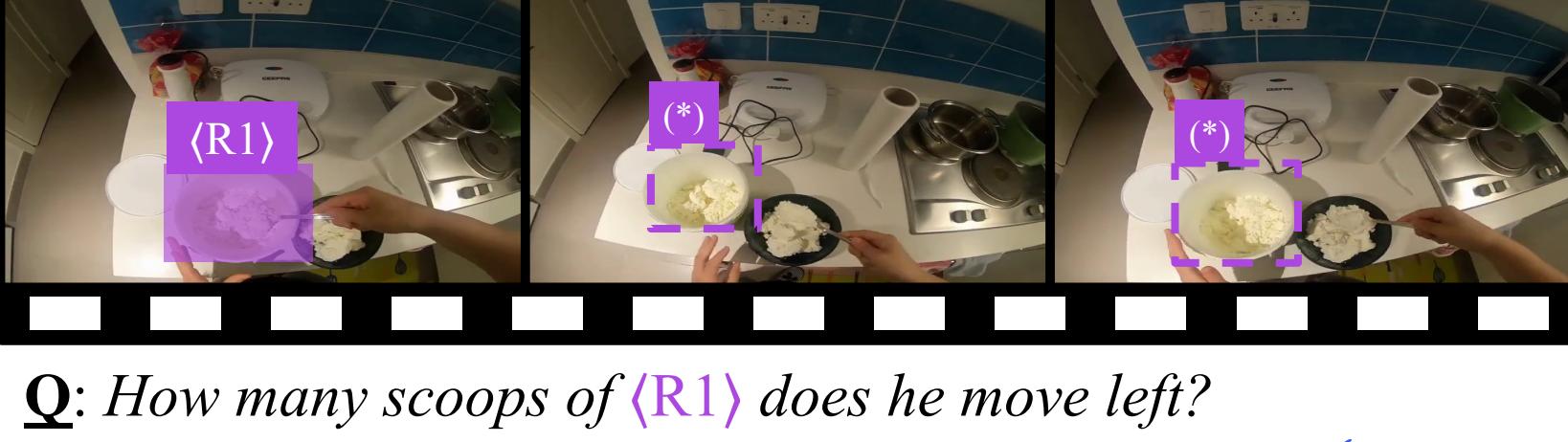


Q: What direction is <R1> moving towards?

✓ Ours: not moving

✓ GPT: not moving

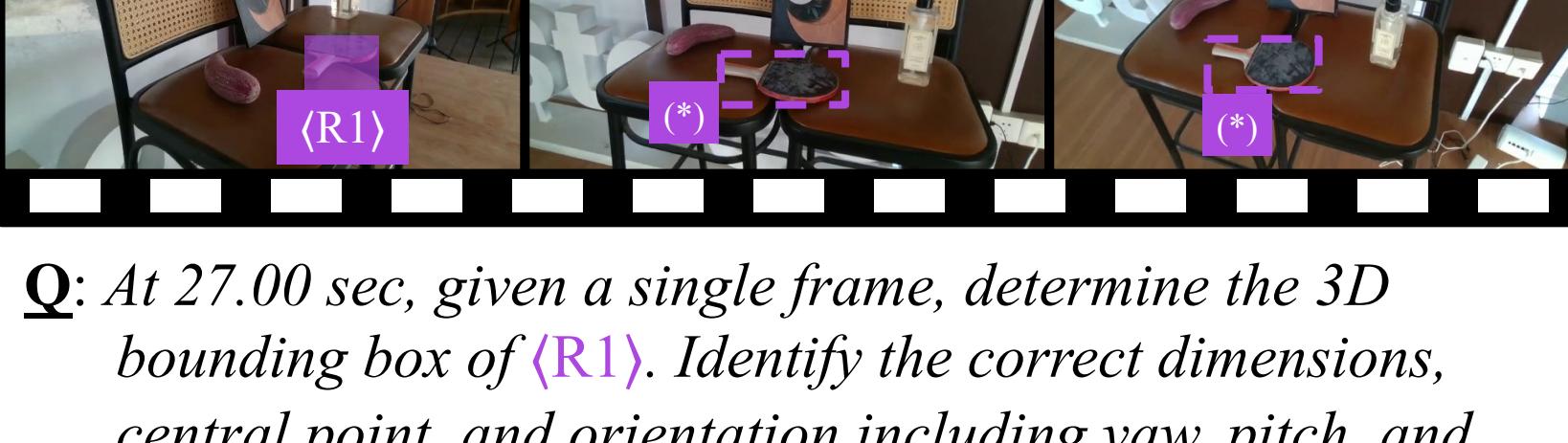


Q: How many scoops of <R1> does he move left?

✗ Ours: 1

✗ GPT: 2

✓ Ans: 0



Q: At 27.00 sec, given a single frame, determine the 3D bounding box of <R1>. Identify the correct dimensions, central point, and orientation including yaw, pitch, and roll. (unit: cm, °)

✓ Ours & GPT:

```
{
    dimensions: [25.62, 2.38, 15.33],
    central_point: [12.91, 2.77, 90.59],
    orientation: {
        yaw: 117.10,
        pitch: 42.61,
        roll: 114.41
    }
}
```



Q: At 18.52 sec, what is the 3D bounding box in camera coordinates of the <R1> from a single randomly selected frame? (unit: m, m/s, m/s^2, °)

✓ Ours:

```
{
    C_lwh:
    [0.86, 1.26, 0.74],
    C_central_point:
    [3.55, 1.33, 2.75],
    C_heading:
    27.51
}
```

✗ GPT:

```
{
    C_lwh:
    [0.86, 1.26, 0.74],
    C_central_point:
    [3.74, 1.39, 2.81],
    C_heading:
    27.20
}
```