Supplementary Material

Xiaotong Chen, Huijie Zhang, Zeren Yu, Anthony Opipari, and Odest Chadwicke Jenkins

University of Michigan, Ann Arbor MI 48109, USA
{cxt,huijiezh,yuzeren,topipari,ocj}@umich.edu

A Dataset Collection Details

The 3D CAD models of objects in the ClearPose dataset are manually created in Blender, with geometry size measurement of real objects. Then during the labeling process of aligning object models to RGB images, the object dimensions are further verified and corrected. Finally, the object models could align well with RGB images in multiple scene sequences.

We used a series of table cloth with different textures as background when collecting the dataset. For each scene, after we randomly put objects in place, a collector will hold the camera and move steadily around the scene with normal room light from relative low, middle, high altitude view angles, for 1 round for each view angle, then another 3 rounds with dim room light and 3 rounds with dim room light plus side light from the lighting board mounted on a tripod on the side. The RealSense camera records aligned RGB-depth image pairs around 30Hz with resolution 1280×720 in scenes with normal room light, and 10-20Hz in other scenes. Most device mentioned above are shown in Figure 1.

A video with detailed annotation steps and results of depth completion and object pose estimation is available at YouTube

Fig. 1. On the left there are lighting board, tripod, and Intel RealSense L515 camera put on the background round table. On the right there are most of background table cloth used in the dataset.

B Object Pose Estimation

More qualitative results with visualization on depth completion and object pose estimation are included in the accompanied video.

Specific to FFB6D, eight 3D sample keypoints on each object are required for the deep neural network to regress. We manually specify keypoints as shown below Figure 2-6. For axial-symmetric objects, all of eight keypoints are generated with equal interval along the symmetric axis. For other objects, keypoints are manually selected to be at feature-rich areas, such as corners or edges.

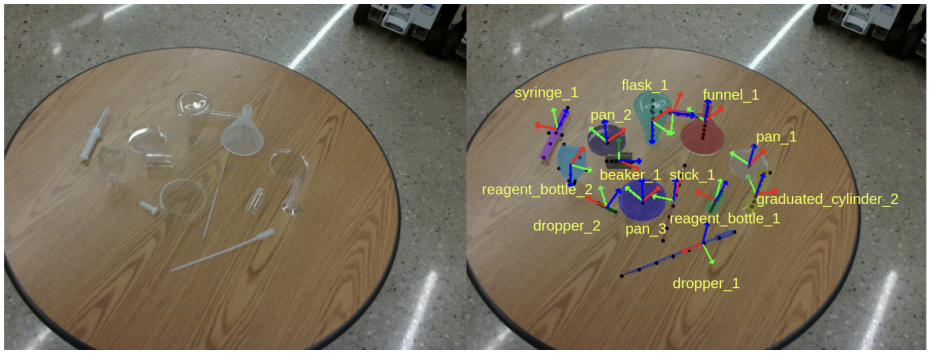

Fig. 2. Object names and labeled keypoints in set 1 of 5.

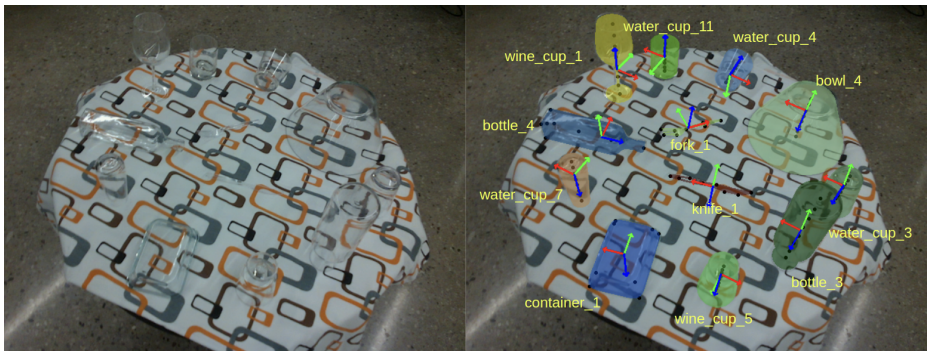

Fig. 3. Object names and labeled keypoints in set 2 of 5.

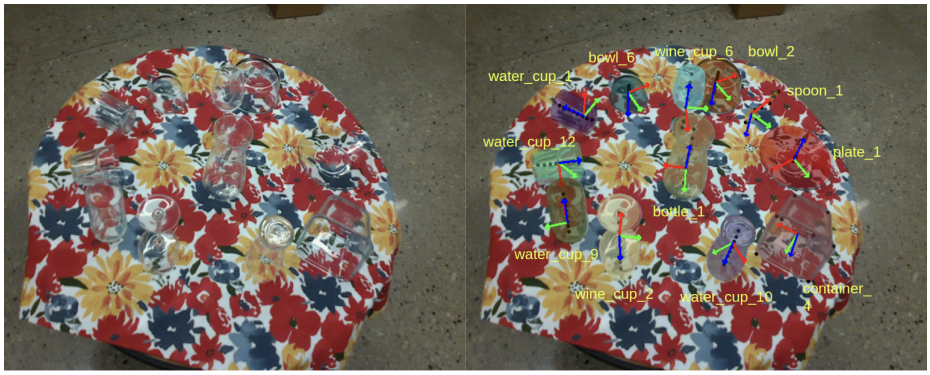

Fig. 4. Object names and labeled keypoints in set 3 of 5.

Fig. 5. Object names and labeled keypoints in set 4 of 5.

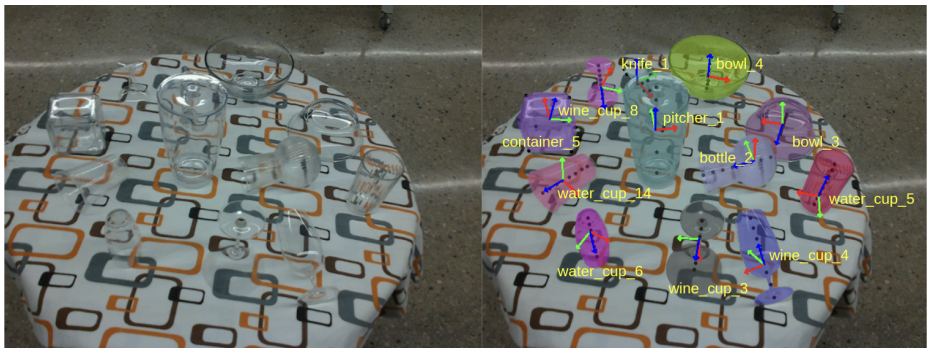

Fig. 6. Object names and labeled keypoints in set 5 of 5.