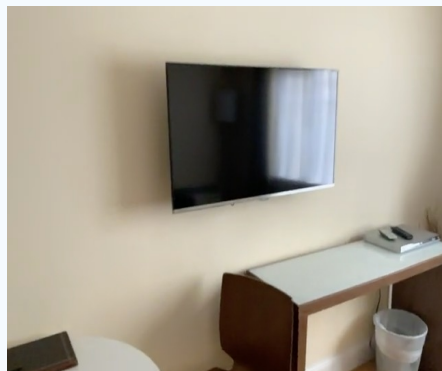


1. Multi-view observations

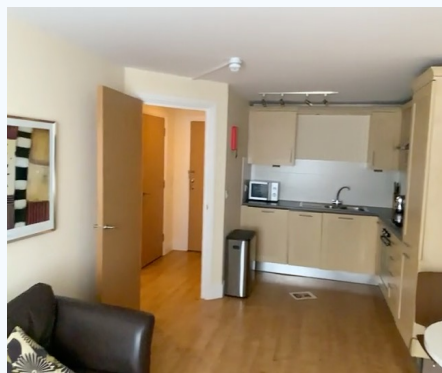
RGB images from different views



View 1



View 2



View 3

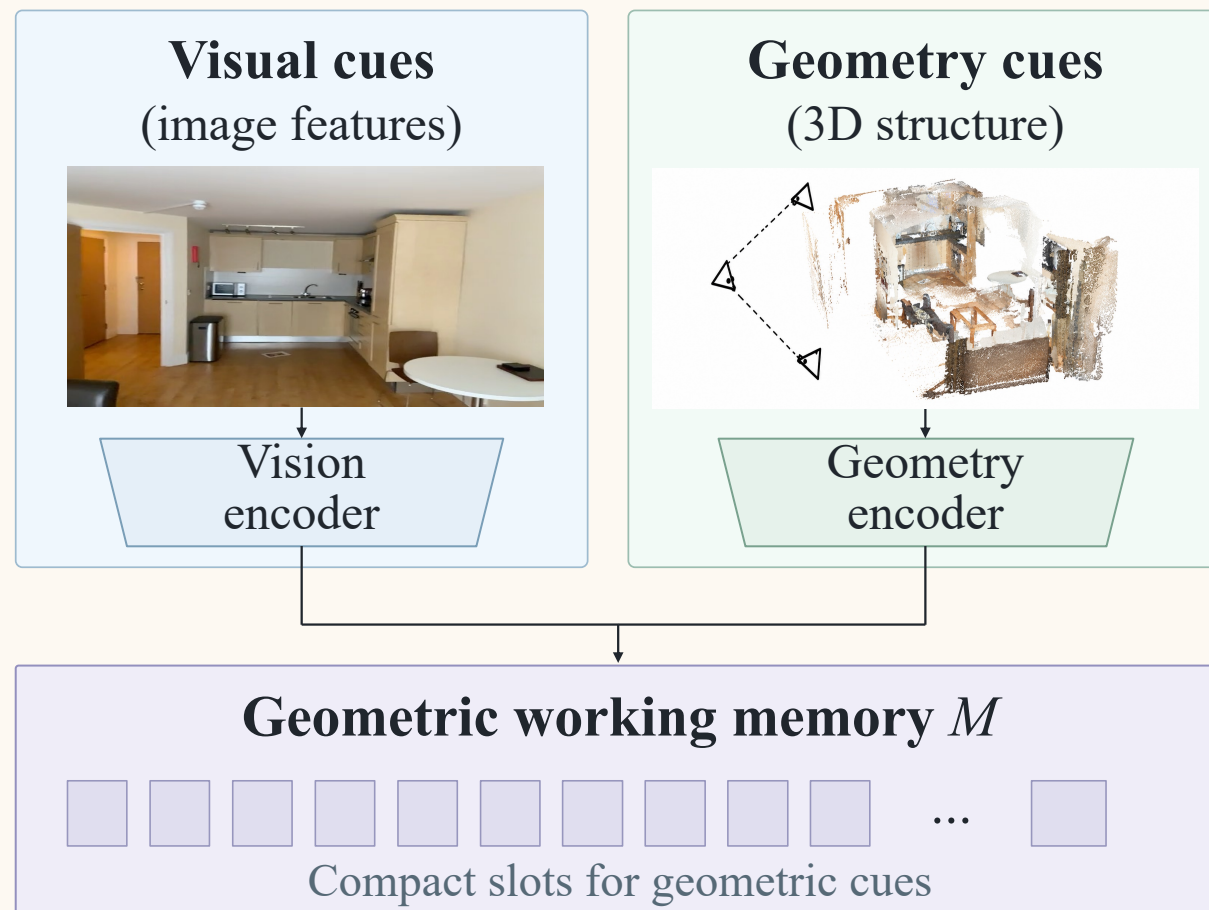


View 4

...

2. Geometry as memory

Turn multi-view observations and 3D geometric cues into a compact geometric working memory



3. Spatial reasoning

Use geometric memory to answer spatial questions

Q: If I am standing by the stove and facing the sofa, is the tv to my front-left, front-right, back-left, or back-right?



A: Front Left