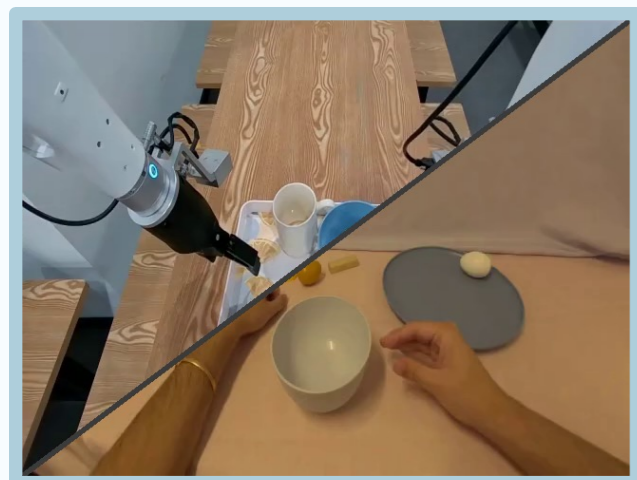
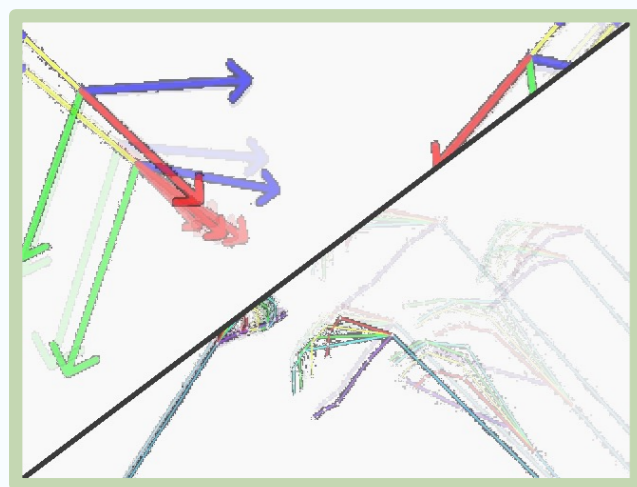


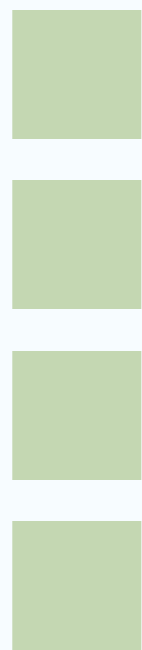
Condition Encoding



First Frame RGB

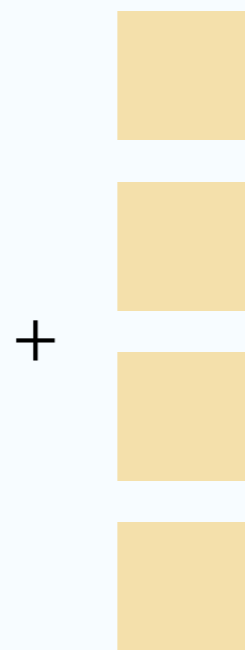


Skeleton Condition



+

Noise



Conditioning Generation



Output

