

# Recognition and Pose Estimation of Rigid Transparent Objects with a Kinect Sensor

Ilya Lysenkov  
Itseez  
ilya.lysenkov@itseez.com

Victor Eruhimov  
Itseez  
victor.eruhimov@itseez.com

Gary Bradski  
Willow Garage  
gary@industrial-perception.com

## Motivation



Mrs. Cupcake's blog, <http://bit.ly/a-lot-of-glass>

Robots should be able to clean out your table with glassware. It means they have to recognize and estimate poses of transparent objects to grasp them.

## Challenges

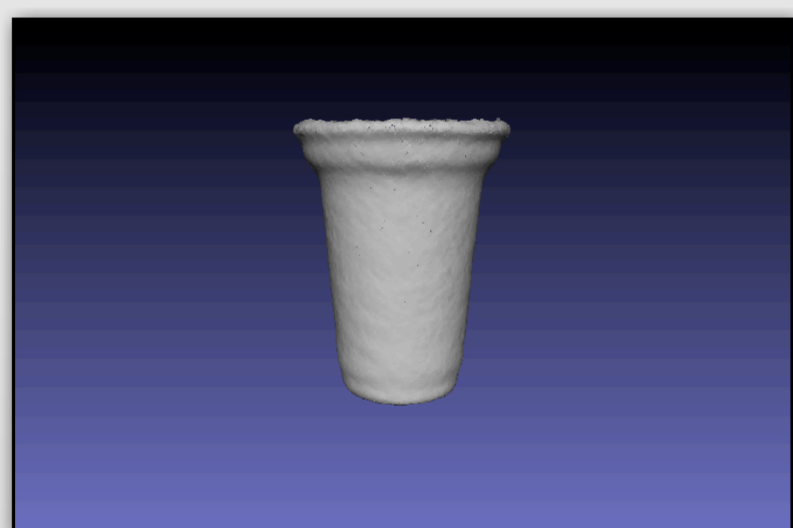
- 3D sensors can't estimate depth or produce point clouds for transparent objects
- Transparent objects appearance heavily depends on background behind them
- Object 2D image is distorted due to refraction, reflection and specularities



K. McHenry, J. Ponce, D. Forsyth, Finding glass, CVPR 2005.

## Proposed approach

### Training stage



Create 3D models of objects with KinectFusion\*

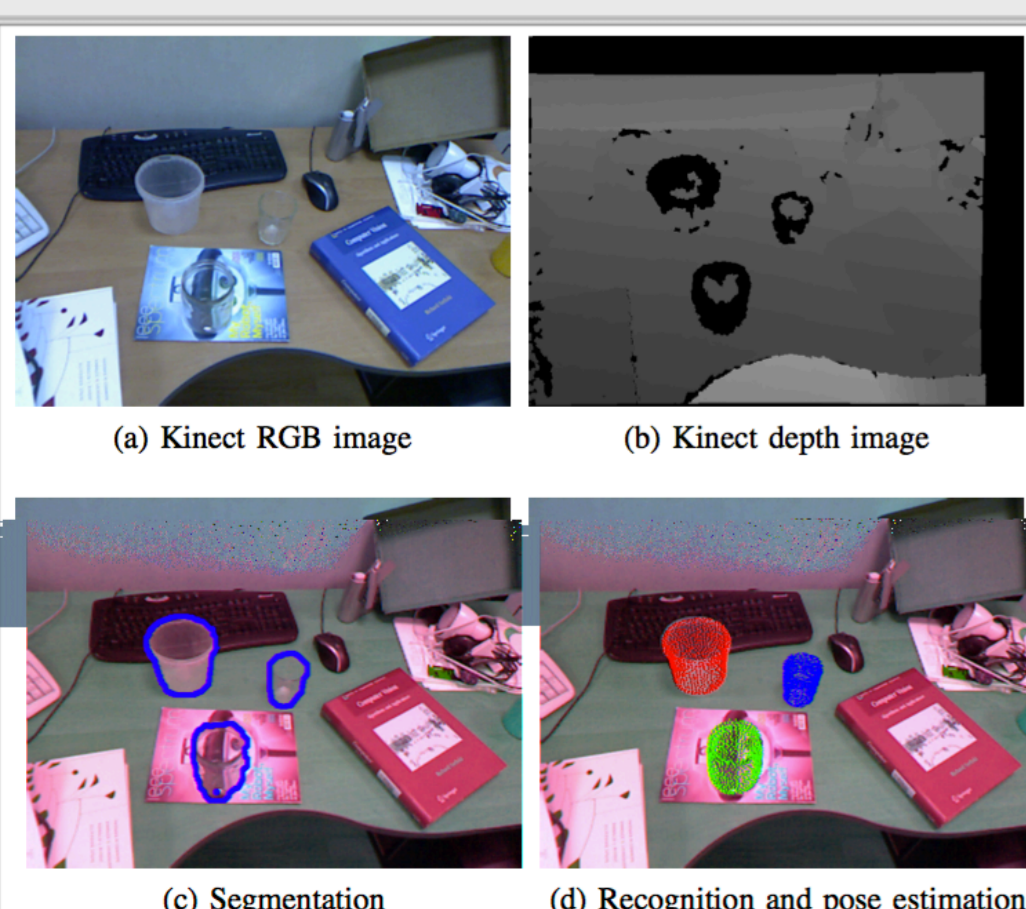
\* Open source PCL implementation of the paper  
R. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. Davison, P. Kohli, J. Shotton, S. Hodges, A. Fitzgibbon, KinectFusion: Real-time dense surface mapping and tracking, ISMAR 2011.

### Testing stage



- **Segmentation:** Kinect can't estimate depth in regions where transparent objects are located —→ the invalid depth mask is used to segment them
- **Pose estimation:** edge fitting to segmented silhouettes
- **Recognition:** based on the fitting quality

## Results



Example of the algorithm working

Object	Success rate	Mean translation error (cm)	
		absolute	relative
bank	0.99	0.3	0.2
bottle	0.99	1.2	0.5
bucket	1.00	0.5	0.5
glass	0.94	0.4	0.3
wine glass	0.97	1.1	0.5
<b>mean</b>	<b>0.98</b>	<b>0.7</b>	<b>0.4</b>

Accuracy of pose estimation



Grasping with the PR2 robot

	Uniform background	Cluttered background
<b>Successful grasps</b>	82.5%	80.0%

Grasping statistics for 8 objects