3D Information

Camera Pos: x, y, z, 𝒦 intrinsic
Objects:
- Countertop: loc: a, b, c
- Toaster: loc: d, e, f
- ...

Choose:
- Rotate Left/Right
- Move Forward
- Teleport to X
- Move object to (p,q,r)

Updated 3D Information

Camera Pos: x', y', z', 𝒦 intrinsic
Objects:
- Countertop: loc: a,b,c
- Toaster: loc: p, q, r
- ...

Generate Question-Answer Pairs

Q: Which direction did the camera taking the picture move?
→ `Check action taken`
Ans: Rotated left

Q: Did any object move from the first frame to the second?
→ `Compute relative movement from 3D Loc`
Ans: Toaster moved left and away from camera