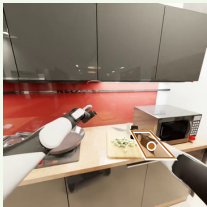


Direction · relative position



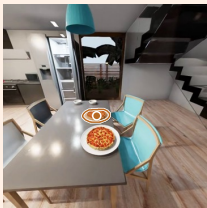
TASK-VQA

From the viewpoint of the current frame, where is carving knife relative to cutting board?

IR / ACTION

Place the carving knife to the right of the cutting board.

Distance · closest



TASK-VQA

Measuring from the center of each object, which object is the closest to the fridge?

IR / ACTION

Reach the pizza plate closest to the fridge on the breakfast table.

Size · ranking



TASK-VQA

Measuring from overall size, which food item on the countertop is the second smallest?

IR / ACTION

Reach the second smallest food item on the countertop.

Direction · locate



TASK-VQA

Where is the tupperware located on the fridge?

IR / ACTION

Place the tupperware in layer 5 of the fridge.

Distance · farthest



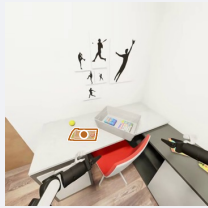
TASK-VQA

Measuring from the center of each object, which object is the farthest from pencil case?

IR / ACTION

Reach the paperback book farthest from the pencil case on the desk.

Size · largest



TASK-VQA

Measuring from overall size, which toy on the desk is the largest?

IR / ACTION

Reach the largest toy on the desk.

Navigation · spatial context



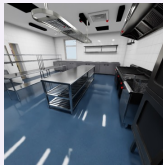
Context 1



Context 2



Context 3



Current view

TASK-VQA

From the viewpoint of the last frame, measuring from object centers, where is the diner booth in the corner relative to the robot?

ACTION

Move to the corner booth.

SPATIAL CONCEPTS

Direction

Relative direction · locate
Spatial ordering / ordinal

Distance

Closest / farthest
Object, camera, or robot reference

Size

Largest / smallest
Ranked size comparisons

OBSERVATION SETTINGS

Single-frame

One view

Multi-frame

Rotation · navigation

VQA EVALUATION

Single-frame 1,078

Multi-frame · rotation 239

Multi-frame · navigation 255

Task-grounded · manipulation 1,082

Task-grounded · navigation 602

TARGET GROUNDING / IR 1,240

ACTION TASKS

14

10

Manipulation tasks Navigation tasks

Manipulation task coverage

Direction · 5

Distance · 5

Size · 4

Direction includes locate and ordering.
14 evaluated tasks; train-only saucer excluded.