Interactive3D: Create What You Want by Interactive 3D Generation

Supplementary Material

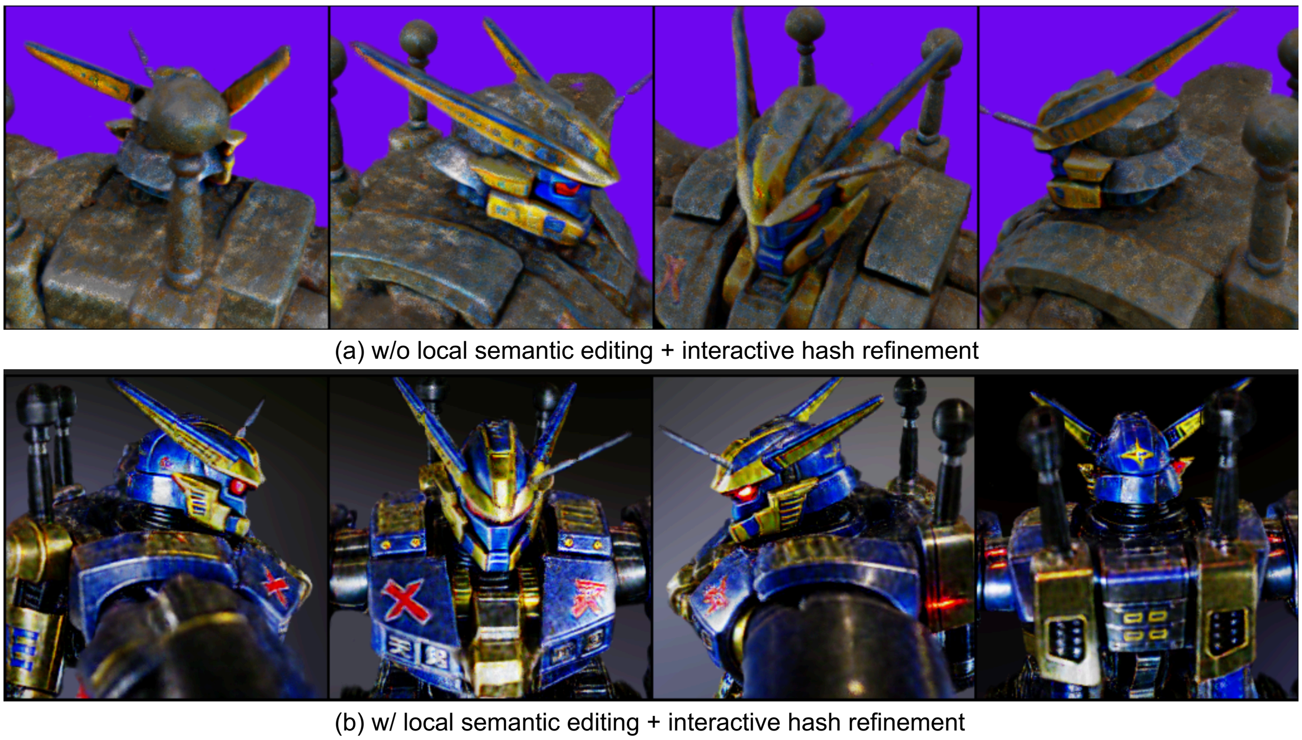

Figure 8. **The effectiveness of the combination refinement.** Refine *a Gundam robot*.

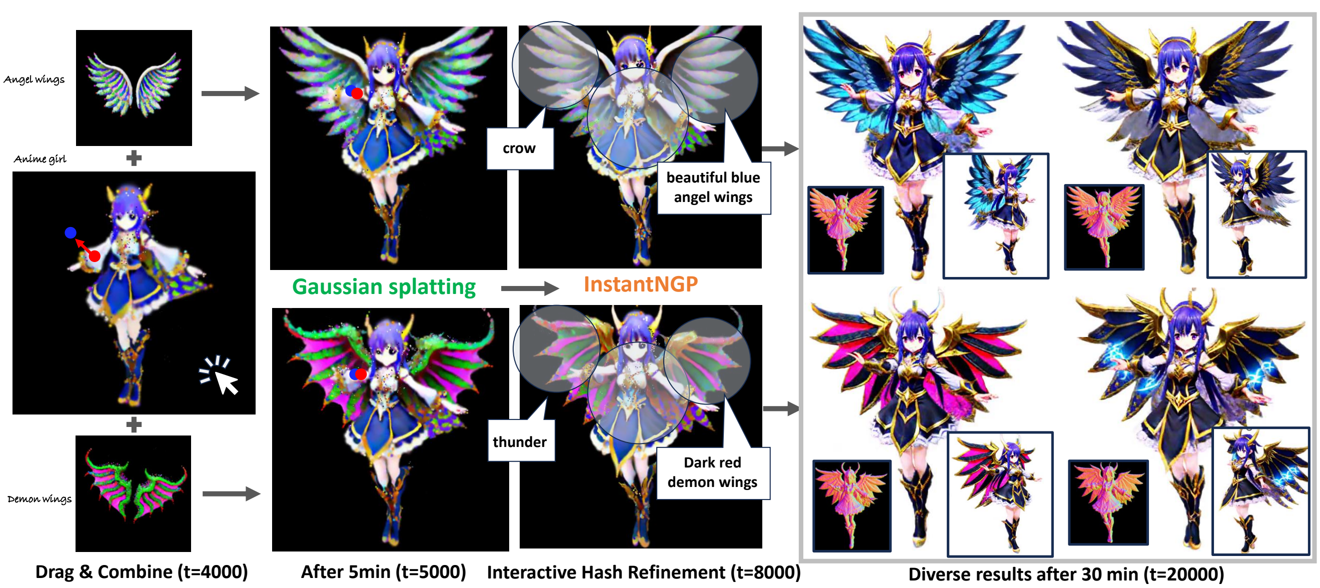

Figure 9. **Qualitative generation results of the proposed Interactive3D.** We show the interaction process of generating angel girls.

Figure 10. **Generation results of the rigid dragging operation**. We use the rigid dragging to open the dragon's mouth.

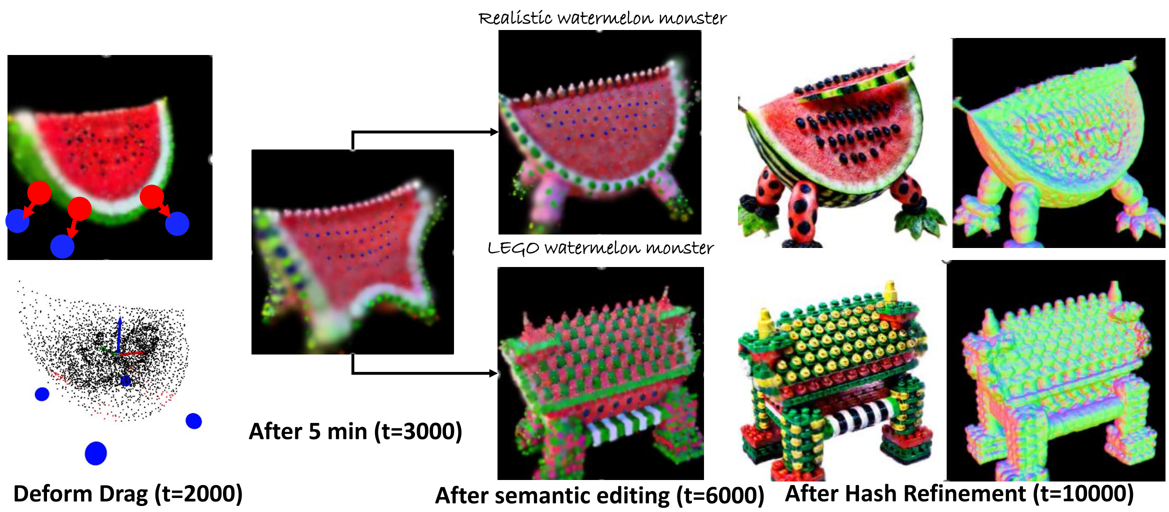

Figure 11. **Results of the proposed deformable dragging and semantic editing operations.** Our Interactive3d allows for flexible geometric deformations and semantic modifications of 3D objects.

In the supplementary material, we first introduce more implementation details in § 5.1, and then provide more results in § 5.2 and the online project page. We also provide additional ablation studies in § 5.3.

5.1. Implementation Details

We train our framework on one NVIDIA A100 GPU for 20k steps. We use the AdamW optimizer with $\beta = (0.9, 0.999)$ and the weight decay $= 0.01$. For Stage I, we utilize 4096 points sampled from Shap-E [10] to initialize the Gaussian blobs following [4]. The learning rate for position, scale, rotation, color, and alpha are set to $0.005, 0.003, 0.003, 0.01, 0.003$, respectively. We use Stable Diffusion 2.1 [20] as our 2D prior. For Stage II, by default, we set the number of levels of the refinement hash table to 8, and set the feature dimension of each position to 2. The capacity of each refinement hash table is set to 2^{19}. We render 256×256 images during training and also use Stable Diffusion 2.1 to provide guidance. Notably, we follow [27] to use the orientation loss as a regularization.

Table 2. Efficiency of using the proposed Interactive SDS loss.

Interactive SDS loss	Average Time
w/o	1.8h
w	**50min**

5.2. More Results

In this section, we give more visualization results of our framework including dragging, removing and adding parts, and interactive hash refinement. In Fig. 9, we show the final results from our entire generation framework (Stage I and Stage II) and the interaction process for generating

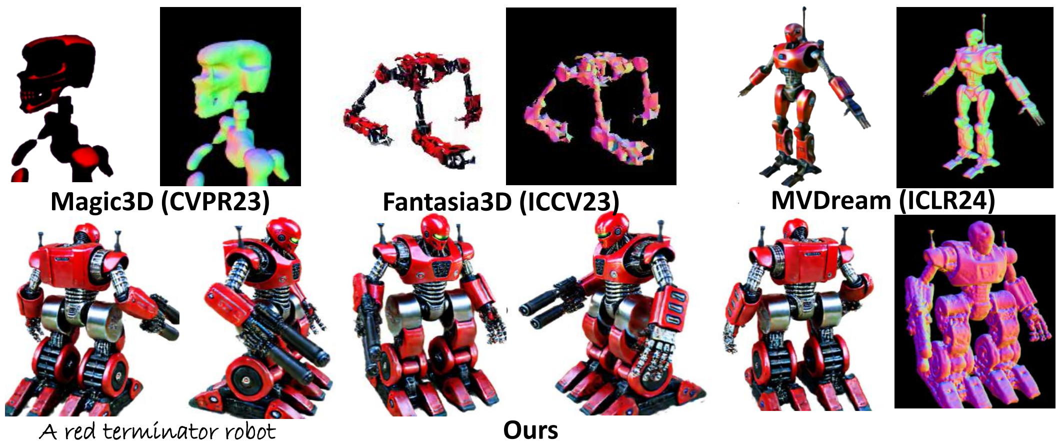

Figure 12. **More comparisons**. We compare Interactive3D with previous state-of-the-art text-to-3D methods.

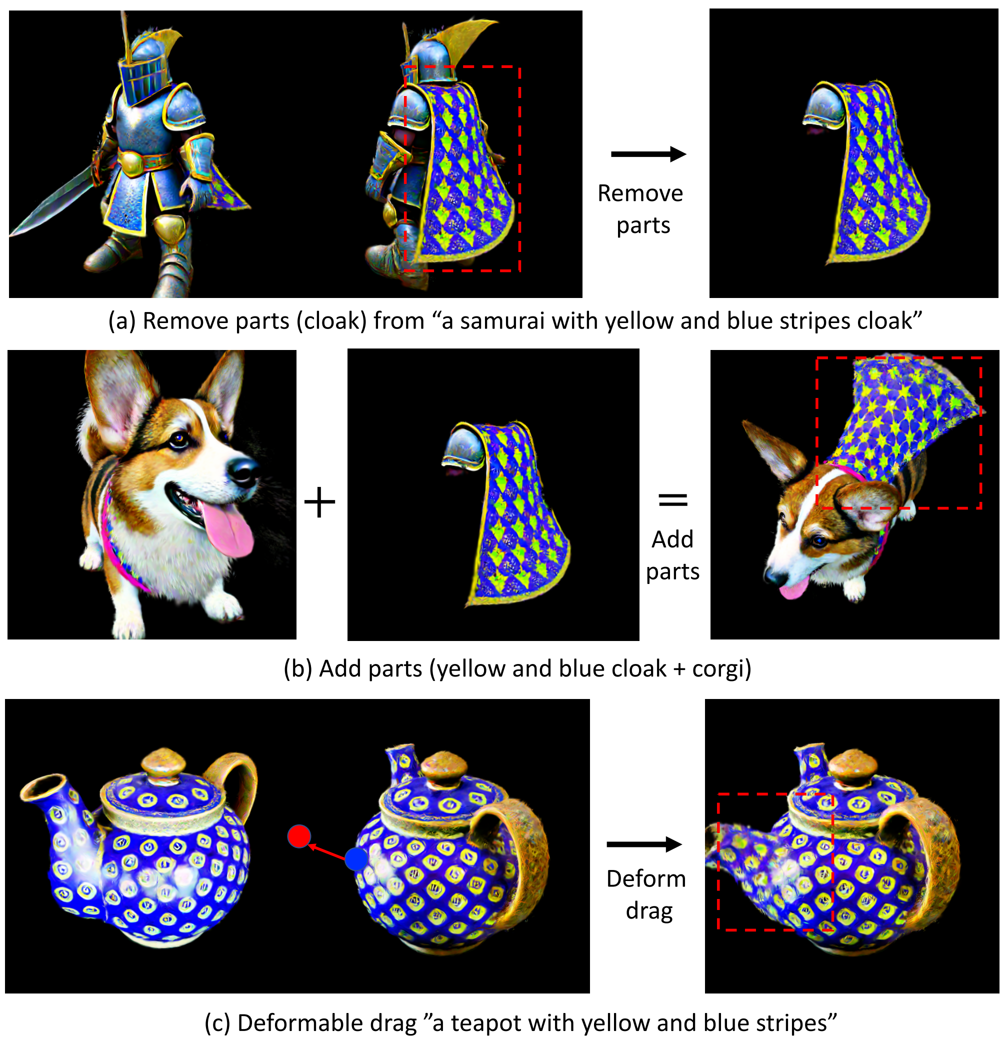

Figure 13. **Results of Interactive3D Stage I**. We show more results of dragging, removing, and adding parts. Note that all the above results after interaction (right side) are from Stage I without Stage II's refinement.

different angel girls. It can be observed that the generated two anime girls have high-quality texture and accurate geometry. In Fig. 10, we can open a dragon's mouth with the rigid dragging operation. In Fig. 11, we use the deformable dragging operation to create a watermelon monster and utilize the semantic editing operation to customize the style. In Fig. 13, we show more generated results of Stage I after user's interaction such as dragging, removing, and adding parts. Fig. 13 (a) shows that we can accurately segment (remove) the cloak from the samurai. Furthermore, the segmented cloak is combined with a corgi in Fig. 13 (b). In Fig. 13 (c), we create a new spout from the teapot by deformable dragging. Fig. 12 shows more comparisons.

In addition, we provide user studies in Tab. 3. Most users agree that interactive3D is convenient and needs fewer attempts. They can create objects better match their expectations than solely using text prompts as in previous methods.

Table 3. User studies (22 pepole).

Method	Preference↑	Avg attempts↓
text-only	4.5%	2.3
Ours	**95.5%**	**1.4**

5.3. More Ablation Results

Effect of Combination Refinement As mentioned in the main paper, semantic editing and interactive refinement can be combined to improve the quality. As demonstrated in Fig. 8, users can perform semantic editing and refinement on any local detail by the combination refinement.

Effect of Interactive SDS Loss We provided qualitative ablation experiments on the Interactive SDS Loss in Fig. 5 from the main paper. We provide a quantitative comparison for the Interactive SDS loss here. As shown in Tab. 2, the Interactive SDS loss significantly improves the efficiency of our whole pipeline.

5.4. Limitations

One limitation of our Interactive3D method is its susceptibility to failure under conditions of excessive and unreasonable manipulation (e.g., in the case of unreasonable dragging operations). This vulnerability highlights the method's reliance on user discretion for optimal outcomes. Additionally, Interactive3D is built upon the foundations of current generative techniques, inheriting their common challenges, including issues related to color saturation. This dependence means that Interactive3D is not immune to the intrinsic limitations of the underlying technologies it utilizes.