Depth Restoration for Hand-Held Transparent Object

Ran Yu, Liguang Ruan

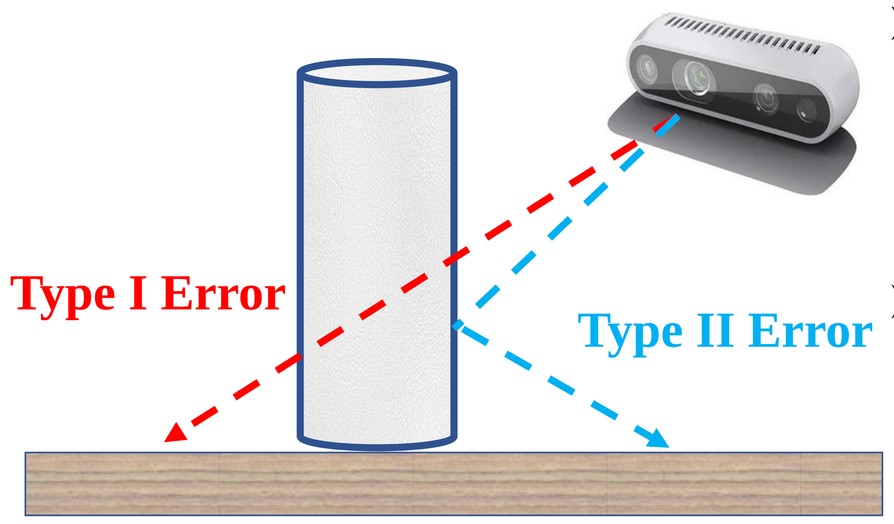
Tsinghua-Berkeley Shenzhen Institute Tsinghua Shenzhen International Graduation School, Tsinghua University

Abstract

- Challenges of home service robot: special object perception (e.g. **transparent**) and human-robot interaction (e.g. **handover**).
- One common scenario: how to allow robot percept and grasp a **transparent object from human's hand**.
- Our approach: a **sim2real method** for hand-held transparent object restoration.
- Two main contributions: synthetic dataset **HandTrans-14K** and **hand-aware depth restoration method**.

Motivation

Two types of error in the depth estimation of transparent objects:

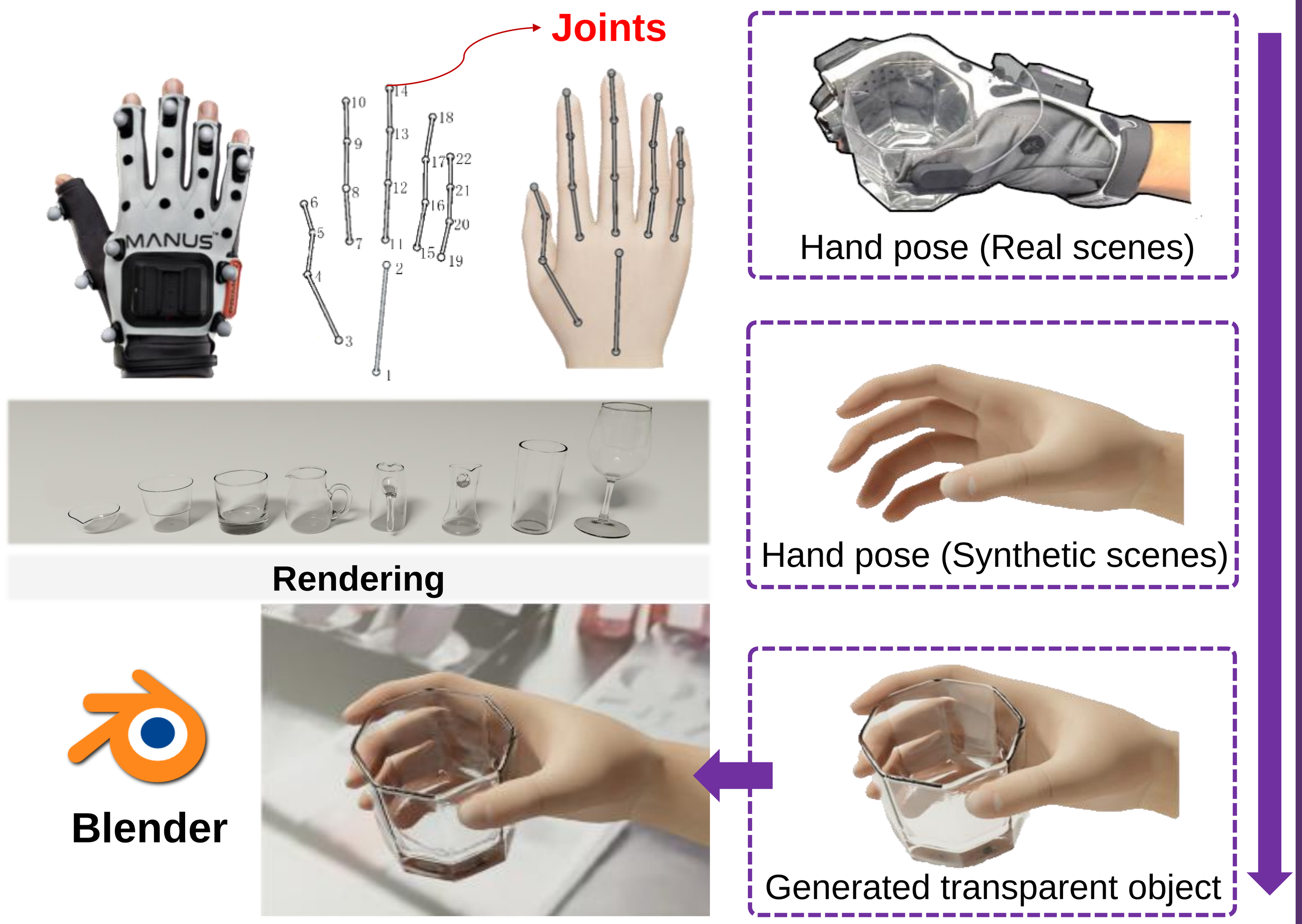


- **Type I Error:** Light refracts through object and reflects back background depth.
- **Type II Error:** Light reflect on transparent surface, which leads to missing depth.

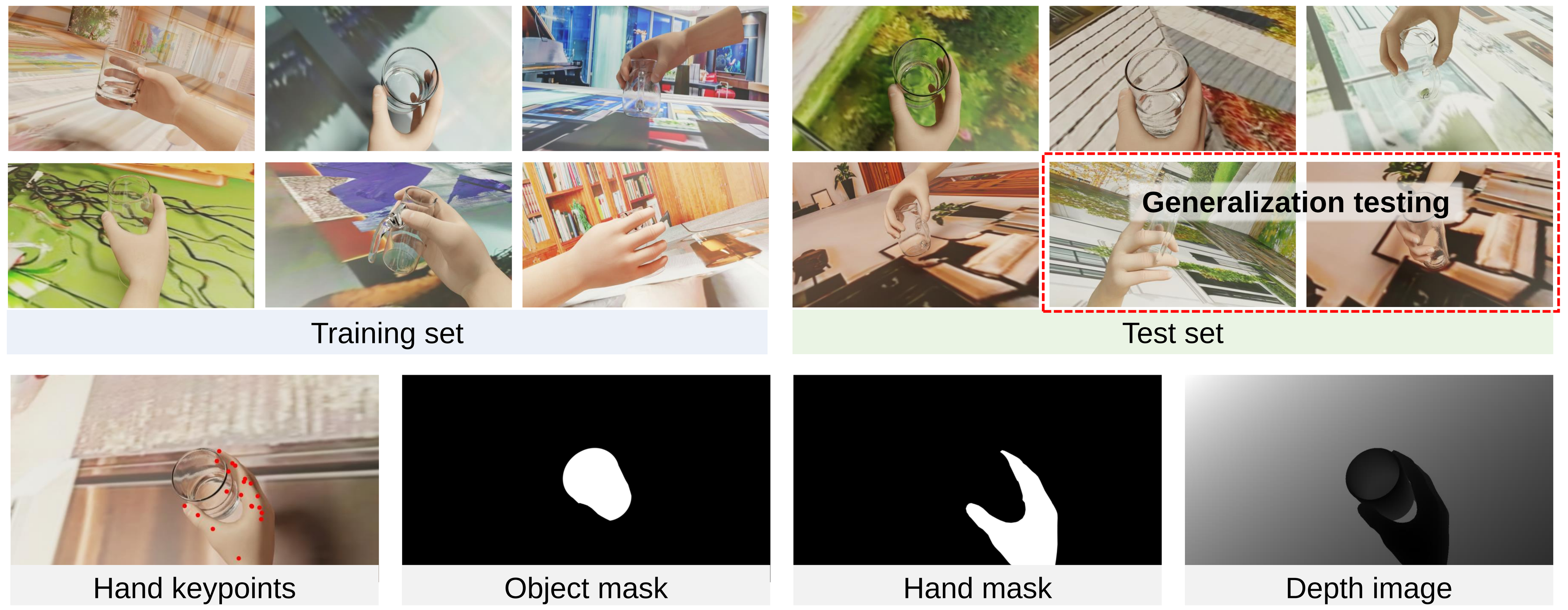
To avoid errors in depth estimation, **depth restoration models** are designed: generating perfect depth based on RGB and corrupted depth input.

Early works [1-3] focus on transparent objects placed on the table. However, it is common for robot to percept grasped transparent object in everyday life. Our work aims to solve this problem.

Real2Sim Pose Generation



Dataset Generation



Method

1. Local Implicit Function

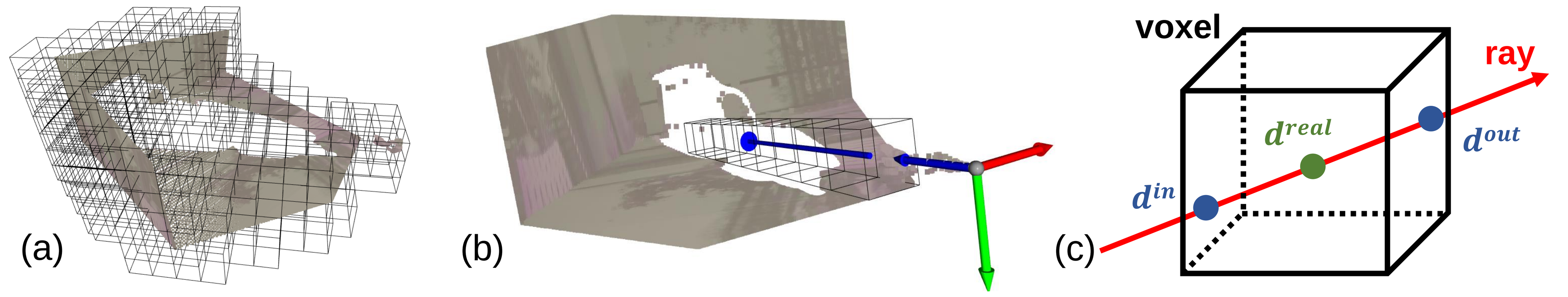
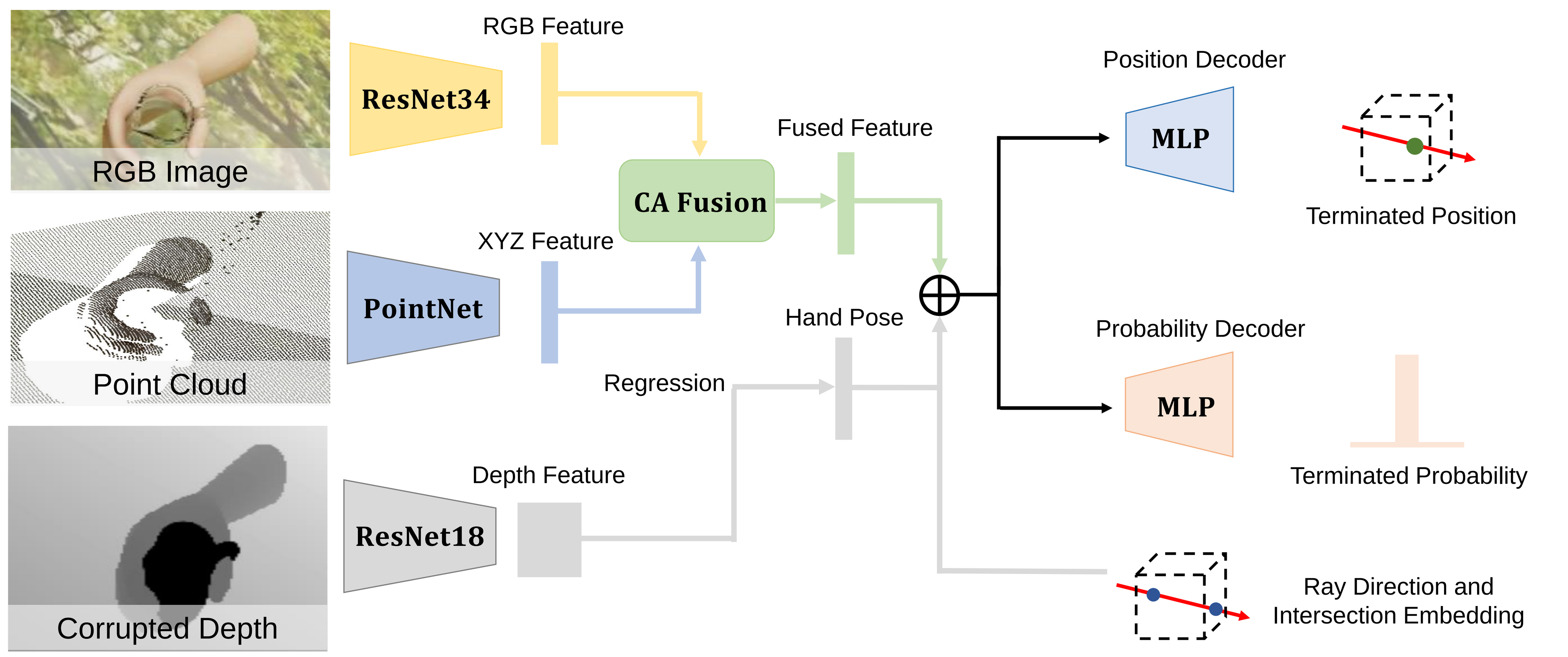


Illustration of local implicit function method. (a) Point cloud placed in voxels. (b) Recording rays and intersecting voxels for missing points. (c) Predicting the position of possible point clouds for given ray-voxel pairs.

2. Hand-aware Depth Restoration



Hand-Aware Implicit Function of Depth Restoration Model

Discussion

- Hand pose information can facilitate the performance of depth restoration, especially for the model generalization ability for unseen objects.
- Future works: better strategy of different modality fusion and sim-to-real experiment to verify real world performance of proposed method.

Results

1. Ablation Test: Influence of different hand feature forms.

Feature Types	RMSE↓	REL↓	MAE↓	$\delta_{1.05}$↑	$\delta_{1.10}$↑	$\delta_{1.25}$↑
No hand feature	0.014	0.029	0.010	85.18	96.08	99.60
2d hand feature	0.013	0.027	0.009	87.12	96.57	99.60
Relative 3d feature	0.013	0.027	0.009	87.54	96.61	99.53
3d hand feature	**0.011**	**0.020**	**0.007**	**92.11**	**97.67**	**99.74**

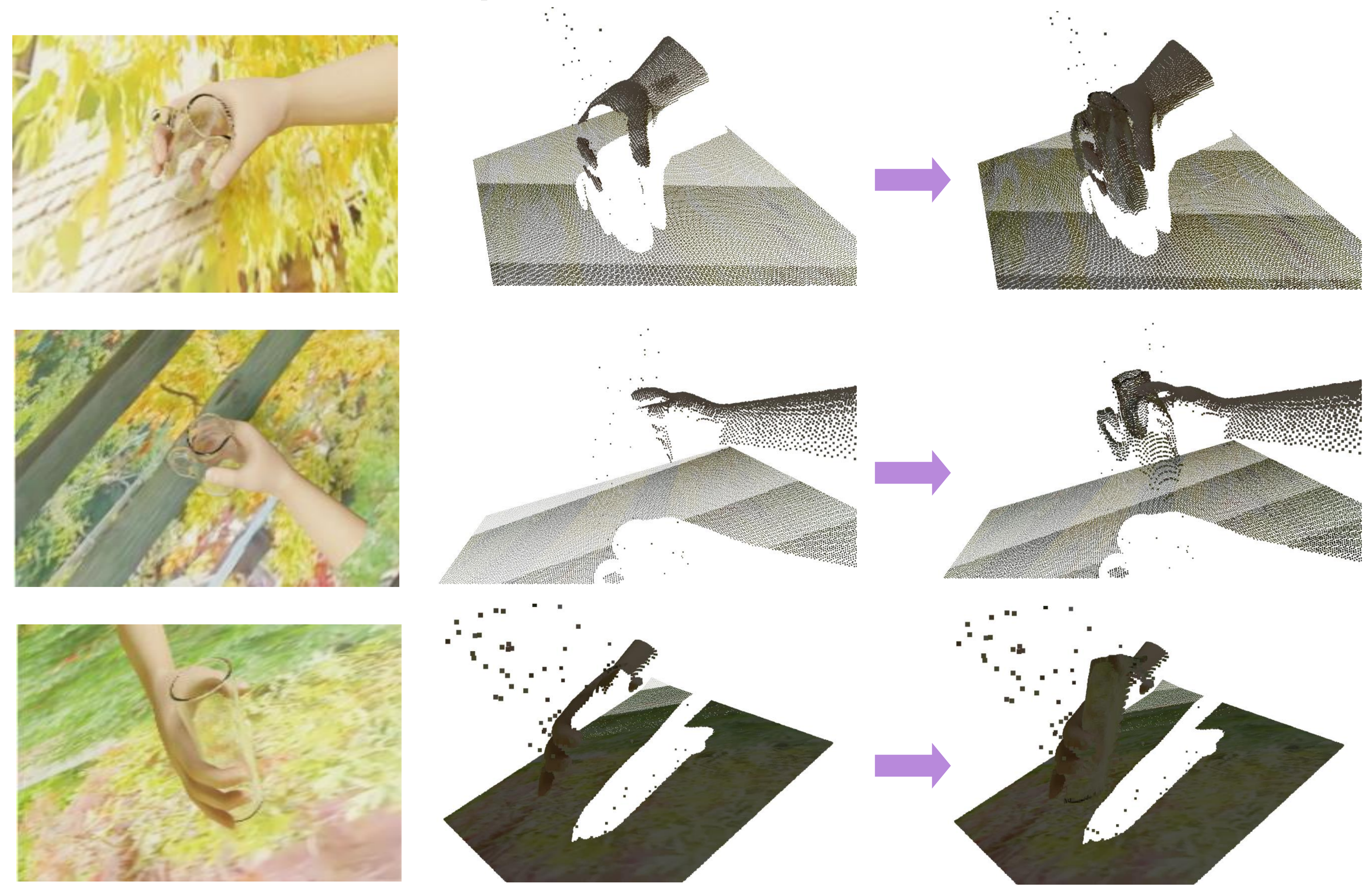
2. Seen Objects Evaluation: Model performance on seen objects with known shapes in HandTrans-14K dataset.

Methods	RMSE↓	REL↓	MAE↓	$\delta_{1.05}$↑	$\delta_{1.10}$↑	$\delta_{1.25}$↑
TODE-Trans	0.024	0.056	0.017	57.03	82.55	98.61
TransCG	0.012	0.026	0.008	86.91	96.96	99.76
SwinDRNet	**0.009**	**0.015**	**0.005**	**94.82**	**98.26**	**99.87**
LIDF	0.014	0.029	0.010	85.18	96.08	99.60
Ours	0.011	0.020	0.007	92.11	97.67	99.74

3. Unseen Objects Evaluation: Model generalization ability on unseen objects with novel shapes in HandTrans-14K dataset.

Methods	RMSE↓	REL↓	MAE↓	$\delta_{1.05}$↑	$\delta_{1.10}$↑	$\delta_{1.25}$↑
TODE-Trans	0.052	0.055	0.037	62.86	89.26	97.21
TransCG	0.027	0.064	0.020	51.54	78.54	98.79
SwinDRNet	0.022	0.049	0.015	65.07	86.79	98.79
LIDF	0.026	0.063	0.021	52.66	78.69	97.90
Ours	**0.019**	**0.042**	**0.014**	**72.74**	**90.06**	**99.05**

4. Visualization of Depth Restoration:



Reference

[1] Zhu L, Mousavian A, et al. RGB-D Local Implicit Function for Depth Completion of Transparent Objects. CVPR, 2021.

[2] Tang Y, Chen J, et al. Depthgrasp: Depth Completion of Transparent Objects Using Self-Attentive Adversarial Network with Spectral Residual for Grasping. IROS, 2021.

[3] Li T, Chen Z, et al. FDCT: Fast Depth Completion for Transparent Objects. IEEE Robotics and Automation Letters, 2023.