1 MORE RESULTS

Scenes	*Ours (Sample 1)*	*Ours (Sample 2)*	*Ours (Sample 3)*

Figure R1: **More Results for Transition Samples.** In each case, transitions are generated by the same prompt with different random seeds. *Best view with Acrobat Reader. Click the images to play videos.*

| Scenes | VQGAN-based | Morphing | SD-based |

Figure R2: **More Results for Compared Transition Methods.** *Best view with Acrobat Reader. Click the images to play videos.*

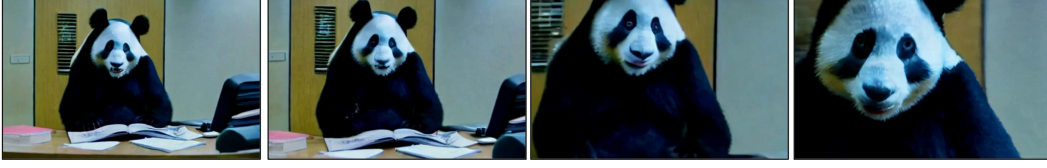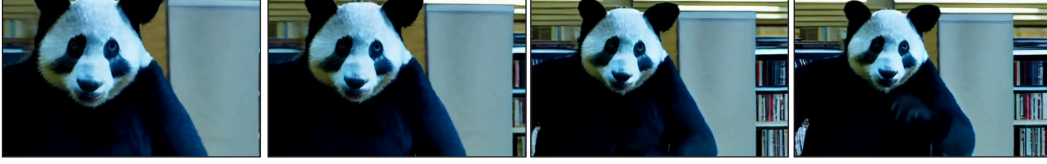| VQGAN-based | Morphing | SD-based |

| Ours (sample 1) | Ours (sample 2) | Ours (sample 3) |

Figure R3: **Comparison of our model with baseline methods.** *Best view with Acrobat Reader. Click the images to play the videos.*

1st prompt: A panda working in the office, camera zoom-in.

2st prompt: A panda working in the library, camera zoom-out.

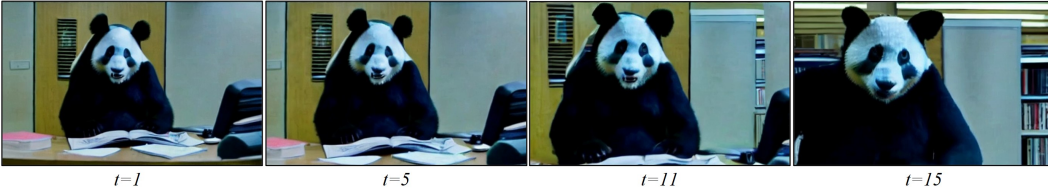*Prompt: A panda working in the office and in the library.*

| t=1 | t=5 | t=11 | t=15 |

Figure R4: **Ablation study for text-control transition.** The first two rows are results with the prompt of "camera zoom-in" and "camera zoom-out", and the last rows are results without "camera zoom-in / zoom-out".
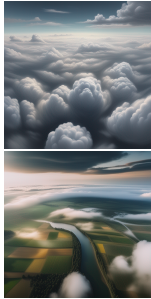

| *Scenes* | *Scene transfer from flower to cherry tree, camera zoom-out* | *Scene transfer from flower to cherry tree, camera pans.* | *Interpolation Results (VQGAN-based)* |

| *Scenes* | *Flying through fantasy landscapes in the cloud.* | *Cloud disappear and a village appear.* | *Interpolation Results (VQGAN-based)* |

Figure R5: **More Results and comparison for text-control transition.** In each row, our results are shown in the first two columns with different prompts, and the last column shows a baseline result of simple interpolation by VQGAN-based model. *Best view with Acrobat Reader. Click the images to play videos.*