Viewpoint-100K

Ego-centric Translation
Object-centric Translation
Ego-centric Rotation

The object captured in these two photos did not move. From the camera/object's perspective, did the position of the camera that captured the second photo move/rotate relative to the first photo's camera position/orientation?

A: Move/Rotate to the right
B: No movement/Rotation
C: Move/Rotate to the left

Human-Assisted CoT Creation

Questions & Detailed Prompt

Gemini

1.Identify the Images ·····
2.Visual Analysis ·····
3.Analyze Specific Movements ···
4.Synthesize the Trajectory ····
5.Answer Question ·····

Coarse CoT Template

Human Modify

- Correct the **movement direction** and **rotate angle**.
- Discuss the **relative relationship** and **scene layout**.
- Conduct the analysis across different **frames of reference**.

Pseudo CoT Generation

- I'll **give you an example**, please follow the example to analyze the camera movement between the two images.

Human-Assisted CoT Template

- Please analyze the given two images using the same approach **as I provided above**.
- Do not use the relative position changes of objects in the images, **consider the changes in 3D space**.

Image Pairs & Questions

The **correct answer** is...

Pseudo CoTs

Gemini

Knowledge Injection

Qwen-VL

SFT

Viewpoint-100K

Pseudo CoTs

90%

10%

Hybrid Cold-Start Initialization

Generalization Enhancement

GRPO

SAT dataset

Actial

System Prompt:
····· You should first think about the reasoning process in the mind ·····
Response:
<think>
Identify the Images ·····
Analyze the Images ·····
Comparison of Images ·····
Identify the Action ·····
Analysis of the Question ·····
</think>
<answer>·····</answer>