3D ShapeNets: A Deep Representation for Volumetric Shapes

Zhirong Wu[†⋆] Shuran Song[†] Aditya Khosla[‡] Fisher Yu[†] Linguang Zhang[†] Xiaoou Tang[⋆] Jianxiong Xiao[†]
[†]Princeton University [⋆]Chinese University of Hong Kong [‡]Massachusetts Institute of Technology

In this supplementary material, we first show detailed analysis of 3D mesh classification by computing the confusion matrix. Then we give a lot more results of shape generation and view-based shape completion. Finally we show additional evaluations of the Next-Best-View experiment in the main submission.

1. 3D Shape Classification

From the main paper, the improvement over baselines for the 40-category experiment is smaller than the improvement over baselines for the 10-category experiment. If we look at the confusion matrix in Figure 2, we can see that it's likely due to that the fact that some objects in 40 categories are just too difficult to be distinguished solely from shapes. For example, "cone" and "tent", "chair" and "bench" , "plant" and "flower pot", "wardrobe" and "dresser" are all very similar in term of 3D shapes. It's the appearance or the context that makes them different. The tasks using 40-categories are a lot harder not only because of more categories, but also because the new categories are more difficult to distinguish. This makes the performance upper bound much lower than the 10-category experiment.

2. 3D Shape Generation

To show our model captures the complex 3D shapes of all classes in arbitrary poses, we generate samples from the model and visualize some examples each class per row. We show the results in **Page 2-5** of this document.

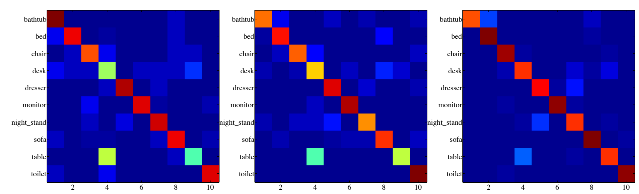
3. View-based Shape Completion

We also give more results on shape completions. For each class, we render 3 instances and give 2 possible completions for each rendering. We show the results in **Page 6-9** of this document.

4. Additional Next-Best-View Evaluation

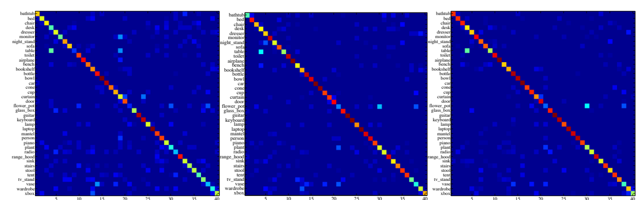
Here we show additional evaluations of the Next-Best-View experiment in the main submission.

Ideally, if the recognition of the original view is confident, adding another view will not make any difference. Therefore, the evaluation of view planning strategies will be more obvious if we look at more difficult cases. In Figure 3, when the uncertainty of recognition is very small, all of the methods achieve almost the same accuracy. However, as uncertainty goes up, bad strategies suffer greatly while our method achieves the best.



(a) SPH (b) LFD (c) Ours

Figure 1: **Confusion matrix** for the 10-category experiment in the main paper.



(a) SPH (b) LFD (c) Ours

Figure 2: **Confusion matrix** for the 40-category experiment.

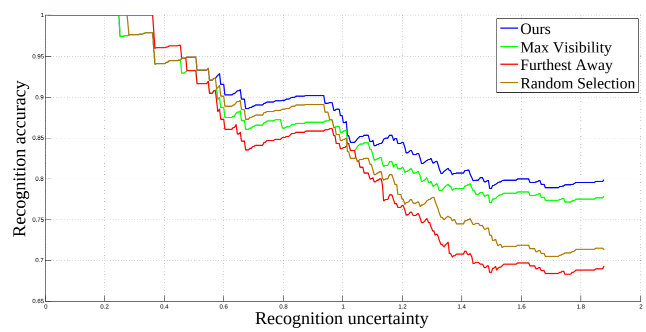
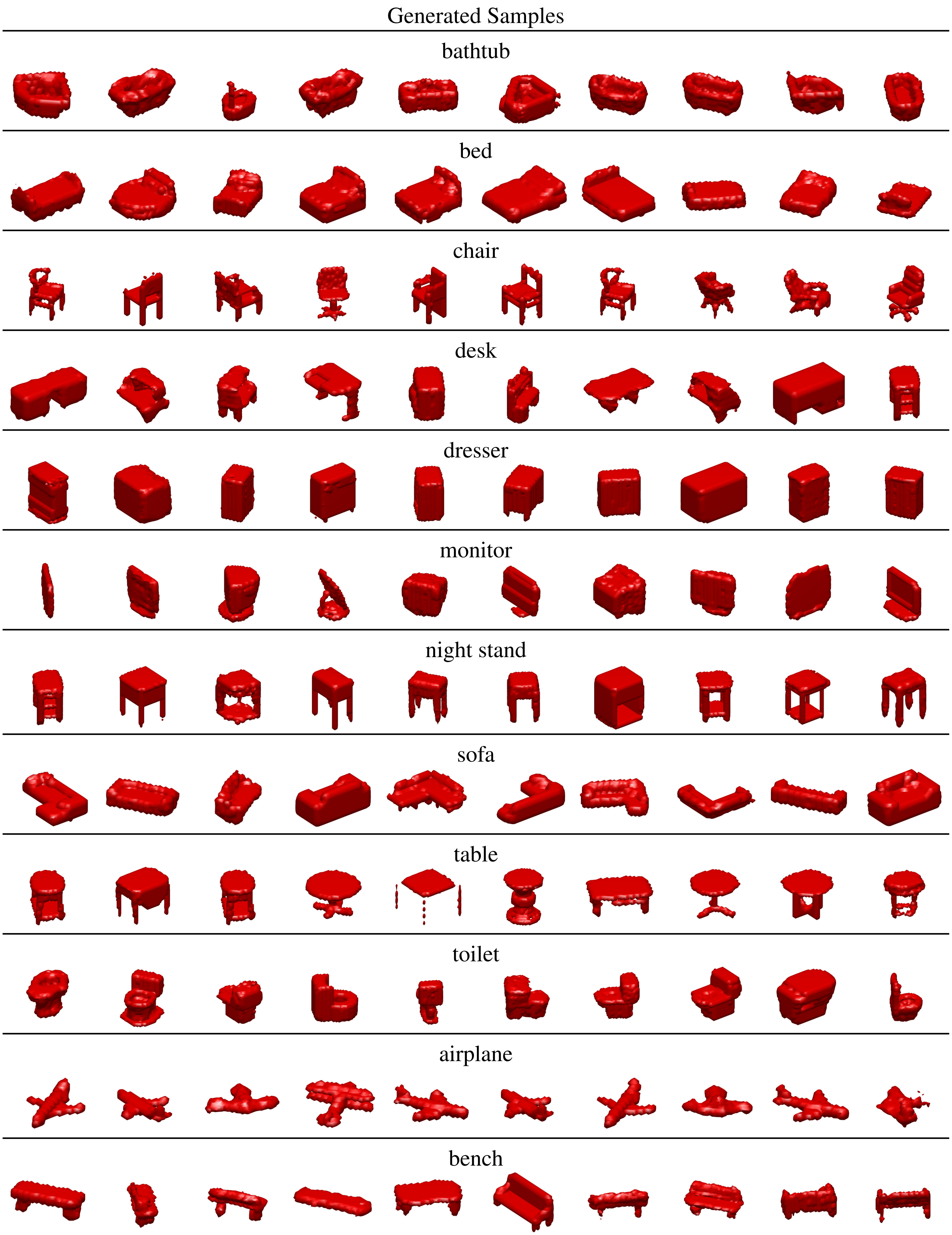
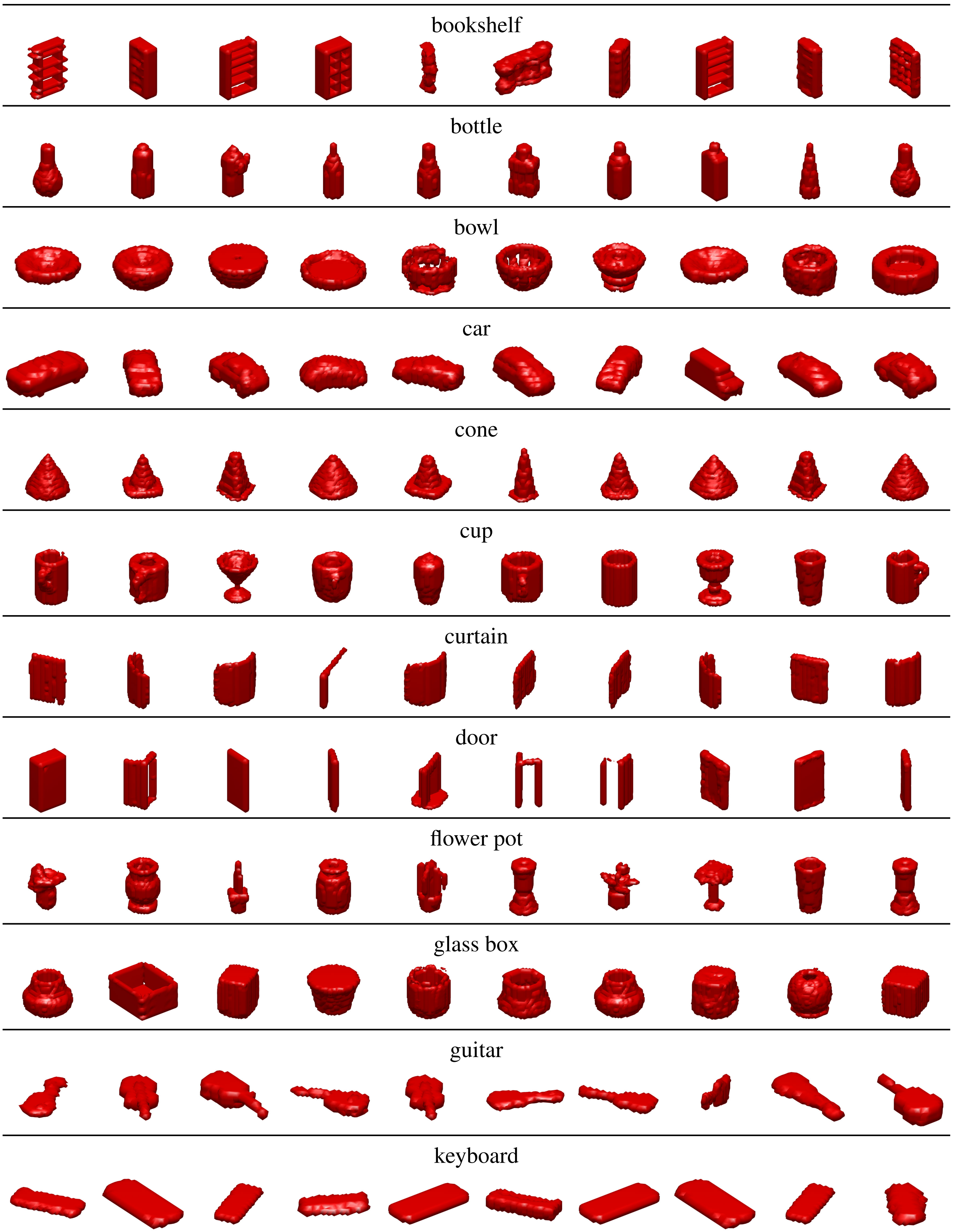
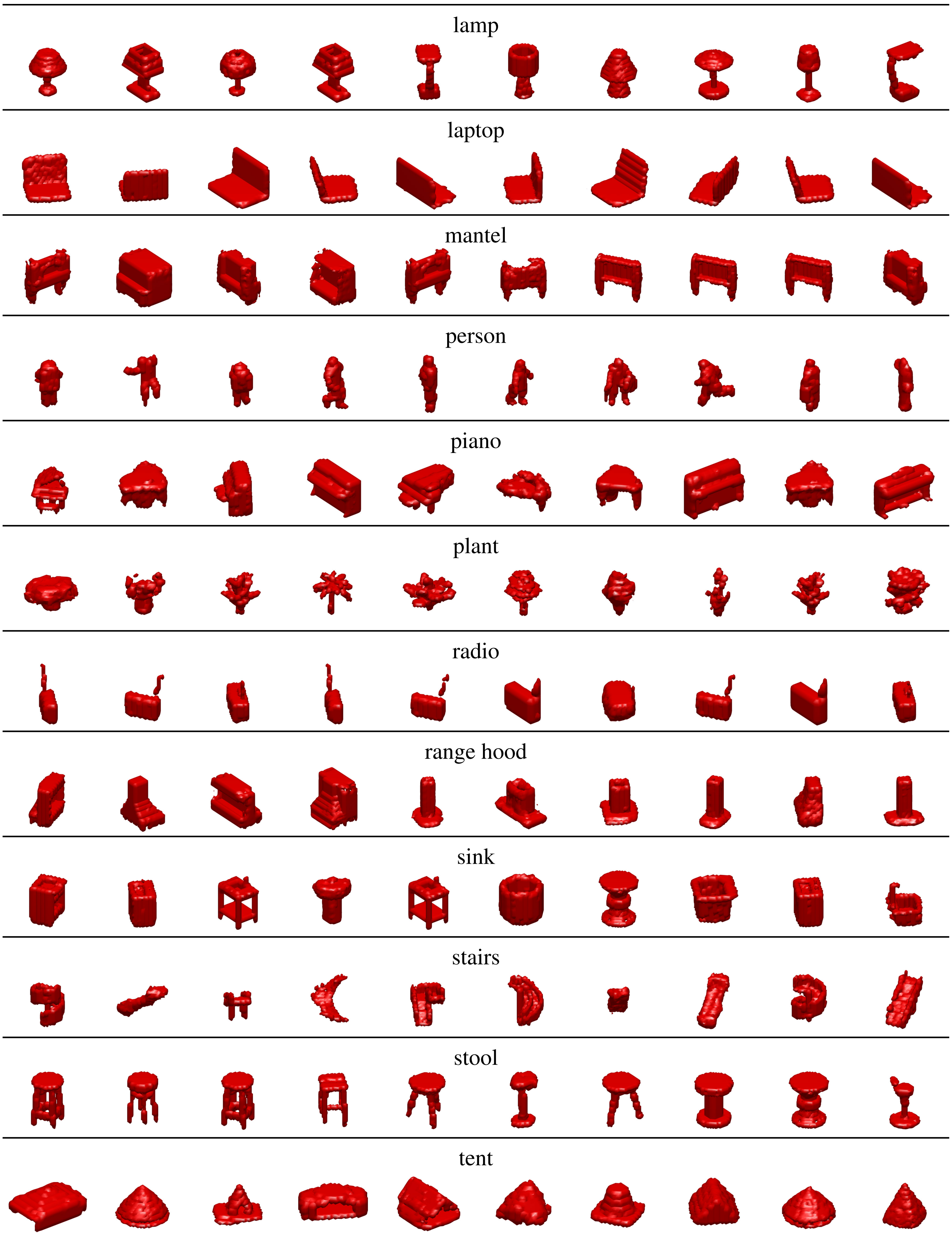
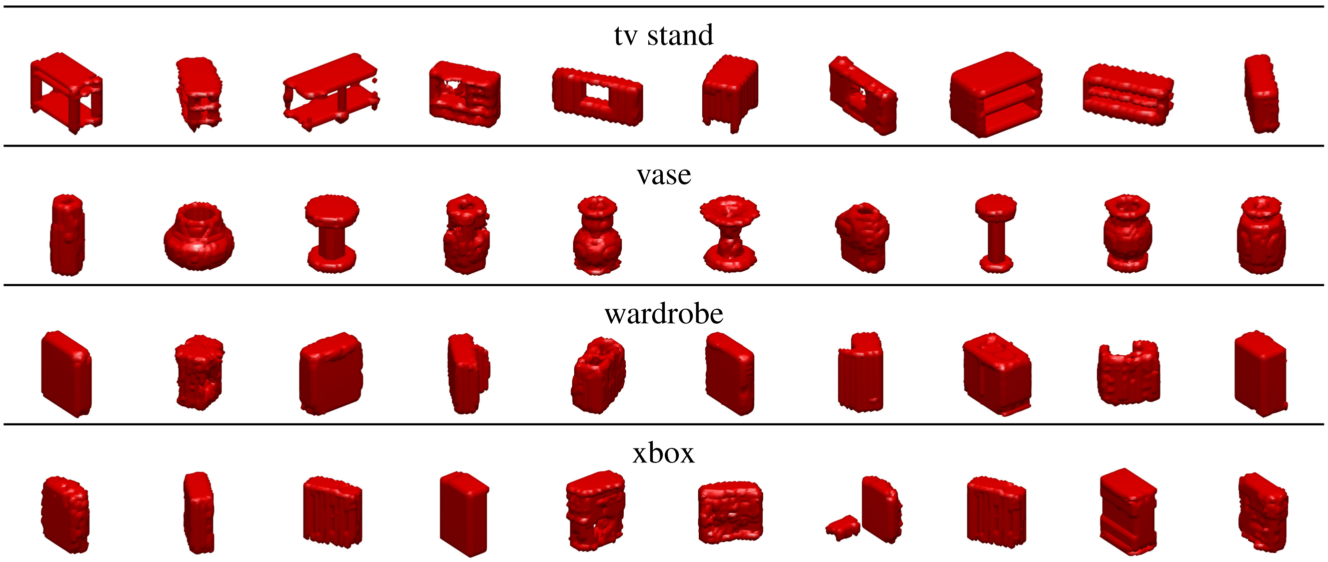


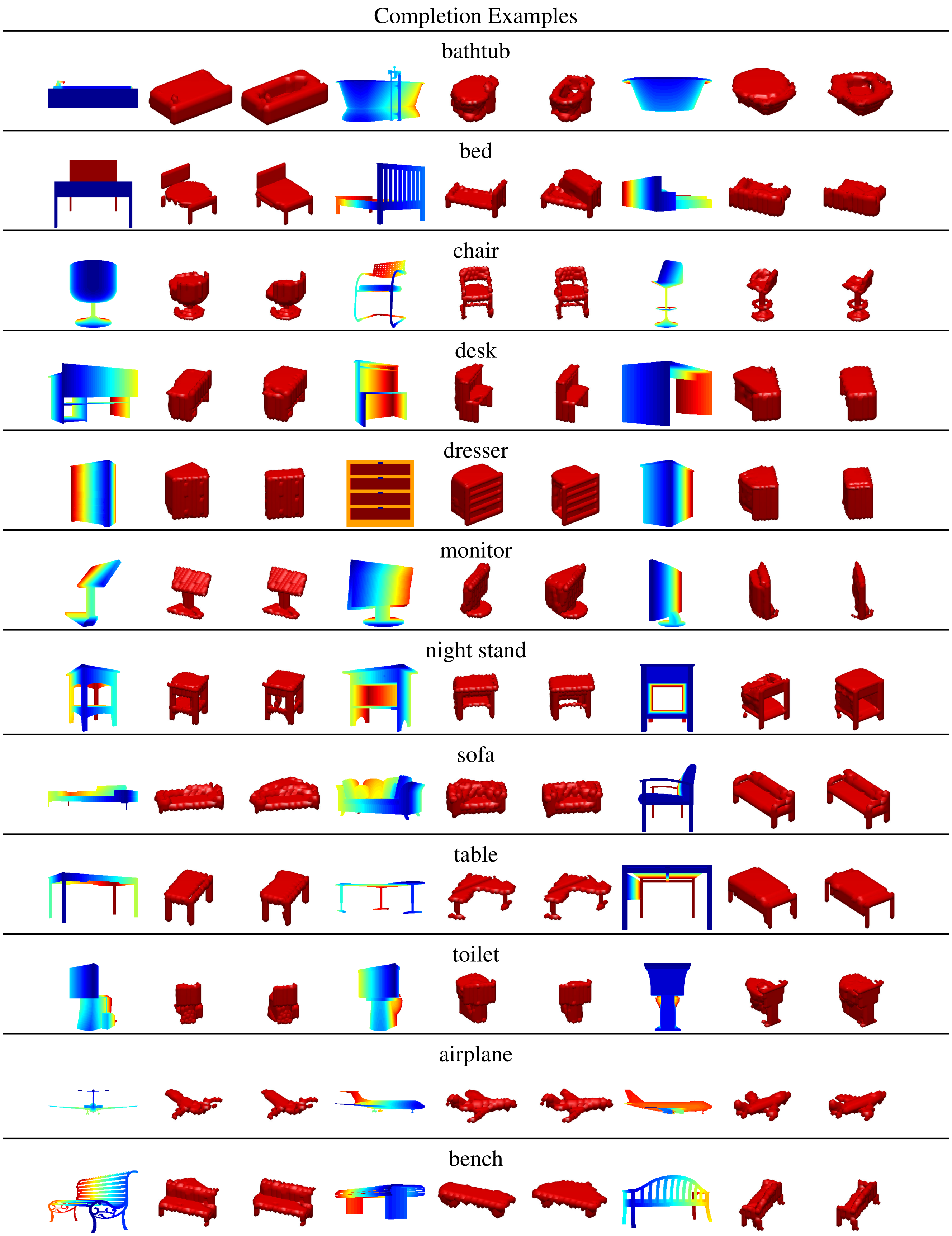
Figure 3: When the uncertainty of observation goes higher, the recognition accuracy drops.

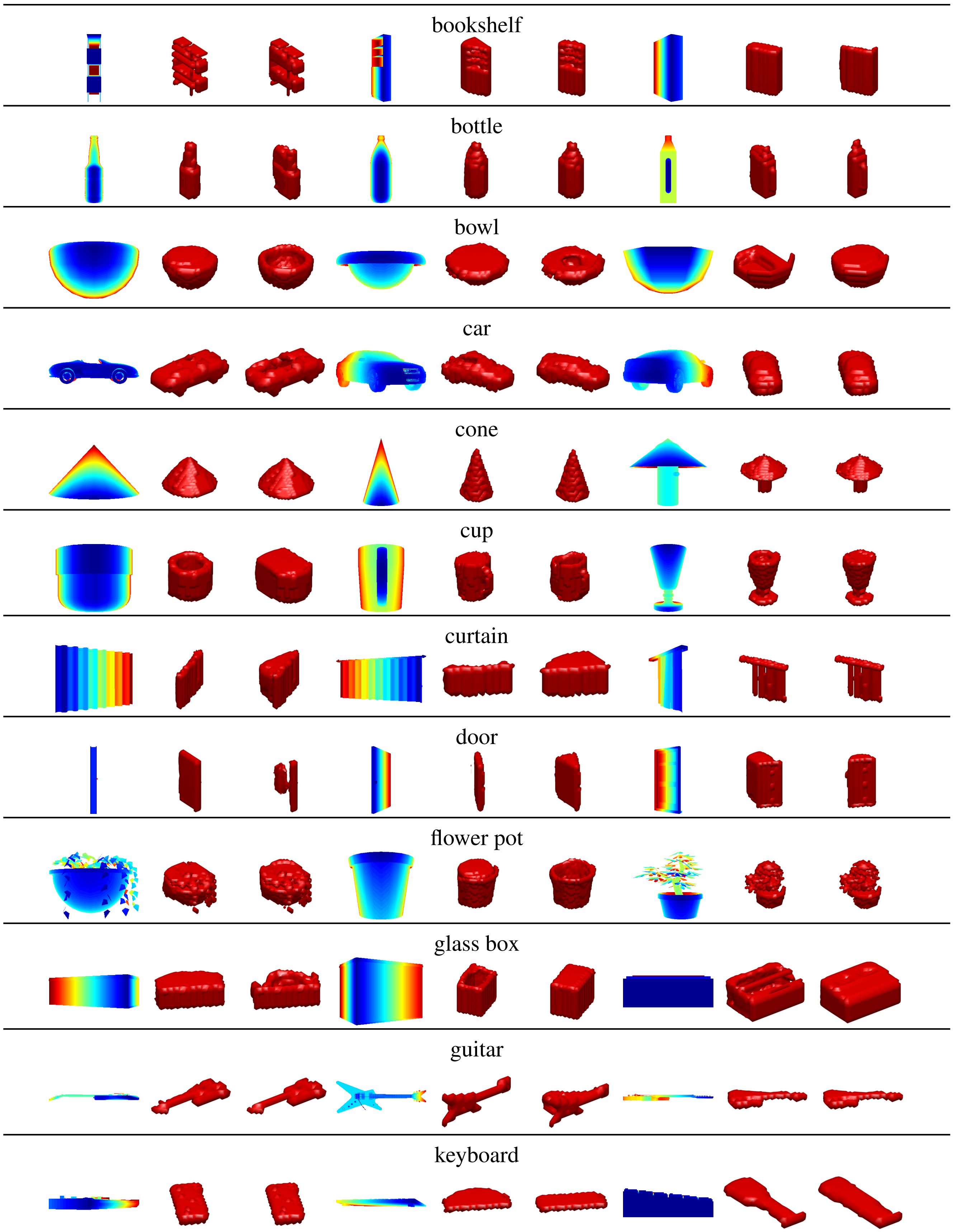

Generated Samples
bathtub
bed
chair
desk
dresser
monitor
night stand
sofa
table
toilet
airplane
bench

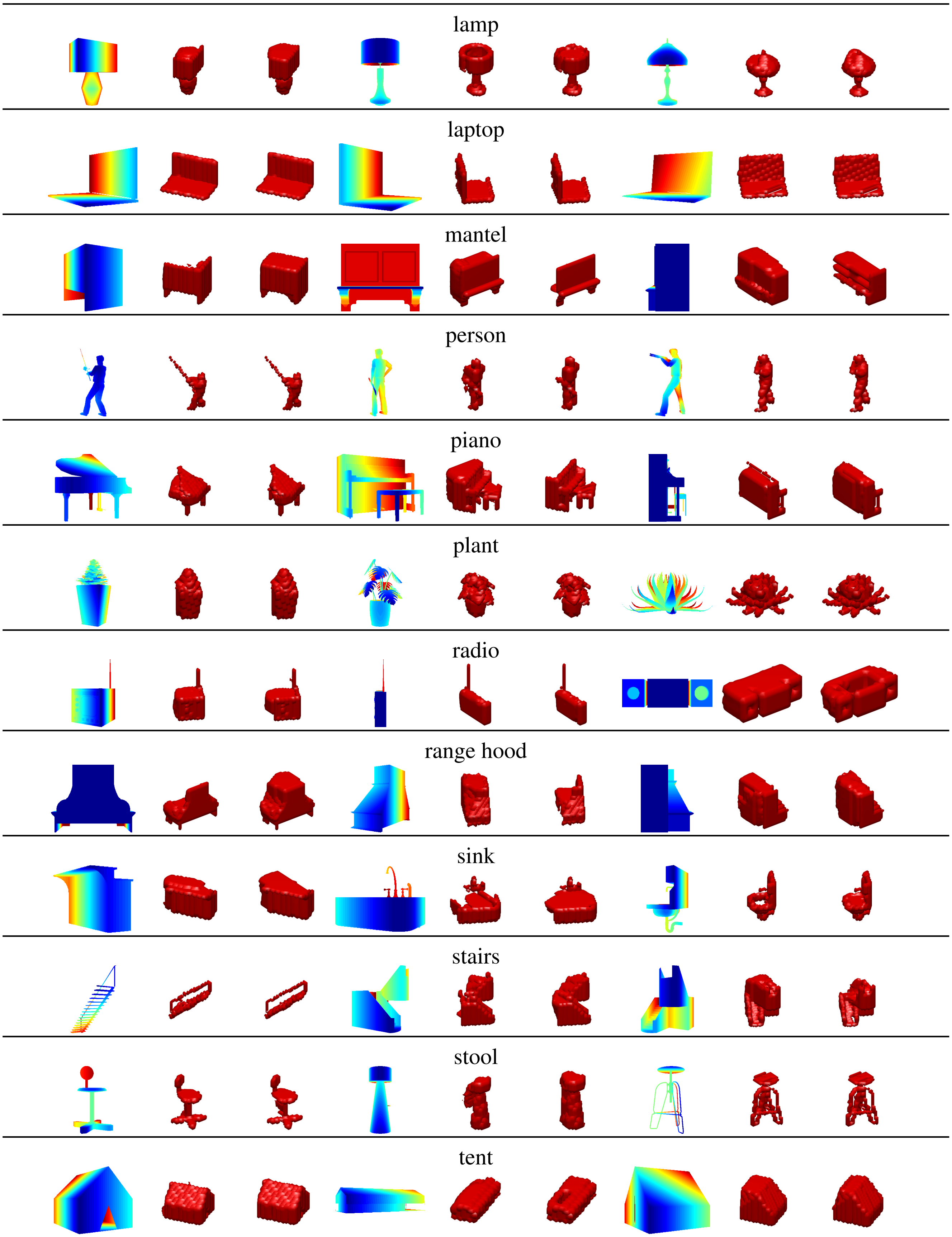

bookshelf
bottle
bowl
car
cone
cup
curtain
door
flower pot
glass box
guitar
keyboard

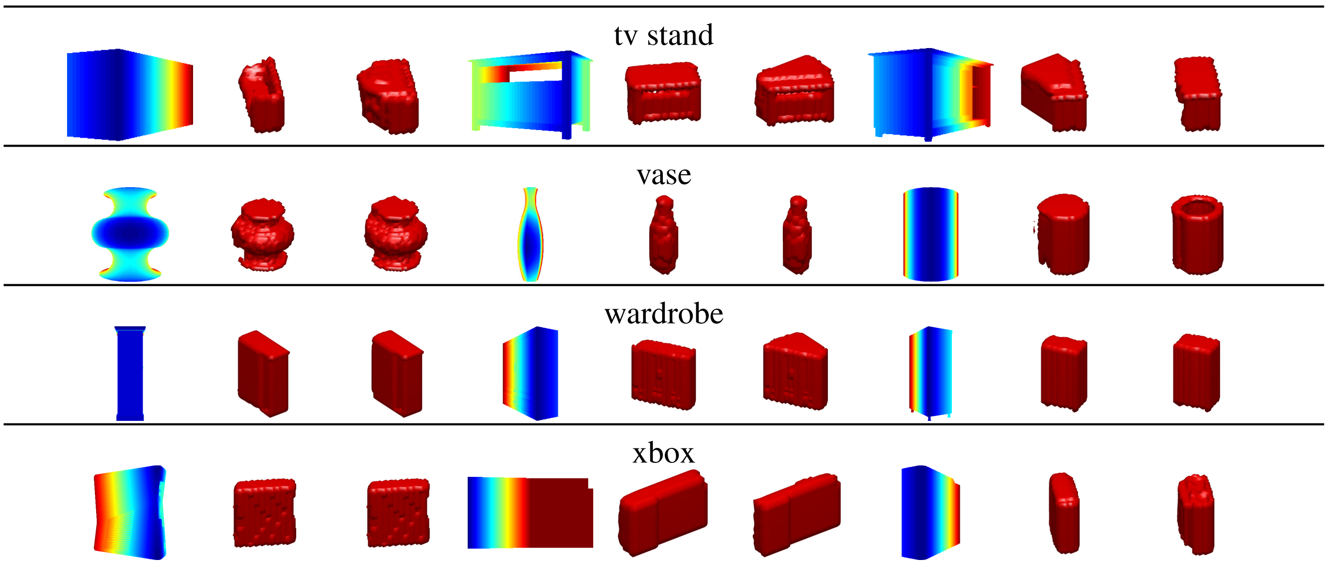

lamp
laptop
mantel
person
piano
plant
radio
range hood
sink
stairs
stool
tent


tv stand
vase
wardrobe
xbox


Completion Examples
bathtub
bed
chair
desk
dresser
monitor
night stand
sofa
table
toilet
airplane
bench


bookshelf
bottle
bowl
car
cone
cup
curtain
door
flower pot
glass box
guitar
keyboard


lamp
laptop
mantel
person
piano
plant
radio
range hood
sink
stairs
stool
tent


tv stand
vase
wardrobe
xbox