Transfusion: A Novel SLAM Method Focused on Transparent Objects

Anonymous ICCV submission

Paper ID 3859

In this supplementary material, we provide more details about the Trans-SLAM database which mentioned in the main paper. Besides, there are more results of contrast experiments on this database show in this supplementary material.

1. Database Details

Scenes in the existing RGB-D SLAM database do not contain transparent objects. To evaluate our algorithm, we collect a new RGB-D database called Trans-SLAM, which contains 25 video sequences, and each sequence contains at least one transparent object.

1.1. Data Acquisition Devices

We acquired a large set of data recordings containing both the RGB-D data from the RealSense D435i RGB-D camera and the ground truth estimates from Quality motion capture system.

RealSense D435i RGB-D camera has been calibrated with the kalibr toolbox before data recording. The RGB-D data is recorded at the frame rate of 30 Hz and sensor resolution of 640x480. Besides, the scale factor of depth map is 5000.

Quality motion capture system is composed of eight high speed cameras. The sampling rate of the motion capture system is 250 Hz, we put the reflective markers on the RGB-D camera to increase the accuracy of pose estimation.

The sampling rate of the camera and the motion capture system are different. So we use timestamp alignment algorithm to align the images and pose information.

1.2. Scenes Composition

As indicated in the figure 1, we use three opaque objects and five transparent objects to set up the scenes in our database. In each scene,there are three opaque objects behind one transparent object. For each transparent object, we capture five video sequences, and in each sequence the opaque objects are placed in different orders. Figure 2 shows the RGB image and depth map of all scenes in our database.

1.3. Database Format

The original data has been recorded as a ROS bag file. In total, we collected 77 GB of RealSense data, divided into separate 25 sequences. Then we convert them to the PNG image format. The name of images is the timestamp when they are captured. For convenience, our trajectory format uses TUM [2] data format.

(a) **(b)** **(c)** **(d)**

Figure 1. **Tools used in data collecting**(a) Three cameras of motion capture system. (b) the RealSense D435i RGB-D camera with reflective markers. (c) Objects used in the scenes. (d) Typical scene

Scene	Sequence	ATE			R.RPE			T.RPE		
		RGBD-SLAM V2	Elasticfusion	Ours	RGBD-SLAM V2	Elasticfusion	Ours	RGBD-SLAM V2	Elasticfusion	Ours
Long-Necked Vase	seq1	0.443	0.341	**0.178**	0.156	0.089	**0.017**	0.219	0.189	**0.021**
	seq2	0.429	0.370	**0.156**	0.111	0.113	**0.020**	0.217	0.194	**0.026**
	seq3	0.432	0.298	**0.139**	0.172	0.148	**0.018**	0.220	0.193	**0.023**
	seq4	0.457	0.329	**0.157**	0.082	0.110	**0.021**	0.221	0.194	**0.024**
	seq5	0.467	0.331	**0.159**	0.154	0.158	**0.016**	0.221	0.192	**0.025**
Cylindrical Glass	seq1	0.355	0.249	**0.110**	0.256	0.139	**0.086**	0.258	0.128	**0.024**
	seq2	0.354	0.315	**0.175**	0.216	0.144	**0.074**	0.262	0.138	**0.018**
	seq3	0.347	0.284	**0.140**	0.238	0.142	**0.090**	0.249	0.119	**0.026**
	seq4	0.349	0.318	**0.122**	0.247	0.123	**0.084**	0.244	0.141	**0.031**
	seq5	0.352	0.272	**0.108**	0.270	0.166	**0.086**	0.254	0.137	**0.024**
Angular glass	seq1	0.348	0.355	**0.117**	0.132	0.135	**0.014**	0.226	0.294	**0.041**
	seq2	0.352	0.352	**0.108**	0.137	0.166	**0.018**	0.213	0.308	**0.032**
	seq3	0.353	0.340	**0.155**	0.133	0.129	**0.030**	0.235	0.273	**0.033**
	seq4	0.344	0.360	**0.140**	0.128	0.143	**0.010**	0.212	0.202	**0.030**
	seq5	0.347	0.345	**0.121**	0.143	0.155	**0.008**	0.228	0.189	**0.037**
Belly Glass Vase	seq1	0.441	0.420	**0.115**	0.137	0.171	**0.023**	0.394	0.384	**0.049**
	seq2	0.411	0.602	**0.205**	0.173	0.183	**0.022**	0.406	0.366	**0.072**
	seq3	0.511	0.527	**0.158**	0.146	0.186	**0.022**	0.354	0.410	**0.023**
	seq4	0.551	0.589	**0.163**	0.175	0.231	**0.021**	0.345	0.435	**0.060**
	seq5	0.414	0.580	**0.206**	0.240	0.249	**0.020**	0.371	0.423	**0.011**
Glass Fish Tank	seq1	1.706	2.228	**0.819**	0.255	0.141	**0.035**	0.838	0.523	**0.411**
	seq2	2.037	2.545	**0.956**	0.235	0.127	**0.018**	0.836	0.492	**0.476**
	seq3	1.877	2.673	**0.964**	0.312	0.144	**0.031**	0.810	0.596	**0.430**
	seq4	1.824	2.101	**1.016**	0.289	0.122	**0.022**	0.821	0.590	**0.366**
	seq5	2.171	2.325	**0.839**	0.274	0.143	**0.036**	0.819	0.518	**0.433**
Mean		0.707	0.778	**0.301**	0.192	0.150	**0.034**	0.379	0.305	**0.110**

Table 1. Tracking Result on All sequence

2. Results

We compared our method with Elasticfusion [3] and RGB-D SLAM v2 [1]. Due to the space limitation, we only list the average tracking results and one reconstruction result in the main paper. All the tracking results are listed in table 1, and the reconstruction results are shown in figure 3 4 5 6 7.

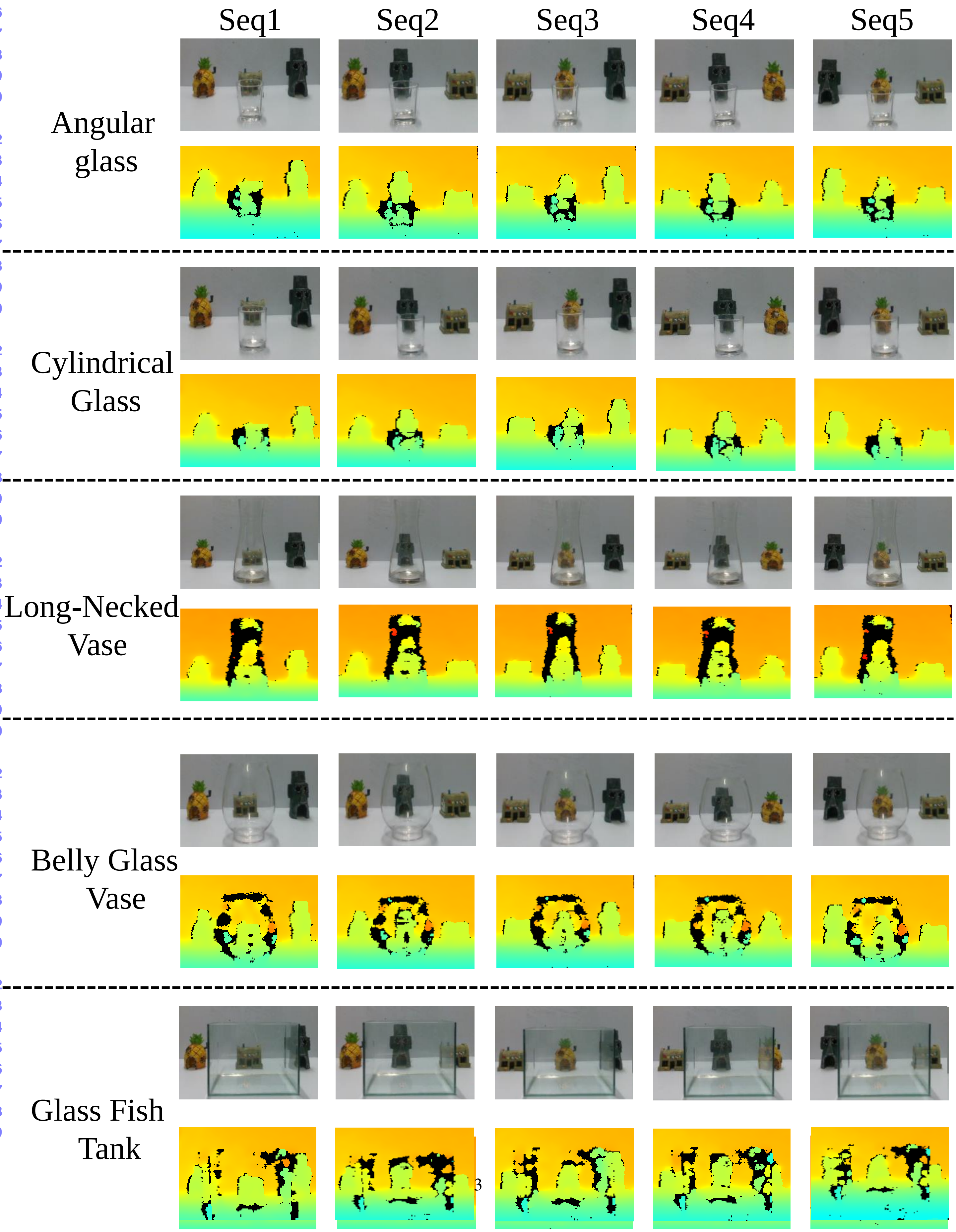

Figure 2. **Database Overview**

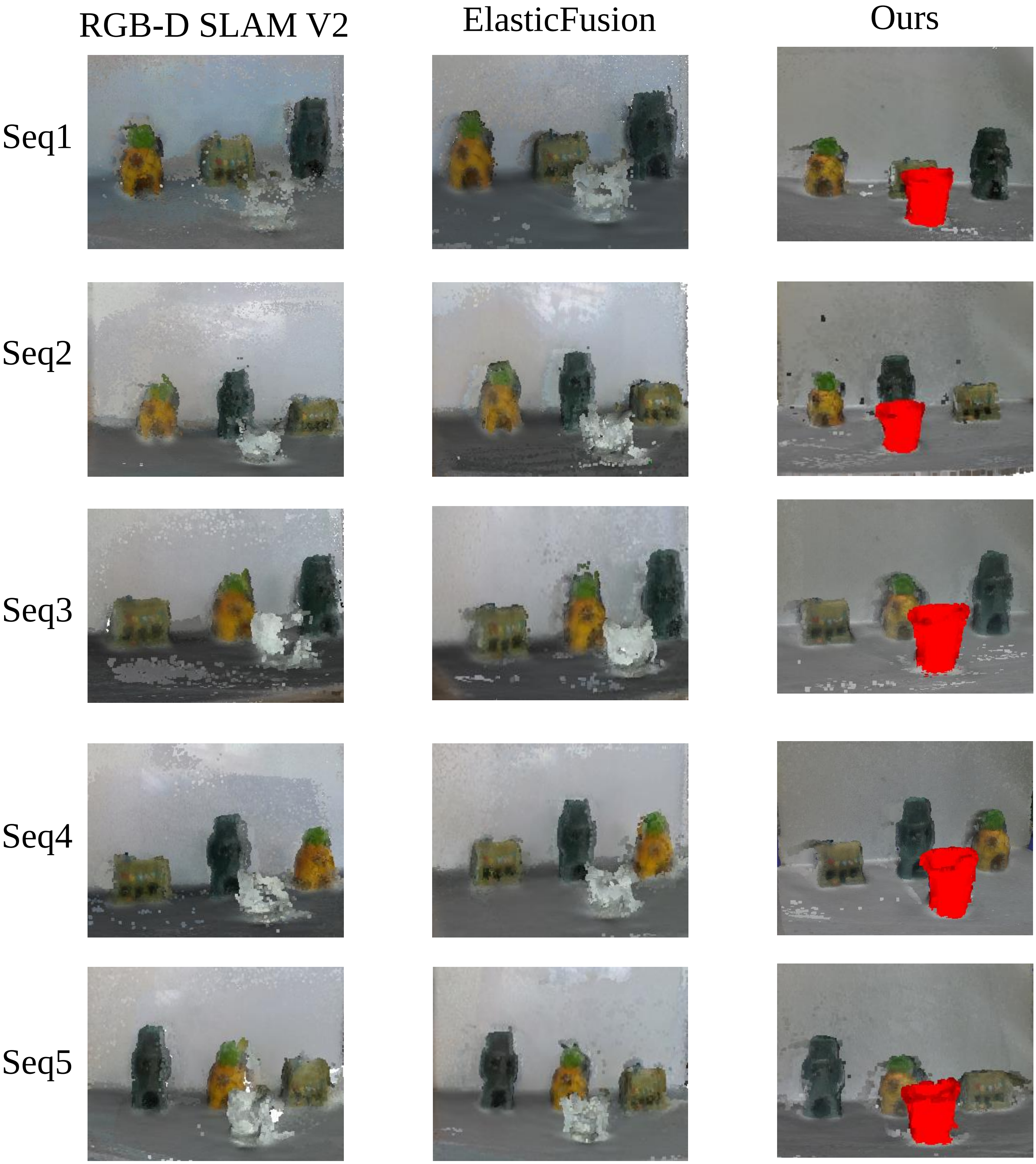

Figure 3. **Reconstruction result of Angular glass**

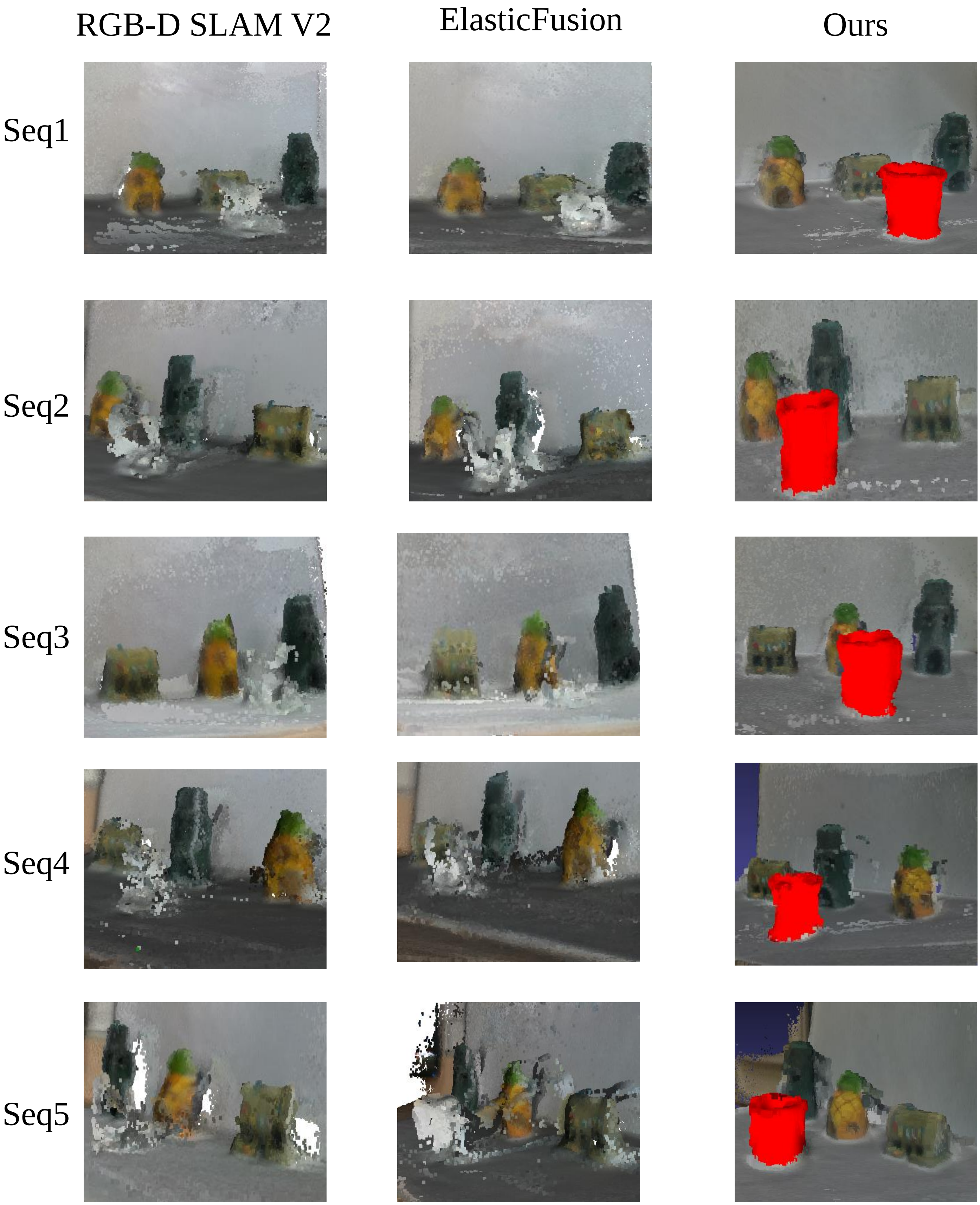

Figure 4. **Reconstruction result of Cylindrical Glass**

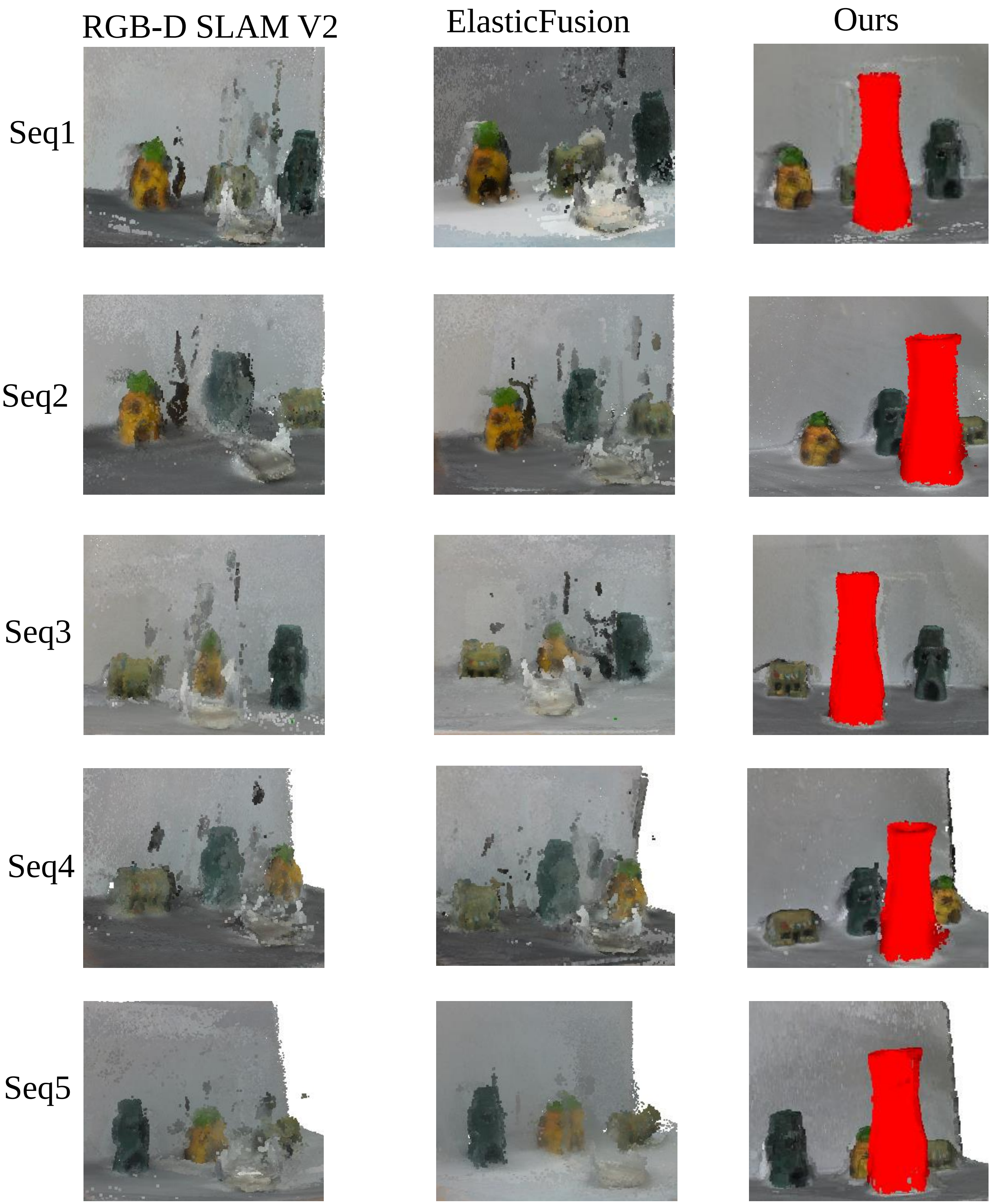

Figure 5. **Reconstruction result of Long-Necked Vase**

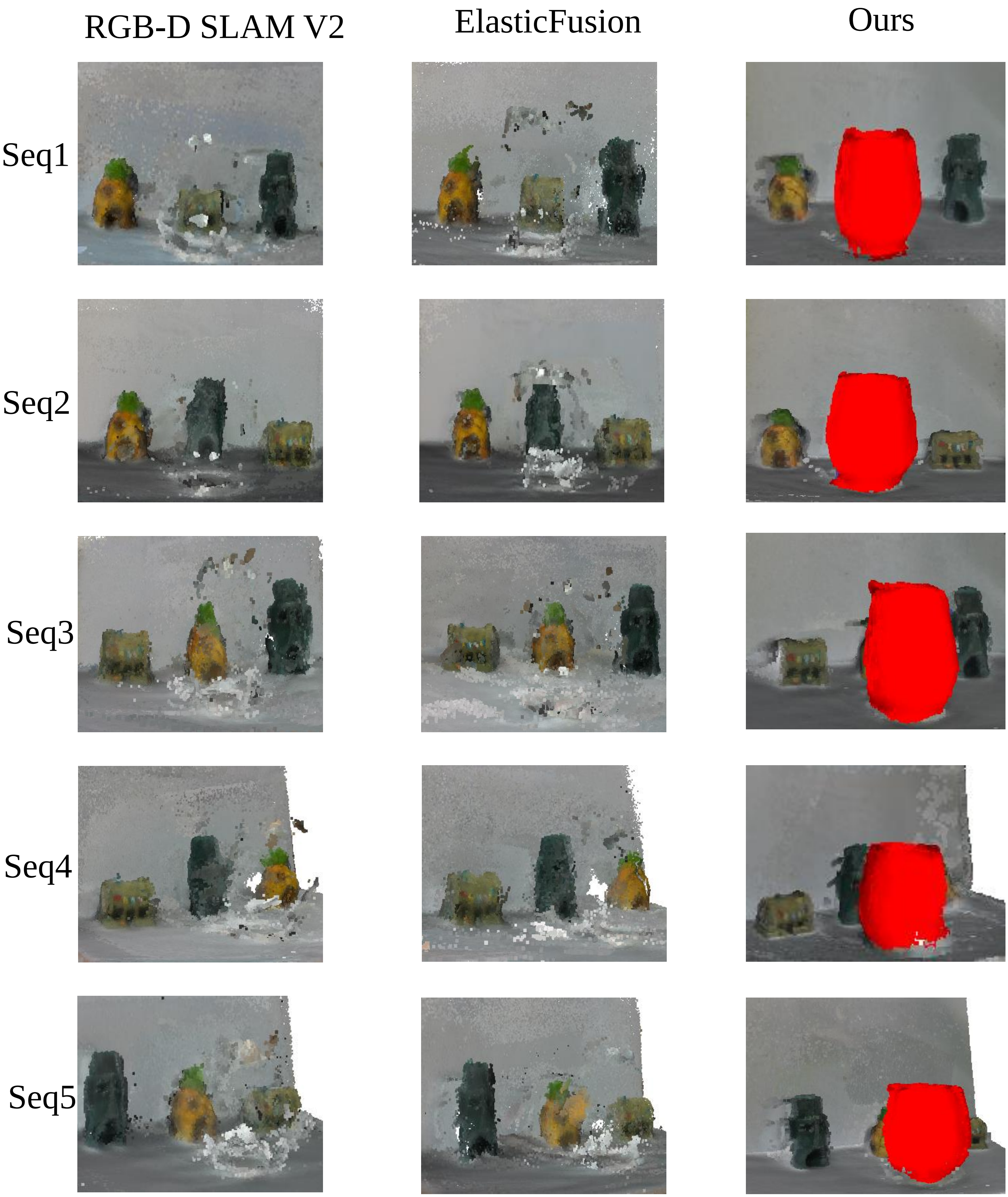

Figure 6. **Reconstruction result of Belly Glass Vase**

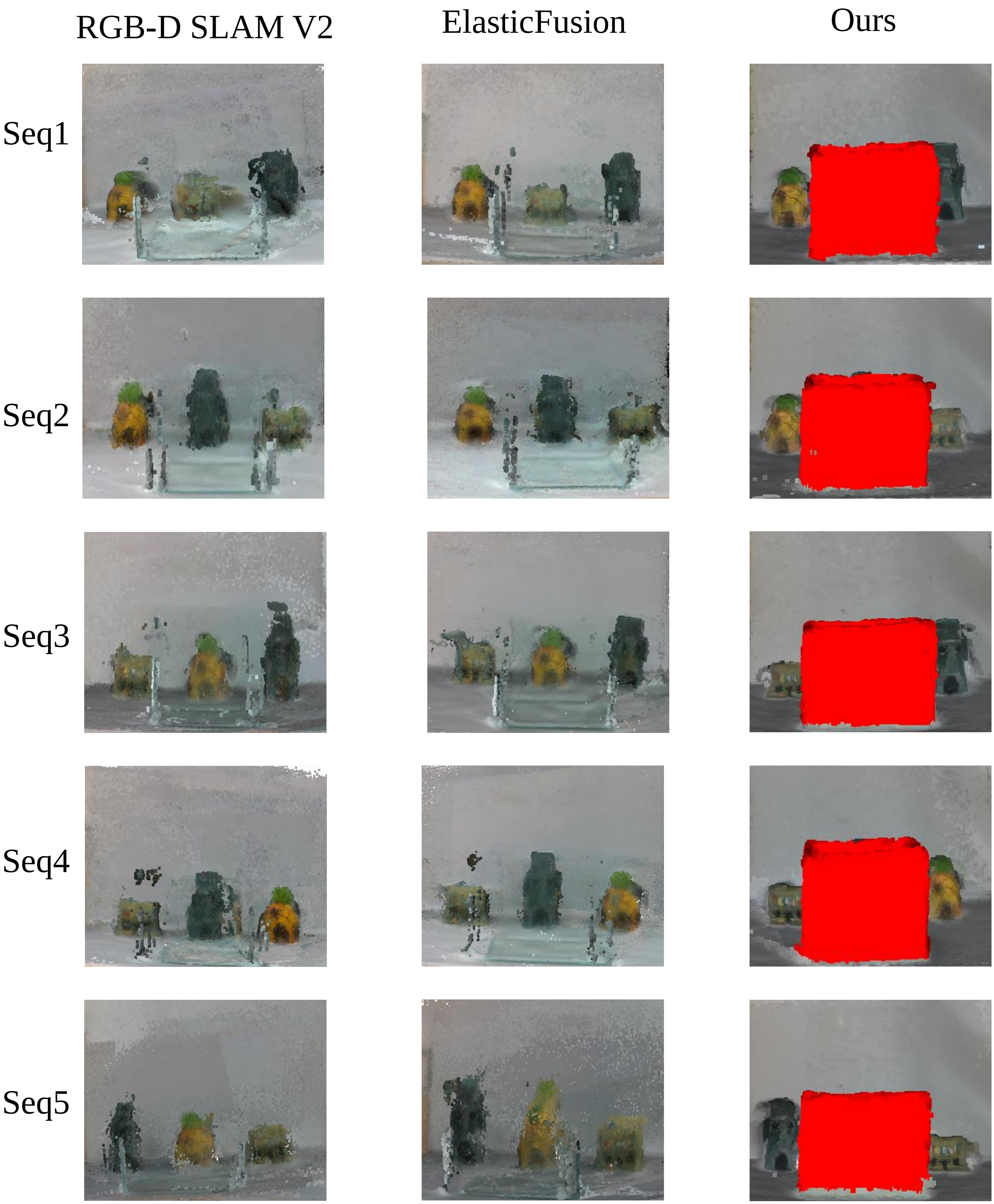

Figure 7. **Reconstruction result of Glass Fish Tank**

References

[1] Felix Endres, Jürgen Hess, Jürgen Sturm, Daniel Cremers, and Wolfram Burgard. 3-d mapping with an rgb-d camera. *IEEE transactions on robotics*, 30(1):177–187, 2013. 2

[2] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers. A benchmark for the evaluation of rgb-d slam systems. In *Proc. of the International Conference on Intelligent Robot Systems (IROS)*, Oct. 2012. 1

[3] Thomas Whelan, Stefan Leutenegger, Renato Salas Moreno, Ben Glocker, and Andrew Davison. ElasticFusion: Dense SLAM Without A Pose Graph. In *Robotics: Science and Systems XI*. Robotics: Science and Systems Foundation, July 2015. 2