

User Input: Scene Description

A 30 m² dining room containing 10 chairs of different types. Add additional furniture to achieve 50% room occupancy. The camera movement mimics a handheld recording style and capture as many objects as possible ...

InfiniBench

LLM

Iterative refinement
of scene constraints

Constraints

Other settings

Object cluster-
based optimization

3D scene

Camera trajectory
optimization

Image/video
streams

Compositional Complexity

Relational Complexity

Observational Complexity

Low

Numbers

Diversity

...

Occupancy

Symmetry

...

Occlusion

Camera Height

...

5 Chairs

One type of chairs

...

Low

Grind-like table arrangement

...

No Occlusion

Drone view

...

10 Chairs

Two types of chairs

...

Medium

Semi-regular arrangement

...

Partial Occlusion

Humanoid robot view

...

15 Chairs

Four types of chairs

...

High

Random Table arrangement

...

Full Occlusion

Robot vacuum view

...

Scene Complexity

High