Dragin3D: Image Editing by Dragging in 3D Space

Supplementary Material

7. Supplementary

7.1. More Qualitative Visualization

As illustrated in Fig. 12, we present additional visualization results of Dragin3D, both on our test set (Fig. 12 (a)) and a series of real-world images (Fig. 12 (b)). Given an input image, the selected object for editing, and a specified rotation trajectory, we can generate an image of the object rotated according to the defined trajectory from various viewpoints. These results demonstrate that, for a given object and its specified rotation trajectory, Dragin3D effectively rotates the object from any initial position to the desired final orientation, ensuring coherence throughout the intermediate stages of generation. Notably, in the fourth and fifth rows of Fig. 12 (b), for the *wolf* and *tiger*, we achieve precise head rotations in any direction with remarkable fidelity.

7.2. User Study

We conducted a user study to evaluate the subjective quality of images generated by Dragin3D and SV3D, focusing on human perception. Participants provided statistical votes based on three criteria: image fidelity, 3D effect, and generalization. The voting results demonstrate that Dragin3D outperforms SV3D by 52%, 16%, and 70% in these respective metrics, as shown in Fig. 11.

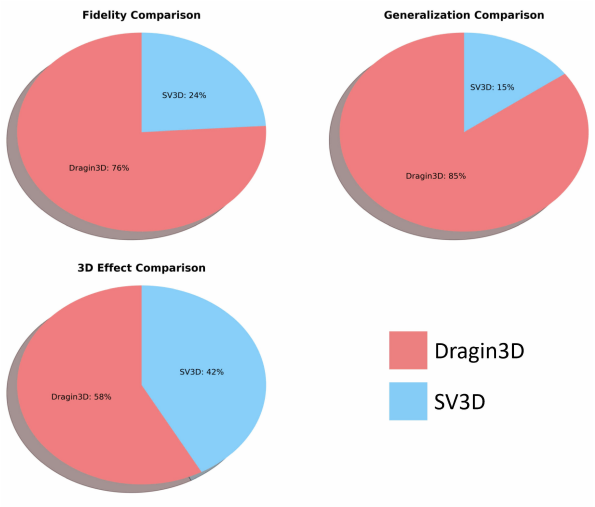


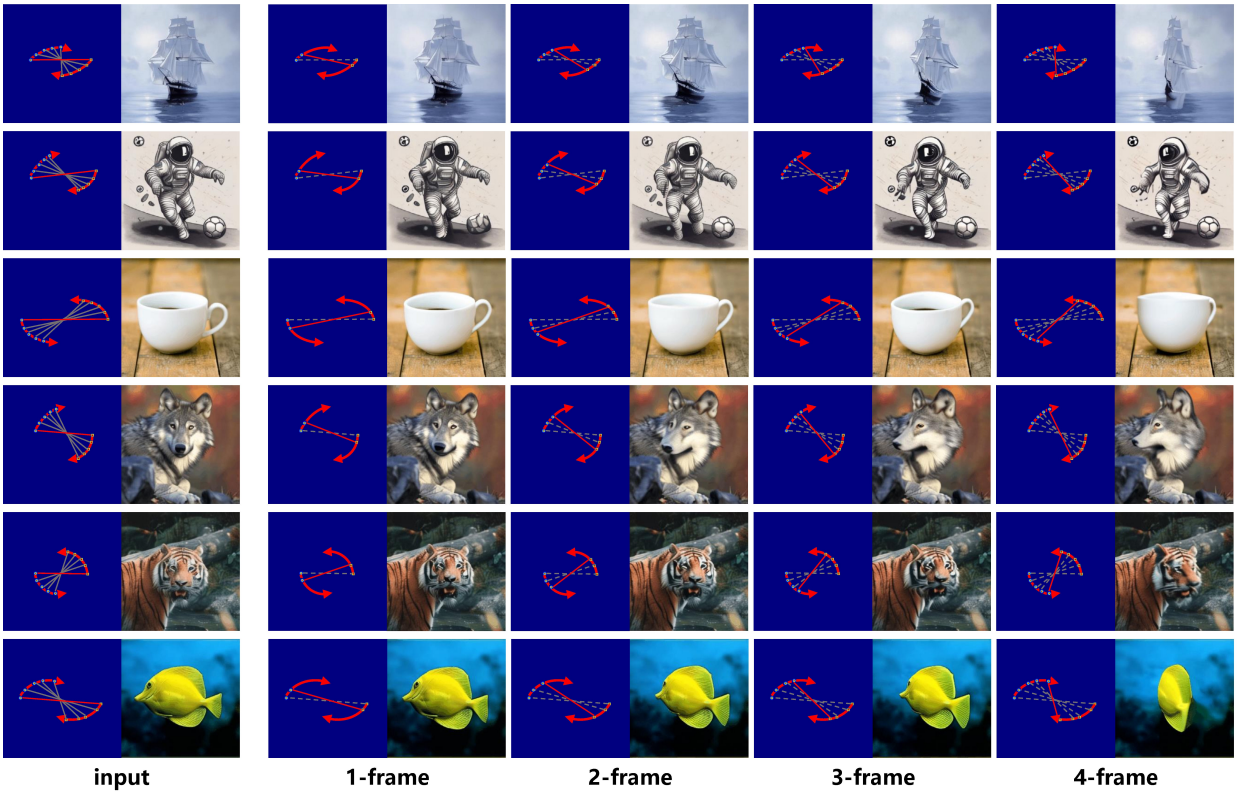
Figure 11. **User study.** We performed a subjective human evaluation to compare the image generation performance of Dragin3D and SV3D. The results indicate that Dragin3D outperforms SV3D in three key areas: fidelity, 3D effect, and generalization, with improvements of 52%, 16%, and 70%, respectively.

7.3. More Ablation Visualization

Effectiveness of shape distinguish strategy (SDGS). As demonstrated in Fig. 13 and Fig. 14, we present the results of additional ablation experiments for the shape distinguish strategy (SDGS). In these experiments, the object is given a trajectory that is rotated to a symmetric view, where the two representation points of the object approximately align (highlighted by the green box). Specifically, in Fig. 13, the object undergoes a substantial deformation rotation (highlighted by the red box), while in Fig. 14, the object fails to continue its trajectory beyond the symmetric view, indicating a rotation limit (highlighted by the yellow box). This limit occurs approximately at the mirror-symmetry position of the input object. These results demonstrate that the SDGS point differentiation strategy effectively mitigates the representation ambiguity caused by specular symmetries, thereby significantly enhancing the model's robustness to large-angle rotations.

Effectiveness of control condition inject module (CCIM). As illustrated in Fig. 15, we present the results of additional ablation experiments evaluating the control condition inject module (CCIM). The results demonstrate that when only the Gaussian map of the original positional information is used, without the introduction of CCIM, the object exhibits a minimal change at the initial frame (highlighted by the red box), remaining nearly identical to the input image. Subsequently, larger deformations and artifacts emerge, resulting in a significant loss of fidelity compared to the original image (highlighted by the yellow box). These results indicate the critical role of CCIM in enhancing the model's capacity to learn and integrate object features during training, ensuring both high fidelity in the generated images and coherence in motion across frames.

(a) Visualization of experimental results on our test set.



(b) Visualization of experimental results on real-world images.

Figure 12. **More visible results.** Given an input image and a set of trajectory point pairs, we obtained Gaussian maps of the trajectories as input (first column), which is then used to generate the corresponding objects at the specified trajectory points (subsequent columns). We show the cases both on our test set and on the real-world images. It is demonstrated that Dragin3D can edit objects to novel views with a random input view and rotate trajectories.

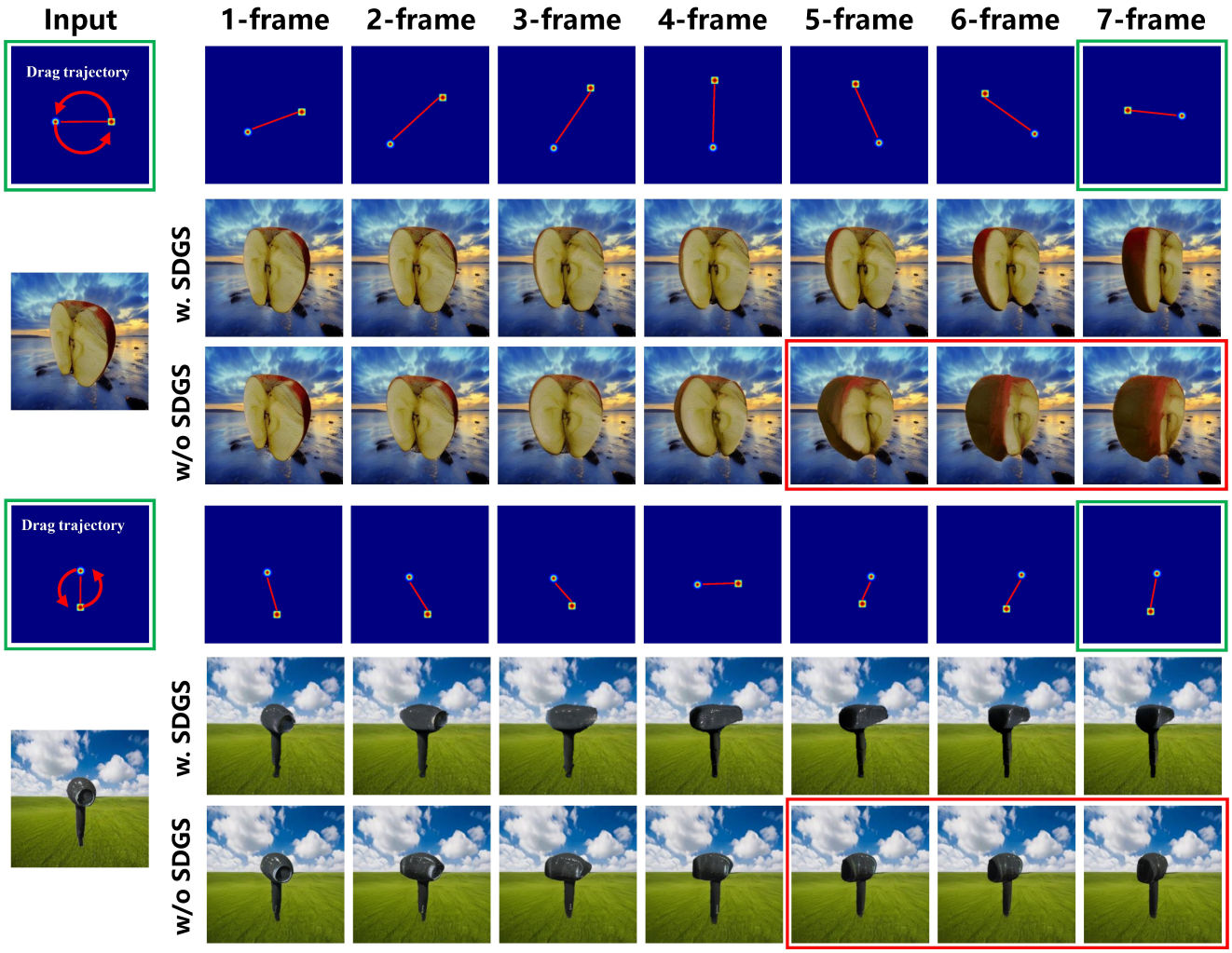


Figure 13. **Effectiveness of shape distinguish strategy (SDGS).** Here we provided two cases for ablation experiments on SDGS. The first row displays the trajectory points projected onto Gaussian maps, while the second and third rows show the objects from the corresponding view of these trajectory points, generated with and without SDGS, respectively. The positions of the two trajectory points in the first frame closely correspond to those in the last frame (green box). In the latter half of the sequence, the object exhibits significant distortions and artifacts due to the ambiguity introduced by the specular Gaussian representation (red box).

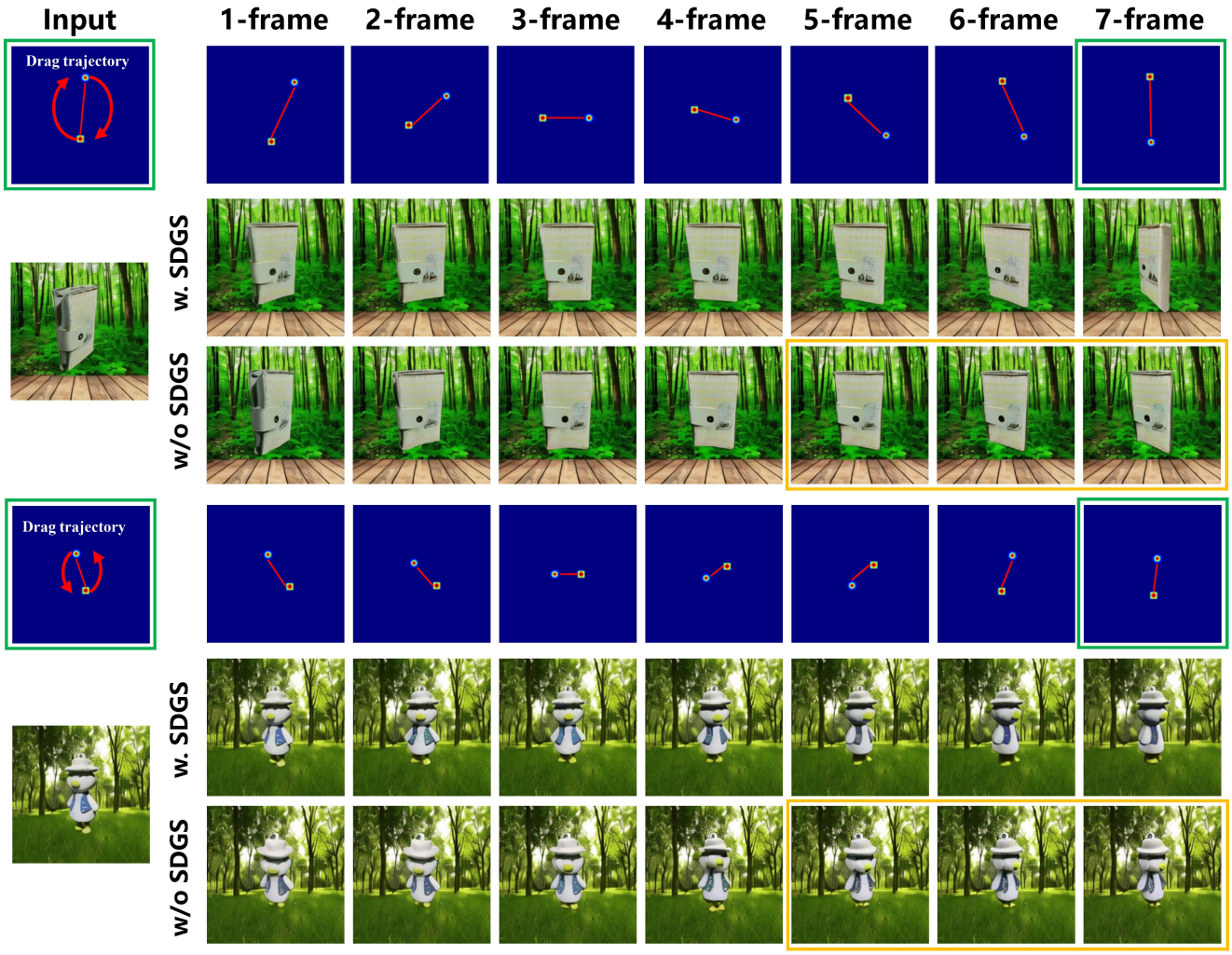


Figure 14. **Effectiveness of shape distinguish strategy (SDGS).** Here we provided two cases for ablation experiments on SDGS. The first row displays the trajectory points projected onto Gaussian maps, while the second and third rows show the objects from the corresponding view of these trajectory points, generated with and without SDGS, respectively. The positions of the two trajectory points in the first frame closely correspond to those in the last frame (green box). In the latter half of the sequence, the object imposes a restriction on the rotation angle, preventing further rotation upon reaching the symmetry view of the input image (yellow box).

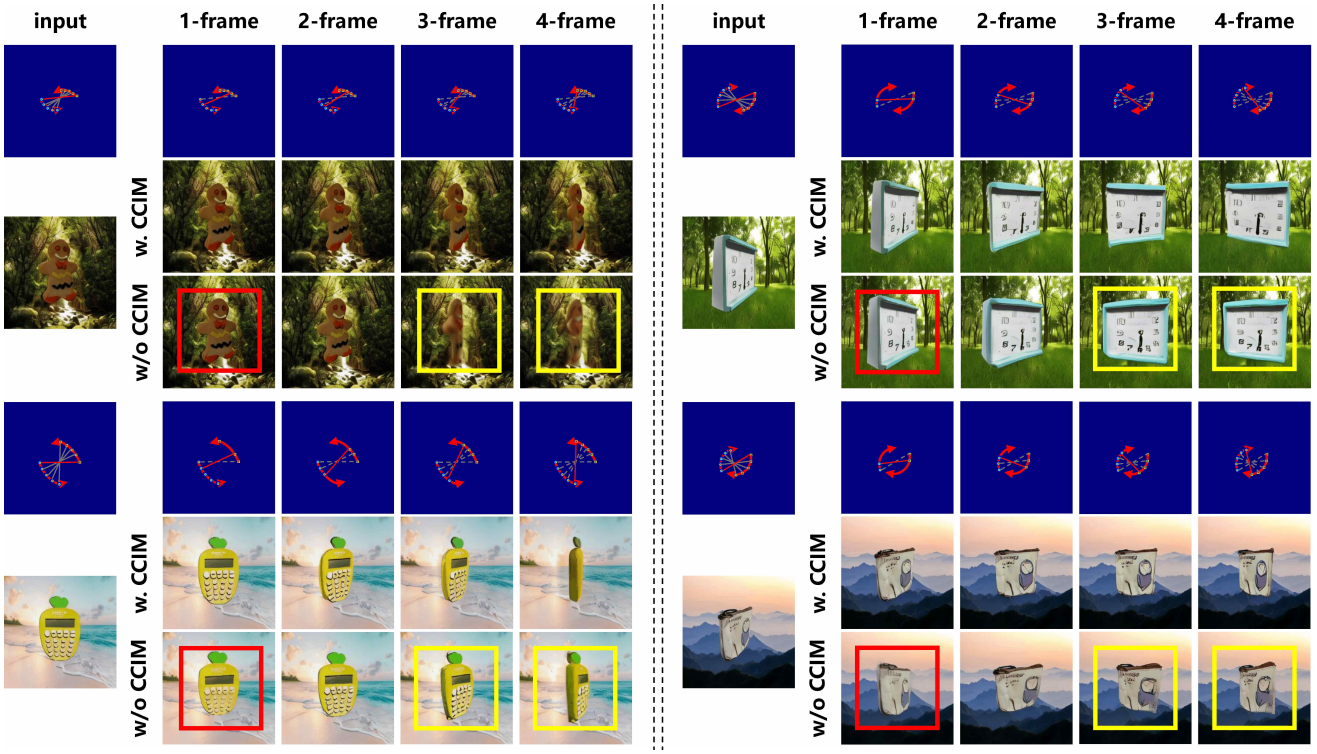


Figure 15. **Effectiveness of control condition inject module (CCIM).** We present the resulting images for more cases. The first row displays the trajectory points projected onto Gaussian maps, while the second and third rows show the objects from the corresponding view of these trajectory points, generated with and without CCIM, respectively. Without CCIM, the image remains largely unchanged in the early stages (red box), but later undergoes a sudden distortion, causing significant fidelity loss (yellow box).