FastGrasp: Efficient Grasp Synthesis with Diffusion

Supplementary Material

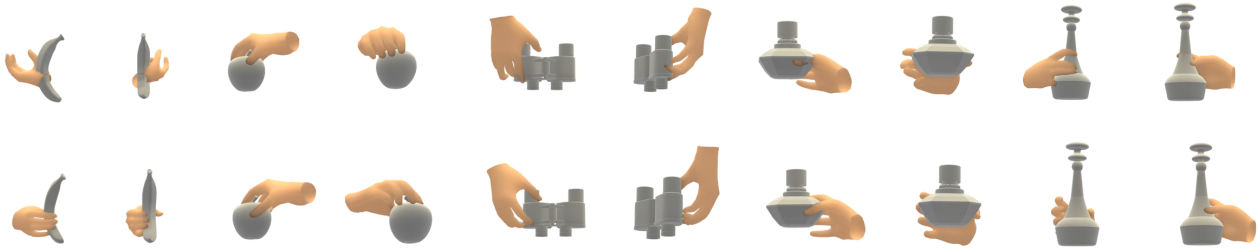

Figure 1. To assess the impact of a physically constrained loss function, we compare model performance with and without it. Each pair of columns shows generated grasps from two distinct views. The first row uses only the reconstruction loss, while the second row presents results from our proposed pipeline. Our method significantly reduces object penetration compared to using the reconstruction loss alone.

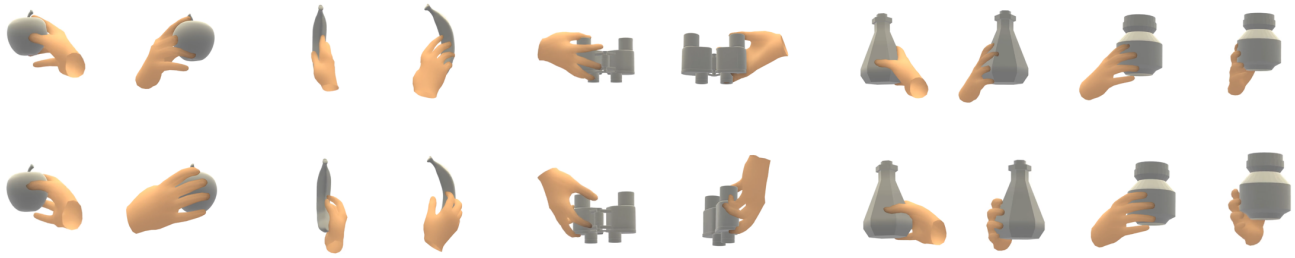

Figure 2. To evaluate the necessity of hand vertices as inputs, we visualize the model's output using both hand parameters and hand vertices. Each pair of columns shows generated grasps from two different views. The first row presents results with hand parameter input, while the second row displays results from our pipeline. Our method enhances performance by capturing hand joint details and improving rotational accuracy, which reduces object penetration.

OakInk	Simulation Displacement ↓	Penetration Distance ↓	Penetration Volume ↓	Contact Ratio ↑
No-physical-loss	1.91	0.93	4.76	96
Hand param	**1.39**	**0.91**	5.91	**98**
Ours	1.83	**0.91**	**2.39**	**98**

Table 1. We conducted ablation experiments to evaluate the impact of the physical constraints loss function and hand vertices.

1. Overview of Material

The supplementary material comprehensively details our experiments, results, and visualizations. Tab. 1 examines the impact of physical constraints during autoencoder training and compares the effects of hand verts versus hand parameters as inputs. Sec. 2.3 offers additional visualizations to enhance understanding of our model.

2. More Autoencoder Experimental Results

In training the autoencoder, we use hand vertices as input and apply both reconstruction and physical loss functions. Sec. 2.1 and Sec. 2.2 examine the effects of training the model with hand vertices and reconstruction loss alone versus using MANO parameters with both reconstruction and physical loss functions in Tab. 1.

2.1. Training Using Reconstruction Loss

The model is trained using hand vertices h_v as input and relies solely on the reconstruction loss function, without incorporating any physical loss function. As shown in Fig. 1, experiments reveal that using only the reconstruction loss often results in significant penetration and displacement issues in hand-object interactions. However, as demonstrated in Tab. 1, incorporating a physical constraint loss function improves the model's ability to capture these details, reducing physical collisions and enhancing grasp stability.

2.2. Training Using Mano Parameter

The model is trained using hand parameters h_p as input. Our experiments indicate that using hand vertices instead of MANO parameters results in less physical volume intrusion. As shown in Fig. 2 and Tab. 1, this is attributed to the Hand vertices providing a more robust data representation than MANO parameters, reducing the model's sensitivity to input variations and thus improving training effectiveness.

2.3. Autoencoder Visulization Result

To validate the effectiveness of our autoencoder model, we provide extensive visualizations in Fig. 3 and 4.

Fig. 3 illustrates two grasping poses for randomly selected test objects. This demonstrates that our model adheres to physical constraints in hand-object interactions for various grasps of the same object. Fig. 4 showcases grasping poses for objects with diverse geometric shapes from the test set, highlighting our model's ability to generate effective grasps across different objects consistently.

Figure 3. In the visualization results of the autoencoder, we selected two different grasping poses for each object, each shown from two different perspectives.

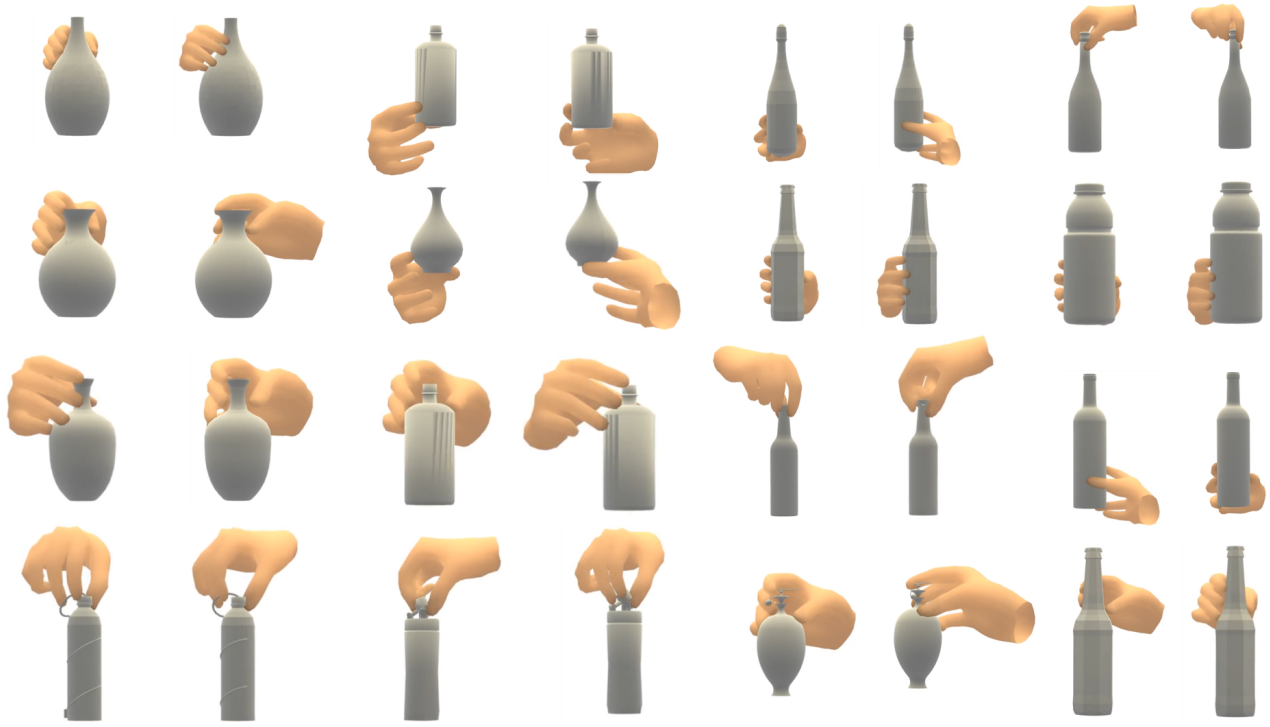

Figure 4. In the autoencoder visualization results, we randomly selected grasping poses, each shown from two different perspectives.