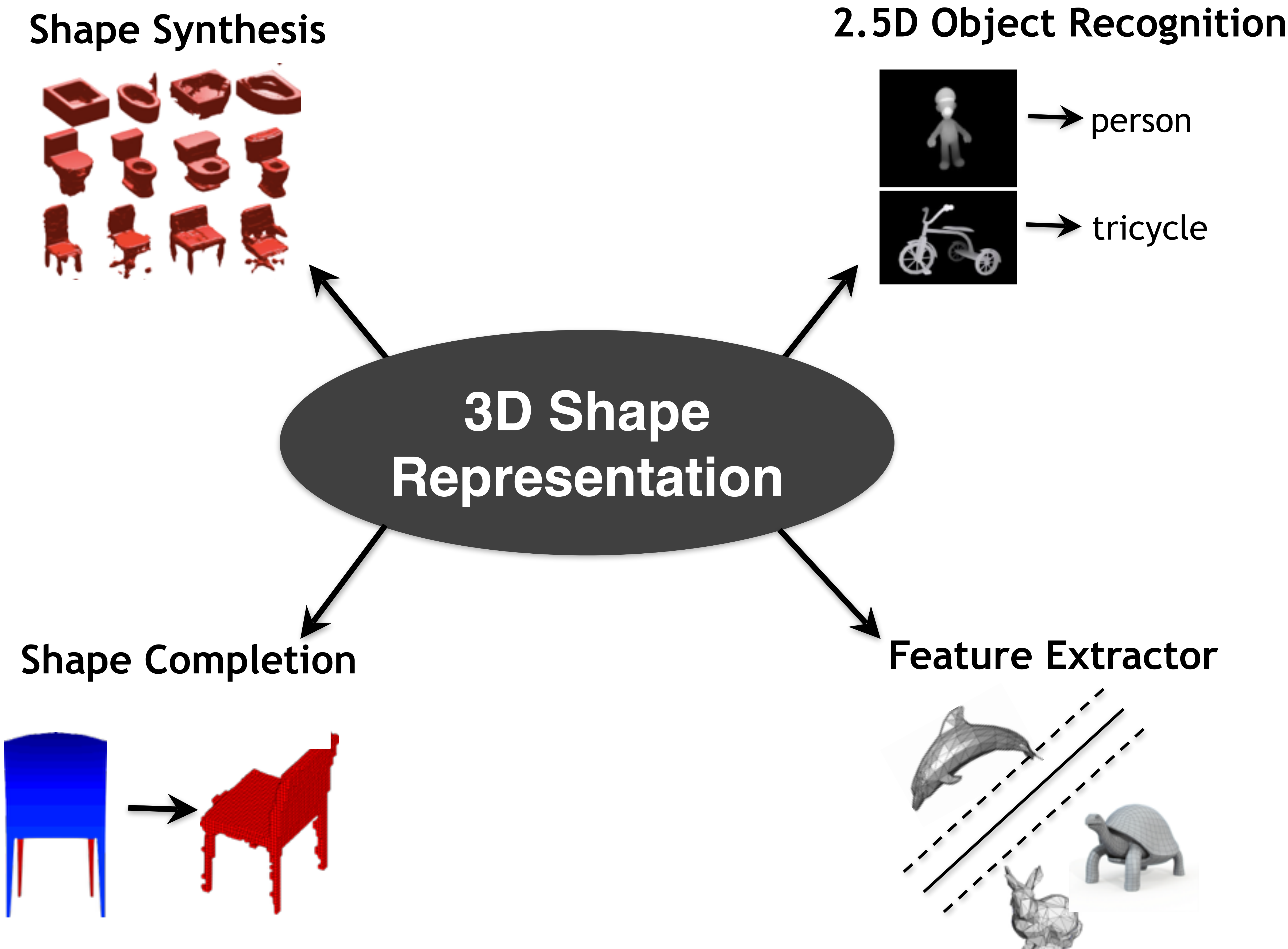
Shape Completion



Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015



Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015



Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Desirable Properties

What is a Desirable
3D Shape Representation?

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

What is a Desirable 3D Shape Representation?

Data-driven

Generic

Compositional

Versatile

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

What is a Desirable 3D Shape Representation?

Data-driven

Data-driven

Compositional

Versatile

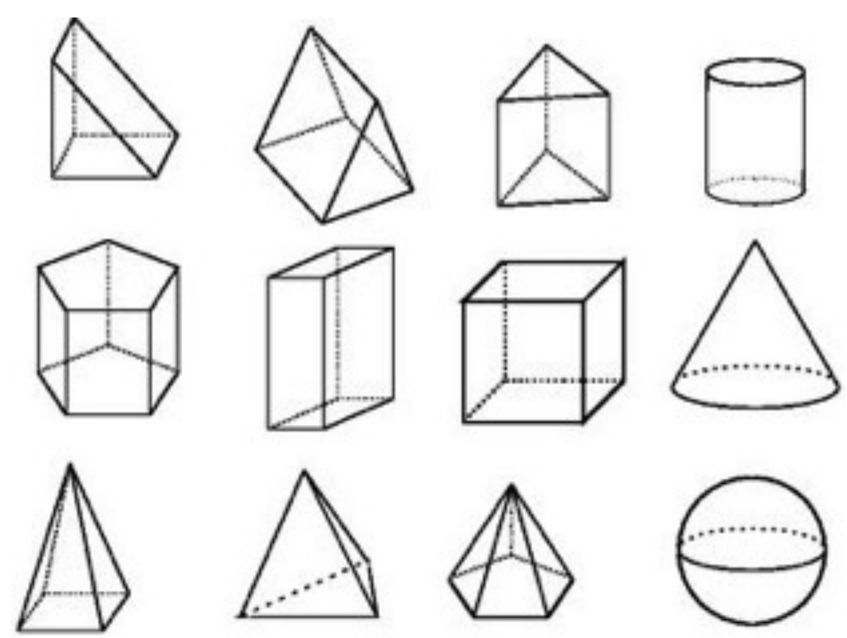
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

What is a Desirable 3D Shape Representation?

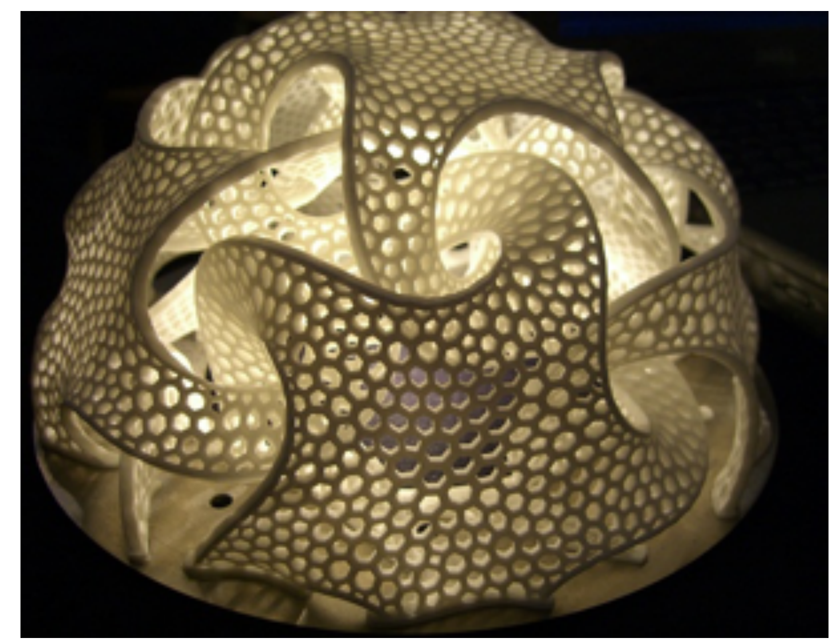
Data-driven

Generic

Versatile



Simple Shapes



Complex Shapes

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

What is a Desirable 3D Shape Representation?

Data-driven

Generic

Compositional

Compositional

building blocks

full object

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

What is a Desirable 3D Shape Representation?

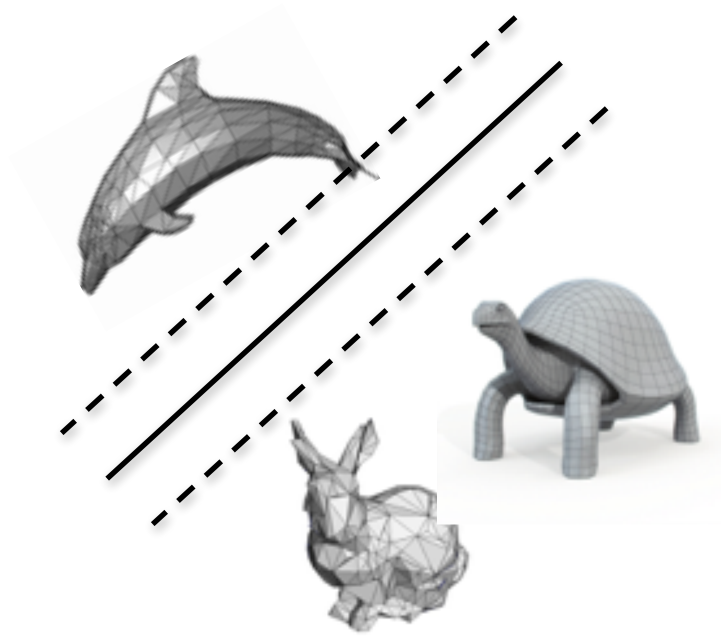
Data-driven

Generic

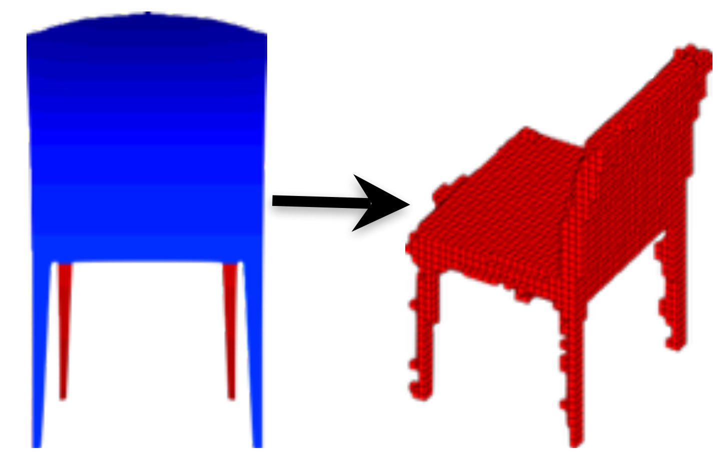
Compositional

Versatile

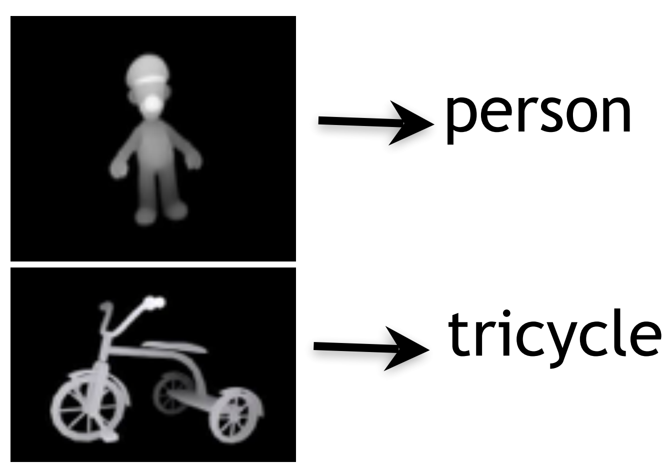
Versatile



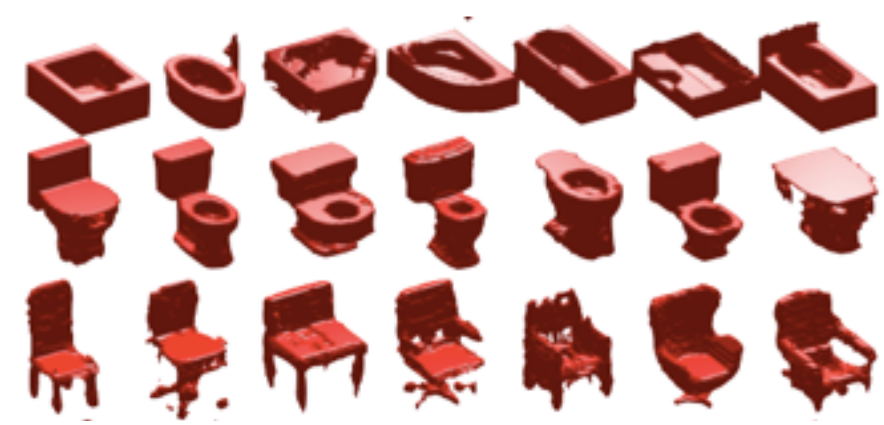
mesh classification



shape completion



2.5D object recognition



shape generation

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Architecture

3D Deep Learning

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D Deep Learning

3D Shape as Volumetric Representation

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D Deep Learning

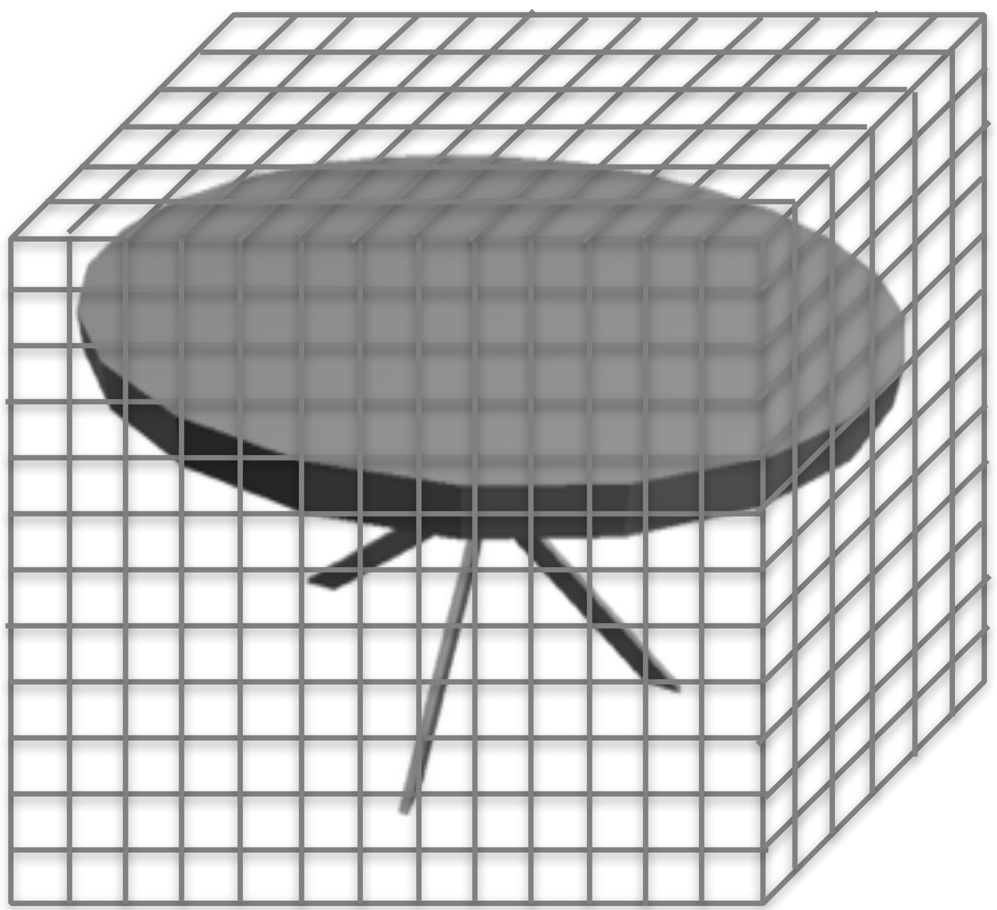
3D Shape as Volumetric Representation

mesh

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D Deep Learning

3D Shape as Volumetric Representation

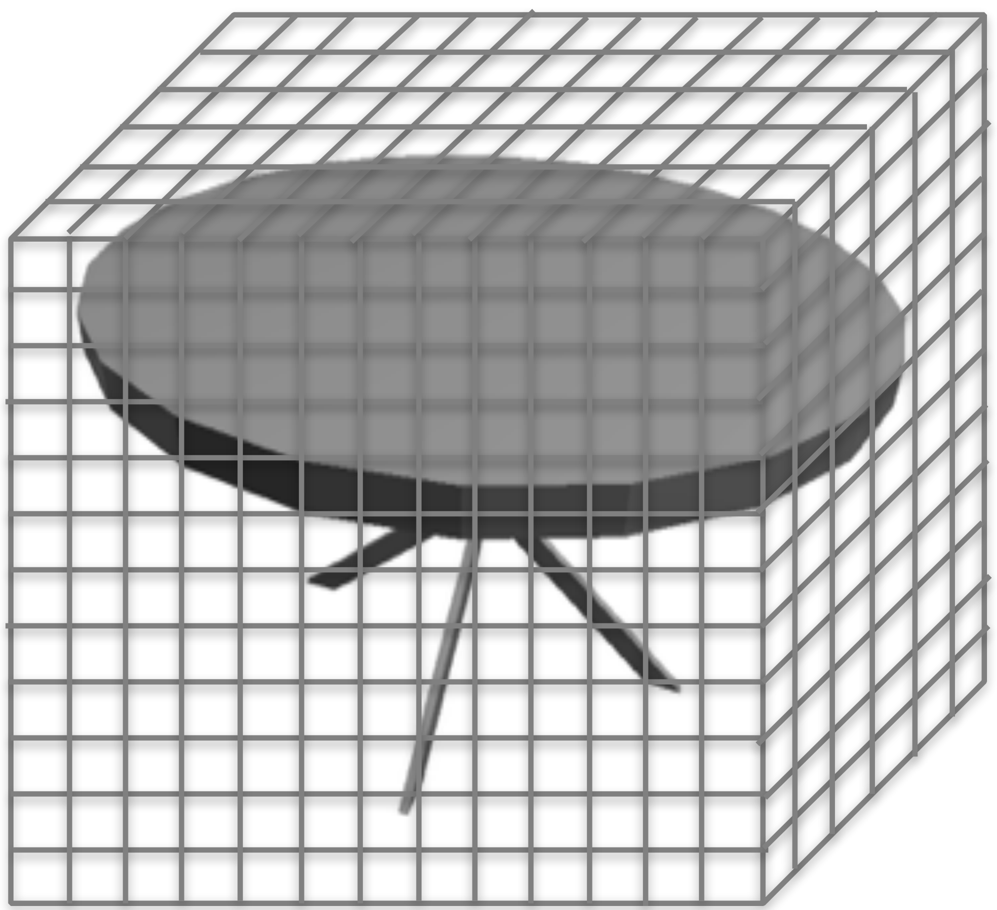


mesh

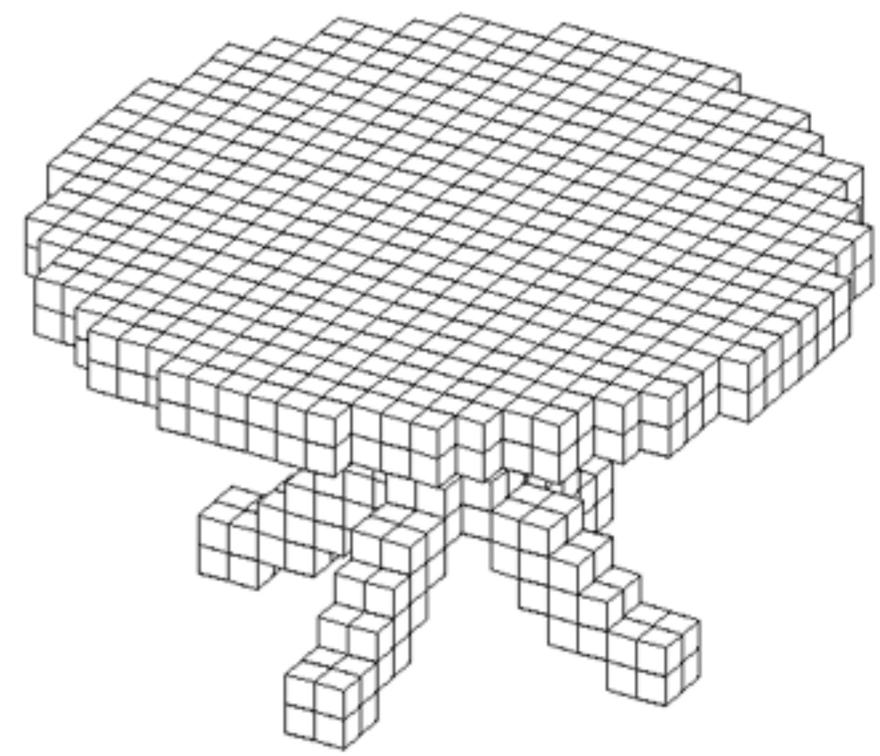
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D Deep Learning

3D Shape as Volumetric Representation



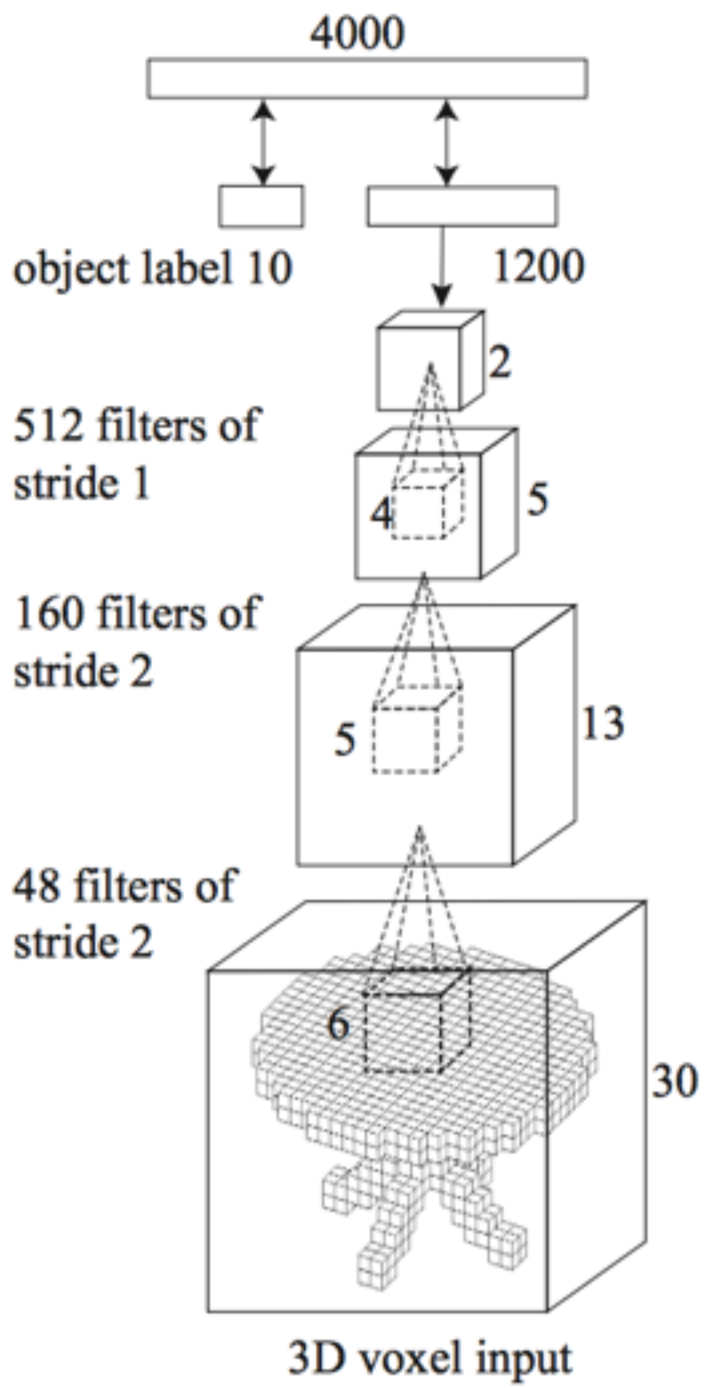
mesh



binary voxel

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

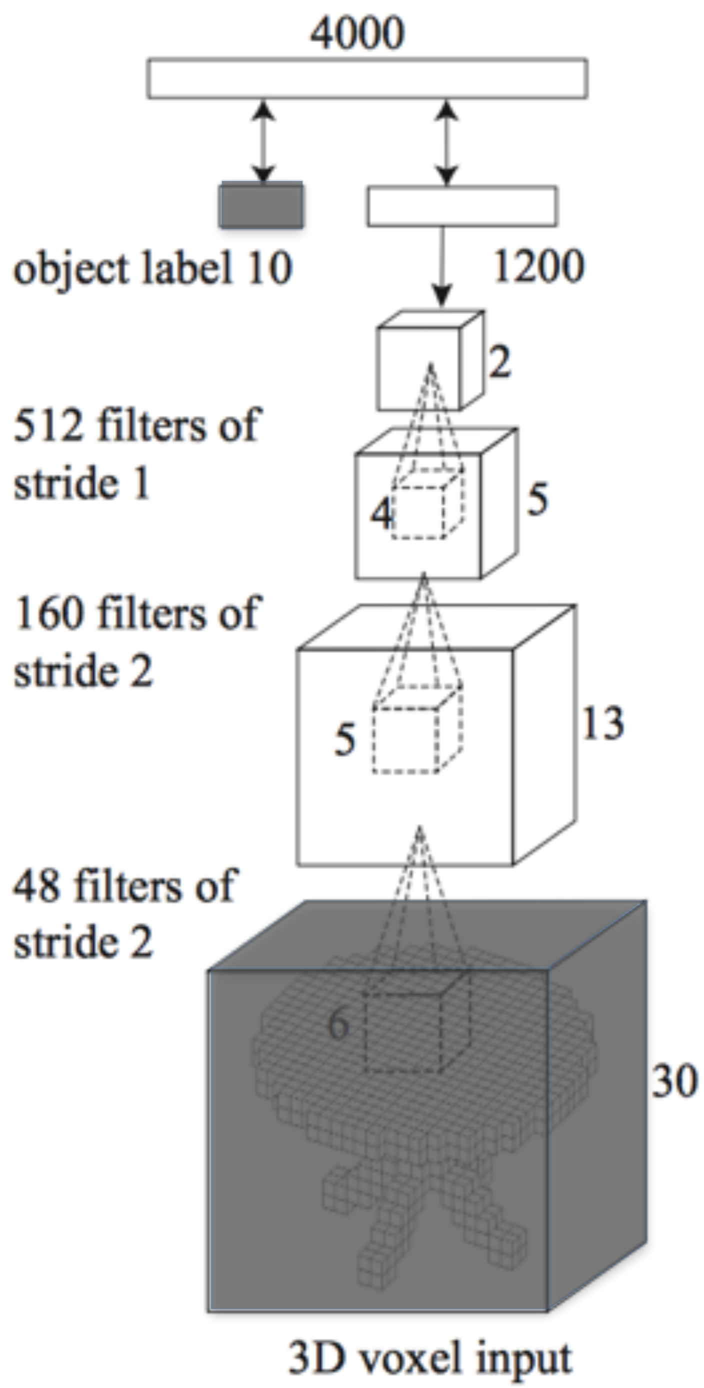
3D ShapeNets



Convolutional Deep Belief Network $p(\mathbf{x}, y)$

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D ShapeNets

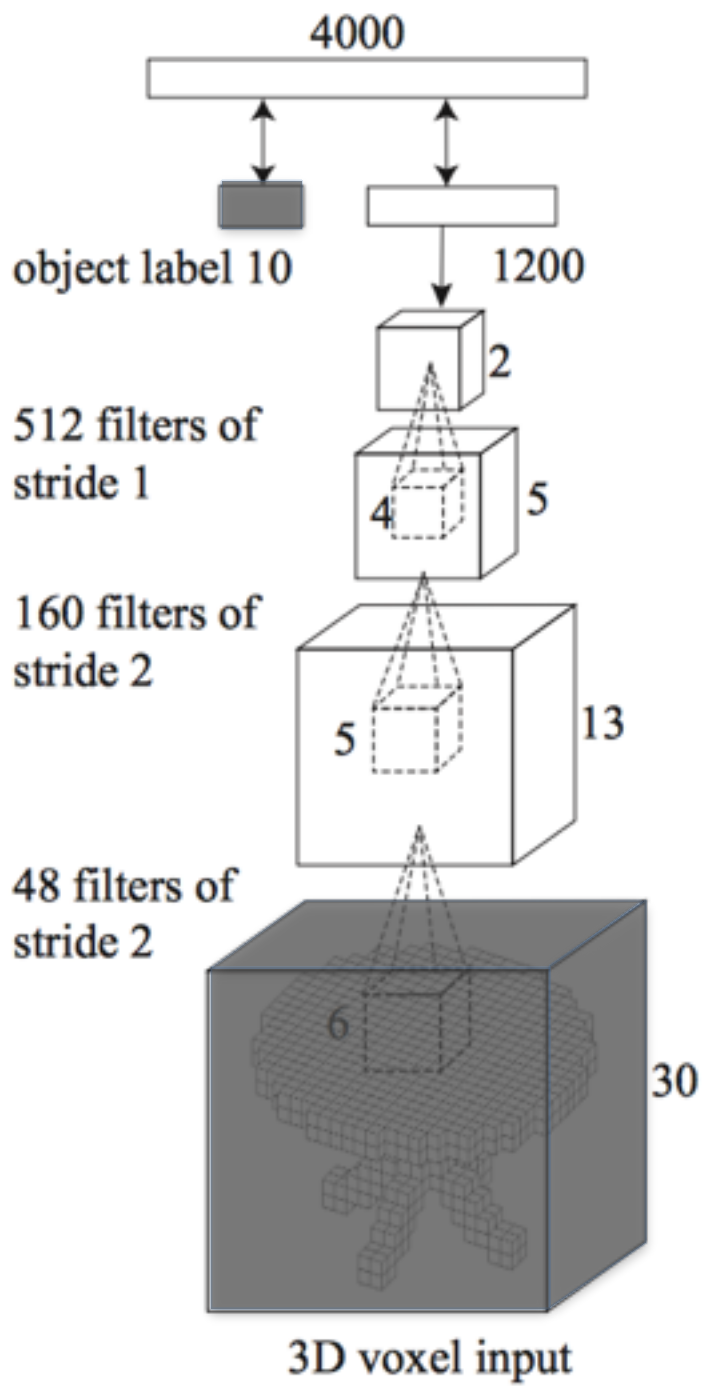


Convolutional Deep Belief Network $p(\mathbf{x}, y)$

A **Deep Belief Network** is a generative graphical model that describes the distribution of input x over class y.

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D ShapeNets



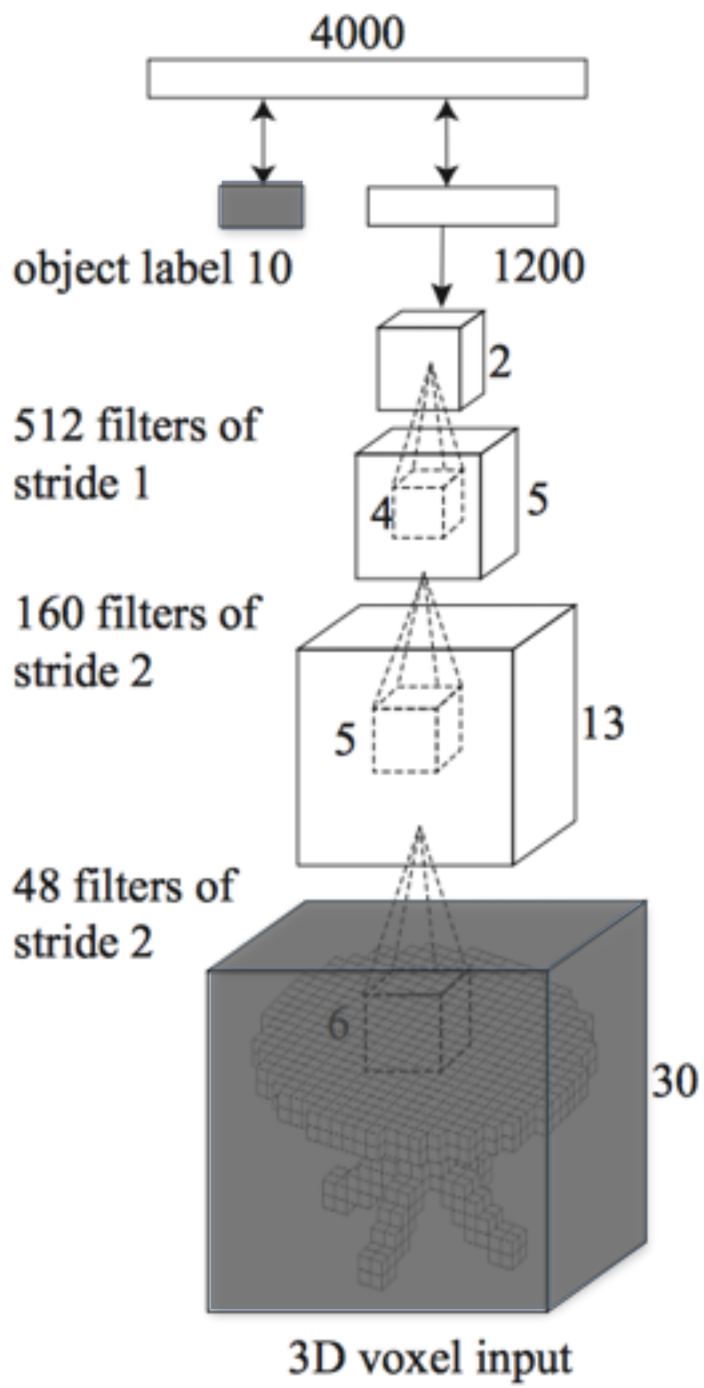
Convolutional Deep Belief Network $p(\mathbf{x}, y)$

A **Deep Belief Network** is a generative graphical model that describes the distribution of input x over class y.

- Convolution to enable compositionality
- No pooling to reduce reconstruction error

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D ShapeNets



Convolutional Deep Belief Network $p(\mathbf{x}, y)$

A **Deep Belief Network** is a generative graphical model that describes the distribution of input x over class y.

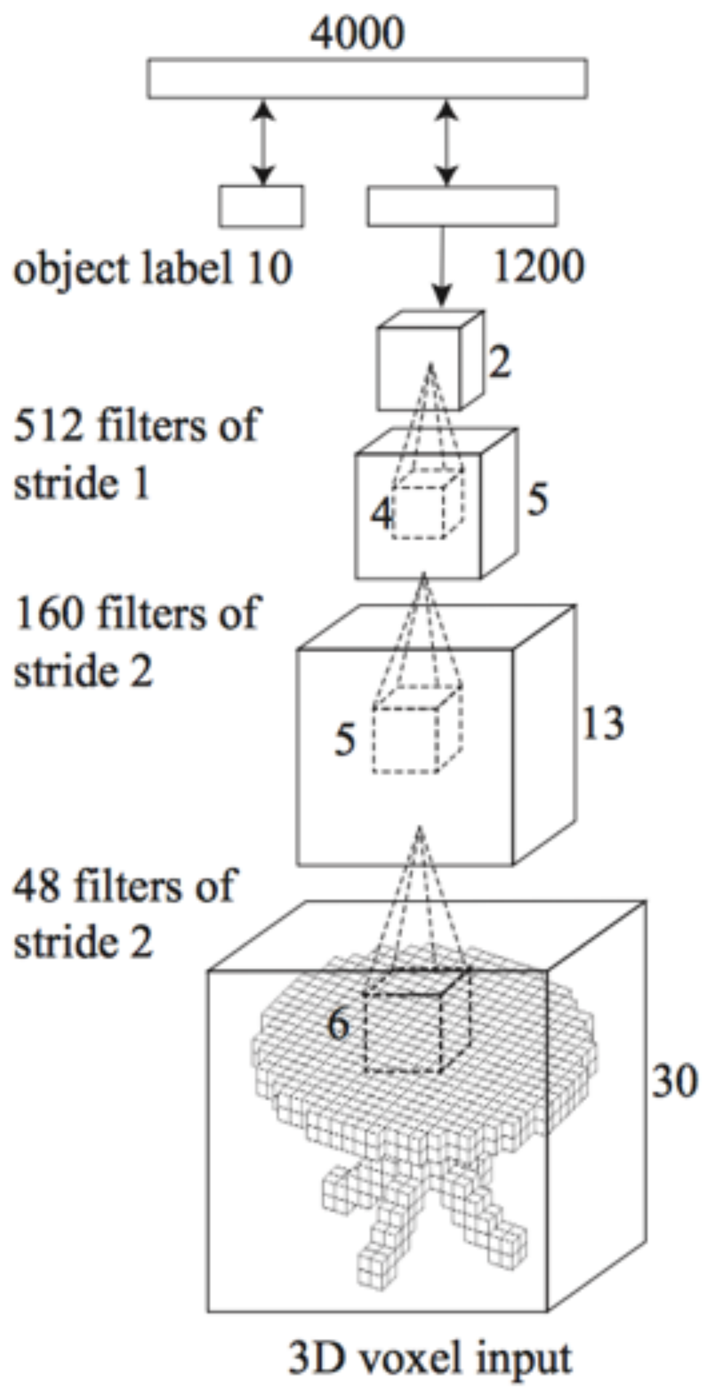
- Convolution to enable compositionality
- No pooling to reduce reconstruction error

configurations

Layer 1-3	convolutional RBM
Layer 4	fully connected RBM
Layer 5	multinomial label + Bernoulli feature form an associate memory

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

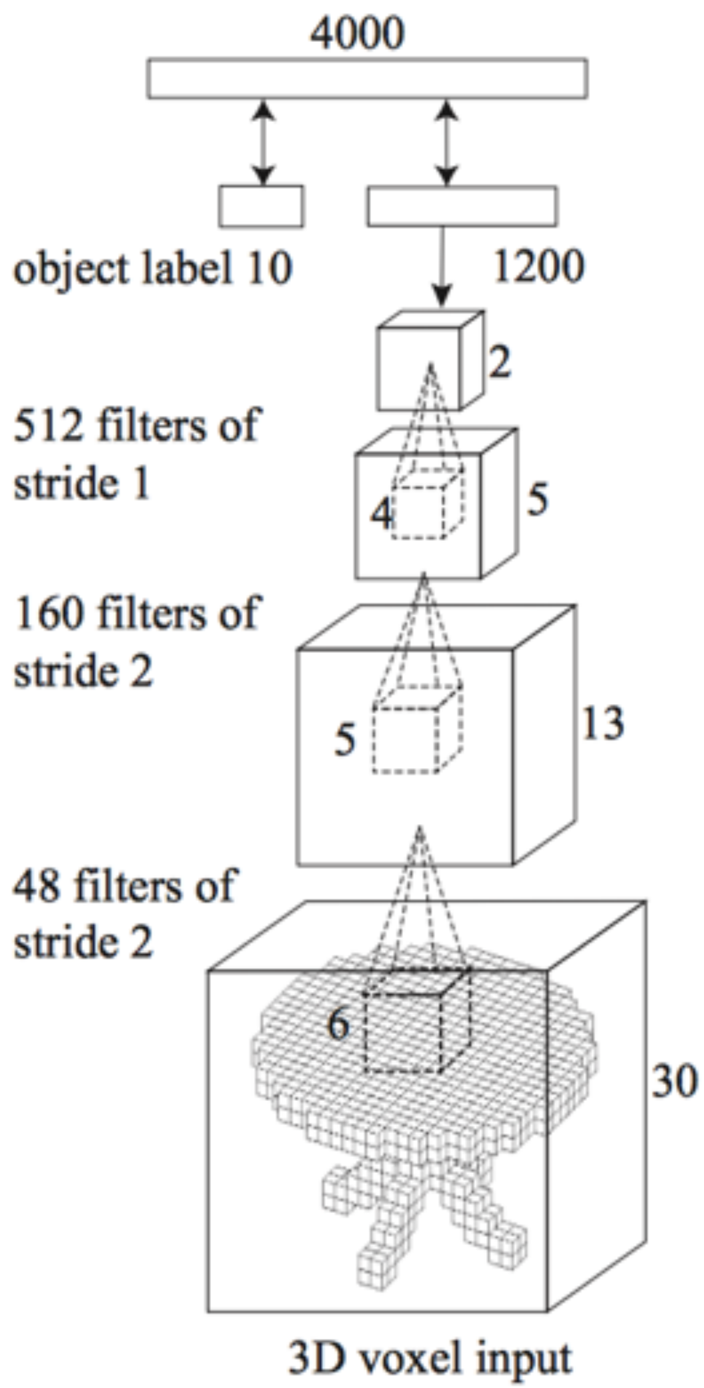
3D ShapeNets



Convolutional Deep Belief Network $p(\mathbf{x}, y)$

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D ShapeNets

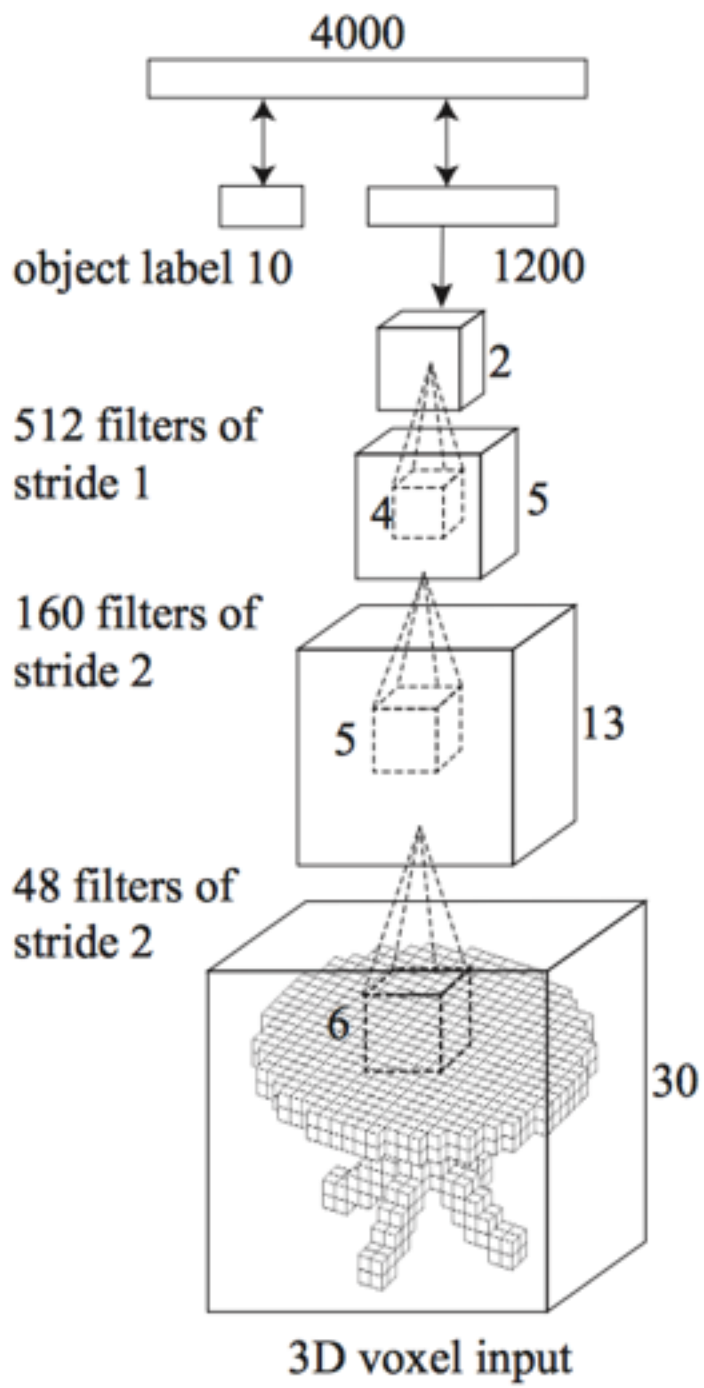


Convolutional Deep Belief Network $p(\mathbf{x}, y)$

3D ShapeNets ≠ CNNs

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D ShapeNets



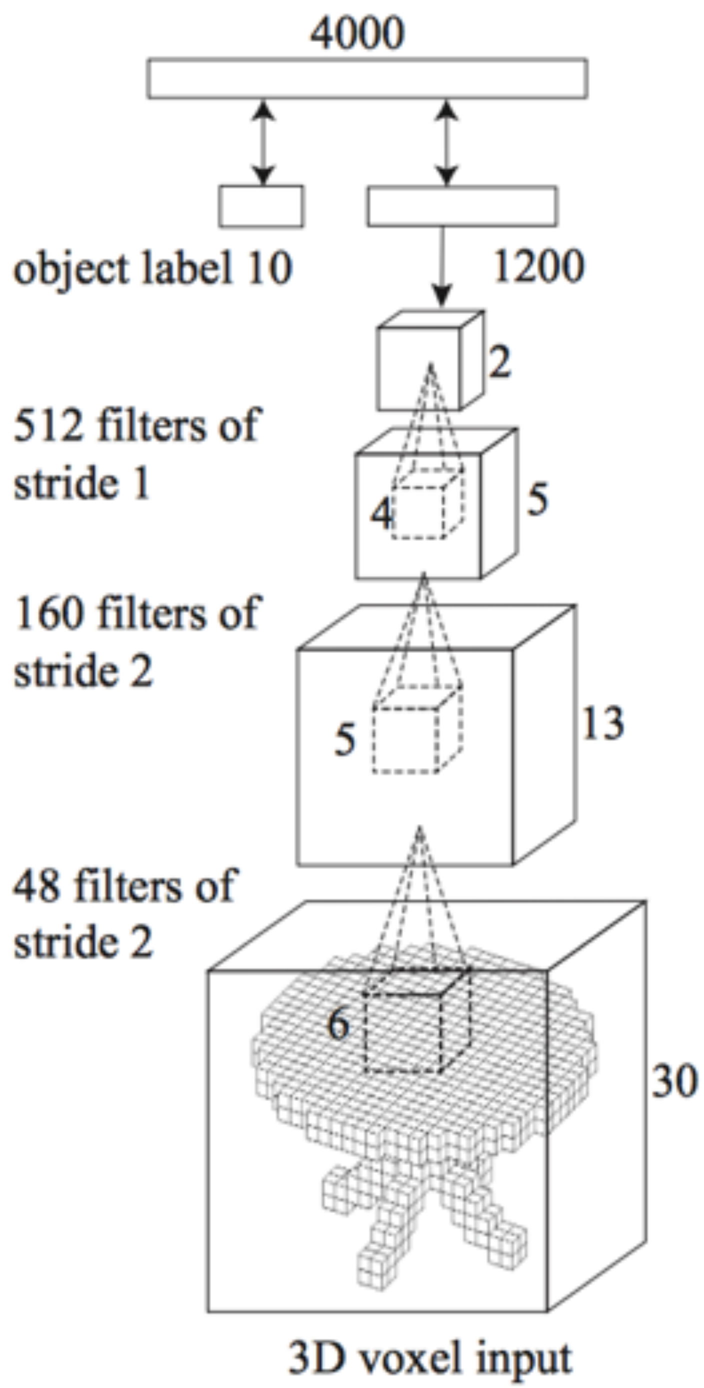
Convolutional Deep Belief Network $p(\mathbf{x}, y)$

3D ShapeNets ≠ CNNs

$$p(y|x)$$

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D ShapeNets



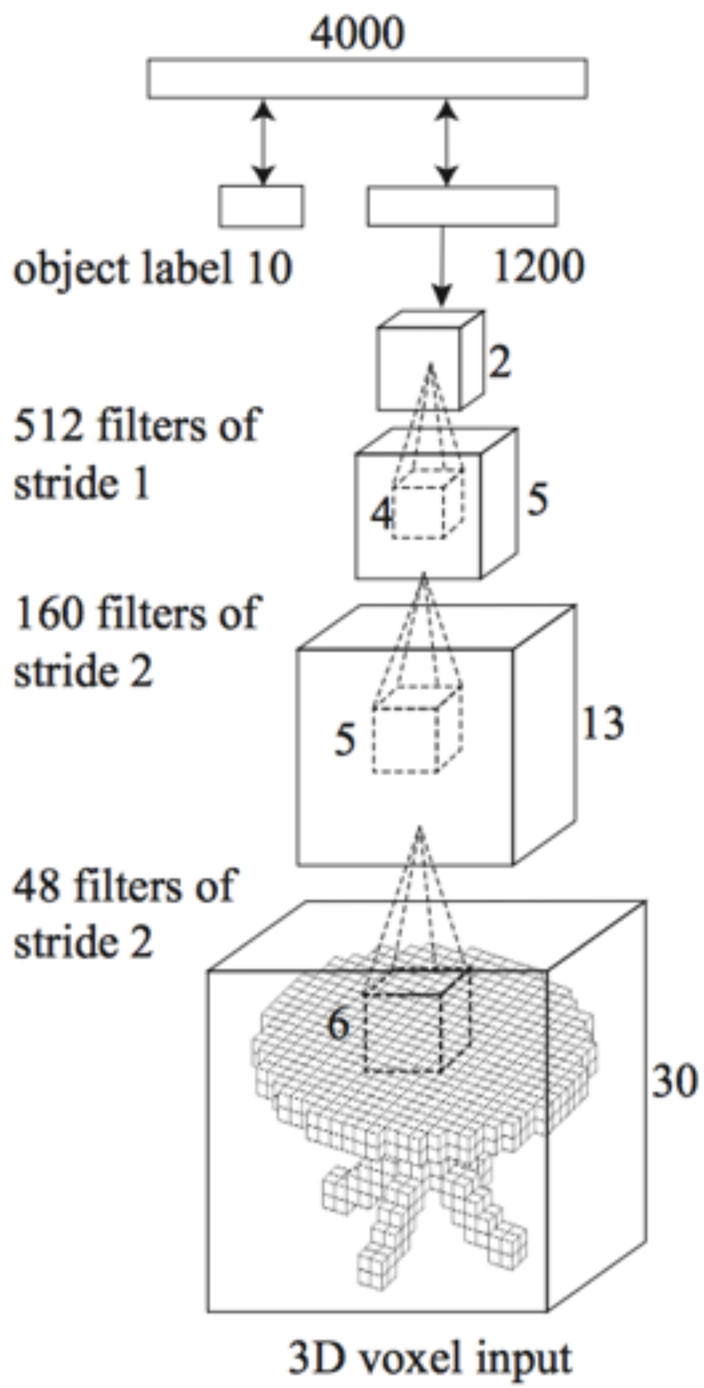
Convolutional Deep Belief Network $p(\mathbf{x}, y)$

3D ShapeNets ≠ CNNs

$p(x, y)$ $p(y|x)$

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D ShapeNets



Convolutional Deep Belief Network $p(\mathbf{x}, y)$

3D ShapeNets $\neq$ CNNs

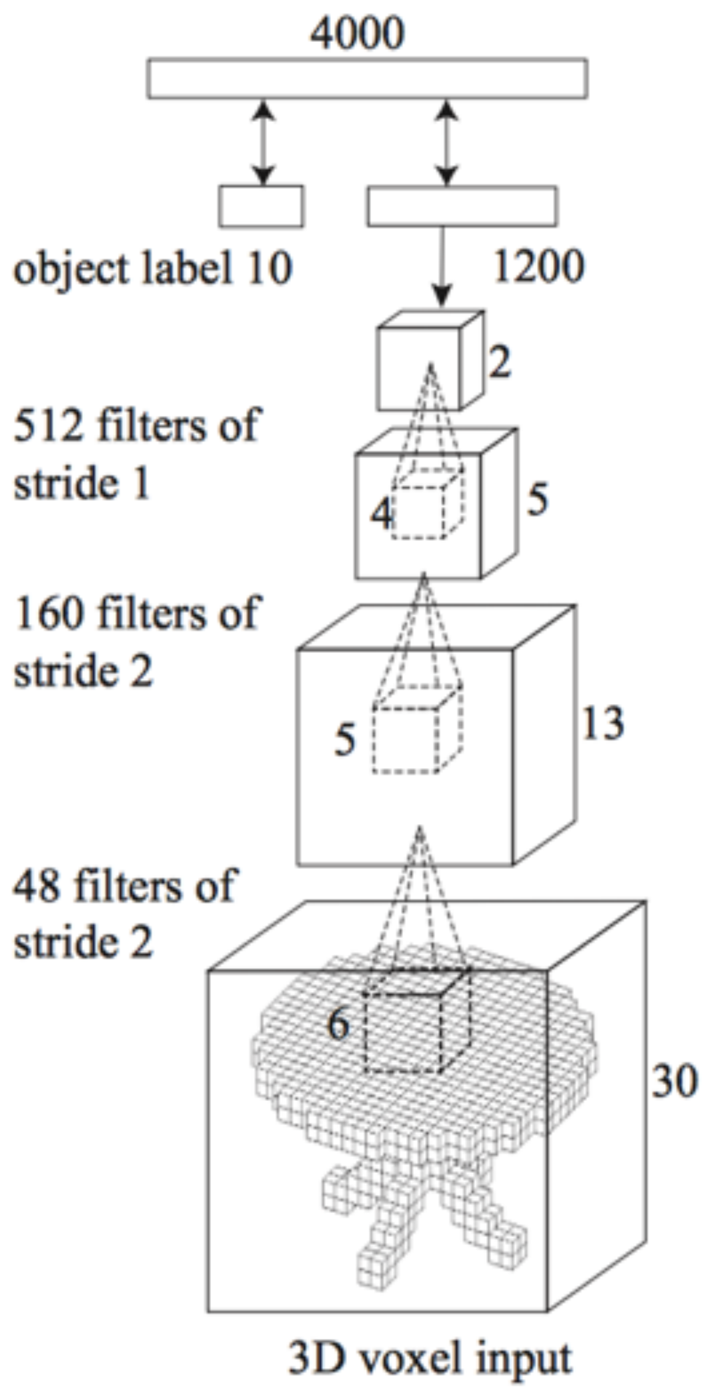
$$p(x, y) \qquad\qquad p(y|x)$$

$$p(y|x)$$ discriminative process

$$p(x|y)$$ generative process

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

3D ShapeNets



Convolutional Deep Belief Network $p(\mathbf{x}, y)$

3D ShapeNets ≠ CNNs

$p(x, y)$ $\qquad$ $p(y|x)$

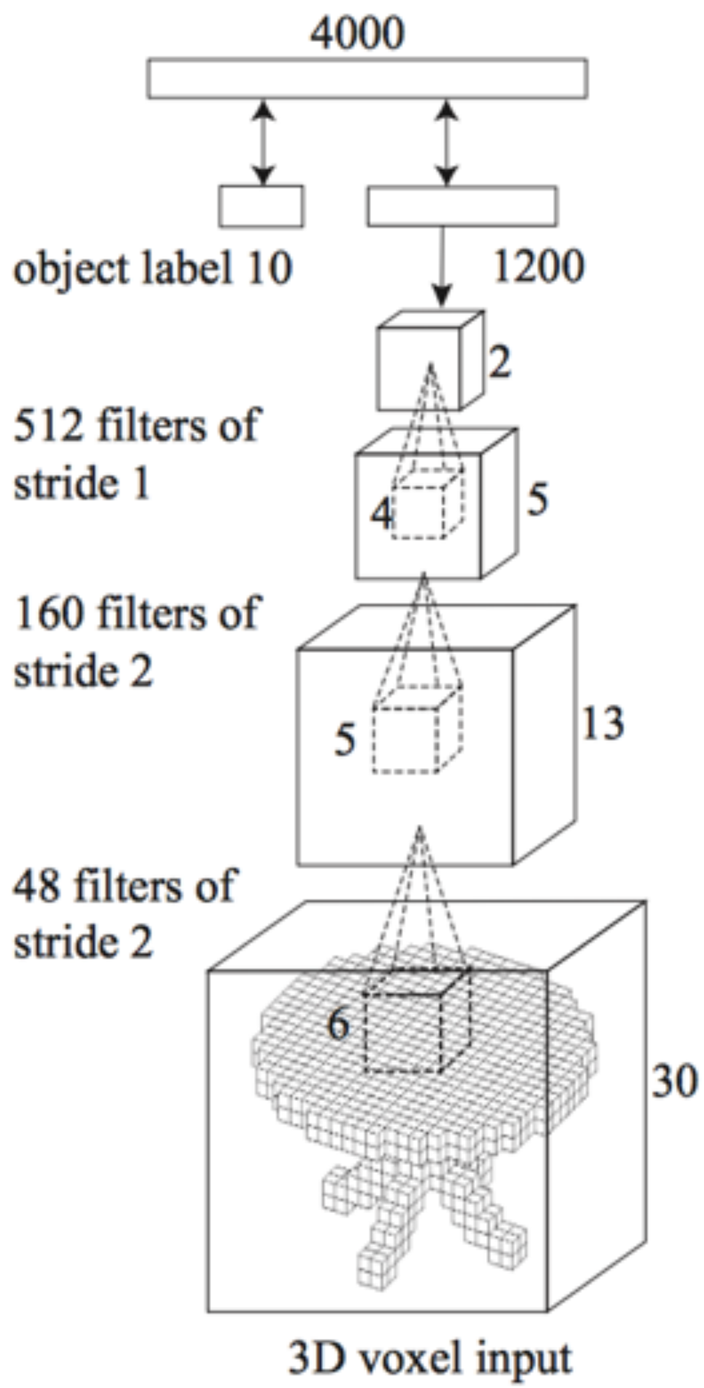
$p(y|x)$ discriminative process

$p(x|y)$ generative process

* 3D ShapeNets can be converted into a CNN, and discriminatively trained with back-propagation.

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Training

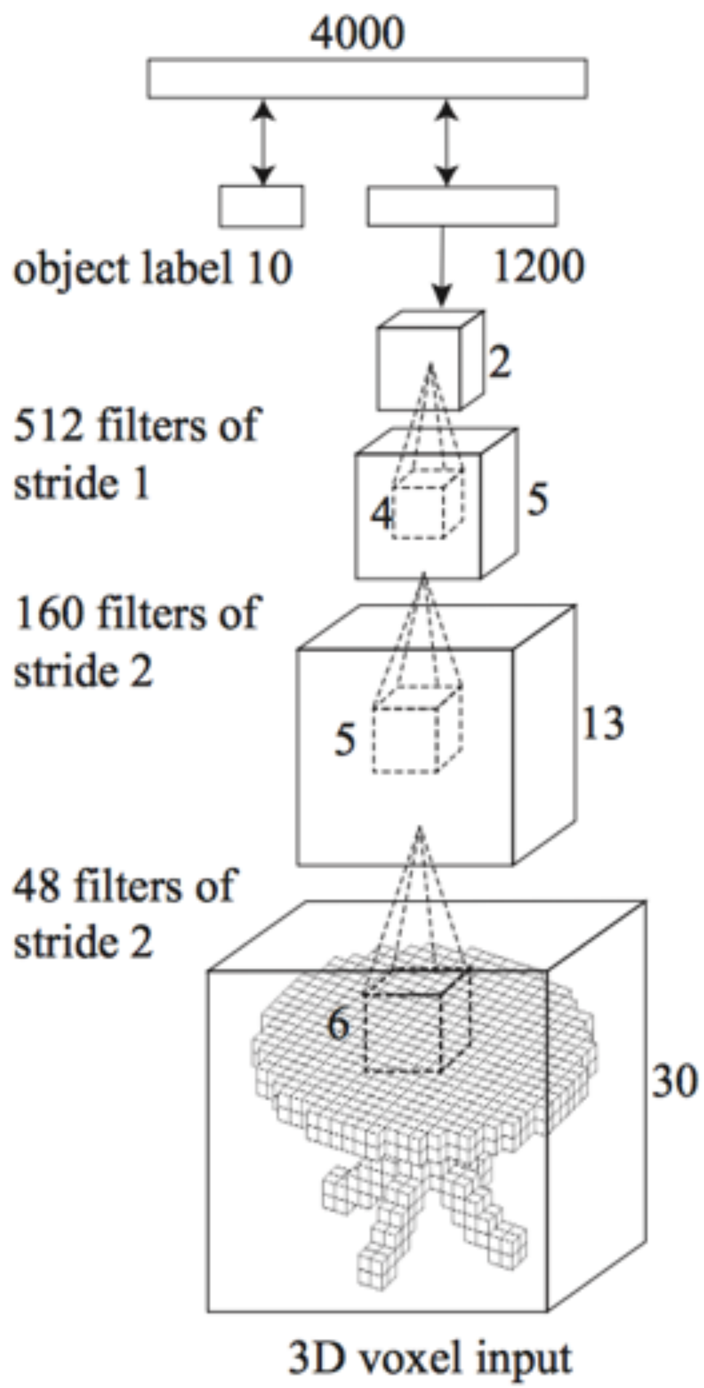


Maximum Likelihood Learning

Convolutional Deep Belief Network $p(\mathbf{x}, y)$

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Training



Convolutional Deep Belief Network $p(\mathbf{x}, y)$

Maximum Likelihood Learning

Layer-wise pre-training:

Lower four layers are trained by CD

Last layer is trained by FPCD[1]

Fine-tuning:

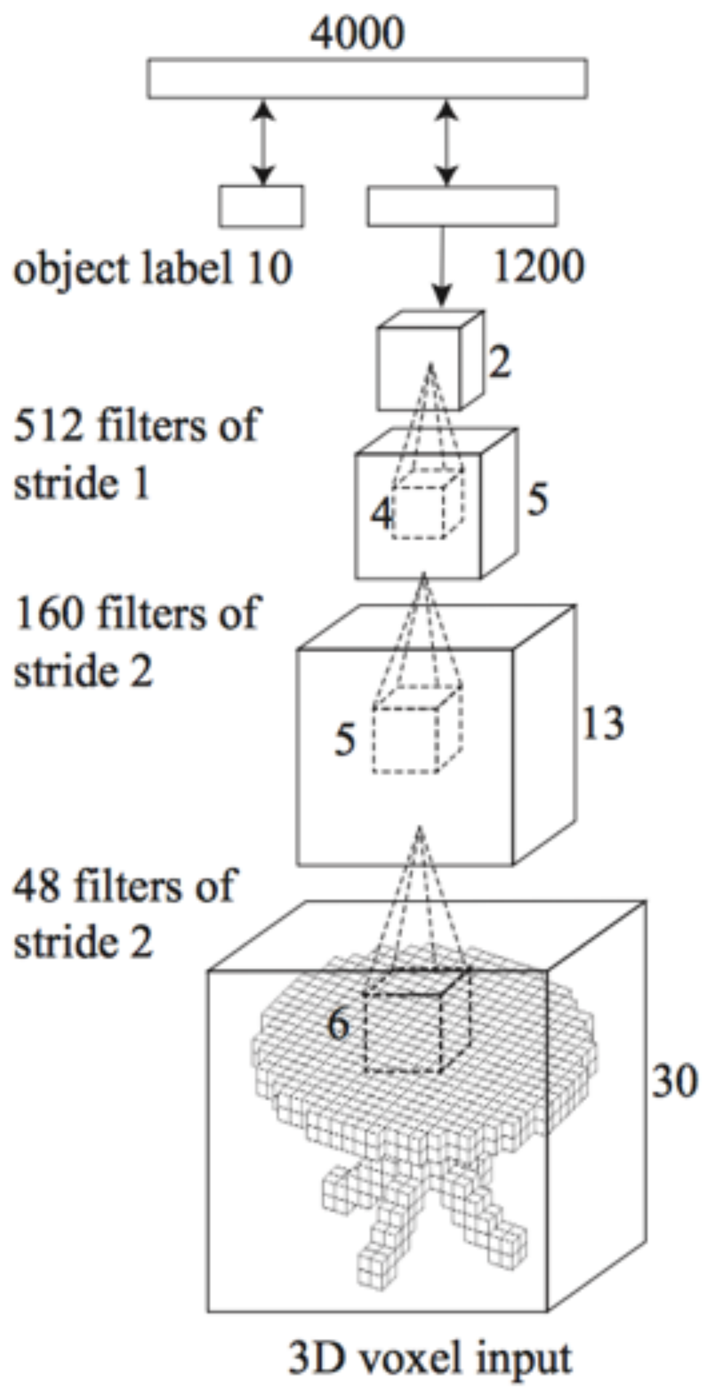
Wake sleep[2] but keep weights tied

[1] Tijmen, et al. "Using fast weights to improve persistent contrastive divergence."

[2] Hinton, et al "A fast learning algorithm for deep belief nets." Neural computation

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Training



Convolutional Deep Belief Network $p(\mathbf{x}, y)$

Maximum Likelihood Learning

Layer-wise pre-training:

- Lower four layers are trained by CD
- Last layer is trained by FPCD[1]

Fine-tuning:

- Wake sleep[2] but keep weights tied

[1] Tijmen, et al. *"Using fast weights to improve persistent contrastive divergence."*

[2] Hinton, et al *"A fast learning algorithm for deep belief nets."* *Neural computation*

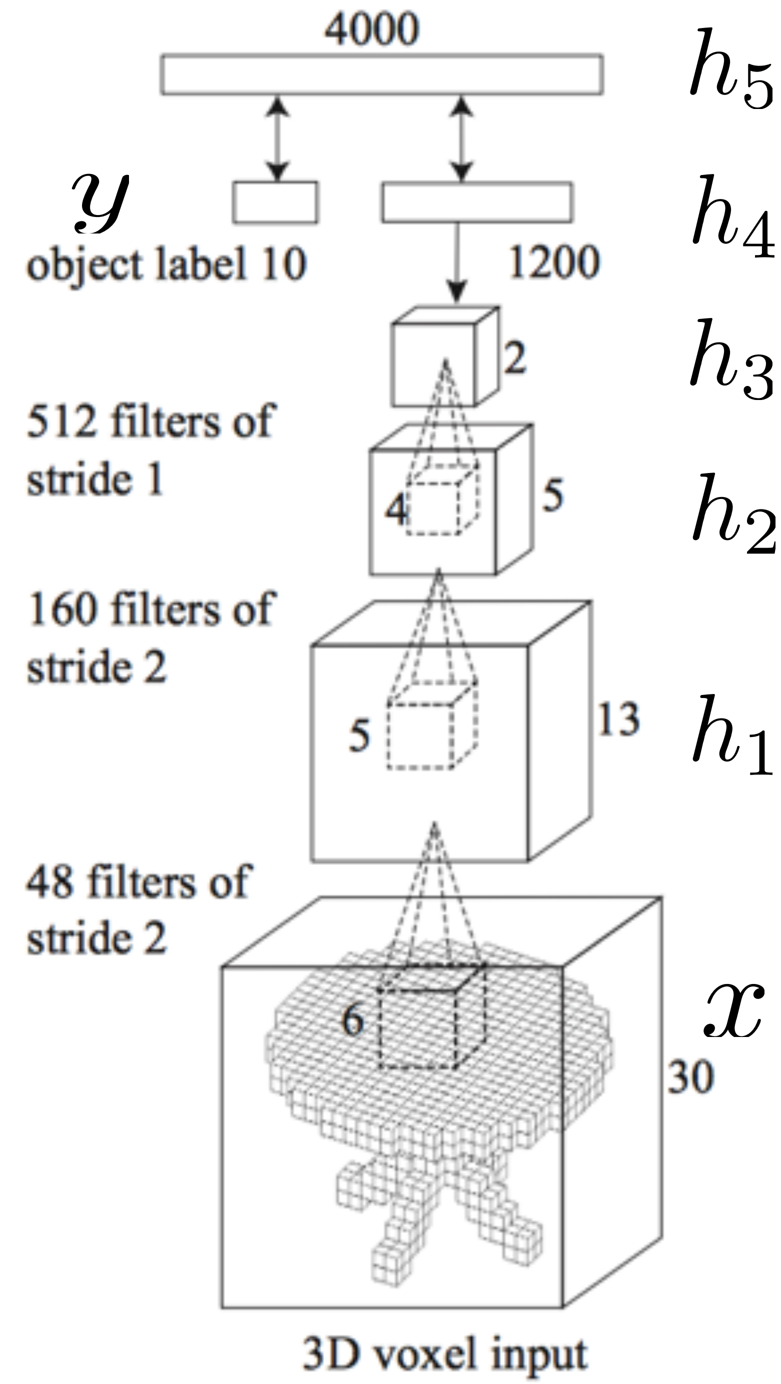
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Sampling

generation process:

Gibbs Sampling

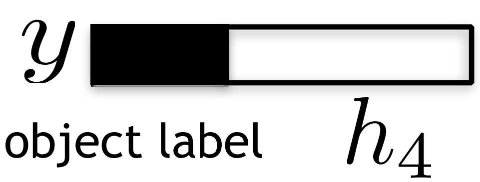
$$p(x, h_1 ... h_5 | y) = p(h_4, h_5 | y) \cdot \prod_{i=1}^{4} p(h_{i-1} | h_i)$$



Convolutional Deep Belief Network $p(\mathbf{x}, y)$

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Sampling



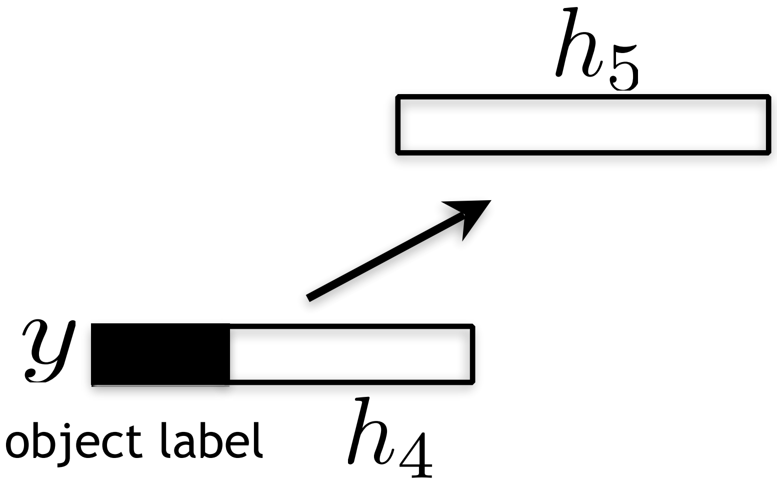
generation process:

Gibbs Sampling

$$p(x, h_1 ... h_5 | y) = p(h_4, h_5 | y) \cdot \prod_{i=1}^{4} p(h_{i-1} | h_i)$$

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Sampling



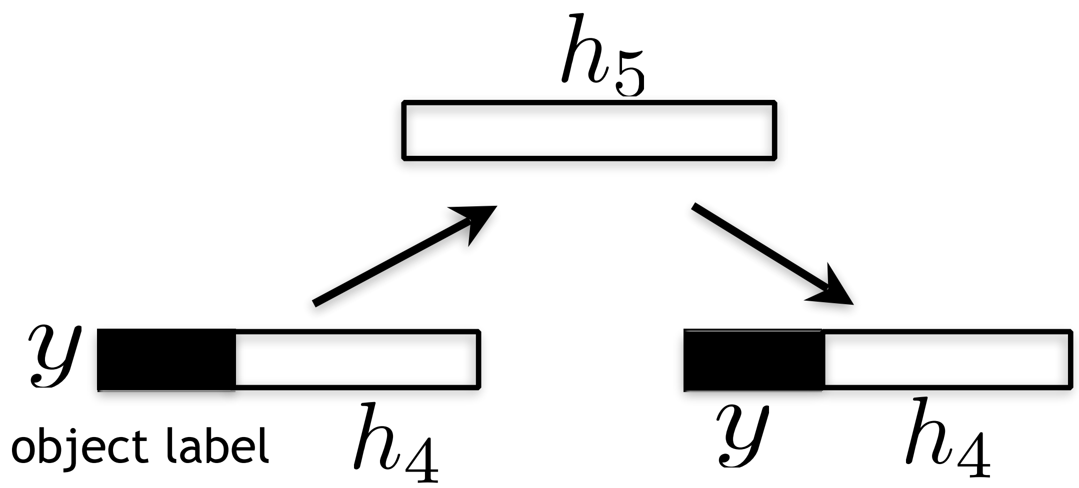
generation process:

Gibbs Sampling

$$p(x, h_1 ... h_5 | y) = p(h_4, h_5 | y) \cdot \prod_{i=1}^{4} p(h_{i-1} | h_i)$$

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Sampling



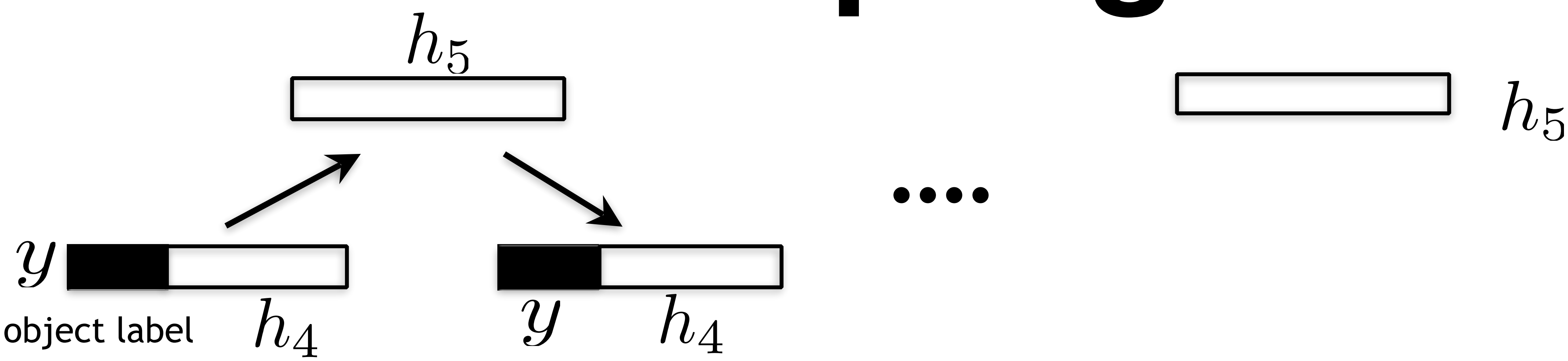
generation process:

Gibbs Sampling

$$p(x, h_1 ... h_5 | y) = p(h_4, h_5 | y) \cdot \prod_{i=1}^{4} p(h_{i-1} | h_i)$$

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Sampling



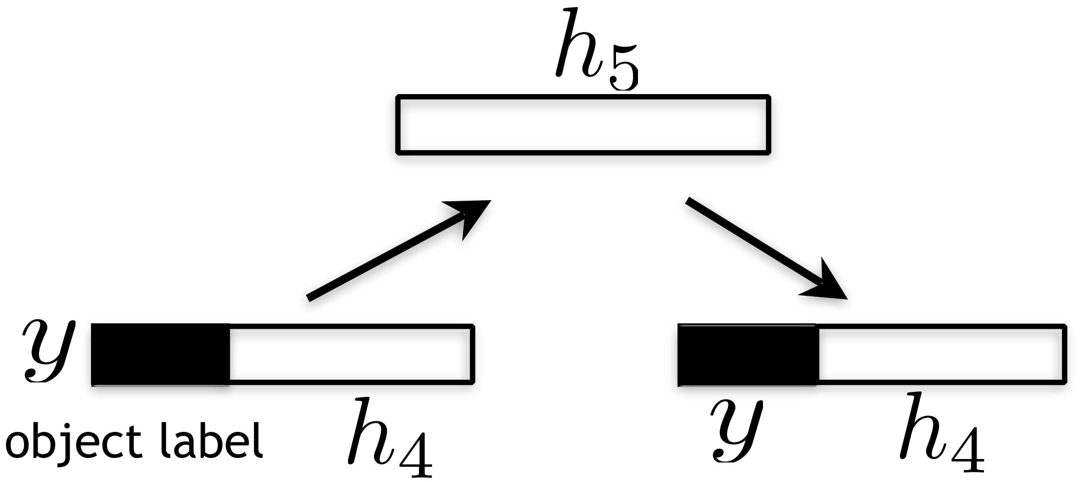
generation process:

Gibbs Sampling

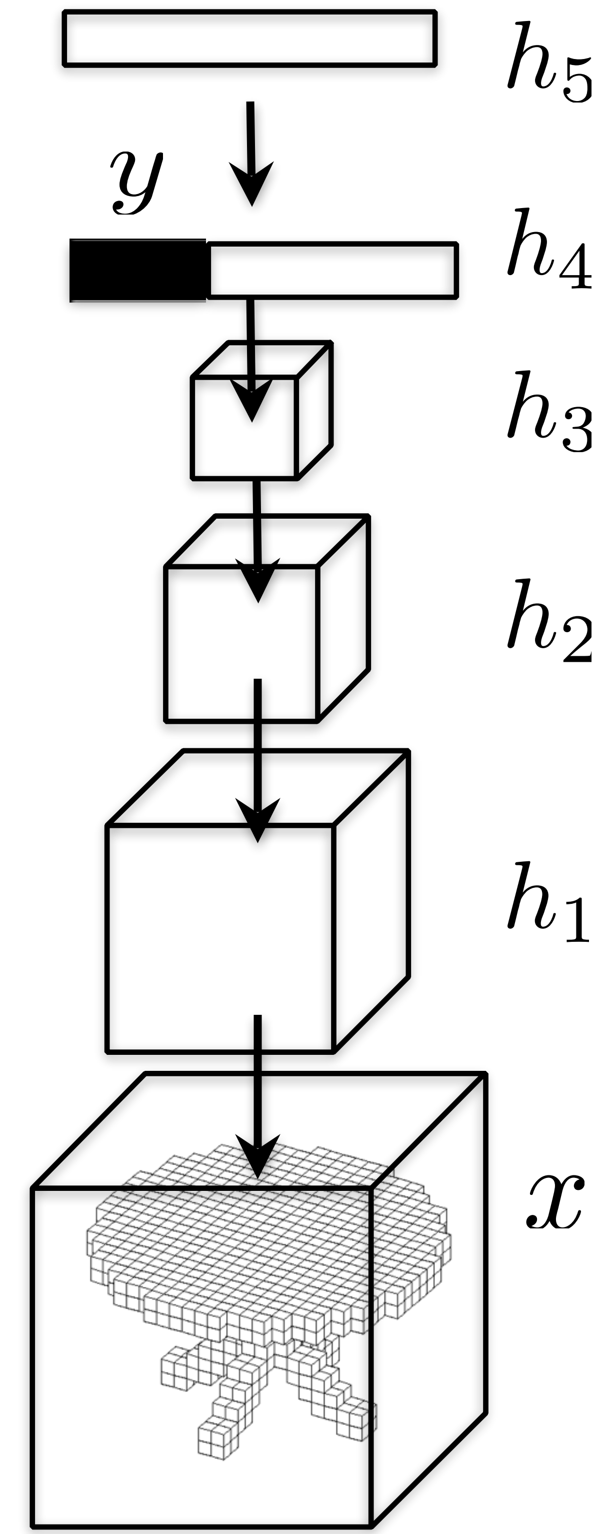
$$p(x, h_1 ... h_5 | y) = p(h_4, h_5 | y) \cdot \prod_{i=1}^{4} p(h_{i-1} | h_i)$$

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Sampling



....



generation process: Gibbs Sampling

$$p(x, h_1 \ldots h_5 | y) = p(h_4, h_5 | y) \cdot \prod_{i=1}^{4} p(h_{i-1} | h_i)$$

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Dataset

Big 3D Data

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Big 3D Data

Query Keyword: common object categories from the SUN database that contain no less than 20 object instances per category

3D Warehouse

yobi 3D

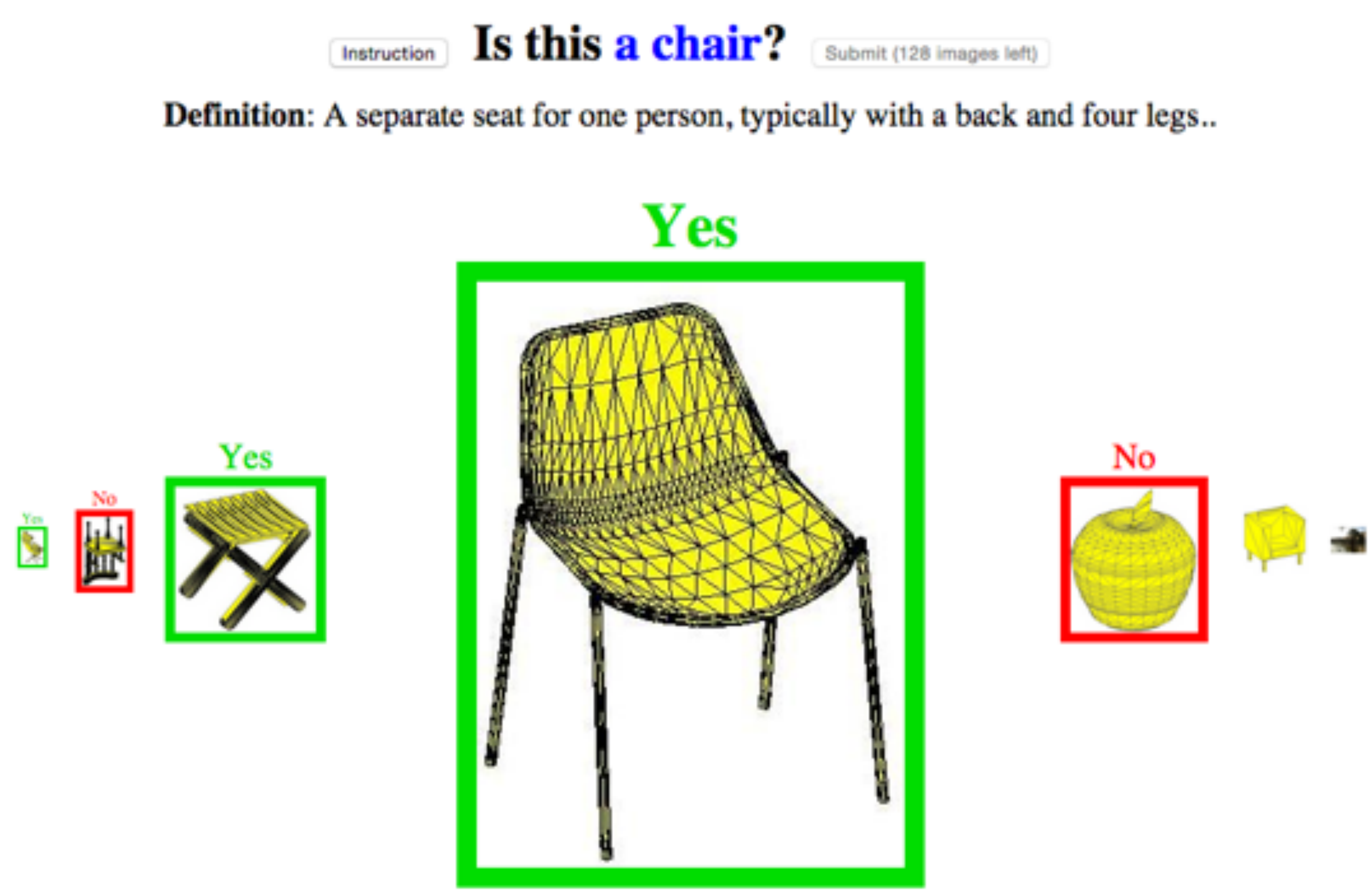
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Big 3D Data

Query Keyword: common object categories from the SUN database that contain no less than 20 object instances per category

3D Warehouse

yobi 3D



Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Big 3D Data

151,128 models 660 categories

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Applications

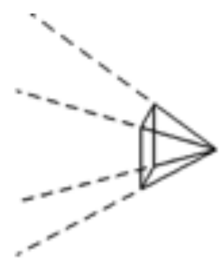
2.5D Completion & Recognition

Slide Credit: Wu et al

2.5D Completion & Recognition

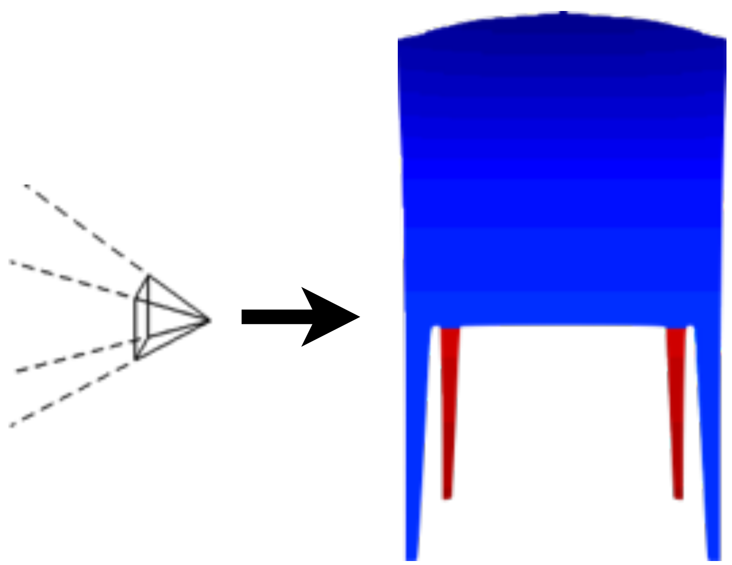
Slide Credit: Wu et al

2.5D Completion & Recognition



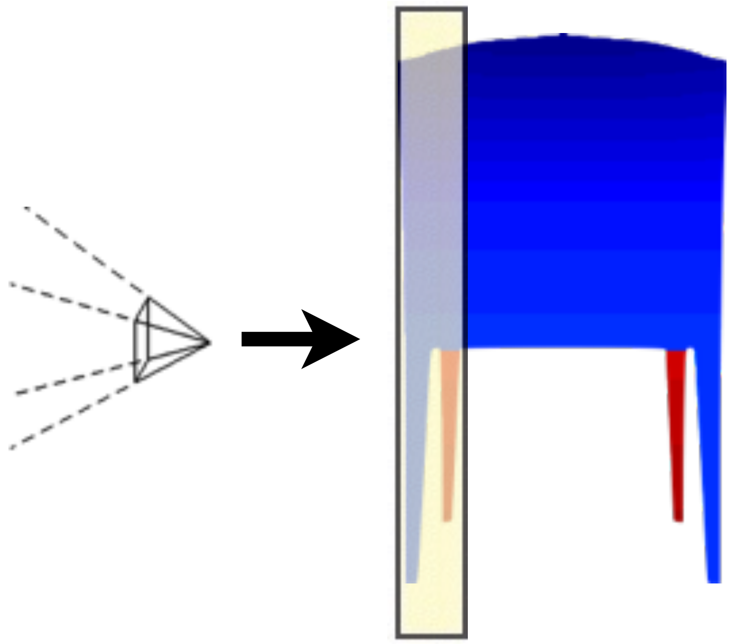
Slide Credit: Wu et al

2.5D Completion & Recognition



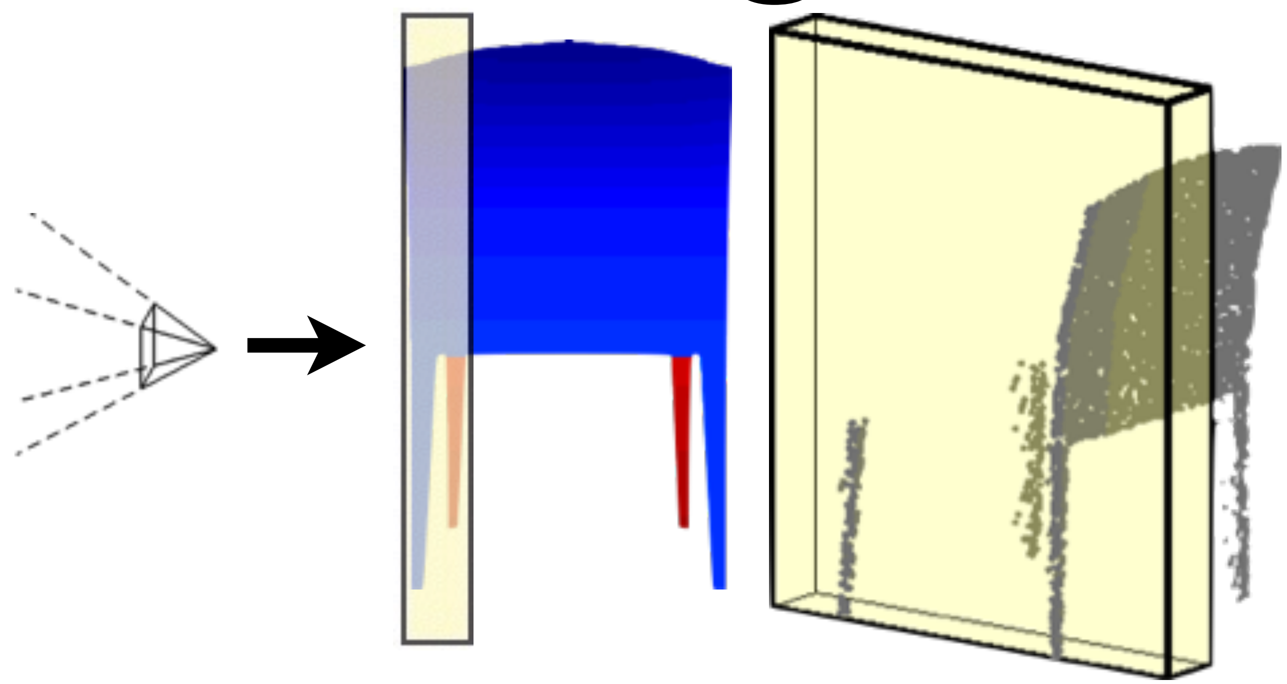
Slide Credit: Wu et al

2.5D Completion & Recognition



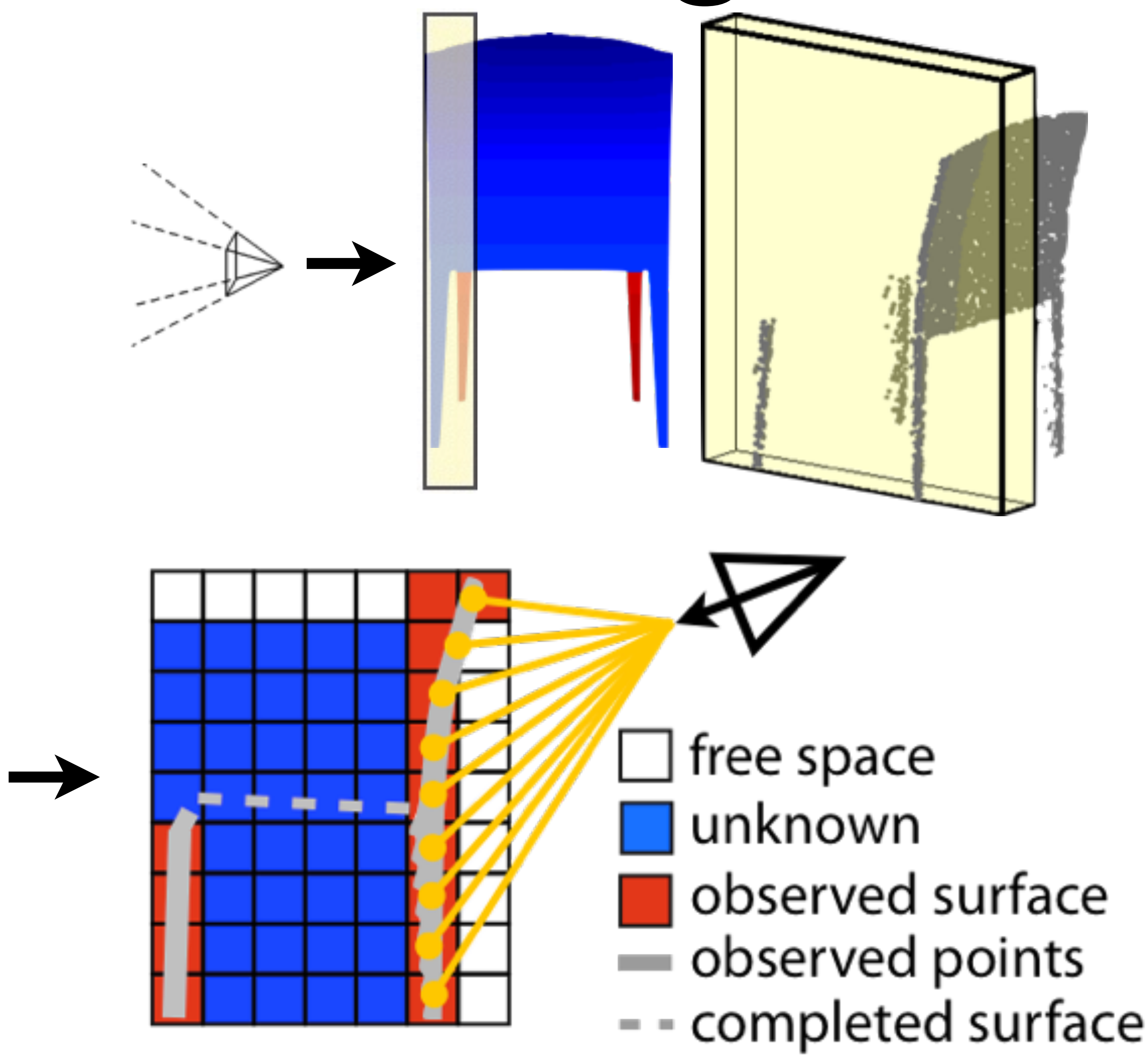
Slide Credit: Wu et al

2.5D Completion & Recognition



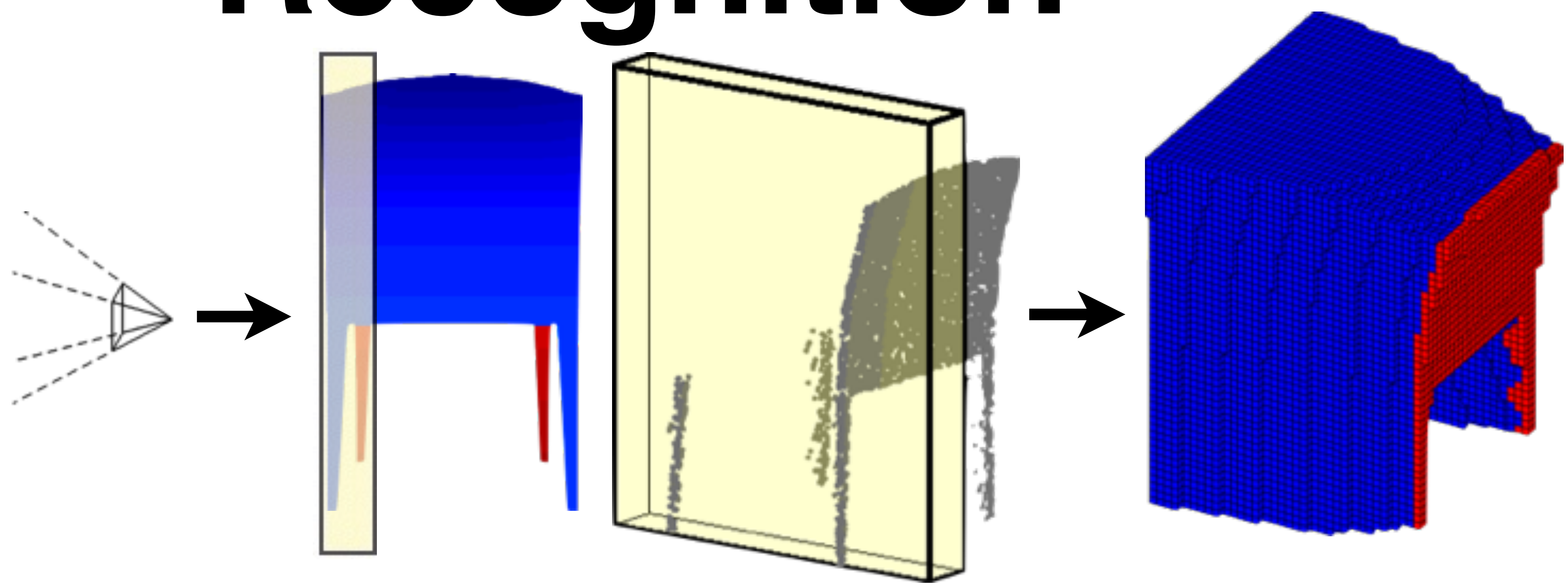
Slide Credit: Wu et al

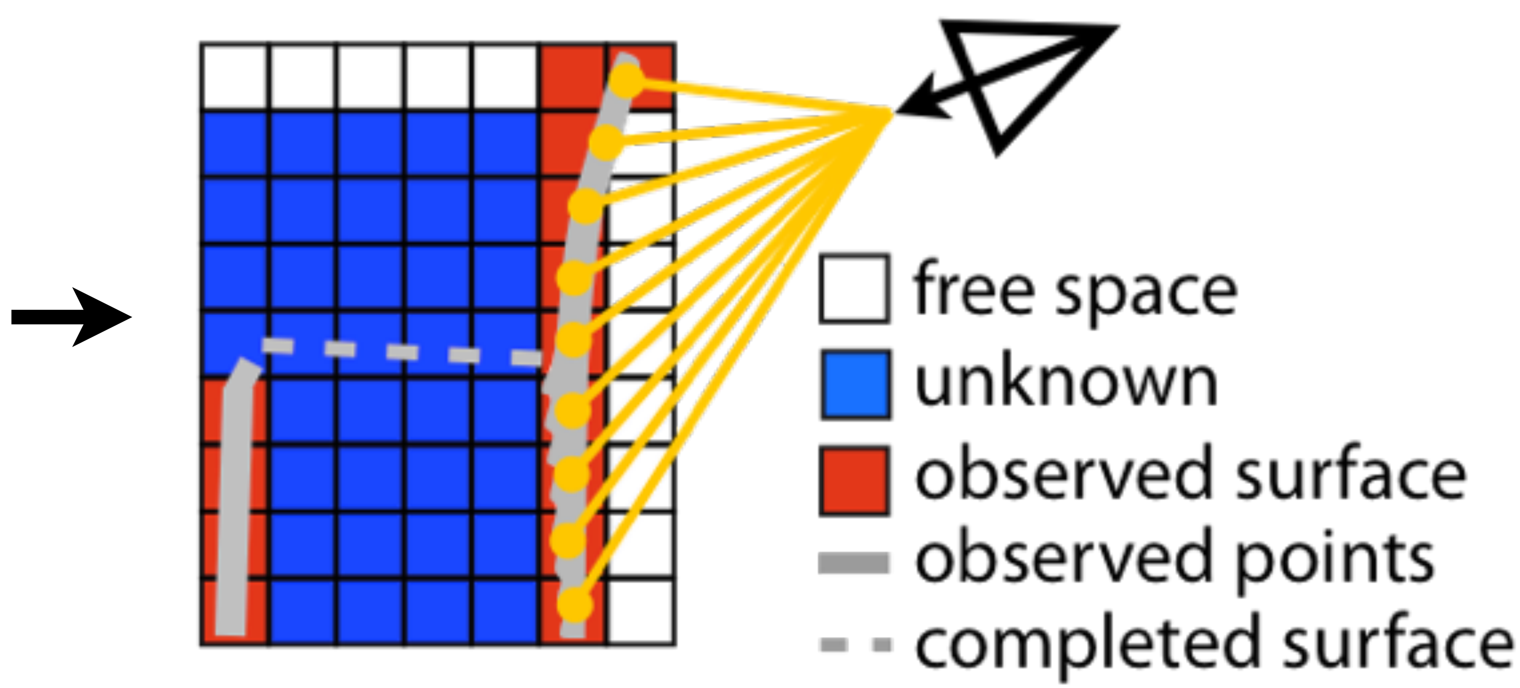
2.5D Completion & Recognition



Slide Credit: Wu et al

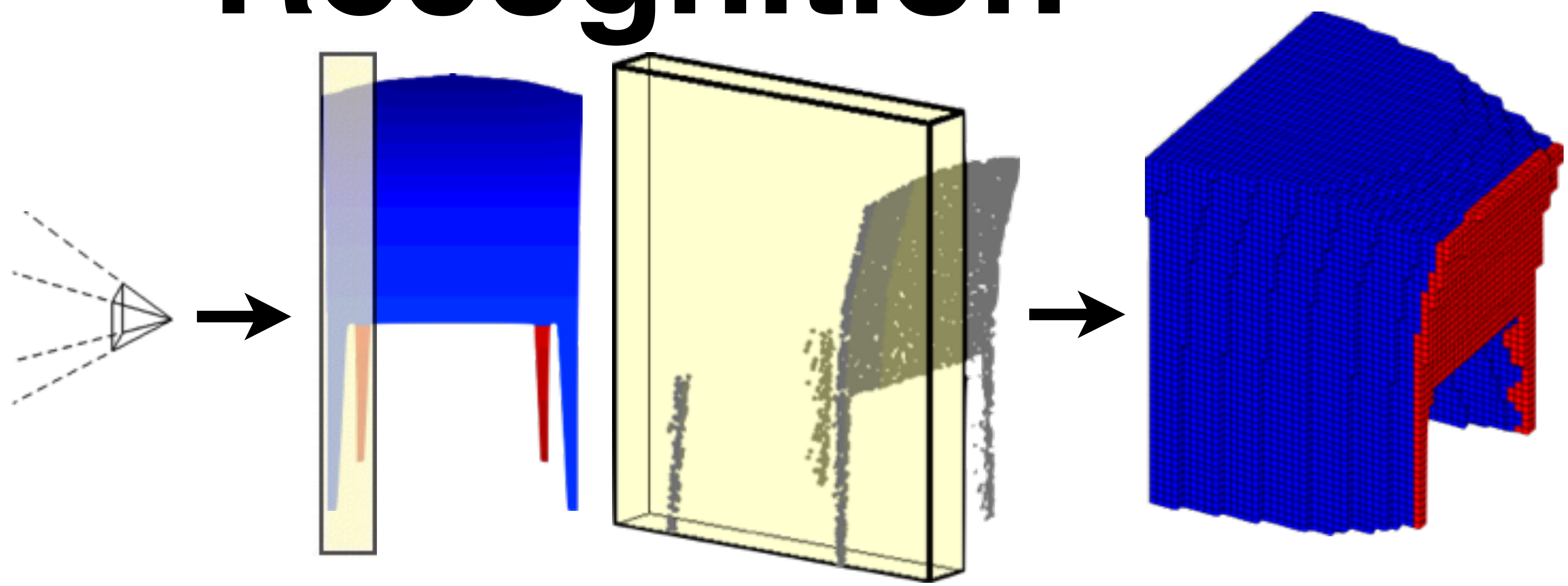
2.5D Completion & Recognition

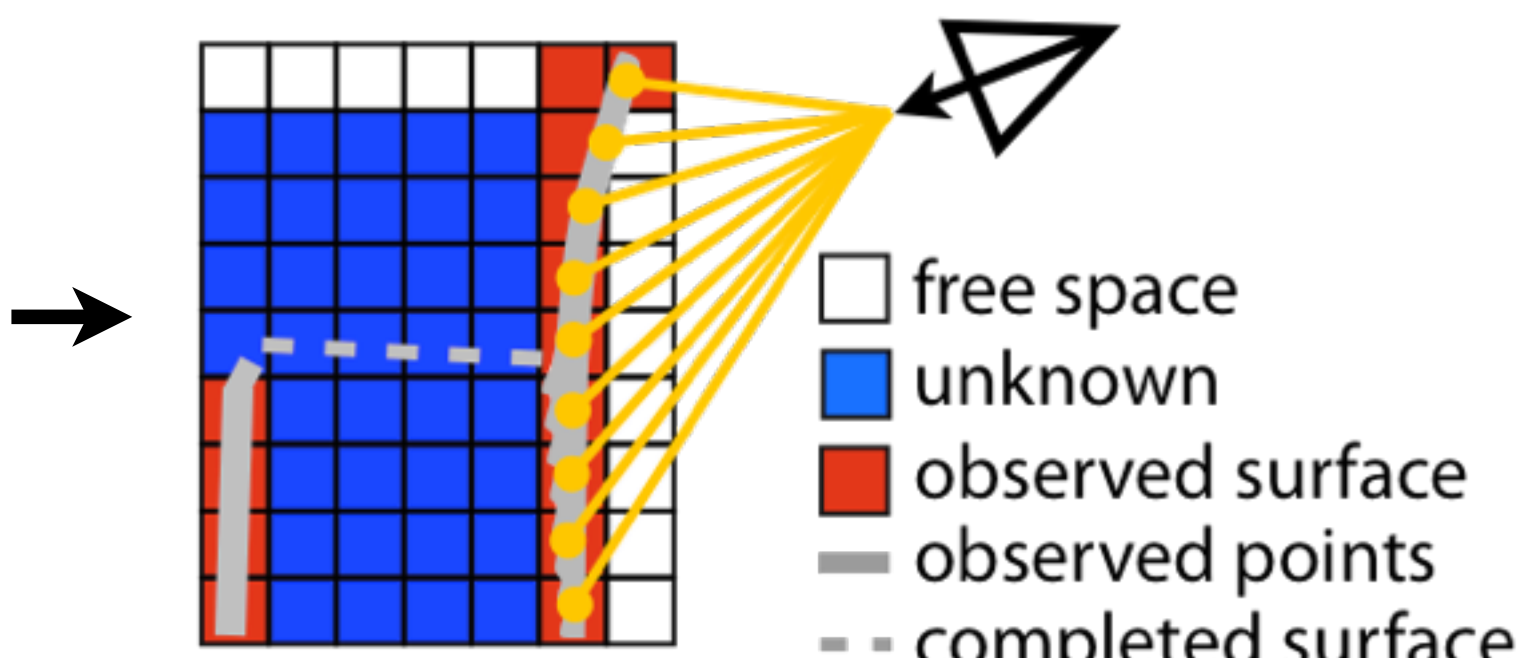




Slide Credit: Wu et al

2.5D Completion & Recognition

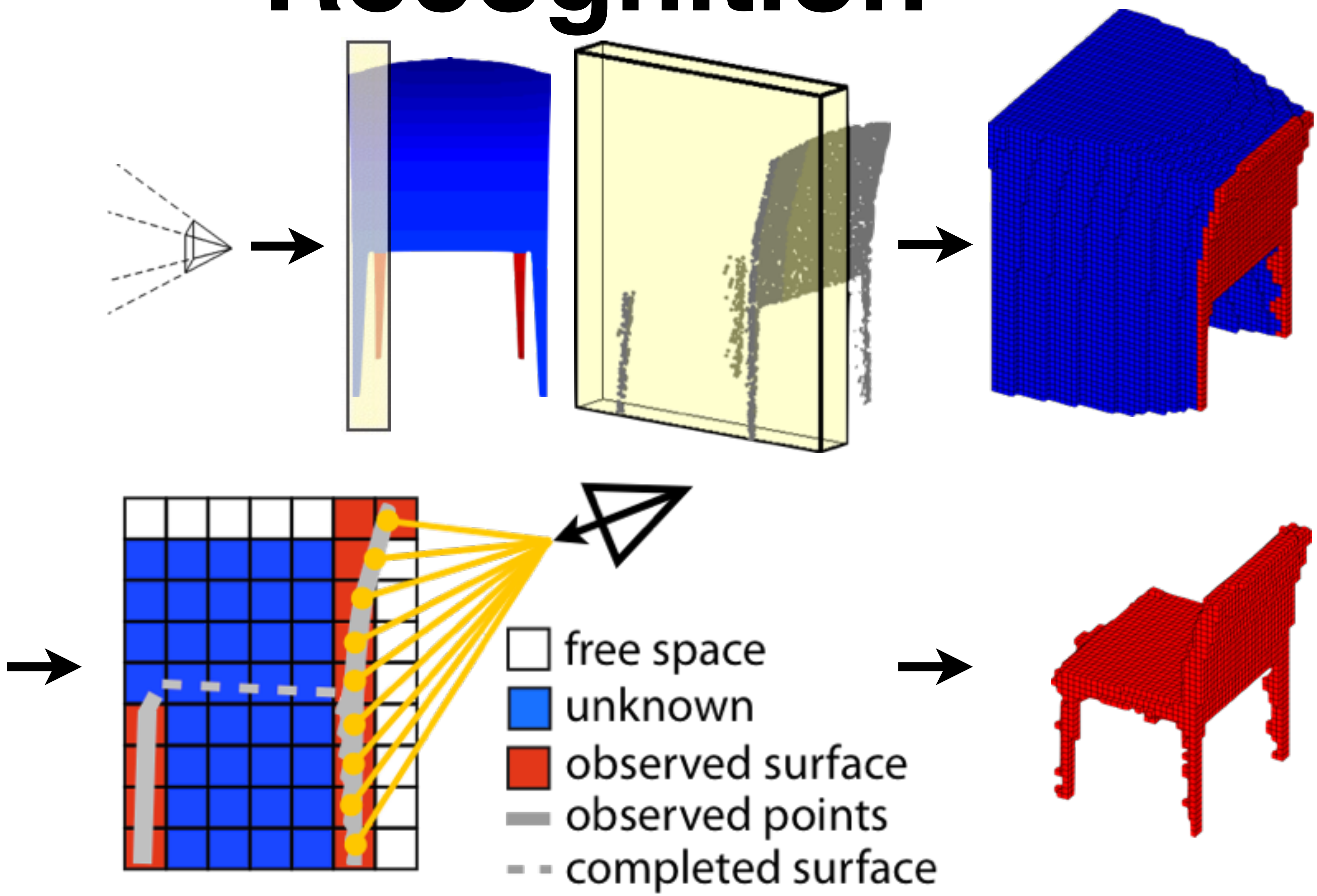




$\mathbf{x} = (\mathbf{x}_u, \mathbf{x}_o)$ $p(y, \mathbf{x_u} | \mathbf{x_o})$ Gibbs sampling with clamping

Slide Credit: Wu et al

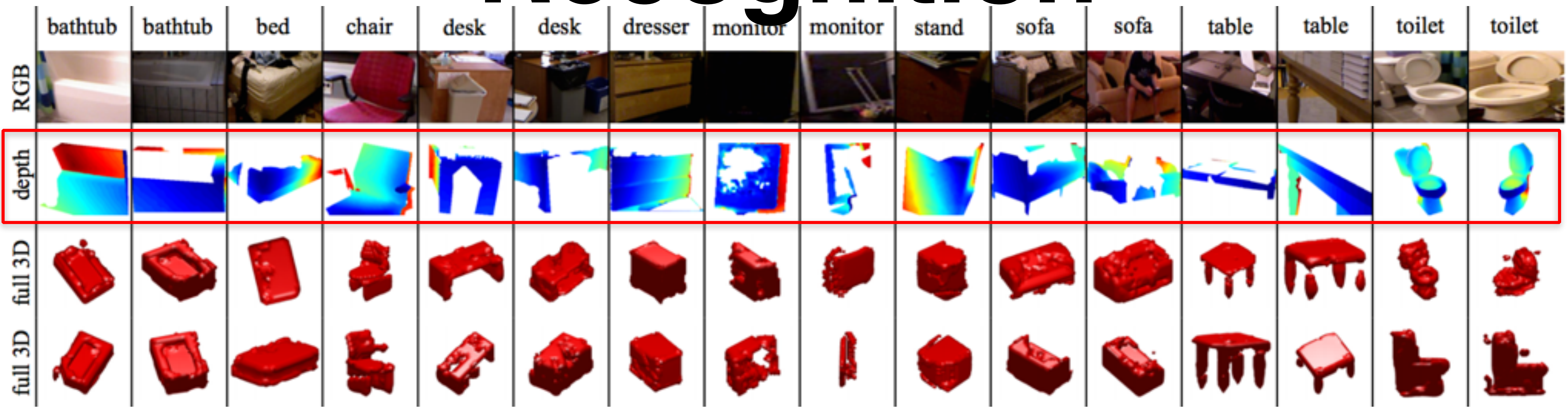
2.5D Completion & Recognition



$\mathbf{x} = (\mathbf{x}_u, \mathbf{x}_o)$ $p(y, \mathbf{x_u} | \mathbf{x_o})$ Gibbs sampling with clamping

Slide Credit: Wu et al

2.5D Completion & Recognition



Training on CAD models and no discriminative tuning!

	all
[29] Depth	0.376
NN	0.374
ICP	0.471
3D ShapeNets	0.437
3D ShapeNets fine-tuned	**0.579**
[29] RGB	0.334
[29] RGBD	0.448

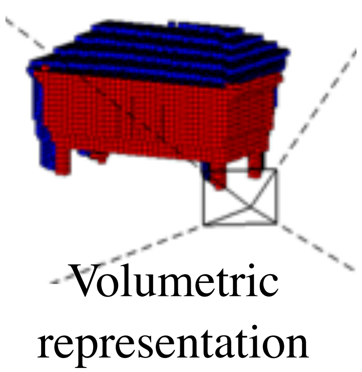
[29] R. Socher, B. Huval, B. Bhat, C. D. Manning, and A. Y. Ng. Convolutional-recursive deep learning for 3d object classification. In NIPS 2012.

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

View Planning for Recognition

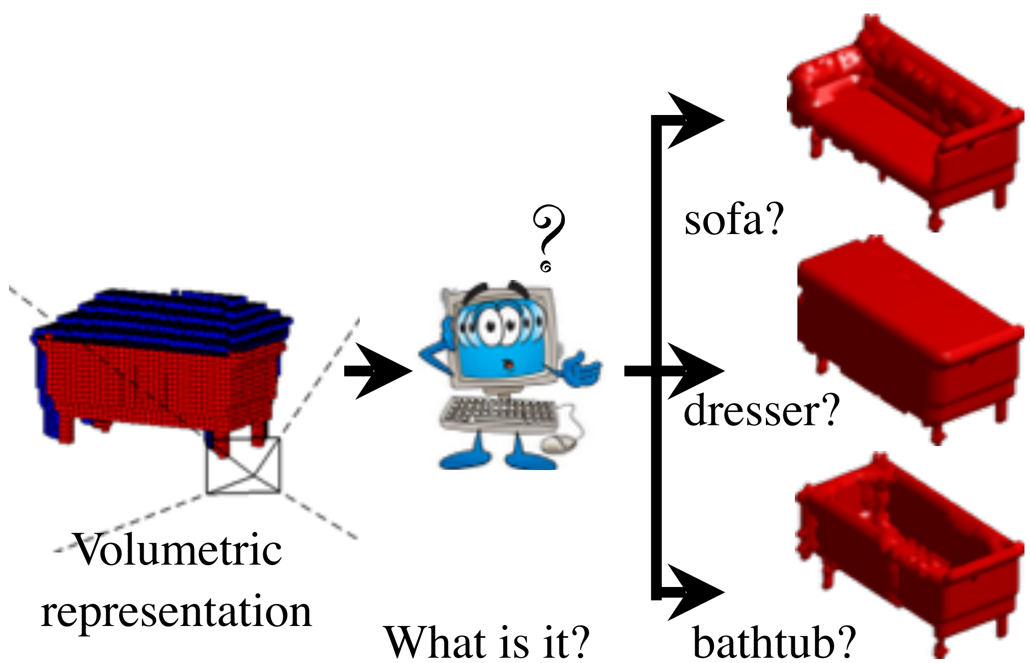
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

View Planning for Recognition



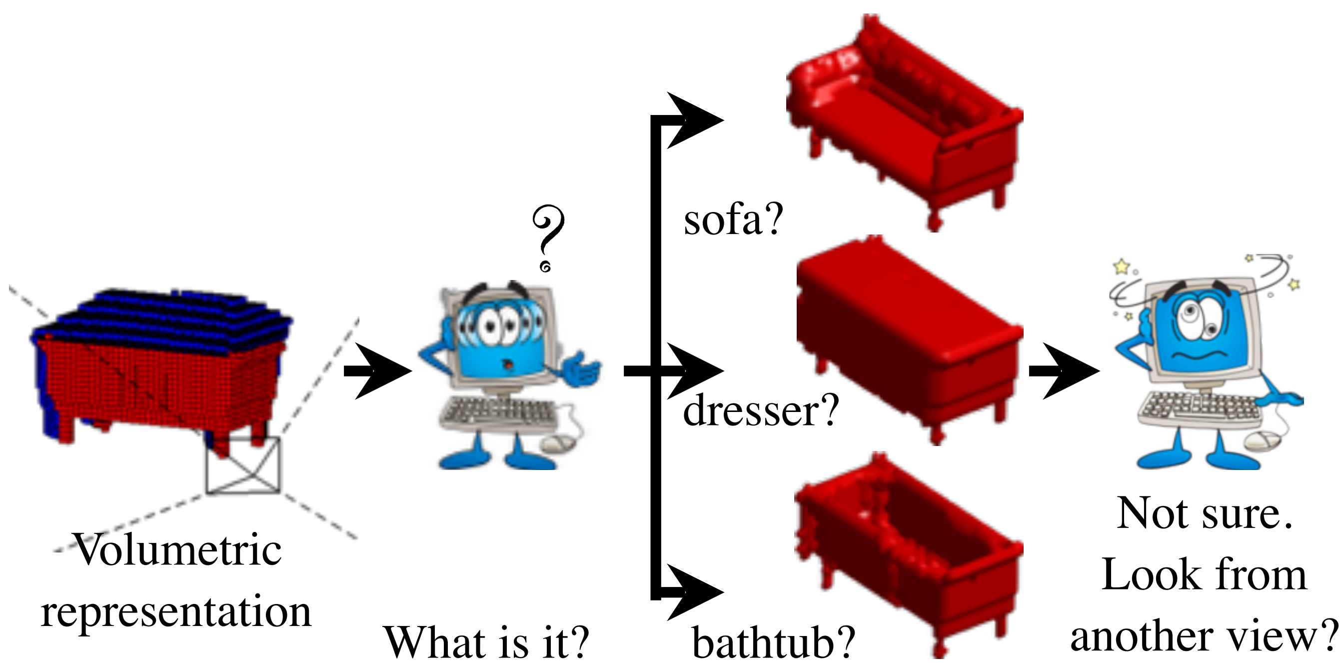
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

View Planning for Recognition



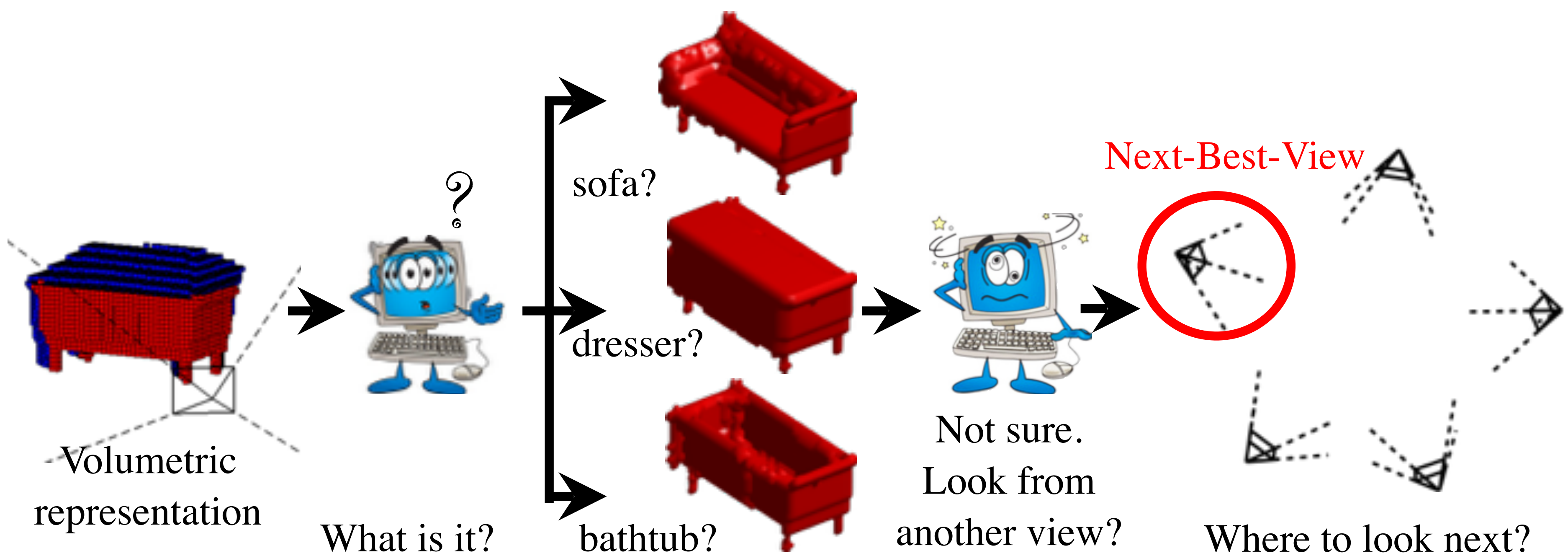
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

View Planning for Recognition



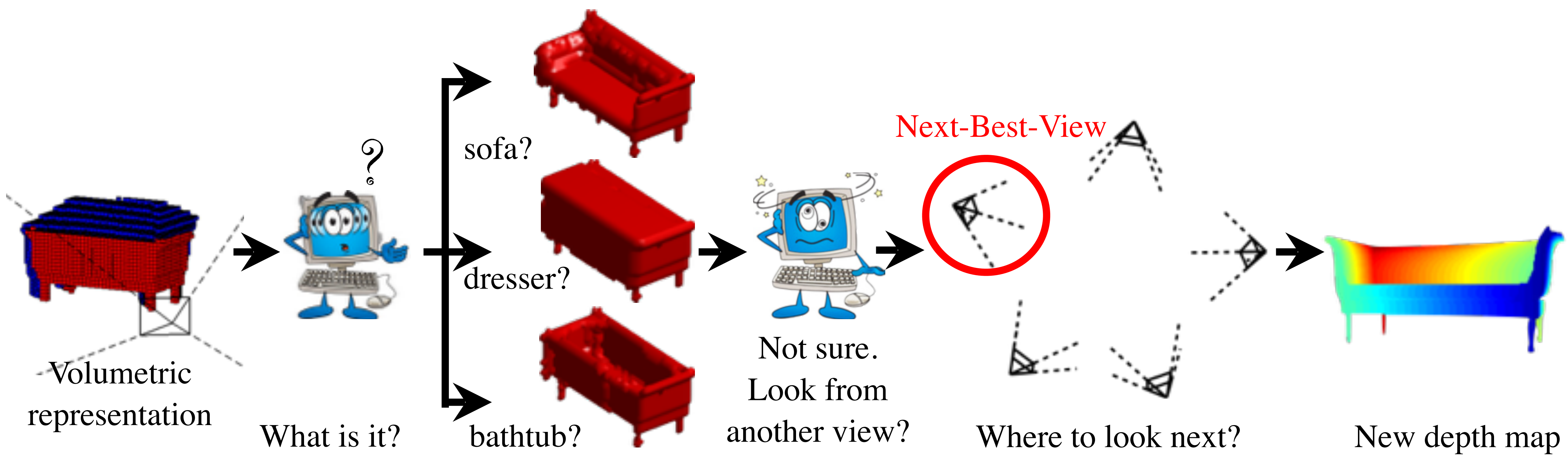
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

View Planning for Recognition



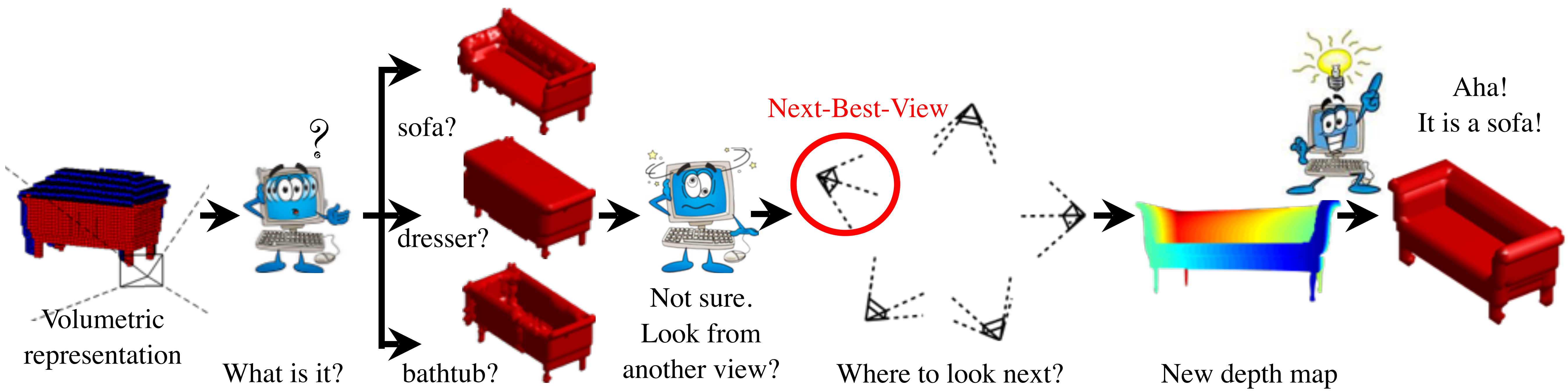
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

View Planning for Recognition



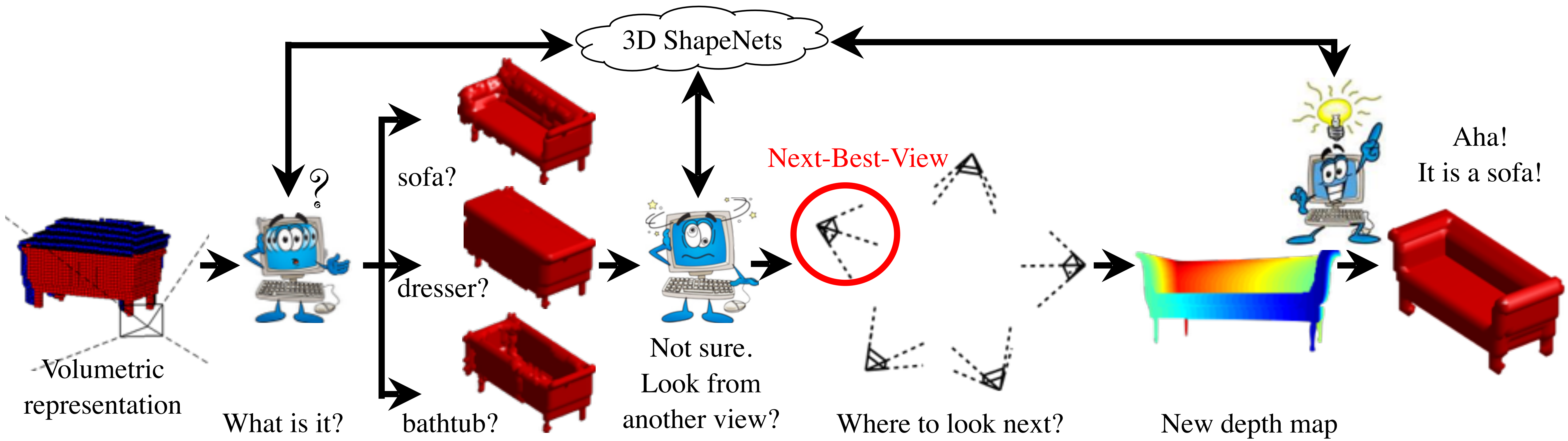
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

View Planning for Recognition



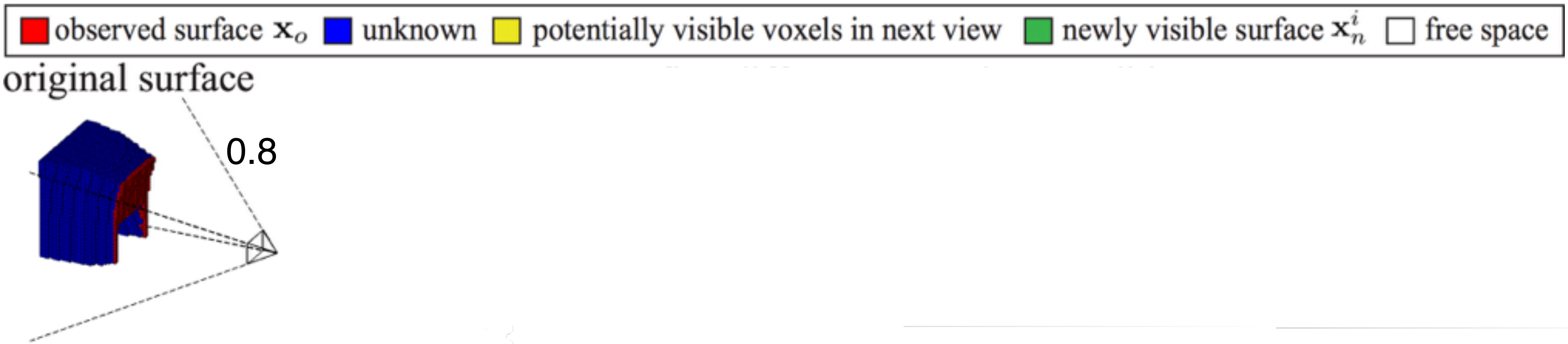
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

View Planning for Recognition



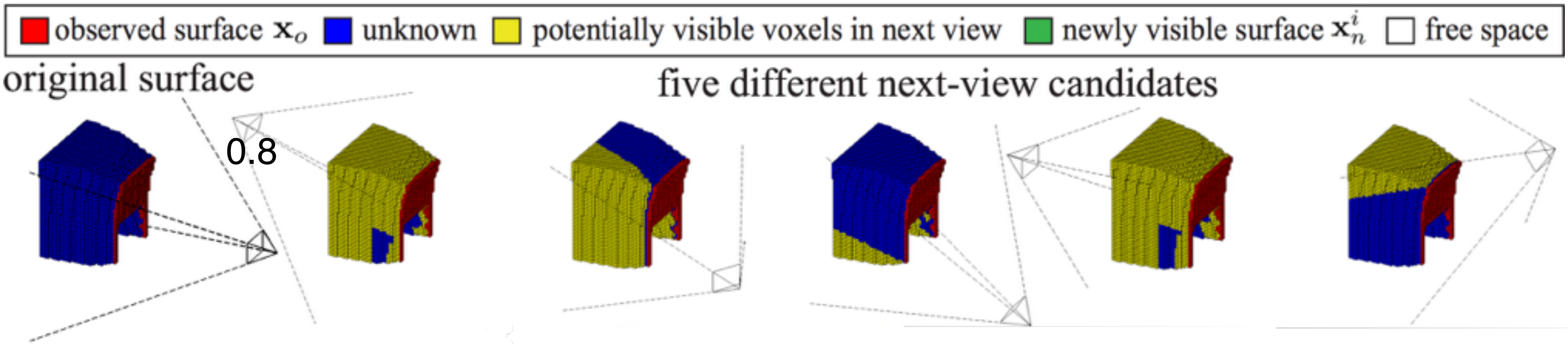
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Deep View Planning



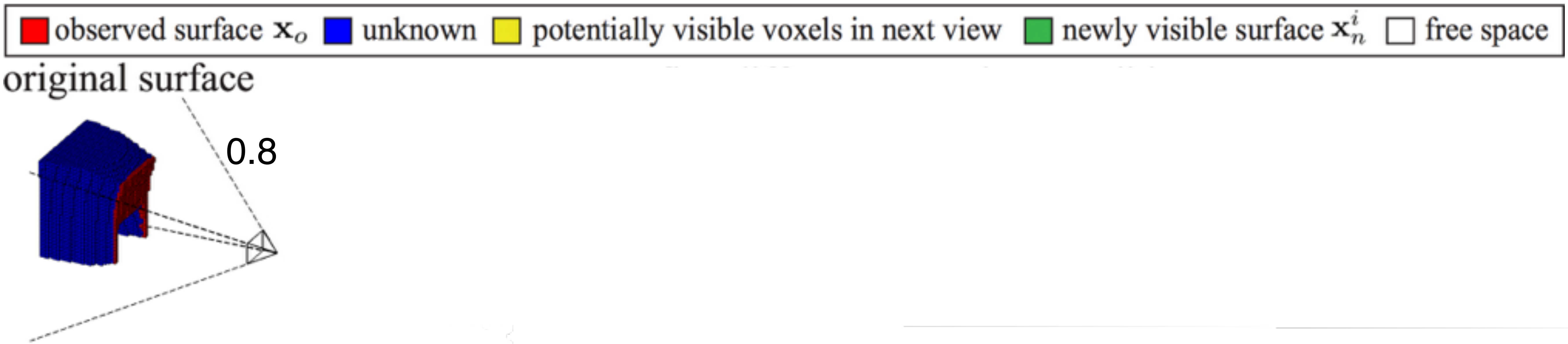
Slide Credit: Wu et al

Deep View Planning



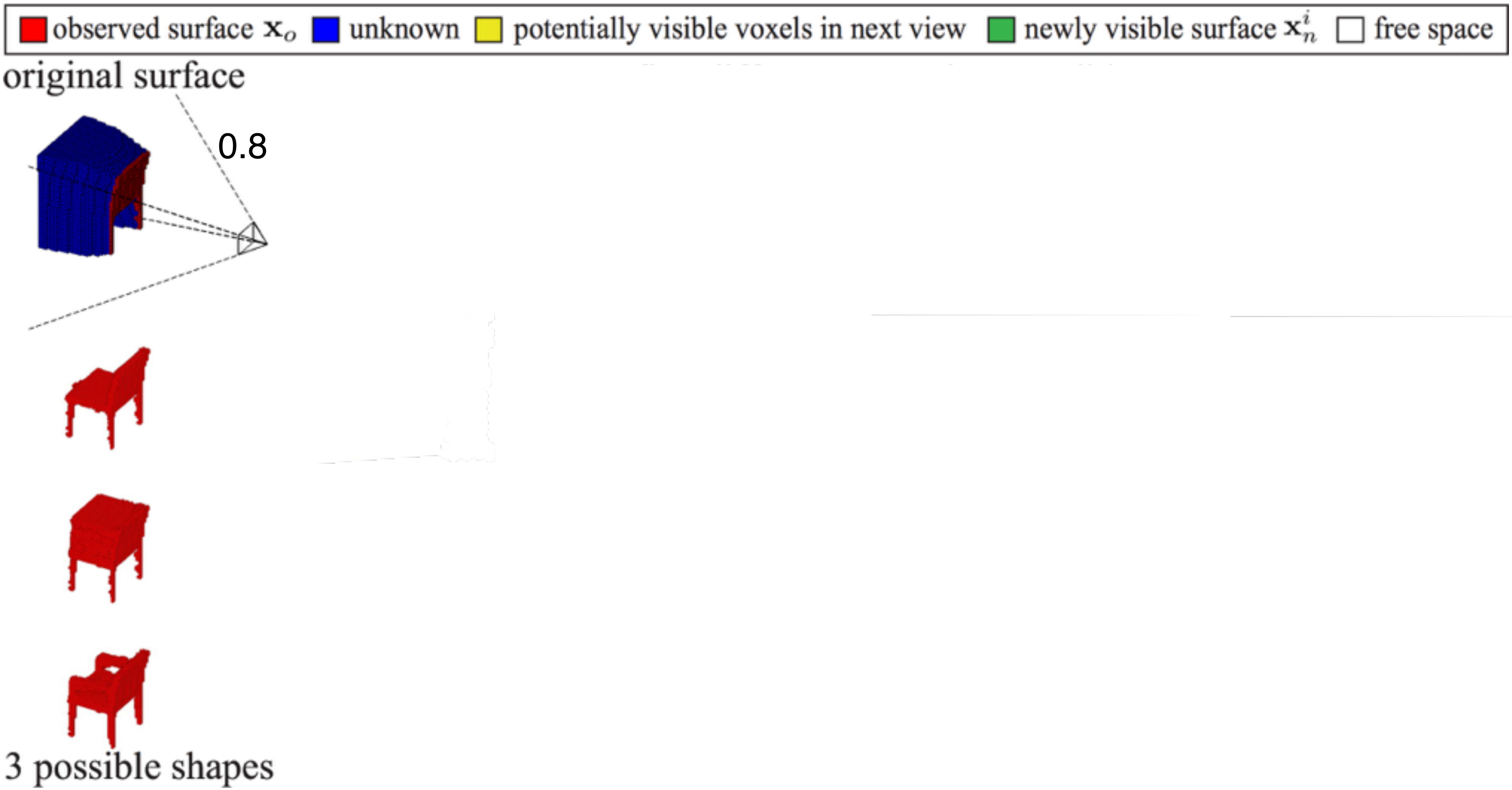
Slide Credit: Wu et al

Deep View Planning



Slide Credit: Wu et al

Deep View Planning

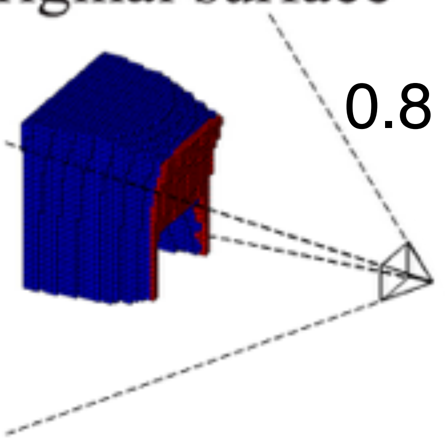


Slide Credit: Wu et al

Deep View Planning

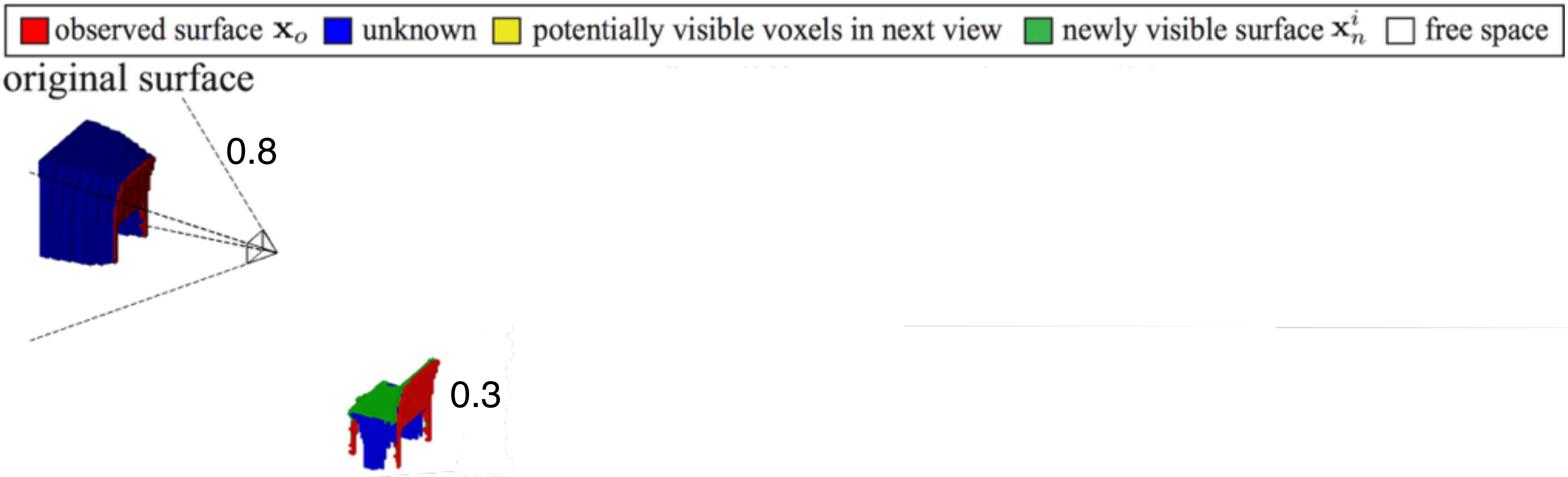
observed surface $\mathbf{x}_o$ unknown potentially visible voxels in next view newly visible surface $\mathbf{x}_n^i$ free space

original surface



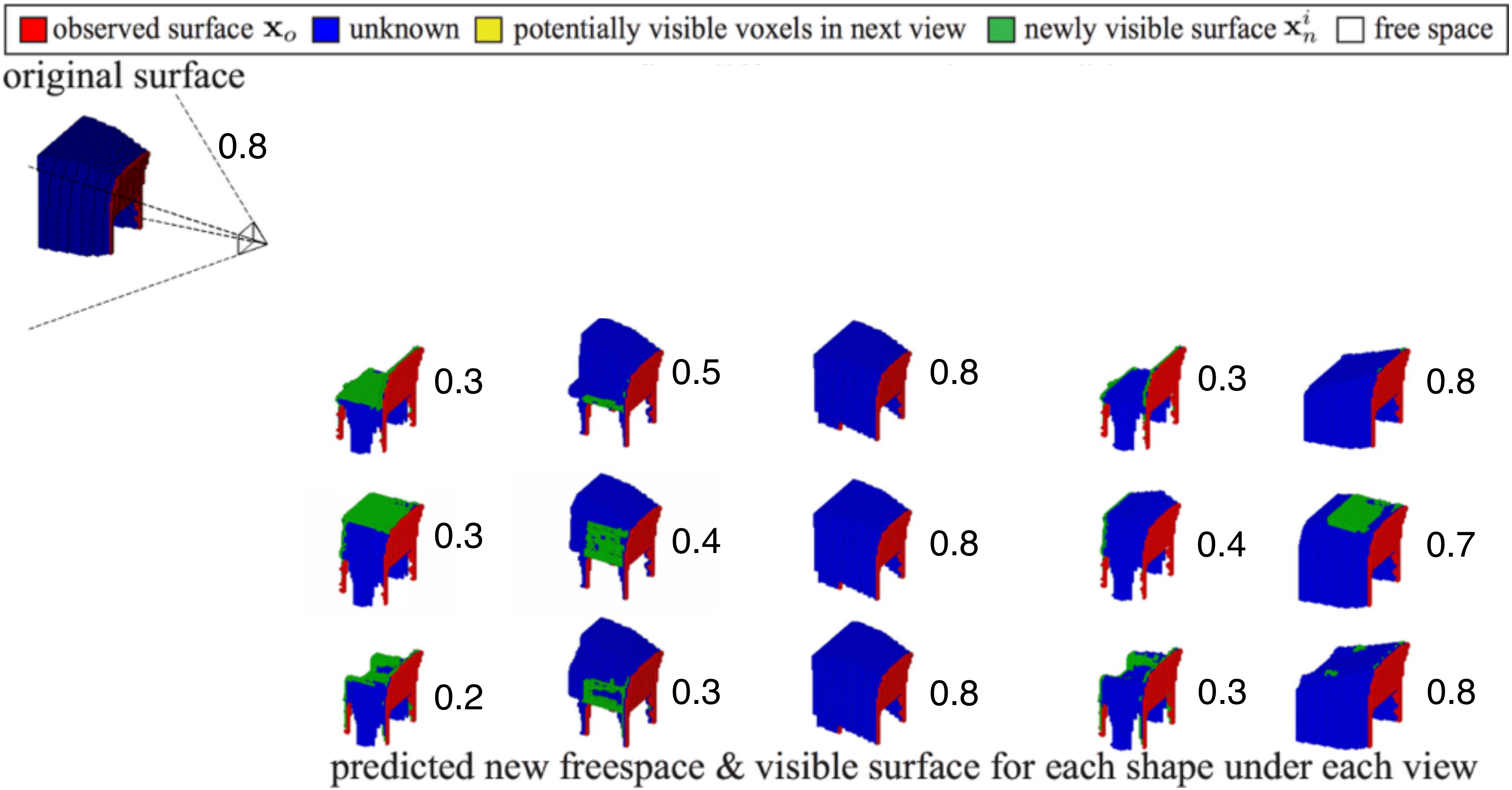
Slide Credit: Wu et al

Deep View Planning



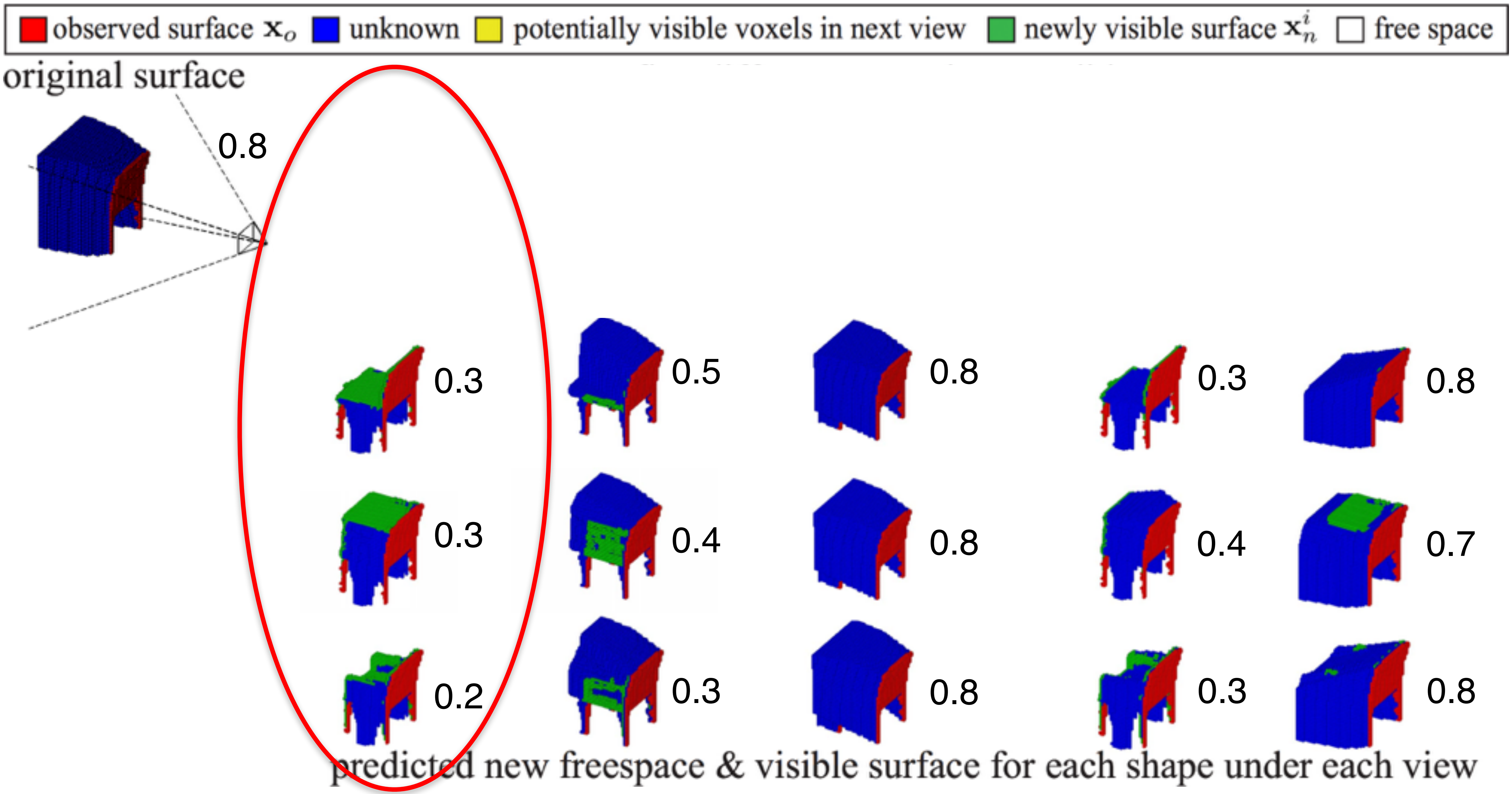
Slide Credit: Wu et al

Deep View Planning



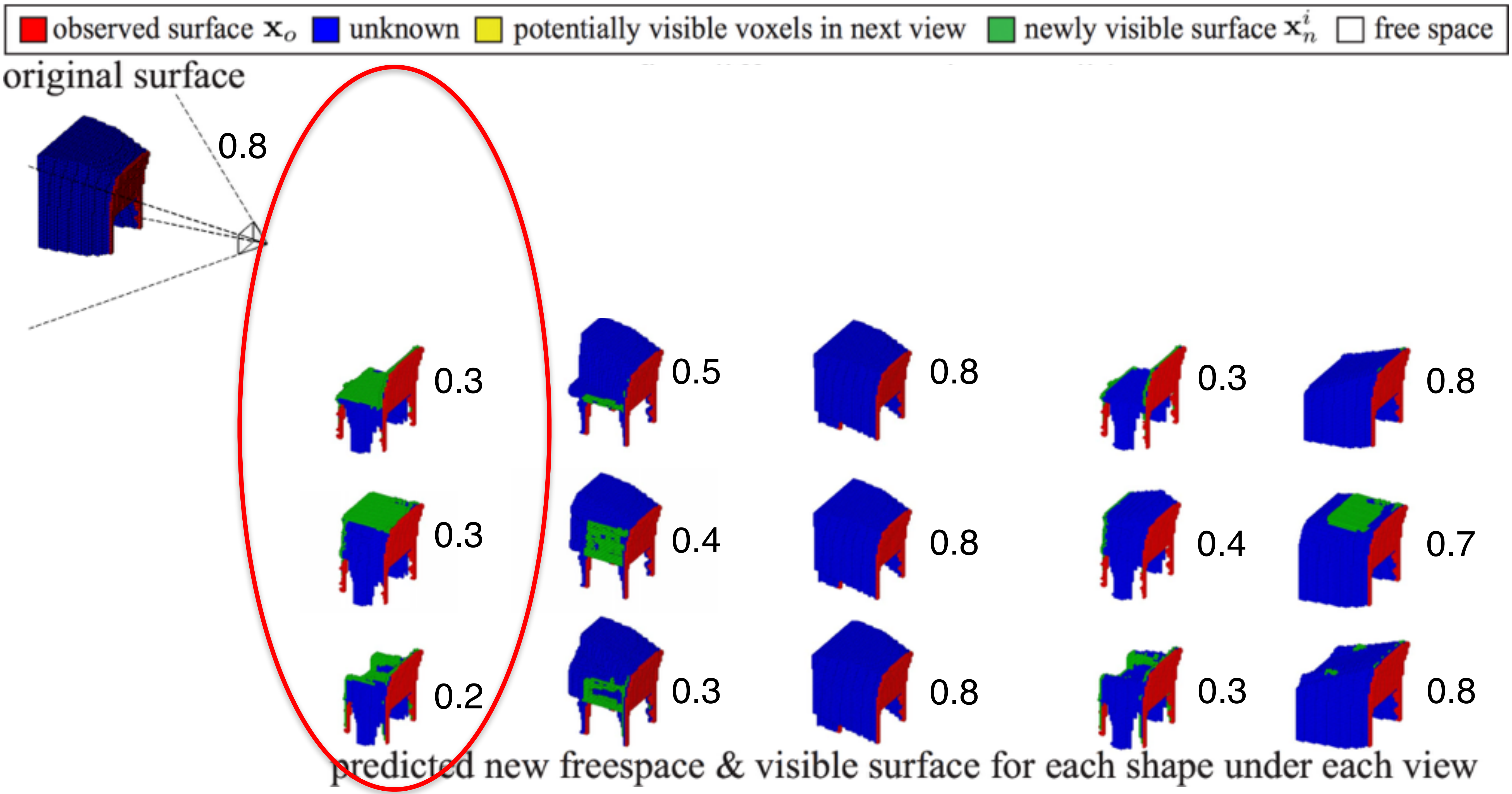
Slide Credit: Wu et al

Deep View Planning



Slide Credit: Wu et al

Deep View Planning



Mathematically, this is equivalent to evaluate the conditional mutual information:

$$I(y; \mathbf{x}_n^i | \mathbf{x}_o = x_o)$$

Slide Credit: Wu et al

Deep View Planning

	all
Ours	**0.80**
Max Visibility	0.78
Furthest Away	0.69
Random Selection	0.72

Recognition Accuracy from Two Views.

Based on the algorithms' choice, we obtain the actual depth map for the next view and recognize the objects using two views by our 3D ShapeNets.

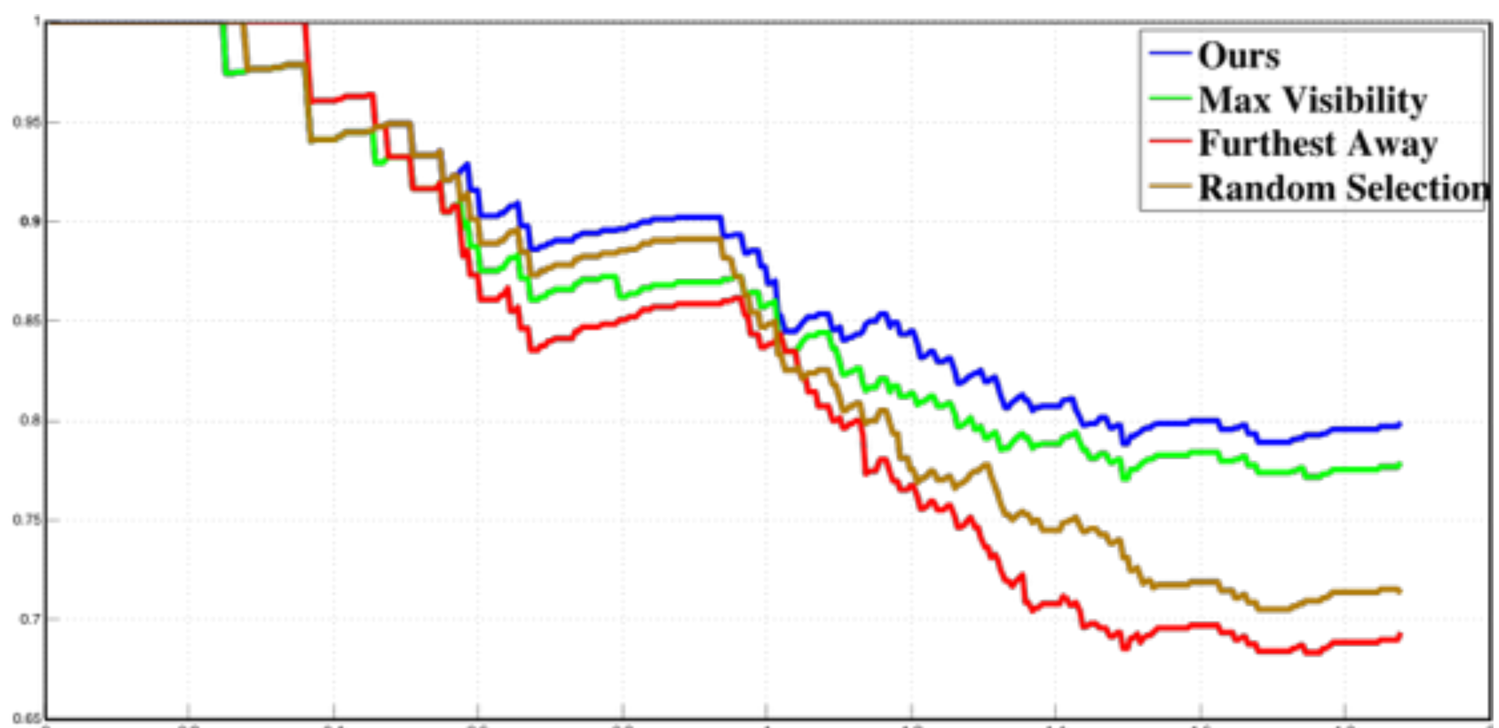
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Deep View Planning

	all
Ours	**0.80**
Max Visibility	0.78
Furthest Away	0.69
Random Selection	0.72

Recognition Accuracy from Two Views.

Based on the algorithms' choice, we obtain the actual depth map for the next view and recognize the objects using two views by our 3D ShapeNets.

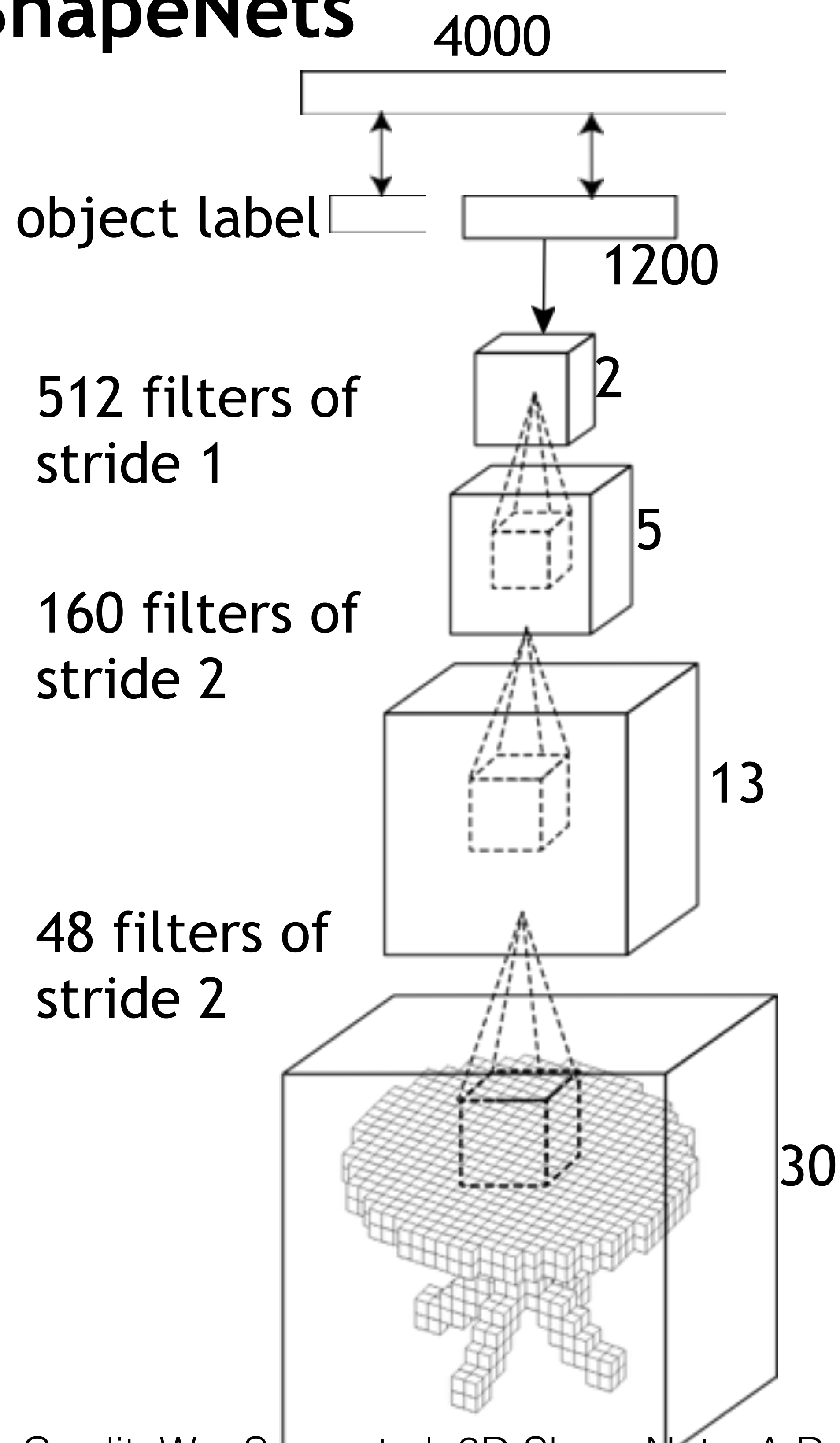


Our algorithm stands out as the uncertainty of the first view increases

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Back Propagation Fine-tuning

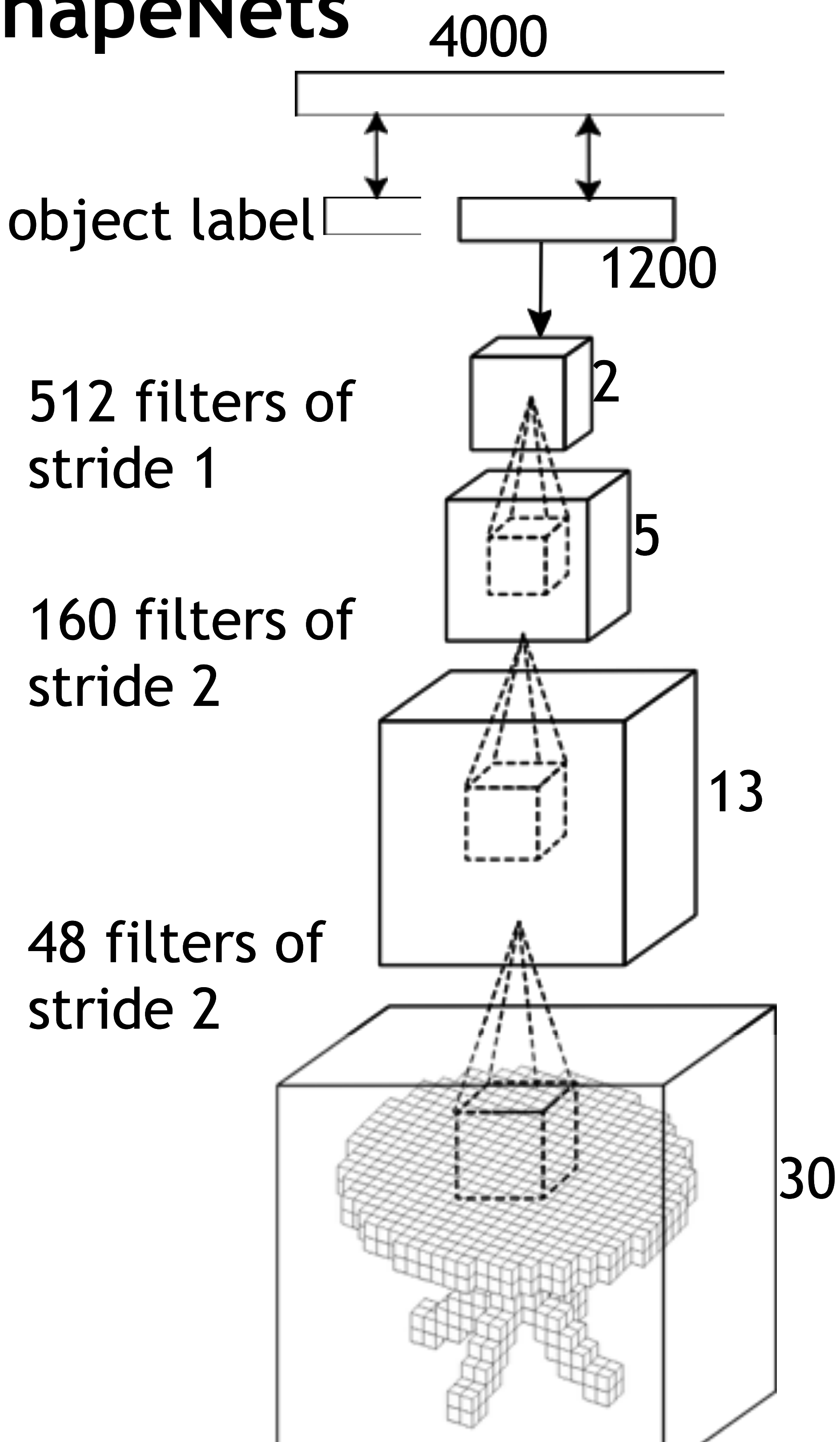
3D ShapeNets



Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Back Propagation Fine-tuning

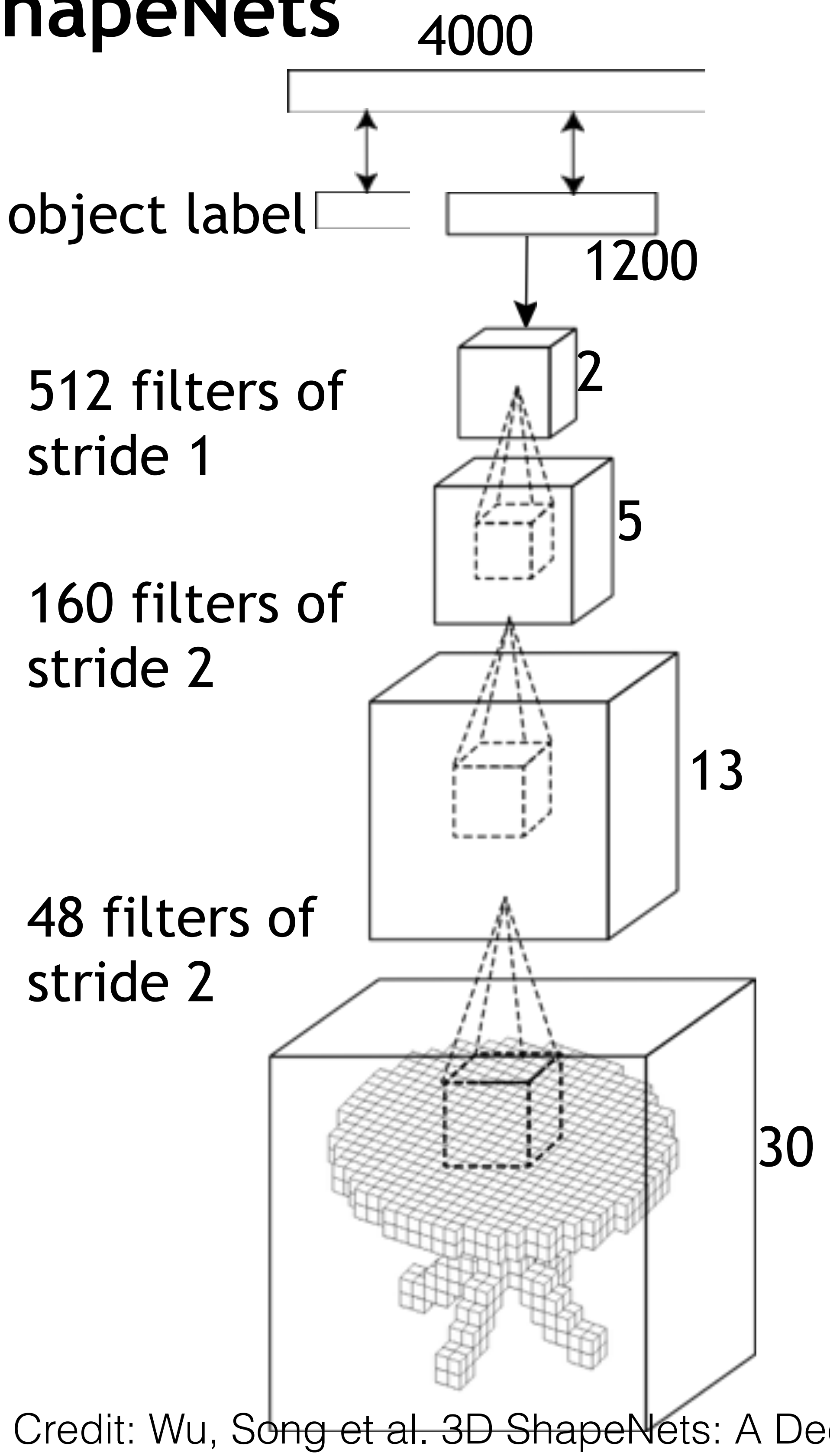
3D ShapeNets

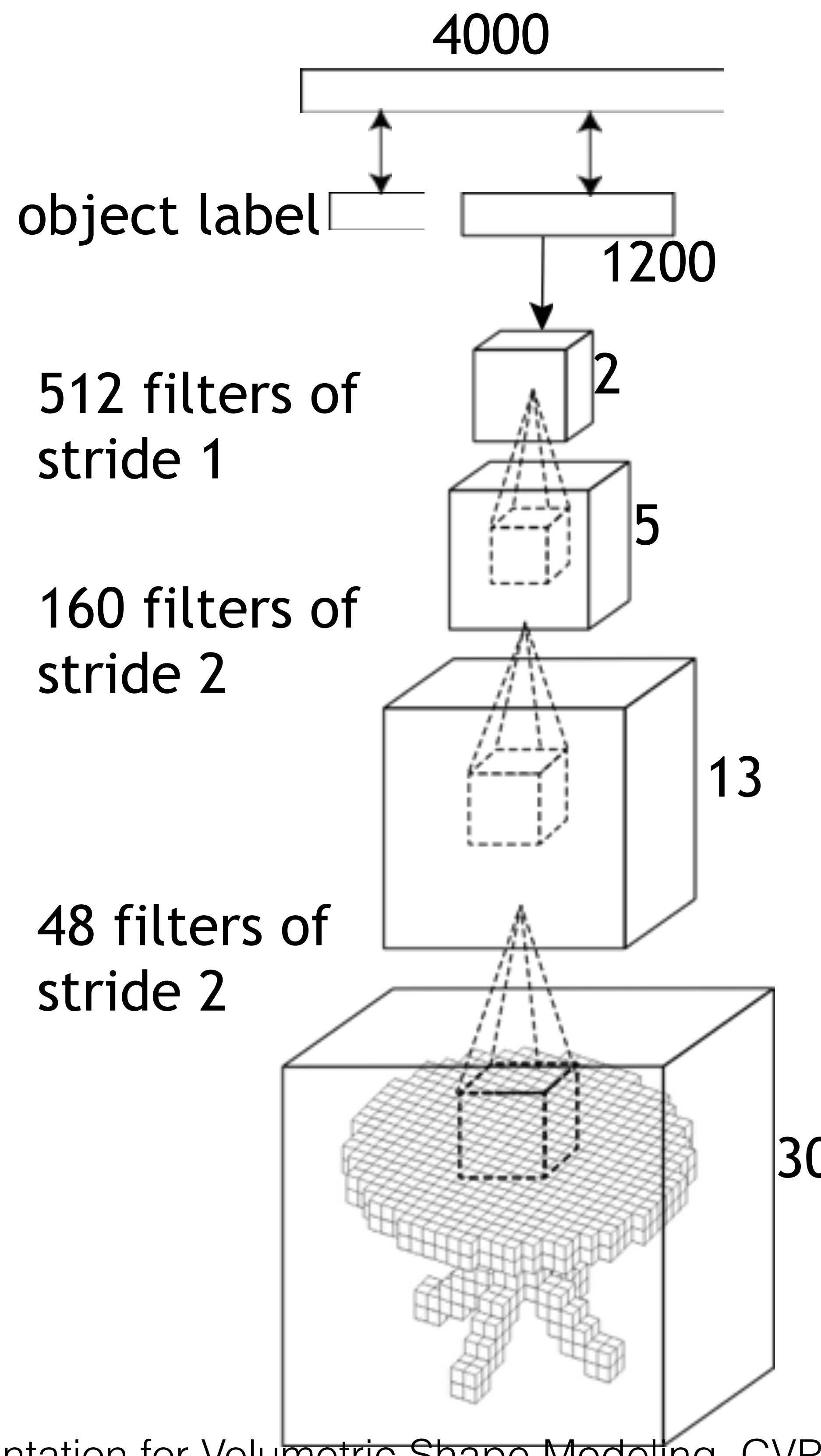


Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Back Propagation Fine-tuning

3D ShapeNets

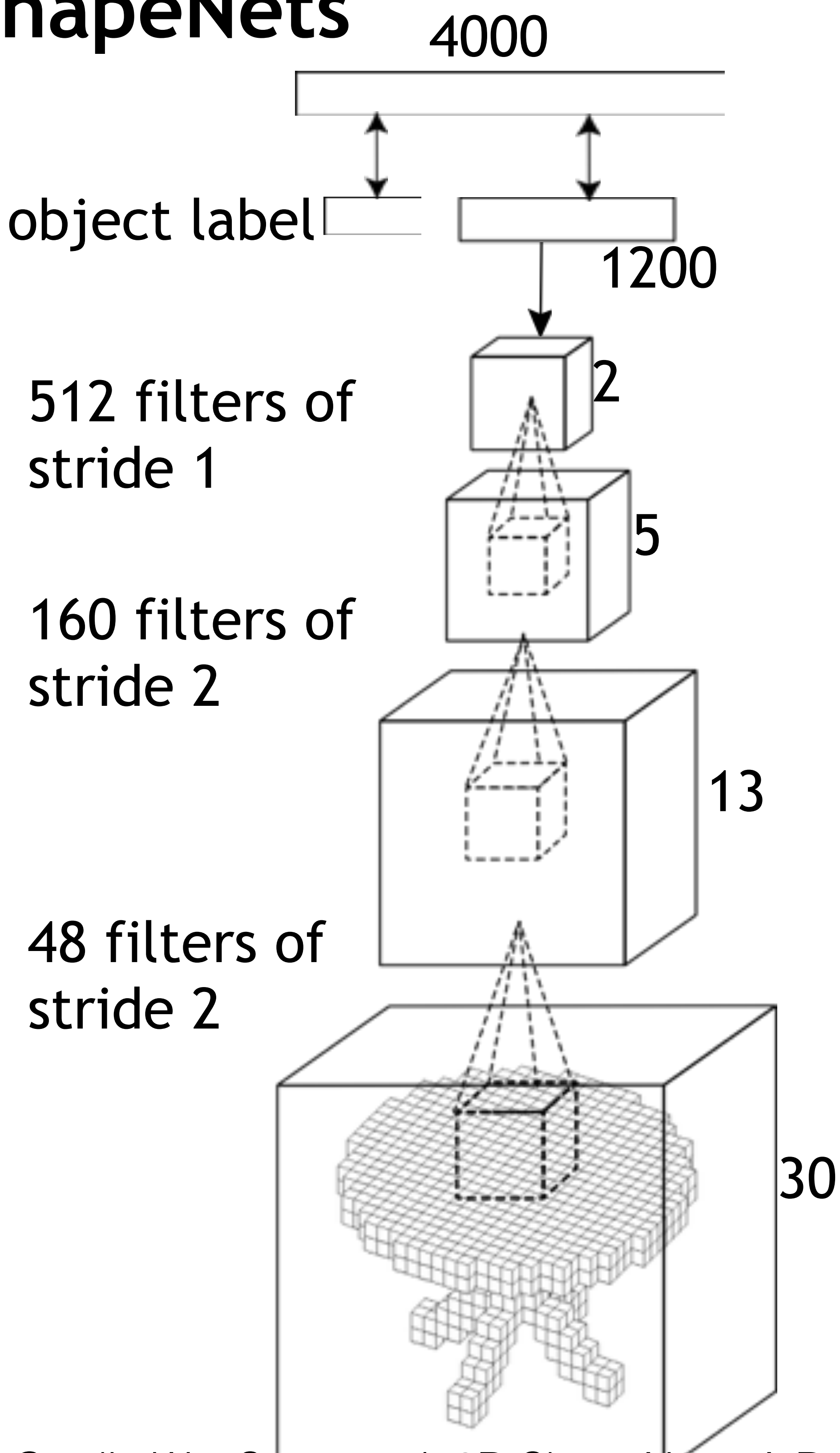


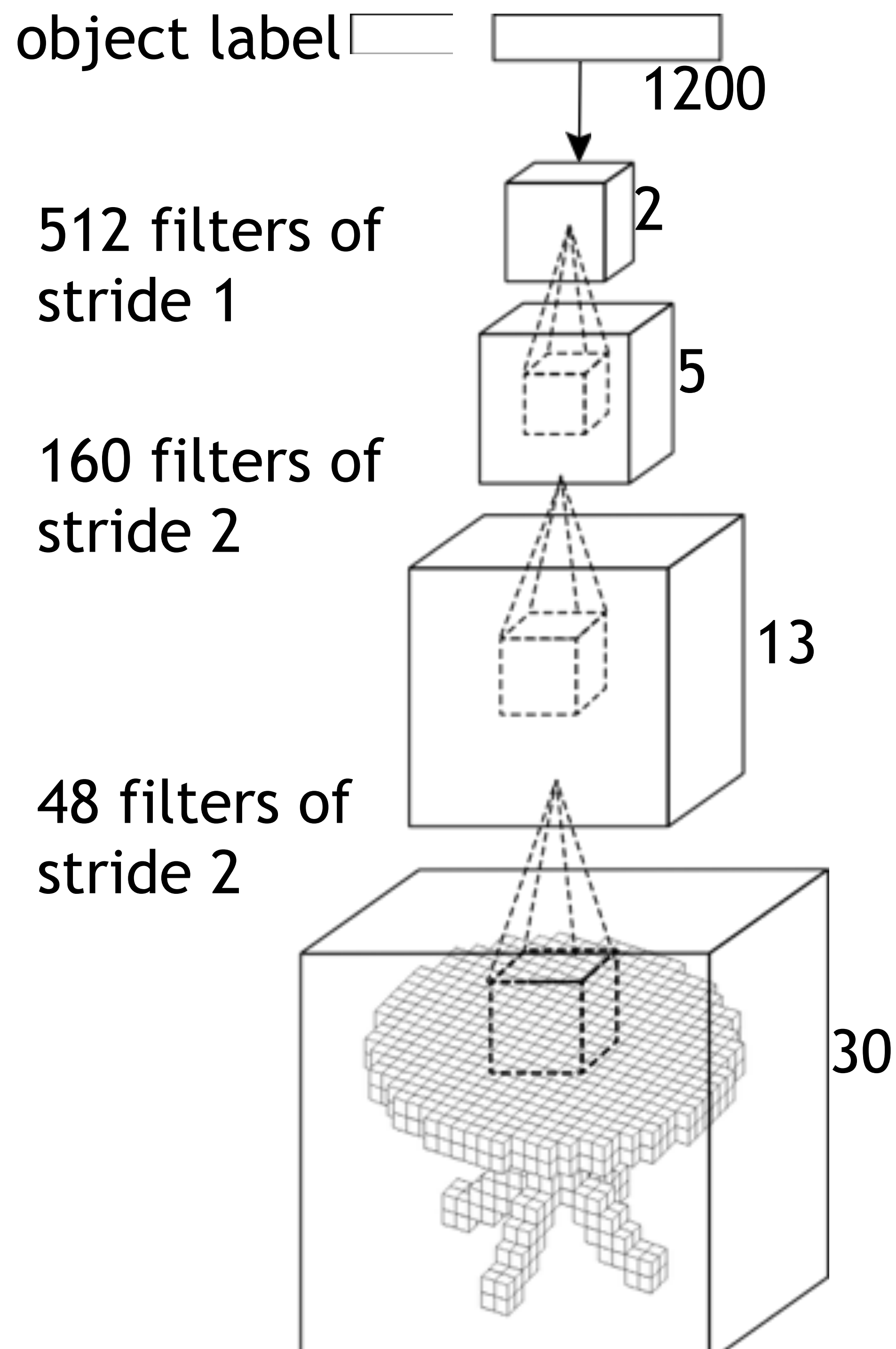


Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Back Propagation Fine-tuning

3D ShapeNets

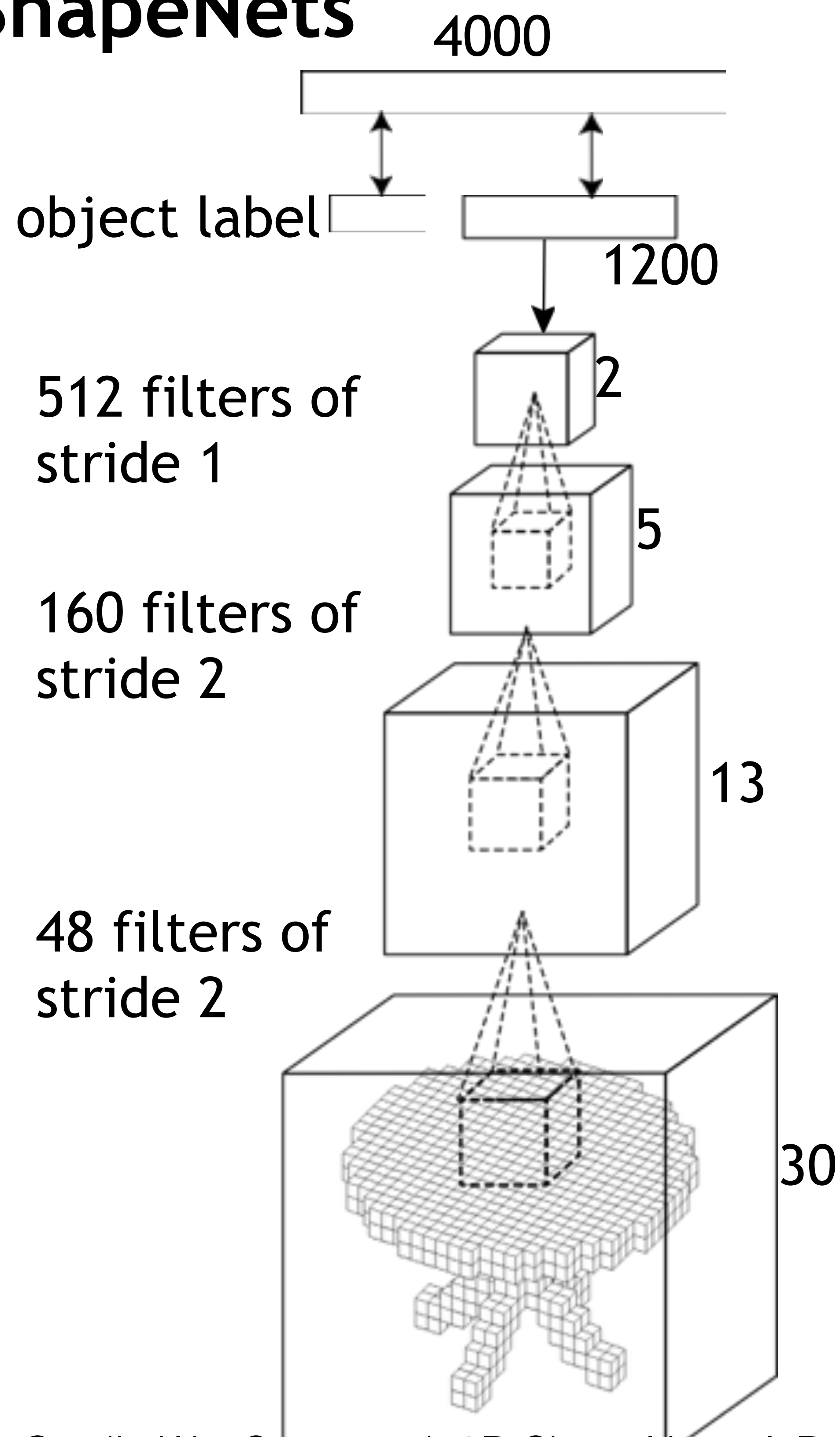


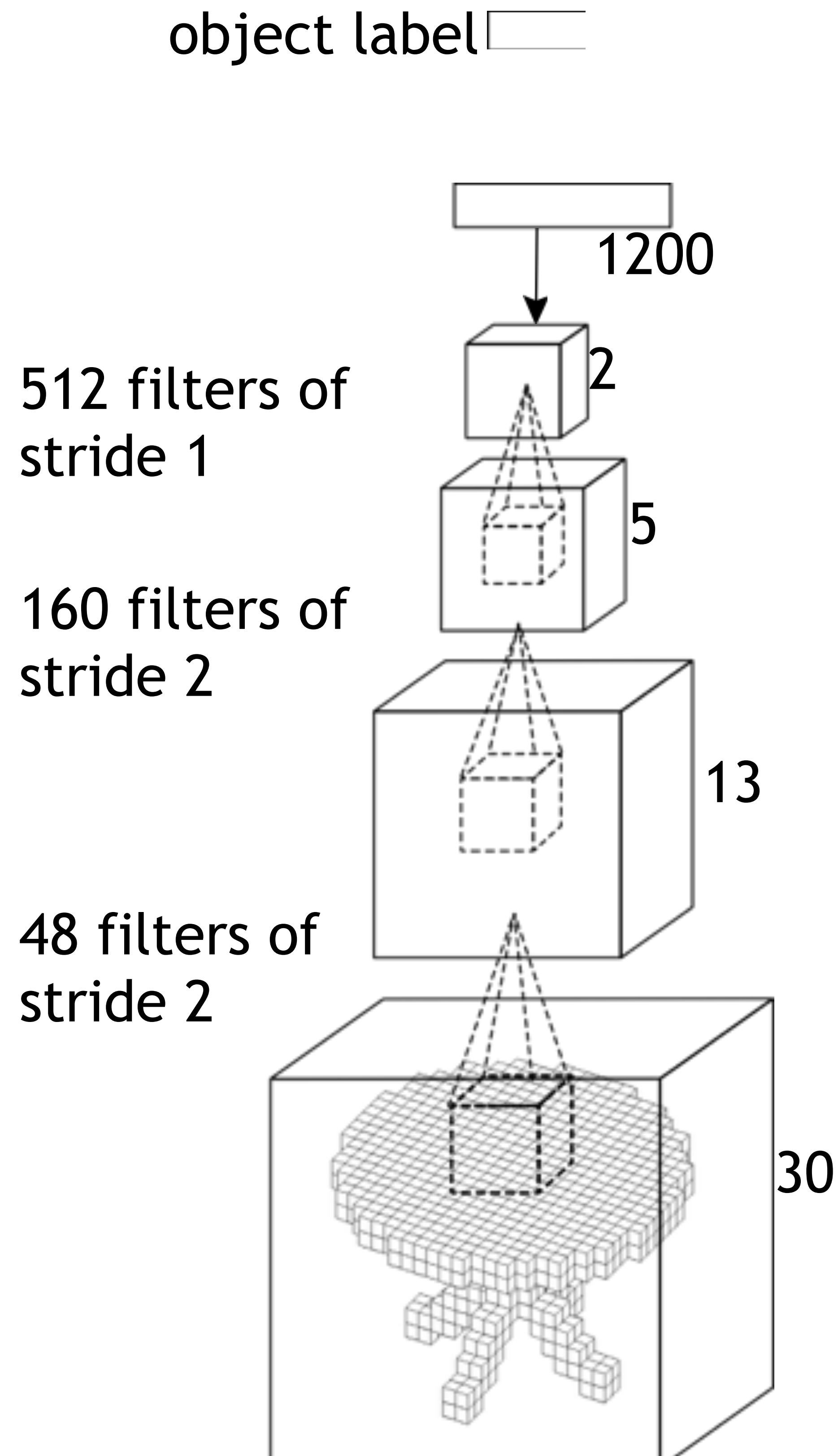


Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Back Propagation Fine-tuning

3D ShapeNets

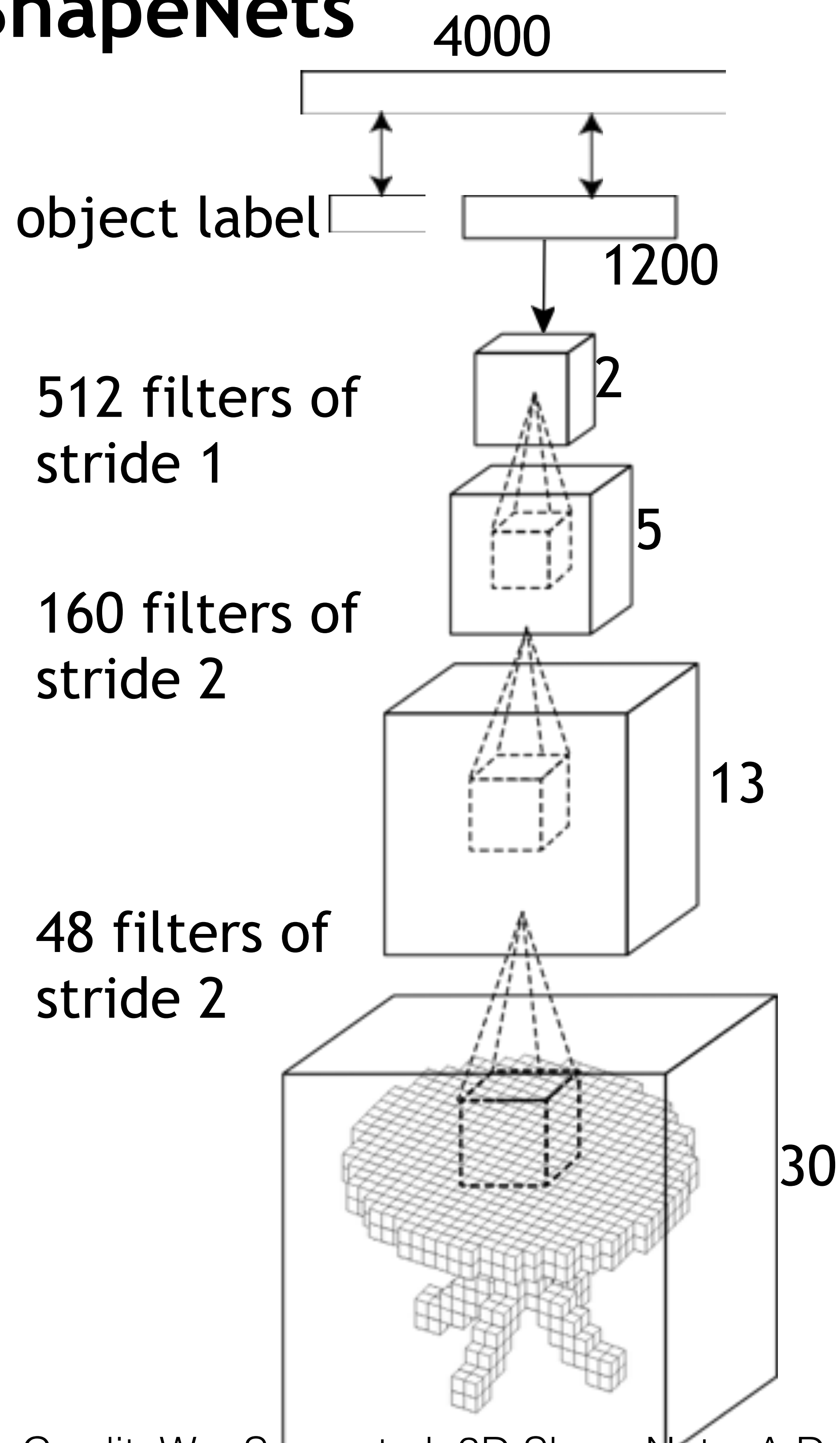


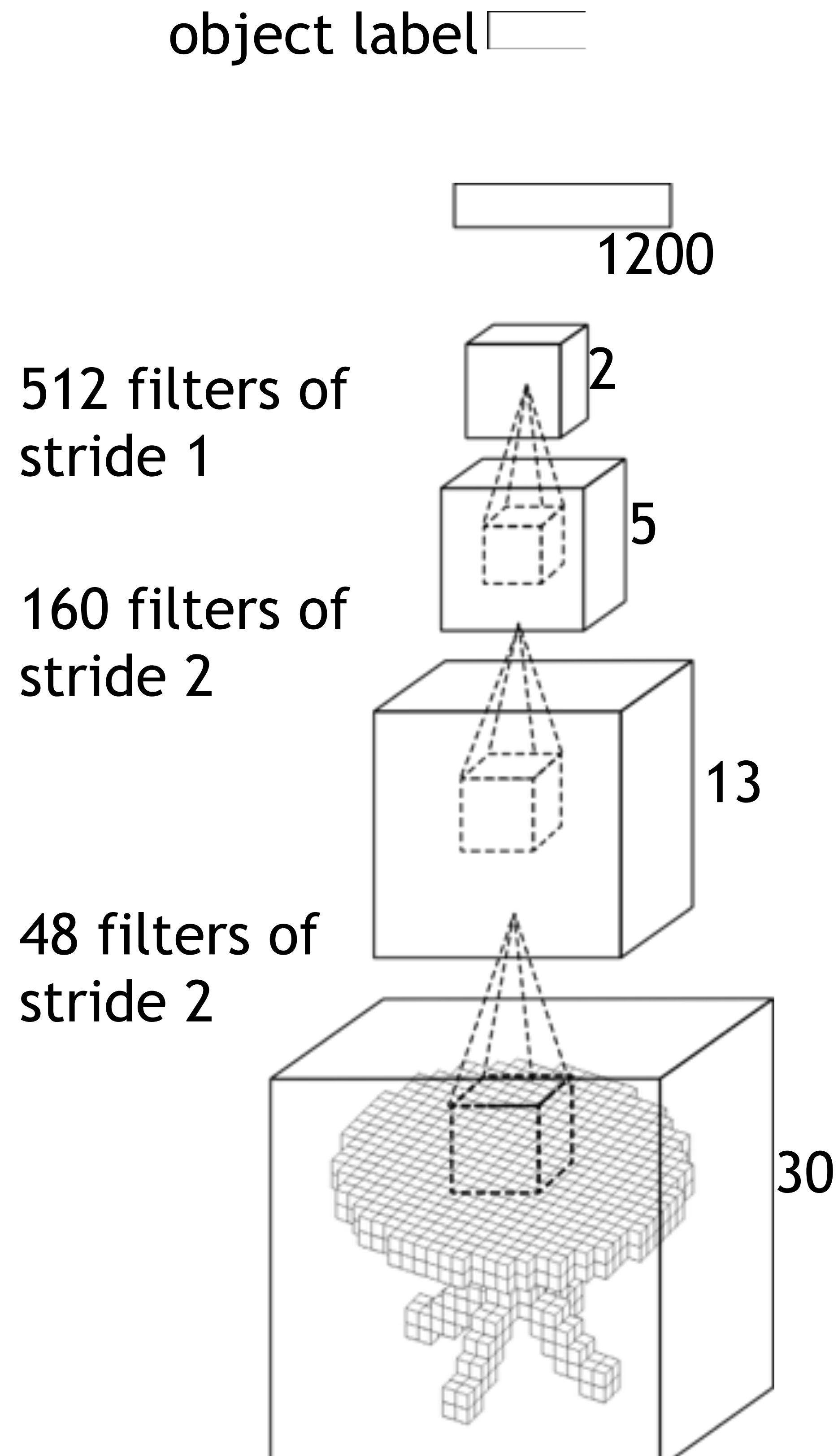


Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Back Propagation Fine-tuning

3D ShapeNets

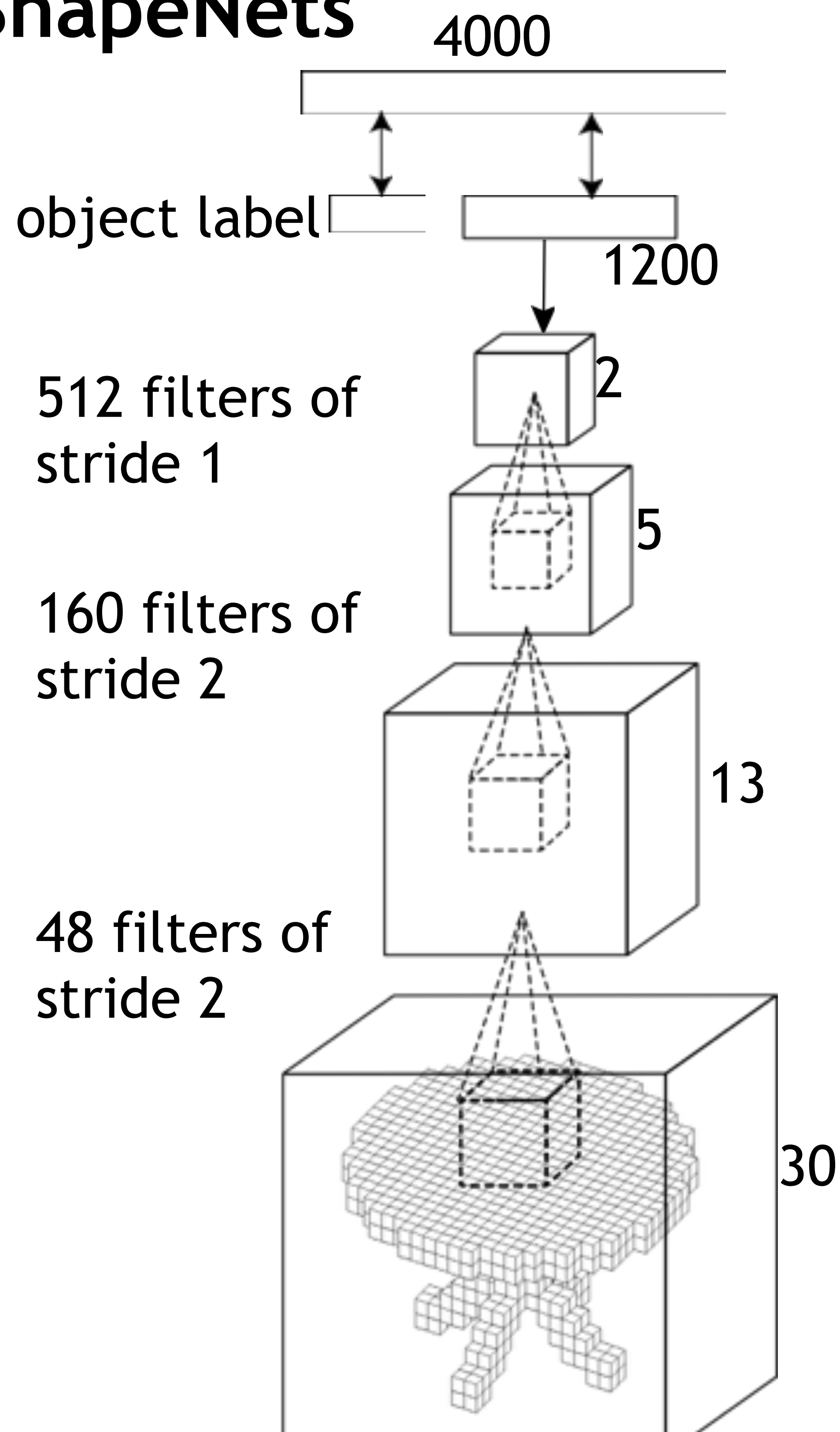


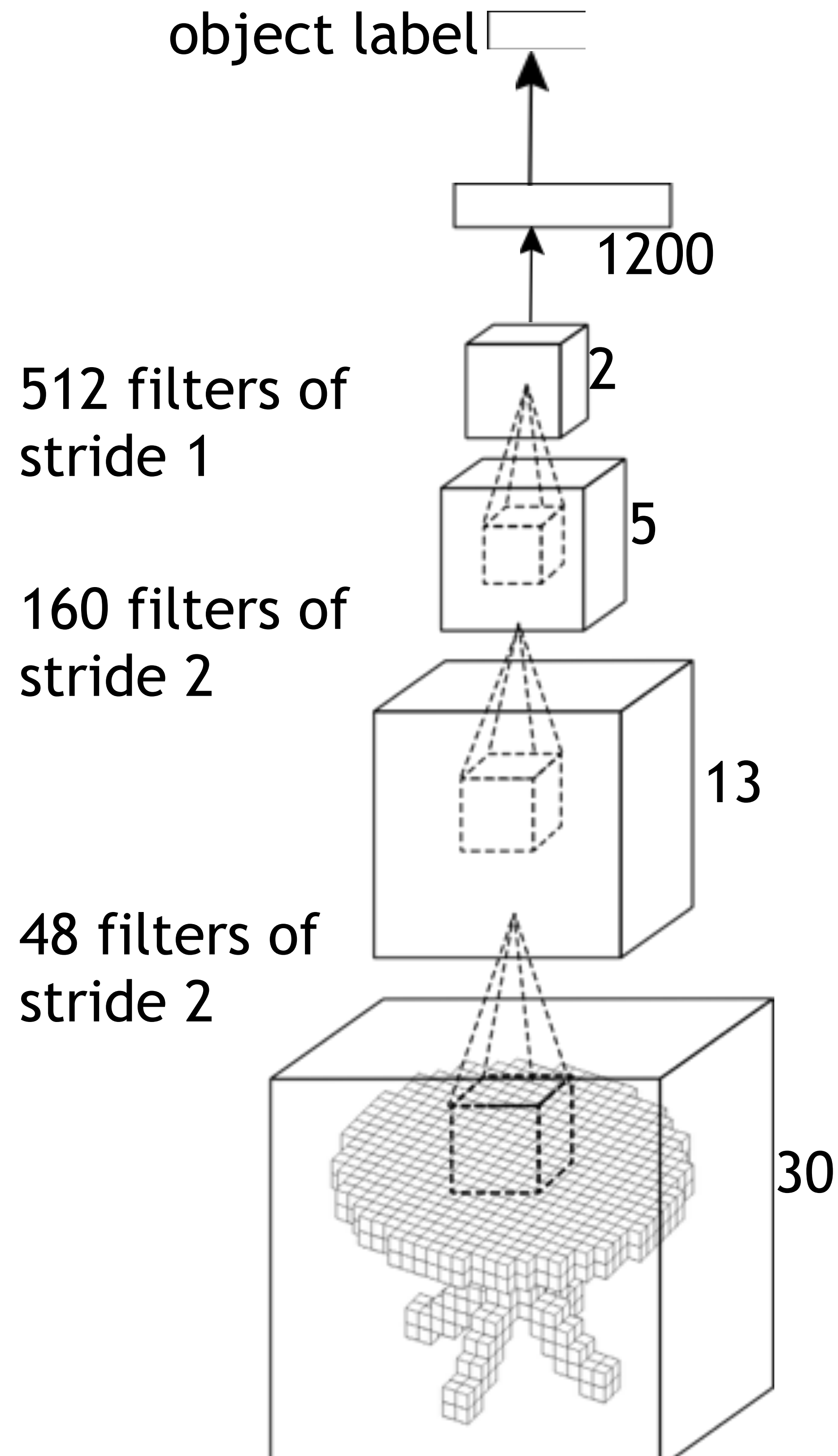


Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Back Propagation Fine-tuning

3D ShapeNets

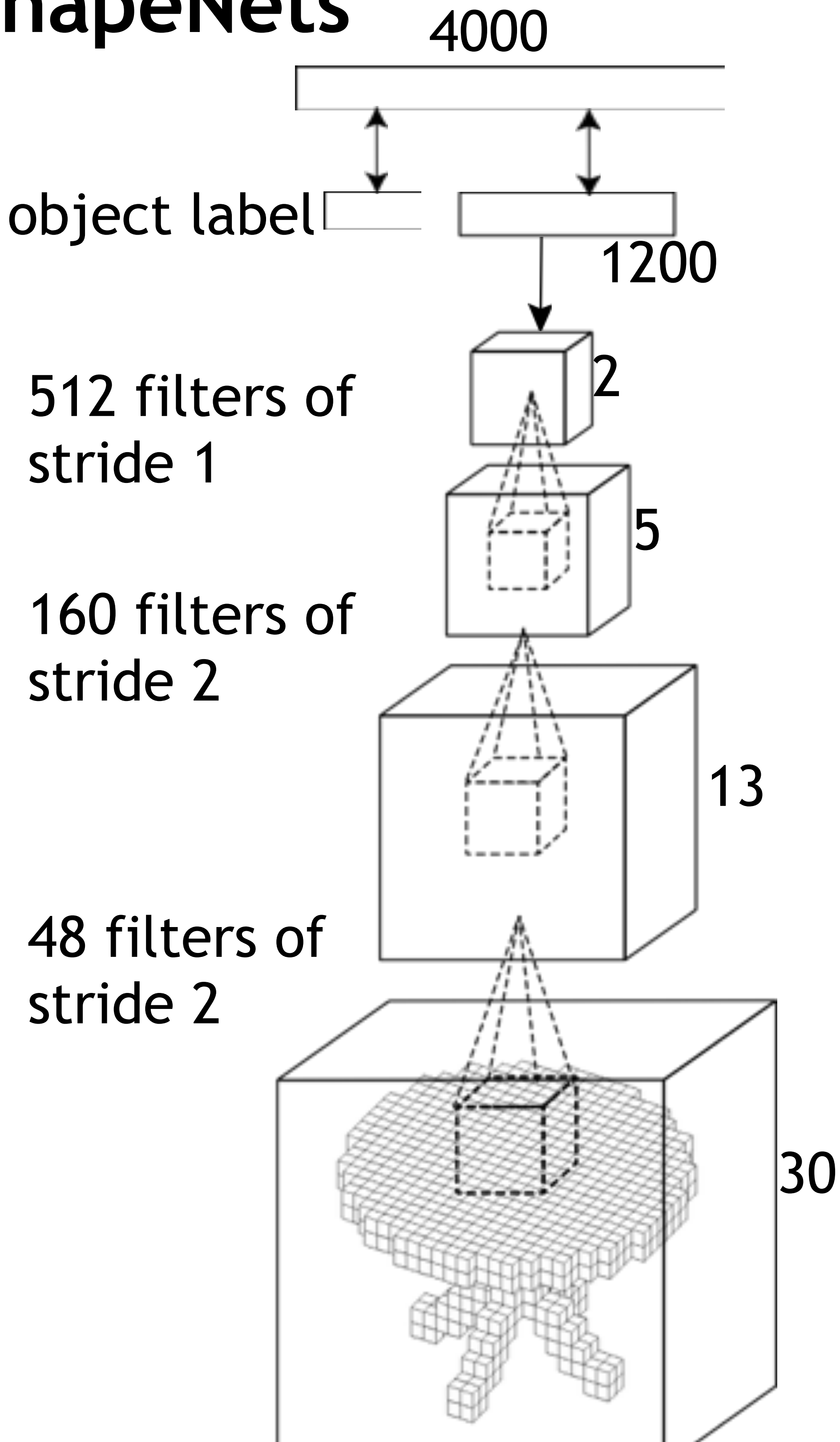




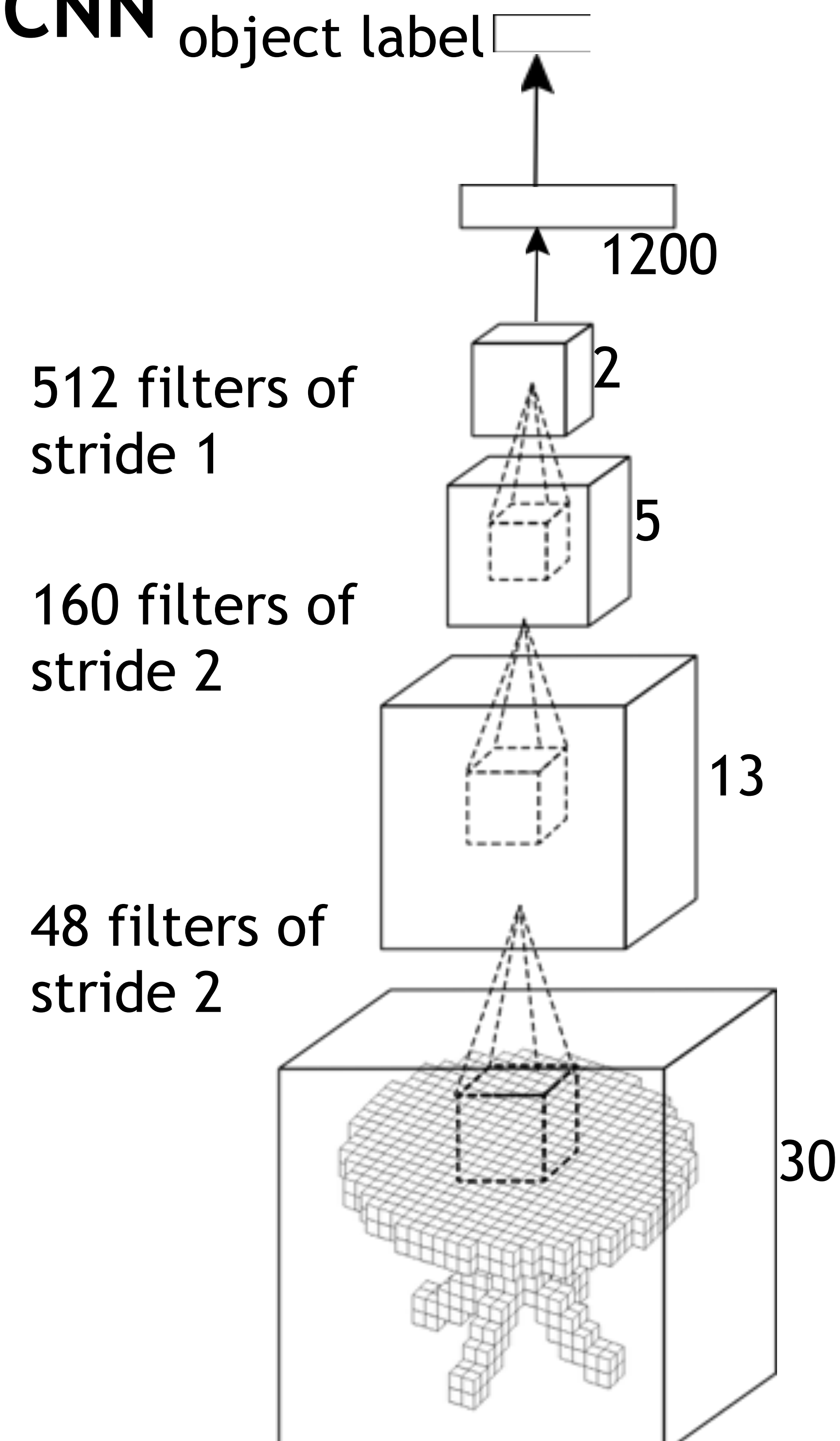
Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Back Propagation Fine-tuning

3D ShapeNets



3D CNN



Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

As a 3D Feature Extractor

Slide Credit: Wu et al

As a 3D Feature Extractor

Mesh Classification & Retrieval

10 classes	SPH [18]	LFD [8]	Ours
classification	79.79 %	79.87 %	**83.54%**
retrieval AUC	45.97%	51.70%	**69.28%**
retrieval MAP	44.05%	49.82%	**68.26%**
40 classes	SPH [18]	LFD [8]	Ours
classification	68.23%	75.47%	**77.32%**
retrieval AUC	34.47%	42.04%	**49.94%**
retrieval MAP	33.26%	40.91%	**49.23%**

Slide Credit: Wu et al

As a 3D Feature Extractor

Mesh Classification & Retrieval

10 classes	SPH [18]	LFD [8]	Ours
classification	79.79 %	79.87 %	**83.54%**
retrieval AUC	45.97%	51.70%	**69.28%**
retrieval MAP	44.05%	49.82%	**68.26%**
40 classes	SPH [18]	LFD [8]	Ours
classification	68.23%	75.47%	**77.32%**
retrieval AUC	34.47%	42.04%	**49.94%**
retrieval MAP	33.26%	40.91%	**49.23%**

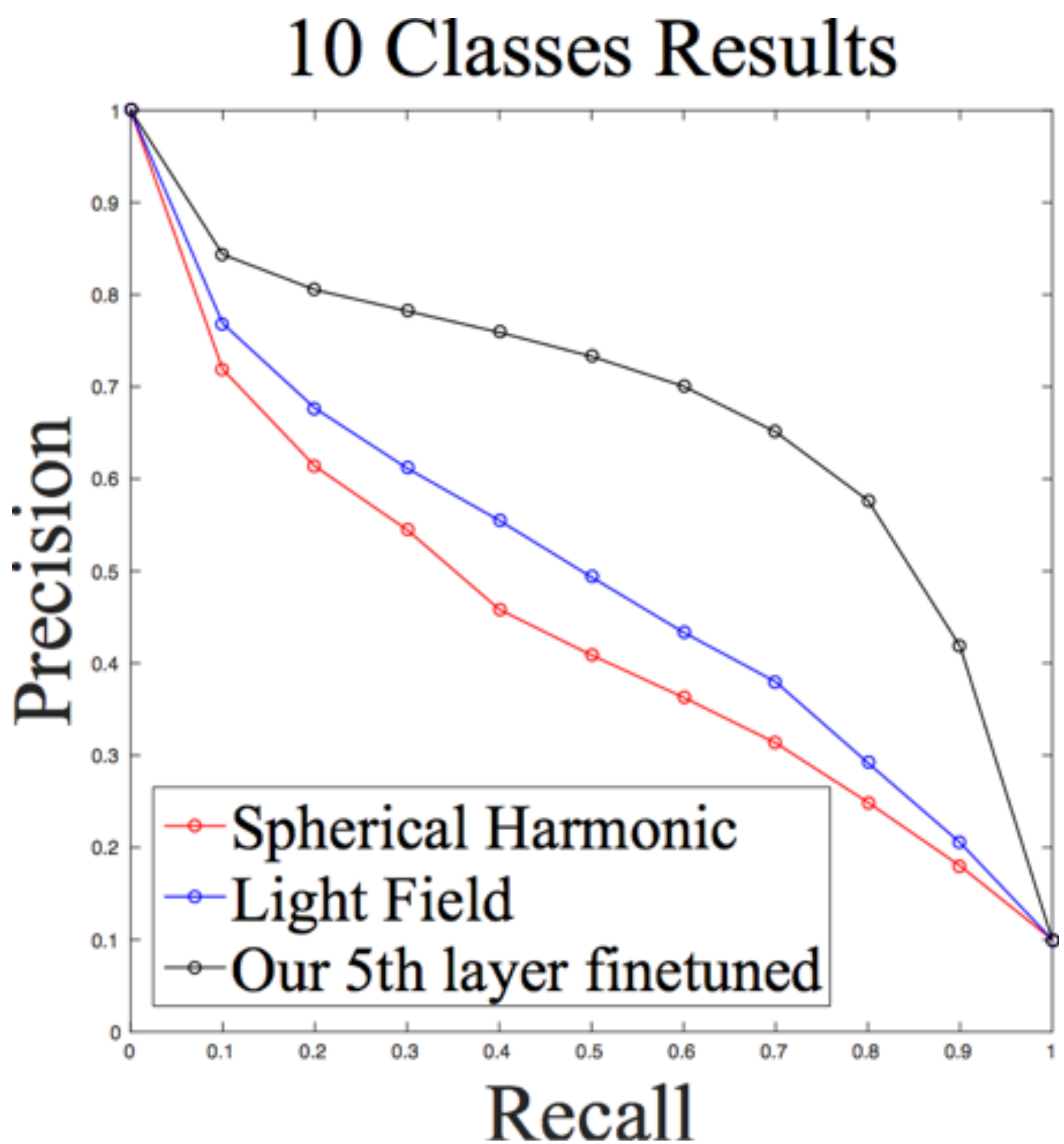
2.5D object recognition

	all
[29] Depth	0.376
NN	0.374
ICP	0.471
3D ShapeNets	0.437
3D ShapeNets fine-tuned	**0.579**
[29] RGB	0.334
[29] RGBD	0.448

[29] R. Socher, B. Huval, B. Bhat, C. D. Manning, and A. Y. Ng. Convolutional-recursive deep learning for 3d object classification. In NIPS 2012.

Slide Credit: Wu et al

As a 3D Feature Extractor



mesh retrieval

Slide Credit: Wu, Song et al. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling, CVPR 2015

Extensions

- Include RGB information in representation
- 3D Segmentation
- Improve for non-rigid 3D objects

Discussion Points

- Is the network deep enough?
    - 30x30x30 = 27000 vs 256x256 = 65000 for Image Net
    - 150K training examples vs millions for Image Net

- Won't removal of max-pooling layers hurt performance on classification tasks?

Algorithm	ModelNet40 Classification (Accuracy)	ModelNet40 Retrieval (mAP)	ModelNet10 Classification (Accuracy)	ModelNet10 Retrieval (mAP)
MVCNN [3]	90.1%	79.5%		
VoxNet [2]	83%		92%	
DeepPano [4]	77.63%	76.81%	85.45%	84.18%
3DShapeNets [1]	77%	49.2%	83.5%	68.3%

[1] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang and J. Xiao. 3D ShapeNets: A Deep Representation for Volumetric Shapes. CVPR2015.
[2] D. Maturana and S. Scherer. VoxNet: A 3D Convolutional Neural Network for Real-Time Object Recognition. IROS2015.
[3] H. Su, S. Maji, E. Kalogerakis, E. Learned-Miller. Multi-view Convolutional Neural Networks for 3D Shape Recognition. ICCV2015.
[4] B Shi, S Bai, Z Zhou, X Bai. DeepPano: Deep Panoramic Representation for 3-D Shape Recognition. Signal Processing Letters 2015.

http://3dshapenets.cs.princeton.edu/

- Any other systems that use binary units with approximate training and inference techniques rather than standard back-prop?
    - Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh. "A fast learning algorithm for deep belief nets." Neural computation 18.7 (2006): 1527-1554
    - Salakhutdinov, Ruslan, Andriy Mnih, and Geoffrey Hinton. "Restricted Boltzmann machines for collaborative filtering." Proceedings of the 24th international conference on Machine learning. ACM, 2007.

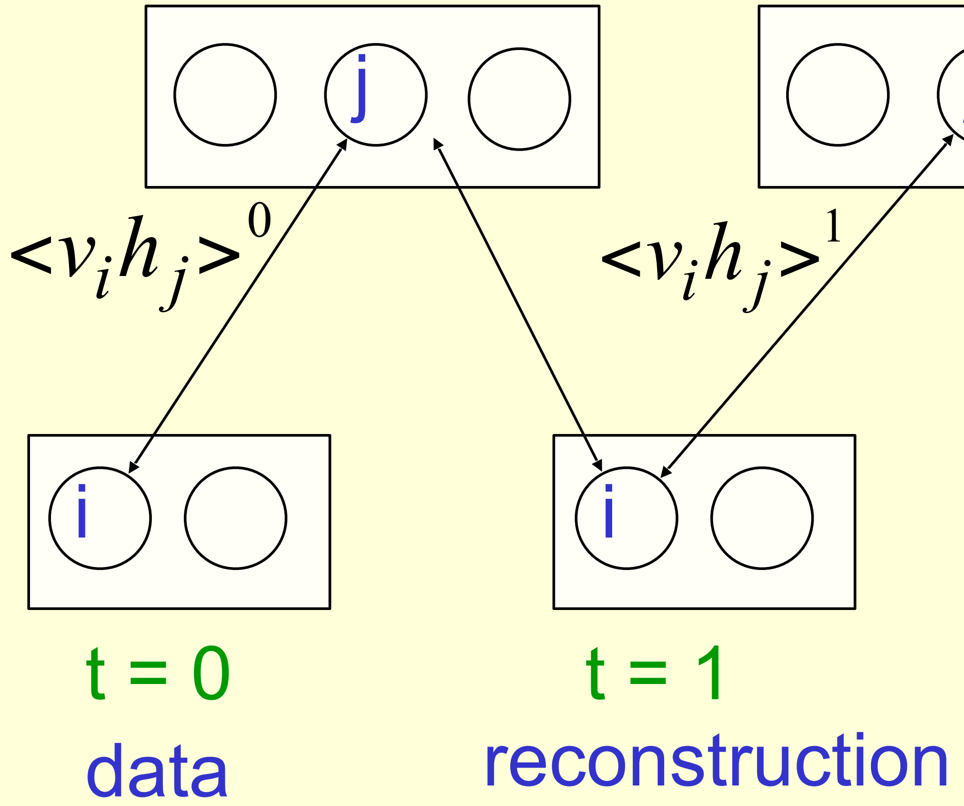
- Are there better ways for representing 3D Shapes. In particular, doesn’t the voxel representation have the bottleneck of cubic dependency on grid size?
    - Yes. Su, Majhi et al that tries to recognize 3D shapes from multiple 2D views instead of voxel representation and get better results for classification .

H. Su, S. Maji, E. Kalogerakis, E. Learned-Miller. Multi-view Convolutional Neural Networks for 3D Shape Recognition. ICCV2015.

- Are there other 3D CAD model datasets
    - 3D Warehouse. https://3dwarehouse.sketchup.com/
- Manually removing clutter from 3D CAD models a problem
- Did not address non-rigid objects sufficiently.
    - Even the 40 model classification dataset seemed to contain only 4 non-rigid categories — persons, plant, sofas, curtains.

Appendix

Contrastive divergence learning: A quick way to learn an RBM



Start with a training vector on the visible units.

Update all the hidden units in parallel

Update all the visible units in parallel to get a "reconstruction".

Update all the hidden units again.

$$\Delta w_{ij} = \varepsilon \, (<v_i h_j>^0 - <v_i h_j>^1)$$

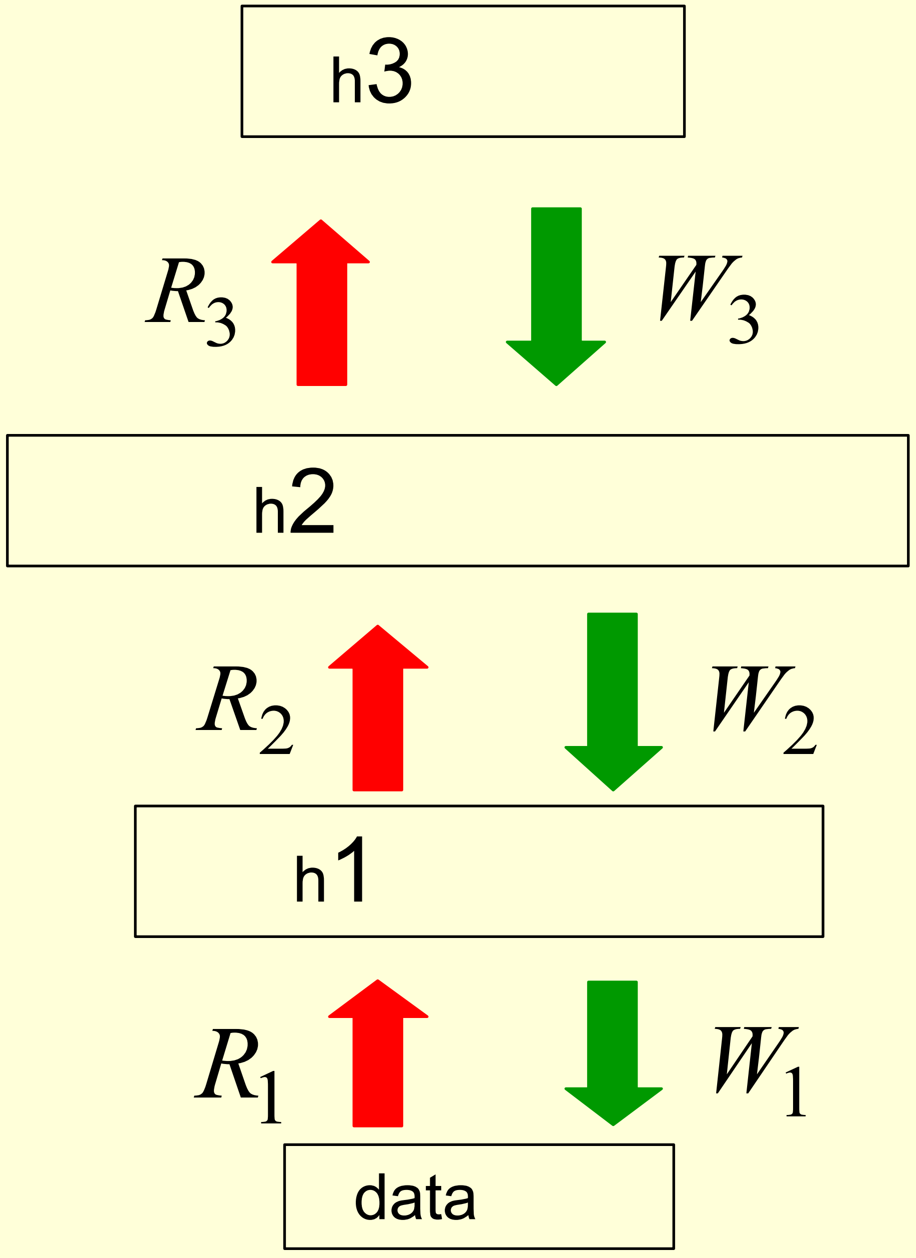
This is not following the gradient of the log likelihood. But it works well.

It is approximately following the gradient of another objective function.

Slide Credit: Geoffery Hinton

The wake-sleep algorithm for an SBN

- Wake phase: Use the recognition weights to perform a bottom-up pass.
  - Train the generative weights to reconstruct activities in each layer from the layer above.
- Sleep phase: Use the generative weights to generate samples from the model.
  - Train the recognition weights to reconstruct activities in each layer from the layer below.



Slide Credit: Geoffery Hinton