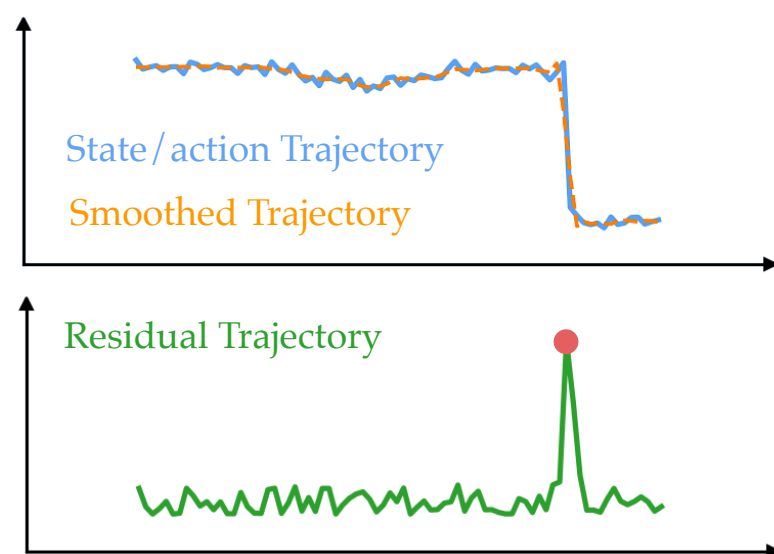


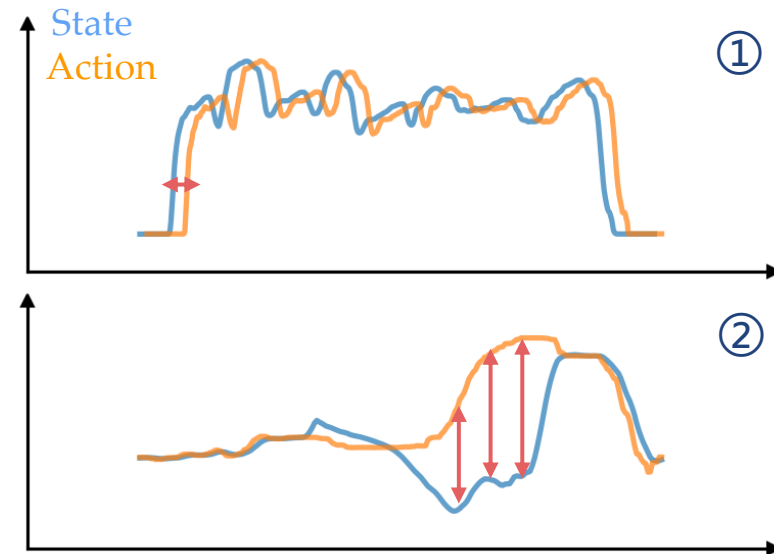
### S1. Sudden Change Detection



Filter Items:

- Residual (1<sup>st</sup> diff.)
- Acceleration (2<sup>nd</sup> diff.)
- Jerk (3<sup>rd</sup> diff.)

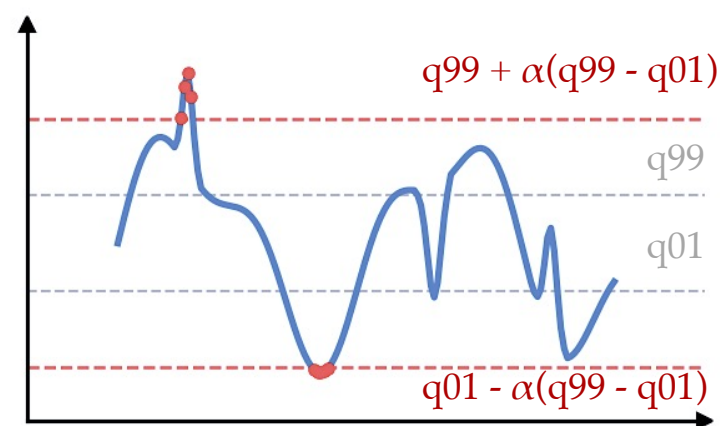
### S2. State-Action Trend Alignment



Two Wrong Cases:

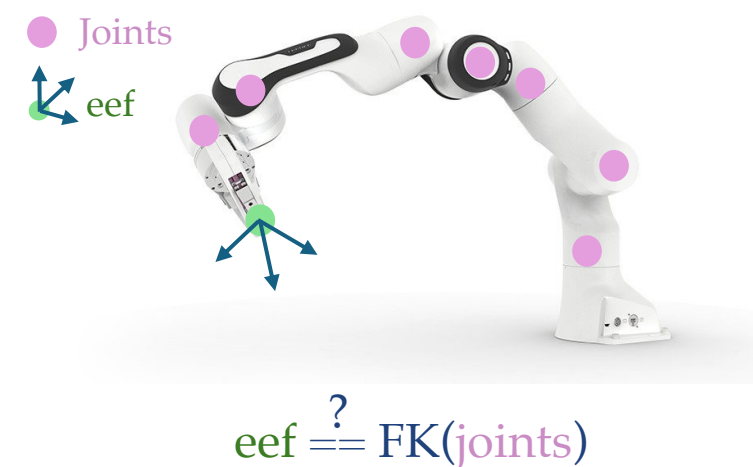
- ① State is before Action
- ② State-Action Trend Misaligned

### S3. Extreme Value Detection



- For each machine type in each dataset, q01 and q99 are computed independently.
- Frames with any dimension outside the band are excluded.

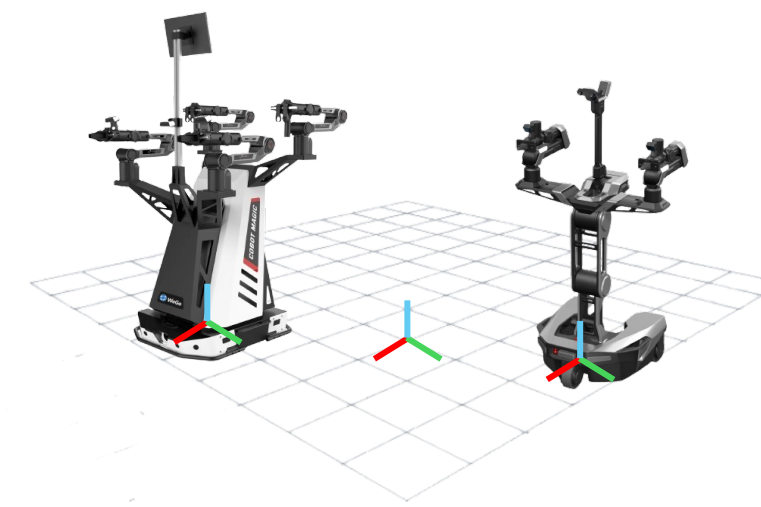
### S4. Kinematic Consistency



Resolve:

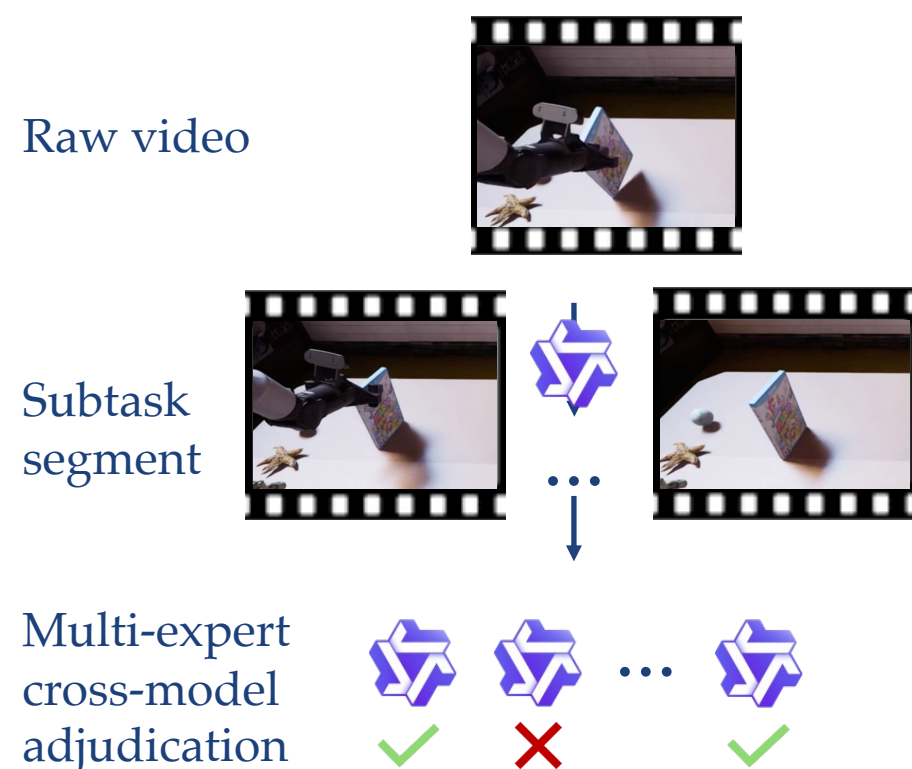
- TCP offset
- Joint offset
- Rotation Representation
- ...

### S5. Orientation Alignment

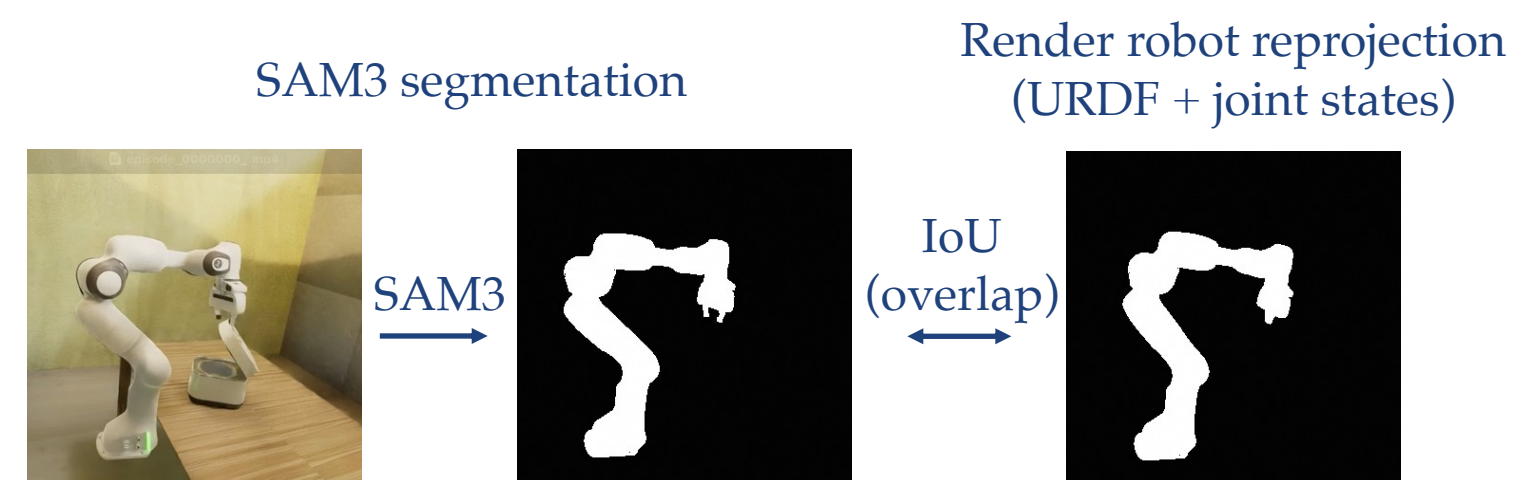


Robots across different datasets share the same orientation and are aligned to a common world frame.

### C1. Instruction Consistency



### C2. Video-State Consistency



When IoU < threshold:

- If caused by camera parameters: optimize camera parameters
- Else: exclude the episodes

### C3. Video Quality Filtering

