

Visual-Tactile Sensing for In-Hand Object Reconstruction

Wenqiang Xu*^{1,2}, Zhenjun Yu*¹, Han Xue¹, Ruolin Ye³, Siqiong Yao¹, Cewu Lu^{1,2}

CVPR 2023

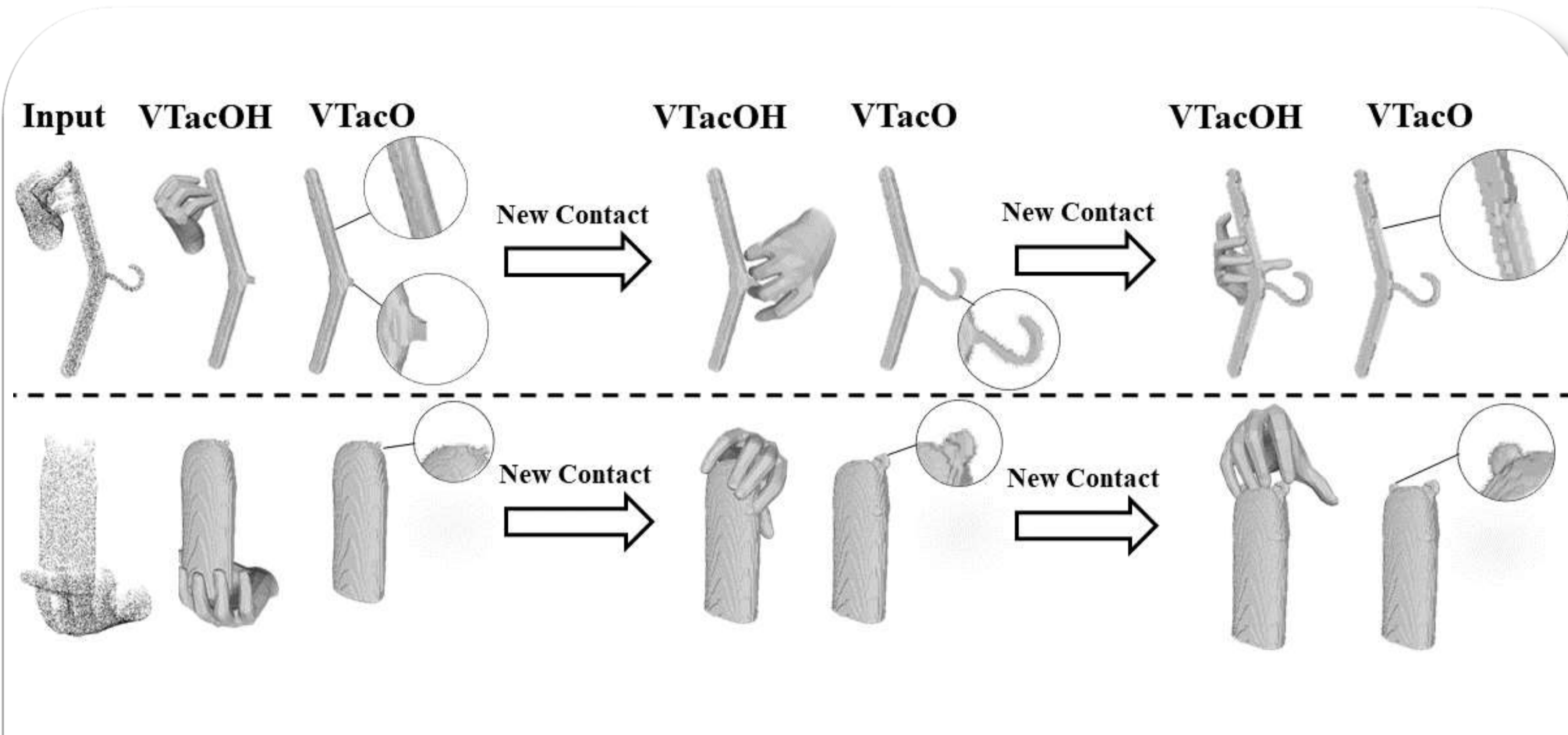
¹Shanghai Jiao Tong University ²Shanghai Qi Zhi institute ³Cornell University



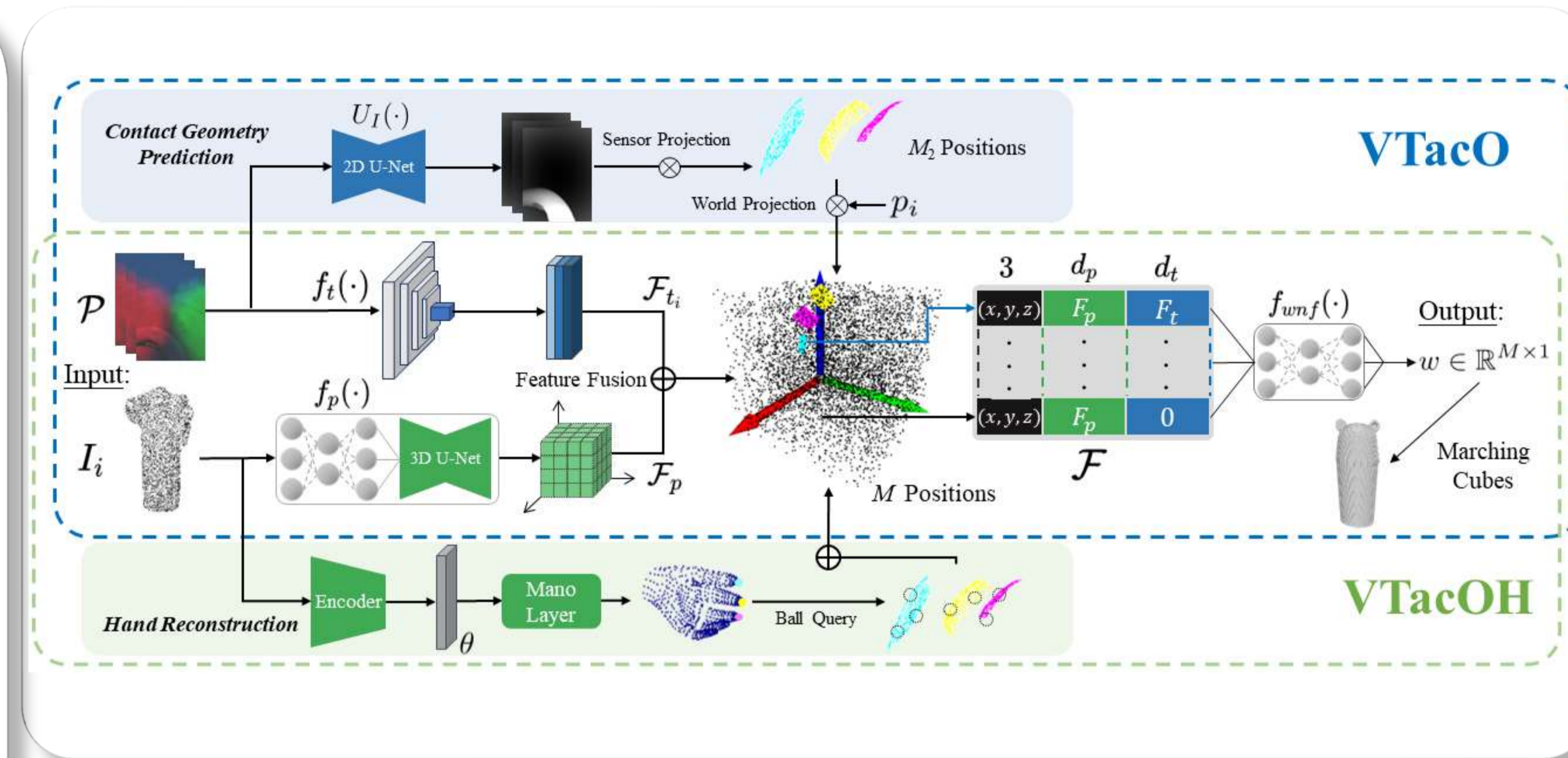
* indicates equal contributions

Visual-Tactile Sensing for In-Hand Object Reconstruction

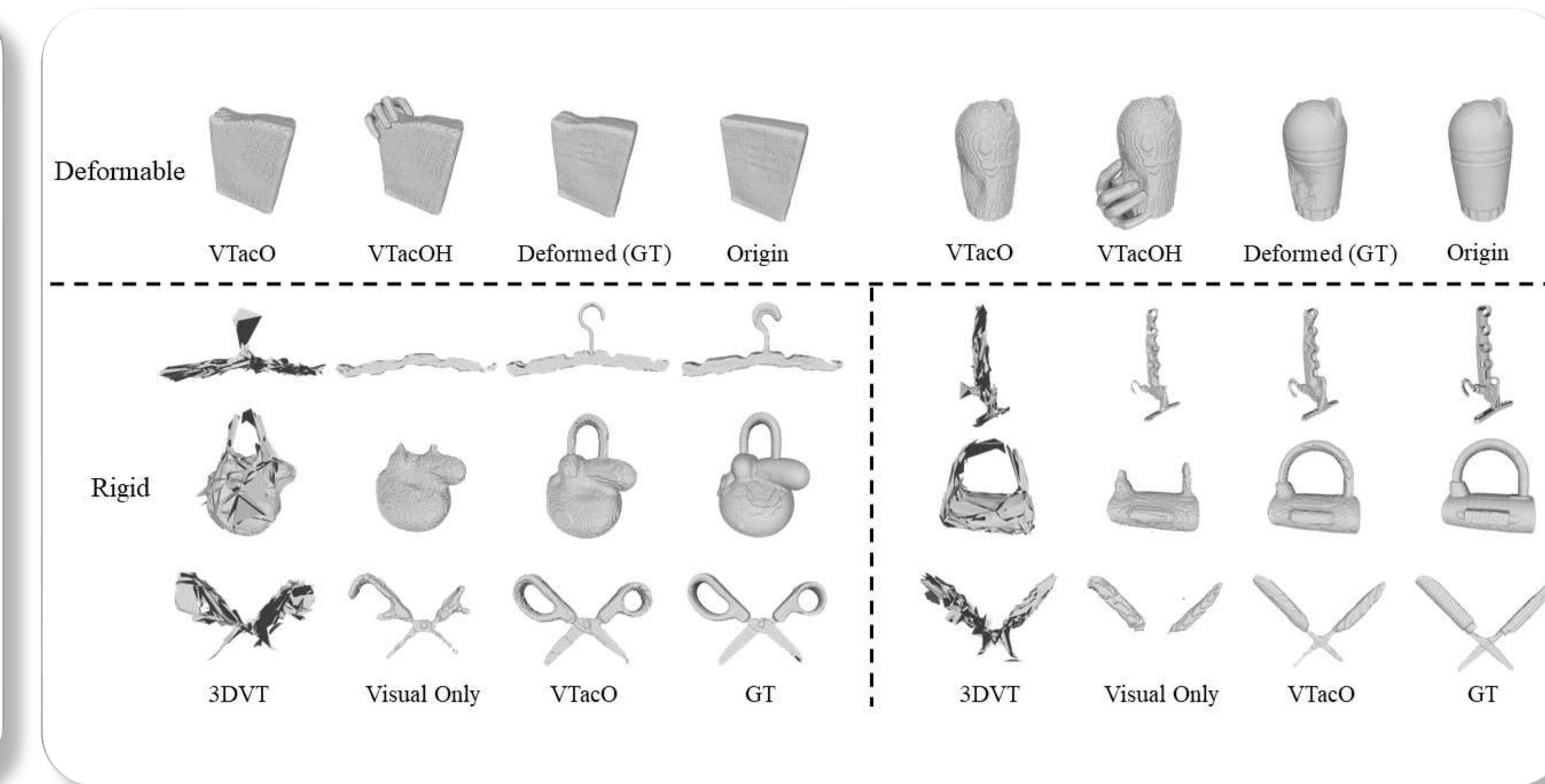
Introduction



Framework of VTacO and VTacOH



Qualitative Results



Main contributions:

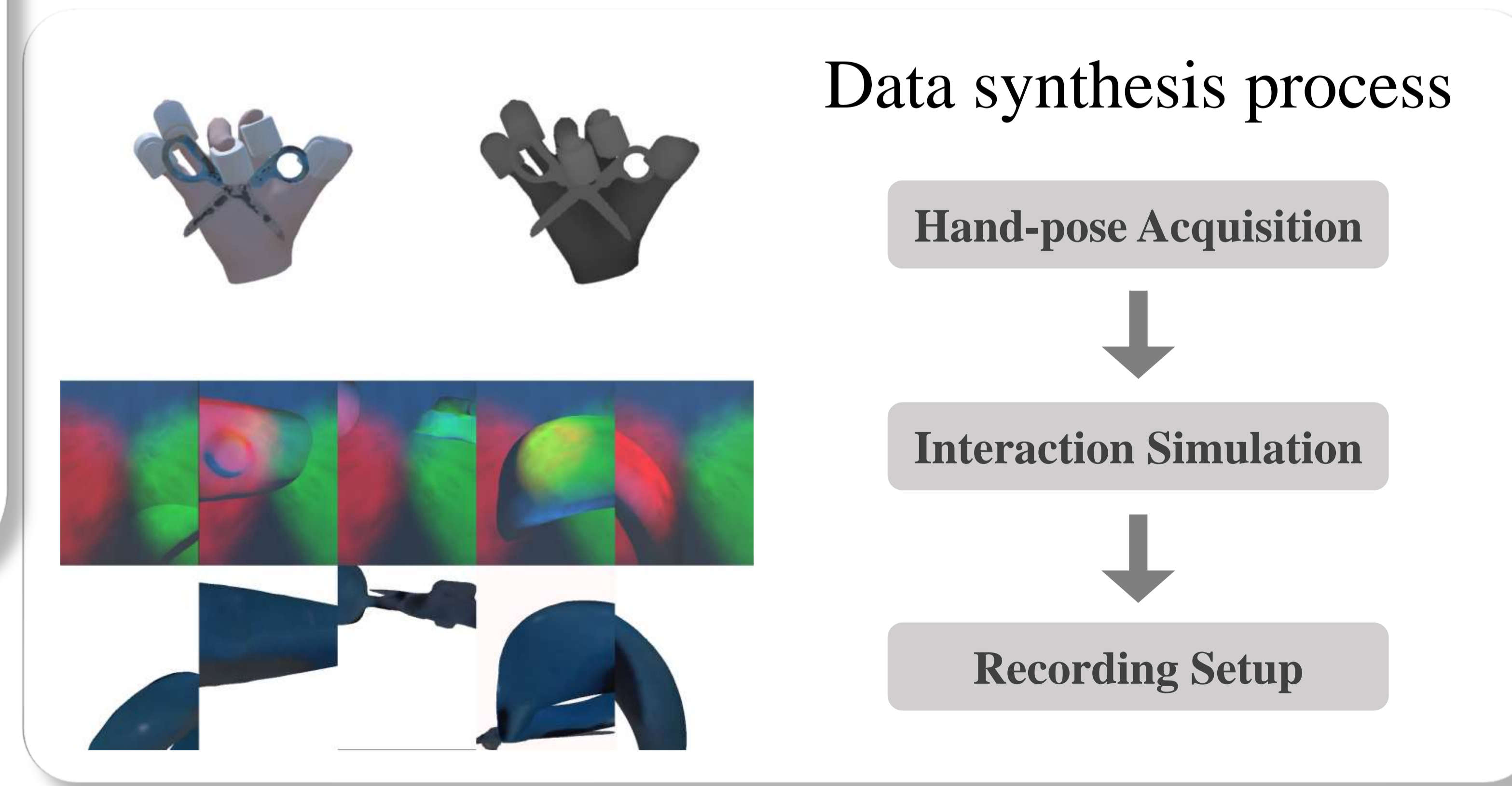
Visual-tactile Learning Framework

- Visual-Tactile in-hand Object reconstruction: VTacO
- Hand-Object version VTacOH

Simulation Environment: VT-Sim

- Based on Unity
- Simulation of hand-object interaction
- Acquisition of training samples $T = \{T_i\}_{i=1}^N$, including WNF, depth images, tactile signals, etc.

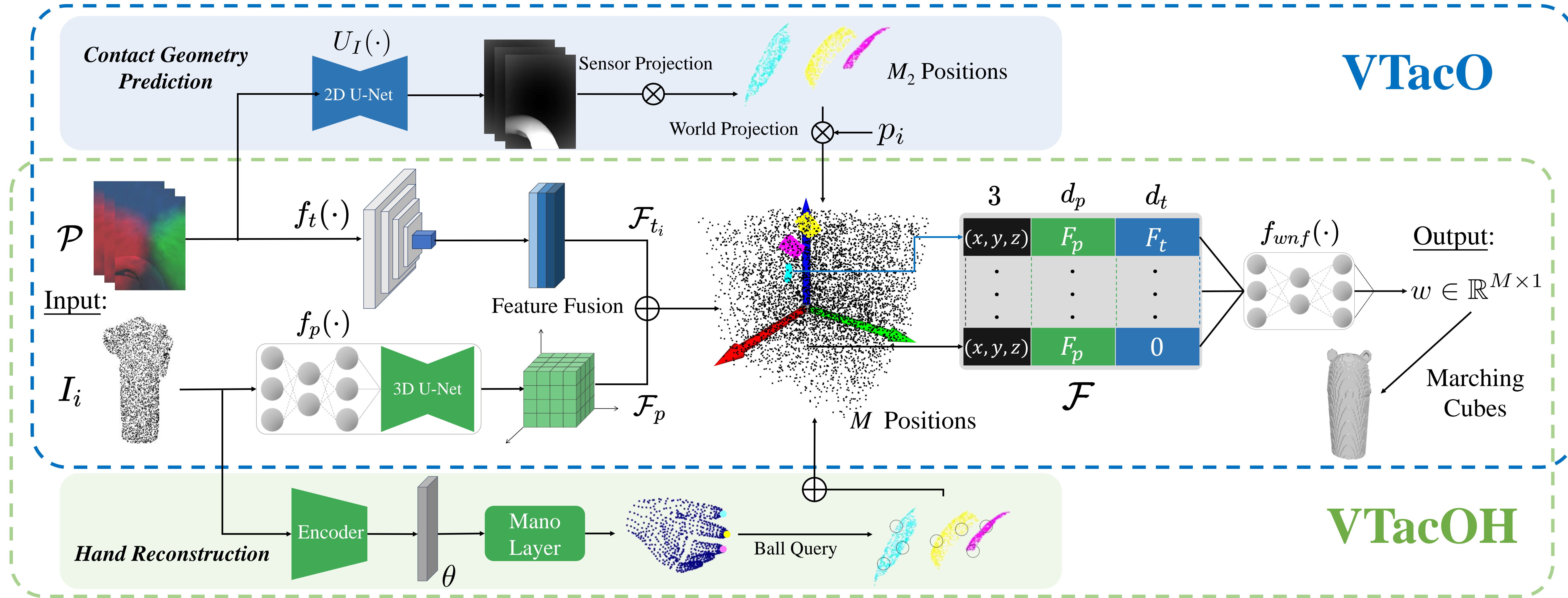
VT-Sim



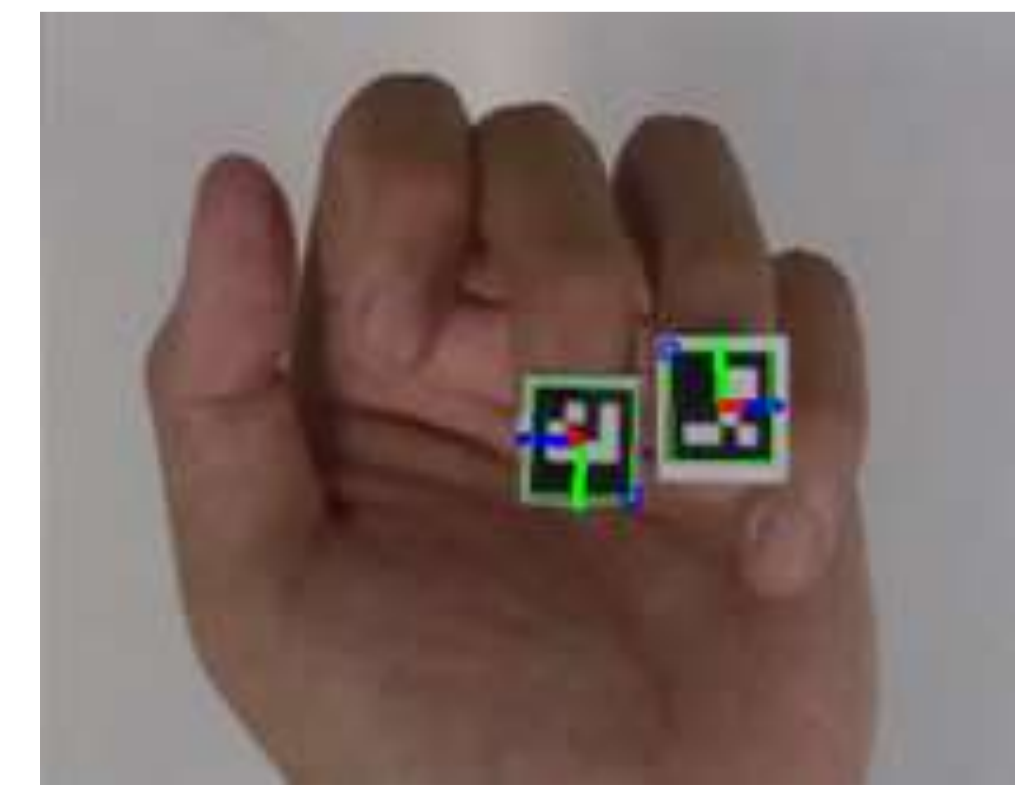
Quantitative Results

Metrics	Method	bottle	box	foldingrack	lock	scissor	mean
IoU	3D Vision and Touch	*	*	0.746	0.468	0.691	0.622
	Ours (Vision only)	0.884	0.934	0.837	0.877	0.763	0.857
	Ours (VTacO)	0.887	0.950	0.782	0.916	0.777	0.860
	Ours (VTacOH)	0.886	0.947	0.765	0.911	0.772	0.855
CD	3D Vision and Touch	*	*	0.242	2.631	3.206	1.268
	Ours (Vision only)	1.109	0.459	1.579	1.140	7.549	1.472
	Ours (VTacO)	0.936	0.305	1.360	0.932	0.894	0.798
	Ours (VTacOH)	0.948	0.312	1.432	0.955	0.945	0.916
EMD	3D Vision and Touch	*	*	0.052	0.175	0.081	0.090
	Ours (Vision only)	0.054	0.028	0.026	0.083	0.110	0.051
	Ours (VTacO)	0.059	0.018	0.028	0.012	0.052	0.028
	Ours (VTacOH)	0.056	0.024	0.027	0.035	0.068	0.042

Framework of VTacO and VTacOH



VTacOH: Extended Version of VTacO



Simulation Environment: VT-Sim

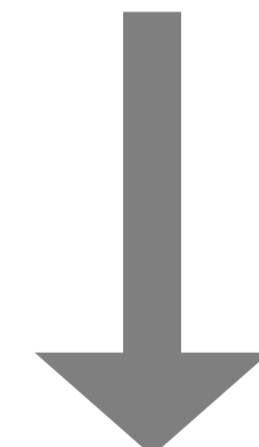
Data synthesis process



Hand-pose Acquisition



Interaction Simulation



Recording Setup

GraspIt!

RigidBody

Obi SoftBody

Point cloud

Tactile image

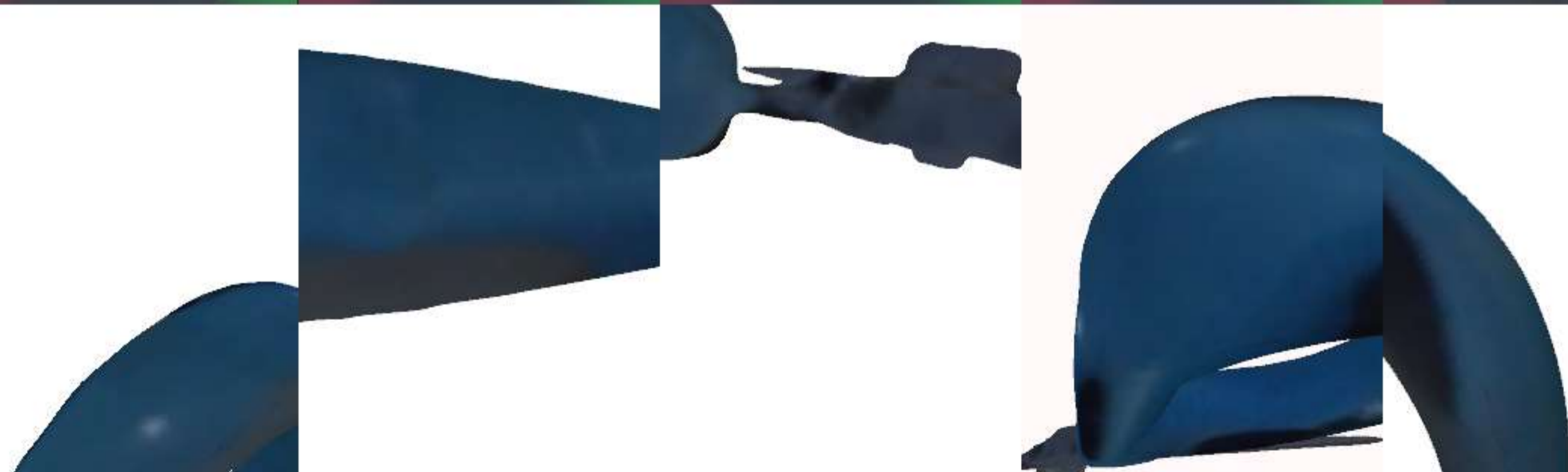
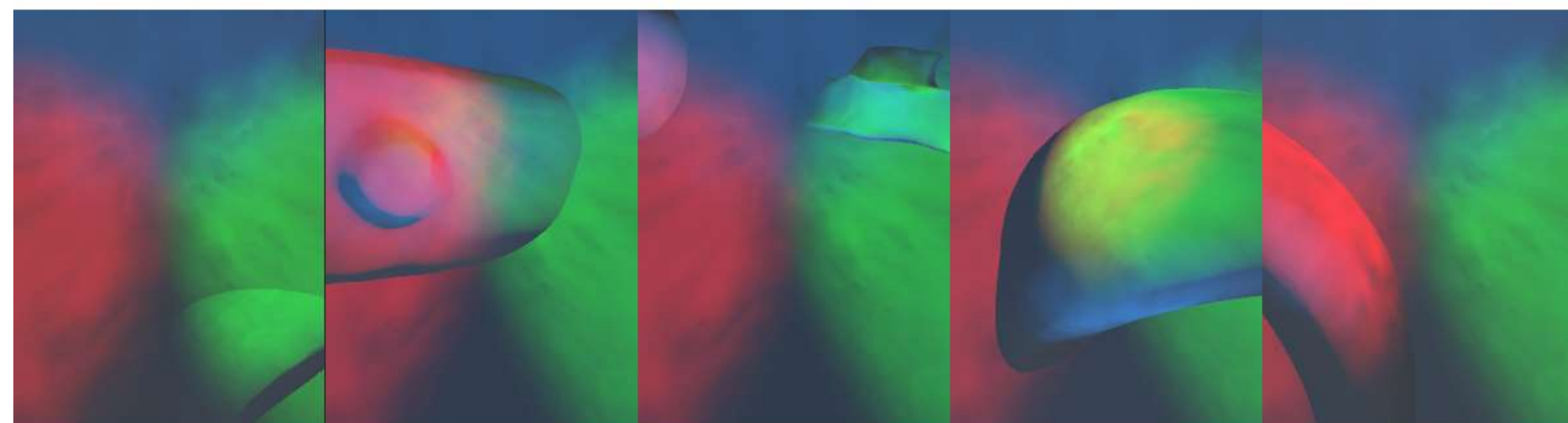
WNF



YCB Objects



AKB-48 Objects

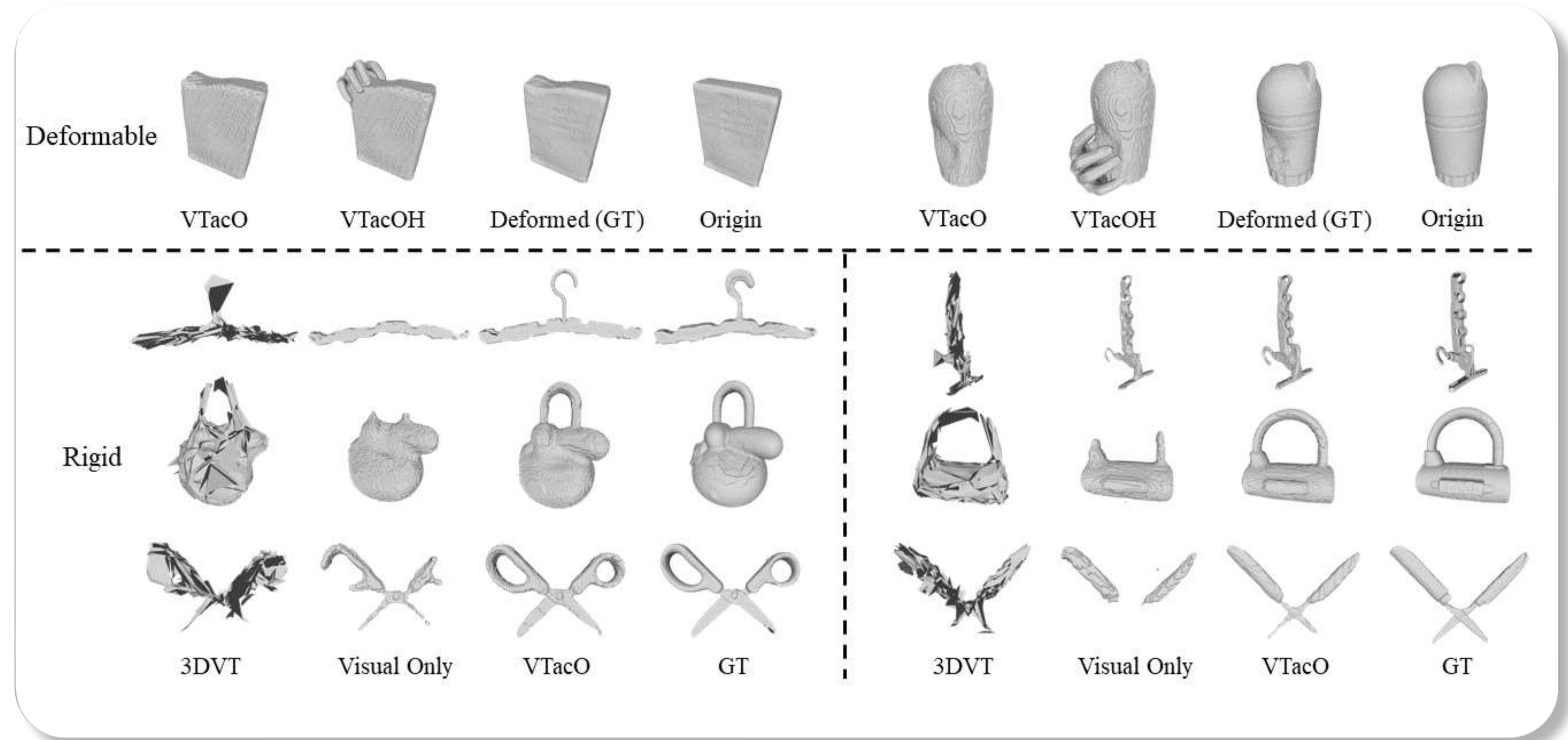


Experiments & Ablation Study

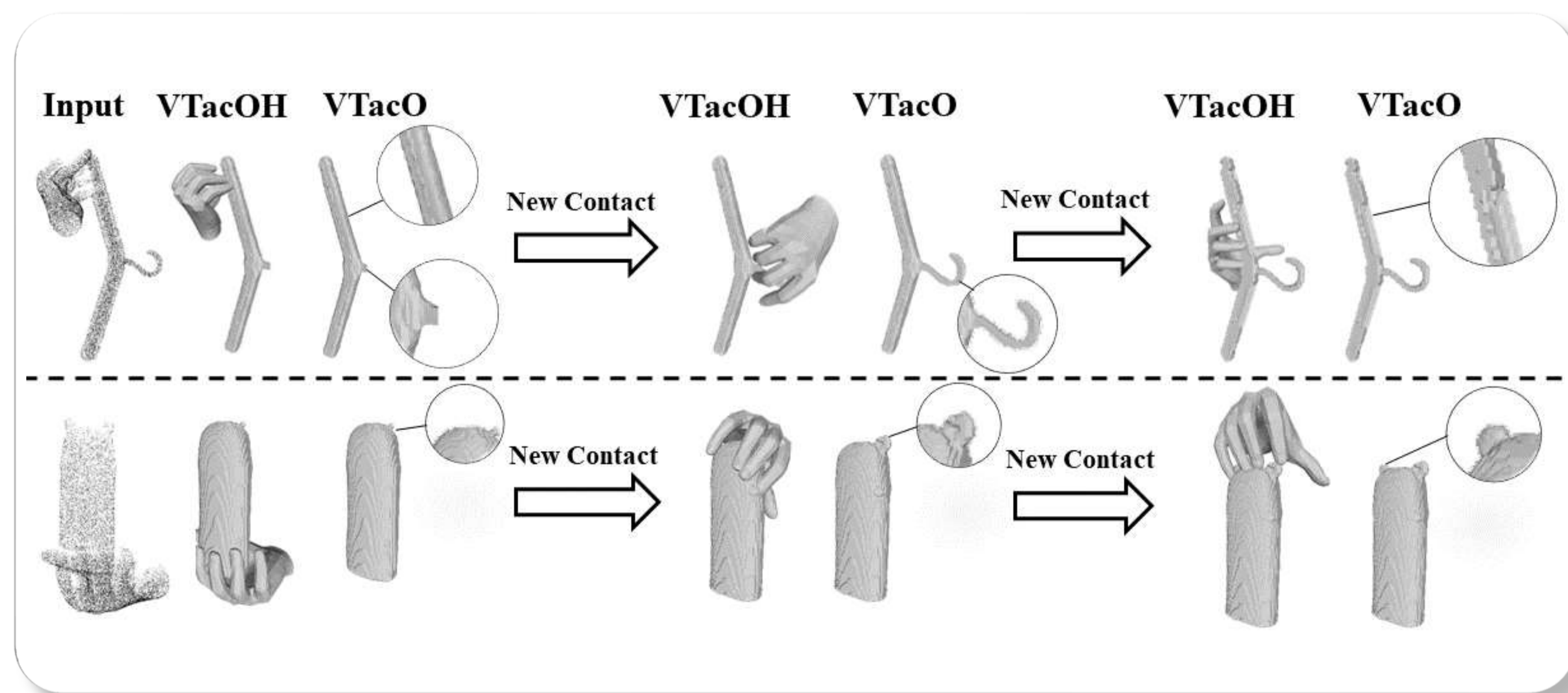
Quantitative Results

Metrics	Method	bottle	box	foldingrack	lock	scissor	mean
IoU	3D Vision and Touch	*	*	0.746	0.468	0.691	0.622
	Ours (Vision only)	0.884	0.934	0.837	0.877	0.763	0.857
	Ours (VTacO)	0.887	0.950	0.782	0.916	0.777	0.860
	Ours (VTacOH)	0.886	0.947	0.765	0.911	0.772	0.855
CD	3D Vision and Touch	*	*	0.242	2.631	3.206	1.268
	Ours (Vision only)	1.109	0.459	1.579	1.140	7.549	1.472
	Ours (VTacO)	0.936	0.305	1.360	0.932	0.894	0.798
	Ours (VTacOH)	0.948	0.312	1.432	0.955	0.945	0.916
EMD	3D Vision and Touch	*	*	0.052	0.175	0.081	0.090
	Ours (Vision only)	0.054	0.028	0.026	0.083	0.110	0.051
	Ours (VTacO)	0.059	0.018	0.028	0.012	0.052	0.028
	Ours (VTacOH)	0.056	0.024	0.027	0.035	0.068	0.042

Qualitative Results



Procedural Tactile Embedding



Comparison with pure vision

Procedural Tactile Embedding

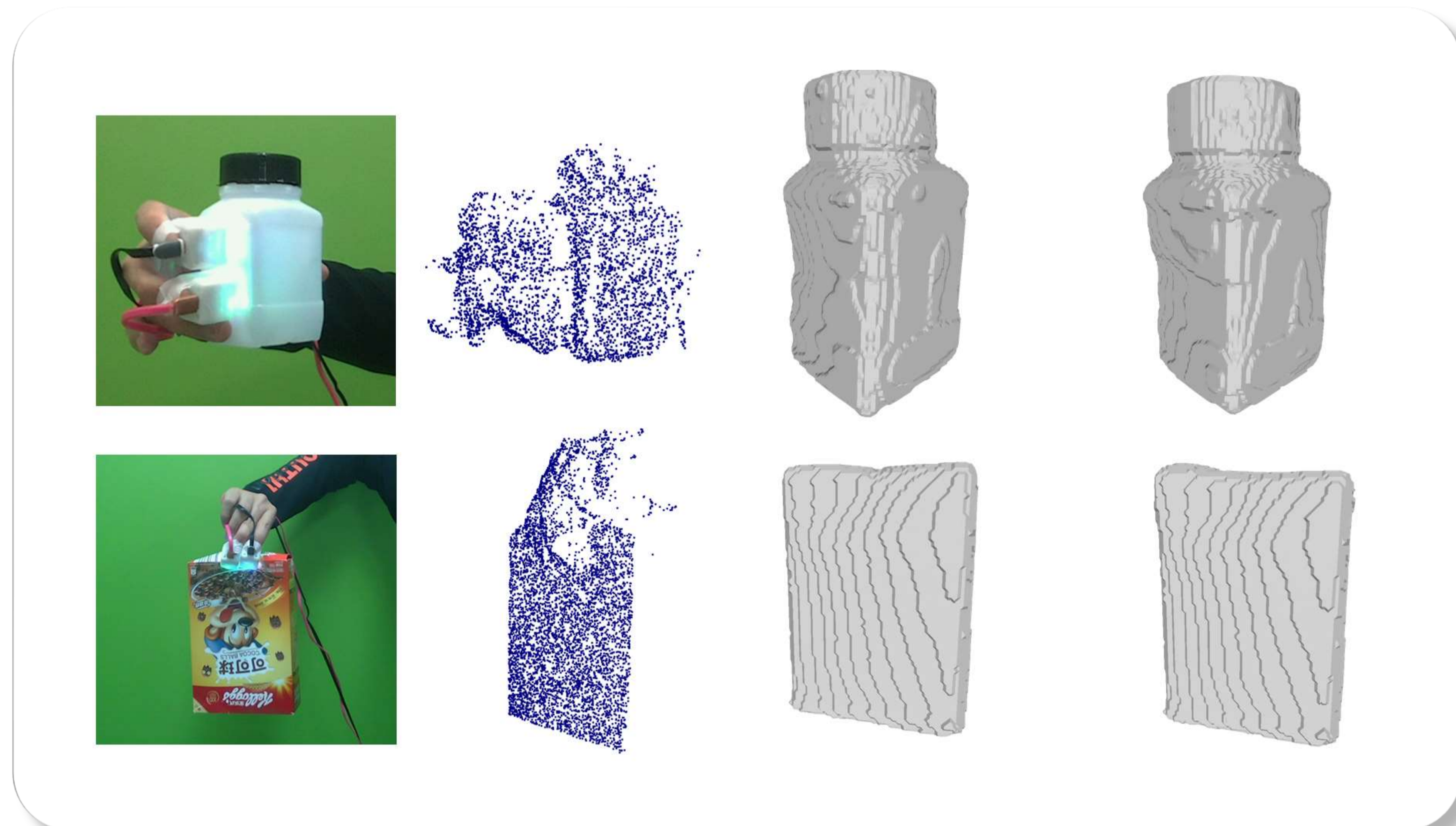
Ablation Study

Visual Feature Encoder

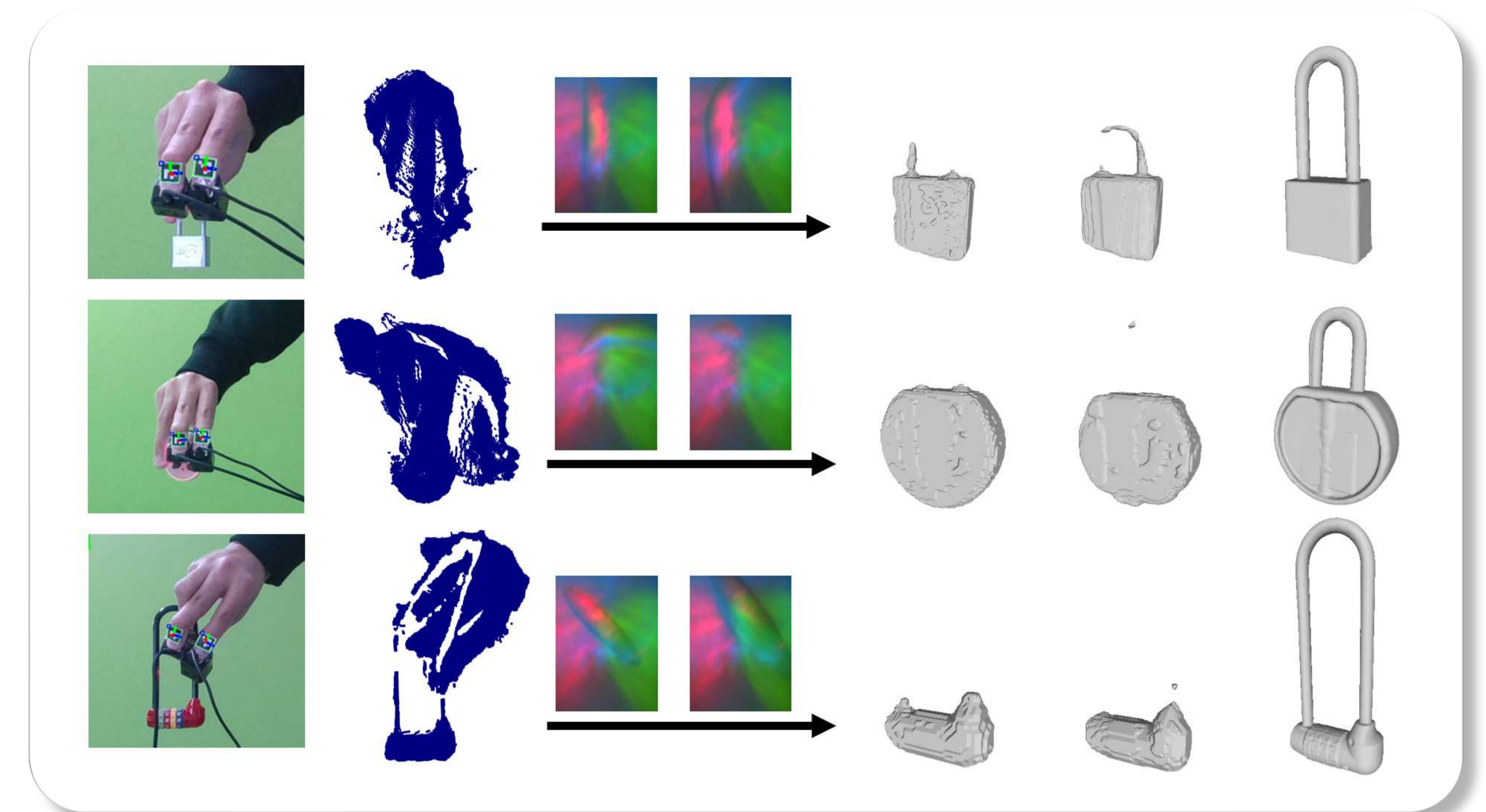
Feature Fusion Strategies

Real-world Experiment

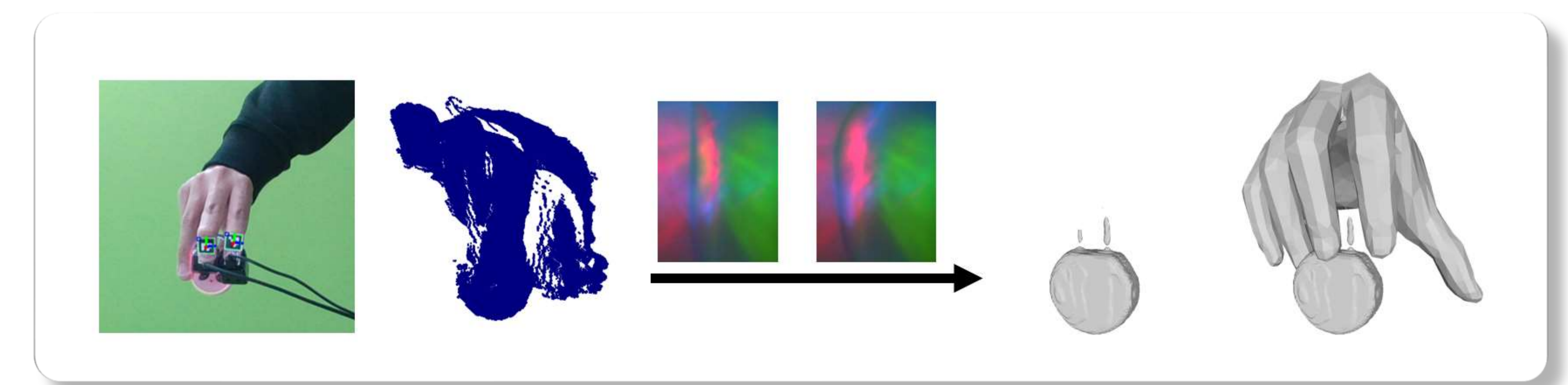
VTacO in real world with boxes and bottles



VTacO in real world with locks



VTacOH with locks

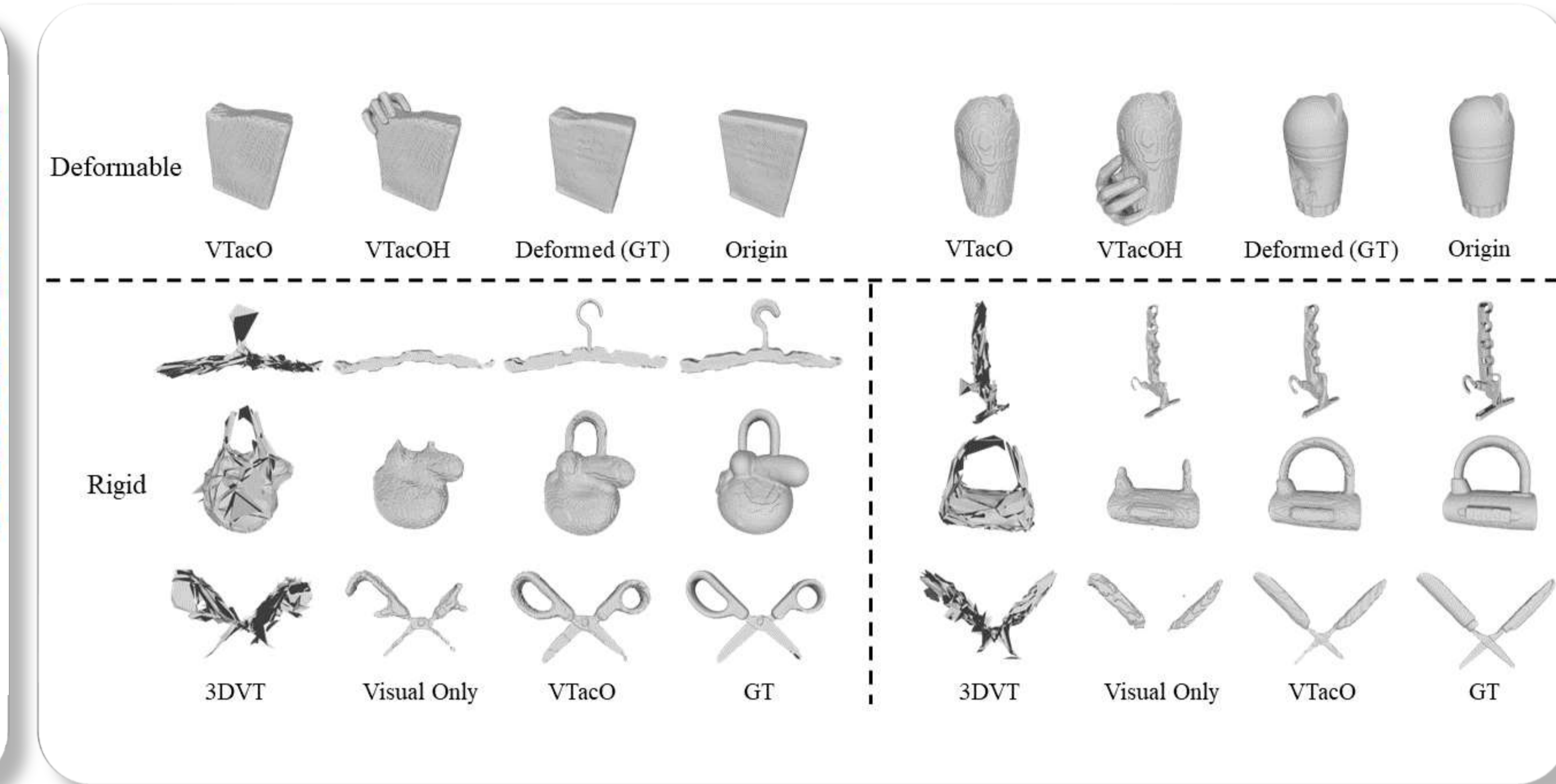
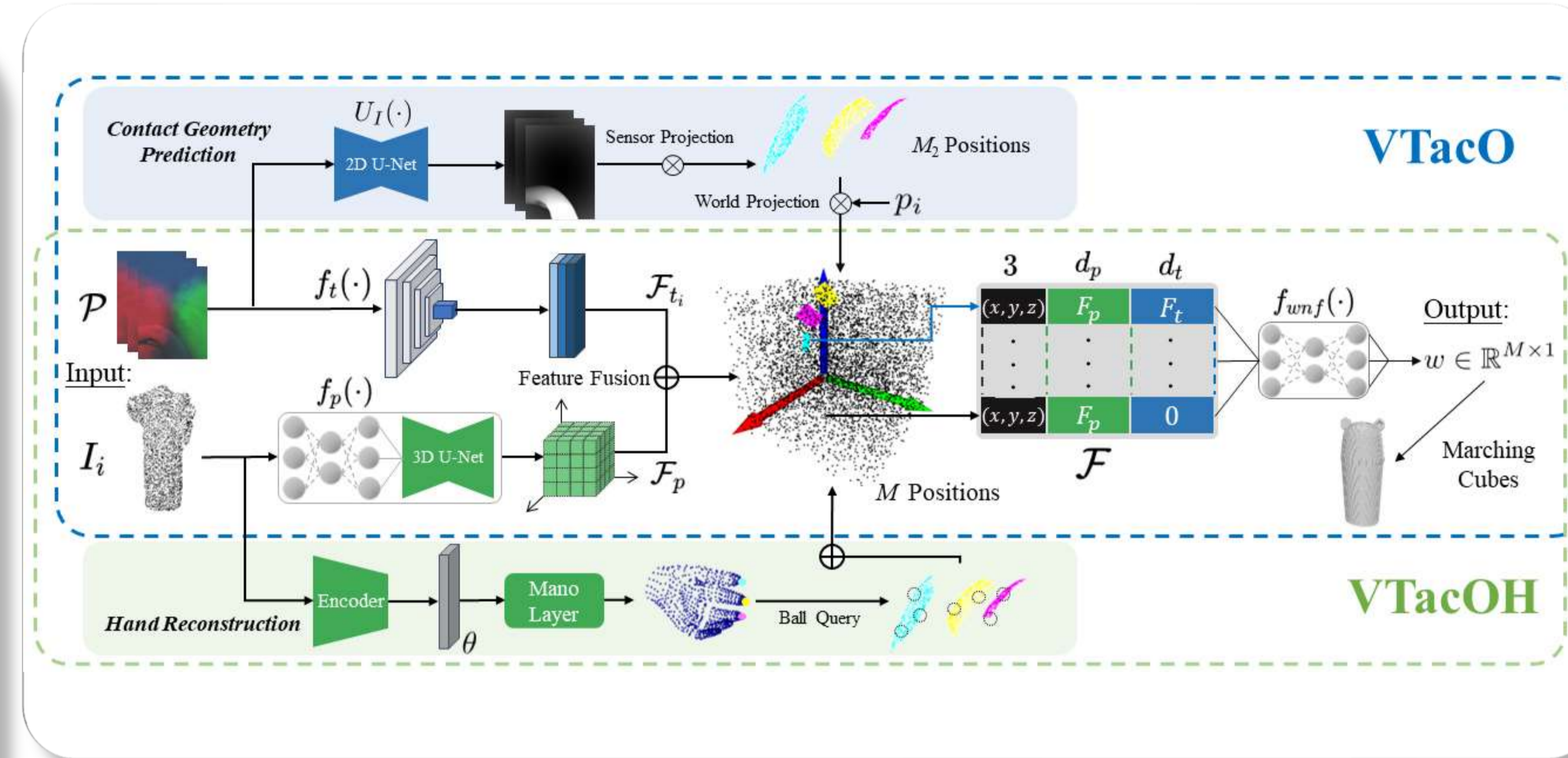
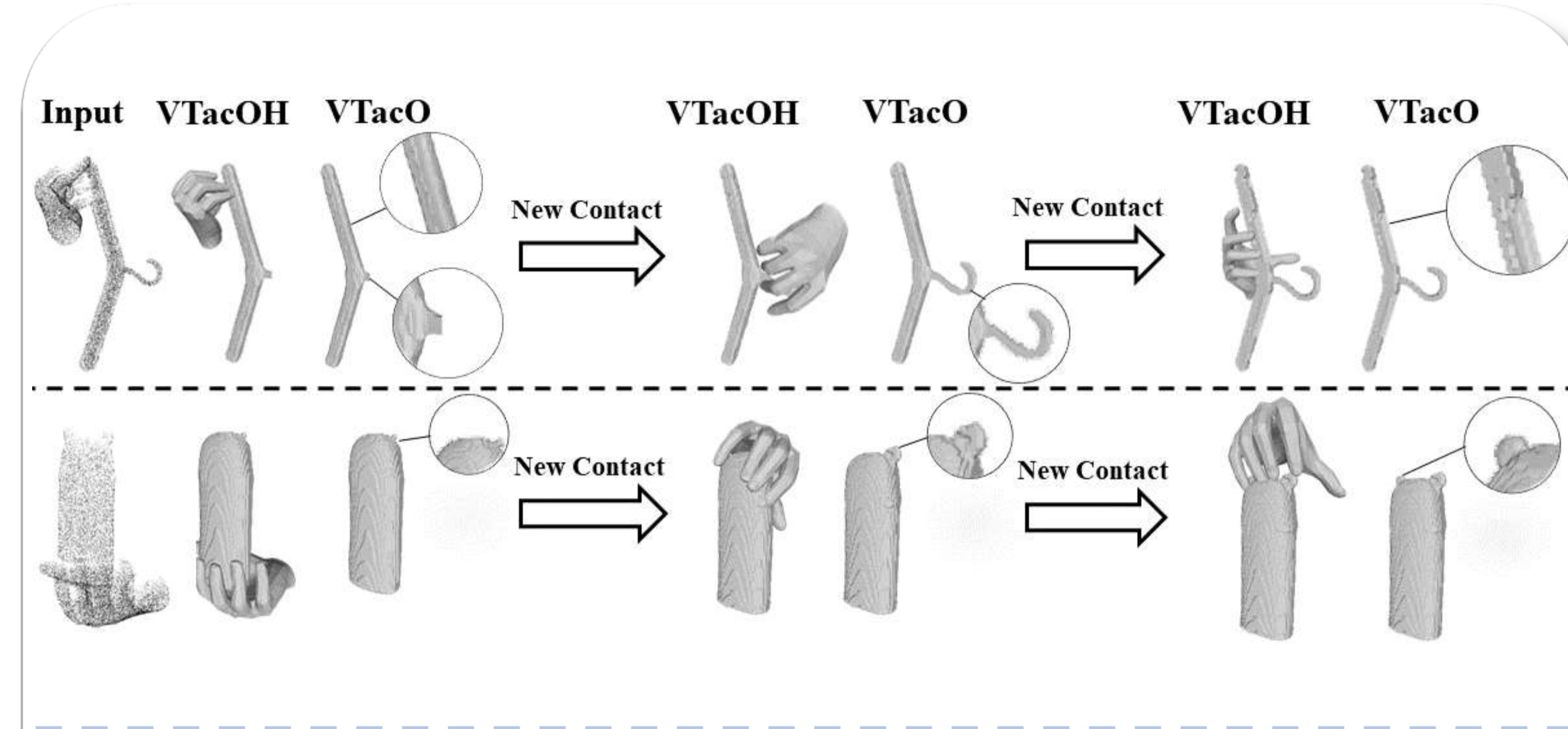




Introduction

Framework of VTacO and VTacOH

Qualitative Results



Main contributions:

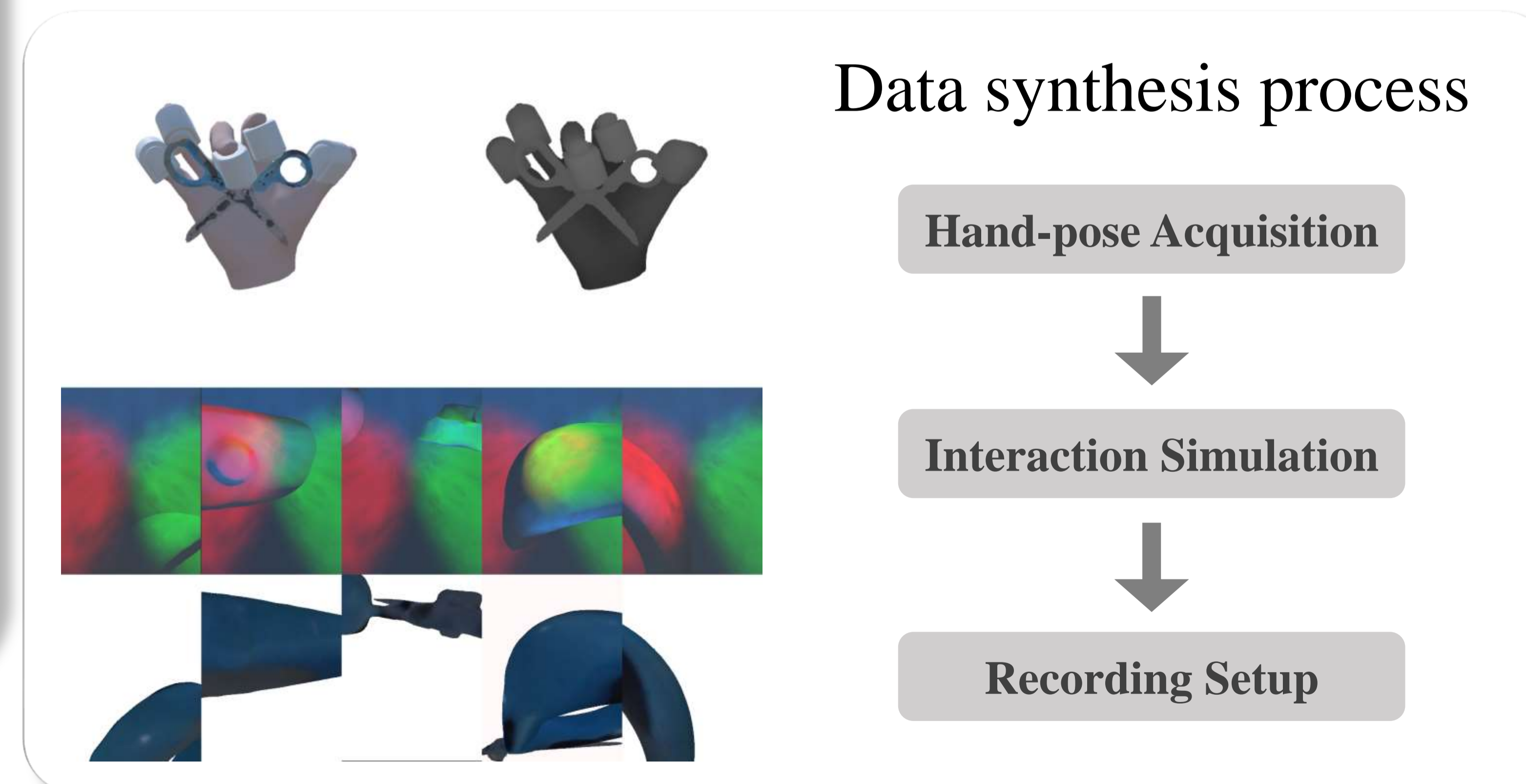
Visual-tactile Learning Framework

- Visual-Tactile in-hand Object reconstruction: **VTacO**
- Hand-Object version **VTacOH**

Simulation Environment: VT-Sim

- Based on **Unity**
- Simulation of **hand-object interaction**
- Acquisition of **training samples** $T = \{T_i\}_{i=1}^N$, including WNF, depth images, tactile signals, etc.

VT-Sim



Quantitative Results

Metrics	Method	bottle	box	foldingrack	lock	scissor	mean
IoU	3D Vision and Touch	*	*	0.746	0.468	0.691	0.622
	Ours (Vision only)	0.884	0.934	0.837	0.877	0.763	0.857
	Ours (VTacO)	0.887	0.950	0.782	0.916	0.777	0.860
	Ours (VTacOH)	0.886	0.947	0.765	0.911	0.772	0.855
CD	3D Vision and Touch	*	*	0.242	2.631	3.206	1.268
	Ours (Vision only)	1.109	0.459	1.579	1.140	7.549	1.472
	Ours (VTacO)	0.936	0.305	1.360	0.932	0.894	0.798
	Ours (VTacOH)	0.948	0.312	1.432	0.955	0.945	0.916
EMD	3D Vision and Touch	*	*	0.052	0.175	0.081	0.090
	Ours (Vision only)	0.054	0.028	0.026	0.083	0.110	0.051
	Ours (VTacO)	0.059	0.018	0.028	0.012	0.052	0.028
	Ours (VTacOH)	0.056	0.024	0.027	0.035	0.068	0.042

* indicates equal contributions