Supplementary Material

A1. Comprison between Bottom-k & Top-k

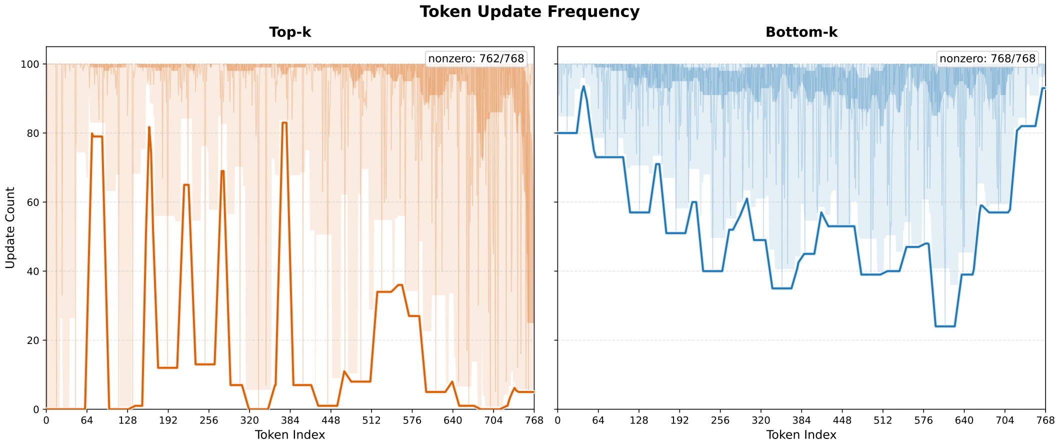

Figure 1: **Visualization of distinct strategies.** We employ `Top-k` and `Bottom-k` strategies separately, and tally the state tokens that are updated in each input frame. Results show some typical results that `Bottom-k` not only achieves a higher update frequency but also yields a more balanced updates across all state tokens.

`Top-k` updates the most-aligned tokenscreating a positive feedback loop that a small set of high-score tokens is repeatedly selected and reinforced, while the rest receive few updates and gradually become stale. This behavior reduces the effectiveness of memory diversity and utilization. In contrast, `Bottom-k` updates the least-aligned tokens, which naturally spreads writes across the state over time and improves overall memory coverage. For CUT3R(w. MeMix) on 7-Scenes, `Top-k` leaves a distinct subset of tokens nearly unupdated, whereas Bottom-k yields far more uniform token updates overall.

A2. Pose Estimation

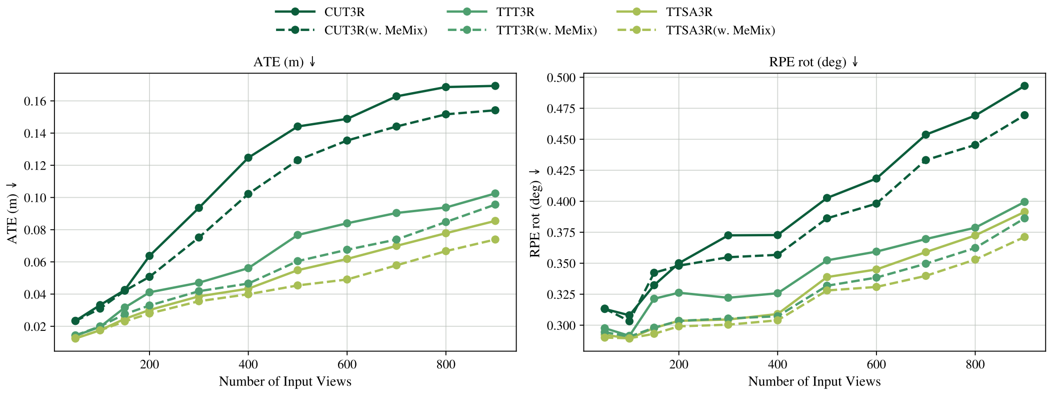

Figure 2: **Evaluation on Long-Sequence Pose Estimation.** We compare CUT3R, TTT3R and TTSA3R with their MeMix variants on long input streams. MeMix consistently improves pose estimation quality, achieving better scales cross three different baselines.

A3. Video Depth Estimation

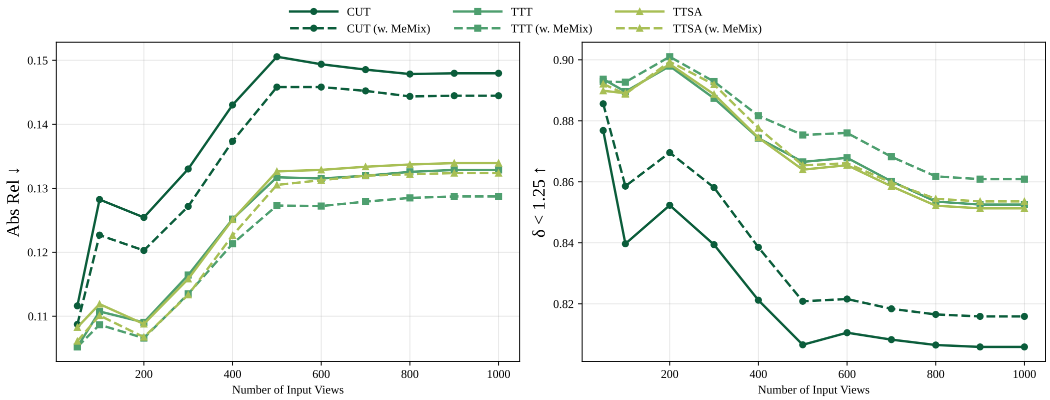

Figure 3: **Evaluation on Video Depth Estimation.** We compare CUT3R, TTT3R and TTSA3R with their w/o.MeMix version under long input streams. MeMix generally improves depth estimation quality from 50 to 1000 frames input. Notably the outcomes are largely depends on capacity of original model.

A4. 3D Reconstruction

Table 1: **3D Reconstruction Results on 7-Scenes and NRGBD.** We test MeMix on 7-Scenes and NRGBD, with sample every one frames **(Dense)**.Green boxes indicate improved or unchanged performance over the base model (w/o MeMix) under the same input length.

Model	**MeMix**	**Input**	**7-Scenes-D**						**NRGBD-D**					
			Acc. ↓		Comp. ↓		NC ↑		Acc. ↓		Comp. ↓		NC ↑	
			Mean	Med.	Mean	Med.	Mean	Med.	Mean	Med.	Mean	Med.	Mean	Med.
VGGT *(Offline)*	–	*300*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*
	–	*400*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*
	–	*500*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*
StreamVGGT	–	*300*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*
	–	*400*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*
	–	*500*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*	*OOM*
CUT3R	×	*300*	0.099	0.062	0.048	0.014	0.542	0.562	0.137	0.092	0.066	0.024	0.572	0.609
	✓		0.076	0.045	0.039	0.010	0.549	0.573	0.113	0.081	0.060	0.035	0.578	0.618
	×	*400*	0.150	0.093	0.090	0.037	0.531	0.543	0.225	0.155	0.119	0.076	0.554	0.579
	✓		0.117	0.071	0.056	0.015	0.536	0.552	0.196	0.128	0.098	0.062	0.572	0.609
	×	*500*	0.165	0.114	0.094	0.039	0.522	0.531	0.313	0.203	0.202	0.148	0.554	0.580
	✓		0.146	0.094	0.067	0.022	0.528	0.541	0.273	0.173	0.162	0.110	0.568	0.602
TTT3R	×	*300*	0.030	0.016	0.019	**0.004**	0.558	0.588	0.057	0.035	0.016	**0.003**	0.595	0.650
	✓		0.030	0.016	0.019	**0.004**	**0.559**	0.589	0.052	0.032	0.015	**0.003**	0.599	0.656
	×	*400*	0.044	0.026	0.024	**0.004**	0.551	0.577	0.093	0.053	0.018	**0.003**	0.587	0.635
	✓		0.039	0.023	0.025	**0.004**	0.552	0.578	0.078	0.042	0.016	**0.003**	0.592	0.644
	×	*500*	0.068	0.046	0.033	0.009	0.542	0.562	0.127	0.061	0.033	**0.003**	0.586	0.635
	✓		0.057	0.039	0.030	0.008	0.546	0.568	0.105	0.048	0.026	0.004	0.586	0.633
TTSA3R	×	*300*	0.023	0.011	0.018	**0.004**	0.558	0.588	0.039	**0.022**	0.011	**0.003**	**0.606**	**0.669**
	✓		**0.022**	**0.009**	**0.017**	**0.004**	**0.559**	**0.588**	**0.037**	**0.022**	**0.010**	**0.003**	0.605	0.668
	×	*400*	0.030	0.016	0.022	**0.004**	0.553	0.580	0.060	**0.027**	**0.010**	**0.003**	**0.598**	**0.655**
	✓		**0.025**	**0.012**	**0.021**	**0.004**	**0.554**	**0.581**	**0.059**	**0.027**	**0.010**	**0.003**	0.596	0.651
	×	*500*	0.045	0.029	0.025	**0.004**	0.545	0.567	0.085	0.034	0.020	**0.003**	**0.596**	**0.651**
	✓		**0.035**	**0.021**	**0.023**	**0.004**	**0.548**	**0.571**	**0.081**	**0.032**	**0.014**	**0.003**	0.595	0.649

A5. Visualization

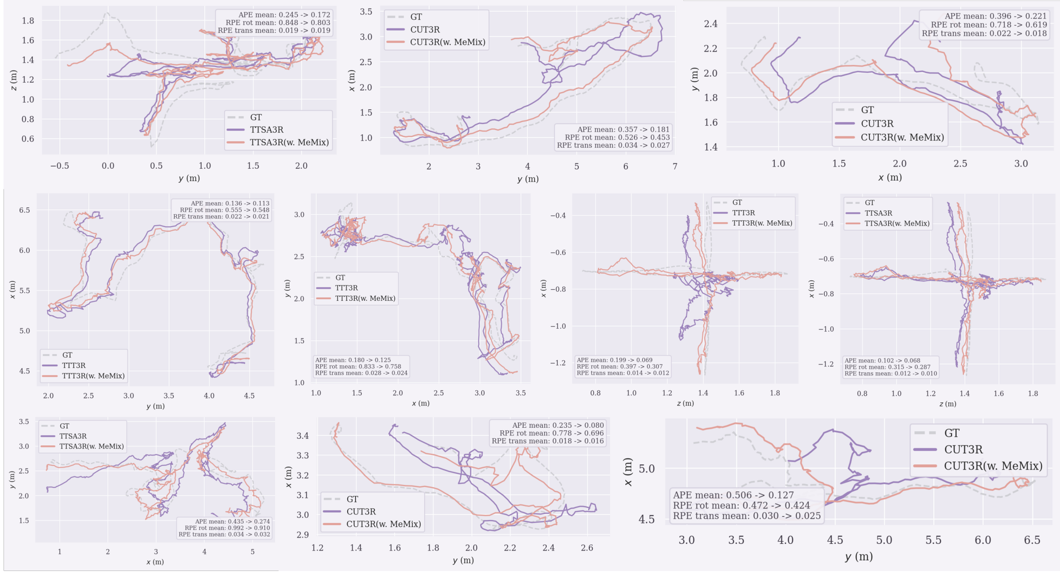

Figure 4: **Visualization of Estimated Camera Trajectories - Long Sequence.** The trajectories are plotted along the two axes with the highest variance to capture the most significant motion. Our estimated camera trajectory CUT3R(w. MeMix) deviates less from the ground truth GT compared to the baseline CUT3R.