SUPPLEMENTARY MATERIAL OF OPTIMAL CAMERA PLACEMENT FOR DYNAMIC SCENES VIA REINFORCEMENT LEARNING IN VIRTUAL ENVIRONMENTS

Author Names

Address info

In this supplementary material of our work, we provide additional experimental evaluation of the proposed Reinforcement Learning (RL) framework. For this purpose, a simple virtual environment of a four-walled room is exploited.

1. EXPERIMENTAL EVALUATION IN A SQUARE SHAPED ROOM

In this set of experiments, the performance of the proposed RL framework is evaluated in a squeared-sized room of $7\mathrm{m}$ length and $3\mathrm{m}$ height. Two virtual cameras with a diagonal FoV, $f = 20^0$, were placed in opposing walls. The Field-of-View (FoV) for the cameras was intentionally chosen to be narrow, in order to provide a more challenging environment for the evaluated RL agents and gain intuition on the achieved results upon convergence. The goal is for the deployed RL agent to optimally configure the cameras to maximize coverage despite their limited FoV. In this set of experiments, a single avatar is used, with its position randomly assigned within the 3D space at the start of each episode. The avatar's movement is contraint to remain within the boundaries of the four walls.

1.1. Twin Delayed Deep Deterministic

The final positions and orientations for the cameras after training with the Twin Delayed Deep Deterministic (TD3) policy gradient agent are presented in Table 1, using the same notation as in Figure 1. For the red camera, the horizontal environment figure is rotated $180°$ to align with the horizontal axis Figure 1, while the vertical figure is flipped along the y-axis to match the vertical axis of the same figure.

The episodic return chart of the training is illustrated in Figure 3. The camera positions and orientations at the beginning, middle and end of the training can be seen in Figure 2, where the lines on the diagrams indicate the cameras view frustrum. Doted lines indicate that the frustrum line has made contact and extends through a bounding object which is either the floor or a wall in this experiment. In the beginning of the experiments both cameras are placed randomly in the walls of the room. In the top view of the environment, the RL agent eventually places the cameras in adjacent vertices and rotates them towards the center of the room. From the side

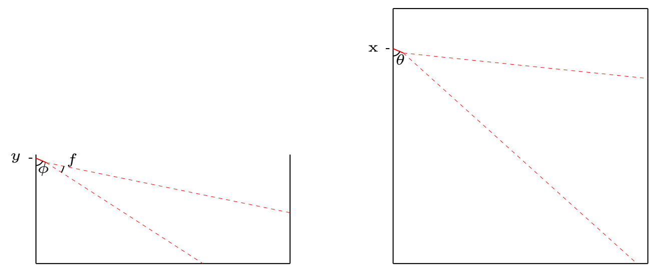

Fig. 1: The slice-view (left) and top-view (right) showing a camera and its parameters for position (x, y), pitch ϕ and yaw θ. The lower and upper boundaries for pitch and yaw are $(30°,90°)$, $(60°,120°)$ respectively.

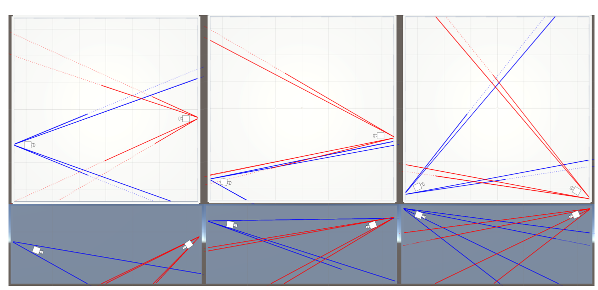

Fig. 2: Top (row 1) and side (row2) view for TD3 in environment 1 with the reward of Equation **??**. (row 1) Top view of the camera positions at the start (left), middle and end (right) of training. (row 2) Side view of the camera positions at the start, middle and end of training.

camera	x	ϕ	y	θ
Blue	5.83	68	2.90	79
Red	6.90	70	2.90	60

Table 1: Camera positions and orientations at the end of training with TD3, corresponding to the right column of rows 1 and 2 of Figure 2.

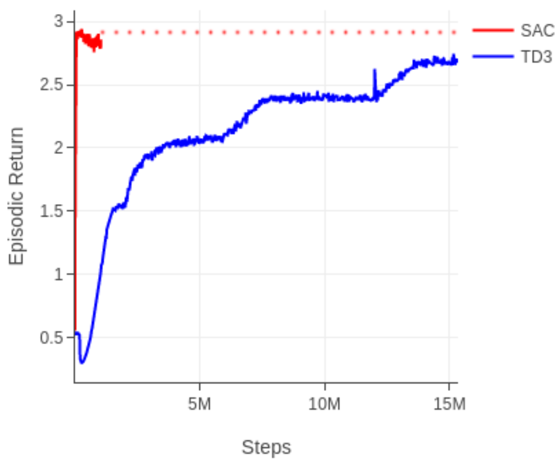

Fig. 3: Moving average of episodic return over the last 10K values, for 15M steps corresponding to the experiments of Figure 2 and Figure 4.

view, both cameras are placed near the ceiling of the room and tilted slightly downwards while facing the opposite wall. Note that cameras look relatively high in the opposing wall so that head keypoints in avatars appearing at a height of 1.8m are still visible.

1.2. Soft Actor Critic

The position and orientation of the cameras at the end of training with Soft Actor Critic (SAC) agent are on Table 2. The episodic return chart of the training is depicted in Figure 3. As shown in Figure 4, the cameras start from random configurations on opposing walls and are adjusted until convergence with a good coverage of the room. Contrary to TD3 they end up in opposing vertices of the room and have less overlap. Note also that while the blue camera is placed lower compared to the red one, its is still above 1.8 m so that it keeps in sight keypoints on the head of avatars.

camera	x	ϕ	y	θ
1	6.90	81	1.82	60
2	6.89	72	2.89	61

Table 2: Camera positions and orientations at the end of training with SAC, corresponding to the right column of Figure 4.

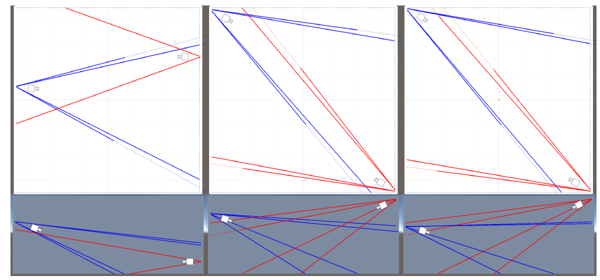

Fig. 4: Top (row 1) and side (row 2) view for SAC with in environment 1 with the reward of Equation **??**. (row 1) Top view of the camera positions at the start, middle and end of training. (row 2) Side view of the camera positions at the start, middle and end of training.

agent	steps	time
TD3	15M	3 days
SAC	180K	3 hours

Table 3: Experiment 1. Number of training steps and training time for each TD3 and SAC, where SAC clearly converges faster.

Table 3 contains the number of training steps and the training time needed for convergence of TD3 and SAC, where the latter has an obvious advantage. This is also shown in Figure 3, where SAC aside converging faster has also a higher episodic return than TD3. Nonetheless with both SAC and TD3, cameras are reconfigured to progressively better positions, as can be seen in the middle and right diagrams, based on the reward received by the respective RL agent. From the top view we can see that the cameras try to cover as much as possible the area that the other camera leaves uncovered. In the side view we see that the cameras converge with a downward pitch as they do not need to look for keypoints higher then 1.80m which is the height of the avatars.